



HAL
open science

Contributions à l'identification en boucle fermée et à la commande (Application aux lecteurs de disques DVD)

Bohumil Hnilicka

► **To cite this version:**

Bohumil Hnilicka. Contributions à l'identification en boucle fermée et à la commande (Application aux lecteurs de disques DVD). Automatique / Robotique. Université Joseph-Fourier - Grenoble I, 2004. Français. NNT: . tel-00169911

HAL Id: tel-00169911

<https://theses.hal.science/tel-00169911>

Submitted on 5 Sep 2007

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

UNIVERSITÉ JOSEPH FOURIER-GRENOBLE
and
BRNO UNIVERSITY OF TECHNOLOGY

Ph.D. THESIS
for obtaining the degree of
DOCTEUR de l'UJF (Ph.D.)

Special field: Automatique-Productique
prepared at the
Laboratoire d'Automatique de Grenoble
and
Department of Control and Instrumentation
Faculty of Electrical Engineering and Communication

in the context of the École Doctorale
Électronique, Électrotechnique, Automatique, Télécommunications, Signal

presented and sat by
Bohumil Hnilička

the
12th of July 2004

**CONTRIBUTIONS ON IDENTIFICATION IN CLOSED-LOOP
AND CONTROL DESIGN
(APPLICATION TO DVD PLAYERS)**

Ph.D. supervisors: Alina Voda
Petr Pivoňka

JURY

Mr.	D. Georges	President	Professor at the INPG-LAG, FR
Mr.	P. Borne	Reviewer	Professor at the Ecole Centrale de Lille, FR
Mr.	P. Dostál	Reviewer	Professor at the Tomas Bata University in Zlín, CZ
Mr.	D. Georges	Examiner	Professor at the INPG-LAG, FR
Mr.	P. Bláha	Examiner	Researcher at the Brno University of Technology, CZ
Mr.	H.-J. Schröder	Examiner	Project leader at STMicroelectronics, FR
Mr.	L. Bejček	Examiner	Assoc. Professor at the Brno University of Technology, CZ
Mrs.	A. Voda	Ph.D. supervisor	Maître de conférences à l'UJF-LAG, FR
Mr.	P. Pivoňka	Ph.D. supervisor	Professor at the Brno University of Technology, CZ

Date of doctoral state exam: May 28, 2002

Date of Ph.D. thesis submission: May 12, 2004

Study specialization: Cybernetics, Control and Measurements

This thesis is at one's proposal:

Laboratoire d'Automatique de Grenoble

ENSIEG, INPG, BP. 46

38402 Saint Martin d'Hères Cedex

France

Tel: +33.4.76.82.62.44

Fax: +33.4.76.82.63.88

E-mail: lag@lag.ensieg.inpg.fr

Faculty of Electrical Engineering and Communication

Brno University of Technology

Údolní 53

602 00 Brno

Czech Republic

Tel: +420 541 141 172

Fax: +420 541 141 123

E-mail: uamt@feec.vutbr.cz

To my parents.

Acknowledgement

There are so many people, whom I should thank to that I cannot mention them all. However, I will at least try.

First of all, I would like to express my sincere gratitude to my supervisors, Assoc. Prof. Alina Voda and Prof. Petr Pivoňka, whose advice has always encouraged me and stimulated my devotion to the topic of my work. Their profound knowledge of automatic control and their practical experiences have helped me pierce into this field. I feel very happy that I have had a chance to cooperate with them.

Furthermore, I also would like to thank to all the people that have helped me during very instructive visits in the STMicroelectronics application laboratory. I thank to Heinz-Jörg Schröder, the DVD recorder technical advisor, and to Pascal Nonier, the DVD Front-End application team manager, whose assistance and advice have always been helpful to me when I had to face practical problems related to the industrial system implementation. Giampaolo Filardi has helped me by realizing modifications of the experimental set-up.

I would like to thank to Gabriel Buche, the lead engineer of the RABBIT project, for his valuable help on the experimental set-up of the bipedal robot.

I cannot indeed forget to mention the support and the precious help given by all the colleagues and friends that work in the different departments: Jonathan Soen, Petr Dub, Jiří Štěpánek, Igor Mat'ko, Hynek Raisigel, Kamil Švancara, Michal Bezděk, Michal Hrouzek. Thank you!

Special thanks belong to Hynek Procházka for his valuable comments and discussions regarding various parts of this work.

I also would like to thank to the rest of the people in a group of Automatic Control in Laboratoire d'Automatique de Grenoble and Department of Control and Instrumentation in Brno University of Technology because their genuine generosity has made my being a part of these groups so pleasant.

Finally, I cannot miss the occasion to express my love to my parents, sister and brother, who have always supported me throughout all my life and who have been missing my presence, as well as I have been missing them, during last four years.

Bohumil Hnilička
Grenoble, May 2004

This thesis has been partially supported by the European Commission (Marie Curie Fellowship) and by the Czech Grant Agency GAČR under the grant number GAČR 102/01/1485 Environment for Development, Modelling and Application of Heterogenous Systems and by the MŠMT research plan CEZ MSM: 260000013 Automation of Technological and Manufacturing Processes.

Contents

1	Introduction	1
1.1	Background	2
1.2	Motivation and Goals of the Thesis	3
1.3	Thesis Outline	7
2	State-space Identification for SISO/MIMO Systems	11
2.1	Introduction	11
2.2	Problem Description	12
2.3	State of the Art	14
2.4	Model Structures	15
2.4.1	Deterministic and Stochastic State-space Models	15
2.4.2	State-space Representation—Canonical Forms	18
2.4.3	Model Structural Identifiability	21
2.5	Optimization	22
2.6	State-space Identification Algorithms for SISO Systems	24
2.6.1	The SISO GM-2 Algorithm	27
2.6.2	The SISO IGM-1 and IGM-2 Algorithms	28
2.6.3	The SISO RLS-2 Algorithm	30
2.7	State-space Identification Algorithms for MIMO Systems	33
2.7.1	The MIMO GM-2 Algorithm	36
2.7.2	The MIMO IGM-1 and IGM-2 Algorithms	37
2.7.3	The MIMO RLS-2 Algorithm	39
2.7.4	A General Structure of PAA for State-space Systems	41
2.8	Properties of Parameter Adaptation Algorithms for State-space Systems	42
2.8.1	Convexity of Criterion Function	42
2.8.2	Stability of PAA in a Deterministic Environment	43
2.8.3	Parametric Convergence Analysis in a Stochastic Environment . .	47
2.8.4	Frequency Distribution of the Asymptotic Bias	52
2.9	Simulation Experiments for SISO Systems	54
2.9.1	Simulation Experiment 1	55
2.9.2	Simulation Experiment 2	60
2.10	Conclusions	60
3	The DVD Player System Description	67
3.1	Introduction	67
3.2	DVD Drive Architecture	67
3.3	The Optics	68
3.3.1	Optical Pick-up Organization	68
3.3.2	Light Intensity Distribution	70

3.3.3	Focus Error Signal Generation: Modelling Problem Description	73
3.4	The Mechanical Servo System	76
3.4.1	Optical Pick-up Unit Control Loops	79
3.5	Disturbances Sources	83
3.5.1	Optical Imperfections	83
3.5.2	Internal Disturbances	85
3.5.3	External Disturbances	86
3.5.4	Disk Surface Defects	86
3.5.5	Disturbances Summary	87
3.6	Control Loops Specifications	89
3.6.1	Focus Control Loop	89
3.6.2	Radial Control Loop	90
3.7	Performance Limitations	90
3.7.1	Control Difficulties of the CD and DVD Players	92
3.7.2	Control Problem Description	92
3.8	Conclusions	93
4	Control System Design	95
4.1	Introduction	95
4.2	State of the Art	95
4.3	Repetitive Disturbances Modelling	96
4.3.1	Focus Loop	96
4.3.2	Radial Loop	97
4.4	Control Design Methodology	98
4.4.1	Combined Pole Placement/Sensitivity Function Shaping	98
4.4.2	Controller Order Reduction	100
4.4.3	Generalized Stability Margin	100
4.5	Control System Design	100
4.5.1	Plant Model in Radial Loop	101
4.5.2	Physical Model Validation	101
4.5.3	Nominal and Uncertainty Model	101
4.5.4	Standard Controller	102
4.5.5	New Controller Design	102
4.6	Simulation and Experimental Results	105
4.6.1	Focus Loop: Simulation Experiments	105
4.6.2	Focus Loop: Real-time Measurements	107
4.6.3	Radial Loop: Simulation Experiments	110
4.6.4	Radial Loop: Real-time Measurements	113
4.7	Conclusions	120
5	Conclusions	121
5.1	Review	121
5.2	Contributions of Presented Thesis	121
5.3	Perspectives	123

A	Closed-loop Output-Error Identification Method ..	125
A.1	Introduction	125
A.2	Basic Equations of Plant, Fixed Predictor and Output-Error	126
A.3	The Family of Closed-loop Output-Error Methods	129
A.4	Parameter Adaptation Algorithms	133
A.5	Properties of Parameter Adaptation Algorithms	142
A.6	Conclusions	146
B	State-space Identification Experiments	149
B.1	Simulation Experiments for SISO Systems	149
B.2	Simulation Experiments for MIMO Systems	153
B.3	Real-time Experiment for SISO System	167
C	Modelling of the Focus Error Signal Generation	173
C.1	Introduction	173
C.2	State of the Art	173
C.3	Principle of the Astigmatic Method	174
C.4	Photodetector Model by Astigmatic Method	175
C.4.1	Thin Lens	175
C.4.2	Diverging Cylindrical Lens	176
C.4.3	Optical System Separation	178
C.4.4	Position of the Focal Lines	178
C.4.5	Spot Size	184
C.4.6	Light Intensity Distribution	186
C.4.7	Modelling Summary	190
C.5	Influence of Different Model Parameters	190
C.6	Estimation of Unknown Model Parameters	192
C.7	Possible Applications of Photodetector Characteristics	198
C.7.1	Simulation of Start-up Procedure	198
C.7.2	Feedback Focus Control Loop	199
C.7.3	New Photodetector Characteristics Design	199
C.8	Conclusions	201
D	Radial Error Signal Generation	203
D.1	Modelling of the Radial Error Signal Generation	205
E	Read-out Signal Generation	211
E.1	Modelling of the Read-out Signal Generation	213
	Curriculum Vitae	233

List of Figures

1	Closed-loop scheme for SISO systems.	xx
2	Block scheme of the DVD optical system.	xxi
3	Thesis outline.	xxvi
1.1	Closed-loop scheme for SISO systems.	3
1.2	Block scheme of the DVD optical system.	5
1.3	Thesis outline.	9
2.1	Equivalent feedback representation of PAA associated to the output error predictor (OE plant model) for MIMO systems.	45
2.2	Closed-loop identification scheme for MIMO systems.	45
2.3	Transformed equivalent feedback systems associated to the PAA with time-varying gain for MIMO systems.	47
2.4	Evolution of the parametric distance for one simulation run (SISO experiment 1).	57
2.5	Evolution of the estimated parameters for one simulation run (RLS-2, SISO experiment 1).	58
2.6	Evolution of the estimated parameters for one simulation run (IGM-2, SISO experiment 1).	59
2.7	Evolution of the parametric distance for one simulation run (SISO experiment 2).	62
2.8	Evolution of the estimated parameters for one simulation run (RLS-2, SISO experiment 2).	63
2.9	Evolution of the estimated parameters for one simulation run (IGM-2, SISO experiment 2).	64
3.1	Schematic view of the DVD architecture.	68
3.2	Optical pick-up unit organization of the DVD mechanism for a single-layer single-side optical disk.	69
3.3	Schematic view of the DVD impressed structure.	71
3.4	Simplified view of the disk impressed structure.	71
3.5	The light intensity distribution of the Airy disk (solid line) and Gaussian (bold dashed line) focused spot profiles illustrating the measurement of d_{Airy} and $d_{\text{Airy(FWHM)}}$ of the Airy disk profile, and $d_{e^{-2}}$ of the Gaussian profile.	73
3.6	The focused laser spot and its light intensity distribution.	74
3.7	An example of measured S-curve, in the time-domain, from an industrial DVD-video player.	75
3.8	Astigmatic method for focus error signal generation.	76

3.9	Photodetector characteristic (S-curve). The case “S ₁ ” shows the situation where the objective lens is too far from the disk information layer and the case “S ₉ ” shows the situation where the objective lens is too close to the disk information layer.	77
3.10	The shapes of laser spots in the photodetector plane if diverging cylindrical lens is used.	77
3.11	An industrial DVD mechanical servo system.	78
3.12	Case A: Representation of the DVD drive servo system mechanical construction. Case B: Schematic cross section of the DVD drive actuators.	79
3.13	Beam focusing upon the disk information layer.	81
3.14	Control of the DVD player, the focus and tracking loop, and main physical components.	82
3.15	Block diagram of the focus control loop of fig. 3.14, and main signals.	82
3.16	Block diagram of the tracking control loop of fig. 3.14, and main signals.	83
3.17	Sources of optical disk noises.	84
3.18	Case A: Optical disk and actuators of the CD player. Case B: Optical disk and actuators of the DVD player.	85
3.19	Disk eccentricity as the maximal value of the disk displacement on the spindle.	85
3.20	Disturbance sources acting in the focus control loop.	88
3.21	Disturbance sources acting in the radial control loop.	88
3.22	Specification of the control loops, $N = 1$. Case A: Focus loop. Case B: Radial loop.	91
4.1	Geometry of the vertical deviation sources that are mainly caused by the disk warping and disk displacement on the spindle.	98
4.2	Geometry of the radial deviation source that is mainly caused by disk displacement on the spindle; <i>i.e.</i> the disk is not ideally placed on the spindle.	98
4.3	Closed-loop system with RS controller.	99
4.4	Desired template for the modulus of the output sensitivity function $ S_{yp} $ for radial tracking (radial loop) of the DVD, $N = 1.5$	104
4.5	Output sensitivity function, focus loop.	106
4.6	Input sensitivity function, focus loop.	106
4.7	Complementary sensitivity function, focus loop.	107
4.8	The measured and the simulated magnitude of the output sensitivity function obtained for the 3rd order controller K_{RS3} and the actual controller K_{RS3} , focus loop.	108
4.9	The measured power spectrum density of the focus error signal e_F for the 3rd order designed controller K_{RS3} and actual implemented controller K_{act} . The test disk has very small disk eccentricity x_{d_max} , but with high disk vertical deviation at the disk outer edge $z_{d_max} = 0.5$ mm, $f_{rot} = 15$ Hz.	109
4.10	The same as in fig. 4.9 but for $f_{rot} = 33$ Hz.	109
4.11	Output sensitivity function, radial loop.	110
4.12	Input sensitivity function, radial loop.	111
4.13	Complementary sensitivity function, radial loop.	111
4.14	Summary: Output sensitivity function, radial loop.	112

4.15	Summary: Step response of the closed-loop, radial loop.	112
4.16	Envelopes of the output sensitivity functions, actual controller, radial loop.	114
4.17	Envelopes of the output sensitivity functions, RS2 controller, radial loop.	114
4.18	Envelopes of the output sensitivity functions, RS3 controller, radial loop.	115
4.19	Envelopes of the radial closed-loop step responses, actual controller. . . .	115
4.20	Envelopes of the radial closed-loop step responses, RS2 controller.	116
4.21	Envelopes of the radial closed-loop step responses, RS3 controller.	116
4.22	Magnitude of the output sensitivity function obtained from measurements and the envelopes of output sensitivity functions from simulation for K_{act} , radial loop.	117
4.23	Magnitude of the output sensitivity function obtained from measurements and the envelopes of output sensitivity functions from simulation for K_{RS2} , radial loop.	118
4.24	Magnitude of the output sensitivity function obtained from measurements and the envelopes of output sensitivity functions from simulation for K_{RS3} , radial loop.	118
4.25	The measured power spectrum density of the radial error signal e_R for the 2nd/3rd order designed controllers (K_{RS2} , K_{RS3}) and actual implemented controller K_{act} . The tested disk has very small disk vertical deviation at the disk outer edge $z_{d_max} \approx 0 \mu\text{m}$ and very small disk eccentricity $x_{d_max} \approx 0 \mu\text{m}$, $f_{rot} = 15 \text{ Hz}$	119
4.26	The same as in fig. 4.25 but for $f_{rot} = 33 \text{ Hz}$	119
A.1	Closed-loop identification scheme for SISO systems.	126
A.2	Equivalent feedback representation of PAA associated to the output-error predictor (OE plant model).	143
A.3	Transformed equivalent feedback systems associated to the PAA with time- varying gain for the output-error predictor (OE plant model).	145
B.1	Evolution of the parametric distance for one simulation run (SISO experi- ment 3).	151
B.2	Evolution of the estimated parameters for one simulation run (RLS-2, SISO experiment 3).	151
B.3	Evolution of the estimated parameters for one simulation run (IGM-2, SISO experiment 3).	152
B.4	Evolution of the parametric distance for one simulation run (MIMO ex- periment 4). Case A: $F(0) = 0.1I$. Case B: $F(0) = 2I$	158
B.5	Evolution of the estimated parameters for one simulation run (RLS-2, SISO experiment 4). Case A: $F(0) = 0.1I$	159
B.6	Evolution of the estimated parameters for one simulation run (RLS-2, MIMO experiment 4). Case B: $F(0) = 2I$	160
B.7	Evolution of the estimated parameters for one simulation run (IGM-2, MIMO experiment 4).	161
B.8	Evolution of the parametric distance for one simulation run (MIMO ex- periment 5).	163
B.9	Evolution of the estimated parameters for one simulation run (IGM-2, MIMO experiment 5).	163

B.10	Evolution of the parametric distance for one simulation run (MIMO experiment 6). Case A: $\sigma_{ei} = 0$. Case B: $\sigma_{ei} = 0.000001$	165
B.11	Evolution of the estimated parameter for one simulation run (IGM-2, MIMO experiment 6). Case A: $\sigma_{ei} = 0$. Case B: $\sigma_{ei} = 0.000001$	166
B.12	The biped prototype RABBIT's experimental setup. Schematics of the prototype RABBIT with measurement conventions are illustrated on the right side.	167
B.13	Autocorrelation of residuals for output and crosscorrelation for input and output residuals (SISO experiment 7). Case A: ARX221. Case B: ARMAX2221. Case C: OE221.	170
B.14	Frequency response of identified models.	171
C.1	Block-scheme of the focus error generation model.	174
C.2	The thin lens and its Lensmaker's equation.	176
C.3	The converging spherical lens with coordinates system. If incident beams are parallel to the optical axis z than the focal point is placed in lens focal length behind of the spherical lens.	176
C.4	The diverging cylindrical lens with coordinates system. If we assume parallel incident beams with optical axis z than apparent focal line is placed in focal length in front of the cylindrical lens.	177
C.5	The marginal rays of the cylindrical diverging lens if light source (disk) is in focus.	177
C.6	Arrangement in the yz plane. Diverging cylindrical lens is approximated by air.	179
C.7	Arrangement in the xz plane. Diverging cylindrical lens is approximated by diverging thin lens.	179
C.8	System formed by two centered thin lens in the yz plane.	180
C.9	System formed by two centered thin lenses in the xz plane.	182
C.10	Eduction of R_{spotcyl} and half-axis b in the yz plane.	184
C.11	Eduction of half-axis a in the xz plane.	185
C.12	The geometrical arrangement of the photodetector.	187
C.13	A photo of the real photodetector arrangement.	187
C.14	Special position of the photodetector.	189
C.15	Block-scheme of the focus error generation model.	190
C.16	Effect of photodetector size d_1	192
C.17	Effect of photodetector displacement Δx_{det}	193
C.18	Effect of photodetector displacement Δy_{det}	193
C.19	Effect of photodetector displacement $\Delta x_{\text{det}}, \Delta y_{\text{det}}; \Delta x_{\text{det}} = \Delta y_{\text{det}}$	194
C.20	Effect of photodetector gap size d_5	194
C.21	Effect of photodetector angle φ_{det}	195
C.22	Effect of Gauss height ratio H_R	195
C.23	S-curves of the real system and mathematical models.	196
C.24	Weighted nonlinear least-squares criterion for the analytical model.	197
C.25	Weighted nonlinear least-squares criterion for the analytical model.	198
C.26	Block diagram of the focus control loop, focus processing procedure to automatically close the focus loop.	200

C.27	Focus processing procedure to automatically close the focus loop obtained from simulation where the created analytical model of the S-curve has been included.	200
C.28	Focus processing procedure to automatically close the focus loop obtained by measurement on a real DVD-video player from STMicroelectronics. . .	201
D.1	An example of DTD radial error signal generation used in DVD players (DTD4-method).	204
D.2	Case A: An example of radial error signal, in the time-domain, from an industrial DVD-video player—simply opening the radial control loop without jump back, $f_{\text{rot}} = 33$ Hz. Case B: Zoom.	204
D.3	Schematic drawing illustrating the far field pattern generated by an optical disk with a regular information pattern.	206
D.4	The radial error signals generated by DTD2 and DTD4 method with different amounts of de-focus. Draw line: $\Delta z = -0.903 \mu\text{m}$, triangles: $\Delta z = 0 \mu\text{m}$, crosses: $\Delta z = +0.903 \mu\text{m}$	208
D.5	Block-scheme of the radial error generation model.	208
D.6	An example of the radial error signal generation modelling.	210
E.1	Example of HF generation for a pit/land impressed structure.	212

List of Tables

2.1	Summary of the non-recursive and recursive algorithm criterions.	42
2.2	Settings of the SISO experiment 1.	56
2.3	Results of the SISO experiment 1.	56
2.4	Results of the SISO experiment 1, mean values and standard deviations of the individual parameters.	57
2.5	Settings of the SISO experiment 2.	61
2.6	Results of the SISO experiment 2.	61
2.7	Results of the SISO experiment 2, mean values and standard deviations of the individual parameters.	61
3.1	Physical parameters of CD and DVD. $N = 1$	80
3.2	An overview of the disturbance frequencies, DVD player, radial loop, $N = 1$	87
3.3	Focus servo specification for the DVD, $N = 1$, compensated open-loop.	90
3.4	Radial servo specification for the DVD, $N = 1$, compensated open-loop.	91
3.5	Requirements on the typical CD/DVD players, radial loop.	92
4.1	Values of the physical parameters for the radial actuator.	101
4.2	Values of the radial actuator physical parameters together with their maximum percentage variation.	102
4.3	Comparison of the controller order reduction and generalized stability margin, $N = 1.5$, focus loop.	107
4.4	Comparison of the various reduced controllers, $N = 1.5$, focus loop.	108
4.5	Comparison of the controller order reduction and generalized stability margin, $N = 1.5$, radial loop.	113
4.6	Comparison of the various reduced controllers, $N = 1.5$, radial loop.	113
A.1	Summary of closed-loop output-error methods.	133
B.1	Settings of the SISO experiment 3.	150
B.2	Results of the SISO experiment 3.	150
B.3	Results of the SISO experiment 3, mean values and standard deviations of the individual parameters.	150
B.4	Settings of the MIMO experiment 4.	157
B.5	Results of the MIMO experiment 4.	157
B.6	Results of the MIMO experiment 4, mean values and standard deviations of the individual parameters.	157
B.7	Settings of the MIMO experiment 5.	162
B.8	Results of the MIMO experiment 5.	163
B.9	Results of the MIMO experiment 5, mean value and standard deviation of the individual parameter.	163
B.10	Settings of the MIMO experiment 6.	164

B.11	Results of the MIMO experiment 6.	165
B.12	Results of the MIMO experiment 6, mean value and standard deviation of the individual parameter.	165
B.13	Statistical parameters of measured and simulated model output.	169
C.1	Unknown model parameters vector θ and known model parameters vector p for the analytical/numerical model.	190
C.2	Nominal parameters of system and models.	191
C.3	Statistical parameters of the curve-fitting procedure.	197
C.4	Nominal parameters of real system.	198
C.5	Identified parameters.	199

Résumé

Chapitre 1: Introduction

Les méthodes et techniques de l'automatique sont largement utilisées pour forcer les systèmes dynamiques à se comporter d'une manière satisfaisante. La notion large de *système dynamique* peut alors faire allusion à de nombreux processus, allant des systèmes biologiques à ceux du domaine de l'ingénierie. Les applications de l'automatique peuvent être trouvées dans de nombreux procédés industriels complexes pour obtenir un fonctionnement correct.

Parallèlement à l'aspect *commande* de l'automatique, l'identification est très utilisée pour élucider les dynamiques des systèmes. Cette procédure permet de prédire le comportement d'un système inconnu à partir d'une observation antérieure du système. De cette façon, la connaissance de la dynamique d'un procédé industriel ou d'un système mécanique se fonde sur des expériences et sur la construction d'un modèle mathématique adéquate. Dans l'ingénierie, la connaissance d'un modèle est souvent cruciale et, aujourd'hui, la conception des systèmes de commande se base en général sur ces objets mathématiques. Ils sont par ailleurs utilisés dans des domaines non nécessairement techniques, tels que la biologie, l'écologie ou l'économie.

Lorsqu'elles sont connues, les lois de la physique gouvernant le comportement des systèmes peuvent être utilisées pour la construction des modèles appelés *boîtes-blanches*. Dans le cas des modèles *boîtes-blanches*, tous les paramètres et toutes les variables peuvent être interprétés de manière physique et toutes les constantes sont connues *a priori*. A l'opposé, se trouve la modélisation de type *boîte-noire*. Dans cette approche, les modèles sont construits à partir des données entrée-sortie du système. Les connaissances *a priori* sur le système sont écartées. En conséquence, le modèle et ses paramètres n'ont pas forcément de signification physique.

Les structures de modèle linéaires ont été largement utilisées pour la modélisation *boîte-noire* parce qu'elles sont mathématiquement attractives. Cette classe de modèles est assez riche pour couvrir un grand nombre d'applications. En effet, la commande des systèmes s'effectue souvent autour d'un point de fonctionnement, aux alentours duquel, le comportement du système est considéré comme linéaire.

Il existe une troisième approche, la modélisation *boîte grise*, combinant ces deux extrêmes. Cette approche exploite les connaissances physiques *a priori* sur le procédé, sans que la structure du modèle et les paramètres ne soient complètement déterminés. Les paramètres du modèle sont alors estimés comme dans le cas des modèles *boîte-noire*, par des méthodes d'identification. Un cas classique d'identification de type *boîte grise* se produit lorsque la structure du modèle peut être déterminée par les relations physiques mises en jeu - d'où le terme *modèle physique*. Dans ce cas, l'identification est effectuée avec une structure de modèle fixe et seule la valeur des paramètres inconnus est estimée.

Cette thèse porte à la fois sur les aspects commande et identification des systèmes, dans le but de concevoir un système de commande basé un modèle de type *boîte blanche* du procédé. Cette approche est en partie illustrée sur l'exemple industriel du lecteur DVD (Digital Versatile Disk-video), contenant un système opto-électro-mécanique de positionnement haute-précision.

L'étude contient deux parties principales. La première traite des aspects théoriques et méthodologiques de *l'identification en boucle fermée des systèmes boîtes-noires*, en utilisant la représentation d'état comme structure de modèle. La deuxième partie de cette thèse présentera la *modélisation et la synthèse de la commande* pour le système de lecture de DVD.

Généralités

Dans cette thèse, nous allons tout d'abord étudier des méthodes pour l'identification des systèmes dynamiques linéaires et invariants dans le temps, à partir de données échantillonnées et d'une structure donnée du modèle. Actuellement, il existe différentes méthodes pour identifier les modèles linéaires en temps discret, *e.g.* la méthode de la prédiction de l'erreur [Lju99; LLM97], les approches de type sous-espace [OM96], le filtre de Kalman étendu [WP97]. Nous nous focaliserons plus spécialement sur les méthodes de prédiction de l'erreur pour l'identification des systèmes opérant en boucle fermée et connues sous le nom de méthodes CLOE (Closed-Loop Output-Error), voir *e.g.* [LK97a], [LLM97] et [LK97b].

L'identification en boucle fermée a souvent été suggérée comme un outil pour l'identification de modèles convenables pour la commande et donc appelée *identification pour la commande* [HGdB96], [Lju99], [FL99], [GdH01], [Ver93], [dHS95] et [LLM97]. Dans la situation typique, le procédé à identifier est cours de fonctionnement et soumis à une commande par retour. L'objectif de cette commande peut être de réduire le temps de réponse ou rendre le système moins sensible aux bruits ou aux perturbations. Dans ce cas, il est possible d'effectuer soit une identification en boucle ouverte en déconnectant le correcteur, soit une identification en boucle fermée. Par contre, si le système est instable ou doit être régulé pour des raisons économiques ou de sécurité, une identification en boucle fermée est inévitable.

Les méthodes couramment utilisées pour l'identification en boucle ouverte ne permettent pas en général d'exploiter les données obtenues en boucle fermée. C'est notamment le cas de la méthode des variables instrumentales, de l'analyse spectrale et de nombreuses méthodes basées sur les sous-espaces par exemple [FL99]. Ces méthodes ne peuvent pas être appliquées en raison de la corrélation entre l'entrée du système et le bruit en sortie, non mesurable et inévitable en boucle fermée.

Puisqu'on ne peut pas, en général, appliquer directement aux données acquises en boucle fermée les méthodes développées pour l'identification en boucle ouverte, il faut utiliser des méthodes spécifiques. Les méthodes de prédiction de l'erreur peuvent, elles, être adaptées, pour palier au problème de corrélation entre entrée et bruit. Une certaine classe d'algorithmes issus de ces nouvelles méthodes est basée sur l'erreur de sortie et a été développée pour l'identification des modèles entrée-sortie. Ces algorithmes sont connus sous le nom d'algorithmes CLOE [LK97b].

Voici maintenant quelques généralités concernant la modélisation et la conception des systèmes de commande des lecteurs DVD.

Le problème relatif à la commande du mécanisme optique d'un lecteur de DVD est similaire à celui qui se pose pour les lecteurs CD (Compact Disk). Il s'agit de garantir que le faisceau laser, utilisé pour lire les données, suit convenablement les pistes du disque. Contrairement aux *anciennes* platines des disques vinyles, où transducteur piézo-électrique ou magnétique est guidé physiquement par le sillon, dans les systèmes optiques, seule la lumière touche le disque. Le suivi de la piste optique doit alors être réalisé par une commande par retour. Dans les réalisations industrielles actuelles, ce retour est assuré par un simple PID. Or, depuis quelques années, les mécanismes utilisés pour la lecture des supports optiques sont utilisés dans un nombre croissant d'applications, de plus en plus exigeantes. Le CD, support audio, est devenu CD-ROM (Read Only memory), support multimédia, alors que se développait le DVD Vidéo puis le DVD-ROM multimédia. Les nouvelles applications multimédia requièrent des niveaux de performance supérieurs à ceux exigés pour les systèmes de lecture des CD audio. En effet, on souhaite augmenter petit à petit les densités de stockage, tout en exigeant des temps d'accès aux données plus faibles [ECM96], [ECM01].

Jusqu'à maintenant, pour les lecteurs DVD, les niveaux de performances désirés ont pu être atteints par une commande heuristique et par un réglage plus fin, et ce, au prix de coûteuses améliorations des outils de production. Ce choix (l'amélioration des outils de production plutôt que celle système de commande lui-même) est vraisemblablement du au fait que les modèles du système optique actuellement disponibles, notamment ceux concernant la génération des signaux nécessaires à la commande, sont trop complexes, [BW87], [BBH⁺85], et [Bra98].

Les modèles proposés dans cette thèse seront donc un compromis entre complexité et précision. Ce travail envisagera donc la possibilité d'améliorer les performances du système en basant la conception du système de commande sur un modèle physique.

Motivation et but de la thèse

Le contenu de ce manuscrit peut donc être divisé en deux parties. La première partie concerne l'identification en boucle fermée des modèles *boîtes-noires*, en utilisant la représentation d'état pour définir la structure du modèle. La deuxième partie se focalise, elle, sur la description physique des lecteurs DVD, sur l'analyse mathématique de la génération des signaux de lecture et de commande et enfin sur la conception du système de commande.

Partie I

La figure 1 présente la boucle de commande standard pour un système monovariable (SISO). Elle est composée du correcteur K et du modèle de procédé G_S ; $r(t)$ est la consigne, $e(t)$ et $u(t)$ sont l'entrée et la sortie du correcteur, $y(t)$ est la sortie du procédé, $p(t)$ est la perturbation en sortie et $w(t)$ est l'excitation extérieure pouvant être ajoutée à la sortie $u(t)$ du correcteur (ou à la consigne $r(t)$) pour l'identification du système.

Afin d'identifier les procédés fonctionnant en boucle fermée, voir fig. 1, dont la structure du modèle est donnée, des algorithmes d'identification sous forme de représentation d'état en temps discret ont récemment été développés [Bez01; Bez04]. Ils sont basés

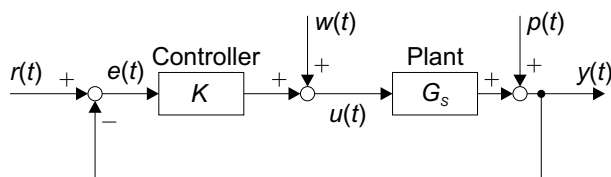


Figure 1: Closed-loop scheme for SISO systems.

sur la méthode d'identification de l'erreur de sortie en boucle fermée (CLOE) introduite précédemment, et développés en premier lieu pour les systèmes SISO *boîtes-noires*, voir *e.g.* [LK97a] et [LLM97]. Ces algorithmes donnent de bons résultats pour l'identification en boucle fermée. Cependant, l'analyse de leurs propriétés n'apparaît pas dans ces références. Elle sera effectuée dans cette partie. De plus, des exemples d'application seront étudiés en simulation.

L'objectif premier de cette thèse était d'utiliser ces algorithmes pour identifier les paramètres physiques inconnus du lecteur DVD, afin de proposer une nouvelle méthode de synthèse du système de commande du DVD. Malheureusement, cette approche n'a pu être retenue. En effet, les contraintes d'expérimentation associées à ce type d'identification se sont trouvées incompatibles avec les possibilités actuelles du banc d'essais industriel du lecteur DVD étudié [Fil03]. L'alternative a donc été de construire à partir des spécifications un modèle linéaire de type *boîte blanche* des actionneurs du lecteur DVD, et de l'utiliser pour la synthèse de la commande.

Les buts de cette partie peuvent être résumés de la façon suivante:

- Présenter de nouveaux algorithmes pour l'identification en boucle fermée de modèles d'état en temps discret et donnés sous forme de représentation d'état.
- Analyser les propriétés de ces algorithmes.
- Appliquer ces algorithmes à des exemples de systèmes mono et multivariables en simulation.

Partie II

Pour des applications telles que la lecture d'informations numériques stockées sur supports optiques (CD, DVD,...), le contrôle en temps réel de la position de l'objectif est impératif. Le faisceau laser utilisé pour lire les données enregistrées doit être focalisé sur la surface et suivre la piste avec une grande précision. Cette tâche est réalisée par les boucles de contrôle de position qui minimisent, à chaque instant, les signaux d'erreur de position générés par le dispositif optique. On distingue en général trois boucles de commande principales dans un lecteur DVD:

- **Une boucle de focalisation**, pour réguler la distance entre l'objectif et le support, voir fig. 3.12.
- **Une boucle de suivi de piste primaire (Basse Fréquence)**, pour positionner grossièrement le dispositif de lecture optique (OPU, Optical Pic-up Unit) en

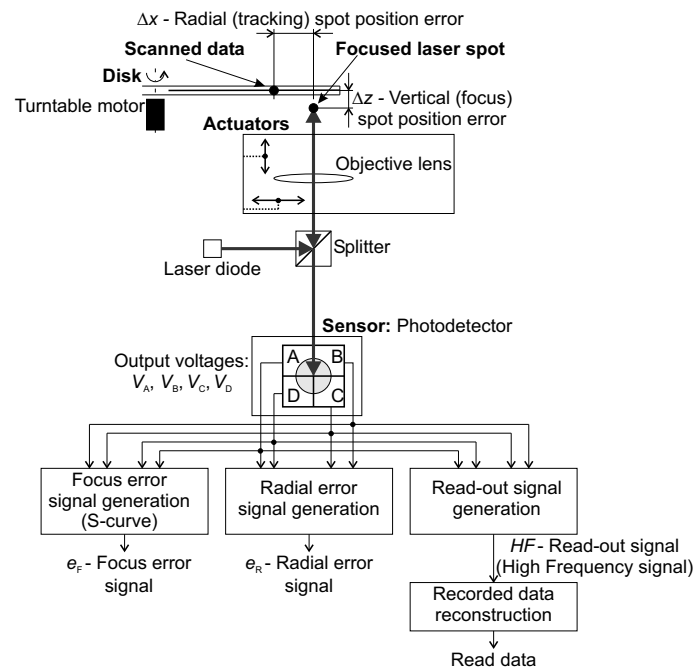


Figure 2: Block scheme of the DVD optical system.

amenant le *chariot* au voisinage de la piste recherchée, voir fig. 3.12, il s'agit là d'une procédure de saut.

- **Une boucle de suivi de piste fin (haute Fréquence)** pour maintenir le faisceau LASER sur la piste désirée.

Les deux boucles les plus critiques sont celles de focalisation et de suivi de piste fin (plus tard appelée boucle **radiale** ou de suivi de piste).

La figure 2 donne le schéma-bloc du système optique du DVD. Les quatre photodiodes (A, B, C, D) reçoivent la lumière réfléchiée par la surface du disque et génèrent les tensions de sortie du photorécepteur V_A , V_B , V_C , V_D . Ces tensions sont utilisées pour *restituer les données enregistrées sur le disque*, mais également pour *générer les signaux d'erreur de focalisation et d'erreur de position radiale* qui permettent de déterminer la position du spot par rapport à la piste. De nombreuses méthodes pour générer les signaux d'erreurs de focalisation et de suivi de piste, à partir du faisceau LASER réfléchi, ont été développées, voir [Sta98] et [Bra98]. La minimisation de ces signaux d'erreur par le système de commande permet d'assurer la qualité du signal de lecture des données.

Deux problèmes seront traités dans cette partie afin d'améliorer les performances globales du système de lecture du lecteur DVD:

1. La modélisation de la génération des signaux d'erreur radiale, d'erreur de focalisation et de lecture.
2. La conception du système de commande pour les boucles de commande de focalisation et positionnement radial.

Modélisation de la génération des différents signaux d'erreur

Pour l'analyse et la synthèse de la commande de la *boucle de focalisation*, une des étapes particulièrement importante est de déterminer la relation qui lie le signal d'erreur de focalisation $e_F(\Delta z)$ à l'erreur de position verticale Δz du spot.

Ce signal d'erreur en focalisation peut être obtenu s'il existe une asymétrie dans le chemin optique du faisceau laser. De nombreux principes ont été proposés pour mesurer l'erreur de positionnement dans le cas des lecteurs CD (*e.g.* la méthode astigmatique, la méthode de l'angle de Foucault, la méthode de l'angle de réflexion critique, *etc.*) [BBH⁺85]. Dans le cas des lecteurs DVD, seule la méthode astigmatique est largement utilisée. Son principe n'est, en général, expliqué que très schématiquement [BBH⁺85], [NO92], [Poh95], [Sin88], [Sta98] et [Wil94]. Au final, on montre que la relation entre le signal d'erreur de focalisation $e_F(\Delta z)$ et l'erreur de position verticale Δz est une fonction non-linéaire, connue sous le nom de *S-curve*.

Malgré son intérêt pratique, il n'existe pas, à notre connaissance, de modèle analytique ou numérique de la fonction *S-curve*, utilisable pour la synthèse de la commande. En pratique, une approximation très grossière de la relation entre le signal d'erreur et l'erreur réelle, est utilisée. En effet, développer d'un modèle numérique ou analytique fait appel à beaucoup de notions complexes de la théorie de l'optique et nécessite des calculs laborieux. Pour cette raison, un des objectifs de ce travail est de présenter une nouvelle méthode pour modéliser la génération du signal d'erreur de focalisation.

Dans la *boucle de commande radiale*, le point clé est cette fois la méthode de génération du signal d'erreur radiale. La relation entre le signal d'erreur de position radiale du spot $e_R(\Delta x)$ et l'erreur de position radiale Δx effective est aussi une fonction non-linéaire. Des méthodes particulières ont été développées [BBH⁺85], [Sta98] et [Poh95] pour mesurer l'erreur Δx .

La méthode la plus utilisée dans les lecteurs DVD est une variante des méthodes de Détection de la Différence de Phase. Similairement au problème relatif au signal d'erreur de focalisation, seulement quelques études de cette variante sont disponibles, voir [Bra98] et références incluses. Cette méthode sera donc présentée et analysée dans cette partie.

Les données numériques enregistrées sur le DVD sont restituées à partir du *signal de lecture*, disponible si le faisceau laser est correctement focalisé et suit avec une grande précision la piste. Pour étudier la qualité de ce signal, de nombreux articles sont disponibles, voir *e.g.* [MCW⁺96], [Mil98], [MU99], [UAM⁺00] et [UM01]. Ils traitent des perturbations induites par la structure du disque (interférences radiales, interférence inter-symbole, et signal de gîte résultant). Les algorithmes proposés sont souvent complexes mais permettent d'analyser la compatibilité entre lecteurs CD et lecteurs DVD. Deux méthodes basées sur la théorie de la diffraction seront brièvement résumées. Elles permettent de modéliser la génération du signal de lecture.

La première méthode utilise les séries de Fourier pour obtenir un modèle symbolique de la génération du signal de lecture et étudier sa qualité [Hop79]. La deuxième méthode est basé sur un modèle numérique global, obtenu en utilisant la transformée de Fourier rapide (FFT) pour résoudre les équations de propagation du champs optique entre deux plans arbitraires, voir [BW87], [Sta86] et [Nut97].

Les objectifs de cette sous-partie sont donc les suivants:

- Trouver une nouvelle méthode de modélisation de la génération du signal d'erreur de focalisation.
- Présenter la méthode de modélisation de la génération du signal d'erreur en position radiale.
- Présenter les méthodes de modélisation de la génération du signal de lecture.
- Analyser l'influence des variations de paramètres, sur la qualité du signal d'erreur de focalisation.

Conception du système de commande

Les spécifications de performances imposées aux systèmes de lecture des supports optiques sont de plus en plus strictes. Leur respect doit passer par une amélioration de l'asservissement du système de lecture. D'une part, le contrôle de la position du faisceau laser doit être plus précis pour faire face aux vitesses de rotation et aux densités de stockage croissantes. D'autre part, le système de commande doit être tolérant aux variations de paramètres, liées à la production de masse [VSA⁺01], et rendre le système insensible aux perturbations extérieures telles que les chocs et les vibrations.

La commande par retour, basée sur une description mathématique du comportement du système, est l'outil généralement utilisé pour satisfaire à ces contraintes. L'expérience montre que la synthèse de la commande est plus fiable et plus robuste si elle est réalisée à partir d'une connaissance précise de la dynamique du procédé à contrôler, plutôt qu'une connaissance moins bonne et l'utilisation d'une méthodologie d'ajustements successifs.

La conception du système de commande sera donc basée sur les modèles mathématiques présentés dans la première partie. L'amélioration des performances obtenue grâce aux correcteurs proposés répond à la fois aux exigences industrielles et au souci de proposer une solution innovante dans le cadre d'un travail de recherche. Notamment, une des fortes contraintes industrielles concerne l'implémentation du correcteur. En effet, la procédure de conception du correcteur doit fournir des correcteurs de complexité réduite.

Des techniques de conception de systèmes de commande pour les mécanismes de lecture de CD ont déjà été proposées dans différents articles ou thèses ([Det01], [DPS98], [DS99], [DS00], [DS02], [dCdH96], [DSdH95], [KON95], [Sta99], [Lee98], [SGSB94], [SSB96], [Ste00], [VSA⁺02] et [VSA⁺03]), mais peu de travaux ont été publiés en ce qui concerne les techniques de conception de commande appliquées au lecteur DVD.

Les filtres coupe-bande (notch filters) sont utilisés pour l'estimation de la vitesse de rotation, afin de concevoir des schémas de commande rejetant sélectivement les perturbations bandes-étroites [BS98]. L'article [ZKS02] présente un système de commande par mode glissant pour traiter les chocs et les vibrations. Le travail [YCC⁺02] propose une commande robuste basée sur un observateur du signal d'erreur en suivi de piste. Enfin, une nouvelle méthode est proposée dans [KYT⁺03]. La commande de type feedback/feedforward est alors basée sur l'annulation de l'erreur de phase.

La plupart de ces méthodes conduisent à des correcteurs de structure complexe. Des composants particuliers doivent être utilisés pour les implémenter - DSP (Digital Signal Processeur) haute performance. Il existe cependant des études récentes sur les correcteurs d'ordre réduits appliquées aux lecteurs de DVD ([BDCS01], [Fil03]).

L'étude [BDCS01] propose une architecture de suivi de piste utilisant un filtrage de type notch et une commande multi-échantillonnée, tandis que la thèse [Fil03] utilise une description mathématique du procédé et la commande robuste. L'amélioration des performances et l'atténuation des perturbations sont obtenues par une synthèse de correcteur de type norme. L'influence des variations paramétriques sur les performances du système et la robustesse sont analysées par la théorie μ .

Dans ce contexte, l'objectif est ici de présenter une méthode de synthèse de correcteurs pour l'asservissement en position du faisceau laser d'un lecteur de DVD industriel, utilisant également une description mathématique du procédé. La méthodologie de synthèse proposée permet d'obtenir des correcteurs d'ordre réduit répondant aux spécifications imposées. L'amélioration des performances en focalisation/suivi et celles du rejet des perturbations sont obtenues par un placement de pôles robuste suivi d'une réduction d'ordre du correcteur. Le placement de pôles est ici bien adapté pour le rejet des perturbations répétitives dans une certaine bande passante.

Les limitations liées à l'implémentation sont évaluées en simulation, avant la phase de test. Enfin, les résultats expérimentaux, réalisés sur le banc de test industriel de STMicroelectronics, sont présentés, ainsi que la comparaison entre la solution industrielle actuelle et les correcteurs proposés dans cette étude.

Les objectifs de cette partie peuvent donc être résumés de la façon suivante:

- Trouver de nouveaux algorithmes de commande pour les boucles de commande en focalisation et suivi de piste, utilisable pour un grand nombre de lecteurs.
- Améliorer les performances du système de commande, malgré les contraintes d'implémentations et les variations possibles du système de lecture, liées aux aléas de fabrication.

Organisation de la thèse

Partie I

Le chapitre 2 est divisé en cinq parties. Dans la première, les objectifs de l'identification sont décrits. Dans la deuxième, les algorithmes d'identification en boucle fermée des systèmes monovariables (SISO) sous forme de représentation d'état et basés sur l'erreur de sortie sont présentés. Ces algorithmes sont étendus, dans la troisième partie, aux systèmes multivariables (MIMO). Dans la quatrième partie, les propriétés des algorithmes présentés sont données. Des résultats expérimentaux pour des systèmes monovariables concluent ce chapitre. D'autres résultats expérimentaux sont donnés en annexe B, pour des systèmes monovariables et multivariables.

Partie II

Modélisation de la génération des signaux d'erreur

Dans le chapitre trois, le mécanisme général de commande des lecteurs DVD est présenté. Il se compose d'une partie *optique* qui reconstitue les données et génère les signaux d'erreur, et d'un *système de commande mécanique*, pour l'asservissement de la position du faisceau laser par rapport à la surface du disque.

Puisqu'un modèle précis de la génération des signaux d'erreurs est utile pour construire un meilleur simulateur du système de commande, le principe du système optique et les modèles simplifiés connus (analytiques et numériques) sont présentés. Ensuite, les problèmes de modélisation des photodétecteurs sont exposés, la formulation du problème de commande des lentilles optiques ainsi que les contraintes liées à l'implémentation sont décrites.

L'annexe C traite de la modélisation des caractéristiques non linéaires des photodétecteurs de lecteurs DVD, en utilisant les théories de l'optique. Tout d'abord, des bases théoriques de l'optique y sont rappelées. Un modèle analytique et un modèle numérique de la génération du signal d'erreur de focalisation sont développés. Ils sont basés sur la méthode astigmatique et sur l'analyse opto-géométrique. Ensuite, la validité du modèle est illustrée par une comparaison avec un signal d'erreur de focalisation réel, enregistré lors d'une procédure de démarrage (de boucle ouverte à boucle fermée).

Le modèle de génération du signal d'erreur radiale est présenté en annexe D et le modèle de génération du signal de lecture est donné en annexe E.

Conception du système de commande

Le chapitre 4 traite de la conception du système de commande. En premier lieu, des modèles simples de perturbations répétitives sont donnés, ainsi que les bases théoriques de la méthode de placement de pôle. Les propriétés physiques des actionneurs électromécaniques des lecteurs DVD et les modèles physiques du procédé sont brièvement présentés. En deuxième lieu, la synthèse d'un correcteur numérique à deux degrés de liberté (RST) est réalisée, en imposant des gabarits sur les fonctions de sensibilité du système en boucle fermée. Ces gabarits sont fixés à partir des spécifications des performances à atteindre et les fonctions de sensibilité sont calculées à partir d'un modèle simplifié du procédé. Après une réduction d'ordre du correcteur, la solution obtenue est implémentée et testée sur le système industriel. Finalement, les résultats sont comparés aux solutions actuellement implémentées. Une analyse simple des incertitudes paramétriques clôt ce chapitre.

La conclusion et les perspectives de ce travail sont présentées dans le chapitre 5.

La structure de la thèse est représentée schématiquement en figure 3.

Résumé du chapitre 2

Ce chapitre introduit une nouvelle approche au problème d'identification en boucle fermée sous forme de représentation d'état. Dans cette approche, la famille des algorithmes de l'erreur de sortie en boucle fermée est considérée comme point de départ.

Ces algorithmes permettent l'estimation des paramètres d'un modèle de type boîte-noire ordinaire, en utilisant un schéma d'identification en boucle fermée. Ceci est d'une importance particulière pour la commande adaptative. Il a été montré dans [LK97b] que les algorithmes d'identification en boucle fermée sont capables de fournir un meilleur prédicteur que ceux opérant en boucle ouverte.

Dans le cadre de ce travail, l'approche de l'erreur de sortie en boucle fermée est étendue afin d'identifier le comportement des systèmes sous forme d'un modèle d'état en temps discret avec une structure donnée. Ce sujet est traité de manière très générale.

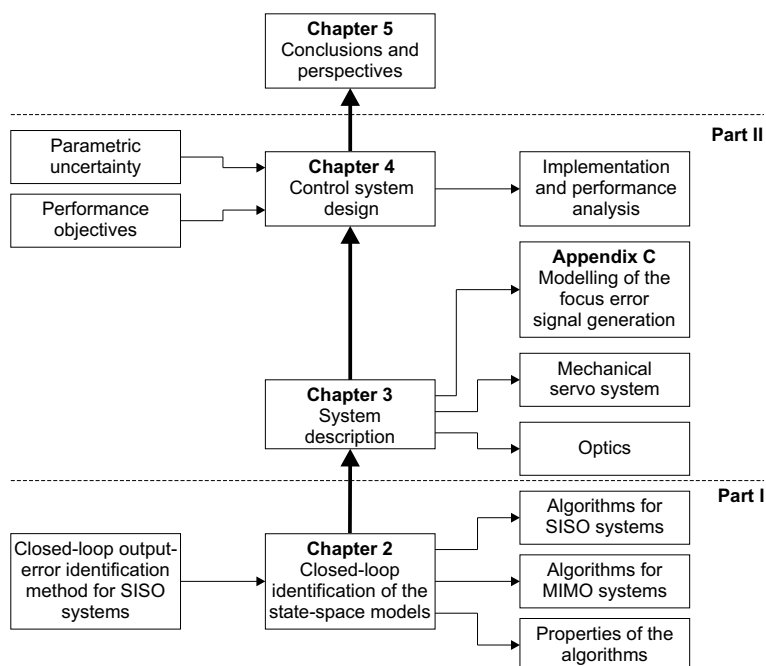


Figure 3: Thesis outline.

Plusieurs nouveaux algorithmes récursifs d'adaptation des paramètres sont proposés. L'algorithme RLS-2 est dérivé du critère des moindres carrés. Il incorpore une composante d'optimisation numérique représentée par le calcul de la pseudo-inverse d'une matrice. Alors que les algorithmes IGM-1, IGM-2 et GM-2 utilisent la technique du gradient pour minimiser un critère quadratique. Ils ont l'avantage d'avoir une forme analytique bouclée.

L'étude du comportement de ces nouveaux algorithmes a été effectuée de manière extensive en simulation sur des procédés mono et multivariables. Les résultats et interprétations font l'objet des paragraphes 2.9, B.2 et B.3. Il est monté qu'en général, l'algorithme RLS-2 converge le plus rapidement. La force des algorithmes IGM-1 et IGM-2 repose dans leur robustesse, *i.e.* dans leur insensibilité aux réalisations du bruit et aux conditions initiales (réglage du gain d'adaptation et valeur initiale du vecteur des paramètres à estimer).

Tous ces algorithmes ont les inconvénients classiques des méthodes de gradient itératives. Leur convergence et leur stabilité ne sont pas garanties et ils ne sont pas capables d'éviter le piège des minima locaux. La stabilité dans un environnement déterministe, la convergence paramétrique et la distribution en fréquence du biais asymptotique sont analysées. Ces propriétés sont liées à la valeur courante du modèle d'état et à sa paramétrisation.

Le cas particulier de paramétrisation, où la fonction de transformation est linéaire par rapport au vecteur des paramètres, est analysé de façon détaillée. Les propriétés des différents algorithmes sont exposées pour cette paramétrisation.

De cette étude, on peut conclure d'une part, que la représentation d'état canonique est le meilleur cas théorique, et d'autre part, que l'observabilité de chaque paramètre à estimer doit être considérée pour la construction de la paramétrisation du modèle d'état. Cette dernière dépend en général de deux conditions: L'observabilité des coefficients de la

fonction de transfert et la détermination des paramètres estimés à partir des coefficients de la fonction de transfert.

Résumé du chapitre 3

Ce chapitre présente une description du lecteur DVD, propose une modélisation des caractéristiques du photodétecteur et introduit la problématique de la commande.

Dans un premier temps, l'architecture du système de lecture est présentée et une description détaillée des optiques permet de clarifier les principes utilisés pour la génération du signal de lecture et des signaux d'erreur en position, nécessaires à la commande. La modélisation effective de la génération du signal d'erreur en focalisation, du signal d'erreur en position radiale et celle du signal de lecture sont données en annexes C, D et E.

La seconde partie de ce chapitre est consacrée à la description du système électromécanique de commande, à la définition des sources de perturbation et à la définition du problème de commande. L'objectif principal est de limiter l'amplitude des signaux d'erreur en position du spot (Δz , Δx) dans les directions radiales et verticales (z , x), en présence de perturbations périodiques dont la fréquence varie avec la vitesse de rotation du disque.

Résumé du chapitre 4

Ce chapitre présente la méthodologie de synthèse de la commande du système de lecture du lecteur DVD (focalisation et du suivi de piste). L'objectif est d'améliorer les performances de la commande vis-à-vis du rejet des perturbations périodiques et d'améliorer la robustesse aux aléas de fabrication.

L'étude montre que les perturbations périodiques sont principalement dues à la géométrie du disque. Des modèles simples de perturbation sont donnés avant la phase de synthèse de la commande.

La méthodologie de synthèse du correcteur utilisée est le placement pôle robuste avec calibrage des fonctions de sensibilité. Les gabarits de performance imposés aux fonctions de sensibilité sont obtenus à partir des spécifications du cahier des charges de l'industriel. Une réduction d'ordre du correcteur est effectuée pour rendre possible l'implémentation du correcteur numérique. Un ensemble d'incertitude sur le modèle, basé sur une description paramétrique, est considéré, afin d'étudier l'influence des variations des paramètres physiques du procédé, sur la robustesse en performance et en stabilité de la solution proposée. Une analyse simple de robustesse montre que le système reste stable et les performances sont conservées pour des incertitudes importantes sur les paramètres des actionneurs.

Comparée au correcteur actuellement implémenté, la solution proposée améliore les performances et la robustesse du système. Les résultats obtenus montrent, en outre, que les limites en terme de réduction d'ordre du correcteur sont atteintes, pour les performances spécifiées et les contraintes d'implémentation actuelles.

Chapitre 5: Conclusion

Rappels

Les trois principales motivations de cette thèse se résument de la façon suivante:

1. Développer et analyser les propriétés des méthodes d'identification en boucle fermée sous forme de représentation d'état.
2. Trouver les outils théoriques de l'optique permettant de développer les modèles de la génération des signaux d'erreur en position dans la partie optique des lecteurs DVD.
3. Mettre en place une méthodologie conception, issue de l'automatique avancée, pour la synthèse de la commande du système optique d'un lecteur de DVD industriel.

Contributions de cette thèse

Compte-tenu de la structure présentée ci-dessus, les aspects suivants ont été traités dans ce travail:

Première Partie - Identification

- Différents algorithmes récents d'adaptation récursive des paramètres ont été présentés pour l'identification des modèles de procédés sous forme de représentation d'état discrète avec une structure donnée (paramétrisation). Ils font partis de la classe des algorithmes de l'erreur de sortie et peuvent être considérés comme des algorithmes de type régression pseudo-linéaire récursive. L'algorithme RLS-2 incorpore une composante d'optimisation numérique représentée par le calcul de la pseudo-inverse d'une matrice. Alors que les algorithmes IGM-1, IGM-2 et GM-2, qui utilisent tous la technique du gradient pour minimiser un critère quadratique, ont l'avantage d'avoir une forme analytique bouclée.
- En général, l'algorithme RLS-2 converge le plus rapidement. La force des algorithmes IGM-1 et IGM-2 repose dans leur robustesse, *i.e.* dans leur insensibilité aux réalisations du bruit et aux conditions initiales (réglage du gain d'adaptation et valeur initiale du vecteur des paramètres à estimer). De cette étude, on peut conclure l'algorithme RLS-2 offre les meilleures propriétés.
- Les conditions suffisantes de stabilité dans un environnement déterministe et de convergence dans un environnement stochastique sont toujours liées à la valeur courante du modèle d'état et à sa paramétrisation. Seulement le cas particulier de paramétrisation, où la fonction de transformation est linéaire par rapport au vecteur des paramètres, a été analysé de façon détaillée. Les propriétés des différents algorithmes ont été exposées pour cette paramétrisation. Elles sont liées à une condition réelle positive d'une fonction de type fonction de sensibilité. Cette condition peut être relaxée par un filtrage des données ou une adaptation proportionnelle. De cette étude, on peut conclure que la représentation d'état canonique est le meilleur cas théorique.

- Il a été montré que les algorithmes développés conduisent, d'une part, à une distribution du biais non influencée par le bruit et contenant, d'autre part, un filtre de pondération fréquentielle implicite, relié à un critère de performance de la commande robuste. Ces propriétés font de ces algorithmes des outils bien adaptés à l'identification des systèmes pour la commande.
- Les parties théoriques nouvelles des algorithmes développés et les expérimentations sur des systèmes mono et multivariables ont été présentées dans le chapitre 2 et dans l'annexe B. Les résultats de cette étude ont été publiés dans [BBVH03] et [HBVBP04], respectivement.

Deuxième Partie - Lecteur DVD: Modélisation et commande

Modélisation de la génération des signaux d'erreurs de lecture et de focalisation radiale et verticale

- Une nouvelle approche a été adoptée pour la modélisation des propriétés des lecteurs DVD. Un modèle analytique et un modèle numérique de la génération du signal d'erreur de focalisation ont été développés en se basant sur la méthode astigmatique et une analyse opto-géométrique.
- Pour estimer les paramètres inconnus du modèle, une méthode d'ajustement de courbe ("curve fitting") est appliquée, en utilisant des données mesurées sur un lecteur de DVD industriel.
- Les performances du modèle analytique sont inférieures à celles du modèle numérique, mais ce modèle est intéressant car il s'agit d'un *modèle analytique complet*. L'obtention d'un tel modèle est d'autant plus importante d'un point de vue identification. En effet, les paramètres non-mesurables ou difficilement mesurables peuvent être identifiés beaucoup plus rapidement que dans le cas du modèle numérique. Ce modèle analytique est particulièrement utile dans le cas de la boucle commande de focalisation du lecteur DVD. De plus, le modèle analytique de la fonction "S-curve" $e_F(\Delta z)$ permet de réduire de façon significative les temps de simulation.
- La comparaison avec les données réelles, enregistrées sur un lecteur de DVD pendant une procédure de démarrage, illustre la qualité du modèle analytique.
- La partie théorique des modèles développés et les résultats expérimentaux sont discutés dans l'annexe C. Les résultats de cette étude ont été présentés dans [HBVSdF02], [HBVSd03] et [HVSd04].
- Le modèle de la génération du signal d'erreur en position radiale est présenté dans l'annexe D, le modèle de génération du signal de lecture dans l'annexe E. Ces modèles sont utiles pour tester la fonctionnalité du lecteur de DVD et pour les améliorations futures de ce système.

Conception du système de commande

- La conception d'un nouveau système de commande, basée sur une méthodologie associant placement de pôle et mise en forme des fonctions de sensibilité, a été proposée pour la synthèse de correcteurs d'ordre réduit. Les correcteurs obtenus améliorent les performances des boucles de focalisation et de suivi de piste et permettent, notamment, de réduire l'effet des perturbations répétitives.
- La réduction d'ordre des correcteurs est effectuée pour permettre l'implémentation pratique de ceux-ci.
- Un ensemble d'incertitude sur le modèle basé une description paramétrique est considéré, afin d'analyser l'influence des variations des paramètres physiques du procédé sur les performances et la robustesse de la solution proposée. Une analyse simple de la robustesse montre que les correcteurs synthétisés conduisent à des boucles restant stables même pour des grandes variations sur les paramètres physiques des actionneurs et les performances sont conservées.
- Cette méthodologie de conception du système de commande peut être appliquée à d'autres familles de lecteurs de DVD (DVD-ROM, DVD-audio, DVD enregistrables) ou à la nouvelle génération de lecteurs DVD haute densité, à laser bleu.
- La conception du système de commande et les résultats sont donnés dans le chapitre 4. Les résultats sont publiés dans [HBVFSd03a], [HBVFSd03b], [HVP03], [HBVF04] et [HVFS04].

Cette thèse répond donc aux objectifs fixés en introduction.

Perspectives

Première Partie - Identification

Dans cette thèse, le cas particulier de paramétrisation, pour lequel la fonction de transformation de l'espace d'état à l'opérateur de transfert entrée-sortie est linéaire vis-à-vis du vecteur paramétrisation, a été traité. Ensuite, les algorithmes proposés ont été étendus aux systèmes multivariables. Néanmoins, la fonction de paramétrisation est donnée d'une manière très générale. Le développement d'une méthodologie plus générale pour la construction d'une paramétrisation convenable des modèles en représentation d'état pourrait donc être une voie intéressante à explorer.

De plus, l'estimation des paramètres physiques via l'identification d'un modèle d'état en temps discret introduit une transformation non-triviale entre modèle en temps continu et modèle à temps discret. Pour cette raison, des algorithmes d'optimisation plus sophistiqués doivent être utilisés pour minimiser un critère non-linéaire dépendant des paramètres du modèle.

Deuxième Partie - Lecteur DVD: Modélisation et commande

Modélisation de la génération des signaux d'erreurs et de lecture

Les modèles proposés sont un compromis entre complexité et précision. Une analyse

complète de la propagation en terme de faisceau gaussien a été effectuée analytiquement pour améliorer la précision des modèles proposés, autour de la position de focalisation.

Des aspects intéressants liés à la diffraction au niveau des couches du disque optique comme, par exemple, les interférences entre signaux de suivi de piste et de focalisation et les perturbations de signaux dans les supports multicouches, pourrait être étudiés. Pour ce, la compréhension des théories complexes de la diffraction et l'accès à un environnement de simulation sont nécessaires.

Enfin, des aberrations d'ordre supérieur de la lentille astigmatique et des lentilles sphériques utilisées dans l'unité de lecture optique et leurs tolérances optiques pourraient être incluses dans des modèles plus complexes. Un compromis entre complexité du modèle et précision doit être fait en raison de la complexité et le temps de la simulation.

Conception du système de commande

Cette thèse a montré que la synthèse de la commande basée sur un modèle et appliquée au système physique que représente le lecteur DVD, nécessite:

1. La connaissance des perturbations affectant le système.
2. La connaissance des incertitudes sur le modèle pour évaluer la robustesse en performance.
3. L'utilisation d'une méthodologie systématique pour synthétiser un correcteur robuste.

Dans ce travail, les incertitudes relatives à une à la modélisation, plus ou moins précise, de l'actionneur électromécanique du lecteur ont été considérées.

Cependant, les performances des boucles de focalisation et de suivi de piste sont aussi fonction de la précision du correcteur, limitée par l'implémentation même de celui-ci. La méthodologie de synthèse du correcteur pourrait donc être étendue pour prendre en compte cette restriction.

Un autre problème, souvent rencontré dans la production de masse des lecteurs optiques, est la sensibilité du comportement dynamique du système aux conditions d'utilisation et au vieillissement. Il existe un besoin net de modèles prenant en compte ce type de modifications de comportement, pour un grand nombre de systèmes, afin d'améliorer la robustesse du système de commande.

Dans ce travail, l'ensemble d'incertitude sur le modèle, pour l'analyse de l'influence des variations de paramètres sur la robustesse et les performances, est basé sur les spécifications techniques d'un type d'unité de lecture du lecteur de DVD-vidéo industriel. Une étude plus approfondie des incertitudes paramétriques mises en jeu dans les unités de lecture serait utile pour améliorer la robustesse.

Feedback control is a widely accepted and frequently used technique in order to enforce a dynamical system to behave in a satisfactory manner. The broad concept of a dynamical system can thereby refer to many biological or engineering processes. The application of automatic control can be found in many complex industrial processes or sophisticated mechanical systems to attain a properly operating dynamical system.

Next to feedback control, system identification is used repeatedly to elucidate the dynamical aspects of the system. This procedure enables one to predict the dynamical behavior of an unknown system on the basis of foregoing observations of the system. In this way, knowledge of the dynamical aspects of an industrial process or of a mechanical system is acquired on the basis of experiments and construction of adequate mathematical models. Models are of rapidly increasing importance in engineering and today all control designs are more or less based on them. Models are also extensively used in other, nontechnical areas such as biology, ecology, and economy.

If the physical laws governing the behavior of the system are known we can use them to construct so called *white-box models* of the system. In white-box model, all parameters and variables can be interpreted in terms of physical entities and all constants are known *a priori*. The opposite extreme is known as *black-box modelling*. In this approach the models are constructed from measured input-output data. *A priori* knowledge about the model structure is eliminated and therefore the model and the parameters of it usually have little physical significance.

Linear model structures have been widely used for black-box modelling because they are mathematically attractive. This class of models is rich enough to cover a large number of applications, since systems are often controlled around an operating point and can be considered to behave linearly for small variations of the input.

A third approach is a combination of two extremes, and is called *grey-box modelling*. This approach exploits the *a priori* physical knowledge about the process, but the model structure and the parameters are not assumed to be completely known. The parameters of the model are estimated as for black-box models, by using identification methods. A typical grey box situation is when the model structure is determined by the physical relations of the process, hence the term *physical models*. In this situation, the identification is done with a fixed model structure (only the unknown parameters are estimated).

This thesis focuses on both control design and system identification, with the aim to design a control for a given system on the basis of *white-box model*. This approach is partly illustrated for an industrial Digital Versatile Disk-video player (DVD-video player), a high accuracy opto-electro-mechanical positioning system.

The thesis contains two main parts: The first part deals with theoretical and methodological part on *system identification* in closed-loop of black-box models using state-space representation to define a model structure. The second part of this thesis presents one case-study part on the *modelling and control system design* of a DVD player.

1.1 Background

In this thesis we will study methods for identification of linear, time-invariant dynamical systems by using discrete-time data and given model structure. By now, several general methods exist for identifying discrete-time linear models, *e.g.* the prediction error methods [Lju99; LLM97], the subspace approaches [OM96], the extended Kalman filter [WP97]. We will focus especially on prediction error methods for identifying systems operating in closed-loop known as a Closed-Loop Output-Error (CLOE) identification method, see *e.g.* [LK97a], [LLM97] and [LK97b].

Closed-loop identification has often been suggested as a tool for identification of models that are suitable for control, so called *identification for control* [HGdB96], [Lju99], [FL99], [GdH01], [Ver93], [dHS95] and [LLM97]. It results when the identification is performed in closed-loop (the system is controlled by some feedback controller). Controllers are frequently used for changing the behavior of dynamical systems. The objective may be, for instance, to increase the speed of response of the system or to make it less sensitive to disturbances, noises. In such cases it is possible to choose to perform the identification experiment in open-loop (if the controller is disconnected) or in closed-loop. However, if the system is unstable or has to be controlled for economic or safety reasons, closed-loop experiments are unavoidable.

A problem with closed-loop data is that many of the common identification methods, that work well in open-loop, fail when applied directly to measured input-output data. This is true for instrumental variables, spectral analysis, and many subspace methods, for instance [FL99]. The reason why these methods fail is the correlation between the system input and the unmeasurable output noise, that is inevitable in a closed-loop situation.

Instead of applying open-loop type identification methods on data acquired in closed-loop, if the output noise is important, than one has to use specific identification methods for closed-loop data. Prediction-error methods can be adapted for this case, in order to deal with the input and noise correlation. One class of such new methods is based on output-error and has been developed for identifying input-output models. They are known as CLOE algorithms [LK97b].

Let us now give some background material on modelling and control system design of a DVD player.

The control problem for a DVD player is similar to the one defined for the Compact Disk player (CD player) mechanism, and it consists in guaranteeing that the laser beam used to read the data follows the track on the disk. Contrary to the vinyl disks players, where the piezo-electric or magnetic transducer (device that convert vibration into electric signal) is guided by the tracks along the disk radius, in optical disk drives only the light touches the disk. Track following should be thus guaranteed by means of feedback control, that in the current industrial applications is achieved with simple PID controllers. In the last years, however, the optical storage device mechanisms have been used for an increasing number of new applications, like the Compact Disk-Read Only Memory (CD-ROM), Photo and Video CD, Digital Versatile Disk-Read Only Memory (DVD-ROM) and Video DVD. These new applications require higher performance levels than the original audio system conceived for CD, since higher data density on the disk and shorter data access time are demanded [ECM96], [ECM01].

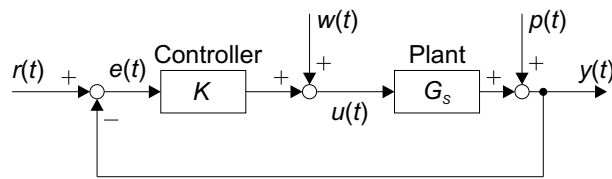


Figure 1.1: Closed-loop scheme for SISO systems.

Up to now, in the DVD-video player devices, the desired levels of performance have been usually achieved by heuristic control design and tuning, and through expensive improvements in the system manufacturing. This is partially caused by the fact that models of the relevant signals generation are very complex, as in [BW87], [BBH⁺85], and [Bra98].

Therefore, the proposed models in this thesis make a compromise between complexity and accuracy. In this work we will investigate the possibility of improving the control system performance using physical models.

1.2 Motivation and Goals of the Thesis

The material in this thesis divides naturally into two parts. Part I concerns identification in closed-loop of black-box models using state-space representation to define a model structure and part II focuses on the physical system description of the DVD-video player, mathematical analysis of servo and read-out signals generation together with control system design.

Part I

Fig. 1.1 illustrates a general closed-loop scheme for SISO systems, and it is composed of the controller K and plant model G_s ; $r(t)$ is the reference, $e(t)$ is the controller input, $u(t)$ is the controller output, $y(t)$ is the plant output, $p(t)$ is the output disturbance and $w(t)$ is the external excitation that can be added to the controller output $u(t)$ (or the reference $r(t)$) for the system identification.

To identify systems operating in closed-loop, see fig. 1.1, in case of given model structure, algorithms for discrete-time state-space model have been developed recently in [Bez01] and [Bez04]. They are based on the Closed-Loop Output-Error (CLOE) identification algorithms developed for black-box Single-Input Single-Output (SISO) systems, see *e.g.* [LK97a] and [LLM97], used successfully in closed-loop scheme.

Nevertheless, the analysis of their properties has not been done there. The aim of this work is to present these algorithms and analyze their properties. Moreover, several simulation cases also have been studied here.

The original objective of the thesis was also to identify the unknown physical parameters of a DVD player by these algorithms in order to design new controllers. Nevertheless, this approach was not successful in this case due to the constraints of the identification experiment on the industrial DVD player benchmark, see [Fil03], and the identification algorithms useability. For this reason, alternative plan has been applied: The linear

white-box model of the DVD player actuator obtained from the specification has been used in control design procedure.

Despite the fact that the developed and analyzed algorithms did not lead to the new control system design for the DVD player, in general, they are new and applicable on other examples, and therefore they are presented in this thesis.

The goals of this part can be summarized as follows:

- Present new algorithms for the identification of discrete-time models, given by state-space representation, in closed-loop.
- Analyze the properties of these algorithms.
- Apply the algorithms on SISO and MIMO simulated systems.

Part II

The need for a real-time control of the objective lens position is imperative for applications such as reading digital information from an optical medium as a DVD. The laser beam which is used to read the recorded data from the disk must be focused on its surface and follow the track very accurately. This task is accomplished by the position control loops which minimize the position error signals generated by the optical device.

In general, three main control loops can be distinguished in DVD players:

- **A focus loop** to ensure the distance between the objective lens and the disk, see fig. 3.12.
- **A coarse (low frequency) tracking loop** to position the Optical Pick-up Unit (OPU) roughly pulling the so-called **sledge** in the vicinity of the desired tracks, see fig. 3.12; this is done by a long jump procedure.
- **A fine (high frequency) tracking loop** to lock the focused laser beam onto the track position.

The most critical control loops are particularly the focus and the fine tracking loop (called shortly the **tracking** or **radial loop** in next text).

Fig. 1.2 shows the block scheme of the DVD optical system. The four photodiodes (A, B, C, D) receive the light reflected from the disk surface and generate the output voltages of the photodetector V_A , V_B , V_C , V_D that are used to retrieve both *data recoded on the disk*, and *position error signals (focus and radial errors)*, used to measure the spot position with respect to the track position. Many focus/tracking methods to generate focus/track error signal from the reflected laser beam have been developed, see *e.g.* [Sta98] and [Bra98]. These error signals are minimized by the controllers in corresponding closed-loops in order to ensure the read-out signal quality.

In this thesis we will study two problems of DVD players to improve the system performance:

1. Modelling the focus error, radial error and read-out signals generation.
2. Control system design for the focus control loop and the radial control loop.

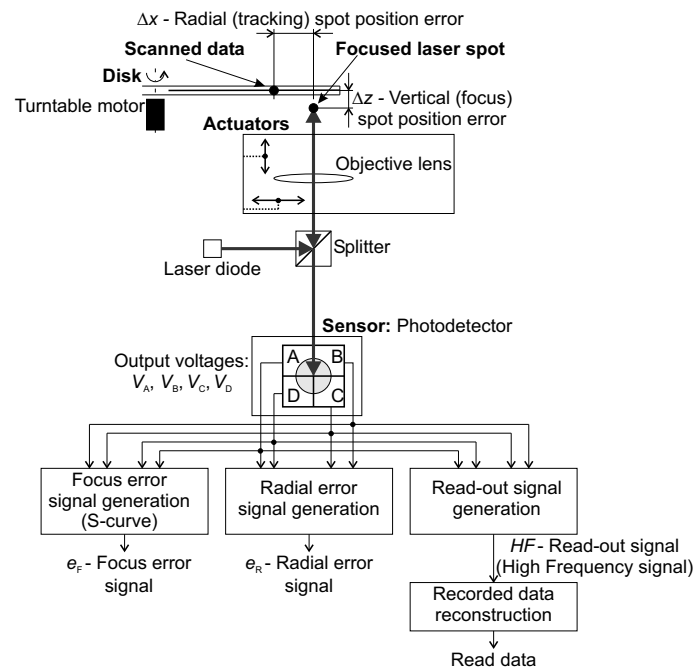


Figure 1.2: Block scheme of the DVD optical system.

Modelling the focus error, radial error and read-out signals generation

In the *focus control loop*, one very important element is the dependence between the focus error signal $e_F(\Delta z)$ and the vertical spot position error Δz which is a non-linear function known under the name “S-curve”.

A focus error signal can be obtained if some asymmetry is present in the optical path. Numerous schemes have been proposed to measure the spot position error for CD players (*e.g.* the astigmatic method, Foucault knife edge method, critical angle reflection method, *etc.*) as in [BBH⁺85]. However, only the astigmatic method is widely used in DVD players. Nevertheless, its principle is often explained, only schematically, as in [BBH⁺85], [NO92], [Poh95], [Sin88], [Sta98] and [Wil94].

Despite its present practical importance, to our knowledge, no useful model (analytical or numerical) of the S-curve is available. Practically, a very rough approximation of the function between the focus error signal and the vertical spot position error is used. Indeed, the development of an analytical or numerical model of this function involves a large amount of the optical theory and some tedious computations. For this reason, the aim of this work is to present a new method for modelling the focus error signal generation.

In the *radial control loop*, a key element is the radial error signal generation method. It defines the dependence between the radial error signal $e_R(\Delta x)$ and the radial spot position error Δx which is also a non-linear function. Special methods like those presented in [BBH⁺85], [Sta98] and [Poh95] have been developed.

Currently, one of the most widely used method in DVD players is a variant of the Differential Phase Detection (DPD) methods. Despite its practical importance, to our knowledge, only few modelling approaches of this method are available, as in [Bra98] and

references therein. Therefore, the Differential Phase Detection (DPD) method that has been adopted as the standardized method for the radial error signal generation in DVD players will be analyzed in this thesis.

The recorded digital data on the DVD are retrieved from a *read-out signal* which is available if the laser beam is focused on the disk data layer surface and follows the track, both with high precision. To study the read-out signal quality, many papers have been presented, see *e.g.* [MCW⁺96], [Mil98], [MU99], [UAM⁺00] and [UM01]. They deal with the disturbances that become from the disk structure (the radial cross-talk, inter-symbol interference and the resulting signal jitter). The proposed algorithms are very complex for this study but they enable to analyze a compatibility between CD and DVD drives. In this context, our aim is to summarize briefly two useful methods, based on the diffraction theory for modelling of the read-out signal generation.

The first method uses a Fourier series to create symbolic model of the read-out signal generation to study its quality, see [Hop79]. The second method is based on the general numerical model, using Fast Fourier Transformation (FFT) to solve the optical field propagation between two arbitrary planes, see *e.g.* [BW87], [Sta86] and [Nut97].

The goals of this section can be summarized as follows:

- Find a new method for modelling the focus error signal generation.
- Present the method for modelling the radial error signal generation.
- Present the methods for modelling the read-out signal generation.
- Analyze the influence the varying model parameters on the focus error signal quality.

Control system design

The strict performance specifications imposed on optical disk drive systems demand an enhancement of tracking control behavior. In addition, due to the higher storage capacity and disk rotational speed, the spot position control system must be more accurate to cope with parameter tolerances due to mechanism mass production, as outlined in [VSA⁺01], and guarantee an insensitivity to external disturbances like shocks and vibrations.

Feedback control design, usually based on a mathematical description of system behavior, is a tool used to satisfy these requirements. From the experience, it is well known that a robust and reliable control design is achieved if an accurate knowledge of the plant dynamics is taken into account during the synthesis, rather than repeatedly applying fine-tuning methodologies.

Accordingly to the industrial objectives and to research requirements of finding innovative solutions, an enhanced control system performance is realized by the designed controllers, on the basis of a mathematical parametric model of the system. An important specification for implementation purposes is that a controller design procedure should deliver controllers of reduced complexity.

Control system design techniques applied to CD mechanisms have already been exposed in several papers and Ph.D. theses as in [Det01], [DPS98], [DS99], [DS00], [DS02], [dCdH96], [DSdH95], [KON95], [Sta99], [Lee98], [SGSB94], [SSB96], [Ste00], [VSA⁺02]

and [VSA⁺03], but only few works have been published about control design techniques applied to a DVD player.

The notch filters are used for the rotational speed estimation in order to implement a control scheme which selectively cancels narrow-band disturbances in [BS98]. The paper [ZKS02] presents a control system using sliding mode control to handle shock and vibration disturbances. The work [YCC⁺02] is devoted to robust tracking control using observer of the tracking error signal. A new tracking control method consists of a feedback controller and a feed-forward controller that employs the zero phase error tracking method is proposed in [KYT⁺03].

Most of these methods result in controllers with a complex structure and a special device (high performance Digital Signal Processing/Processor (DSP)) is necessary to implement such controllers. However, there are recently developed studies on reduced order controller, applied to an industrial DVD player, like [BDCS01], [Fil03].

The work [BDCS01] proposes a control architecture for track following using the notch filtering and multirate control. On the other hand, the thesis [Fil03] is focused on usage of a mathematical description of the plant to compute robust controllers. An enhanced tracking performance and disturbance attenuation have been obtained via norm-based control design, and the influence of drive parametric uncertainties on the system performance and robustness has been analyzed by means of μ -theory.

In the same context, our aim is to present controllers design methodology applied to an industrial DVD-video player spot positioning control system, also by using a mathematical description of the plant. This synthesis is proposed to compute restricted complexity controllers that are able to comply with tough DVD specifications. An enhanced focus/tracking performance and disturbance attenuation are obtained with pole placement design followed by controller order reduction. Pole placement method is adapted here to realize repetitive disturbance rejection in a certain bandwidth.

Design limitations are evaluated in simulation before the implementation step. Experimental results, obtained on the STMicroelectronics industrial benchmark, are finally presented to compare the current industrial solution and the computed controllers.

The goals of this section can be summarized as follows:

- Find a new control algorithms for the focus/radial control loops that may be used for a large quantity of DVD-players.
- Improve the control system performance to compensate drives manufacturing tolerances and improve disk playability under implementation constraints.

1.3 Thesis Outline

Part I

Chapter 2 is divided into five parts substantially. In the first part, the identification objectives are described. In the second part, the single-input single-output state-space versions of closed-loop output-error algorithm are presented. The algorithms are further extended in the third part to incorporate Multi-Input Multi-Output (MIMO) systems. In the fourth part, the properties of the presented algorithms are given. Experimental

results on the SISO systems conclude this chapter. Another experimental results on the SISO and MIMO systems are presented in appendix B.

Part II

Modelling the focus error, radial error and read-out signals generation

In chapter 3, a general DVD player servo mechanism is presented as composed of two main parts: The *optics* that retrieves the data and generates the error signals and the *mechanical servo system* that has to control the laser beam position with respect to the disk surface.

Since an accurate model of position error signals generation can be useful to build a more precise servo system simulator, optical system principles and known simplified analytical and numerical models of optics are presented here. Then, the modelling problem description for the photodetector and the control problem formulation for the optical lens together with the implementation constraints are described.

Appendix C treats the modelling of the non-linear characteristics of a photodetector in a DVD player by using optical theory. Firstly, theoretical background on optical theory is given. An analytical and a numerical models of the focus error signal generation are developed here based on the astigmatic method and the opto-geometrical analysis. Secondly, model quality is illustrated by a comparison with the real focus error signal acquired during the start-up procedure (from open-loop to closed-loop).

Model of the radial error signal generation is presented in appendix D and model of the read-out signal generation is given in appendix E.

Control system design

Chapter 4 deals with the control design of the DVD-video player. Firstly, simple repetitive disturbance models and theoretical background on pole placement method are given. Physical properties of the DVD electro-mechanic actuator and plant physical modelling are presented briefly. Secondly, a two degree of freedom digital controller (RST controller) design is performed by imposing frequency templates on the system closed-loop sensitivity functions together with the performance specifications and simplified plant model knowledge. After the controller order reduction, the achieved solution is implemented and tested in the industrial system. These results are finally compared to those obtained with the actual implemented industrial controller. Simple parametric uncertainty analysis closes this chapter.

Conclusions and perspectives of the work are presented in chapter 5.

The thesis structure is schematically represented in fig. 1.3.

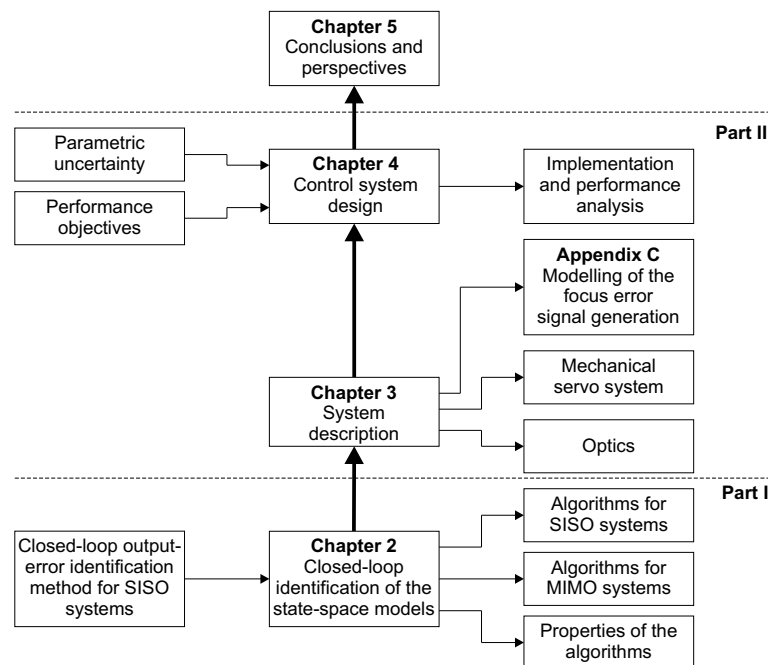


Figure 1.3: Thesis outline.

State-space Identification for SISO/MIMO Systems

2

2.1 Introduction

This chapter is devoted to the problem of state-space identification in closed-loop. The family of closed-loop output-error algorithms, presented in appendix A, is considered as a starting point.

These algorithms allow to estimate the parameters of an ordinary black-box model, utilizing the closed-loop identification scheme. In order to identify various discrete-time state-space plant models with a given structure, the closed-loop output-error method is expanded here.

Algorithms for the closed-loop identification of discrete-time state-space models are presented, using the results from [Bez01] and [Bez04]. Nevertheless, the analysis of their properties has not been done there.

Our contribution consists of the algorithm properties analysis, developed in [Bez01] and [Bez04], namely, the convexity of criterion function, the stability of Parameter Adaptation Algorithm (PAA) in deterministic environment, the parametric convergence and the frequency distribution of the asymptotic bias. The same techniques as in appendix A are used to analyze their properties. In addition, the experimental results are illustrated to compare the developed identification algorithms and study their behavior.

The structure of this chapter is the following: The description of the state-space identification problem is presented in section 2.2. In section 2.3, a brief state of the art is presented. Some theoretical background on the model structures is given in section 2.4. An overview on the optimization algorithms is described in section 2.5. The state-space identification algorithms for SISO systems are presented in section 2.6. In section 2.7 we give an extension of these algorithms for MIMO systems. The properties of the proposed parameter adaptation algorithms are shown in section 2.8. The convexity of criterion function, the stability of parameter adaptation algorithm in a deterministic environment, the parametric convergence and the frequency distribution of the asymptotic bias are studied here.

The algorithms are illustrated on simulation experiments which involve a closed-loop identification of a SISO and MIMO model. Two simulation experiments on the SISO systems are shown in section 2.9 and another experiment is presented in appendix B.1. The simulation experiments on the MIMO systems are presented in appendix B.2.

Finally, in appendix B.3, real-time experiment on the SISO system, a bipedal robot (RABBIT), is illustrated. Unfortunately, presented algorithms have not been used in this case due to the constraints of the identification experiments on the bipedal robot. Therefore, another identification technique has been applied to estimate unknown parameters of black-box model. This model will be used in future to eliminate some of the actual constraints, and therefore it is presented in this thesis.

We draw some conclusions in section 2.10.

2.2 Problem Description

For most physical systems it is easier to construct models with physical insight into continuous-time because most physical laws (Newton's laws of motion, relationships between electrical quantities *etc.*) are expressed in continuous-time. This means that modelling leads to a state-space representation:

$$\dot{x}(t) = F_0(\theta_c)x(t) + G_0(\theta_c)u(t), \quad (2.1)$$

where $F_0(\theta_c)$ and $G_0(\theta_c)$ are the matrices of appropriate dimensions ($n \times n$ and $n \times r$, respectively, for an n -dimensional state and an r -dimensional input) and θ_c is a vector of parameters that typically corresponds to unknown values of physical coefficients in continuous-time model (material constants, *etc.*). Such a selection of the parameter vector θ_c is called the *physical parametrization* of a plant model. The model (2.1) also defines the vector of state variables x which usually have the physical significance (position, velocity, charge, current *etc.*). The measured outputs are known combinations of the states.

The corresponding discrete-time state-space model can be obtained by sampling the input and output signals of the parametrized continuous-time model (2.1):

$$x(kT_s + T_s) = A_0(\theta_c)x(kT_s) + B_0(\theta_c)u(kT_s), \quad (2.2)$$

where T_s denotes the sampling period and $f_s = 1/T_s$ is the sampling frequency. Briefly:

$$x(k+1) = A_0(\theta_c)x(k) + B_0(\theta_c)u(k). \quad (2.3)$$

The discrete-time state-space matrices $A_0(\theta_c)$ and $B_0(\theta_c)$ can be expressed using the following formulae:

$$\begin{aligned} A_0(\theta_c) &= e^{F_0(\theta_c)T_s}, \\ B_0(\theta_c) &= \int_0^{T_s} e^{F_0(\theta_c)\tau} G_0(\theta_c) d\tau. \end{aligned} \quad (2.4)$$

Under certain conditions it is possible to replace this complex relationship by a first-order approximation obtained if higher-order terms in the power series expansion of the matrix exponential are neglected:

$$\begin{aligned} A_0(\theta_c) &\approx I + T_s F_0(\theta_c), \\ B_0(\theta_c) &\approx T_s G_0(\theta_c). \end{aligned} \quad (2.5)$$

Considering the formulae (2.4) or (2.5), the discrete-time representation has a disadvantage that the matrices $A_0(\theta_c)$ and $B_0(\theta_c)$ are more complicated functions of the parameter vector θ_c than the matrices $F_0(\theta_c)$ and $G_0(\theta_c)$.

Adding the output equation, the standard discrete-time state-space model is obtained:

$$\begin{aligned} x(k+1) &= A_0(\theta_c)x(k) + B_0(\theta_c)u(k), \\ y(k) &= C_0(\theta_c)x(k) + D_0(\theta_c)u(k). \end{aligned} \quad (2.6)$$

The matrices $A_0(\theta_c)$, $B_0(\theta_c)$, $C_0(\theta_c)$, $D_0(\theta_c)$ of the model (2.6) can depend on the parameter vector θ_c in different ways. A typical case is when certain elements of matrices $A_0(\theta_c)$, $B_0(\theta_c)$, $C_0(\theta_c)$, $D_0(\theta_c)$ may be known or fixed values. The reason may be that

the values of the parameters are *a priori* known physical constants; or, we would like to impose a certain structure on the model. If model (2.6) is a canonical form then the parameter vector θ_c consists of the parameters of the original input-output transfer operator, [Kai80].

Although sampling of input and output signals of a continuous-time model is a natural way to get the discrete-time model (2.6), for certain applications it could be also given directly in discrete-time, with the matrices $A_0(\theta)$, $B_0(\theta)$, $C_0(\theta)$, $D_0(\theta)$ parametrized in terms of the parameter vector θ , rather than via (2.4) or (2.5).

For this assumption, only few specific identification methods in closed-loop are available, [Mel94], [Lju99]. Therefore, new identification algorithms for the discrete-time state-space models have been developed recently, see [Bez01] and [Bez04]. Nevertheless, useful analysis of their properties is missing. This is important for their use and the following improvement.

For these reasons, our study is only focused on the particular case of the identification algorithms where the state-space matrices $A_0(\theta)$, $B_0(\theta)$, $C_0(\theta)$, $D_0(\theta)$ are parametrized in terms of the parameter vector θ . Our contribution in this chapter is an analysis of the state-space identification algorithm properties developed in [Bez01] and [Bez04].

This analysis naturally depends on the parametrization of the matrices $A_0(\theta)$, $B_0(\theta)$, $C_0(\theta)$, $D_0(\theta)$. In fact, if all matrix elements of $A_0(\theta)$, $B_0(\theta)$, $C_0(\theta)$, $D_0(\theta)$ are assumed to be unknown, uniqueness cannot hold because the input-output transfer operator of the model is the same for the different state-space realizations. The multiplicity of realizations of a given system is well illustrated by the fact that given realization $A_0(\theta)$, $B_0(\theta)$, $C_0(\theta)$, $D_0(\theta)$, [Kai80]:

$$\begin{aligned} x(k+1) &= A_0(\theta)x(k) + B_0(\theta)u(k), \\ y(k) &= C_0(\theta)x(k) + D_0(\theta)u(k), \end{aligned} \tag{2.7}$$

we can form to another realization by a change of variables (note the order: old variables = $T \cdot$ new variables):

$$x(k) = Tx_{\text{new}}(k), \quad \det T \neq 0, \tag{2.8}$$

so that, say:

$$\begin{aligned} x_{\text{new}}(k+1) &= T^{-1}A_0(\theta)Tx_{\text{new}}(k) + T^{-1}B_0(\theta)u(k), \\ y(k) &= C_0(\theta)Tx_{\text{new}}(k) + D_0(\theta)u(k). \end{aligned} \tag{2.9}$$

Since there is a multiple infinity of nonsingular matrices T , there is clearly multitude realization. Matrix T is known as *similarity matrix*.

Therefore, to get a unique model it is necessary to impose restrictions of some form on the matrix elements, *e.g.* to set a few matrix elements to constant values or to find smaller number of estimated parameters in parameter vector θ or to find the different connections between estimated parameters of the parameter vector θ , *etc.* All restrictions have one goal, to uniquely estimate the parameter vector θ of the state-space model $A_0(\theta)$, $B_0(\theta)$, $C_0(\theta)$, $D_0(\theta)$ from the model input-output transfer operator. This is known under the name *structural identifiability*.

To sum up, when identifying a model of the form (2.7), the following *a priori* knowledge is required:

- The structural parameters, *i.e.* the dimension n of the state vector, the number r of inputs and the number s of outputs.

- The parametrization itself, *i.e.* the way in which the parameter vector θ enters the matrices $A_0(\theta)$, $B_0(\theta)$, $C_0(\theta)$, $D_0(\theta)$.

2.3 State of the Art

Identification of SISO and MIMO systems has been a topic of big interest in the feedback control and signal processing literature for several years already (for some recent references, see *e.g.* [Zhu01] and [CP01] and the references therein). This is given by the still practical importance of plant model identification for: Firstly, to obtain better model for new controller design (redesign) and secondly, to identify unknown physical model parameters for the following system development.

A natural way, how to construct models with physical insight is to use the continuous-time state-space model (2.1). Nevertheless, to avoid the optimization problem during parameter vector θ_c estimation, the identification is usually performed using the discrete-time state-space model (2.6) which directly corresponds to the original continuous-time state-space model (2.1). Moreover, the measured data are acquired in discrete-time. In this case the parameter vector θ_c that corresponds to the continuous-time model unknown parameters is often transformed to another parameter vector θ , and therefore the discrete-time state-space model for identification takes the form (2.7).

To identify the discrete-time state-space model (2.7), the subspace methods based on the geometrical properties of the input and output sequences have been proposed, initially developed in the open-loop context, [OM96]. Nevertheless, the practical importance of closed-loop identification caused that they have been extended to the closed-loop case, see *e.g.* [OM97] and [Ver93]. The subspace methods are numerically robust, they do not require iterative optimization and their adjustment can be reduced to the selection of the model order. Moreover, the subspace identification methods have been recently extended, see [VV02], to deal with multivariable linear parameter-varying state-space systems. Nevertheless, the subspace methods use fully parametrized state-space models which means that all elements of the matrices $A_0(\theta)$, $B_0(\theta)$, $C_0(\theta)$, $D_0(\theta)$ are estimated parameters, [Lju99], [McK95]. In addition, usually the physical insight of each parameter is lost.

On the other hand, a different approach based on fractional representation of the SISO and MIMO systems in state of the state-space model description (2.7) has been proposed for identification in closed-loop in [dHdC96] and [dC98]. The main motivation to use fractional model representations is due to the fact that both stable or unstable systems can be represented by (the ratio of two) stable factors. The advantage of this method is ability to estimate not only a nominal model but also the sets of models tuned towards a robust control design. Nevertheless, this fractional approach does not allow to incorporate the physical insight of estimated parameters and given model structure.

Therefore, a statistic method, directly concerning to the problem of the physical SISO models identification, has been proposed in [Mel94]. The model has been based on the frame of stochastic differential equations that allow to account random phenomena explicitly in order to obtain the desired accuracy in the modelling. The identification approach has been based on the approximation of the partial differential equation by set of ordinary differential equations and an estimation of the parameters has been done in the time domain. The maximum likelihood method, [Lju99], has been considered for the

choice of the estimator for identification only in the open-loop. The advantage of this approach is that it is easy to extend it to more complex systems and even consider nonlinearities in the model, [Mel94]. Moreover, the parameters of the original equation can easily be calculated from the parameters of the ordinary differential equations afterwards if that is of interest. The main drawback of this approach is the computation efforts involved, *e.g.* the maximization of the likelihood function during optimization has to be done by using a numerical method of iterative character.

To the contrary, another approach to estimate the state-space model (2.7) known as an extended Kalman filter, [WP97], obtains popularity from its remarkable simplicity of implementation. It offers to identify the state vector $x(k)$ of the system together with the elements of matrices $A_0(\theta)$, $B_0(\theta)$, $C_0(\theta)$. It has been used in a large number of practical applications even though it presents some hard-to-comprehend divergence phenomena, [WP97]. In the special case of stationary linear input models, the introduction of corrective term to the parameters makes the convergence properties of the filter identical to those of the estimator obtained with a maximum likelihood approach, [Lju79]. However, the state-space model identification without the state estimation is not possible.

From this bibliographic search one can conclude that there is still need to find identification algorithms to estimate some system matrices elements, using the advantage of working in the closed-loop identification scheme.

For this reason, the new identification algorithms for discrete-time state-space model (2.7), using given model structure, have been developed in [Bez01] and [Bez04]. This kind of algorithms is based on the the CLOE algorithms developed for black-box SISO systems, identified in closed-loop, as presented in appendix A. Nevertheless, the analysis of their properties that is important from practical point of view is missing.

Therefore, our aim in this chapter is to analyze their properties and to compare their different variants on the simulation experiments. Before going further in treating the identification algorithms presentation, the model structure and optimization of criterion function is discussed in section 2.4 and section 2.5, respectively.

2.4 Model Structures

The model of a system is a rule to compute, from quantities known *a priori* or measured from the system, other quantities that we are interested in and which we hope will resemble their actual values in the system.

The choice of a structure for the mathematical model is called *model-structure selection*, or more consistently *characterization* and it is a crucial step in the modelling. The following sections, (2.4.1 - 2.4.3), give a brief theoretical background on the various classes of state-space models and their properties.

2.4.1 Deterministic and Stochastic State-space Models

A deterministic model describes the outputs y as if it was uniquely determined by the inputs u . This is often unrealistic, because of the various *perturbations* that act on the system or corrupt measurements. It is then necessary to find the description of these perturbations. A statistical description is usually used, where perturbations are described as stochastic process.

Let's consider the following SISO stochastic state-space model:

$$\begin{aligned} x(k+1) &= A_0(\theta)x(k) + B_0(\theta)u(k) + v(k), \\ y(k) &= C_0(\theta)x(k) + p(k), \end{aligned} \quad (2.10)$$

where $x(k)$ is an n -dimensional state vector, $p(k)$ is an output disturbance noise and $v(k)$ is an n -dimensional random quantity acting on the state vector $x(k)$:

$$v(k) = [v_1(k) \ \dots \ v_n(k)]^T. \quad (2.11)$$

Disturbance vector $v(k)$ is often called *process noise*. In addition, we assume that $D_0 = 0$, *i.e.* there is no direct connection from the input to the output.¹

Furthermore, the transfer operators $S_{yu}(q^{-1})$ and $S_{yv}(q^{-1})$ are introduced by the equations:

$$y(k) = S_{yu}(q^{-1})u(k), \quad (2.12)$$

$$y(k) = S_{yv}(q^{-1})v(k), \quad (2.13)$$

where the vector transfer operator $S_{yv}(q^{-1})$ is defined as follows:

$$S_{yv}(q^{-1}) = [S_{yv_1}(q^{-1}) \ \dots \ S_{yv_n}(q^{-1})], \quad (2.14)$$

$$y_i(k) = S_{yv_i}(q^{-1})v_i(k) \quad \text{for } i = 1, \dots, n. \quad (2.15)$$

Using (2.12) and (2.13) for the state-space representation (2.10), one can easily obtain the well-known transfer operators $S_{yu}(q^{-1})$ and $S_{yv}(q^{-1})$:

$$\begin{aligned} S_{yu}(q^{-1}) &= q^{-1}C_0(I - q^{-1}A_0)^{-1}B_0 = \frac{q^{-1}C_0 \operatorname{adj}(I - q^{-1}A_0)B_0}{\det(I - q^{-1}A_0)} = \\ &= \frac{B(q^{-1})}{A(q^{-1})}, \end{aligned} \quad (2.16)$$

$$\begin{aligned} S_{yv}(q^{-1}) &= q^{-1}C_0(I - q^{-1}A_0)^{-1} = \frac{q^{-1}C_0 \operatorname{adj}(I - q^{-1}A_0)}{\det(I - q^{-1}A_0)} = \\ &= \left[\frac{B_1(q^{-1})}{A(q^{-1})} \ \dots \ \frac{B_n(q^{-1})}{A(q^{-1})} \right]. \end{aligned} \quad (2.17)$$

One can clearly see that the transfer operators $S_{yu}(q^{-1})$ and $S_{yv_i}(q^{-1})$ have a common denominator $A(q^{-1})$.

The aim of this section is to analyze the way to transform the process noise $v(k)$ of the state-space model (2.10) to the output $y(k)$. It leads to the modified output disturbance noise $p(k) \rightarrow p_{\text{new}}(k)$ of the following state-space model:

$$\begin{aligned} x(k+1) &= A_0(\theta)x(k) + B_0(\theta)u(k), \\ y(k) &= C_0(\theta)x(k) + p_{\text{new}}(k), \end{aligned} \quad (2.18)$$

This analysis is useful because the basic CLOE algorithm, presented in appendix A and extended here, uses only the output disturbance noise $p(k)$.

¹This corresponds to the black-box model given by (A.1). Taking into account the definitions (A.2) and (A.3), the relative order of $G(q^{-1})$ is always bigger than zero, therefore there is also no direct input-output relation.

Without any loss of generality one can assume:

$$v(k) = [C_1(q^{-1}) \dots C_n(q^{-1})]^T e(k), \quad (2.19)$$

where $e(k)$ is a zero-mean Gaussian white noise and $C_1(q^{-1}), \dots, C_n(q^{-1})$ are polynomials representing different “coloring” Finite Impulse Response (FIR) filters. Introducing (2.17) and (2.19) into (2.13), one can express the effect of the process noise $v(k)$ on the scalar output $y(k)$:

$$\begin{aligned} \bar{v}(k) &= S_{yv}(q^{-1})v(k) = \begin{bmatrix} B_1 & & B_n \\ A & & A \end{bmatrix} [C_1 \dots C_n]^T e(k) = \\ &= \frac{B_1 C_1 + \dots + B_n C_n}{A} e(k), \end{aligned} \quad (2.20)$$

and defining:

$$B_1(q^{-1})C_1(q^{-1}) + \dots + B_n(q^{-1})C_n(q^{-1}) = C_{\bar{v}}(q^{-1}), \quad (2.21)$$

one obtains:

$$\bar{v}(k) = \frac{C_{\bar{v}}(q^{-1})}{A(q^{-1})} e(k). \quad (2.22)$$

This means that the process noise $v(k)$ defined by (2.19) can be transformed to the output as an additional disturbance noise $\bar{v}(k)$ given by (2.22).

If the output disturbance noise $p(k)$ in model (2.10) is assumed in the following form:

$$p(k) = \frac{C_p(q^{-1})}{A(q^{-1})} e(k), \quad (2.23)$$

one can finally express the aggregate output disturbance model:

$$p_{\text{new}}(k) = p(k) + \bar{v}(k) = \left(\frac{C_p(q^{-1})}{A(q^{-1})} + \frac{C_{\bar{v}}(q^{-1})}{A(q^{-1})} \right) e(k), \quad (2.24)$$

and denoting:

$$C(q^{-1}) = C_p(q^{-1}) + C_{\bar{v}}(q^{-1}), \quad (2.25)$$

the resulting output disturbance model is obtained:

$$p_{\text{new}}(k) = p(k) + \bar{v}(k) = \frac{C(q^{-1})}{A(q^{-1})} e(k). \quad (2.26)$$

One can see from table A.1 that the output disturbance model (2.26) is equivalent to the one considered in the X-CLOE algorithm. Therefore, this algorithm should be used in the presence of the process noise. Correspondingly, if Infinite Impulse Response (IIR) “coloring” filters are assumed, which leads to the $\frac{C}{AD}e(k)$ noise model, the G-CLOE algorithm will be applied.

Nevertheless, in section 2.6, only the basic CLOE algorithm is considered to illustrate the development of the new output-error state-space identification algorithms.

2.4.2 State-space Representation—Canonical Forms

As presented in section (2.2), a natural way to represent linear systems is by the state-space representation (2.7). The state-space representation describes the internal behavior of the system and in that sense is richer than the input-output transfer operator $S_{yu}(q^{-1})$ given by (2.16).

The problem of finding a state-space representation such that (2.16) holds for a given $S_{yu}(q^{-1})$ is called the realization problem. In this case one looks in general for a minimal realization *a state space representation of minimal order*. In this section we shall present an overview of the systems representations which are well adapted to the study of the structural properties of linear systems. The reachability and observability concepts are recalled [Kai80] and their canonical forms that use the similarity matrix T are introduced in general case, only for MIMO state-space systems.

Reachability Indices and Reachable Canonical Form

Assume a state-space model (2.7) and denote the *reachable subspace*, [Kai80]:

$$\mathcal{R} = \text{Im}[B_0, A_0B_0, \dots, A_0^{n-1}B_0]. \quad (2.27)$$

If we denote $\mathcal{B} = \text{Im}(B_0)$ the image of the linear map B_0 we can equivalently write (2.27) as:

$$\mathcal{R} = \mathcal{B} + A_0\mathcal{B} + \dots + A_0^{n-1}\mathcal{B}, \quad (2.28)$$

where “+” stands here for the subspace addition. Then we can introduce the subspaces:

$$\mathcal{R}_0 = 0, \quad \mathcal{R}_k = \mathcal{B} + A_0\mathcal{B} + \dots + A_0^{k-1}\mathcal{B}, \quad k = 1, \dots, n, \quad (2.29)$$

with the properties:

$$\mathcal{R}_0 \subset \mathcal{R}_1 \dots \subset \mathcal{R}_n = \langle A_0 | \mathcal{B} \rangle. \quad (2.30)$$

Let us define the following list of integers:

$$\bar{\rho}_i = \dim(\mathcal{R}_i / \mathcal{R}_{i-1}), \quad i = 1, \dots, n \quad \text{with} \quad \mathcal{R}_0 = 0. \quad (2.31)$$

If we consider a finite list of positive integers decreasingly ordered, we can associate with this list a “dual” list. Denote by $\bar{\rho}_1, \dots, \bar{\rho}_k$ the initial list with the property that $\bar{\rho}_1 \geq \bar{\rho}_2 \dots \geq \bar{\rho}_k$ and $\bar{\rho}_1 > 0$, $i = 1, \dots, k$. The dual list is obtained as:

$$\rho_i = \text{number of } \bar{\rho}_j \geq i. \quad (2.32)$$

Using definition $\bar{\rho}_1 = \dim \mathcal{B} = \text{rank}[B_0]$ and

$$\sum_{\bar{\rho}_i \geq 0} \bar{\rho}_i = \dim \langle A_0 | \mathcal{B} \rangle, \quad (2.33)$$

the dual list ρ_i will have $\bar{\rho}_i = \text{rank}[B_0] = m'$ non zero elements called the *reachability (controllability) indices* of (A_0, B_0) and

$$\sum_{i=1}^{m'} \rho_i = \dim \langle A_0 | \mathcal{B} \rangle. \quad (2.34)$$

form (similar statement can be made for the controllability form) in the multivariable case.

We will not prove this result here but we give the practical procedure for getting (A_c, B_c) starting with a reachable pair (A_0, B_0) (see [Pop72] and [Kai80] for more details). It contains the following steps:

Step 1: Compute the reachability matrix:

$$R = [B_0, A_0B_0, \dots, A_0^{n-1}B_0]. \quad (2.39)$$

By assumption this matrix has rank n . This matrix is also called the *controllability* or the *model controllability* matrix.

Step 2: Extract from R a non singular matrix by the following procedure: Select columns in R starting from the left and keeping only the columns which are independent of the columns already selected. Call this matrix P .

Step 3: Reorder the columns of P in order to obtain:

$$Q = [b_1, A_0b_1, \dots, A_0^{\rho_1-1}b_1, b_2, A_0b_2, \dots, A_0^{\rho_2-1}b_2, \dots, b_m, A_0b_m, \dots, A_0^{\rho_m-1}b_m]. \quad (2.40)$$

This may imply a renumbering of inputs to get the ordered list of reachability indices.

Step 4: Compute Q^{-1} and denote l_k its k th row. Denote:

$$\beta_j = \sum_{i=1}^j \rho_i \quad j = 1, \dots, m. \quad (2.41)$$

Form the matrix:

$$S = \begin{bmatrix} l_{\beta_1} \\ A_0 l_{\beta_1} \\ \vdots \\ A_0^{\rho_1-1} l_{\beta_1} \\ l_{\beta_2} \\ \vdots \\ A_0^{\rho_2-1} l_{\beta_2} \\ \vdots \\ l_{\beta_m} \\ \vdots \\ A_0^{\rho_m-1} l_{\beta_m} \end{bmatrix}. \quad (2.42)$$

Step 5: The required change of basis is $T = S^{-1}$. Compute the reachable (controller) canonical form:

$$A_c = T^{-1}A_0T, \quad B_c = T^{-1}B_0, \quad C_c = C_0T. \quad (2.43)$$

Observability Indices and Observer Canonical Form

We could introduce *observability indices* similarly as for the reachability (controllability) indices using the subspaces \mathcal{N}_k but we prefer to use duality.

Define that dual realization $\bar{A}_0, \bar{B}_0, \bar{C}_0$ of the triplet (A_0, B_0, C_0) of system (2.6) as follows: $(\bar{A}_0, \bar{B}_0, \bar{C}_0) = (A_0^T, B_0^T, C_0^T)$. It is easy to prove that (A_0, B_0, C_0) is reachable (resp. observable) if and only if (A_0^T, B_0^T, C_0^T) is observable (resp. reachable). The observability indices of the pair (C_0, A_0) , σ_i , $i = 1, \dots, \rho'$ where $\rho' = \text{rank}(C_0)$, can be computed as the reachability (controllability) indices of the pair (A_0^T, C_0^T) .

The *observer canonical form* of the pair (C_0, A_0) denoted $(C_{\text{ob}}, A_{\text{ob}})$ can be computed for an observable pair with C_0 full row rank by the following procedure:

Step 1: Compute first the reachable (controller) canonical form (\bar{A}_c, \bar{B}_c) of (A_0^T, B_0^T) .

Step 2: It turns out that $A_{\text{ob}} = \bar{A}_c^T$ and $C_{\text{ob}} = \bar{B}_c^T$.

2.4.3 Model Structural Identifiability

Once a model structure has been chosen, its properties should be studied. As a matter of fact, this study should if possible be done before the parameters estimation, to detect protectional problems before collecting data. These problems can become if the estimated parameters $\hat{\theta}$ do not correspond uniquely to the parameters of the model θ as presented in section 2.2.

For this reason, the identifiability of a model has to be analyzed, see [WP97]. It is defined as follows:

Let \mathcal{M} and $\mathcal{M}(\theta)$ respectively denote the model structure and the model with structure \mathcal{M} and parameters θ . The prior feasible set for θ will be denoted by Θ . In what follows, Θ will usually be either \mathbb{R}^{n_θ} or a subset of \mathbb{R}^{n_θ} defined by a finite set of inequality contrariness.

We wish to know whether the identical input-output behavior implies that the parameters $\hat{\theta}$ of the model equal those of the process θ^* . More precisely, parameter θ_i will be *structurally globally (uniquely) identifiable* if for almost any θ^* in Θ ,

$$\mathcal{M}(\hat{\theta}) = \mathcal{M}(\theta^*) \Rightarrow \hat{\theta}_i = \theta^*.$$

The structure \mathcal{M} will be structurally globally identifiable if all its parameters are be structurally globally identifiable.

When one cannot prove that the structure considered is globally identifiable, one may try to establish that it is at least locally. The parameter θ_i will be *structurally locally identifiable* if for almost any θ^* in Θ , there exists a neighborhood $\mathbb{V}(\theta^*)$ such that

$$\hat{\theta} \in \mathbb{V}(\theta^*) \quad \text{and} \quad \mathcal{M}(\hat{\theta}) = \mathcal{M}(\theta^*) \Rightarrow \hat{\theta}_i = \theta^*.$$

Local identifiability is therefore a necessary condition for global identifiability. The structure \mathcal{M} will be structurally locally identifiable if all its parameters are be structurally locally identifiable.

The parameter θ_i will be *structurally unidentifiable* if for almost any θ^* in Θ , there is no neighborhood $\mathbb{V}(\theta^*)$ such that

$$\hat{\theta} \in \mathbb{V}(\theta^*) \quad \text{and} \quad \mathcal{M}(\hat{\theta}) = \mathcal{M}(\theta^*) \Rightarrow \hat{\theta}_i = \theta^*.$$

The structure \mathcal{M} will be structurally unidentifiable if one at least of its parameters is structurally unidentifiable.

In the following section, the choice of the adjustable predictor used in optimization algorithm is discussed.

2.5 Optimization

The identification of the model can be formalized as an optimization problem [Lju99], *i.e.*:

$$\hat{\theta}_N = \arg_{\theta \in D_{\mathcal{M}}} \min J_N(\theta, Z^N), \quad (2.44)$$

where $\hat{\theta}_N$ is the estimated parameter vector, Z^N is the data set of the length N , $D_{\mathcal{M}}$ is the set of models $\mathcal{M}(\theta)$ and $J_N(\theta, Z^N)$ is the criterion function of the model parameter θ . In general, it is defined by the following norm:

$$J_N(\theta, Z^N) = \frac{1}{N} \sum_{t=1}^N \mathcal{L}(\varepsilon_F(k, \theta)), \quad (2.45)$$

where $\mathcal{L}(\cdot)$ is a scalar-valued function (typically the quadratic norm L_2 or L_1 norm or L_∞ norm) obtained from the filtered prediction error $\varepsilon_F(\cdot)$ as follows:

$$\varepsilon_F(k, \theta) = L(q^{-1})\varepsilon(k, \theta), \quad (2.46)$$

where $L(q^{-1})$ is a stable linear filter and $\varepsilon(k, \theta)$ is the prediction error by a certain model \mathcal{M} at time sample k .

In general, the function (2.44) cannot be minimized by analytical methods. The solution then has to be found by iterative, numerical techniques. This is usually done according to:

$$\hat{\theta}^{(i+1)} = \hat{\theta}^{(i)} + \mu^{(i)} f^{(i)}, \quad (2.47)$$

where $\hat{\theta}^{(i)}$ specifies the current iterate (number i), $f^{(i)}$ is the *search direction* and $\mu^{(i)}$ is the step size. The iterations are continued until $\hat{\theta}^{(i+1)}$ is believed to be sufficiently close to the θ .

Remark 1 *In this context, notation i stands for **iteration**, while further, in the context of recursive algorithms, notation i stands for **number of sample**.*

The family of minimization methods can be divided into three groups, [Lju99]:

1. Descent method or *steepest descent* method: Gradient, conjugate gradient, IGM, Back-propagation, *etc.*
2. Quadratic convergence rate method: Newton, Gauss-Newton, RLS (Gauss-Newton where the Hessian is solved by very well known matrix inversion lemma), *etc.*
3. Combination of them: Levenberg-Marquard.

In many cases it is necessary, or useful, to have a model of the system available on-line while the system is operating. The model should be then based on the observations up to the current time. The need for such an on-line model construction typically arises for either of two reasons, see *e.g.* [WP97], [LS83], [Lju99]:

1. There may be too many data to be stored simultaneously in the computer.
2. One may wish to use the results of the identification to take immediate decisions from the measurements performed so far, without having to wait until all data have been collected.

Identification techniques that are used in these cases are called *recursive identification methods* or on-line identification or adaptive parameter estimation, since the measured input-output data are processed recursively as they become available. Apart from the use of recursive methods in adaptive schemes, the algorithms can also turn out to be quite competitive alternatives for parameter estimation in off-line (non-recursive) situations.

The similarities and differences between the off-line (non-recursive) and the recursive identification algorithms have been investigated for the open-loop identification in [TCPG02] and the final results are presented here.

This analysis is focused on the convergence properties of the identification algorithms by using Ordinary Different Equation (ODE) method, [Lju99]. Only the case of the linear discrete-time black-box SISO systems (A.1) $\left(G_S(q^{-1}) = \frac{q^{-d}B(q^{-1})}{A(q^{-1})} \right)$ has been studied there, using two different model adjustable predictor structures, [LLM97]:

1. Recursive Least Squares (“Equation Error”) predictor:

$$\hat{y}(k+1) = \hat{\theta}^T \phi(k), \quad (2.48)$$

where $\hat{\theta}^T$ denotes the estimated parameter vector:

$$\hat{\theta}^T = \left[\hat{a}_1 \dots \hat{a}_{n_A} \quad \hat{b}_1 \dots \hat{b}_{n_B} \right], \quad (2.49)$$

and $\phi(k)$ is the *predictor regressor vector*:

$$\phi(k) = [-y(k) \dots -y(k-n_A+1) \quad u(k-d) \dots u(k-n_B+1-d)]^T. \quad (2.50)$$

2. Output-Error predictor (a parallel adjustable predictor known also as parallel Model Reference Adaptive System):

$$\hat{y}(k+1) = \hat{\theta}^T \phi(k, \hat{\theta}), \quad (2.51)$$

where $\hat{\theta}^T$ denotes the estimated parameter vector:

$$\hat{\theta}^T = \left[\hat{a}_1 \dots \hat{a}_{n_A} \quad \hat{b}_1 \dots \hat{b}_{n_B} \right], \quad (2.52)$$

and $\phi(k, \hat{\theta})$ is the *predictor regressor vector*:

$$\phi(k, \hat{\theta}) = [-\hat{y}(k) \dots -\hat{y}(k-n_A+1) \quad u(k-d) \dots u(k-n_B+1-d)]^T. \quad (2.53)$$

One can see from (2.50) and (2.53) that only the past values of the predictor output \hat{y} are replaced by the corresponding values of the output y in case of the output-error predictor. This is a fundamental difference between both approaches that leads to different convergence properties.

For this reason, a quadratic criterion (non-recursive) for the ‘‘Equation Error’’ predictor becomes:

$$\min_{\hat{\theta}(k)} J(k) = \sum_{i=1}^k [y(i) - \hat{y}(i)]^2 = \sum_{i=1}^k [y(i) - \hat{\theta}^T \phi(i)]^2, \quad (2.54)$$

a quadratic criterion (non-recursive) for the Output-Error predictor is:

$$\min_{\hat{\theta}(k)} J(k) = \sum_{i=1}^k [y(i) - \hat{y}(i)]^2 = \sum_{i=1}^k [y(i) - \hat{\theta}^T \phi(i, \hat{\theta})]^2 \quad (2.55)$$

and it corresponds to the criterion of the recursive algorithm as follows, [LLM97]:

$$\min_{\hat{\theta}(k)} J(k) = \sum_{i=1}^k \left\{ y(i) - \hat{\theta}^T(k) \phi(i-1, \hat{\theta}) \right\}^2 + \left\{ \theta - \hat{\theta}(0) \right\}^T F(0)^{-1} \left\{ \theta - \hat{\theta}(0) \right\} \quad (2.56)$$

where F is the adaptation gain.

The aim of the analysis was to compare the off-line (non-recursive) method to minimize the quadratic criterion (2.54) and the recursive method to minimize the criterion (2.56).

One can conclude that if the criterion (2.54), minimized by the non-recursive method, is used one can find secondary minima, [TCPG02]. Only for the zero mean value of the output noise p and infinity number of the measured samples $k \rightarrow \infty$, the non-recursive method converges to the true parameter vector $\hat{\theta}(k) \rightarrow \theta^*$.

On the other hand if the criterion (2.56) is minimized by the recursive method, one can find a condition of the global convergence, [TCPG02]:

$$Re \left\{ \frac{L(q^{-1})}{A(q^{-1})} - \frac{1}{2} \right\} > 0, \quad (2.57)$$

where, in general, $L(q^{-1})$ is a filter used to obtain the filtered error $\varepsilon_F(k, \theta) = L(q^{-1})\varepsilon(k, \theta)$ (see (2.46)) if it is necessary. The adaptive algorithm, proposed in [LK97b], it is the only case when the filter is $L(q^{-1}) = \hat{A}(\hat{\theta}(k-1), q^{-1})$.

Practically, a useful choice of the filter $L(q^{-1}) \rightarrow A(q^{-1})$, to satisfy the global convergence guaranty, is an advantage of the recursive output-error method against the non-recursive method using the ‘‘Equation Error’’ predictor. Therefore, the output-error predictor has been used for the new state-space identification algorithms.

2.6 State-space Identification Algorithms for SISO Systems

As discussed in section 2.4.1, the basic CLOE algorithm is taken here to illustrate the development of the output-error identification algorithm of the state-space systems.

Some results of this chapter have been presented in [BBVH03] and [HBVBP04]. This work is based on the results [Bez04] and [Bez01].

The aim of the original output-error state-space identification algorithms, which are presented in this section, is to estimate the parameters of the SISO discrete-time state-space model:

$$\boxed{\begin{aligned} x(k+1) &= A_0(\theta)x(k) + B_0(\theta)u(k-d), \\ y(k) &= C_0(\theta)x(k) + p(k), \end{aligned}} \quad (2.58)$$

where $x(k)$ is an n -dimensional state vector, $p(k)$ is an output disturbance noise and θ is a d_0 -dimensional parameter vector:

$$\theta = [\theta_1 \ \dots \ \theta_{d_0}]^T. \quad (2.59)$$

One should note that the number of parameters d_0 is not anyhow related to the model order n .

To simplify notation, *backward shift operator* q^{-1} will be omitted in some terms. In the sequel *parameter vector* θ and *estimated parameter vector* $\hat{\theta}$ are also omitted in some terms whenever there is no risk of confusion.

The state-space model (2.58) replaces the black-box model (A.1) in the closed-loop identification scheme (see fig. A.1). Hence, the time delay of d sampling periods is introduced into the model. In addition, it is assumed that there is no process noise ($v(k) \equiv 0$) and $D_0(\theta) \equiv 0$, as explained in section 2.4.1.

The input-output transfer operator is given by:

$$S_{yu}(q^{-1}) = q^{-d-1}C_0(I - q^{-1}A_0)^{-1}B_0 = \frac{q^{-d-1}C_0 \operatorname{adj}(I - q^{-1}A_0)B_0}{\det(I - q^{-1}A_0)}. \quad (2.60)$$

Denoting the numerator and the denominator:

$$S_{yu}(q^{-1}) = \frac{q^{-d}B(q^{-1})}{A(q^{-1})}, \quad (2.61)$$

$$B(q^{-1}) = b_1q^{-1} + \dots + b_nq^{-n} = q^{-1}B^*(q^{-1}), \quad (2.62)$$

$$A(q^{-1}) = 1 + a_1q^{-1} + \dots + a_nq^{-n} = 1 + q^{-1}A^*(q^{-1}), \quad (2.63)$$

one obtains:

$$\begin{aligned} B(q^{-1}) &= q^{-1}C_0 \operatorname{adj}(I - q^{-1}A_0)B_0, \\ A(q^{-1}) &= \det(I - q^{-1}A_0), \end{aligned} \quad (2.64)$$

$$\begin{aligned} B^*(q^{-1}) &= C_0 \operatorname{adj}(I - q^{-1}A_0)B_0, \\ A^*(q^{-1}) &= q(\det(I - q^{-1}A_0) - 1), \end{aligned} \quad (2.65)$$

and the output of the plant is given by:

$$\begin{aligned} y(k+1) &= S_{yu}(q^{-1})u(k+1) + p(k+1), \\ &\vdots \\ y(k+1) &= -A^*y(k) + B^*u(k-d) + Ap(k+1). \end{aligned} \quad (2.66)$$

Taking into account that state-space matrices $A_0(\theta)$, $B_0(\theta)$, $C_0(\theta)$ are functions of the parameter vector θ , the equations (2.64) or (2.65) also define the transfer function coefficients b_1, \dots, b_n and a_1, \dots, a_n as functions of θ . Therefore using (2.64) or (2.65),

one can determine the function $\Gamma(\theta)$, which transforms the parameter vector θ to a vector of the transfer function coefficients:

$$\Gamma^T(\theta) = [a_1(\theta) \ \dots \ a_n(\theta) \ b_1(\theta) \ \dots \ b_n(\theta)]. \quad (2.67)$$

This transformational function plays a key role in all the newly proposed algorithms. One can clearly see that $\Gamma(\theta)$ is a $2n$ -dimensional vector function of a d_0 -dimensional parameter vector θ . It is generally non-linear, the estimated parameters can appear in sums, products, ratios and other aggregate terms. It is assumed differentiable.

Using (2.67), (2.66) can be rewritten to the regressor form:

$$y(k+1) = \Gamma^T(\theta)\varphi(k) + Ap(k+1), \quad (2.68)$$

where $\varphi(k)$ denotes the regressor vector:

$$\varphi(k) = [-y(k) \ \dots \ -y(k-n+1) \ u(k-d) \ \dots \ u(k-n+1-d)]^T, \quad (2.69)$$

and the closed-loop predictor is given by:

$$\hat{y}(k+1) = \Gamma^T(\hat{\theta})\phi(k), \quad (2.70)$$

where $\hat{\theta}$ denotes the estimated parameter vector:

$$\hat{\theta} = [\hat{\theta}_1 \ \dots \ \hat{\theta}_{d_0}]^T, \quad (2.71)$$

and $\phi(k)$ the predictor regressor vector:

$$\phi(k) = [-\hat{y}(k) \ \dots \ -\hat{y}(k-n+1) \ \hat{u}(k-d) \ \dots \ \hat{u}(k-n+1-d)]^T. \quad (2.72)$$

Replacing the fixed predictor of the closed-loop (2.70) by an adjustable predictor, one obtains *a priori* predicted output:

$$\hat{y}^\circ(k+1) = \hat{y} [k+1|\hat{\theta}(k)] = \Gamma^T [\hat{\theta}(k)] \phi(k), \quad (2.73)$$

and *a posteriori* predicted output:

$$\hat{y}(k+1) = \hat{y} [k+1|\hat{\theta}(k+1)] = \Gamma^T [\hat{\theta}(k+1)] \phi(k). \quad (2.74)$$

Consequently, the *a priori* prediction error can be defined as:

$$\varepsilon_{\text{CL}}^\circ(k+1) = y(k+1) - \hat{y}^\circ(k+1), \quad (2.75)$$

and the *a posteriori* prediction error as:

$$\varepsilon_{\text{CL}}(k+1) = y(k+1) - \hat{y}(k+1). \quad (2.76)$$

The development of the parameter adaptation algorithms will be shown now. One should follow the procedures described in section A.4. Firstly, the state-space algorithm based on the gradient method is given. It has been named as GM-2 algorithm in the following text. Then two variant of the improved gradient algorithms, namely IGM-1 and IGM-2 are developed. Finally, the state-space algorithm based on the recursive least squares method is presented, named the RLS-2 algorithm. It is discussed more precisely than previous algorithms (GM-2, IGM-1, IGM-2) to demonstrate some specific problems related to the $\Gamma(\theta)$ transformational function.

2.6.1 The SISO GM-2 Algorithm

Let's consider a quadratic criterion in terms of the *a priori* prediction error:

$$\min_{\hat{\theta}(k)} J(k+1) = \frac{1}{2} [\varepsilon_{\text{CL}}^{\circ}(k+1)]^2. \quad (2.77)$$

To minimize this criterion, the gradient technique is used:

$$\hat{\theta}(k+1) = \hat{\theta}(k) - F \frac{\partial J(k+1)}{\partial \hat{\theta}(k)}, \quad (2.78)$$

where F is a constant matrix adaptation gain, and:

$$\frac{\partial J(k+1)}{\partial \hat{\theta}(k)} = \frac{\partial \varepsilon_{\text{CL}}^{\circ}(k+1)}{\partial \hat{\theta}(k)} \varepsilon_{\text{CL}}^{\circ}(k+1). \quad (2.79)$$

Since the *a priori* prediction error is given by:

$$\varepsilon_{\text{CL}}^{\circ}(k+1) = y(k+1) - \hat{y}^{\circ}(k+1) = y(k+1) - \Gamma^{\text{T}} [\hat{\theta}(k)] \phi(k), \quad (2.80)$$

one can continue:

$$\frac{\partial \varepsilon_{\text{CL}}^{\circ}(k+1)}{\partial \hat{\theta}(k)} = -\Gamma_{\theta}^{\prime \text{T}} [\hat{\theta}(k)] \phi(k) - \frac{\partial \phi(k)}{\partial \hat{\theta}(k)}^{\text{T}} \Gamma [\hat{\theta}(k)], \quad (2.81)$$

where $\partial \phi(k)/\partial \hat{\theta}(k)$ is the Jacobian matrix of the predictor regressor vector defined by:

$$\frac{\partial \phi(k)}{\partial \hat{\theta}(k)} = \begin{bmatrix} -\frac{\partial \hat{y}(k)}{\partial \hat{\theta}_1(k)} & \cdots & -\frac{\partial \hat{y}(k)}{\partial \hat{\theta}_{d_0}(k)} \\ \vdots & \ddots & \vdots \\ -\frac{\partial \hat{y}(k-n+1)}{\partial \hat{\theta}_1(k)} & \cdots & -\frac{\partial \hat{y}(k-n+1)}{\partial \hat{\theta}_{d_0}(k)} \\ \frac{\partial \hat{u}(k-d)}{\partial \hat{\theta}_1(k)} & \cdots & \frac{\partial \hat{u}(k-d)}{\partial \hat{\theta}_{d_0}(k)} \\ \vdots & \ddots & \vdots \\ \frac{\partial \hat{u}(k-n+1-d)}{\partial \hat{\theta}_1(k)} & \cdots & \frac{\partial \hat{u}(k-n+1-d)}{\partial \hat{\theta}_{d_0}(k)} \end{bmatrix}. \quad (2.82)$$

As explained in section A.4, the Jacobian matrix $\partial \phi(k)/\partial \hat{\theta}(k)$ cannot be generally considered null in the closed-loop identification scheme, since $\hat{y}(k)$ is a function of $\hat{\theta}(k)$ and, moreover, $\hat{u}(k-d)$ can also become a function of $\hat{\theta}(k)$ under certain circumstances. However, this function is controller-dependent, therefore it must be separately determined in each particular case.

Introducing (2.81) into (2.79), the parameter adaptation algorithm (2.78) becomes:

$$\hat{\theta}(k+1) = \hat{\theta}(k) + F \left(\Gamma_{\theta}^{\prime \text{T}} [\hat{\theta}(k)] \phi(k) + \frac{\partial \phi(k)}{\partial \hat{\theta}(k)}^{\text{T}} \Gamma [\hat{\theta}(k)] \right) \varepsilon_{\text{CL}}^{\circ}(k+1). \quad (2.83)$$

If the matrix adaptation gain is assumed in the form $F = \alpha I$, where $\alpha > 0$, an equivalent algorithm is obtained when replacing the matrix adaptation gain F by a scalar adaptation gain α .

Finally, the GM-2 state-space parameter adaptation algorithm for SISO systems can be formulated:

$$\begin{aligned} \hat{\theta}(k+1) &= \hat{\theta}(k) + \alpha \left(\Gamma_{\theta}^{\prime T} [\hat{\theta}(k)] \phi(k) + \frac{\partial \phi(k)^T}{\partial \hat{\theta}(k)} \Gamma [\hat{\theta}(k)] \right) \varepsilon_{\text{CL}}^{\circ}(k+1) \\ \varepsilon_{\text{CL}}^{\circ}(k+1) &= y(k+1) - \Gamma^T [\hat{\theta}(k)] \phi(k) \end{aligned} \quad (2.84)$$

One should note that the GM-2 algorithm have the advantage of a closed analytic form.

2.6.2 The SISO IGM-1 and IGM-2 Algorithms

In order to improve the performance and stability properties of the GM-2 algorithm, the gradient technique is used again but a criterion in terms of the *a posteriori* prediction error is considered:

$$\min_{\hat{\theta}(k+1)} J(k+1) = \frac{1}{2} [\varepsilon_{\text{CL}}(k+1)]^2. \quad (2.85)$$

The gradient of the criterion becomes:

$$\frac{\partial J(k+1)}{\partial \hat{\theta}(k+1)} = \frac{\partial \varepsilon_{\text{CL}}(k+1)}{\partial \hat{\theta}(k+1)} \varepsilon_{\text{CL}}(k+1), \quad (2.86)$$

and since the *a posteriori* prediction error is given by:

$$\varepsilon_{\text{CL}}(k+1) = y(k+1) - \hat{y}(k+1) = y(k+1) - \Gamma^T [\hat{\theta}(k+1)] \phi(k), \quad (2.87)$$

one gets:

$$\frac{\partial \varepsilon_{\text{CL}}(k+1)}{\partial \hat{\theta}(k+1)} = -\Gamma_{\theta}^{\prime T} [\hat{\theta}(k+1)] \phi(k). \quad (2.88)$$

In this case, the Jacobian matrix $\partial \phi(k)/\partial \hat{\theta}(k+1)$ is null as explained in section 2.

Introducing (2.88) into (2.86), the parameter adaptation algorithm becomes:

$$\hat{\theta}(k+1) = \hat{\theta}(k) + F_1 \Gamma_{\theta}^{\prime T} [\hat{\theta}(k+1)] \phi(k) \varepsilon_{\text{CL}}(k+1), \quad (2.89)$$

where F_1 is a constant matrix adaptation gain. This algorithm depends on $\varepsilon_{\text{CL}}(k+1)$ and $\Gamma_{\theta}^{\prime} [\hat{\theta}(k+1)]$, which are functions of $\hat{\theta}(k+1)$. For implementing the algorithm, these terms must be expressed as functions of $\hat{\theta}(k)$ instead.

One can rewrite (2.87) as:

$$\varepsilon_{\text{CL}}(k+1) = y(k+1) - \Gamma^T [\hat{\theta}(k)] \phi(k) - \left\{ \Gamma [\hat{\theta}(k+1)] - \Gamma [\hat{\theta}(k)] \right\}^T \phi(k), \quad (2.90)$$

or simply:

$$\varepsilon_{\text{CL}}(k+1) = \varepsilon_{\text{CL}}^{\circ}(k+1) - \Delta \Gamma^T(k+1) \phi(k). \quad (2.91)$$

Hence, one needs to express the increment of the vector of transfer function coefficients $\Delta \Gamma(k+1)$ in terms of the prediction error. Two different approaches can be used to achieve this objective.

1. Although the algorithm (2.89) directly controls the parameter vector $\hat{\theta}$, one can assume that the evolution of the vector of transfer function coefficients $\Gamma \left[\hat{\theta}(k) \right]$ can be approximated by the evolution of the parameter vector controlled by the primary improved gradient algorithm (see section 2). Therefore using (A.75), one obtains:

$$\Delta\Gamma(k+1) \approx F_2\phi(k)\varepsilon_{\text{CL}}(k+1), \quad (2.92)$$

where F_2 is a constant matrix adaptation gain ($\dim F_2 \neq \dim F_1$).

Introducing (2.92) into (2.91), the *a posteriori* prediction error is obtained in the following form:

$$\varepsilon_{\text{CL}}(k+1) = \frac{\varepsilon_{\text{CL}}^\circ(k+1)}{1 + \phi^\text{T}(k)F_2^\text{T}\phi(k)}, \quad (2.93)$$

and the algorithm (2.89) becomes:

$$\hat{\theta}(k+1) = \hat{\theta}(k) + \frac{F_1\Gamma_\theta'^\text{T} \left[\hat{\theta}(k+1) \right] \phi(k)\varepsilon_{\text{CL}}^\circ(k+1)}{1 + \phi^\text{T}(k)F_2^\text{T}\phi(k)}. \quad (2.94)$$

2. Under the assumption that the Jacobian matrix $\Gamma'_\theta(\theta)$ varies slowly and/or the parameter vector increments $\Delta\hat{\theta}(k+1)$ keep small, linearization can be applied:

$$\Delta\Gamma(k+1) \approx \Gamma'_\theta \left[\hat{\theta}(k) \right] \Delta\hat{\theta}(k+1). \quad (2.95)$$

From (2.89) one obtains:

$$\Delta\hat{\theta}(k+1) = F_1\Gamma_\theta'^\text{T} \left[\hat{\theta}(k+1) \right] \phi(k)\varepsilon_{\text{CL}}(k+1), \quad (2.96)$$

and therefore:

$$\Delta\Gamma(k+1) \approx \Gamma'_\theta \left[\hat{\theta}(k) \right] F_1\Gamma_\theta'^\text{T} \left[\hat{\theta}(k+1) \right] \phi(k)\varepsilon_{\text{CL}}(k+1). \quad (2.97)$$

Introducing (2.97) into (2.91), the *a posteriori* prediction error is obtained in the following form:

$$\varepsilon_{\text{CL}}(k+1) = \frac{\varepsilon_{\text{CL}}^\circ(k+1)}{1 + \phi^\text{T}(k)\Gamma'_\theta \left[\hat{\theta}(k+1) \right] F_1^\text{T}\Gamma_\theta'^\text{T} \left[\hat{\theta}(k) \right] \phi(k)}, \quad (2.98)$$

and the algorithm (2.89) becomes:

$$\hat{\theta}(k+1) = \hat{\theta}(k) + \frac{F_1\Gamma_\theta'^\text{T} \left[\hat{\theta}(k+1) \right] \phi(k)\varepsilon_{\text{CL}}^\circ(k+1)}{1 + \phi^\text{T}(k)\Gamma'_\theta \left[\hat{\theta}(k+1) \right] F_1^\text{T}\Gamma_\theta'^\text{T} \left[\hat{\theta}(k) \right] \phi(k)}. \quad (2.99)$$

One can see that both algorithms (2.94) and (2.99) still contain the term $\Gamma'_\theta \left[\hat{\theta}(k+1) \right]$. Under the assumption stated above in order to apply the linearization (2.95), one can resolve:

$$\Gamma'_\theta \left[\hat{\theta}(k+1) \right] \approx \Gamma'_\theta \left[\hat{\theta}(k) \right]. \quad (2.100)$$

Finally, introducing the approximation (2.100) into (2.94) and using the scalar adaptation gains α_1 and α_2 , the IGM-1 state-space parameter adaptation algorithm for SISO systems can be summarized:

$$\boxed{\begin{aligned}\hat{\theta}(k+1) &= \hat{\theta}(k) + \frac{\alpha_1 \Gamma_{\theta}^{\prime T} [\hat{\theta}(k)] \phi(k) \varepsilon_{\text{CL}}^{\circ}(k+1)}{1 + \alpha_2 \phi^T(k) \phi(k)} \\ \varepsilon_{\text{CL}}^{\circ}(k+1) &= y(k+1) - \Gamma^T [\hat{\theta}(k)] \phi(k)\end{aligned}} \quad (2.101)$$

Correspondingly, the IGM-2 algorithm is obtained from (2.99):

$$\boxed{\begin{aligned}\hat{\theta}(k+1) &= \hat{\theta}(k) + \frac{\alpha_1 \Gamma_{\theta}^{\prime T} [\hat{\theta}(k)] \phi(k) \varepsilon_{\text{CL}}^{\circ}(k+1)}{1 + \alpha_1 \phi^T(k) \Gamma_{\theta}^{\prime} [\hat{\theta}(k)] \Gamma_{\theta}^{\prime T} [\hat{\theta}(k)] \phi(k)} \\ \varepsilon_{\text{CL}}^{\circ}(k+1) &= y(k+1) - \Gamma^T [\hat{\theta}(k)] \phi(k)\end{aligned}} \quad (2.102)$$

One should note that the IGM-1 and IGM-2 algorithms have the advantage of a closed analytic form.

2.6.3 The SISO RLS-2 Algorithm

One starts with the *least squares* (LS) criterion:

$$\min_{\hat{\theta}(k)} J(k) = \sum_{i=1}^k \left\{ y(i) - \hat{y} [i | \hat{\theta}(k)] \right\}^2 = \sum_{i=1}^k \left\{ y(i) - \Gamma^T [\hat{\theta}(k)] \phi(i-1) \right\}^2. \quad (2.103)$$

To get the value of $\hat{\theta}(k)$, which minimizes the criterion (2.103), one seeks the value that cancels $\partial J(k) / \partial \hat{\theta}(k)$:

$$\frac{\partial J(k)}{\partial \hat{\theta}(k)} = -2 \sum_{i=1}^k \left\{ y(i) - \Gamma^T [\hat{\theta}(k)] \phi(i-1) \right\} \Gamma_{\theta}^{\prime T} [\hat{\theta}(k)] \phi(i-1) = 0, \quad (2.104)$$

where the Jacobian matrix $\Gamma_{\theta}^{\prime}(\theta)$ is introduced by:

$$\Gamma_{\theta}^{\prime}(\theta) = \frac{\partial \Gamma(\theta)}{\partial \theta} = \begin{bmatrix} \frac{\partial a_1(\theta)}{\partial \theta_1} & \cdots & \frac{\partial a_1(\theta)}{\partial \theta_{d_0}} \\ \vdots & \ddots & \vdots \\ \frac{\partial a_n(\theta)}{\partial \theta_1} & \cdots & \frac{\partial a_n(\theta)}{\partial \theta_{d_0}} \\ \frac{\partial b_1(\theta)}{\partial \theta_1} & \cdots & \frac{\partial b_1(\theta)}{\partial \theta_{d_0}} \\ \vdots & \ddots & \vdots \\ \frac{\partial b_n(\theta)}{\partial \theta_1} & \cdots & \frac{\partial b_n(\theta)}{\partial \theta_{d_0}} \end{bmatrix}. \quad (2.105)$$

From (2.104) one obtains:

$$\sum_{i=1}^k \Gamma^T [\hat{\theta}(k)] \phi(i-1) \Gamma_{\theta}^{\prime T} [\hat{\theta}(k)] \phi(i-1) = \sum_{i=1}^k y(i) \Gamma_{\theta}^{\prime T} [\hat{\theta}(k)] \phi(i-1), \quad (2.106)$$

and transposing both sides:

$$\sum_{i=1}^k \left\{ \Gamma^T \left[\hat{\theta}(k) \right] \phi(i-1) \phi^T(i-1) \right\} \Gamma'_\theta \left[\hat{\theta}(k) \right] = \sum_{i=1}^k \left\{ y(i) \phi^T(i-1) \right\} \Gamma'_\theta \left[\hat{\theta}(k) \right]. \quad (2.107)$$

One can see that the Jacobian matrix $\Gamma'_\theta \left[\hat{\theta}(k) \right]$ is ready to be canceled out of (2.107). However, since this matrix is not generally square, its rigid inverse cannot be found. Therefore, the resulting equation (2.108) is not equivalent, but it can be understood as a sufficient condition (which is not always necessary):

$$\sum_{i=1}^k \Gamma^T \left[\hat{\theta}(k) \right] \phi(i-1) \phi^T(i-1) = \sum_{i=1}^k y(i) \phi^T(i-1). \quad (2.108)$$

Re-transposing (2.108), the following form is obtained:

$$\left[\sum_{i=1}^k \phi(i-1) \phi^T(i-1) \right] \Gamma \left[\hat{\theta}(k) \right] = \sum_{i=1}^k y(i) \phi(i-1), \quad (2.109)$$

and therefore:

$$\Gamma \left[\hat{\theta}(k) \right] = \left[\sum_{i=1}^k \phi(i-1) \phi^T(i-1) \right]^{-1} \sum_{i=1}^k y(i) \phi(i-1) = F(k) \sum_{i=1}^k y(i) \phi(i-1), \quad (2.110)$$

in which:

$$F(k)^{-1} = \sum_{i=1}^k \phi(i-1) \phi^T(i-1). \quad (2.111)$$

Replacing $\Gamma \left[\hat{\theta}(k) \right]$ by $\hat{\theta}(k)$, these formulae become equivalent to (A.83) and (A.84). Hence, the procedure shown in section 2 can be used to obtain the recursive algorithm:

$$\Gamma \left[\hat{\theta}(k+1) \right] = \Gamma \left[\hat{\theta}(k) \right] + F(k+1) \phi(k) \varepsilon_{\text{CL}}^\circ(k+1), \quad (2.112)$$

$$F(k+1) = F(k) - \frac{F(k) \phi(k) \phi^T(k) F(k)}{1 + \phi^T(k) F(k) \phi(k)}. \quad (2.113)$$

This result can be easily interpreted. The algorithm is a traditional recursive least squares algorithm for an estimation of black-box model parameters, in which the vector of transfer function coefficients is represented as $\Gamma(\theta)$. However, the aim is to estimate the parameter vector θ . Therefore, the algorithm must be expanded to contain the direct θ estimation.

For this purpose, linearization is applied. Using (2.112), one obtains:

$$\Delta \Gamma(k+1) = \Gamma \left[\hat{\theta}(k+1) \right] - \Gamma \left[\hat{\theta}(k) \right] = F(k+1) \phi(k) \varepsilon_{\text{CL}}^\circ(k+1). \quad (2.114)$$

Under the assumption that the higher-order terms in the Taylor series expansion of $\Gamma \left[\hat{\theta}(k) \right]$ can be neglected, the difference $\Delta \Gamma(k+1)$ can be approximated by a first-order formula:

$$\Delta \Gamma(k+1) \approx \Gamma'_\theta \left[\hat{\theta}(k) \right] \Delta \hat{\theta}(k+1), \quad (2.115)$$

where:

$$\Delta\hat{\theta}(k+1) = \hat{\theta}(k+1) - \hat{\theta}(k). \quad (2.116)$$

The equation (2.115) represents a system of $2n$ linear equations with d_0 unknowns, in which $\Delta\hat{\theta}(k+1)$ is a vector of unknowns and $\Gamma'_\theta \left[\hat{\theta}(k) \right]$ is a $2n \times d_0$ dimensional equation system matrix. Hence, three different cases can be distinguished:

- $2n = d_0$, *i.e.* the number of transfer function coefficients is equal to the number of estimated parameters. The Jacobian matrix $\Gamma'_\theta \left[\hat{\theta}(k) \right]$ is square, and if it is nonsingular, the inverse $\Gamma'_\theta \left[\hat{\theta}(k) \right]^{-1}$ can be found and the system (2.115) has a unique solution.
- $2n < d_0$, *i.e.* there are more parameters to be estimated than transfer function coefficients. The system (2.115) may be underdetermined and infinity solutions possibly exist. However, one of them can be selected using an additional criterion.
- $2n > d_0$, *i.e.* there are less parameters to be estimated than transfer function coefficients. The system (2.115) may be overdetermined and it possibly does not have any exact solution. However, an optimal solution in the least square sense can be always found.

The appropriate algorithms for solving under- and overdetermined systems of equations are implemented in MATLAB[®] *slash operator*, for example. Thus, the solution of (2.115) can be obtained:

$$\Delta\hat{\theta}(k+1) \approx \Gamma'_\theta \left[\hat{\theta}(k) \right]^{(-1)} \Delta\Gamma(k+1), \quad (2.117)$$

where $\Gamma'_\theta \left[\hat{\theta}(k) \right]^{(-1)}$ denotes so-called *pseudoinverse* or *generalized inverse* of the Jacobian matrix $\Gamma'_\theta \left[\hat{\theta}(k) \right]$ (see [SS89] for theoretical treatment of this issue).

Putting (2.114), (2.117) and (2.116) together, the recursive formula for parameter vector estimate is gained:

$$\hat{\theta}(k+1) = \hat{\theta}(k) + \Gamma'_\theta \left[\hat{\theta}(k) \right]^{(-1)} F(k+1)\phi(k)\varepsilon_{\text{CL}}^\circ(k+1), \quad (2.118)$$

and replacing (2.113) by (A.102), the RLS-2 state-space parameter adaptation algorithm for SISO systems can be summarized in the following three equations:

$$\begin{array}{l} \hat{\theta}(k+1) = \hat{\theta}(k) + \Gamma'_\theta \left[\hat{\theta}(k) \right]^{(-1)} F(k+1)\phi(k)\varepsilon_{\text{CL}}^\circ(k+1) \\ F(k+1) = \frac{1}{\lambda_1(k)} \left[F(k) - \frac{F(k)\phi(k)\phi^T(k)F(k)}{\frac{\lambda_1(k)}{\lambda_2(k)} + \phi^T(k)F(k)\phi(k)} \right] \\ \varepsilon_{\text{CL}}^\circ(k+1) = y(k+1) - \Gamma^T \left[\hat{\theta}(k) \right] \phi(k) \end{array} \quad (2.119)$$

One should note that when transforming the vector of transfer function coefficients $\Gamma \left[\hat{\theta}(k+1) \right]$ to the parameter vector $\hat{\theta}(k+1)$, two sources of inaccuracy arise:

1. The linearization used in (2.115) if $\Gamma(\theta)$ is a non-linear function of θ .
2. The pseudoinverse used in (2.117) if the system (2.115) is overdetermined.

Therefore, the stability and convergence properties of the algorithm are difficult to explore. They will always depend on the parametrization itself.

One can see that a numerical approach needed to compute the pseudoinverse matrix in the RLS-2 algorithm is avoided in the GM-2, IGM-1 and IGM-2 algorithms. Hence, the closed analytic form is an advantage of the GM-2, IGM-1, IGM-2 algorithms over the RLS-2 algorithm. However, stability and convergence still cannot be assured. In addition, gradient algorithm generally suffers from lower performance.

In the following section, we give an extension of these algorithms (GM-2, IGM-1, IGM-2, RLS-2) for MIMO systems.

2.7 State-space Identification Algorithms for MIMO Systems

Let's consider a MIMO discrete-time state-space model now:

$$\boxed{\begin{aligned} x(k+1) &= A_0(\theta)x(k) + B_0(\theta)u(k), \\ y(k) &= C_0(\theta)x(k) + p(k), \end{aligned}} \quad (2.120)$$

where $x(k)$ is an n -dimensional state column vector:

$$x(k) = [x_1(k) \ \dots \ x_n(k)]^T, \quad (2.121)$$

$u(k)$ is an r -dimensional input column vector:

$$u(k) = [u_1(k) \ \dots \ u_r(k)]^T, \quad (2.122)$$

$y(k)$ is an s -dimensional output column vector:

$$y(k) = [y_1(k) \ \dots \ y_s(k)]^T, \quad (2.123)$$

and $w(k)$ is an s -dimensional output disturbance column vector:

$$p(k) = [p_1(k) \ \dots \ p_s(k)]^T. \quad (2.124)$$

Hence, one has an n th-order MIMO model with r inputs and s outputs, and the state-space matrices are of the following dimensions:

$$\begin{aligned} \dim A_0(\theta) &= n \times n, \\ \dim B_0(\theta) &= n \times r, \\ \dim C_0(\theta) &= s \times n. \end{aligned}$$

As usual, θ is a d_0 -dimensional parameter vector:

$$\theta = [\theta_1 \ \dots \ \theta_{d_0}]^T. \quad (2.125)$$

The state-space model (2.120) replaces the black-box model in the closed-loop identification scheme (see fig. A.1). One should note that all the quantities marked in fig. A.1 become vectors of appropriate dimensions.

The input-output transfer operator of the model (2.120) is an $s \times r$ dimensional matrix given by:

$$S_{yu}(q^{-1}) = q^{-1}C_0(I - q^{-1}A_0)^{-1}B_0 = \frac{q^{-1}C_0 \operatorname{adj}(I - q^{-1}A_0)B_0}{\det(I - q^{-1}A_0)}. \quad (2.126)$$

Denoting the numerator and the denominator:

$$S_{yu}(q^{-1}) = \frac{B(q^{-1})}{A(q^{-1})}, \quad (2.127)$$

$$B(q^{-1}) = q^{-1}B^*(q^{-1}), \quad (2.128)$$

$$A(q^{-1}) = 1 + q^{-1}A^*(q^{-1}), \quad (2.129)$$

one obtains:

$$\begin{aligned} B(q^{-1}) &= q^{-1}C_0 \operatorname{adj}(I - q^{-1}A_0)B_0, \\ A(q^{-1}) &= \det(I - q^{-1}A_0), \end{aligned} \quad (2.130)$$

$$\begin{aligned} B^*(q^{-1}) &= C_0 \operatorname{adj}(I - q^{-1}A_0)B_0, \\ A^*(q^{-1}) &= q(\det(I - q^{-1}A_0) - 1), \end{aligned} \quad (2.131)$$

and the output of the plant is given by:

$$\begin{aligned} y(k+1) &= S_{yu}(q^{-1})u(k+1) + p(k+1), \\ &\vdots \\ y(k+1) &= -A^*y(k) + B^*u(k) + Ap(k+1). \end{aligned} \quad (2.132)$$

As expected, the equations (2.130)–(2.131) derived for MIMO systems formally correspond to the equations (2.64)–(2.65) derived for SISO systems. In this case, the numerator of the transfer operator $S_{yu}(q^{-1})$ is an $s \times r$ dimensional polynomial matrix of the following form:

$$B(q^{-1}) = \begin{bmatrix} B^{11}(q^{-1}) & \dots & B^{1r}(q^{-1}) \\ \vdots & \ddots & \vdots \\ B^{s1}(q^{-1}) & \dots & B^{sr}(q^{-1}) \end{bmatrix}, \quad (2.133)$$

where $B^{ji}(q^{-1})$ (for $j = 1, \dots, s$ and $i = 1, \dots, r$) denotes the numerator of the transfer operator between the i th input and the j th output:

$$B^{ji}(q^{-1}) = b_1^{ji}q^{-1} + \dots + b_n^{ji}q^{-n}. \quad (2.134)$$

On the contrary, the denominator is a common polynomial:

$$A(q^{-1}) = 1 + a_1q^{-1} + \dots + a_nq^{-n}. \quad (2.135)$$

Since the numerator $B(q^{-1})$ contains nrs coefficients and the denominator $A(q^{-1})$ contains n coefficients, there are $n(rs+1)$ different transfer function coefficients in the MIMO case.

Taking into account that state-space matrices $A_0(\theta)$, $B_0(\theta)$, $C_0(\theta)$ are functions of the parameter vector θ , the equations (2.130) or (2.131) also define all the transfer function coefficients as functions of θ . Therefore using (2.130) or (2.131), one can determine the function $\Gamma(\theta)$, which transforms the parameter vector θ to a vector of the transfer function coefficients:

$$\Gamma(\theta) = [a(\theta) \quad b^{11}(\theta) \quad \dots \quad b^{1r}(\theta) \quad \dots \quad b^{s1}(\theta) \quad \dots \quad b^{sr}(\theta)]^T, \quad (2.136)$$

where:

$$a(\theta) = [a_1(\theta) \quad \dots \quad a_n(\theta)], \quad (2.137)$$

$$b^{ji}(\theta) = [b_1^{ji}(\theta) \quad \dots \quad b_n^{ji}(\theta)]. \quad (2.138)$$

In this case, $\Gamma(\theta)$ is generally a non-linear $n(rs + 1)$ -dimensional vector function of a d_0 -dimensional parameter vector θ . It is assumed differentiable again.

Using (2.136), (2.132) can be rewritten to the regressor form:

$$y(k + 1) = \varphi(k)\Gamma(\theta) + Ap(k + 1), \quad (2.139)$$

where $\varphi(k)$ denotes the regressor matrix:

$$\varphi(k) = \begin{bmatrix} -Y_1(k) & U_1(k) & \dots & U_r(k) & & 0 \\ \vdots & & & & \ddots & \\ -Y_s(k) & & 0 & & U_1(k) & \dots & U_r(k) \end{bmatrix}, \quad (2.140)$$

using the notation:

$$Y_j(k) = [y_j(k) \quad \dots \quad y_j(k - n + 1)], \quad (2.141)$$

$$U_i(k) = [u_i(k) \quad \dots \quad u_i(k - n + 1)]. \quad (2.142)$$

The closed-loop predictor is given by:

$$\hat{y}(k + 1) = \phi(k)\Gamma(\hat{\theta}), \quad (2.143)$$

where $\hat{\theta}$ denotes the estimated parameter vector:

$$\hat{\theta} = [\hat{\theta}_1 \quad \dots \quad \hat{\theta}_{d_0}]^T, \quad (2.144)$$

and $\phi(k)$ the predictor regressor matrix:

$$\phi(k) = \begin{bmatrix} -\hat{Y}_1(k) & \hat{U}_1(k) & \dots & \hat{U}_r(k) & & 0 \\ \vdots & & & & \ddots & \\ -\hat{Y}_s(k) & & 0 & & \hat{U}_1(k) & \dots & \hat{U}_r(k) \end{bmatrix}, \quad (2.145)$$

$$\hat{Y}_j(k) = [\hat{y}_j(k) \quad \dots \quad \hat{y}_j(k - n + 1)], \quad (2.146)$$

$$\hat{U}_i(k) = [\hat{u}_i(k) \quad \dots \quad \hat{u}_i(k - n + 1)]. \quad (2.147)$$

Replacing the fixed predictor of the closed-loop (2.143) by an adjustable predictor, one obtains *a priori* predicted output:

$$\hat{y}^\circ(k + 1) = \hat{y} [k + 1 | \hat{\theta}(k)] = \phi(k)\Gamma [\hat{\theta}(k)], \quad (2.148)$$

and *a posteriori* predicted output:

$$\hat{y}(k+1) = \hat{y} \left[k+1 | \hat{\theta}(k+1) \right] = \phi(k) \Gamma \left[\hat{\theta}(k+1) \right]. \quad (2.149)$$

Consequently, the *a priori* and the *a posteriori* prediction error can be defined in the usual way:

$$\varepsilon_{\text{CL}}^{\circ}(k+1) = y(k+1) - \hat{y}^{\circ}(k+1), \quad (2.150)$$

$$\varepsilon_{\text{CL}}(k+1) = y(k+1) - \hat{y}(k+1). \quad (2.151)$$

One should note that $\varepsilon_{\text{CL}}^{\circ}$ and ε_{CL} are s -dimensional vectors this time.

Now, the MIMO state-space parameter adaptation algorithms will be derived in a concise way. One should follow the pattern given in section 2.6 to gain a complete image.

2.7.1 The MIMO GM-2 Algorithm

As usual, one minimizes the quadratic criterion:

$$\min_{\hat{\theta}(k)} J(k+1) = \frac{1}{2} \varepsilon_{\text{CL}}^{\circ\text{T}}(k+1) \varepsilon_{\text{CL}}^{\circ}(k+1), \quad (2.152)$$

using the gradient technique:

$$\hat{\theta}(k+1) = \hat{\theta}(k) - F \frac{\partial J(k+1)}{\partial \hat{\theta}(k)}, \quad (2.153)$$

where F is a constant matrix adaptation gain, and the gradient is given by:

$$\frac{\partial J(k+1)}{\partial \hat{\theta}(k)} = \frac{\partial \varepsilon_{\text{CL}}^{\circ}(k+1)^{\text{T}}}{\partial \hat{\theta}(k)} \varepsilon_{\text{CL}}^{\circ}(k+1). \quad (2.154)$$

Since the *a priori* prediction error is defined as:

$$\varepsilon_{\text{CL}}^{\circ}(k+1) = y(k+1) - \hat{y}^{\circ}(k+1) = y(k+1) - \phi(k) \Gamma \left[\hat{\theta}(k) \right], \quad (2.155)$$

one can continue:

$$\frac{\partial \varepsilon_{\text{CL}}^{\circ}(k+1)^{\text{T}}}{\partial \hat{\theta}(k)} = -\Gamma_{\theta}^{\text{T}} \left[\hat{\theta}(k) \right] \phi^{\text{T}}(k) - \frac{\partial \phi(k)}{\partial \hat{\theta}(k)} \odot \Gamma \left[\hat{\theta}(k) \right], \quad (2.156)$$

where $\partial \phi(k) / \partial \hat{\theta}(k)$ is a tensor-like structure obtained by differentiating each element of the predictor regressor matrix $\phi(k)$ defined by (2.145)–(2.147) with respect to the parameter vector $\hat{\theta}(k)$. Hence, the dimension of $\partial \phi(k) / \partial \hat{\theta}(k)$ can be symbolically expressed as:

$$\dim \frac{\partial \phi(k)}{\partial \hat{\theta}(k)} = s \times n(rs+1) \times d_0.$$

The operator \odot realizes dot product along rows of the same dimension, which is $n(rs+1)$, so that the resulting dimension is compatible with the other terms in (2.156):

$$\dim \frac{\partial \phi(k)}{\partial \hat{\theta}(k)} \odot \Gamma \left[\hat{\theta}(k) \right] = d_0 \times s.$$

As observed in section A.4, the derivative $\partial\phi(k)/\partial\hat{\theta}(k)$ is generally nonzero in the closed-loop identification scheme, and it is controller-dependent.

Introducing (2.156) into (2.154), the parameter adaptation algorithm (2.153) can be rewritten:

$$\hat{\theta}(k+1) = \hat{\theta}(k) + F \left(\Gamma_{\theta}^{\prime T} [\hat{\theta}(k)] \phi^T(k) + \frac{\partial\phi(k)}{\partial\hat{\theta}(k)} \odot \Gamma [\hat{\theta}(k)] \right) \varepsilon_{\text{CL}}^{\circ}(k+1). \quad (2.157)$$

Using the scalar adaptation gain α , the MIMO version of the GM-2 state-space parameter adaptation algorithm can be summarized:

$$\begin{array}{l} \hat{\theta}(k+1) = \hat{\theta}(k) + \alpha \left(\Gamma_{\theta}^{\prime T} [\hat{\theta}(k)] \phi^T(k) + \frac{\partial\phi(k)}{\partial\hat{\theta}(k)} \odot \Gamma [\hat{\theta}(k)] \right) \varepsilon_{\text{CL}}^{\circ}(k+1) \\ \varepsilon_{\text{CL}}^{\circ}(k+1) = y(k+1) - \phi(k)\Gamma [\hat{\theta}(k)] \end{array} \quad (2.158)$$

2.7.2 The MIMO IGM-1 and IGM-2 Algorithms

A quadratic criterion in terms of the *a posteriori* prediction error is taken:

$$\min_{\hat{\theta}(k+1)} J(k+1) = \frac{1}{2} \varepsilon_{\text{CL}}^T(k+1) \varepsilon_{\text{CL}}(k+1). \quad (2.159)$$

In order to minimize the criterion (2.159), the gradient technique is used. The gradient is given by:

$$\frac{\partial J(k+1)}{\partial\hat{\theta}(k+1)} = \frac{\partial\varepsilon_{\text{CL}}(k+1)}{\partial\hat{\theta}(k+1)} \varepsilon_{\text{CL}}^T(k+1). \quad (2.160)$$

The *a posteriori* prediction error is defined as:

$$\varepsilon_{\text{CL}}(k+1) = y(k+1) - \hat{y}(k+1) = y(k+1) - \phi(k)\Gamma [\hat{\theta}(k+1)], \quad (2.161)$$

and one obtains:

$$\frac{\partial\varepsilon_{\text{CL}}(k+1)}{\partial\hat{\theta}(k+1)} \varepsilon_{\text{CL}}^T(k+1) = -\Gamma_{\theta}^{\prime T} [\hat{\theta}(k+1)] \phi^T(k). \quad (2.162)$$

Introducing (2.162) into (2.160), the parameter adaptation algorithm becomes:

$$\hat{\theta}(k+1) = \hat{\theta}(k) + F_1 \Gamma_{\theta}^{\prime T} [\hat{\theta}(k+1)] \phi^T(k) \varepsilon_{\text{CL}}(k+1), \quad (2.163)$$

where F_1 is a constant matrix adaptation gain. One has to express the terms $\varepsilon_{\text{CL}}(k+1)$ and $\Gamma_{\theta}^{\prime T} [\hat{\theta}(k+1)]$ as functions of $\hat{\theta}(k)$ now.

One can rewrite (2.161) as:

$$\varepsilon_{\text{CL}}(k+1) = y(k+1) - \phi(k)\Gamma [\hat{\theta}(k)] - \phi(k) \left\{ \Gamma [\hat{\theta}(k+1)] - \Gamma [\hat{\theta}(k)] \right\}, \quad (2.164)$$

or simply:

$$\varepsilon_{\text{CL}}(k+1) = \varepsilon_{\text{CL}}^{\circ}(k+1) - \phi(k)\Delta\Gamma(k+1). \quad (2.165)$$

As shown in section 2.6.2, two different ways are proposed to express $\Delta\Gamma(k+1)$ in terms of the prediction error.

- Using the results obtained for the primary improved gradient algorithm (see section 2), one can similarly assume in the MIMO case:

$$\Delta\Gamma(k+1) \approx F_2\phi^T(k)\varepsilon_{\text{CL}}(k+1), \quad (2.166)$$

where F_2 is a constant matrix adaptation gain ($\dim F_2 \neq \dim F_1$).

Introducing (2.166) into (2.165), the *a posteriori* prediction error is obtained in the following form:

$$\varepsilon_{\text{CL}}(k+1) = [I + \phi(k)F_2\phi^T(k)]^{-1} \varepsilon_{\text{CL}}^{\circ}(k+1), \quad (2.167)$$

and the algorithm (2.163) becomes:

$$\begin{aligned} \hat{\theta}(k+1) = & \hat{\theta}(k) + F_1\Gamma_{\theta}'^T [\hat{\theta}(k+1)] \phi^T(k) [I + \\ & + \phi(k)F_2\phi^T(k)]^{-1} \varepsilon_{\text{CL}}^{\circ}(k+1). \end{aligned} \quad (2.168)$$

- Applying linearization, one can assume:

$$\Delta\Gamma(k+1) \approx \Gamma_{\theta}' [\hat{\theta}(k)] \Delta\hat{\theta}(k+1). \quad (2.169)$$

From (2.163) one obtains:

$$\Delta\hat{\theta}(k+1) = F_1\Gamma_{\theta}'^T [\hat{\theta}(k+1)] \phi^T(k)\varepsilon_{\text{CL}}(k+1), \quad (2.170)$$

and therefore:

$$\Delta\Gamma(k+1) \approx \Gamma_{\theta}' [\hat{\theta}(k)] F_1\Gamma_{\theta}'^T [\hat{\theta}(k+1)] \phi^T(k)\varepsilon_{\text{CL}}(k+1). \quad (2.171)$$

Introducing (2.171) into (2.165), the *a posteriori* prediction error is obtained in the following form:

$$\varepsilon_{\text{CL}}(k+1) = \left\{ I + \phi(k)\Gamma_{\theta}' [\hat{\theta}(k)] F_1\Gamma_{\theta}'^T [\hat{\theta}(k+1)] \phi^T(k) \right\}^{-1} \varepsilon_{\text{CL}}^{\circ}(k+1), \quad (2.172)$$

and the algorithm (2.163) becomes:

$$\begin{aligned} \hat{\theta}(k+1) = & \hat{\theta}(k) + F_1\Gamma_{\theta}'^T [\hat{\theta}(k+1)] \phi^T(k) \left\{ I + \right. \\ & \left. + \phi(k)\Gamma_{\theta}' [\hat{\theta}(k)] F_1\Gamma_{\theta}'^T [\hat{\theta}(k+1)] \phi^T(k) \right\}^{-1} \varepsilon_{\text{CL}}^{\circ}(k+1). \end{aligned} \quad (2.173)$$

Furthermore, the following approximation is utilized to finish the development of (2.168) and (2.173):

$$\Gamma_{\theta}' [\hat{\theta}(k+1)] \approx \Gamma_{\theta}' [\hat{\theta}(k)]. \quad (2.174)$$

Hence, using the scalar adaptation gains α_1 and α_2 in (2.168), the MIMO version of the IGM-1 state-space parameter adaptation algorithm is obtained:

$\begin{aligned} \hat{\theta}(k+1) = & \hat{\theta}(k) + \\ & + \alpha_1\Gamma_{\theta}'^T [\hat{\theta}(k)] \phi^T(k) [I + \alpha_2\phi(k)\phi^T(k)]^{-1} \varepsilon_{\text{CL}}^{\circ}(k+1) \\ \varepsilon_{\text{CL}}^{\circ}(k+1) = & y(k+1) - \phi(k)\Gamma [\hat{\theta}(k)] \end{aligned} \quad (2.175)$

Correspondingly, the MIMO version of the IGM-2 algorithm can be obtained from (2.173):

$$\begin{aligned}
 \hat{\theta}(k+1) &= \hat{\theta}(k) + \alpha_1 \Gamma_{\theta}'^T [\hat{\theta}(k)] \phi^T(k) \left\{ I + \right. \\
 &\quad \left. + \alpha_1 \phi(k) \Gamma_{\theta}' [\hat{\theta}(k)] \Gamma_{\theta}'^T [\hat{\theta}(k)] \phi^T(k) \right\}^{-1} \varepsilon_{\text{CL}}^{\circ}(k+1) \\
 \varepsilon_{\text{CL}}^{\circ}(k+1) &= y(k+1) - \phi(k) \Gamma [\hat{\theta}(k)]
 \end{aligned} \tag{2.176}$$

2.7.3 The MIMO RLS-2 Algorithm

The least squares criterion takes the following form in the MIMO case:

$$\begin{aligned}
 \min_{\hat{\theta}(k)} J(k) &= \sum_{i=1}^k \left\{ y(i) - \hat{y} [i|\hat{\theta}(k)] \right\}^T \left\{ y(i) - \hat{y} [i|\hat{\theta}(k)] \right\} = \\
 &= \sum_{i=1}^k \left\{ y(i) - \phi(i-1) \Gamma [\hat{\theta}(k)] \right\}^T \left\{ y(i) - \phi(i-1) \Gamma [\hat{\theta}(k)] \right\}.
 \end{aligned} \tag{2.177}$$

To minimize the criterion (2.177), one solves:

$$\frac{\partial J(k)}{\partial \hat{\theta}(k)} = -2 \sum_{i=1}^k \Gamma_{\theta}'^T [\hat{\theta}(k)] \phi^T(i-1) \left\{ y(i) - \phi(i-1) \Gamma [\hat{\theta}(k)] \right\} = 0, \tag{2.178}$$

where the Jacobian matrix $\Gamma_{\theta}'(\theta)$ is introduced by:

$$\Gamma_{\theta}'(\theta) = \frac{\partial \Gamma(\theta)}{\partial \theta} = \begin{bmatrix} \frac{\partial a_1(\theta)}{\partial \theta_1} & \cdots & \frac{\partial a_1(\theta)}{\partial \theta_{d_0}} \\ \vdots & \ddots & \vdots \\ \frac{\partial a_n(\theta)}{\partial \theta_1} & \cdots & \frac{\partial a_n(\theta)}{\partial \theta_{d_0}} \\ \frac{\partial b_1^1(\theta)}{\partial \theta_1} & \cdots & \frac{\partial b_1^1(\theta)}{\partial \theta_{d_0}} \\ \vdots & & \vdots \\ \vdots & \ddots & \vdots \\ \vdots & & \vdots \\ \frac{\partial b_n^r(\theta)}{\partial \theta_1} & \cdots & \frac{\partial b_n^r(\theta)}{\partial \theta_{d_0}} \end{bmatrix}. \tag{2.179}$$

From (2.178) one obtains:

$$\Gamma_{\theta}'^T [\hat{\theta}(k)] \sum_{i=1}^k \phi^T(i-1) \phi(i-1) \Gamma [\hat{\theta}(k)] = \Gamma_{\theta}'^T [\hat{\theta}(k)] \sum_{i=1}^k \phi^T(i-1) y(i), \tag{2.180}$$

and after canceling out the term $\Gamma_{\theta}'^T [\hat{\theta}(k)]$, a sufficient condition is obtained:

$$\left[\sum_{i=1}^k \phi^T(i-1) \phi(i-1) \right] \Gamma [\hat{\theta}(k)] = \sum_{i=1}^k \phi^T(i-1) y(i). \tag{2.181}$$

Therefore:

$$\begin{aligned}\Gamma \left[\hat{\theta}(k) \right] &= \left[\sum_{i=1}^k \phi^T(i-1)\phi(i-1) \right]^{-1} \sum_{i=1}^k \phi^T(i-1)y(i) = \\ &= F(k) \sum_{i=1}^k \phi^T(i-1)y(i),\end{aligned}\quad (2.182)$$

in which:

$$F(k)^{-1} = \sum_{i=1}^k \phi^T(i-1)\phi(i-1). \quad (2.183)$$

These formulae are analogous to the formulae obtained in the SISO case. Therefore, the well-known procedure can be utilized to derive a recursive form. However, one should remember that ϕ is a matrix and y is a vector now.

The resulting recursive algorithm can be formulated:

$$\Gamma \left[\hat{\theta}(k+1) \right] = \Gamma \left[\hat{\theta}(k) \right] + F(k+1)\phi^T(k)\varepsilon_{\text{CL}}^{\circ}(k+1), \quad (2.184)$$

$$\begin{aligned}F(k+1) &= F(k) - F(k)\phi^T(k) \left[I + \right. \\ &\quad \left. + \phi(k)F(k)\phi^T(k) \right]^{-1} \phi(k)F(k).\end{aligned}\quad (2.185)$$

One can see that this algorithm is only able to estimate the transfer function coefficients represented as $\Gamma(\theta)$. To include the direct θ estimation, linearization is applied. Using (2.184) to express:

$$\Delta\Gamma(k+1) = \Gamma \left[\hat{\theta}(k+1) \right] - \Gamma \left[\hat{\theta}(k) \right] = F(k+1)\phi^T(k)\varepsilon_{\text{CL}}^{\circ}(k+1), \quad (2.186)$$

and denoting:

$$\Delta\hat{\theta}(k+1) = \hat{\theta}(k+1) - \hat{\theta}(k), \quad (2.187)$$

the difference $\Delta\Gamma(k+1)$ can be approximated by a first-order formula:

$$\Delta\Gamma(k+1) \approx \Gamma'_{\theta} \left[\hat{\theta}(k) \right] \Delta\hat{\theta}(k+1). \quad (2.188)$$

The equation (2.188) represents a system of $n(rs+1)$ linear equations with d_0 unknowns, in which $\Delta\hat{\theta}(k+1)$ is a vector of unknowns and $\Gamma'_{\theta} \left[\hat{\theta}(k) \right]$ is a $n(rs+1) \times d_0$ dimensional equation system matrix. Hence, the system (2.188) may generally become under- or overdetermined. The solution can be expressed using a pseudoinverse:

$$\Delta\hat{\theta}(k+1) \approx \Gamma'_{\theta} \left[\hat{\theta}(k) \right]^{(-1)} \Delta\Gamma(k+1), \quad (2.189)$$

where $\Gamma'_{\theta} \left[\hat{\theta}(k) \right]^{(-1)}$ is the pseudoinverse of the Jacobian matrix $\Gamma'_{\theta} \left[\hat{\theta}(k) \right]$ (see [SS89] for theoretical treatment of this issue).

Putting (2.186), (2.189) and (2.187) together, the recursive formula for parameter vector estimate is gained:

$$\hat{\theta}(k+1) = \hat{\theta}(k) + \Gamma'_{\theta} \left[\hat{\theta}(k) \right]^{(-1)} F(k+1)\phi^T(k)\varepsilon_{\text{CL}}^{\circ}(k+1), \quad (2.190)$$

and taking the generalized matrix adaptation gain instead of (2.185), the MIMO version of the RLS-2 state-space parameter adaptation algorithm can be given:

$$\begin{aligned}
 \hat{\theta}(k+1) &= \hat{\theta}(k) + \Gamma'_\theta \left[\hat{\theta}(k) \right]^{(-1)} F(k+1) \phi^T(k) \varepsilon_{\text{CL}}^\circ(k+1) \\
 F(k+1) &= \frac{1}{\lambda_1(k)} \left\{ F(k) - F(k) \phi^T(k) \left[\frac{\lambda_1(k)}{\lambda_2(k)} I + \right. \right. \\
 &\quad \left. \left. + \phi(k) F(k) \phi^T(k) \right]^{-1} \phi(k) F(k) \right\} \\
 \varepsilon_{\text{CL}}^\circ(k+1) &= y(k+1) - \phi(k) \Gamma \left[\hat{\theta}(k) \right]
 \end{aligned} \tag{2.191}$$

2.7.4 A General Structure of PAA for State-space Systems

One can consider as a general structure the following PAA (integral type), which is given in the similar way as in section 2:

$$\begin{aligned}
 \hat{\theta}(k+1) &= \hat{\theta}(k) + \Gamma'_\theta \left[\hat{\theta}(k) \right]^{(-1)} F(k+1) \phi^T(k) \varepsilon_{\text{CL}}(k+1), \\
 F(k+1)^{-1} &= \Lambda_1(k) F(k)^{-1} + \Lambda_2(k) \phi(k) \phi^T(k), \\
 \Lambda_1 &= \lambda_1 I, \\
 \Lambda_2 &= \lambda_2 I, \\
 0 < \lambda_1(k) \leq 1, & \quad 0 \leq \lambda_2(k) < 2, \\
 F(0) > 0, & \quad F(k)^{-1} > \alpha F(0)^{-1}, \quad 0 < \alpha < \infty, \\
 F(k+1) &= \frac{1}{\lambda_1(k)} \left\{ F(k) - F(k) \phi^T(k) \left[\frac{\lambda_1(k)}{\lambda_2(k)} I + \right. \right. \\
 &\quad \left. \left. \phi(k) F(k) \phi^T(k) \right]^{-1} \phi(k) F(k) \right\}, \\
 \varepsilon_{\text{CL}}(k+1) &= \varepsilon_{\text{CL}}^\circ(k+1) \left\{ I + \phi^T(k) F(k) \phi(k) \right\}^{-1},
 \end{aligned} \tag{2.192}$$

where the *a priori* predicted output is given by:

$$\hat{y} \left[k+1 | \hat{\theta}(k) \right] = \phi(k) \Gamma \left[\hat{\theta}(k) \right],$$

the *a priori* closed-loop prediction error is:

$$\varepsilon_{\text{CL}}^\circ(k+1) = y(k+1) - \hat{y} \left[k+1 | \hat{\theta}(k) \right],$$

and the *a posteriori* closed-loop prediction error is:

$$\varepsilon_{\text{CL}}(k+1) = y(k+1) - \hat{y} \left[k+1 | \hat{\theta}(k+1) \right].$$

The different criterion functions that are summarized in table 2.1 and minimization techniques lead to the different adaptation gain matrices in PAA.

One can see that GM-2, IGM-1 and IGM-2 algorithms correspond to the RLS-2 algorithm using a constant adaptation gain matrix $F(k+1) = F(k) = F > 0$. Therefore, the properties of RLS-2 algorithm, studied in the following sections, are assumed to be closed to the properties of GM-2, IGM-1 and IGM-2 algorithms.

Remark 2 *One step criterion used in the GM-2, IGM-1 and IGM-2 algorithms means that at any instant k we have a different criterion with a different minimum. As time goes on and the recursive algorithm is applied we hope to find the true minimum θ^* .*

Table 2.1: Summary of the non-recursive and recursive algorithm criterions.

	Name	Criterion
Non-recursive	GM	$\min_{\hat{\theta}(k)} J(k+1) = \frac{1}{2} \varepsilon_{\text{CL}}^{\circ\text{T}}(k+1) \varepsilon_{\text{CL}}^{\circ}(k+1)$
	IGM	$\min_{\hat{\theta}(k+1)} J(k+1) = \frac{1}{2} \varepsilon_{\text{CL}}^{\text{T}}(k+1) \varepsilon_{\text{CL}}(k+1)$
	LS	$\min_{\hat{\theta}(k)} J(k) = \sum_{i=1}^k \left\{ \left\{ y(i) - \hat{y} \left[i \hat{\theta}(k) \right] \right\}^{\text{T}} \left\{ y(i) - \hat{y} \left[i \hat{\theta}(k) \right] \right\} \right\}$
Recursive	GM	$\min_{\hat{\theta}(k)} J(k+1) = \frac{1}{2} \varepsilon_{\text{CL}}^{\circ\text{T}}(k+1) \varepsilon_{\text{CL}}^{\circ}(k+1)$
	IGM	$\min_{\hat{\theta}(k+1)} J(k+1) = \frac{1}{2} \varepsilon_{\text{CL}}^{\text{T}}(k+1) \varepsilon_{\text{CL}}(k+1)$
	RLS	$\min_{\hat{\theta}(k)} J(k) = \sum_{i=1}^k \left\{ \left\{ y(i) - \hat{y} \left[i \hat{\theta}(k) \right] \right\}^{\text{T}} \left\{ y(i) - \hat{y} \left[i \hat{\theta}(k) \right] \right\} \right\} +$ $+ \left\{ \Gamma[\theta] - \Gamma \left[\hat{\theta}(0) \right] \right\}^{\text{T}} F(0)^{-1} \left\{ \Gamma[\theta] - \Gamma \left[\hat{\theta}(0) \right] \right\}$

2.8 Properties of Parameter Adaptation Algorithms for State-space Systems

The same techniques, as presented in section A.5, have been used to analyze the properties of the parameter adaptation algorithms for state-space systems. This analysis is given in appendix 2.8.

2.8.1 Convexity of Criterion Function

The criterion function of the recursive least squares (RLS-2) algorithm as a function of the parameter vector $\hat{\theta}$ (see (2.177) and (2.144)), is given by:

$$J(\hat{\theta}) = \sum_{i=1}^k \left(y(i) - \phi(i-1)\Gamma(\hat{\theta}) \right)^{\text{T}} \left(y(i) - \phi(i-1)\Gamma(\hat{\theta}) \right). \quad (2.193)$$

The same criterion function is obtained also for the other developed algorithms (GM-2, IGM-1, IGM-2) if the time horizon is one sample, *i.e.* if $i = 1$. Therefore, all new algorithms can be studied together, using the same theoretical analysis.

Assuming a particular parametrization of $\Gamma(\hat{\theta})$ function, $\Gamma(\hat{\theta}) = M\hat{\theta}$, where M is the square matrix given from the parametrization and $\hat{\theta}$ is the estimated parameter vector, then the criterion function $J(\hat{\theta})$, given by (2.193), is a convex function.

Remark 3 *The condition $\Gamma(\hat{\theta}) = M\hat{\theta}$ means that the transformation function $\Gamma(\hat{\theta})$, given by (2.136), is linear towards the parameter vector $\hat{\theta}$. In general, this restriction is difficult to be satisfied for MIMO systems.*

Proof: Any convex function $\Upsilon(X) : R^l \rightarrow R$ has to verify the following condition, [SW70] and [Roc82]:

$$\begin{aligned} \Upsilon(\lambda X_1 + (1 - \lambda)X_2) &\leq \lambda \Upsilon(X_1) + (1 - \lambda) \Upsilon(X_2) \\ \forall \lambda &\in \langle 0, 1 \rangle; \forall X_1, X_2 \in R^l. \end{aligned}$$

In this case $J(\hat{\theta}) \equiv \Upsilon(X)$ and the convexity condition becomes for the criterion function (2.193) as follows:

$$\begin{aligned} \sum_{i=1}^k \left\{ y(i) - \phi(i-1)M(\lambda X_1 + (1-\lambda)X_2) \right\}^T \left\{ y(i) - \phi(i-1)M(\lambda X_1 + (1-\lambda)X_2) \right\} \leq \\ \lambda \sum_{i=1}^k \left\{ y(i) - \phi(i-1)M\lambda X_1 \right\}^T \left\{ y(i) - \phi(i-1)M\lambda X_1 \right\} + \\ (1-\lambda) \sum_{i=1}^k \left\{ y(i) - \phi(i-1)M\lambda X_2 \right\}^T \left\{ y(i) - \phi(i-1)M\lambda X_2 \right\} \end{aligned} \quad (2.194)$$

Using basic matrix operations, the inequality (2.194) becomes:

$$\begin{aligned} \lambda^2 \sum_{i=1}^k \left\{ \phi(i-1)MX_1 \right\}^T \left\{ \phi(i-1)MX_1 \right\} + \lambda \sum_{i=1}^k \left\{ \phi(i-1)MX_1 \right\}^T \left\{ \phi(i-1)MX_2 \right\} + \\ \lambda^2 \sum_{i=1}^k \left\{ -\phi(i-1)MX_1 \right\}^T \left\{ \phi(i-1)MX_2 \right\} + \lambda \sum_{i=1}^k \left\{ \phi(i-1)MX_2 \right\}^T \left\{ \phi(i-1)MX_1 \right\} + \\ \lambda \sum_{i=1}^k \left\{ -\phi(i-1)MX_2 \right\}^T \left\{ \phi(i-1)MX_2 \right\} + \lambda^2 \sum_{i=1}^k \left\{ -\phi(i-1)MX_2 \right\}^T \left\{ \phi(i-1)MX_1 \right\} + \\ \lambda^2 \sum_{i=1}^k \left\{ -\phi(i-1)MX_2 \right\}^T \left\{ \phi(i-1)MX_2 \right\} + \lambda \sum_{i=1}^k \left\{ -\phi(i-1)MX_1 \right\}^T \left\{ \phi(i-1)MX_1 \right\} \leq 0 \end{aligned} \quad (2.195)$$

The term $(\lambda^2 - \lambda)$ can be separated as follows:

$$\begin{aligned} (\lambda^2 - \lambda) \left\{ \sum_{i=1}^k \left\{ \phi(i-1)MX_1 \right\}^T \left\{ \phi(i-1)MX_1 \right\} - \left\{ \phi(i-1)MX_1 \right\}^T \left\{ \phi(i-1)MX_2 \right\} - \right. \\ \left. \left\{ \phi(i-1)MX_2 \right\}^T \left\{ \phi(i-1)MX_1 \right\} + \left\{ \phi(i-1)MX_2 \right\}^T \left\{ \phi(i-1)MX_2 \right\} \right\} \leq 0 \end{aligned} \quad (2.196)$$

Finally, the inequality (2.196) becomes:

$$\lambda(\lambda - 1) \sum_{i=1}^k \left\{ \phi(i-1)MX_1 - \phi(i-1)MX_2 \right\}^T \left\{ \phi(i-1)MX_1 - \phi(i-1)MX_2 \right\} \leq 0, \quad (2.197)$$

which is always true because the term inside of the summation is a quadratic form and $(\lambda - 1) \leq 0$.

2.8.2 Stability of PAA in a Deterministic Environment

Equivalent Feedback Representation of PAA

In the case of recursive least squares or of the improved gradient algorithm, the following *a posteriori* predictor has been used, see (2.149):

$$\hat{y}(k+1) = \hat{y} \left[k+1 | \hat{\theta}(k+1) \right] = \phi(k)\Gamma \left[\hat{\theta}(k+1) \right]. \quad (2.198)$$

Assume that the transformation function $\Gamma(\hat{\theta})$, given by (2.136), is linear in the parameter vector $\hat{\theta}$, *i.e.* $\Gamma(\hat{\theta}) = M\hat{\theta}$. Then (2.149) becomes:

$$\hat{y}(k+1) = \hat{y} \left[k+1 | \hat{\theta}(k+1) \right] = \phi(k)M\hat{\theta}(k+1). \quad (2.199)$$

The PAA has the following form, see (2.192):

$$\hat{\theta}(k+1) = \hat{\theta}(k) + \Gamma'_\theta \left[\hat{\theta}(k) \right]^{(-1)} F(k+1)\phi^T(k)\varepsilon_{\text{CL}}(k+1) \quad (2.200)$$

and if $\Gamma(\hat{\theta}) = M\hat{\theta}$ then $\Gamma'_\theta \left[\hat{\theta}(k) \right] = M$. This enables to write (2.200) as:

$$\hat{\theta}(k+1) = \hat{\theta}(k) + M^{-1}F(k+1)\phi^T(k)\varepsilon_{\text{CL}}(k+1). \quad (2.201)$$

A vector of the parameter errors is defined as follows:

$$\tilde{\theta}(k) = \hat{\theta}(k) - \theta. \quad (2.202)$$

Subtracting θ in both sides of (2.201) and taking into account (2.202), one obtains:

$$\tilde{\theta}(k+1) = \tilde{\theta}(k) + M^{-1}F(k+1)\phi^T(k)\varepsilon_{\text{CL}}(k+1). \quad (2.203)$$

From definition of the *a posteriori* prediction error $\varepsilon_{\text{CL}}(k+1)$ given by (2.151) and taking into account (2.139) for zero output disturbance noise $p(k) \equiv 0$ (deterministic case) and (2.202), one gets:

$$\varepsilon_{\text{CL}}(k+1) = -H_r(q^{-1})\tilde{\theta}^T(k+1)\phi(k), \quad (2.204)$$

where $H_r(q^{-1})$ is a transfer function matrix given from the closed-loop predictor. It will be defined in the following. Using (2.203), one can write:

$$\tilde{\theta}^T(k+1)\phi(k) = \tilde{\theta}(k)^T\phi(k) - \phi(k)M^{-1}F(k)\phi(k)^T\varepsilon_{\text{CL}}(k+1). \quad (2.205)$$

Equation (2.203) through (2.205) defines an equivalent feedback system illustrated in fig. 2.1. Equation (2.204) defines a linear block with constant parameters on the feed-forward path, whose input is $-\tilde{\theta}^T(k+1)\phi(k)$ in the similar way as in the case of SISO black-box model identification in section A.5. This block is characterized by a transfer function matrix $H_r(q^{-1})$ in the case of output error predictor (OE plant model). Equations (2.203) and (2.205) define a non-linear time-varying block in the feedback path.

The equivalent feedback representation associated to PAA is also formed by two blocks, namely the linear time-invariant block and the non-linear time-varying block, as presented in section A.5. Therefore, the passivity (hyperstability) properties have been used for the stability analysis of such feedback system.

Stability of PAA Using the Equivalent Feedback Representation

Assuming that the closed-loop system (with the controller transfer function matrix K) is stable and the function $\Gamma(\hat{\theta})$ is linear in the parameter vector $\hat{\theta}$, *i.e.* $\Gamma(\hat{\theta}) = M\hat{\theta}$, the recursive parameter estimation algorithm given by (2.192) assures

$$\begin{aligned} \lim_{k \rightarrow \infty} \varepsilon_{\text{CL}}(k+1) &= 0, \\ \lim_{k \rightarrow \infty} \varepsilon_{\text{CL}}^\circ(k+1) &= 0, \\ \|\phi(k)\| &< C, \quad 0 < C < \infty, \quad \forall k \end{aligned}$$

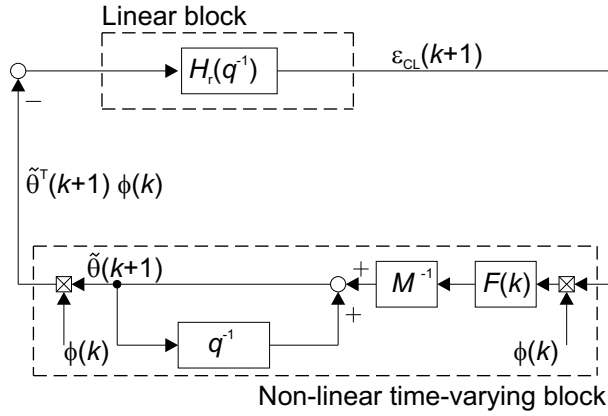


Figure 2.1: Equivalent feedback representation of PAA associated to the output error predictor (OE plant model) for MIMO systems.

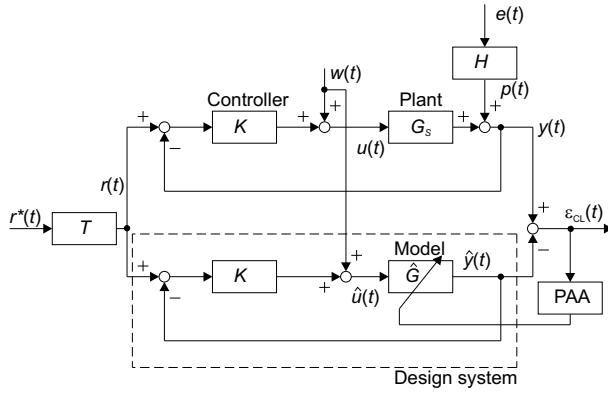


Figure 2.2: Closed-loop identification scheme for MIMO systems.

for all initial conditions $\hat{\theta}(0)$, $\varepsilon_{CL}^o(0)$ and $\phi(0)$ if

$$H'_r(z^{-1}) = H_r(z^{-1}) - \frac{\lambda_2}{2}I = \left[I + z^{-1}(A^* + B^*K) \right]^{-1} - \frac{\lambda_2}{2}I \quad (2.206)$$

is strictly positive real matrix, where: $\max_k(\lambda_2(k)) \leq \lambda_2 < 2$.

Proof: Assume that the plant is described by the input-output transfer operator of the true system $G_S \equiv S_{yu}(q^{-1})$ (see (2.126)) and it is operated in closed-loop with a digital controller $K(q^{-1})$ (without lack of generality) as it is shown in fig. 2.2.

The output of the plant operating in closed-loop is given by:

$$y(k+1) = -A^*y(k) + B^*u(k) + Aw(k+1) + p(k+1). \quad (2.207)$$

In deterministic case, $w(k+1) \equiv 0$, $p(k+1) \equiv 0$. Therefore the output of the plant becomes:

$$y(k+1) = -A^*y(k) + B^*u(k) = \varphi(k)\Gamma(\theta), \quad (2.208)$$

where

$$\begin{aligned} B^*(q^{-1}) &= C_0 \text{adj}(I - q^{-1}A_0)B_0, \\ A^*(q^{-1}) &= q(\det(I - q^{-1}A_0) - 1). \end{aligned} \quad (2.209)$$

For a fixed value of the estimated parameters, the *a posteriori* predictor of the closed-loop is expressed as:

$$\hat{y}(k+1) = \hat{y} \left[k+1 | \hat{\theta}(k+1) \right] = \phi(k) \Gamma \left[\hat{\theta}(k+1) \right]. \quad (2.210)$$

Then the closed-loop *a posteriori* prediction (output) error can be expressed as follows:

$$\varepsilon_{\text{CL}}(k+1) = y(k+1) - \hat{y}(k+1) = \varphi(k) \Gamma(\theta) - \phi(k) \Gamma \left[\hat{\theta}(k+1) \right]. \quad (2.211)$$

The aim of the following mathematical operations is to transform (2.211) to similar form (A.121), presented in section A.5, which is useful for passivity feedback system analysis.

The term $\varphi(k) \Gamma(\theta)$ in (2.211) can be written using (2.208) as follows:

$$\varphi(k) \Gamma(\theta) = -A^* y(k) + B^* u(k) + \hat{A}^* \hat{y}(k) - \hat{B}^* \hat{u}(k) - \hat{A}^* \hat{y}(k) + \hat{B}^* \hat{u}(k). \quad (2.212)$$

Using the following expressions:

$$\begin{aligned} u(k) &= -K y(k) + w(k), \\ \hat{u}(k) &= -K \hat{y}(k) + w(k), \\ \hat{y}(k+1) &= -\hat{A}^* \hat{y}(k) + \hat{B}^* \hat{u}(k) = \phi(k) \Gamma \left[\hat{\theta}(k+1) \right], \end{aligned}$$

the relationship (2.212) becomes:

$$\begin{aligned} \varphi(k) \Gamma(\theta) &= -A^* y(k) - B^* K y(k) + \hat{A}^* \hat{y}(k) + \hat{B}^* K \hat{y}(k) + \phi(k) \Gamma \left[\hat{\theta}(k+1) \right] \\ &= \phi(k) \Gamma \left[\hat{\theta}(k+1) \right] + \hat{A}^* \left(\hat{y}(k) - y(k) \right) + \hat{B}^* K \left(\hat{y}(k) - y(k) \right) \\ &= \phi(k) \Gamma \left[\hat{\theta}(k+1) \right] + \left(\hat{A}^* + \hat{B}^* K \right) \left(\hat{y}(k) - y(k) \right). \end{aligned} \quad (2.213)$$

Assuming that the $A^* = \hat{A}^*$, $B^* = \hat{B}^*$ and $\varepsilon_{\text{CL}}(k) = y(k) - \hat{y}(k)$ the equation (2.213) can be expressed as follows:

$$\varphi(k) \Gamma(\theta) = \phi(k) \Gamma [\theta] - \left(A^* + B^* K \right) \varepsilon_{\text{CL}}(k). \quad (2.214)$$

Then the equation for the *a posteriori* prediction error (2.211) can be written as:

$$\begin{aligned} \varepsilon_{\text{CL}}(k+1) &= \phi(k) \Gamma [\theta] - \left(A^* + B^* K \right) \varepsilon_{\text{CL}}(k) - \phi(k) \Gamma \left[\hat{\theta}(k+1) \right] \\ &= \phi(k) \left\{ \Gamma [\theta] - \Gamma \left[\hat{\theta}(k+1) \right] \right\} - \frac{q}{q} \left(A^* + B^* K \right) \varepsilon_{\text{CL}}(k) \\ &= \phi(k) \left\{ \Gamma [\theta] - \Gamma \left[\hat{\theta}(k+1) \right] \right\} - q^{-1} \left(A^* + B^* K \right) \varepsilon_{\text{CL}}(k+1). \end{aligned} \quad (2.215)$$

One can separate the *a posteriori* prediction error from (2.215) as follows:

$$\varepsilon_{\text{CL}}(k+1) = \left\{ I + q^{-1} (A^* + B^* K) \right\}^{-1} \phi(k) \left\{ \Gamma [\theta] - \Gamma \left[\hat{\theta}(k+1) \right] \right\}. \quad (2.216)$$

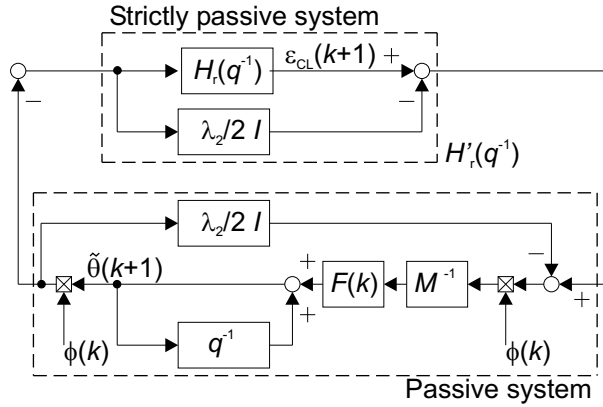


Figure 2.3: Transformed equivalent feedback systems associated to the PAA with time-varying gain for MIMO systems.

Moreover, if there is a unique solution of the function $\Gamma(\hat{\theta}) = M\hat{\theta}$ then (2.216) becomes:

$$\varepsilon_{CL}(k+1) = \left\{ I + q^{-1}(A^* + B^*K) \right\}^{-1} \phi(k)M \left\{ \theta - \hat{\theta}(k+1) \right\}. \quad (2.217)$$

Now, one can use the results, [LS79] and [Lan80], on the global asymptotic stability of the feedback system that is composed from a linear time-invariant feedforward block belonging to the class $L(\Lambda)$ and a linear time-varying feedback block belonging to the class $N(\Lambda)$.

Using the Theorem 2.1 from [Lan80], it can be shown that in a deterministic environment the sufficient condition for

$$\lim_{k \rightarrow \infty} \left\{ \Gamma[\theta] - \Gamma[\hat{\theta}(k+1)] \right\} = 0$$

together with the boundedness of $\varepsilon_{CL}(k+1)$ for any initial condition is that:

$$H_r'(z^{-1}) = H_r(z^{-1}) - \frac{\lambda_2}{2} I = \left[I + z^{-1}(A^* + B^*K) \right]^{-1} - \frac{\lambda_2}{2} I. \quad (2.218)$$

is strictly positive real matrix, where $\max_k(\lambda_2(k)) \leq \lambda_2 < 2$ and the term $H_r(z^{-1})$ is given by:

$$H_r(z^{-1}) = \left[I + z^{-1}(A^* + B^*K) \right]^{-1}. \quad (2.219)$$

The feedback system corresponding to the PAA time-varying gain for MIMO systems is shown in fig. 2.3.

2.8.3 Parametric Convergence Analysis in a Stochastic Environment

If the estimated parameters do not converge to the values corresponding to the deterministic case (under same input richness conditions), the resulting estimates is called *biased* estimates, and the corresponding error is termed *asymptotic bias*.

One of the objectives of closed-loop identification is to obtain asymptotic unbiased estimates in the presence of noise on the plant output. We shall use for this analysis the Ordinary Different Equation (ODE) approach [Lju77] and a specific results for a class of parameter estimation algorithms, [DL80].

Using a decreasing adaptation gain a possible equilibrium point $\hat{\theta} = \theta$ is given by the condition:

$$E\{\phi(k)p(k+1)\} = 0 \quad (2.220)$$

i.e., the observation vector and the image of the stochastic disturbance in the adaptation error equation should be uncorrelated. This observation is used in an averaging method for the analysis of PAA in a stochastic environment.

The Averaging Method for the Analysis of Adaptation Algorithms in a Stochastic Environment

We will try next to further explore the effect of the stochastic disturbances by examining the behavior of the PAA for large k , when the adaptation gain is very small by using an averaging approach. To simplify the analysis, without loss of generality, we will consider $F(k) = \frac{1}{k}$, *i.e.*, the adaptation gain $F(k)$ tends to zero when $k \rightarrow \infty$ (vanishing adaptation gain). Consider an algorithm with scalar decreasing adaptation gain of the form:

$$\hat{\theta}(k+1) = \hat{\theta}(k) + \frac{1}{k}\phi(k)\varepsilon_{\text{CL}}(k+1), \quad (2.221)$$

where $\phi(k)$ is the measurement vector and $\varepsilon_{\text{CL}}(k+1)$ is the *a posteriori* adaptation error. Taking a number of steps N such that $1 \ll N \ll k$, one obtains from (2.221):

$$\begin{aligned} \hat{\theta}(k+N+1) &= \hat{\theta}(k) + \sum_{i=0}^N \frac{1}{k+i}\phi(k+i)\varepsilon_{\text{CL}}(k+1+i) \\ &\approx \hat{\theta}(k) + \left(\sum_{i=0}^N \frac{1}{k+i}\right) \frac{1}{N+1} \sum_{i=0}^N \phi(k+i)\varepsilon_{\text{CL}}(k+1+i) \end{aligned} \quad (2.222)$$

In general, $\phi(k+i)$ and $\varepsilon_{\text{CL}}(k+1+i)$ will depend on the values of $\hat{\theta}(k+i)$. Since $\frac{1}{k+i}$ is very small, $\hat{\theta}(k+i)$ is slowly varying and one can write:

$$\frac{1}{N+1} \sum_{i=0}^N \phi(k+i)\varepsilon_{\text{CL}}(k+1+i) \approx \frac{1}{N+1} \sum_{i=0}^N \phi(k+i, \hat{\theta}(k))\varepsilon_{\text{CL}}(k+1+i, \hat{\theta}(k)), \quad (2.223)$$

where $\phi(k+i, \hat{\theta}(k))$ and $\varepsilon_{\text{CL}}(k+1+i, \hat{\theta}(k))$ are stationary quantities generated when $i \geq 0$, one fixes $\hat{\theta}(k+i) = \hat{\theta}(k)$.

If N is sufficiently large:

$$\begin{aligned} \frac{1}{N+1} \sum_{i=0}^N \phi(k+i, \hat{\theta}(k))\varepsilon_{\text{CL}}(k+1+i, \hat{\theta}(k)) &\approx E\left\{\phi(k+i, \hat{\theta}(k))\varepsilon_{\text{CL}}(k+1+i, \hat{\theta}(k))\right\} \\ &= f(\hat{\theta}(k)) \end{aligned} \quad (2.224)$$

and (2.222) becomes:

$$\hat{\theta}(k + N + 1) - \hat{\theta}(k) \approx \left(\sum_{i=0}^N \frac{1}{k + i} \right) f(\hat{\theta}(k)). \quad (2.225)$$

When $k \rightarrow \infty$ (2.225) can be considered as a discrete time approximation of the Ordinary Different Equation (ODE):

$$\frac{d\hat{\theta}}{d\tau} = f(\hat{\theta}), \quad (2.226)$$

where:

$$\Delta\tau_k^{N+1} \triangleq \sum_{i=0}^N \frac{1}{k + i},$$

$$f(\hat{\theta}) = E \left\{ \phi(k, \hat{\theta}) \varepsilon_{\text{CL}}(k + 1, \hat{\theta}) \right\}. \quad (2.227)$$

The correspondence between the discrete time k and the continuous time τ is given by:

$$\tau = \sum_{i=1}^{k-1} \frac{1}{i}. \quad (2.228)$$

Equation (2.225) can be interpreted as a discretized version of (2.226) with a decreasing sampling time. Equation (2.226) is an *average type equation* since the right hand side corresponds to the averaging of the product of stationary process $\phi(k, \hat{\theta}) \varepsilon_{\text{CL}}(k + 1, \hat{\theta})$ evaluated for a value $\hat{\theta}$ of the estimated parameters.

The asymptotic behavior of (2.226) gives information upon the asymptotic behavior of the parameter adaptation algorithm in stochastic environment, [Lju77]. To use this analogy, there is an assumption of the stationary process $\phi(k + 1, \hat{\theta})$ and $\varepsilon_{\text{CL}}(k + 1, \hat{\theta})$ existence for any possible value $\hat{\theta}(k)$. This signifies that the system generating $\phi(k)$ and $\varepsilon_{\text{CL}}(k + 1)$ are stable for any possible value $\hat{\theta}(k)$ generated by (2.221).

Once this assumption is satisfied, the ODE of (2.226) can be used for the analysis of the asymptotic behavior of the adaptation algorithm of (2.221). The following correspondences between (2.221) and (2.226) have been established:

1. The equilibrium points of (2.226) correspond to the only possible convergence points of the algorithm (2.221).
2. The global (or local) asymptotic stability of an equilibrium points θ^* of (2.226) corresponds to the global (local) convergence with probability 1 of the algorithm (2.221) towards θ^* .

The first step for effectively applying the associated ODE for convergence analysis is the evaluation of $f(\hat{\theta})$ which requires the knowledge of the equation governing $\varepsilon_{\text{CL}}(k + 1, \hat{\theta})$. However, for many recursive schemes used for identification the equation governing $\varepsilon_{\text{CL}}(k + 1, \hat{\theta})$ has a normalized form and therefore, one can obtain the following stochastic theorem which can be used, (Theorem 4.2.1 in [LLM97]):

Theorem 2 Consider the recursive parameter estimation algorithm:

$$\begin{aligned}\hat{\theta}(k+1) &= \hat{\theta}(k) + F(k)\phi(k)\varepsilon_{\text{CL}}(k+1), \\ F(k+1)^{-1} &= F(k)^{-1} + \lambda_2(k)\phi(k)\phi^T(k), \\ 0 \leq \lambda_2(k) < 2, \quad F(0) > 0.\end{aligned}\tag{2.229}$$

- Assume that the stationary process $\phi(k, \hat{\theta})$ and $\varepsilon_{\text{CL}}(k+1, \hat{\theta})$ can be defined for $\hat{\theta}(k) \equiv \hat{\theta}$.
- Assume that $\hat{\theta}(k)$ generated by the algorithm belongs infinitely often to the domain D_s for which the stationary process $\phi(k, \hat{\theta})$ and $\varepsilon_{\text{CL}}(k+1, \hat{\theta})$ can be defined.
- Assume that $\varepsilon_{\text{CL}}(k+1, \hat{\theta})$ is given by an equation of the normalized form:

$$\varepsilon_{\text{CL}}(k+1) = H_r(q^{-1})\phi^T(k, \hat{\theta})[\theta^* - \hat{\theta}] + \omega_r(k+1, \hat{\theta}),\tag{2.230}$$

where $H_r(q^{-1})$ is a discrete transfer operator, ratio of monic polynomials.

- Assume that either:
 - a) $\omega_r(k+1, \hat{\theta})$ is a sequence of independent equally distributed normal random variables $(0, \sigma)$, or
 - b) $E\{\phi(k, \hat{\theta}), \omega_r(k+1, \hat{\theta})\} = 0$.

Define convergence domain: $D_c = \{\hat{\theta} : (\phi^T(k, \hat{\theta}))[\theta^* - \hat{\theta}] = 0\}$.

Then if there is $\lambda_2 : \max_k (\lambda_2(k)) \leq \lambda_2 < 2$ such that:

$$H'_r(z^{-1}) = H_r(z^{-1}) - \frac{\lambda_2}{2}\tag{2.231}$$

is strictly positive real (SPR) discrete transfer function, one has:

$$\text{Prob}\left\{\lim_{k \rightarrow \infty} \hat{\theta}(k) \in D_c\right\} = 1.$$

Corollary 1: If $\phi^T(k, \hat{\theta})(\theta - \hat{\theta}) = 0$ has a unique solution (richness condition) then the condition $H'_r(z^{-1})$ given by (2.231) be strictly positive real implies that

$$\text{Prob}\left\{\lim_{k \rightarrow \infty} \hat{\theta}(k+1) = \theta\right\} = 1.$$

Corollary 2: If the adaptation gain given by (2.229) is replaced by $F(k) = \frac{1}{k}F$, $F > 0$.

Then the condition of (2.231) is replaced by the condition that $H_r(z^{-1})$ be a strictly positive real discrete transfer function.

The Closed-Loop Output-Error Algorithm

If a particular case of parametrization is assumed $\Gamma(\hat{\theta}) = M\hat{\theta}$, where M is a square matrix, then the *a posteriori* prediction error in the presence of output disturbance noise is given by:

$$\begin{aligned} \varepsilon_{\text{CL}}(k+1) &= \left\{ I + q^{-1}(A^* + B^*K) \right\}^{-1} \phi(k)M \left\{ \theta - \hat{\theta}(k+1) \right\} \\ &\quad + \left\{ I + q^{-1}(A^* + B^*K) \right\}^{-1} Ap(k+1). \end{aligned} \quad (2.232)$$

Consider the recursive parameter estimation algorithm given by (2.192) with $\lambda_1(k) = 1$. Define

$$\phi(k, \hat{\theta}) \triangleq \phi(k)|_{\hat{\theta}(k)=\hat{\theta}=\text{const.}}, \quad (2.233)$$

$$\varepsilon_{\text{CL}}(k+1, \hat{\theta}) \triangleq \varepsilon_{\text{CL}}(k+1)|_{\hat{\theta}(k)=\hat{\theta}=\text{const.}}, \quad (2.234)$$

$$D_s \triangleq \{ \hat{\theta} : \hat{A}(z^{-1})S(z^{-1}) + z^{-d}B(z^{-1})R(z^{-1}) = 0 \Rightarrow |z| < 1 \} \quad (2.235)$$

Assume that $\hat{\theta}(k)$ generated by the algorithm belongs infinitely often to the domain D_s for which the stationary process $\phi(k, \hat{\theta})$ and $\varepsilon_{\text{CL}}(k+1, \hat{\theta})$ can be defined.

Assume that $p(k)$ is a zero-mean stochastic process with finite moments independent of the reference sequence $r^*(k)$.

If

$$H'_r(z^{-1}) = H_r(z^{-1}) - \frac{\lambda_2}{2}I = \left[I + z^{-1}(A^* + B^*K) \right]^{-1} - \frac{\lambda_2}{2}I. \quad (2.236)$$

is strictly positive real (SPR) transfer function matrix, where: $\max_t (\lambda_2(t)) \leq \lambda_2 < 2$ then

$$\text{Prob} \left\{ \lim_{k \rightarrow \infty} \hat{\theta}(k) \in D_c \right\} = 1,$$

where $D_c = \{ \hat{\theta} : (\phi^T(k, \hat{\theta}))(\theta - \hat{\theta}) = 0 \}$. If, furthermore, $\phi^T(k, \hat{\theta})(\theta - \hat{\theta}) = 0$ has a unique solution (richness condition) then the condition $H'_r(z^{-1})$ given by (2.236) be strictly positive real matrix implies that

$$\text{Prob} \left\{ \lim_{k \rightarrow \infty} \hat{\theta}(k+1) = \theta \right\} = 1.$$

Proof: Assume that the plant is described by the input-output transfer operator of the true system $G_S \equiv S_{\text{yu}}(q^{-1})$ (see (2.126)) and it is operated in closed-loop with a digital controller $K(q^{-1})$ (without lack of generality) as it is shown in fig. 2.2.

Then the closed-loop *a posteriori* prediction (output) error can be expressed as follows:

$$\varepsilon_{\text{CL}}(k+1) = y(k+1) - \hat{y}(k+1) = \varphi(k)\Gamma(\theta) + Ap(k+1) - \phi(k)\Gamma \left[\hat{\theta}(k+1) \right]. \quad (2.237)$$

Using the results from (2.212) through (2.214) one can write:

$$\begin{aligned} \varepsilon_{\text{CL}}(k+1) &= \phi(k)\Gamma[\theta] - (A^* + B^*K)\varepsilon_{\text{CL}}(k) + Ap(k+1) - \phi(k)\Gamma \left[\hat{\theta}(k+1) \right] \\ &= \phi(k) \left\{ \Gamma[\theta] - \Gamma \left[\hat{\theta}(k+1) \right] \right\} - \frac{q}{q} (A^* + B^*K)\varepsilon_{\text{CL}}(k) + Ap(k+1) \\ &= \phi(k) \left\{ \Gamma[\theta] - \Gamma \left[\hat{\theta}(k+1) \right] \right\} - q^{-1} (A^* + B^*K)\varepsilon_{\text{CL}}(k+1) + Ap(k+1). \end{aligned} \quad (2.238)$$

One can separate the *a posteriori* prediction error from (2.238) as follows:

$$\begin{aligned} \varepsilon_{\text{CL}}(k+1) &= \left\{ I + q^{-1}(A^* + B^*K) \right\}^{-1} \phi(k) \left\{ \Gamma[\theta] - \Gamma[\hat{\theta}(k+1)] \right\} + \\ &\quad \left\{ I + q^{-1}(A^* + B^*K) \right\}^{-1} Ap(k+1). \end{aligned} \quad (2.239)$$

Moreover, if $\Gamma(\hat{\theta}) = M\hat{\theta}$ then (2.239) becomes:

$$\begin{aligned} \varepsilon_{\text{CL}}(k+1) &= \left\{ I + q^{-1}(A^* + B^*K) \right\}^{-1} \phi(k) M \left\{ \theta - \hat{\theta}(k+1) \right\} + \\ &\quad \left\{ I + q^{-1}(A^* + B^*K) \right\}^{-1} Ap(k+1). \end{aligned} \quad (2.240)$$

From (2.240), it results that, for $\hat{\theta}(k) \equiv \theta$,

$$\begin{aligned} \varepsilon_{\text{CL}}(k+1) &= \left\{ I + q^{-1}(A^* + B^*K) \right\}^{-1} \phi(k) M \left\{ \theta - \hat{\theta} \right\} + \\ &\quad \left\{ I + q^{-1}(A^* + B^*K) \right\}^{-1} Ap(k+1). \end{aligned} \quad (2.241)$$

which has the form of (2.230) of Theorem 2. Since $p(k)$ and $u(k)$ are independent, one concludes that $\phi(k, \hat{\theta})$ and $p(k+1, \hat{\theta})$ are independent for $\hat{\theta}(k) \equiv \theta$, and therefore:

$$E \left\{ \phi(k, \hat{\theta}), \omega_r(k+1, \hat{\theta}) \right\} = E \left\{ \phi(k, \hat{\theta}), \left\{ I + q^{-1}(A^* + B^*K) \right\}^{-1} Ap(k+1) \right\} = 0. \quad (2.242)$$

It results that global convergence toward unbiased estimates can be obtained with probability 1 if

$$H'_r(z^{-1}) = \left[I + z^{-1}(A^* + B^*K) \right]^{-1} - \frac{\lambda_2}{2} I \quad (2.243)$$

is strictly positive real transfer function matrix. Notice that a similar convergence condition has been obtained in the deterministic case.

2.8.4 Frequency Distribution of the Asymptotic Bias

The asymptotic expression for the bias distribution of model estimate plays important role in the analysis of different identification method.

One can consider that the various recursive parametric adaptation algorithms try to minimize recursively a quadratic criterion in terms of the plant model prediction error or, in general, in terms of the adaptation error, and the optimal value of the estimated vector is given by:

$$\hat{\theta}^* = \arg \min_{\hat{\theta} \in \mathcal{D}} \lim_{k \rightarrow \infty} \frac{1}{N} \sum_{k=1}^N \varepsilon_{\text{CL}}^2(k, \hat{\theta}) \approx \arg \min_{\hat{\theta} \in \mathcal{D}} E \{ \varepsilon_{\text{CL}}^2(k, \hat{\theta}) \}, \quad (2.244)$$

where \mathcal{D} is the domain of admissible parameters related to the model set and $\varepsilon_{\text{CL}}(k, \hat{\theta})$ is the adaptation error.

Under the basic assumption that:

$$\lim_{k \rightarrow \infty} \frac{1}{N} \sum_{k=1}^N \varepsilon_{\text{CL}}^2(k, \hat{\theta}) < \infty,$$

one can use the Parseval Theorem to get a frequency interpretation of the criterion (2.244):

$$\hat{\theta}^* \approx \arg \min \frac{1}{2\pi} \int_{-\pi}^{\pi} \phi_{\varepsilon_{\text{CL}}}(e^{j\omega}, \hat{\theta}) d\omega,$$

where $\phi_{\varepsilon_{\text{CL}}}(e^{j\omega}, \hat{\theta})$ is the spectral density of the adaptation error (the $\frac{1}{2\pi}$ can be dropped out). The (2.245) allows to assess the possible achievable performances of a given PAA, *e.g.* a frequency distribution of the asymptotic bias. This method has been introduced by Ljung, [Lju99].

Consider the closed-loop output error recursive algorithm where the excitation signal $w(k)$ is added to the controller output. Assume the system described by:

$$y(k) = S_{\text{yu}}(q^{-1})u(k) + H(q^{-1})e(k), \quad (2.245)$$

where $S_{\text{yu}}(q^{-1})$ is the input-output transfer operator of the model (2.120) and $H(q^{-1})$ is the input-output transfer operator of the noise model.

The *a posteriori* prediction error is given by:

$$\begin{aligned} \varepsilon_{\text{CL}}(k) &= y(k) - \hat{y}(k) \\ &= S_{\text{yu}}(q^{-1})u(k) + H(q^{-1})e(k) - \hat{S}_{\text{yu}}(q^{-1})\hat{u}(k) \\ &= S_{\text{yu}}(q^{-1})u(k) + H(q^{-1})e(k) - \hat{S}_{\text{yu}}(q^{-1})\hat{u}(k) + S_{\text{yu}}(q^{-1})\hat{u}(k) - S_{\text{yu}}(q^{-1})\hat{u}(k) \\ &= \left(S_{\text{yu}}(q^{-1}) - \hat{S}_{\text{yu}}(q^{-1}) \right) \hat{u}(k) + S_{\text{yu}}(q^{-1}) \left(u(k) - \hat{u}(k) \right) + H(q^{-1})e(k) \end{aligned} \quad (2.246)$$

Using the following relationship:

$$\begin{aligned} u(k) - \hat{u}(k) &= \left(-K(q^{-1})y(k) + w(k) \right) - \left(-K(q^{-1})\hat{y}(k) + w(k) \right) \\ &= -K(q^{-1}) \left(y(k) - \hat{y}(k) \right), \end{aligned}$$

the expression (2.246) becomes:

$$\begin{aligned} \varepsilon_{\text{CL}}(k) &= \left(S_{\text{yu}}(q^{-1}) - \hat{S}_{\text{yu}}(q^{-1}) \right) \hat{u}(k) - S_{\text{yu}}(q^{-1})K(q^{-1}) \left(y(k) - \hat{y}(k) \right) + H(q^{-1})e(k), \\ &= \left(S_{\text{yu}}(q^{-1}) - \hat{S}_{\text{yu}}(q^{-1}) \right) \hat{u}(k) - S_{\text{yu}}(q^{-1})K(q^{-1})\varepsilon_{\text{CL}}(k) + H(q^{-1})e(k), \end{aligned}$$

which can be write as follows:

$$\left(I + S_{\text{yu}}(q^{-1})K(q^{-1}) \right) \varepsilon_{\text{CL}}(k) = \left(S_{\text{yu}}(q^{-1}) - \hat{S}_{\text{yu}}(q^{-1}) \right) \hat{u}(k) + H(q^{-1})e(k). \quad (2.247)$$

One can separate $\varepsilon_{\text{CL}}(k)$ and (2.247) takes the form:

$$\begin{aligned} \varepsilon_{\text{CL}}(k) &= \left(I + S_{\text{yu}}(q^{-1})K(q^{-1}) \right)^{-1} \left\{ \left(S_{\text{yu}}(q^{-1}) - \hat{S}_{\text{yu}}(q^{-1}) \right) \hat{u}(k) + H(q^{-1})e(k) \right\}, \\ &= S_{\text{yp}}(q^{-1}) \left\{ \left(S_{\text{yu}}(q^{-1}) - \hat{S}_{\text{yu}}(q^{-1}) \right) \hat{u}(k) + H(q^{-1})e(k) \right\}, \end{aligned} \quad (2.248)$$

and using the expression:

$$\hat{u}(k) = \hat{S}_{yp}(q^{-1})w(k),$$

one obtains the following result:

For the case of the closed-loop methods, the formula for the estimated parameter vector when number of data $N \rightarrow \infty$ is

$$\hat{\theta}^* = \arg \min_{\hat{\theta} \in \mathcal{D}} \int_{-\pi}^{\pi} |S_{yp}|^2 \left[|S_{yu} - \hat{S}_{yu}|^2 |\hat{S}_{yp}|^2 \phi_w + |H|^2 \phi_e \right] d\omega,$$

where $S_{yp}(\omega)$ corresponds to the true output sensitivity transfer function matrix between the output disturbances vector p and the plant outputs vector y , $\hat{S}_{yp}(\omega)$ is the estimated output sensitivity transfer function matrix, $\phi_w(\omega)$ corresponds to the spectral density of the external excitation signals vector w , and the $\phi_e(\omega)$ is the spectral density of the noise vector e . This formula shows that the noise does not affect the parameter estimation.

2.9 Simulation Experiments for SISO Systems

In order to compare the different state-space identification algorithms and study their behavior, several simulation experiments are considered. In the case of SISO systems, the experiments are performed in the closed-loop identification scheme (fig. A.1) using the basic CLOE approach. One should note that, independently of the criterion taken, the simulation results are controller-dependent. A SISO test plant \mathcal{S} is operated by a digital PID controller \mathcal{K} which represents a special case of the RST controller depicted in fig. A.1.

A Monte Carlo simulation with N runs is carried out in order to examine the statistical behavior of the algorithm with respect to different realizations of the output disturbance noise $p(k)$. One can introduce the mean value and the standard deviation of the j th estimated parameter $\hat{\theta}_j$, respectively:

$$\bar{\theta}_j = \frac{1}{N} \sum_{i=1}^N \hat{\theta}_j^i, \quad \sigma_j = \sqrt{\frac{1}{N} \sum_{i=1}^N (\hat{\theta}_j^i - \bar{\theta}_j)^2}, \quad (2.249)$$

where $\hat{\theta}_j^i$ is the final estimate of the j th parameter obtained in the i th simulation run.

In order to evaluate the algorithm, the mean value of the parametric error and the standard deviation of the estimated parameter vector are used. These criteria are defined as follows:

$$\bar{\theta} = \frac{1}{d_0} \sum_{j=1}^{d_0} |\bar{\theta}_j - \theta_j|, \quad (2.250)$$

$$\bar{\sigma} = \frac{1}{d_0} \sum_{j=1}^{d_0} \sigma_j, \quad (2.251)$$

where θ_j is the true value of the j th estimated parameter.

In addition, the parametric distance defined by:

$$D_{\theta}^i(k) = \sqrt{\sum_{j=1}^{d_0} [\hat{\theta}_j^i(k) - \theta_j]^2}, \quad (2.252)$$

is used to observe the evolution of the estimated parameter vector during one simulation run i .

The following discrete-time state-space representation of the test plant \mathcal{S} is considered:

$$\begin{aligned} A_{\mathcal{S}} &= \left[\begin{array}{c|c} 0.9986 & 0.0001612 \\ \hline -17.21 & 0.9902 \end{array} \right], & B_{\mathcal{S}} &= \left[\begin{array}{c} 1.648 \cdot 10^5 \\ 0.2031 \end{array} \right], \\ C_{\mathcal{S}} &= \left[\begin{array}{c|c} 1.794 \cdot 10^5 & 0 \end{array} \right]. \end{aligned}$$

This test plant represents a DVD focus actuator second-order system, [BBVH03]:

$$\mathcal{S}: \quad G_{\mathcal{S}}(q^{-1}) = \frac{2.9573q^{-1} + 2.9490q^{-2}}{1 - 1.9887q^{-1} + 0.9915q^{-2}}. \quad (2.253)$$

The controller \mathcal{K} has the following form:

$$\mathcal{K}: \quad K(q^{-1}) = \frac{0.07179 - 0.1382q^{-1} + 0.06654q^{-2}}{1 - 0.8021q^{-1} - 0.1979q^{-2}}. \quad (2.254)$$

The sampling period is $T_s = 0.162$ ms which corresponds to 20 times higher sampling period than is used in an actual DVD drive.

In order to test the proposed state-space identification algorithms, various state-space representations $\{A_0(\theta), B_0(\theta), C_0(\theta)\}$ of the plant \mathcal{S} operated by the controller \mathcal{K} are identified in the following simulation experiments.

A pseudo-random binary sequence (PRBS), generated by a 11-bit register, added to the controller output u is taken as excitation signal w . Its value alternates between $+0.02$ and -0.02 and its clock frequency is equal to $f_s/2$. A zero-mean Gaussian white noise of standard deviation σ_e is used as the output disturbance signal $e(k)$ and a noise filter $H = 1$ has been considered.

2.9.1 Simulation Experiment 1

The following state-space representation of the test plant \mathcal{S} is considered:

$$\begin{aligned} A_0(\theta) &= \left[\begin{array}{c|c} -\theta_1 & -\theta_2 \\ \hline 1 & 0 \end{array} \right], & B_0(\theta) &= \left[\begin{array}{c} 1 \\ 0 \end{array} \right], \\ C_0(\theta) &= \left[\begin{array}{c|c} \theta_3 & \theta_4 \end{array} \right], \end{aligned}$$

for the parameter vector defined by:

$$\theta = [\theta_1, \theta_2, \theta_3, \theta_4]^T.$$

This parametrization (a canonical controller form) corresponds to the elementary case of direct estimation of the transfer function coefficients:

$$\Gamma(\theta) = \theta = \begin{bmatrix} \theta_1 \\ \theta_2 \\ \theta_3 \\ \theta_4 \end{bmatrix}, \quad \Gamma'_{\theta}(\theta) = I = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Table 2.2: Settings of the SISO experiment 1.

Initial parameter vector estimate	$\hat{\theta}(0) = [0, 0, 0, 0]^T$
Standard deviation of white noise	$\sigma_e = 0.00001$
RLS-2 initial matrix adaptation gain	$F(0) = 10000I$
RLS-2 variation profile sequence $\lambda_1(k)$	$\lambda_1(k) \equiv 0.98$
RLS-2 variation profile sequence $\lambda_2(k)$	$\lambda_2(k) \equiv 1$
IGM-1, IGM-2 scalar adaptation gain	$\alpha_1 = \alpha_2 = 300$
GM-2 scalar adaptation gain	$\alpha = 0.12$
Number of simulation steps	$k_{\text{end}} = 2047$
Number of simulation runs	$N = 100$

Table 2.3: Results of the SISO experiment 1.

	$\bar{\theta}[\times 10^{-4}]$	$\bar{\sigma}[\times 10^{-4}]$
RLS-2	9.972	± 2.214
IGM-1	15.416	± 1.216
IGM-2	15.416	± 1.216
GM-2	15706	± 357

This model structure is linear in its inputs but nonlinear in its parameters since the right-hand side of closed-loop predictor (2.70) involves the product of parameters by model outputs that depend on $\hat{\theta}$. Identifiability of this model is satisfied (Laplace transform approach [BÅ70]). The true values of the estimated parameters are straightforward:

$$\theta = \Gamma(\theta) = \begin{bmatrix} -1.9887 \\ 0.9915 \\ 2.9573 \\ 2.9490 \end{bmatrix}.$$

Table 2.2 gives an overview of the simulation settings. One can see that a constant forgetting factor of 0.98 is selected as an adaptation gain variation profile for the RLS-2 algorithm.

The mean value of the parametric error $\bar{\theta}$ and the standard deviation of the estimated parameter vector $\bar{\sigma}$ are compared for the different algorithms in table 2.3. One can see that the RLS-2 algorithm gives the best estimate, as expected. On the other hand, the GM-2 algorithm provides a very inaccurate result. Since the Jacobian matrix $\Gamma'_\theta(\theta)$ is an identity matrix, the algorithms IGM-1 and IGM-2 are equivalent. These facts can be also observed in fig. 2.4, where the evolution of the parametric distance $D_\theta(k)$ is illustrated. The convergence speed of the GM-2 algorithm is very low in comparison with the RLS-2 algorithm. In addition, it becomes unstable if a wrong value of the scalar adaptation gain α is selected. As the examples, the evolutions of the individual parameters for the RLS-2 and IGM-2 algorithms are illustrated in figs. 2.5 and 2.6, respectively. Mean values $\bar{\theta}_j$ and standard deviations of the individual estimated parameters σ_j for the RLS-2 and IGM-2 algorithms are given in table 2.4.

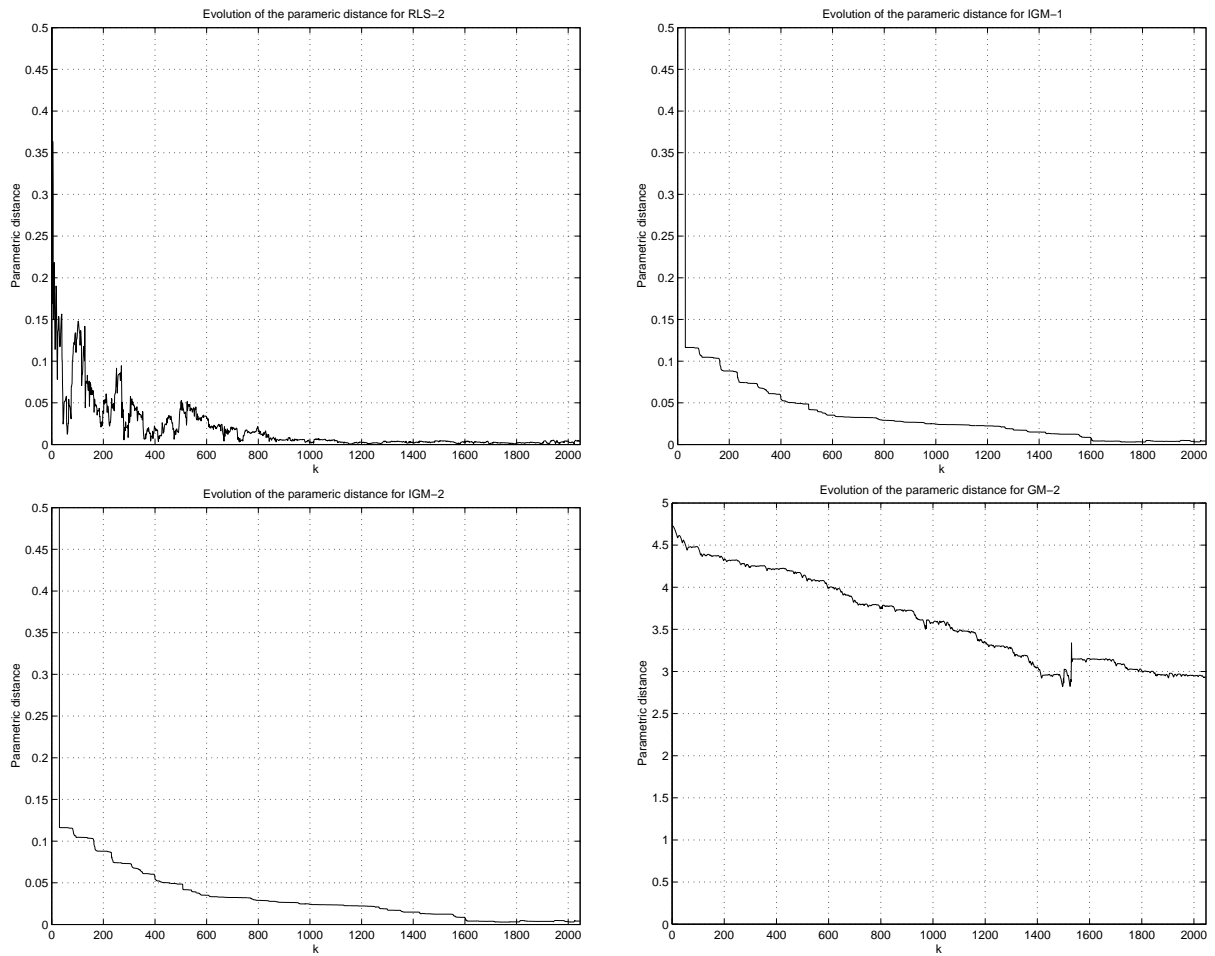


Figure 2.4: Evolution of the parametric distance for one simulation run (SISO experiment 1).

Table 2.4: Results of the SISO experiment 1, mean values and standard deviations of the individual parameters.

	j	1	2	3	4
RLS-2	$\hat{\theta}_j$	-1.9886	0.9914	2.9586	2.9466
	$\sigma_j [\times 10^{-3}]$	± 0.0364	± 0.0316	± 0.3301	± 0.4876
IGM-2	$\hat{\theta}_j$	-1.9891	0.9920	2.9536	2.9506
	$\sigma_j [\times 10^{-3}]$	± 0.1225	± 0.0795	± 0.1293	± 0.1553

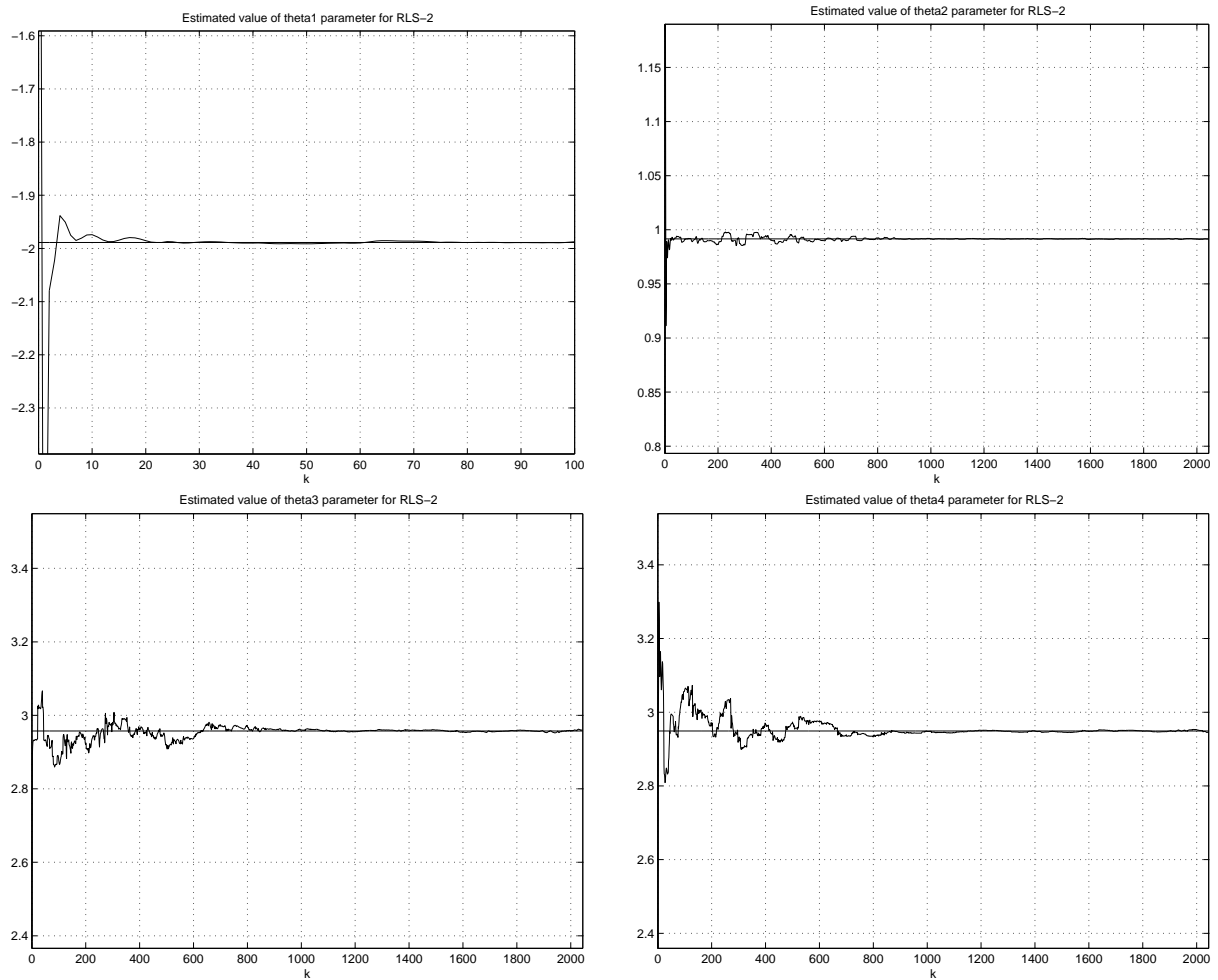


Figure 2.5: Evolution of the estimated parameters for one simulation run (RLS-2, SISO experiment 1).

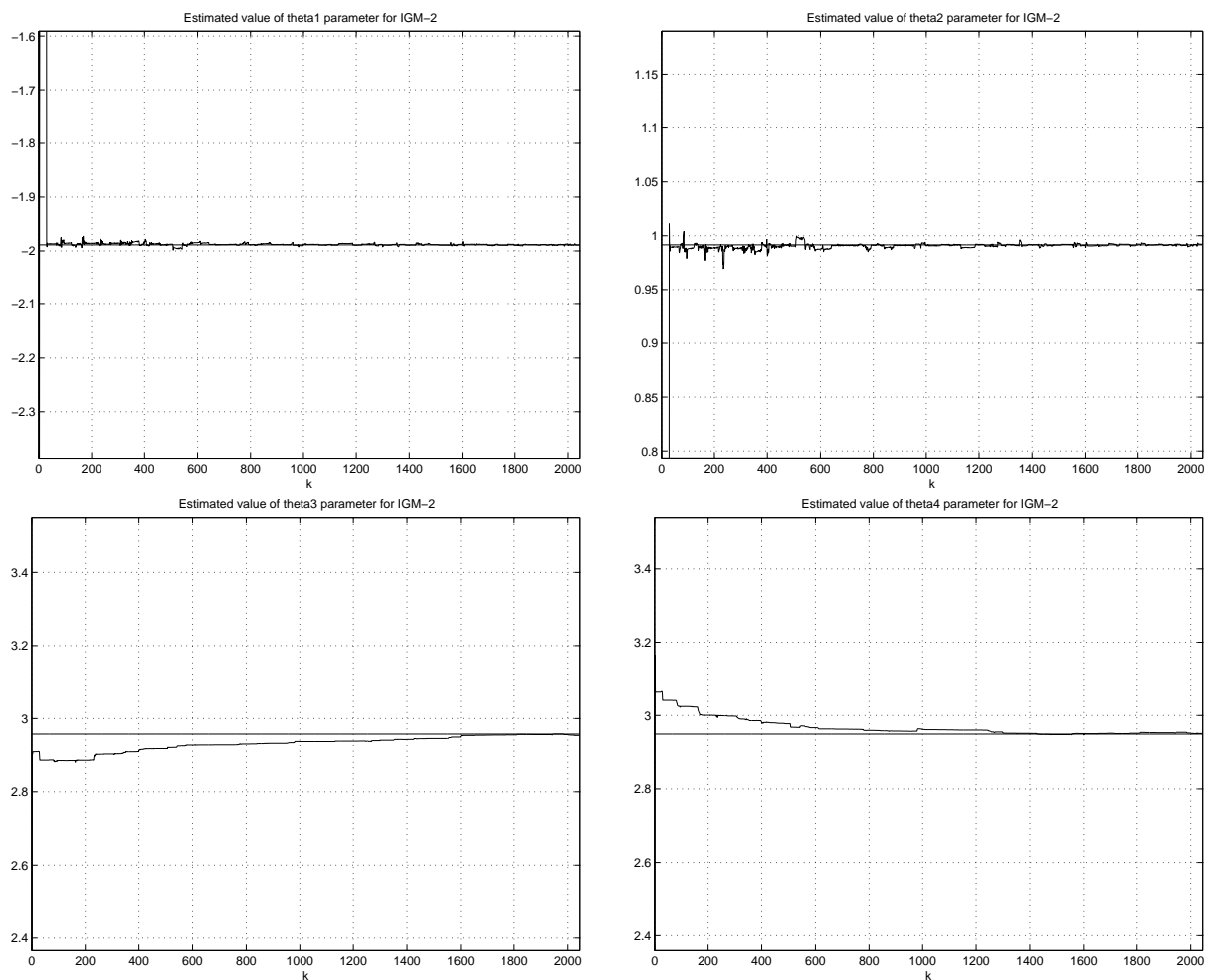


Figure 2.6: Evolution of the estimated parameters for one simulation run (IGM-2, SISO experiment 1).

2.9.2 Simulation Experiment 2

The following state-space representation of the test plant \mathcal{S} is considered:

$$\begin{aligned} A_0(\theta) &= \left[\begin{array}{c|c} -\theta_1 + \theta_2 & -\theta_3 + \theta_4 \\ \hline 1 & 0 \end{array} \right], & B_0(\theta) &= \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \\ C_0(\theta) &= \left[\theta_1 + \theta_2 + \theta_3 + \theta_4 \mid -\theta_3 \right], \end{aligned}$$

for the parameter vector defined by:

$$\theta = [\theta_1, \theta_2, \theta_3, \theta_4]^T.$$

Since the state-space matrices are given in a canonical controller form, one can immediately determine the vector of transfer function coefficients $\Gamma(\theta)$ and its Jacobian matrix $\Gamma'_\theta(\theta)$:

$$\Gamma(\theta) = \begin{bmatrix} \theta_1 - \theta_2 \\ \theta_3 - \theta_4 \\ \theta_1 + \theta_2 + \theta_3 + \theta_4 \\ -\theta_3 \end{bmatrix}, \quad \Gamma'_\theta(\theta) = \begin{bmatrix} 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 1 & 1 & 1 & 1 \\ 0 & 0 & -1 & 0 \end{bmatrix}.$$

This model structure is linear in its inputs but nonlinear in its parameters since the right-hand side of closed-loop predictor (2.70) involves the product of parameters by model outputs that depend on $\hat{\theta}$. Identifiability of this model is satisfied (Laplace transform approach [BÅ70]). The true values of the estimated parameters are calculated to obtain the transfer operator of the test plant \mathcal{S} :

$$\theta = \begin{bmatrix} 3.9291 \\ 5.9179 \\ -2.9490 \\ -3.9406 \end{bmatrix} \Rightarrow \Gamma(\theta) = \begin{bmatrix} -1.9887 \\ 0.9915 \\ 2.9573 \\ 2.9490 \end{bmatrix}.$$

Table 2.5 gives an overview of the simulation settings. One can see that the constants of 0.11 and 300 are selected as the scalar adaptation gains for the IGM-1 and IGM-2 algorithms, respectively. In addition, the four periods of the excitation signal $w(k)$ has been applied on controller output $u(k)$ for the IGM-1, IGM-2 and GM-2 algorithms in order to show their convergence properties in more details.

The simulation results are summarized in tables 2.6 and 2.7. Figure 2.7 shows that the convergence speed of the RLS-2 algorithm remains unbeaten. The RLS-2 algorithm offers the best estimation. The IGM-2 algorithm suffers from a lower convergence speed, the IGM-1 and GM-2 algorithms have the worst convergence speed irrespective of the scalar adaptation gain settings.

Figures 2.8 and 2.9 illustrate the evolution of the individual parameters in the case of the RLS-2 and IGM-2 algorithms, respectively. One can see that the IGM-2 algorithm requires more simulation steps than the RLS-2 algorithm to come near the true values.

The other state-space identification experiments are illustrated in appendix B.

2.10 Conclusions

This chapter introduces a new approach to the problem of state-space identification. As a starting point, the family of closed-loop output-error algorithms has been considered.

Table 2.5: Settings of the SISO experiment 2.

Initial parameter vector estimate	$\hat{\theta}(0) = [0, 0, 0, 0]^T$
Standard deviation of white noise	$\sigma_e = 0.00001$
RLS-2 initial matrix adaptation gain	$F(0) = 10000I$
RLS-2 variation profile sequence $\lambda_1(k)$	$\lambda_1(k) \equiv 0.98$
RLS-2 variation profile sequence $\lambda_2(k)$	$\lambda_2(k) \equiv 1$
IGM-1 scalar adaptation gain	$\alpha_1 = 0.11$
IGM-2 scalar adaptation gain	$\alpha_2 = 300$
GM-2 scalar adaptation gain	$\alpha = 0.015$
Number of simulation steps (RLS-2)	$k_{\text{end}} = 2047$
Number of simulation steps (IGM-1, IGM-2, GM-2)	$k_{\text{end}} = 8185$ (4 periods)
Number of simulation runs	$N = 100$

Table 2.6: Results of the SISO experiment 2.

	$\bar{\theta}[\times 10^{-4}]$	$\bar{\sigma}[\times 10^{-4}]$
RLS-2	21.45	± 4.3247
IGM-1	7868	± 0.4003
IGM-2	159.4	± 1.3768
GM-2	8987	± 13.65

Table 2.7: Results of the SISO experiment 2, mean values and standard deviations of the individual parameters.

	j	1	2	3	4
RLS-2	$\bar{\theta}_j$	3.9274	5.9160	-2.9466	-3.9381
	$\sigma_j[\times 10^{-3}]$	± 0.3510	± 0.3778	± 0.4876	± 0.5132
IGM-2	$\bar{\theta}_j$	3.9153	5.8991	-2.9352	-3.9233
	$\sigma_j[\times 10^{-3}]$	± 0.1763	± 0.1372	± 0.1284	± 0.1087

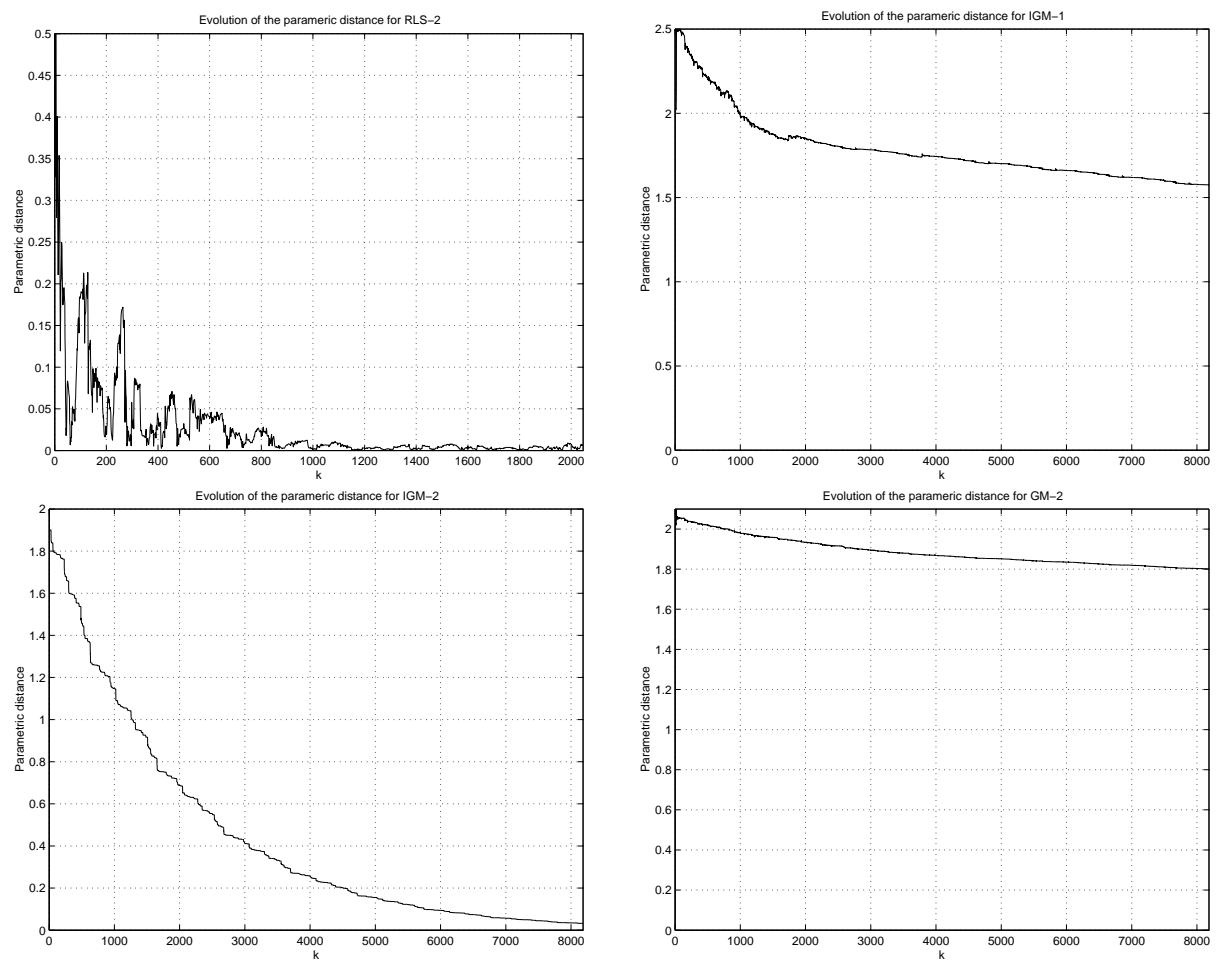


Figure 2.7: Evolution of the parametric distance for one simulation run (SISO experiment 2).

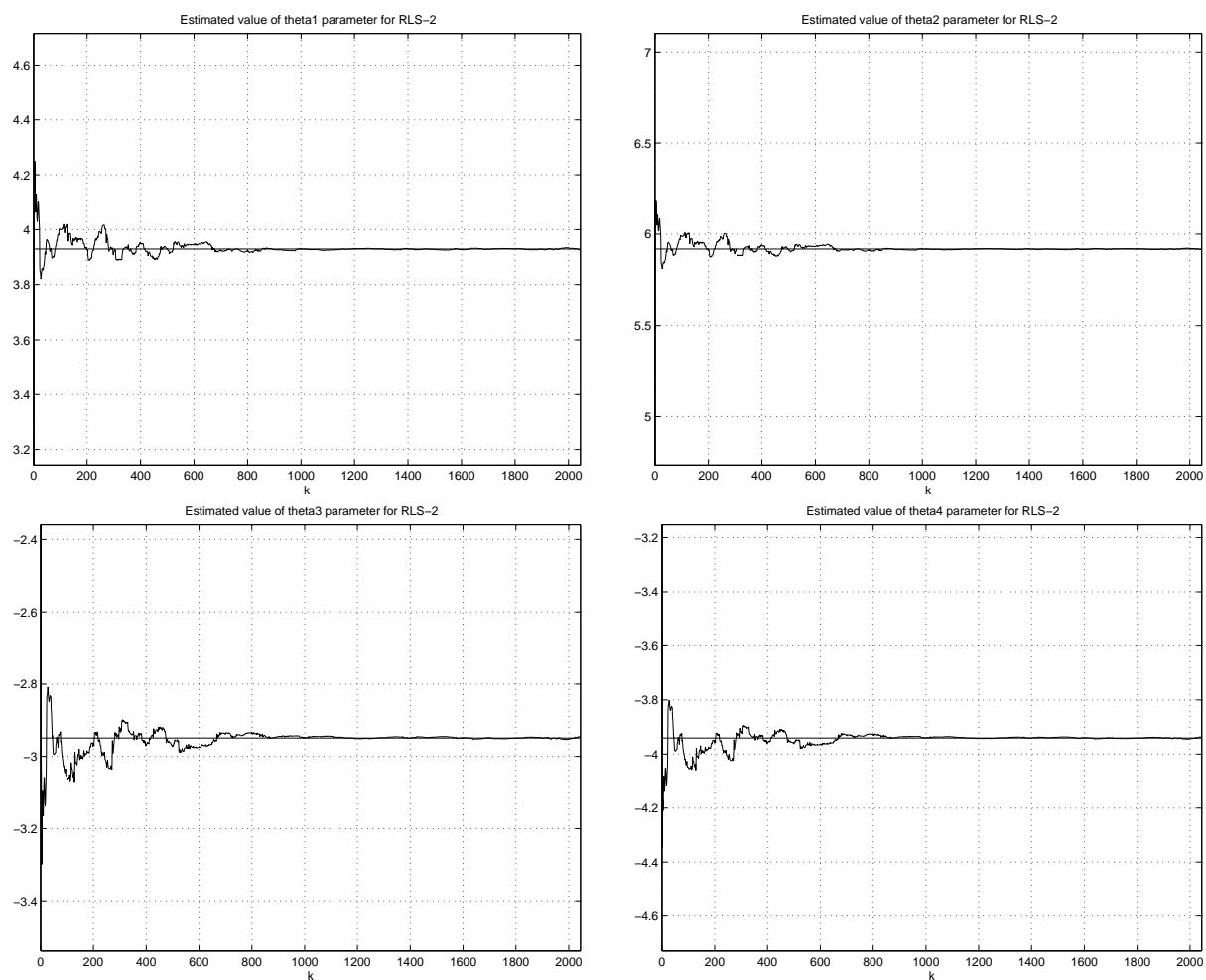


Figure 2.8: Evolution of the estimated parameters for one simulation run (RLS-2, SISO experiment 2).

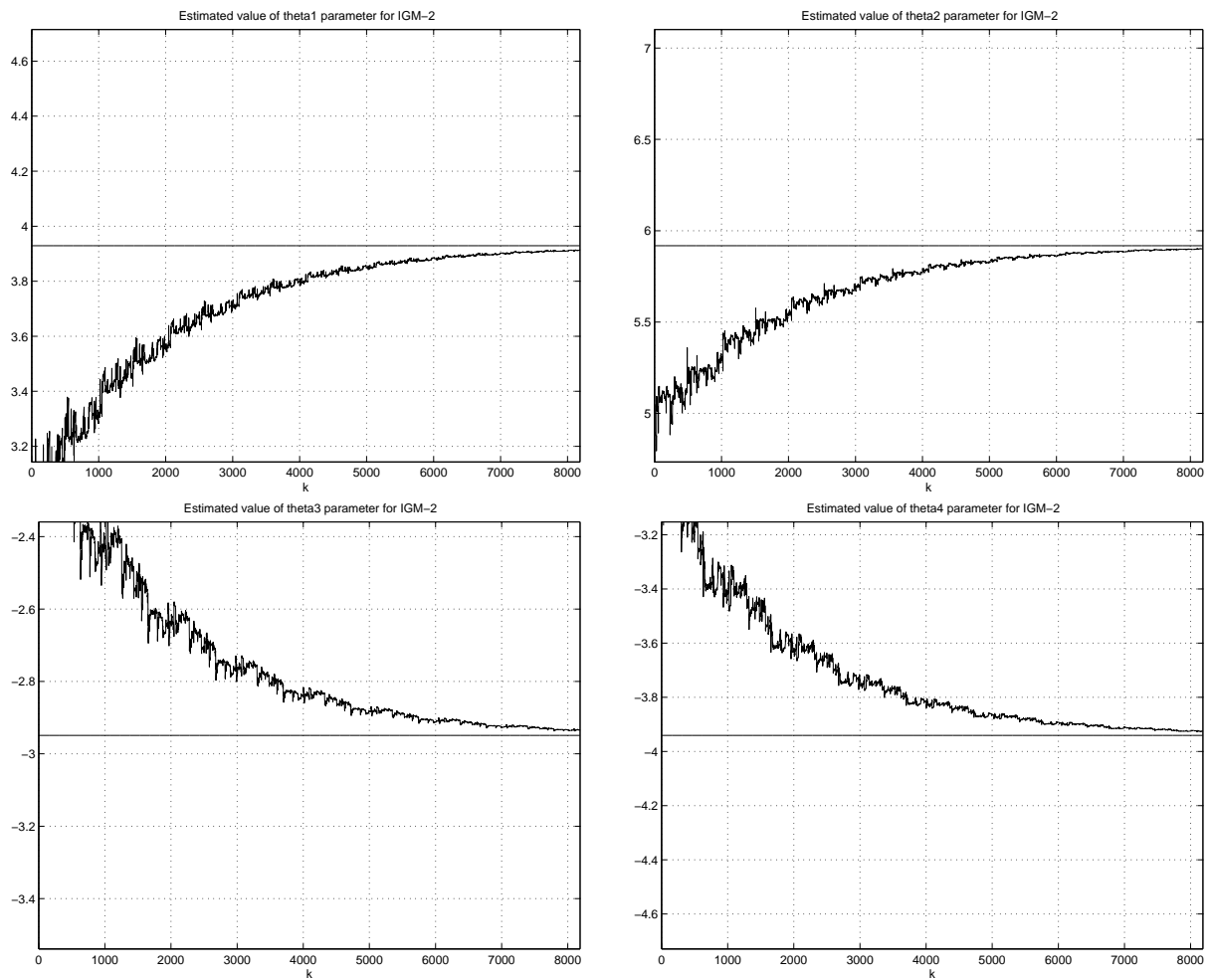


Figure 2.9: Evolution of the estimated parameters for one simulation run (IGM-2, SISO experiment 2).

These algorithms can estimate the parameters of an ordinary black-box model, utilizing the closed-loop identification scheme. This is of particular importance in adaptive control. It has been shown in [LK97b] that the closed-loop identification algorithms are able to provide a better predictor for the closed-loop via an estimation of the plant model in comparison with algorithms operating in the traditional open-loop scheme.

The closed-loop output-error approach has been expanded in order to identify state-space discrete-time models with given structure in this chapter. This issue has been treated in a very general way.

Several novel recursive parameter adaptation algorithms have been proposed for this purpose. The RLS-2 algorithm, which has been derived from the well-known least squares criterion, incorporates a numerical optimization component represented by computation of a matrix pseudoinverse. On the other hand, the IGM-1, IGM-2 and GM-2 algorithms, which all use the gradient technique to minimize a one-step quadratic criterion, have the advantage of a closed analytic form.

Extensive simulation experiments have been performed in order to study the behavior of the newly proposed algorithms. Both SISO and MIMO test plants have been considered. One should refer to sections 2.9, B.2 and B.3 to obtain the particular results and their interpretations.

In general, the RLS-2 algorithm possesses the fastest convergence. The strength of the IGM-1 and IGM-2 algorithms lies in robustness, *i.e.* in their insensitivity to the realization of the disturbance noise and to various initial settings (initial parameter vector estimate, adaptation gain settings).

All the algorithms suffer from the typical disadvantages of gradient iterative methods. Their convergence and stability are not guaranteed and they are not able to avoid the trap of a local minimum of the objective criterion generally. Therefore, the stability in a deterministic environment, parametric convergence and frequency distribution of the asymptotic bias have been analyzed. These properties are always related to the actual state-space model and its parametrization.

The particular case of parametrization, where the transformation function is linear towards the parameter vector, has been assumed. The properties of the algorithms have been shown for this parametrization and one can conclude that the canonical state-space representation is the best theoretical case. Nevertheless, the parametrization function is given in general way and the other parametrization can be of further interest for research.

Furthermore, the identifiability of each individual estimated parameter should be considered when constructing the parametrization of a state-space model. It generally depends on two conditions: Identifiability of the transfer function coefficients and determination of the estimated parameters from the transfer function coefficients.

One should also note that the estimation of physical parameters of continuous-time model via identification of a discrete-time state-space model requires the nontrivial transformation from a continuous-time to a discrete-time model. This transformation is generally a non-linear and therefore, the analysis of the convergence, stability and bias distribution of the related discrete-time state-space model is difficult.

The DVD Player System Description

3.1 Introduction

This chapter is devoted to the description of the DVD player, and to the definition of modelling the photodetector characteristic problem together with the control system problem formulation.

Before going further in treating the modelling and control problem, we have considered necessary to describe the general principles which make the system work. This would render the subject of this work more accessible and easier to understand to the reader.

The contribution of our investigation in this chapter can thus be described in the context of a pure description of the system under study, including the presentation of optics and the description of principles used to generate the servo and the read-out signals.

In section 3.2, the DVD drive architecture is presented. Section 3.3 gives a detailed description of the optical procedure used in industrial systems for generating the position errors and the data read-out signals. The problem of modelling the photodetector characteristic is formulated here. In section 3.4, the servo mechanical system is presented. Section 3.5 describes the disturbance sources coming to the system. Specification requirements on the focus and radial control loops are given in section 3.6.

Finally, in section 3.7, the industrial performance limitation together with the comparison between the CD and DVD player control difficulties are presented. Here, the control problem is also treated. Conclusions are drawn in section 3.8.

3.2 DVD Drive Architecture

Almost all DVD drives rely on the same system architecture. In general, the drive can be divided into a basic engine and a data path, symbolically separated by a control signals bus, as presented in fig. 3.1. Apart from other specific DVD functions, the data path provides also the interface between the basic engine and the host system (usually a personal computer).

The rotating disk is read out without any mechanical contact with its surface. An Optical Pick-up Unit (OPU) generates a laser beam to the disk and receives back the reflected light, optically modulated by the disk geometrical structure. The OPU contains, among other components, a semiconductor laser, optical elements to guide the laser beam and a photodetector used to transform the optical power into electrical current.

By properly processing this current, two servo signals are derived for positioning the laser beam along the disk radius and spiral, respectively. At the same time, a high-

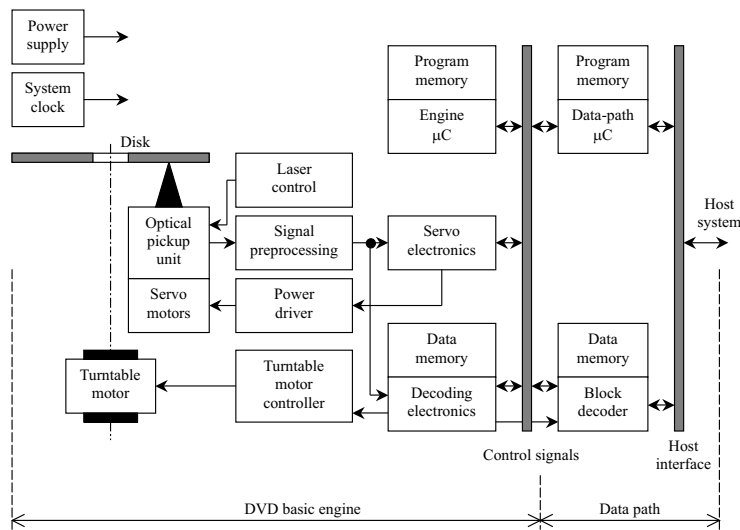


Figure 3.1: Schematic view of the DVD architecture.

frequency signal, carrying the information recorded on the disk, is also extracted and forwarded to the decoding electronics.

The laser beam displacement along the vertical and radial directions, with respect to the disk, is accomplished by two voice-coil motors. These actuators keep the laser beam on track and in focus by executing fine displacements. An additional servo motor is also used to perform large displacements of the laser spot along the disk radial direction. This electromechanical construction is usually called as “two-stage” or “sledge-actuator” servo [Sta98]. The functionalities of all electromechanical components are governed by a firmware running on a micro Controller (μC). The decoding electronics process the incoming high-frequency signal and regenerates the digital data, stored on the disk, that are then processed by the data path and send to the host system. In the following sections, a more detailed description of the DVD optics and servo mechanical subsystem is presented.

3.3 The Optics

3.3.1 Optical Pick-up Organization

In fig. 3.2, a schematic view of the DVD mechanism for a single-layer single-side optical disk is shown. The system is composed of an Optical Pick-up Unit (OPU) that retrieves data from the disk. The optical disk is rotated by a Direct Current (DC) turntable motor with the spindle rotation frequency f_{rot} .

A laser diode (1) located in the OPU emits a laser beam which is guided through the optical elements (2, 3, 4, 5, 7) to the disk information layer (10). An objective lens (7) is the last optical element for laser beam focusing on the disk information layer (10). The objective lens (7) can be moved in vertical direction z , to give focusing action, and in radial direction x , to perform track following. It is suspended by leaf springs and its position is controlled by electromagnets in vertical and radial directions (z , x), respectively.

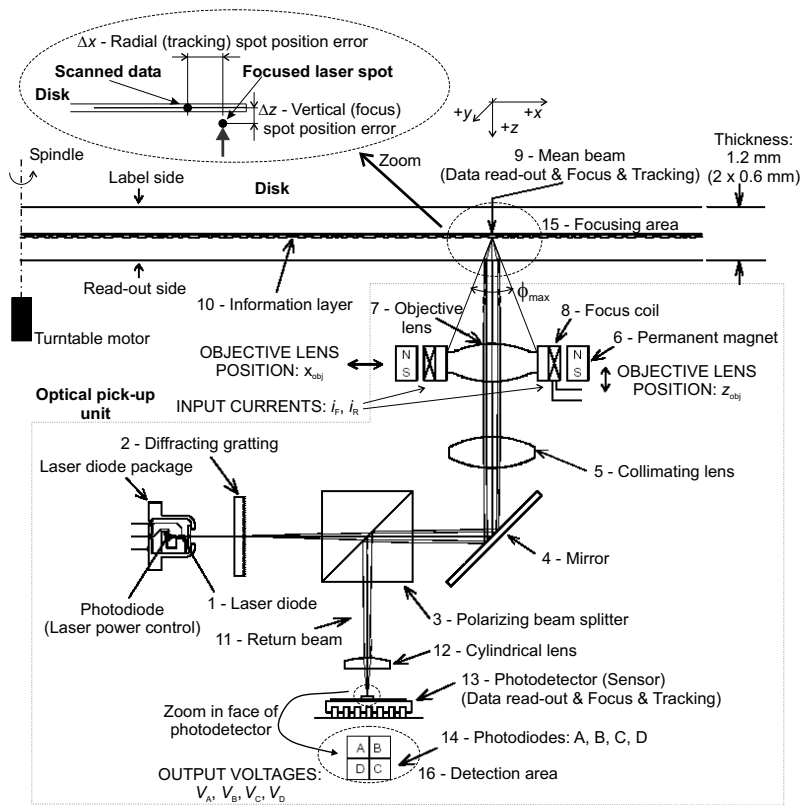


Figure 3.2: Optical pick-up unit organization of the DVD mechanism for a single-layer single-side optical disk.

In fig. 3.2, only the focus electromagnet elements (*i.e.* a focus coil (8) and a permanent magnet (6)) are shown, because the radial electromagnet elements are placed in perpendicular direction and cannot be easily shown on the same figure. To focus the incident beams on the disk information layer (10), the focus coil (8) is placed in the electromagnetic field of the permanent magnet (6), under the influence of the input current i_F .

The mean beam (9) of incident rays meets the information layer (10) at a focusing area (15) whose zoom is also illustrated in fig. 3.2. Here, the focused laser spot position error (in next called shortly the **spot position error** (Δz , Δx)) is the distance between the laser spot and the real track position in vertical and radial directions (z , x).

The return beam (11) passes through the objective lens (7) and the optical elements (5, 4). Then the return beam is splitted by a polarizing beam splitter (3) in perpendicular direction to the beam emitted by the laser diode (1). Therefore the return beam (11) passes through a cylindrical lens (12), shaping the laser spot that is falling on a photodetector (13). The photodetector (13), composed of the four photodiodes (14) (A, B, C, D), is a detection area (16) for the returned beam (11).

Since only light touches the disk information layer (10), the light intensity distribution of the returned beam (11) carries knowledge about the information layer (10) position with respect to the objective lens position (z_{obj} , x_{obj}). The four photodiodes (14) (A, B, C, D) are used to measure this light intensity distribution of the laser spot on the photodetector (13). Therefore to generate the focus and radial error signals (e_F , e_R) for the controllers, the four output voltages from the photodiodes (14) (V_A , V_B , V_C , V_D) are

completely sufficient. The recorded data reconstruction is also realized from the output voltages (V_A, V_B, V_C, V_D).

The goal of the control design is to minimize the focused laser spot position error ($\Delta z, \Delta x$) (measured from the output voltages (V_A, V_B, V_C, V_D)), despite the presence of internal and external disturbances, in order to preserve the read-out signal quality.

Digital Versatile Disk standard

Since the disk itself is part of the optical subsystem of a DVD drive, it will co-determine the way the recorded information is being processed. Fortunately, the disk has been standardized by the five books Book A-E, [ECM01].

A DVD disk is a transparent plastic (polycarbonate) disk carrying a continuous spiral of impressed pits. The impressed surface is covered with a thin metallic layer, on top of which another plastic layer is being used for protection. The laser beam reads the profiled polycarbonate surface from below, perceiving the impressed pits as bumps and being reflected back by the metallic layer.

On a DVD, the digital information is organized as sectors of 2048 bytes plus 12 bytes of header data. Blocks of 16 sectors are error protected using RSPC (Reed Solomon Product Code). In addition, DVD uses an 8 to 16 modulation scheme, giving pit lengths of 3 to 14 (minimum to maximum length), compared with CDs 3 to 11 obtained with EFM (Eight to Fourteen) modulation. This makes the jitter specification (defined as the standard deviation of the time variation of the read-out signal passed through the given filter (equalizer)) slightly tighter for DVDs, see [ECM96] and [ECM01].

Data are physically contained on a spiral-shaped track that evolves from the innermost to the outermost position of the disk. The track is constituted by a sequence of *pits* and *lands* of varying length as shown in fig. 3.3. The shape of the pits is prefixed and their length can be distinguished because of the discrete distribution along the track, as shown in fig. 3.4. Here, the pit length L_{pit} and land length L_{land} are proportional with the channel bit length $T_{\text{pit}} = 0.133 \mu\text{m}$, *i.e.* $L_{\text{pit}} = 3 - 14 T_{\text{pit}}$ and $L_{\text{land}} = 3 - 13 T_{\text{pit}}$ for a DVD. The distance q_{real} of two subsequent track locations along the disk radius (*track pitch*) is equal to $0.74 \mu\text{m}$. The width of the pit is $\gamma \approx 0.4 \mu\text{m}$.

Another important parameter of the pit geometry is its real mechanical depth $d_{\text{Mreal}} = \frac{\lambda}{4n_p} = 104 \text{ nm}$, where $\lambda = 650 \text{ nm}$ is the wavelength of the laser beam in air and $n_p = 1.55$ is the refractive index of the transparent substrate for a DVD. Its value determines the reflected laser beam phase, generating then constructive or destructive beam interferences. The read-out (high frequency) signal is given by the relief of the track that is detected via light intensity measurements.

3.3.2 Light Intensity Distribution

From physical optics it is well known that a light beam passing through an aperture of dimensions which are small relative to the light wavelength λ will give rise to Fraunhofer or far-field diffraction. However, it can be shown mathematically that an aperture situated just before a converging lens will also produce a similar Fraunhofer diffraction in the focal plane of the lens. This situation is also characteristic for CD/DVD players where the converging objective lens is used in the OPUs.

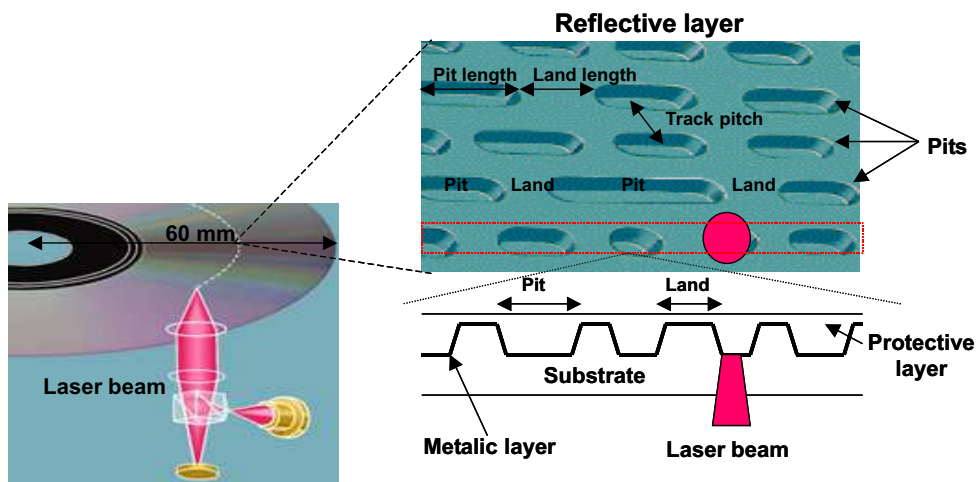


Figure 3.3: Schematic view of the DVD impressed structure.

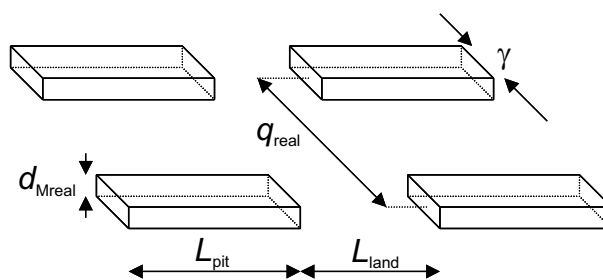


Figure 3.4: Simplified view of the disk impressed structure.

The resulting light intensity of the focused laser spot depends on the incident illumination distribution coming through the optical system from the laser diode. Two interesting cases can be distinguished.

If the incident illumination is *uniform* in distribution, then the normalized light intensity I/I_0 is given by a squared first-order Bessel function divided by a quarter of its own squared argument as described by (3.1), which is well known *Airy disk* formula, see e.g. [BW87] and [NFP00]:

$$\frac{I(r)}{I_0(r=0)} = \left(\frac{2J_1(2\pi \cdot NA \cdot r\lambda^{-1})}{2\pi \cdot NA \cdot r\lambda^{-1}} \right)^2, \quad (3.1)$$

where r is the normalized focused spot radius on DVD (see fig. 3.5), λ is the wavelength of the laser beam, NA is the numerical aperture of the objective lens. The integral representation of the Bessel function is given by:

$$J_n(x) = \frac{j^{-n}}{2\pi} \int_0^{2\pi} e^{jx \cos \alpha} e^{jn\alpha} d\alpha, \quad (3.2)$$

therefore, its first order, reduces to recurrence relation:

$$\int_0^x x' J_0(x') dx' = x J_1(x). \quad (3.3)$$

The Airy disk diameter d_{Airy} is usually taken to be the distance between the first two minima, located on either side of the main peak, and is a quantitative measure of the width of the focused spot. It can be shown that the width of the Airy disk is given by:

$$d_{\text{Airy}} = 1.22 \frac{\lambda f_{\text{obj}}}{R_{\text{obj}}} \approx 1.22 \frac{\lambda}{NA}, \quad (3.4)$$

where R_{obj} is the objective lens radius, f_{obj} is the objective lens focal length.

The numerical aperture is generally defined by:

$$NA = n_r \sin \left(\frac{\phi_{\text{max}}}{2} \right), \quad (3.5)$$

where ϕ_{max} is the opening angle of the objective lens, as shown in fig. 3.2, and n_r is the refractive index of objective space, $n_r = 1$ for the objective lens in air, [Mah98].

Fig. 3.5 illustrates the Airy disk focused spot profile, and the Airy disk diameter d_{Airy} measurement. The numerical value $d_{\text{Airy}} = 1.322 \mu\text{m}$ is given using the following parameters: $\lambda = 650 \text{ nm}$ and $NA = 0.6$.

An alternative expression for the spot size of the Airy disk is that of the Full Width at Half Maximum (FWHM) $d_{\text{Airy(FWHM)}}$, also shown in fig. 3.5. The FWHM of the Airy disk is:

$$d_{\text{Airy(FWHM)}} = 0.6 \frac{\lambda}{NA}. \quad (3.6)$$

On the other hand, if the incident illumination distribution is *Gaussian* and assuming it is untruncated by the aperture of the objective lens then the light intensity on the DVD is given by the Gaussian distribution. The width of the Gaussian focused spot is governed by the width of the incident Gaussian distribution in relation to the aperture

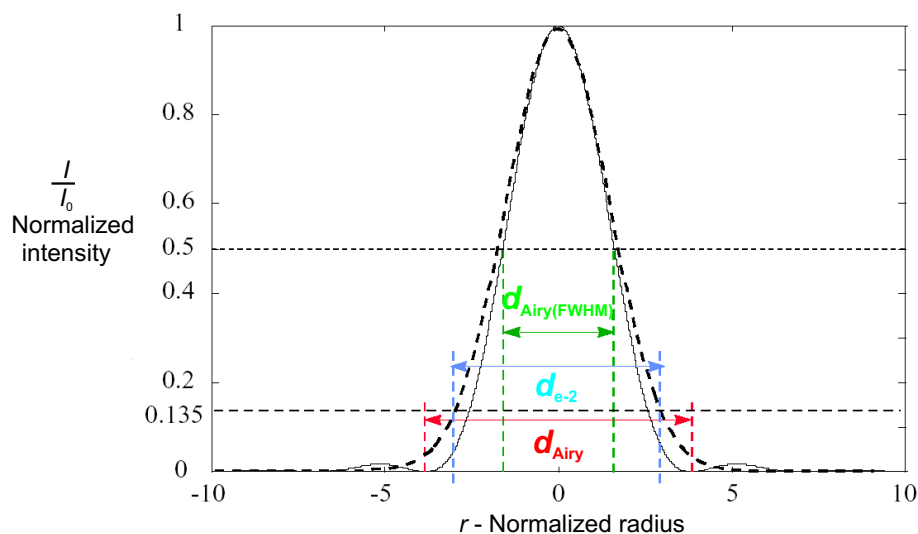


Figure 3.5: The light intensity distribution of the Airy disk (solid line) and Gaussian (bold dashed line) focused spot profiles illustrating the measurement of d_{Airy} and $d_{\text{Airy(FWHM)}}$ of the Airy disk profile, and $d_{e^{-2}}$ of the Gaussian profile.

diameter of the the objective lens d_a . For example, when $d_a = R_{\text{obj}}/2$ (*i.e.* the effectively an untruncated Gaussian illumination), then the e^{-2} diameter of the Gaussian intensity distribution on DVD is given by:

$$d_{e^{-2}} = \frac{4\lambda f_{\text{obj}}}{\pi R_{\text{obj}}} \approx \frac{4\lambda}{\pi \cdot NA} \approx 1.22 \frac{\lambda}{NA}. \quad (3.7)$$

Fig. 3.5 illustrates the Gaussian focused spot profile and the measurement of its diameter $d_{e^{-2}}$.

From the comparison between the Airy disk and Gaussian spot profiles, illustrated in fig. 3.5, one can conclude that the Airy disk spot is narrower than the Gaussian spot, for an untruncated incident Gaussian illumination. Therefore, the highest imaging resolution is observed for a uniform incident objective lens illumination. However, the sidelobes in the Airy disk focused spot can give rise to undesirable imaging phenomena, like the crosstalks of the two adjacent tracks.

For these reasons, a compromise is usually done between these two incident illuminations. Moreover, the optical aberrations of the lens and the optical properties of generated laser beam have to be taken into account during OPU design. In practice, the light intensity distribution generated by the laser diode in the OPUs is close to the Gaussian distribution as it is shown in fig. 3.6.

3.3.3 Focus Error Signal Generation: Modelling Problem Description

Numerous optical properties have been used to generate focus error signal e_F from small vertical spot position error Δx in CD players, as presented in [BBH⁺85] and [Sta98], but for DVD systems, the astigmatic method is the most widely used. The principle of the astigmatic method is based upon an optical aberration, called *astigmatism*. This

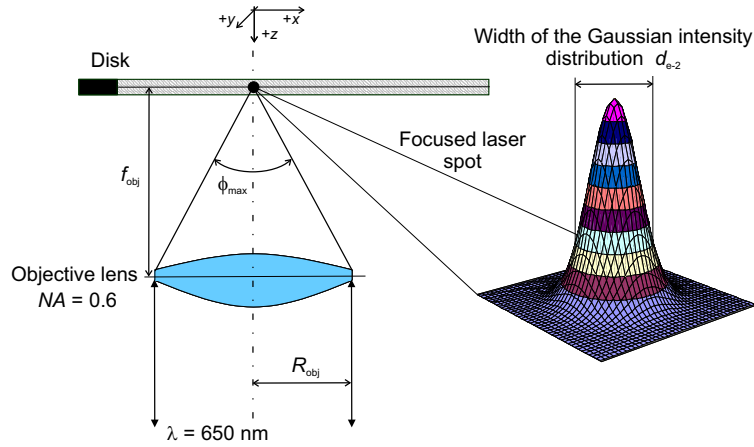


Figure 3.6: The focused laser spot and its light intensity distribution.

distortion is usually introduced by the cylindrical lens (12), see fig. 3.2. Fig. 3.8 represents the simplified model of the reflected beam optical path, presented in fig. 3.2 from the focusing area (15) to the detection area (16).

An astigmatic image is rotated with respect to its optical axis z and a focus error signal e_F can be extracted if a special arrangement of the four photodiodes A, B, C, D is used. The focus error signal is given by:

$$e_F(\Delta z) = (V_A + V_C) - (V_B + V_D), \quad (3.8)$$

where V_A, V_B, V_C, V_D are voltages from the quadrants A, B, C, D of the photodetector and Δz is the vertical spot position error, see fig. 3.2. The focus error signal e_F is feed back to the servo system, to control the actuator fine displacement along the vertical direction z .

The photodetector (13) in fig. 3.2 provides a non-linear bipolar focus error characteristic $e_F(\Delta z)$, usually called *S-curve*, which is used to determine if the laser spot is correctly focused on the disk information layer (10).

In fig. 3.7 an example of the S-curve measured in the time-domain $e_F(t)$ is presented. This characteristic has been obtained from the industrial DVD-video player available in the STMicronics laboratories. As discussed in section 3.3.1, the light reflected from the impressed pit/land grating is falling on a photodetector, by means of optical elements. This spot can have variable size, depending on the de-focus Δz existing between the disk information layer and focused laser spot.

Some preprocessing procedures, described in section C.6, have to be carried out on the measured data to obtain the focus error characteristic (S-curve) in the vertical spot position error domain $e_F(\Delta z)$ from the time domain $e_F(t)$. Fig. 3.9 shows the focus error characteristic $e_F(\Delta z)$ after the preprocessing procedures.

Looking at figs. 3.8, 3.9 and 3.10, the principle used to generate the focus error signal in a DVD player, can be easily resumed as follows:

- In the *optimal* focus condition $\Delta z = 0$, the laser is correctly focused with respect to the disk layer, and all the light reflected by the disk is focused on the photodetector as a circular spot, whose intensity is equally distributed on its four quadrants. In this situation, $(V_A + V_C) - (V_B + V_D) = 0$ and the so-called *focus point* is reached (see point S_5 in fig. 3.9), where the focus control loop can be locked.

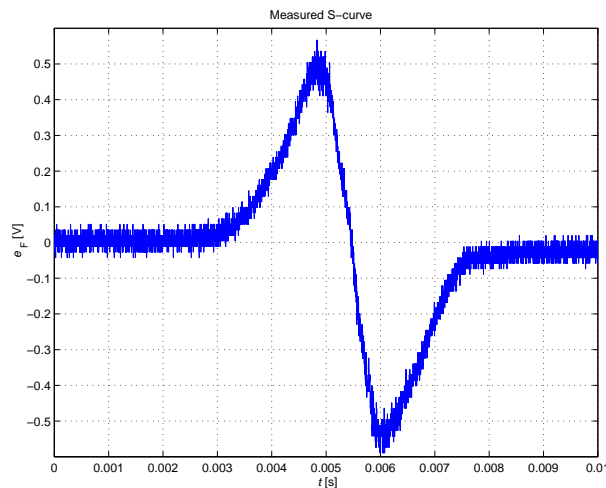


Figure 3.7: An example of measured S-curve, in the time-domain, from an industrial DVD-video player.

- When $\Delta z > 0$, the disk information layer is too far from its ideal position and the reflected light forms on the photodetectors an elliptical shaped spot. The amount of light reflected on the pair A and C is bigger than the one on the pair B and D, so that $(V_A + V_C) - (V_B + V_D) > 0$ and the point S_4 is reached on the focus S-curve.
- When $\Delta z < 0$, the disk information layer is too close from its ideal position and the reflected light forms on the photodetectors an elliptical shaped spot. The amount of light reflected on the pair B and D is bigger than the one on the pair A and C, so that $(V_A + V_C) - (V_B + V_D) < 0$ and the point S_6 is reached on the focus S-curve.
- When the de-focusing Δz becomes bigger (or smaller) than a pre-fixed value, then the laser spot is reflected on the photodetectors as a slanting straight line. In these cases, points S_3 and S_7 of fig. 3.9 delimit the so called *linear zone* of the focus error S-curve. Inside this region the photodetector behavior can be assumed linear, and a gain proportional to the slope of the S-curve, is used to characterize the relation between input and output signals. This is very useful for control design, as will be presented in section 3.4.1.
- When the objective lens is too far from the disk information layer, the spot is said to be “out of focus”, and the generated focus error signal would be zero (the points S_1 and S_2 in fig. 3.9). Similarly, the spot will be “out of focus” in the other direction, when the objective lens is too close to the disk information layer, generating a reflected beam larger than the detector size (the points S_8 and S_9 in fig. 3.9). In both cases the amount of light falling in the photodetector is too weak to retrieve information about focusing. The *optimal* focus condition is recovered by means of software servo algorithms, implemented on a dedicated micro controller.

Parameters characterizing the S-curve are specified in [PEC00] and [SEC01]: The area of the region delimited by its peaks is called *lock-on range*, and the physical distance corresponding to this area represents the voltage range in which the system is able to compensate the maximum actuator displacement to remain still locked. The *acquisition*

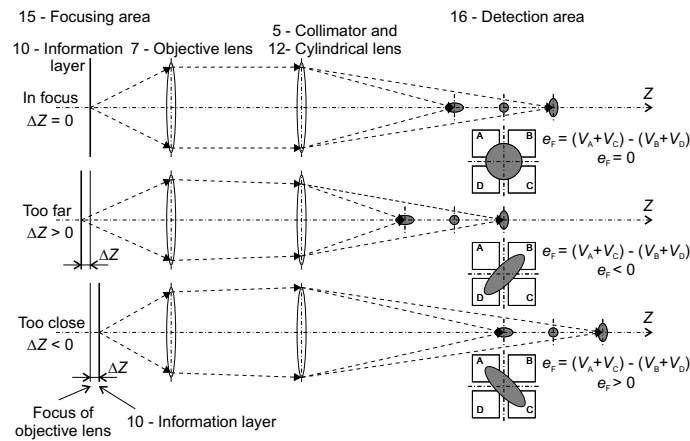


Figure 3.8: Astigmatic method for focus error signal generation.

range is defined as the maximum distance from the focus point S_5 , in which e_F is large enough to still allow the system to perform focusing. Finally, the S-curve symmetry gives an indication on the quality of the error detection procedure. A very asymmetric S-curve would lead to measurement problems, since the system should be electronically adjusted to compensate the fact that the error signal e_F is not equal to zero at the system's best focus.

Problem Description

Despite the S-curve practical importance, no simple analytical or numerical model of the S-curve is available to our knowledge at present. These models could be important for more sophisticated control system design and for testing its functionality (normal playing mode, start-up procedure, calibration, error detection, offset compensation, *etc.*).

More complex models could be devised using diffraction theory to take crosstalk effects into account, see *e.g.* [MCW⁺96], [Mil98], [MU99], [UAM⁺00], [UM01] and [Hop79]. However, the model complexity is an important issue, because the unknown model parameters have to be estimated. Furthermore, the modelling using the diffraction theory is very computational time-consuming.

Therefore, in appendix C, we propose the models of the S-curve, constructed from an opto-geometrical analysis, to achieve compromise between complexity and accuracy.

Radial error signal generation and read-out signal generation are discussed in appendixes D and E, respectively.

3.4 The Mechanical Servo System

The DVD-video player servo mechanics is mainly constituted by three control loops that keep the laser spot in focus on the disk information layer and, during playback, allow the beam to follow the disk spiral. The servo circuitry must also be able to allow high-speed tracks crossing in the radial direction without losing focus, and to find the target location on the disk (long jump or jump- n -track modes).

The three main control loops are in both the CD and DVD players:

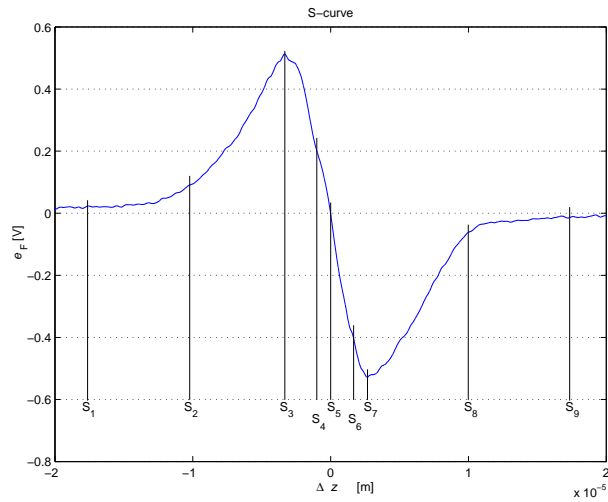


Figure 3.9: Photodetector characteristic (S-curve). The case “ S_1 ” shows the situation where the objective lens is too far from the disk information layer and the case “ S_9 ” shows the situation where the objective lens is too close to the disk information layer.

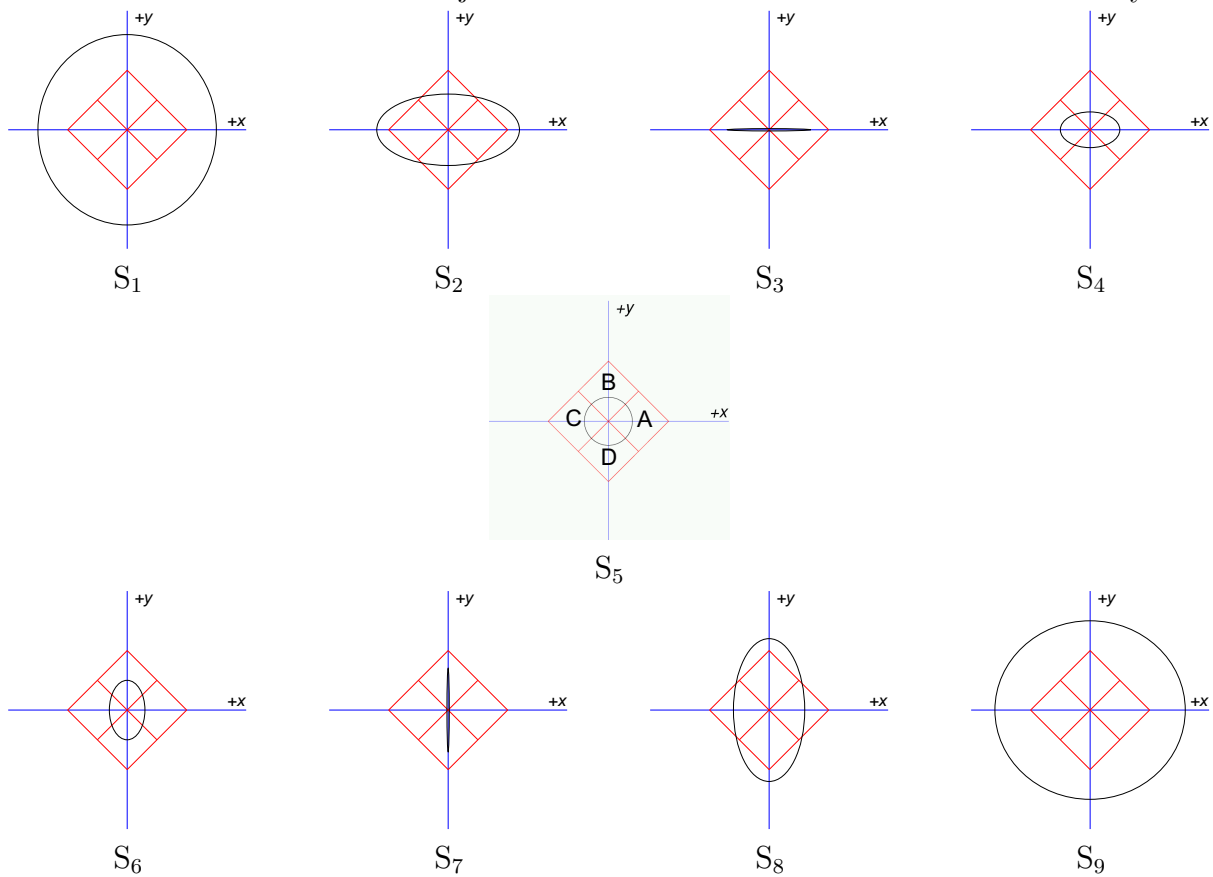


Figure 3.10: The shapes of laser spots in the photodetector plane if diverging cylindrical lens is used.

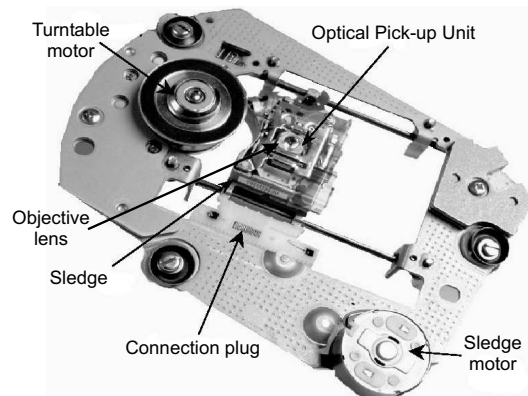


Figure 3.11: An industrial DVD mechanical servo system.

- **A focus loop** to ensure the distance between the objective lens and the disk.
- **A coarse (low frequency) tracking loop** to roughly position the Optical Pick-up Unit (OPU), pulling the so-called **sledge**, in the vicinity of the desired tracks; this is done with a long jump procedure.
- **A fine (high frequency) tracking loop** to lock the focused laser beam onto the track position.

In addition, two others servo loops are used to regulate the speed of the turntable motor (spindle motor control) and to load/unload the disk (tray control), respectively.

In this work we concentrate on the focus and the fine tracking loop (in next called shortly the **tracking** or **radial loop**) used to perform the actuators fine displacements along the vertical and radial directions, since they represent the more complex and critical feedback systems implemented in a DVD drive.

Two actuators are used to perform fine displacement of the laser objective lens along the vertical and the radial direction, in order to keep the laser spot in focus and on track. At the same time the whole system, composed by actuators and optics, is positioned by a sledge at a raw radial location. The sledge, together with the turntable motor, the actuators and the optics, form a rigid body that present the advantage to passively dump unwanted vibrations due to disk rotations.

A picture showing an industrial DVD mechanical servo system is presented in fig. 3.11, where the turntable motor, the Optical Pick-up Unit, the sledge motor, the connection plug, the sledge and the objective lens are shown. A schematic representation of the construction of a DVD drive mechanical servo system is given in fig. 3.12 (case A), where it can be seen that the sledge, the turntable motor and the turntable itself form a rigid body being further consolidated by what is commonly called the baseplate.

On a general DVD drive there are three rotary DC motors: The turntable motor for spinning the disk, the sledge motor for long jump procedure and the tray motor for load/unload the disk. Two other actuators are needed to achieve very fine laser spot positioning on the disk, and they rely on pairs of coils and permanent magnets which can move the objective lens in the vertical or in the radial direction. They are commonly designated as focus and radial actuators and they are presented in fig. 3.12 (case B).

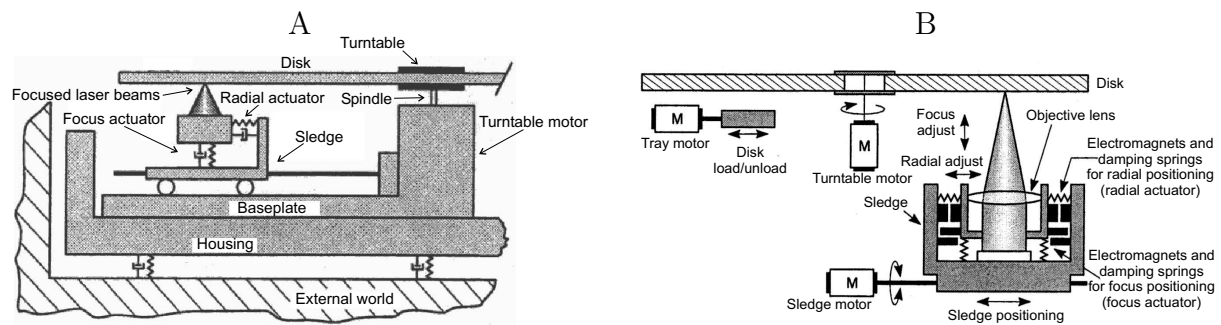


Figure 3.12: Case A: Representation of the DVD drive servo system mechanical construction. Case B: Schematic cross section of the DVD drive actuators.

In optical disk devices, to achieve the highest data capacity on the disk, a constant scan velocity is used for data read-out. This method allows to obtain a constant data density from the inside to the outside of the disk, and it consists in varying the disk rotational frequency $f_{\text{rot}}(t)$ accordingly to the position of the track that is being read.

The disk rotational frequency $f_{\text{rot}}(t)$ is related to the scanning spot velocity v_a and to the desired spot position $x_{\text{ref}}(t)$ along the disk radius, by the following relation:

$$v_a = N2\pi f_{\text{rot}}(t)x_{\text{ref}}(t) = \text{constant}. \quad (3.9)$$

This behavior is known in DVD players as Constant Linear Velocity (CLV), and is achieved using a control loop with the turntable motor. It is interesting to notice that the value of the scan velocity v_a is given in [ECM01], where is said that for the over-speed factor $N = 1$, $v_a = 3.49$ m/s, and $v_a = 3.84$ m/s for a Single Layer (SL) and for a Dual Layer (DL) DVD respectively, as presented in table 3.1. The constant N is a number, usually referred as *over-speed factor* (or X-factor). It expresses the ratio between actual read-out speed and its fundamental value given by specification.

3.4.1 Optical Pick-up Unit Control Loops

The quality of data read-out signal is given by a good positioning of the laser beam on the track, which quantitatively means very small focus and radial errors e_F , e_R .

Focus and radial error

The only measurable signals are the four voltages (V_A , V_B , V_C , V_D) generated by the photodiodes, and these voltages are used to calculate the focus error e_F , radial error e_R and read-out data. The read-out signal is available only during playback mode when the focus error e_F and the radial error e_R are smaller than critical values.

To minimize these errors (e_F , e_R) during playback, despite the presence of disturbances, the presence of a controller in closed-loop is needed. Two main control loops are designed for these purposes: the *focus loop* which maintains the focus point of the laser on the disk layer, and the *radial loop* which allows the laser to follow the tracks.

The focus error e_F is a linear function of the spot position error Δz in vertical direction z only for small values of Δz , as explained in section 3.3.3. For larger values (larger than critical value), the focus error has a nonlinear dependence on Δz . The radial error e_R is

Table 3.1: Physical parameters of CD and DVD. $N = 1$

Symbol	Parameters	CD	DVD
	Layers	single	single/dual
	Substrate thickness	1.2 mm	0.6 mm
	Sides	1	2
	Capacity	0.65 GB	4.7/17 GB
T_{pit}	Channel bit length	266.6 nm	133.3/146.7 nm
q_{real}	Track pitch	$1.6 \mu\text{m}$	$0.74 \mu\text{m}$
	Minimum pit length	$0.83 \mu\text{m}$	$0.4 \mu\text{m}$
v_a	Scan velocity	1.3 m/s	3.49/3.84 m/s
λ	Wavelength	780 nm	635/650 nm
NA	Numerical aperture	0.45	0.6
	Modulation	EFM ¹	8 to 16
	ECC ²	CIRC ³	RSPC ⁴
	Subcode/Tracks	Yes	No

Remark 4 *EFM: Eight to Fourteen Modulation used on every CD for modulation and error correction.*

Remark 5 *ECC: Error Correction Code. CDs use CIRC³, DVDs use RSPC⁴.*

Remark 6 *CIRC: Cross Interleaved Reed-Solomon Code, which adds two dimensional parity information, to correct errors in CDs, and also interleaves the data on the disk to protect from burst errors.*

Remark 7 *RSPC: Reed-Solomon Product Code to correct errors in DVDs.*

also a linear function of the spot position error Δx in radial direction x only for small values of Δx , as presented in section D.

The vertical spot position error $\Delta z = z_{\text{obj}} - z_{\text{ref}}$ is defined as the difference between the objective lens position z_{obj} from the disk information layer and its reference value z_{ref} . In other words, z_{ref} is the desired objective lens position from the disk information layer in vertical direction z , see fig. 3.13 where the solid layers correspond to the desired objective lens position and the dashed layers correspond to the situation where the objective lens is too close the information layer. Using optical theory, one can verify that $z_{\text{ref}} \equiv f'_{\text{obj}}$ and $f'_{\text{obj}} = -f_{\text{obj}}$, where f_{obj} is the objective lens focal length. Since the objective lens focal length f_{obj} is constant, the reference value z_{ref} is also constant.

The radial spot position error $\Delta x = x_{\text{obj}} - x_{\text{ref}}$ is defined as the difference between the objective lens position x_{obj} from the center of the disk and its reference value x_{ref} in radial direction x . In other words x_{ref} is the desired position of the laser spot from the disk center in radial direction x . In general, this reference x_{ref} is not a linear function of time but for the playback operating mode it can be approximated by a linear function of time as follows: $x_{\text{ref}}(t) = v_x t + x_{\text{min}}$, where v_x is the desired velocity of the laser spot in radial direction x and x_{min} is the minimal radius of data zone. A very slowly rising reference value x_{ref} in radial direction x during playback can be ignored because the radial disturbances rejection is more important.

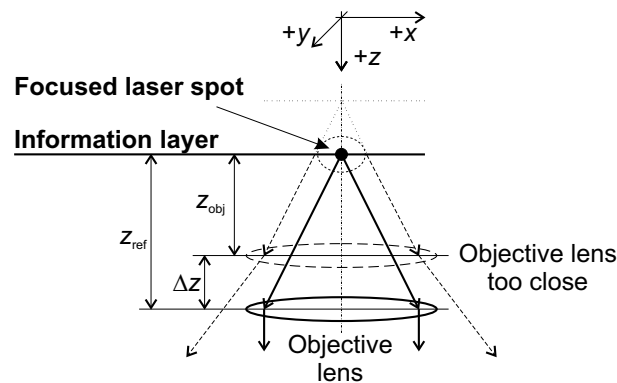


Figure 3.13: Beam focusing upon the disk information layer.

Remark 8 *In the playback mode the laser spot must follow a given track, while the system delivers data to the host interface. What happens in practice is that the accurate objective lens positioning is performed by the fine tracking loop, while the sledge is set to slowly follow the fine actuator movements, by means of the controller in the coarse tracking control loop. In this particular mode, the spot moves towards the outer disk radius, leaving the sledge behind. As the fine actuator displacement is relatively small, the sledge has to advance, but slowly, without following the fast laser spot movements. More detailed information can be found in [Fil03].*

Focus and radial control loops

A complex diagram for the optical pick-up unit control of the DVD player is illustrated in fig. 3.14 as a Two-Input Two-Output (TITO) system. Figures 3.15 and 3.16 show the Single Input Single Output (SISO) focus and tracking control loops, respectively, with the controller on one hand (K_F , K_R) and the plant on the other hand (G_F , G_R).

To move the objective lens (7) (see fig. 3.2) in vertical/radial direction, the focus/radial coils, which are used to generate the electromagnetic field, are supplied by the input currents i_F , i_R (the radial coil is not mentioned in fig. 3.2). These currents i_F , i_R are the *inputs to the mechanical actuators* described by the continuous-time transfer functions $G_{\text{Fact}}(s)$, $G_{\text{Ract}}(s)$. The mean beam (9) of incident rays, generated by the laser diode (1), is reflected by the disk information layer (10) (fig. 3.2). Therefore the light intensity of return beams (11), which is the *input to the sensor*, contains the information about the actual vertical/radial spot position error Δz , Δx , respectively.

The equivalent path between the input currents i_F , i_R and output voltages from the photodetector (13) V_A , V_B , V_C , V_D in fig. 3.2 is realized by two blocks (named *actuators* and *sensor* in fig. 3.14).

In figs. 3.15 and 3.16 the plants outputs y_F , y_R are obtained from the measured voltages V_A , V_B , V_C , V_D . Firstly, the voltages V_A , V_B , V_C , V_D are converted by the Analog to Digital (A/D) convertors. Secondly, some calculations are done to generate the focus/radial error signals e_F , e_R from the measured voltages V_A , V_B , V_C , V_D . The optical gains (A_{FOpt} , A_{ROpt}) are here defined by the slope of the non-linear functions for small spot position errors (Δz , Δx).

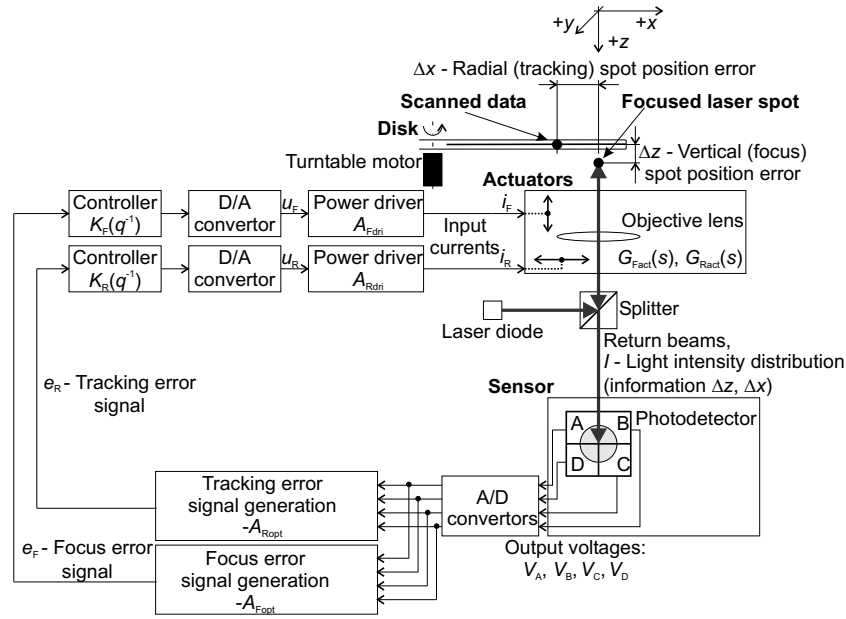


Figure 3.14: Control of the DVD player, the focus and tracking loop, and main physical components.

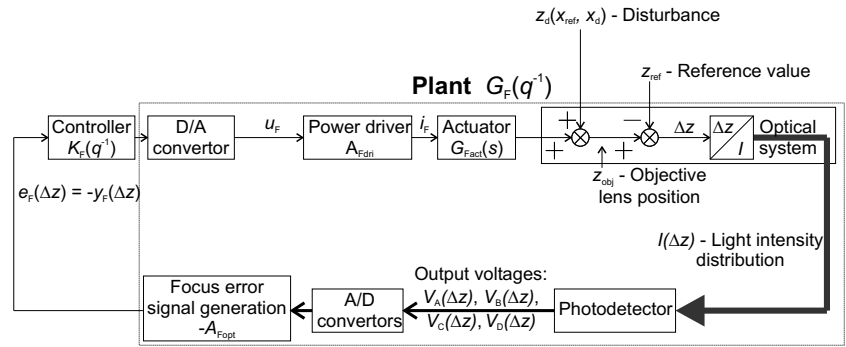


Figure 3.15: Block diagram of the focus control loop of fig. 3.14, and main signals.

The plant outputs y_F, y_R corresponds to the error signals with inverse signs $e_F = -y_F, e_R = -y_R$ in this case because the reference values (z_{ref}, x_{ref}) are inside of the optical system (see figs. 3.15 and 3.16).

$K_F(q^{-1}), K_R(q^{-1})$ denote the digital controllers transfer functions whose digital outputs are converted to analog signals by the Digital to Analog (D/A) converters.

Finally, the power drivers, which gains are A_{Fdri}, A_{Rdri} , amplify the controllers outputs voltages u_F, u_R to supply the focus/radial coils with sufficient currents i_F, i_R .

There exists different kind of disturbance sources which affect the normal behavior of a DVD player. They are described in the following section.

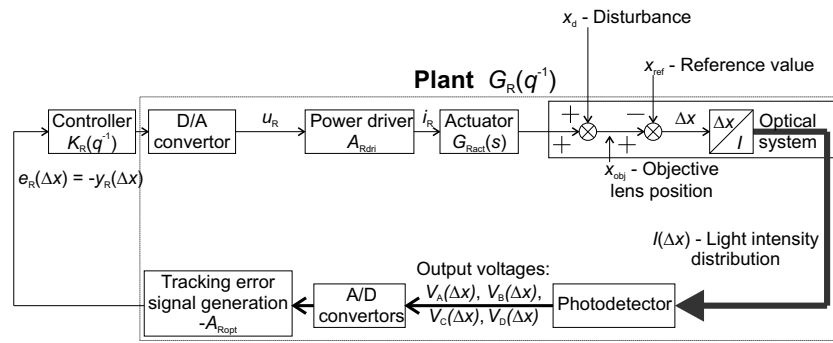


Figure 3.16: Block diagram of the tracking control loop of fig. 3.14, and main signals.

3.5 Disturbances Sources

The disturbances affecting the DVD systems have different sources and they can be divided according to whether they come from the optical imperfections or from the DVD player itself (internal disturbances) or from the environment (external disturbances) or from the disk surface defects. These different kinds of disturbance sources are studied in following sections.

3.5.1 Optical Imperfections

Optical imperfections can be summarized as follows:

1. Laser noise, illustrated in fig. 3.17a, gives a background high frequency noise which may alias into the servo frequencies if an anti-aliasing filter is not used.
2. Optical misalignment and optical skew, shown in fig. 3.17b&c, cause asymmetry and cross coupling between the focus and radial error signal. This means that focus/tracking control loops create a true MIMO system with two inputs i_F, i_R which influence two outputs e_F, e_R .
3. The optical cross coupling between the focus and tracking loops have been minimized for the DVD players. The CD mechanism, shown in fig. 3.18 (case A), usually consists of a radial arm in order to follow the spiral track of the disk (a coarse tracking loop). This induces the optical cross coupling between the focus and tracking loops, caused by a changing of the skew angle, see fig. 3.17c. Instead of the radial arm, all DVD drives use a linear actuator to roughly position the optical head assembly in the vicinity of the desired tracks, see fig. 3.18 (case B). This improvement prevents the changing of the skew angle. Therefore the optical cross coupling between the focus and tracking loops has been suppressed and control design is always done for two decoupled SISO systems.
4. The photodetector is subject dead-zones between photodiodes segments as it is illustrated in fig. 3.17d.
5. Lens and groove imperfections show up as spurious signals on the detectors, see fig. 3.17e.

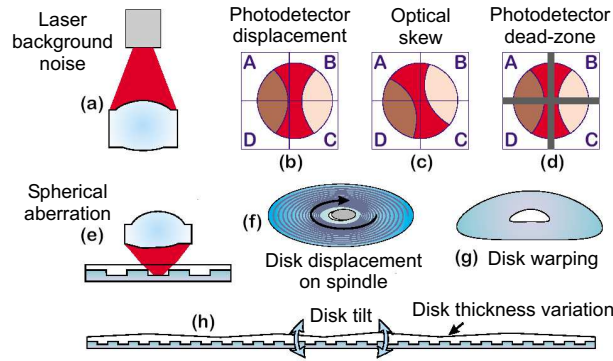


Figure 3.17: Sources of optical disk noises.

The numerical aperture NA , defined by (3.5), is for CD players $NA = 0.45$ and for DVD players $NA = 0.6$. Higher numerical aperture NA for DVD players causes the following phenomena:

1. Decreasing of the focused spot size which corresponds to the increasing the disk capacity. The focus spot size corresponds to the Airy disk diameter, given by $d_{\text{Airy}} \approx 1.22 \frac{\lambda}{NA}$. $\lambda = 780 \text{ nm}$ is used in CD players while $\lambda = 650 \text{ nm}$ is used in DVD systems. Generally, DVDs have higher data recording density than CDs. Capacity of the CD is usually 0.65 GB while capacity of the one-layer one-side DVD is 4.7 GB. Notice that the one-layer one-side DVD of new blue laser DVD format ($\lambda = 405 \text{ nm}$, $NA = 0.65$) has capacity 15 GB.
2. Decreasing of the focus depth, defined by:

$$\Delta z_{\text{max}} = \frac{\lambda}{2NA^2}. \quad (3.10)$$

The focus depth is the maximal spot position error in vertical direction and the controller in the focus control loop should control the objective lens position within $\pm \Delta z_{\text{max}}$ to avoid losing the data read-out signal during playing.

3. Increasing of the coma non-linearly, [BW87]. The coma is the primal aberration of the objective lens. It causes the non-uniform intensity distribution of the focused laser spot which leads to worse returned laser beam quality. Non-uniform intensity distribution induces higher sensitivity of returned laser beam quality to the disk tilt and disk thickness variations.
4. Influences the disk removability because the distance between the disk and objective lens is smaller. DVDs have thinner protective layers than CDs to improve the returned beam quality and preserve the compatibility of CD/DVD reading systems, therefore they are more susceptible to destruction during handling.

These optical imperfections can finally generate the offsets which are added to the output voltages V_A , V_B , V_C , V_D .

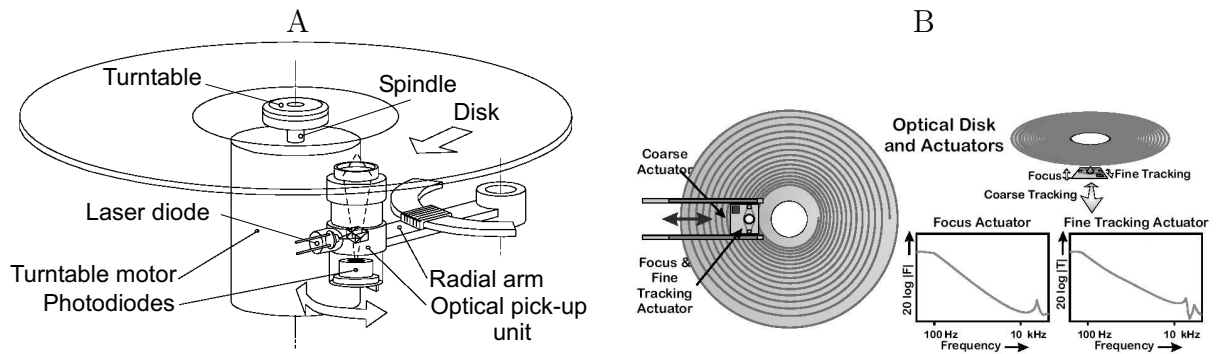


Figure 3.18: Case A: Optical disk and actuators of the CD player. Case B: Optical disk and actuators of the DVD player.

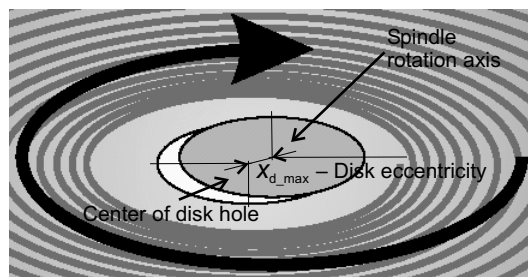


Figure 3.19: Disk eccentricity as the maximal value of the disk displacement on the spindle.

3.5.2 Internal Disturbances

Internal disturbances are mainly due to the spindle rotation frequency f_{rot} and to the reaction force that the actuators develop on the drive baseplate and housing, during playback. Usually, also internal disturbances are synchronous with the spindle rotational frequency f_{rot} , introducing thus harmonics at this frequency. Disk warping and disk eccentricity, known as disk displacement on spindle, belong to this class of disturbances.

Disk warping, shown in fig. 3.17g, is a vertical deviation z_d that puts a large repetitive error into the focus loop. In case of the tracking control loop, repetitive disturbance is caused by a disk misalignment, known as the disk displacement on the spindle x_d .

Disk eccentricity x_{d_max} is the maximal distance between the disk hole center and the spindle rotational axis (*i.e.* the maximal value of disk displacement on the spindle x_d). A detailed zoom on the disk eccentricity x_{d_max} is illustrated in fig. 3.19.

Disk warping $z_d(x_{ref}, x_d)$ and disk displacement on the spindle x_d , are not only synchronous with the spindle frequency f_{rot} but have also higher harmonics of f_{rot} . The power spectrum of the radial error signal e_R , at the given rotational frequency $f_{rot} = 15$ Hz shown in fig. 4.25, has been acquired on the real DVD system in STMicroelectronics laboratories. One can see that the main disturbances are given by the first and the third harmonic components of disk rotational frequency f_{rot} .

The repetitive disturbance modeling will be shown in section 4.3.

The optical head remains relatively large to hold the focus and tracking motors in the optical pick-up unit of DVD/CD players, compared to a hard disk reading head.

Spindle speeds are also slower than in the hard disks. Thus an air turbulence (windage) is negligible in DVD/CD systems, in contrast to the hard disks where it induces vibrations of the magnetic-head. Critical speeds and aerodynamic flutter instability of various optical disks and hard disks are studied in [LKK02]. This analysis shows that it is possible to increase the rotational speed f_{rot} of current DVD/CD drivers up to 233 Hz without self-excited vibration due to aerodynamic flutter.

All DVD and last CD drivers rely on mechanical concepts similar to those depicted in fig. 3.12 (case A). The focus and radial actuators are dedicated to keep the laser spot in focus and on track. They can perform fine displacements along the focus and respectively radial direction relative to the disk while being positioned by a sledge at a raw radial location. The sledge forms a rigid body together with the turntable motor and turntable itself, being consolidated on what is called the baseplate. In this structure, the baseplate and the housing itself are considered as one body. Therefore the internal disturbances that come from the internal vibrations of the mechanical components, are largely stimulated by the spindle rotation of the disk and the actuators reaction forces on the drive baseplate and housing.

Repetitive disturbances are typically dealt with using some form of harmonic cancellation, typically either repetitive control [MLC98] or disturbance observer design [YCC⁺02]. The spectral disturbances which are not repetitive must be handled by the servo loop or removed by some modification of the mechanics. Mechanical methods for compensating such problems involve changing the disk substrate, either with a stiffer material such as glass, or with an internal layer of viscoelastic damping material. The former method has the advantage of having been already tested on small form factor portable drives and the most recent high performance drives.

3.5.3 External Disturbances

The external disturbances affecting drives are typically environmental shocks and vibrations, whether from a moving vehicle, a factory floor environment, a computer under a desk being kicked, or simply the motion of a laptop computer. For streaming media such as DVD/CD, they are usually overcome by data buffering in the portable players or cars. Therefore the practical use of accelerometers to disturbance cancellation in DVD/CD players has been limited. On the other hand the accelerometers are widely used to disturbance cancellation in random access applications, as hard disks.

3.5.4 Disk Surface Defects

The scratches, dust, fingerprints and impressed pit and land imperfections can give spurious signals on the photodetector. Disk thickness variations and disk tilt, shown in fig. 3.17h, also belong to this class of disturbances.

Generally, they have the high-frequency contents and the relevant influence on the performance of controllers that are designed to achieve high bandwidth. An affective way to deal with this problem is increasing the rotational frequency of the disk. As a consequence, the “frequency contents of the surface irregularities” will be shifted to higher frequencies, where the controller has low gain.

Table 3.2: An overview of the disturbance frequencies, DVD player, radial loop, $N = 1$.

Frequency range (Hz)	Cause of the disturbance
output voltages offsets (9.6 – 23.1) and harmonics 30 – 80 $\approx (> 100)$ 5000 – 11000 $\approx (> 5000)$ random	photodetector displacement, optical skew, photodetector dead-zone, spherical aberration disk displacement on the spindle mild mechanical resonance of actuator (construction) scratches, dust, fingerprints, disk thickness variation slight mechanical resonances of actuator (construction) laser background noise, pit structure shocks, external vibration
2.4 kHz 9.6 – 23.1	closed-loop bandwidth spindle rotational frequency

3.5.5 Disturbances Summary

In fig. 3.20 and fig. 3.21 an overview of the disturbance and noise sources acting in both the control loops is shown. In these pictures, it is possible to distinguish additional sources of disturbance entering in the loops, as the noise introduced by the A/D and the D/A converters, the sensing noise, cross coupling phenomena and non-linearities due to the error signals generation methods, as presented in sections 3.3.3 and D.

More detail information on the presented disturbances sources of both the DVD/CD and hard disks systems are in [Abr01] and the references therein.

In this work we consider low-frequency disturbances mainly due to non perfect location of the hole at the center of the disk or non-perfectly orthogonal disk clamping. These imperfections may produce eccentricity in radial direction and vertical deviations. Shocks and vibrations are not taken into account since they are random events that don't often affect the behavior of a home DVD player. Since disk surface defects present high-frequency contents, we don't take them into account in the design of controllers, which are basically conceived to achieve limited bandwidth.

The disturbance frequencies on DVD players are difficult to specify exactly but we can give some ranges. An overview of them for the over-speed factor $N = 1$ is briefly presented in table 3.2.

From the control point of view, the influence of the different components such as power drivers, A/D and D/A converters, sensors and error signal generation blocks, can be included in a high-level blocks, as shown in the schematic diagram of fig. 4.3 in section 4.4.1. Here, we consider the controller, the plant, to simplify the control loop scheme, and define the standard control problem.

The control problem consists in projecting the laser spot with high accuracy onto the track, in the vertical and in the radial direction, in a way that the error between the track and the spot positions should not exceed the its critical values. These values are defined from the desired control loop performances which are presented in the following section.

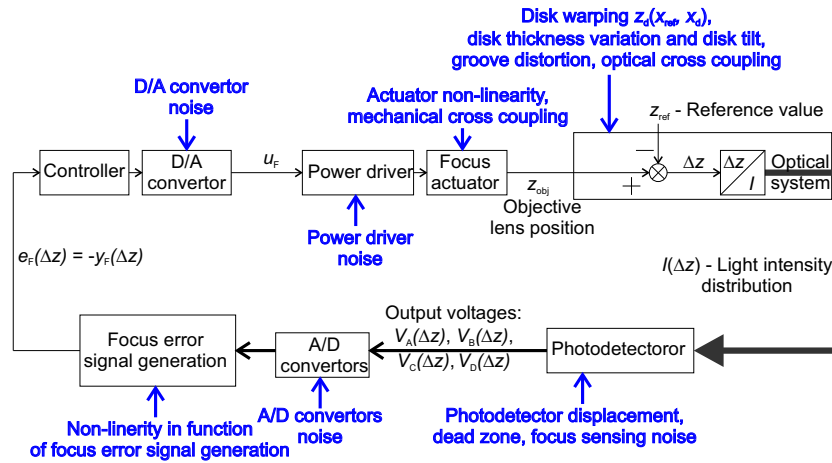


Figure 3.20: Disturbance sources acting in the focus control loop.

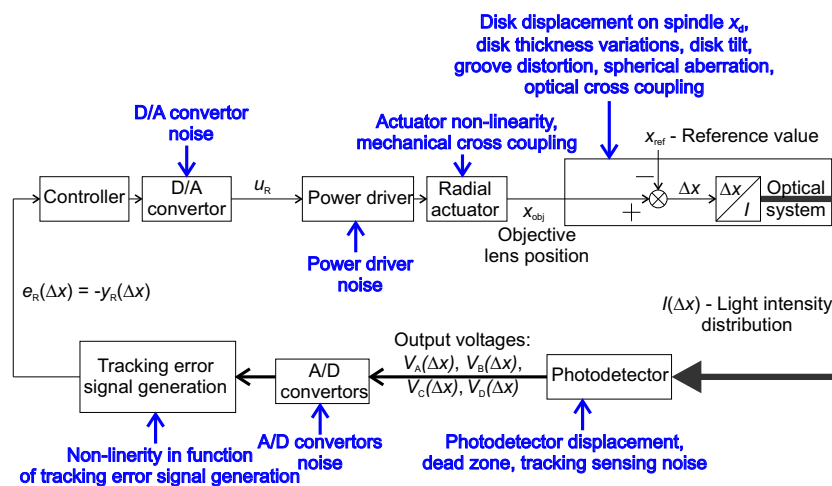


Figure 3.21: Disturbance sources acting in the radial control loop.

3.6 Control Loops Specifications

3.6.1 Focus Control Loop

The laser spot position must follow the disk track (in the focus/radial directions) despite the presence of disturbances. There is a standard specification [ECM01] to enable the DVD disks interchanging between the disk drivers.

This specification is given when an over-speed factor $N = 1$ is considered, and it prescribes values for the maximal deviations of the disk from nominal position for different disk rotational frequencies f_{rot} . Since the complete specification can be found in [ECM01], only important definitions are presented in the following.

The actual (spindle) rotational frequency of the disk f_{rot} , given by (3.11), is not constant for DVD players where a Constant Linear Velocity (CLV) method is used to read the data on the disk.

$$f_{\text{rot}}(t) = \frac{Nv_a}{2\pi x_{\text{ref}}(t)}, \quad (3.11)$$

where N is the over-speed factor, $v_a = 3.49$ m/s is the constant scanning velocity and x_{ref} is the distance between the disk hole center and the falling laser beam (*i.e.* the desired reference position), $x_{\text{ref}}(t) \in \langle x_{\text{min}}, x_{\text{max}} \rangle$, (see figs. 4.1 and 4.2). Here, $x_{\text{min}} = 24$ mm is the minimal radius of the disk data zone and $x_{\text{max}} = 58$ mm is the maximal radius of the disk data zone with respect to the disk hole center.

Since the rotational frequency f_{rot} is not constant during playing, one can determine for the over-speed factor $N = 1$ the maximal rotation frequency of the disk $f_{\text{rot_max}}$ at the beginning of data zone $x_{\text{ref}}(t) = x_{\text{min}}$ and the minimal rotational frequency of the disk $f_{\text{rot_min}}$ at the end of data zone $x_{\text{ref}}(t) = x_{\text{max}}$.

An additional information, concerning the closed-loop performances, can be given from the rise time t_r of the closed-loop step response, defined as the time it takes for the output y_F to first reach 90% of its final value. The rise time t_r usually verifies the following equation:

$$t_r \simeq \frac{2.3}{2\pi f_{\text{cr}}} = \frac{2.3}{2\pi \cdot 2 \times 10^3} = 0.183 \text{ ms}, \quad (3.12)$$

where f_{cr} is the cross-over frequency of the compensated open-loop. The rise time t_r of real closed-loop system has to be smaller than this critical value.

As presented in section 3.5.1, the important parameter for focusing is the focus depth Δz_{max} that establishes the maximal vertical spot position error. Δz_{max} is defined by (3.10) and its numerical value $\Delta z_{\text{max}} \doteq 0.903 \mu\text{m}$ is given using the laser wavelength $\lambda = 650$ nm and the objective lens numerical aperture $NA = 0.6$. The focus servo should control the objective lens position within $\pm \Delta z_{\text{max}}$ to avoid losing the data read-out signal during playing.

The focus servo loop performance, to define the disk parameters for both the manufacturing process and the controller design, is briefly presented in table 3.3.

Finally, fig. 3.22 (case A) shows the upper and lower limits of the disturbance rejection requirements on the output disturbances in frequency domain, for the focus control loop, $N = 1$. Here, the curve L_1 illustrates the requirements for manufacturing of the disk. The curve L_2 is an upper limit and curve L_3 is a lower limit on the output sensitivity function S_{yp} (defined in section 4.4.1) for controller design.

Table 3.3: Focus servo specification for the DVD, $N = 1$, compensated open-loop.

	Parameter	Range	Value
v_a	Scanning velocity		3.49 m/s
$f_{\text{rot_min}}$	Min. spindle frequency		9.6 Hz
$f_{\text{rot_max}}$	Max. spindle frequency		23.1 Hz
f_{cr}	Cross-over frequency of compensated open-loop		2 kHz
x_{min}	Min. radius of data zone	$f_{\text{rot}} = f_{\text{rot_max}}$	24 mm
x_{max}	Max. radius of data zone	$f_{\text{rot}} = f_{\text{rot_min}}$	58 mm
$z_{\text{d_low}}$	Max. deviation from nominal position for low f_{rot}	$f_{\text{rot}} \leq f_{\text{rot_max}}$	± 0.3 mm
$z_{\text{d_high}}$	Max. deviation from nominal position for high f_{rot}	$f_{\text{rot}} > f_{\text{cr}}$	± 0.23 μm
a_{max}	Max. vertical acceleration of end point on data zone	$f_{\text{rot}} > f_{\text{rot_max}}$	8 m/s^2
f^{BW}	Desired closed-loop band-width		$\approx f_{\text{cr}}$
Δz_{max}	Max. vertical spot position error		$0.903 \mu\text{m}$
t_r	Rise time of closed-loop		0.183 ms

3.6.2 Radial Control Loop

The radial servo performance specification is given in a similar way in [ECM01]. Therefore, its overview is presented briefly in the following.

The track pitch, $q_{\text{real}} = 0.74 \mu\text{m}$, is an important parameter for tracking that establishes the distance between the centerlines of a pair of adjacent physical tracks, measured in radial direction. It is difficult to define Δx_{max} as the maximal radial spot position error to avoid losing the data read-out signal during playing. As a rule of thumb normally a deviation 10% of the track pitch is appropriate. This leads to the following formula:

$$\Delta x_{\text{max}} = 0.1 q_{\text{real}} = 0.1 \cdot 0.74 \times 10^{-6} = 0.074 \mu\text{m}. \quad (3.13)$$

In table 3.4, the radial servo specification is briefly shown.

Finally, the upper and lower limits on the output sensitivity function S_{yp} (defined in section 4.4.1) in frequency domain for the radial control loop are given in fig. 3.22 (case B).

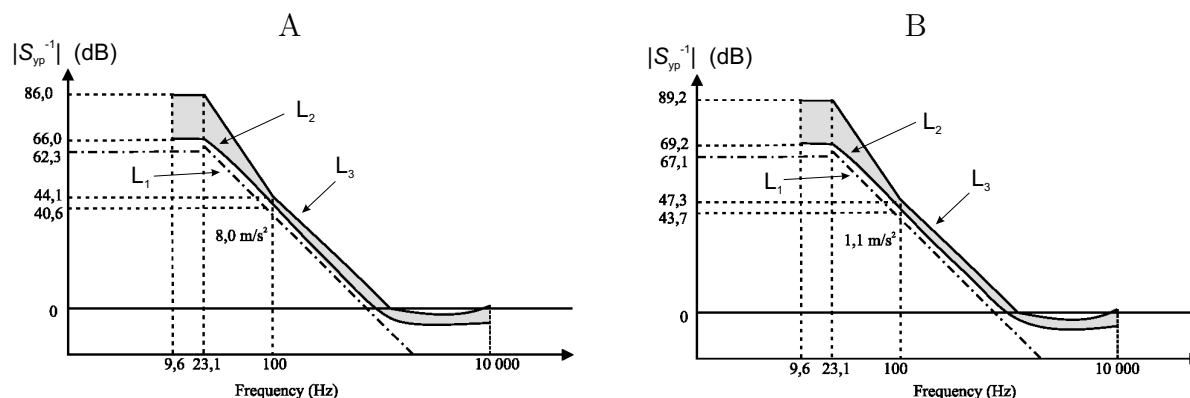
3.7 Performance Limitations

In the actual industrial solution the spot position control loops are implemented by using a dedicated DSP module.

Although this device is very cheap and simple to program, quantization effects due to the A/D converters, finite precision of digital computation and rounding errors pose strong limitations on the complexity of the computational structure and on the choice of coefficients used to implement the digital filters. When implementing digital signal processing systems one must represent signals and coefficients in some digital number

Table 3.4: Radial servo specification for the DVD, $N = 1$, compensated open-loop.

	Parameter	Range	Value
f_{cr}	Cross-over frequency of compensated open-loop		2.4 kHz
x_{d_low}	Max. deviation from nominal position for low f_{rot}	$f_{rot} \leq f_{rot_max}$	$\pm 50 \mu\text{m}$
x_{d_high}	Max. deviation from nominal position for high f_{rot}	$f_{rot} > f_{cr}$	$\pm 0.022 \mu\text{m}$
a_{max}	Max. radial acceleration of end point on data zone	$f_{rot} > f_{rot_max}$	1.1 m/s^2
f^{BW}	Desired closed-loop band-width		$\approx f_{cr}$
Δx_{max}	Max. radial spot position error		$0.074 \mu\text{m}$
t_r	Rise time of closed-loop		0.153 ms

Figure 3.22: Specification of the control loops, $N = 1$. Case A: Focus loop. Case B: Radial loop.

system that must always be of finite precision, since the output samples from the A/D converter are quantized and represented by binary numbers.

As stated in [OS97], the operation of quantizing a number with a finite sequence of bits can be implemented by rounding or by truncation, but in both cases quantization is a non linear operation, which affects the implementation of linear time-invariant discrete-time systems.

Hence, because of the finite precision of digital computation, the digital filters implementation structures must be carefully chosen and the controllers complexity should be limited to reduce the effects of rounding errors and quantization noise, as suggested in [WY02].

In the actual control solution adopted for a DVD-video player, the precision is of 8-bit for coefficients and of 16-bits for data. Controller order reduction and choice of the filter structures, are related to implementation constraints, such as the fixed number of bits used for represent coefficients and data, and to the chosen computational frequency f_s . All these factors have to be taken into account, during the controller implementation phase, when the DSP is programmed by using a dedicated assembler language.

Table 3.5: Requirements on the typical CD/DVD players, radial loop.

	Parameter	CD (1x)	DVD (1x)	CD (52x)	DVD (12x)
f_{rot} (Hz)	Rotational frequency	4 – 8	9.6 – 23.1	208 – 416	115 – 277
f^{BW} (kHz)	CL bandwidth	0.5	2.4	26	28.8
Δx_{max} (μm)	Max. radial spot position error	0.16	0.074	0.16	0.074

3.7.1 Control Difficulties of the CD and DVD Players

Many papers are devoted to the control design of CD players, mostly for the tracking loop. Linear Matrix Inequality (LMI) techniques [Det01], μ -synthesis [SSB96], repetitive control [Ste00], deterministic method for obtaining nominal and uncertainty models [VSA⁺02], applied to a CD mechanism, are the topics treated in the literature up to now.

The important difference between CD and DVD players, in sense of mechanical construction, lies in the course tracking loop. The CD mechanism, usually consists of a radial arm in order to follow the spiral track of the disk, see 3.18 (case A). This induces a non-linear gain variation in the fine tracking loop, [Det01], and also introduces the optical cross coupling between the focus and fine tracking loops. Instead of the radial arm, all DVD drives use a linear actuator to roughly position the optical head assembly in the vicinity of the desired tracks, see 3.18 (case B). This improvement leads to the constant radial actuator gain in the fine tracking loop and it suppresses the optical cross coupling between the focus and fine tracking loop.

Smaller wavelength of laser beam and higher objective numerical aperture of the objective lens, used in DVD drivers than in CD players, increase the DVD capacity, the desired Closed-Loop (CL) bandwidth and cause stronger requirements on the spot position accurate for higher disk rotational frequencies in order to read data correctly. The requirements on CD/DVD players for the typical over-speed factor $N = 52$ (52 times the transfer rate of the original CD) and $N = 12$ (DVD), together with $N = 1$ (CD/DVD) are briefly presented in table 3.5.

An objective comparison of control difficulties of the CD and DVD players is hard. Many technological and software improvements in DVD players construction (the linear radial motor, smaller number of the photodiodes in Optical Pick-up Unit (OPU), different algorithms for radial error signal generation, smaller frequency rotational range for the typical over-speed factor N) simplify some parts of the control design. On the other hand, smaller maximal radial and vertical spot position error tolerances for DVD players to avoid losing the data read-out signal, increase performance requirements.

3.7.2 Control Problem Description

Concerning DVD players, to our knowledge, only recently few papers have been dedicated to a narrow-band disturbance suppression in a high-speed DVD players system. In [BDCS01], a control architecture for track following using the notch filtering and multirate control is proposed. Paper [ZKS02] presents a control system using sliding mode

control to handle shock and vibration disturbances while paper [YCC⁺02] is devoted to robust tracking control using observer of the tracking error signal. In [KYT⁺03], a new tracking control method consisting of a feedback controller and a feed-forward controller that employs the zero phase error tracking method is proposed. All these methods result in controllers with a complex structure, and to implement such controllers, special device (high performance Digital Signal Processing/Processor (DSP)) is necessary. This is still needed to search for simpler controller structure for industrial application.

In the present study, the aim is to find a fixed low-order controller being able to improve the eccentricity suppression in the focus/radial control loops for DVD players. It is shown that this can be achieved with pole placement design, followed by controller order reduction. Pole placement method is here adapted to realize repetitive disturbance rejection in a certain bandwidth. The control system design will be presented in chapter 4.

3.8 Conclusions

In this chapter we have described the DVD player and defined the modelling the photodetector characteristic problem and the control system design problem.

In the first part of this chapter, we have presented the DVD drive architecture, and a detailed description of optics, in order to clarify principles used to generate the servo and the read-out signals. Model of the radial error signal generation has been presented in appendix D while model of the read-out signal generation has been presented in appendix E.

The problem definition of modelling the photodetector characteristic has been introduced in section 3.3.3. In the following chapter C the focus error characteristic modelling will be proposed.

The second part of this chapter is devoted to the description of the electro-mechanical servo system and to the definition of the disturbances sources and spot position control problem together with system's performance specifications.

The main control objective is to impose a hard bound on the time-domain amplitude of the spot position errors $(\Delta z, \Delta x)$, along the radial and the vertical directions (z, x) , in the presence of periodic disturbances whose period varies with the rotational frequency of the disk f_{rot} .

The proposed control design will be presented in chapter 4.

4.1 Introduction

As presented in section 3.7.2, this chapter is devoted to the design of a control system applied in an industrial DVD-video drive.

A model-based control design procedure is proposed in this chapter for the construction of a low complexity controller that achieves an enhanced track following performance for the DVD-video player. As stated in section 3.7, low complexity controllers are motivated by the need of keeping the design cost low, and to make digital controllers implementation feasible in the actual industrial solution.

A combined pole placement/sensitivity function shaping methodology is used for the control design purposes to reduce the effect of repetitive disturbances. Controller order reduction is performed to allow its practical implementation. Experimental results, obtained on a real system in STMicroelectronics laboratories, illustrate the performance of the proposed algorithms for both the focus and radial control loops.

Our contribution consists of providing a general and simple methodology for control design of a DVD-video servo system, that is more suitable for unknown parameter variations and real-time implementation. This methodology leads to the development of controllers which are potentially more accurate, powerful and more simple to implement.

This chapter is organized as follows: In section 4.2, a brief state of the art is presented. Section 4.3 describes the modelling the repetitive disturbances. We show that these disturbance are due to the geometry of the disk. The control methodology is described in section 4.4. Section 4.5 shows the control system design. The simulation and experimental results are given in section 4.6. Finally, in section 4.7, we draw some conclusions.

4.2 State of the Art

In the literature, there exist several studies concerning application of different control synthesis applied to Compact Disk mechanism. In the following, we briefly summarize, by author, the subjects treated in the previous researches.

Many papers are devoted to the control design of CD players, mostly for the tracking loop. Linear Matrix Inequality (LMI) techniques [Det01], μ -synthesis [SSB96], repetitive control [Ste00], deterministic method for obtaining nominal and uncertainty models [VSA⁺02], applied to a CD mechanism, are the topics treated in the literature up to now.

Moreover, in the presence of eccentricity, the controller should be able to deal with this repetitive disturbance, which has a variable frequency (depending on the beam actual position on the track). The resulting controller structure is often quite complex (robust, adaptive).

Concerning DVD players, to our knowledge, only recently few papers have been dedicated to a narrow-band disturbance suppression in a high-speed DVD players system. In [BDCS01], a control architecture for track following using the notch filtering and multirate control is proposed. Paper [ZKS02] presents a control system using sliding mode control to handle shock and vibration disturbances while paper [YCC⁺02] is devoted to robust tracking control using observer of the tracking error signal. In [KYT⁺03], a new tracking control method consisting of a feedback controller and a feed-forward controller that employs the zero phase error tracking method is proposed. All these methods result in controllers with a complex structure, and to implement such controllers, special device (high performance Digital Signal Processing/Processor (DSP)) is necessary. This is still needed to search for simpler controller structure for industrial application.

In [Fil03], an H_∞ norm-based control design procedure has been proposed for construction of a low-complexity controller, able to achieve an enhanced track following performance and periodic disturbance rejection. The same DVD-player benchmark, but different control system design methodologies, have been used in [Fil03] and our work because they have been done in parallel at STMicroelectronics laboratories.

In the present study, the aim is to find a fixed low-order controller being able to improve the eccentricity rejection in the focus/tracking control loops for DVD players. It is shown that this can be achieved with pole placement design, followed by controller order reduction. Pole placement method is here adapted to realize repetitive disturbance rejection in a certain bandwidth.

The control design method is validated on a real system from STMicroelectronics.

In this chapter, we will use combined pole placement/sensitivity shaping method methodology exposed in [LLM97], [LL99], to perform a single-objective controller design and study the system steady-state behavior, its disturbance rejection properties, and simply robustness analysis against structural parametric uncertainty.

4.3 Repetitive Disturbances Modelling

A laser beam is used to read the recorded digital data from the optical disk. In order to retrieve the data correctly, the laser beam must be focused on the data layer disk surface, and must follow the track, both with high precision, as discussed in 3.6. This is difficult in the presence of disturbances and control loops are necessary.

As presented in section 3.4.1, the goal of the focus controller design is to minimize the amplitude of the vertical spot position error Δz , measured by the focus error signal e_F , during playback. On the other hand, as shown in section 3.4.1, the goal of the radial controller design is to minimize the amplitude of the radial spot position error Δx , measured by the radial error signal e_R , during playback.

One of the most important disturbances are the repetitive disturbances which have been introduced in section 3.5.2. In this context, the following sections dealt with the modelling of the repetitive disturbances for both the focus and radial control loop.

4.3.1 Focus Loop

The repetitive disturbances have mainly the rotational frequency f_{rot} of the spindle, that is not constant during playback, (see (3.11)), because a constant linear velocity is used

to read the data recorded on the disk (*i.e.* $v_a = \text{constant}$). The repetitive disturbances are mostly caused by the disk warping and disk displacement on the spindle $x_d(t)$.

Disk warping, illustrated in fig. 4.1, causes the non-zero disk vertical deviation $z_d(t)$ of the turning disk from the reference plane also for the zero disk displacement on the spindle ($x_d(t)=0$). Moreover, the disk vertical deviation $z_d(t)$ depends on the reference value of the objective lens $x_{\text{ref}}(t)$ in radial direction x . To describe disk warping, the linearity of the disk vertical deviation amplitude $z_{dA}(x_{\text{ref}}, t)$ has been assumed along the increasing reference value $x_{\text{ref}}(t)$, see (4.1).

Moreover, in real DVD/CD players, the disk is not ideally placed on the turntable of the spindle. It leads to the non-zero disk eccentricity ($x_{d_max} \neq 0$), defined as the maximal distance between the hole center and the spindle rotation axis, see fig. 3.19. It causes the time varying radial disk displacement on the spindle $x_d(x_{d_max}, t)$ which slightly influences disk warping. Modelling of the radial disk displacement on the spindle $x_d(x_{d_max}, t)$ is explained in section 4.3.2.

An amplitude of the vertical deviation $z_{dA}(x_{\text{ref}}, x_d, t)$, including the disk warping and disk displacement on the spindle, is given by the following expression:

$$z_{dA}(x_{\text{ref}}, x_d, t) = \frac{z_{d_min}}{x_{\text{min}}} \left(x_{\text{ref}}(t) + x_d(t) \right) = \frac{z_{d_max}}{x_{\text{max}}} \left(x_{\text{ref}}(t) + x_d(t) \right), \quad (4.1)$$

where z_{d_min} and z_{d_max} are the minimum/maximum vertical deviations at the minimal/maximal data zone radii for zero disk eccentricity ($x_{d_max} = 0$). The geometry of the vertical deviation sources is shown in fig. 4.1.

Finally, the vertical deviation model can be approximated by:

$$z_d(x_{\text{ref}}, x_d, t) = z_{dA}(t) \left(\sin(2\pi f_{\text{rot}} t + \varphi_F) \right), \quad (4.2)$$

where φ_F is the initial phase of the repetitive disturbance in the focus loop and the other variables have their usual meaning.

One can verify that the minimal vertical deviation z_{d_min} , see fig. 4.1, at the minimal radius of data zone x_{min} (where $f_{\text{rot}} = f_{\text{rot_max}}$) is *smaller* than the maximal vertical deviation z_{d_max} at the maximum radius of data zone x_{max} (where $f_{\text{rot}} = f_{\text{rot_min}}$).

Hence, the different requirements on the disturbance rejection can be given in frequency domain. The sensitivity function shaping method is a useful tool to design controllers satisfying these system requirements.

4.3.2 Radial Loop

The repetitive disturbance is mainly caused by time varying disk displacement on the spindle $x_d(x_{d_max}, t)$. Its amplitude, called the disk eccentricity, is defined by x_{d_max} here, see fig. 3.19. This repetitive disturbance $x_d(x_{d_max}, t)$ is synchronous with the time varying spindle frequency f_{rot} . The geometry of the radial deviation source is illustrated in fig. 4.2.

One can see that the disk displacement amplitude on the spindle x_{d_max} is constant from the starting radius of the data zone x_{min} , (where $f_{\text{rot}} = f_{\text{rot_max}}$), to the maximum radius of data zone x_{max} (where $f_{\text{rot}} = f_{\text{rot_min}}$).

Therefore the disk displacement on the spindle can be approximated by the following relationship:

$$x_d(x_{d_max}, t) = x_{d_max} \left(\sin(2\pi f_{\text{rot}} t + \varphi_R) \right), \quad (4.3)$$

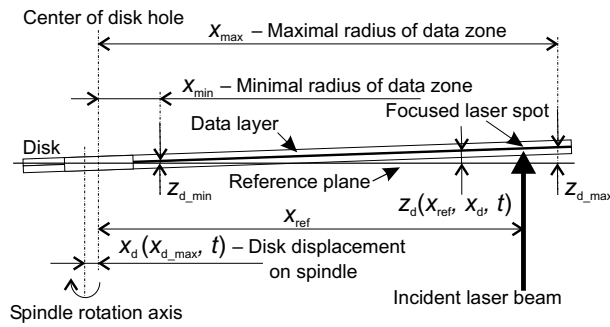


Figure 4.1: Geometry of the vertical deviation sources that are mainly caused by the disk warping and disk displacement on the spindle.

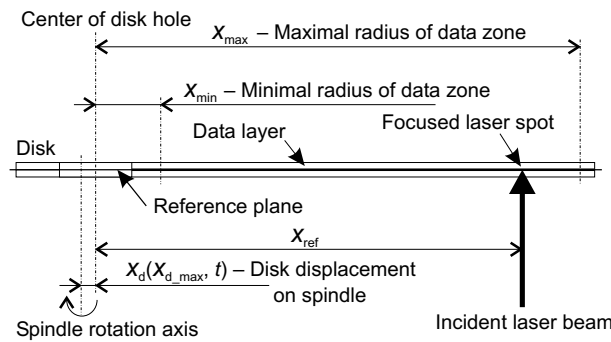


Figure 4.2: Geometry of the radial deviation source that is mainly caused by disk displacement on the spindle; *i.e.* the disk is not ideally placed on the spindle.

where φ_R is the initial phase of the repetitive disturbance in the radial loop and the other variables have been introduced above.

In general, the amount of data on the DVD/CD area increase proportionally to the square of the disk radius $x_{ref}(t) \in \langle x_{min}, x_{max} \rangle$. This means that more recorded data are read from the DVD/CD if the reference value $x_{ref}(t)$ is closed to its maximal value x_{max} . Therefore, the controller design method should also take into account the fact that the disturbance is larger near the maximal radius of data zone x_{max} than near the minimal radius of data zone x_{min} .

Such a method is the sensitivity function shaping method that has been used to design the controller of the radial loop.

4.4 Control Design Methodology

4.4.1 Combined Pole Placement/Sensitivity Function Shaping

The standard digital control configuration obtained with polynomial RS controller, see *e.g.* [LLM97], is presented in fig. 4.3. Part T of RST structure has been omitted because the control design in the focus/tracking loop of DVD/CD players only deals with the disturbance rejection problem. The detailed block diagrams of the focus/tracking control

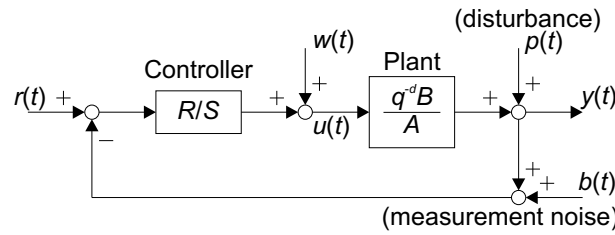


Figure 4.3: Closed-loop system with RS controller.

loops, illustrated in figs. 3.15 and 3.16, are possible to be transformed to the general structure for the digital closed-loop controller that is shown in fig. 4.3.

The linear time invariant model of the plant is in general described by the transfer function:

$$\begin{aligned} G(q^{-1}) &= \frac{q^{-d} B(q^{-1})}{A(q^{-1})} \\ &= \frac{q^{-d} (b_1 q^{-1} + \dots + b_{n_B} q^{-n_B})}{1 + a_1 q^{-1} + \dots + a_{n_A} q^{-n_A}}, \end{aligned} \quad (4.4)$$

where q^{-1} is the backward time shift operator, d is the pure time delay, T_s is the sampling period and $f_s = 1/T_s$ is the sampling frequency.

Pole placement method has been used to design the RS controller which has the following transfer function:

$$\begin{aligned} K(q^{-1}) &= \frac{R(q^{-1})}{S(q^{-1})} = \frac{R'(q^{-1}) H_R(q^{-1})}{S'(q^{-1}) H_S(q^{-1})} \\ &= \frac{r_0 + r_1 q^{-1} + \dots + r_{n_R} q^{-n_R}}{1 + s_1 q^{-1} + \dots + s_{n_S} q^{-n_S}}, \end{aligned} \quad (4.5)$$

where $H_R(q^{-1})$ and $H_S(q^{-1})$ denote the fixed parts of the controller (either imposed by the design or introduced in order to shape the sensitivity functions) and $R'(q^{-1})$, $S'(q^{-1})$ are the solutions of the Bezout equation:

$$AS'H_S + BR'H_R = P, \quad (4.6)$$

where P represents the characteristic polynomial (closed-loop poles).

The sensitivity functions play a crucial role in the robustness analysis of the closed-loop system with respect to modelling errors. These functions are *shaped* in order to assure *nominal performance* for the rejection of the disturbances and the stability of the closed-loop system in the presence of model mismatch.

The *output sensitivity function* $S_{yp}(q^{-1})$ is the transfer function between the output disturbance $p(t)$ and the plant output $y(t)$. It is given by expression:

$$S_{yp}(q^{-1}) = \frac{y(t)}{p(t)} = \frac{AS'H_S}{P}. \quad (4.7)$$

The other sensitivity functions are defined in the same way as follows: The *input sensitivity function* $S_{up}(q^{-1})$ is the transfer function between the output disturbance $p(t)$ and

the plant input $u(t)$. The *output sensitivity function* with respect to an input disturbance $S_{yw}(q^{-1})$ is the transfer function between the input disturbance $w(t)$ and the plant output $y(t)$. The *complementary sensitivity function* $S_{yr}(q^{-1})$ is the transfer function between the reference $r(t)$ and the plant output $y(t)$. This complementary sensitivity function with a negative sign is called the *noise sensitivity function* $S_{yb}(q^{-1})$.

In our case not only robustness (the modulus margin ΔM , delay margin $\Delta\tau$ and phase margin $\Delta\phi$) but also the performances specifications, (described in section 3.6), have to be checked. Therefore the sensitivity function shaping is a useful tool to the controller design in case of DVD/CD players.

4.4.2 Controller Order Reduction

Experience shows that simple linear model of DVD/CD system leads to sufficient good-performance controllers, see [BDCS01], given by specifications presented in section 3.6. More complex controllers, based on high order model, have higher performance but they are more difficult to implement, almost if there are implementation constraints on the controller order.

One useful methodology, that has been used here, is the balanced reduction method in state space domain which is using a Gramian of the balanced state-space realization of the reachable, observable, stable system, [SP96]. If the system is normalized properly, small elements in the balanced Gramian indicate states that can be removed to reduce the controller to lower order.

By applying this algorithm, a 3rd order controller has been obtained, *i.e.* $n_R = 3$, $n_S = 3$ in (4.5), which has been implemented in discrete-time on the industrial benchmark.

4.4.3 Generalized Stability Margin

The resulting reduced order controller should stabilize the nominal model and should give sensitivity functions which are close to the nominal ones in the critical frequency regions, to ensure performance and robustness. One way to verify the stability margin of the whole system is the generalized stability margin $b(K, G)$, [Vin93], defined from all sensitivity functions

$$b(K, G) = \begin{cases} \left\| \mathbf{T}(K, G) \right\|_{\infty}^{-1} & \text{if } (K, G) \text{ is stable,} \\ 0 & \text{otherwise,} \end{cases} \quad (4.8)$$

where

$$\mathbf{T}(K, G) = \begin{bmatrix} S_{yr} & S_{yw} \\ -S_{up} & S_{yp} \end{bmatrix},$$

in which S_{yr} , S_{yw} , S_{up} and S_{yp} have been defined in section 4.4.1. Higher value of $b(K, G)$ generally corresponds to better generalized stability margin of the whole system.

4.5 Control System Design

Some presented results have been published in [HBVFSd03a], [HBVFSd03b], [HVP03], [HBVF04]. An application of this methodology will be presented in more detail for the

Table 4.1: Values of the physical parameters for the radial actuator.

Symbol	Parameter	Nominal value
R_R	Coil resistance	6.5Ω
L_R	Coil inductance	$18 \mu\text{H}$
K_{Re}	Back efm constant	0.061 Vs/m
M_R	Actuator moving mass	$0.33 \times 10^{-3} \text{ kg}$
D_R	Damping constant	0.014 Ns/m
K_{Rs}	Elastic constant	35.2 N/m
K_{Rf}	Force constant	0.061 N/A
A_{Rdri1}	First power driver gain	3.13 V/V
A_{Rdri2}	Second power driver gain	4 V/V
A_{Ropt}	Optical gain & remanent gains	$1.447 \times 10^6 \text{ V/m}$

radial loop. Nevertheless, the final results are given for both the focus and radial control loop.

4.5.1 Plant Model in Radial Loop

A simplified, linear transfer function of the plant in the radial control loop $G_R(s)$ is derived from the physical equations of the radial system as follows, see *e.g.* [Sta98], [Fil03]:

$$G_R(s) = \frac{\frac{K_{Rf}}{M_R L_R} A_{Rdri1} A_{Rdri2} A_{Ropt}}{s^3 + \left(\frac{R_R}{L_R} + \frac{D_R}{M_R}\right) s^2 + \left(\frac{D_R R_R}{M_R L_R} + \frac{K_{Rs}}{M_R} + \frac{K_{Rf} K_{Re}}{M_R L_R}\right) s + \frac{K_{Rs} R_R}{M_R L_R}} \quad (4.9)$$

The values of the physical parameters are presented in table 4.1. The discrete transfer function of the radial system $G_R(q^{-1})$ is given by conversion from the continuous-time to the discrete-time using a zero-order hold and the sampling period $T_s = 8.1 \mu\text{s}$.

4.5.2 Physical Model Validation

In [Fil03] the frequency-domain procedure has been used to identify the plant model of an industrial DVD-video player servo system. SISO local linear models of the moving lens focus/radial actuators have been derived, which are valid around a certain track location, through closed-loop frequency response measurements obtained with a dynamic signal analyzer and curve fitting procedure.

In this chapter the validated physical plant models, given in [Fil03], are used for designing a low complexity focus/radial loop controllers to achieve enhanced disturbance attenuation.

Moreover, the coupling phenomena between the focus and radial loops have been also analyzed in [Fil03], using the measured frequency responses of the system closed-loop transfer functions. It can be pointed out that, in the frequency range of interest for control, the dynamic interaction between the both loops remains relatively weak.

4.5.3 Nominal and Uncertainty Model

To verify the robustness of the proposed controllers, different models of the radial actuator, based on the parameters specification and their variation, have been created. The

Table 4.2: Values of the radial actuator physical parameters together with their maximum percentage variation.

Symbol	Nominal value	Variation
R_R	6.5 Ω	$\pm 15\%$
L_R	18 μH	$\pm 33\%$
f_{Rn}	52 Hz	$\pm 5\%$
M_R	0.33 g	$\pm 10\%$
K_{RDC}	$0.27 \cdot 10^{-3}$	$\pm 20\%$

nominal model of the plant is obtained considering the values of physical parameters of a DVD optical pick-up unit, which is used in the industrial application. Then, the model set is created by taking into account the variation of each physical parameter in an interval of values, as indicated in the optical pick-up unit data-sheet, [FSBVSd03].

The nominal values of the actuator physical parameters, together with their maximum percentage variation, are shown in table 4.2. f_{Rn} and K_{RDC} are the values of the actuator resonance frequency and DC sensitivity, that are used to compute the values of K_{Re} , D_R , K_{Rs} and K_{Rf} in the model transfer function (4.9), see [Fil03] for more details.

4.5.4 Standard Controller

The 2nd order lead-lag controller $K_{act}(q^{-1})$, given by (4.10), is used in many actual DVD/CD applications as a standard controller structure.

$$K_{act}(q^{-1}) = g_{LL} \frac{(1 - z_{LL1} q^{-1})(1 - z_{LL2} q^{-1})}{(1 - p_{LL1} q^{-1})(1 - p_{LL2} q^{-1})} \quad (4.10)$$

This standard 2nd order lead-lag controller $K_{act}(q^{-1})$ for RS structure: $n_R = 2, n_S = 2$ is actually replaced by a 3rd order $K_{RS3}(q^{-1})$: $n_R = 3, n_S = 3$ or a 4th order $K_{RS4}(q^{-1})$: $n_R = 4, n_S = 4$ controller in order to meet higher performance on the disturbance rejection.

The aim of this work is to provide a methodology to design the 3rd and 4th order controllers that improve the actual performance on the disturbance rejection and fulfil the implementation constraints on the actual DVD platform.

4.5.5 New Controller Design

The 3rd order RS controller $K_{RS3}(q^{-1})$, ($n_R = 3, n_S = 3$), and the 4th order RS controller $K_{RS4}(q^{-1})$, ($n_R = 4, n_S = 4$), are designed for an over-speed factor $N = 1.5$.

When the disk rotates at higher speeds, say Nv_a , the minimal, maximal and cross-over frequencies (f_{rot_min} , f_{rot_max} , f_{cr}) have to be linearly shifted by the over-speed factor N , and the focus and radial accelerations have to be multiplied by the factor N^2 .

It means that the maximal spindle rotational frequency is $f_{rot_max} = 23.1 \cdot 1.5 = 34.7$ Hz at the minimal radius of data zone x_{min} while $f_{rot_min} = 14.4$ Hz at the maximal radius of data zone x_{max} . The cross-over frequency of the compensated open-loop is equal to $f_{cr} = 3.6$ kHz in this case.

The main improvement of repetitive disturbance rejection has been done by output sensitivity function shaping at the spindle frequencies $f_{\text{rot}} \in \langle f_{\text{rot_min}}, f_{\text{rot_max}} \rangle$. Generally, more recorded data are read from the DVD/CD if the reference value $x_{\text{ref}}(t)$ is closed to its maximal value x_{max} , (see section 4.3.2). Therefore the main improvement of disturbance rejection has been imposed to the minimal rotation frequency $f_{\text{rot_min}}$ by modification on the output sensitivity function template. For higher spindle frequencies $f_{\text{rot}} \in \langle f_{\text{rot_min}}, f_{\text{rot_max}} \rangle$, the template has been modified with respect to the ratio of the minimal/maximal data zone radii $\left(\frac{x_{\text{min}}}{x_{\text{max}}}\right)$ as it is described in the following.

The minimum output sensitivity function magnitude S_{low} at the low rotational frequencies $f_{\text{rot}} \leq f_{\text{rot_max}}$, for the manufacturing process, is given by:

$$S_{\text{low}} = 20 \log \left(\frac{|x_{\text{d_high}}|}{|x_{\text{d_low}}|} \right) = -67.13 \text{ dB.} \quad (4.11)$$

This slight modification on the $|S_{\text{yp}}|$ template has been proposed to suppress the repetitive disturbances:

1. $f_{\text{rot}} = 34.7 \text{ Hz}$:
 $|S_{\text{yp}}| = S_{\text{low}} - 27 \text{ dB} = -94.13 \text{ dB}$
2. $f_{\text{rot}} = 14.4 \text{ Hz}$:
 $|S_{\text{yp}}| = 20 \log \left(\frac{|x_{\text{d_high}}|}{|x_{\text{d_low}}|} \cdot \frac{x_{\text{min}}}{x_{\text{max}}} \right) - 27 \text{ dB} = -101.80 \text{ dB}$
3. $f_{\text{rot}} \in \langle 14.4 \text{ Hz}, 34.7 \text{ Hz} \rangle$:
 $|S_{\text{yp}}| = \text{linear interpolation between } |S_{\text{yp}}| \text{ at two given frequencies: } f_{\text{rot_min}} = 14.4 \text{ Hz and } f_{\text{rot_max}} = 34.7 \text{ Hz}$
4. $f_{\text{rot}} > 34.7 \text{ Hz}$: $|S_{\text{yp}}|$ is given by specification in [ECM01]

The low limit at the rotational frequency $f_{\text{rot_max}} = 34.7 \text{ Hz}$ has been toughened up with 27 dB. The low limit at the rotational frequency $f_{\text{rot_min}} = 14.4 \text{ Hz}$ has been linearly decreased with respect to the ratio of data zone radii $\left(\frac{x_{\text{min}}}{x_{\text{max}}}\right)$, in order to take into account data distribution on the disk, and also shifted with 27 dB.

The specification requirements, defined for the DVD manufacturing and for the control design by the compensated open-loop (normalized servo) transfer function, together with the modification are illustrated as the output sensitivity function modulus $|S_{\text{yp}}|$ templates in fig. 4.4.

The controller design of the fourth order controller K_{RS4} has been realized using the following specifications:

- P (closed-loop poles, see also (4.6)):
 - a pair of complex poles near the models slowest vibration frequency $f_{\text{D}} = 52 \text{ Hz} \rightarrow 170 \text{ Hz}$ but well damped $\xi_{\text{D}} = 0.068 \rightarrow 0.936$
 - two multiple real poles $\gamma_{\text{F}} = 0.9$ for keeping in the $|S_{\text{yp}}|$, $|S_{\text{up}}|$ templates
 - one complex pole $f_{\text{F}} = 13800 \text{ Hz}$, $\xi_{\text{F}} = 0.927$ to restrain the controller action in higher frequencies where the gain of the system is low
- H_{S} : a pair of complex poles $f_{\text{S}} = 19 \text{ Hz}$, $\xi_{\text{S}} = 0.4$ to ensure disturbance rejection in the frequency range $f_{\text{rot}} \in \langle 14.4 \text{ Hz}, 34.7 \text{ Hz} \rangle$

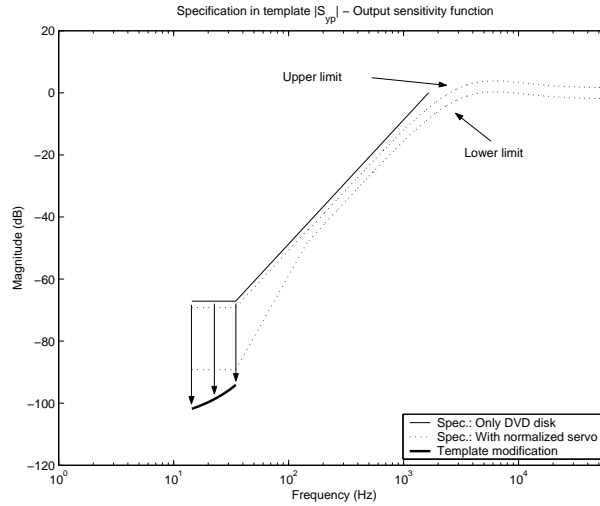


Figure 4.4: Desired template for the modulus of the output sensitivity function $|S_{yp}|$ for radial tracking (radial loop) of the DVD, $N = 1.5$.

- H_R : a real zero $\gamma_R = 0.1$ to lower the magnitude of the input sensitivity function $|S_{up}|$ at high frequencies where the gain of the system is low.
- The resulting controller has the orders $n_R = 5$ and $n_S = 5$. Therefore the balanced reduction method [SP96] has been used to obtain a controller structure $n_R = 4$ and $n_S = 4$.
- Check the sensitivity functions $|S_{yp}|$, $|S_{up}|$, $|S_{yr}|$ again. End the design procedure if the requirements on the sensitivity functions were satisfied.

The same procedure has been used to the 3rd order RS controller K_{RS3} design except balanced reduction. Nevertheless, the restrictions on the resulting controller result in smaller performance in the frequency range $f_{rot} \in \langle 14.4 \text{ Hz}, 34.7 \text{ Hz} \rangle$.

Remark 9 *The 2nd order controller is used as the standard controller in the actual DVD-video industrial player radial loop, i.e. $K_{act}(q^{-1})$, ($n_R = 2, n_S = 2$). To compare the same complexity controller performances together, the 2nd order controller has been also designed as follows.*

Since the pole/placement method together with the balance reduction method do not allow to design a suitable 2nd order controller K_{RS2} , a digital version of the continuous-time PID controller has been used. It is a proportional-summation-derivation structure, using the filtrated derivation part. This controller can be defined by, see e.g. [Piv02]:

$$K_{RS2}(q^{-1}) = K \left(1 + \frac{T_s q^{-1}}{T_i(1 - q^{-1})} + N_d \frac{(1 - q^{-1})}{(1 - e^{-\frac{T_s N_d}{T_d} q^{-1}})} \right), \quad (4.12)$$

where K is the proportional gain, T_i is the integration-time constant, T_d is the derivation-time constant and T_d/N_d is the filtration constant, $N_d \in \langle 3, 20 \rangle$. The equation (4.12)

can be rewritten to general polynomial RS form as follows:

$$K_{RS2} = \frac{(K + KN_d) + (-K e^{-\frac{T_s N_d}{T_d}} - K + \frac{KT_s}{T_i} - 2KN_d) q^{-1} + (K e^{-\frac{T_s N_d}{T_d}} - \frac{K}{T_i} e^{-\frac{T_s N_d}{T_d}} + KN_d) q^{-2}}{1 + (-e^{-\frac{T_s N_d}{T_d}} - 1) q^{-1} + (e^{-\frac{T_s N_d}{T_d}}) q^{-2}}. \quad (4.13)$$

The parameters of the 2nd order controller K_{RS2} have been obtained by using the sensitivity function shaping methodology, presented above, where the well known Ziegler-Nichols's relations for the continuous-time PID controller have been used as the initial parameters setting (for K , T_i , T_d).

Remark 10 On the other hand the 3rd order controller is used as the standard controller in the DVD-video industrial player focus loop, i.e. $K_{act}(q^{-1})$, ($n_R = 3, n_S = 3$). In this case, the designed 2nd order controller, given by (4.13), has not satisfied the performance specifications and therefore the results obtained by using this controller are not presented in the following.

4.6 Simulation and Experimental Results

In this section we present simulation and experimental results obtained when the full and the reduced order controllers are considered.

4.6.1 Focus Loop: Simulation Experiments

The simulation results are presented for the designed RS controllers of the 3rd/4th order (K_{RS3} , K_{RS4}) and the 3rd order controller which is the actual implemented (standard) controller on the DVD-player benchmark (K_{act}).

The disturbance rejection is illustrated by the output sensitivity function modulus $|S_{yp}|$ in fig. 4.5 for the K_{RS4} and K_{act} controller. Notice that the perturbations suppression at $f_{rot} = 22$ Hz has been achieved by the H_S polynomial choice. Fig. 4.6 presents the input sensitivity function modulus $|S_{up}|$ while fig. 4.7 presents the complementary sensitivity function modulus $|S_{yr}|$ for these controllers. The lower peak in $|S_{yp}|$ and lower values of $|S_{up}|$ in low frequencies for the controller K_{RS4} than ones for actual controller K_{act} are seen in fig. 4.5 and fig. 4.6, respectively.

The disturbance rejection of the 3rd order controller K_{RS3} and actual controller K_{act} is shown by the output sensitivity function modulus $|S_{yp}|$ in fig. 4.8.

A good controller order reduction methodology and a good generalized stability margin of the 3rd/4th order designed controllers (K_{RS3} , K_{RS4}) are shown in table 4.3, where the parameters of reduced controllers are n_R , n_S , $b(K, G)$ and the parameters of non-reduced controllers are \widehat{n}_R , \widehat{n}_S , $\widehat{b(K, G)}$.

Finally, the comparison of the various reduced controllers performances is shown in table 4.4. One can see that the specification requirements are satisfied by all controllers. In additional, the lower values of $|S_{yp}|_{14.4}$ and $|S_{up}|_{34.7}$ for the designed controllers than ones for actual controller lead to higher disturbance suppression at low frequencies which correspond to the varying disk rotational frequencies f_{rot} . This advantage is compensated by the higher overshoot values of the closed-loop step response and maximal values of output sensitivity function modulus $|S_{yp}|_{max}$ for the designed controllers in contrast to actual controller.

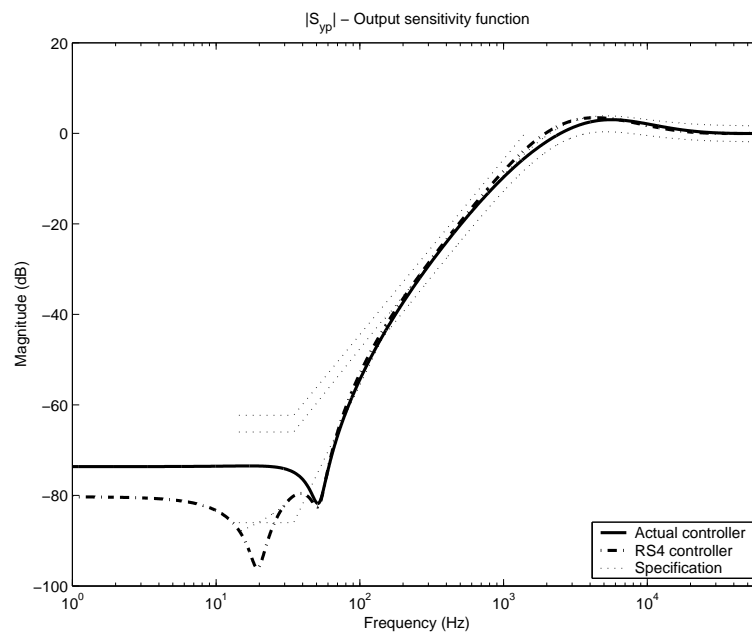


Figure 4.5: Output sensitivity function, focus loop.

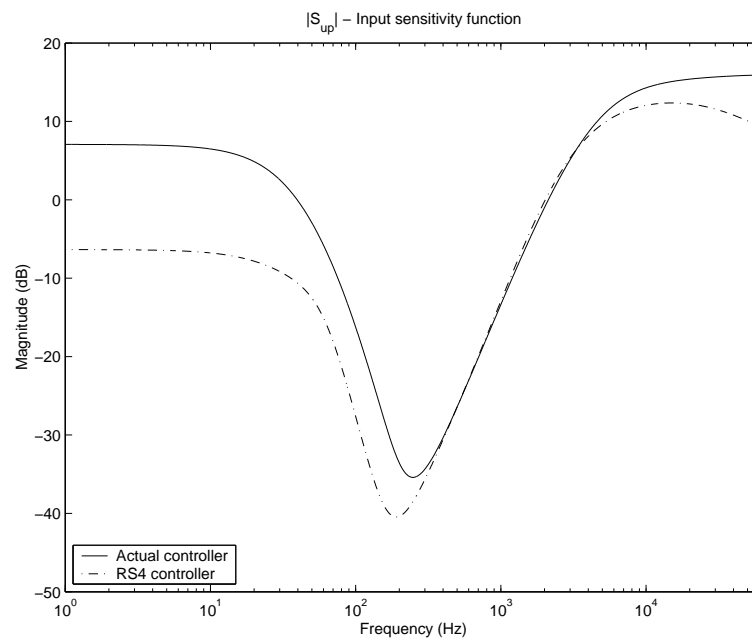


Figure 4.6: Input sensitivity function, focus loop.

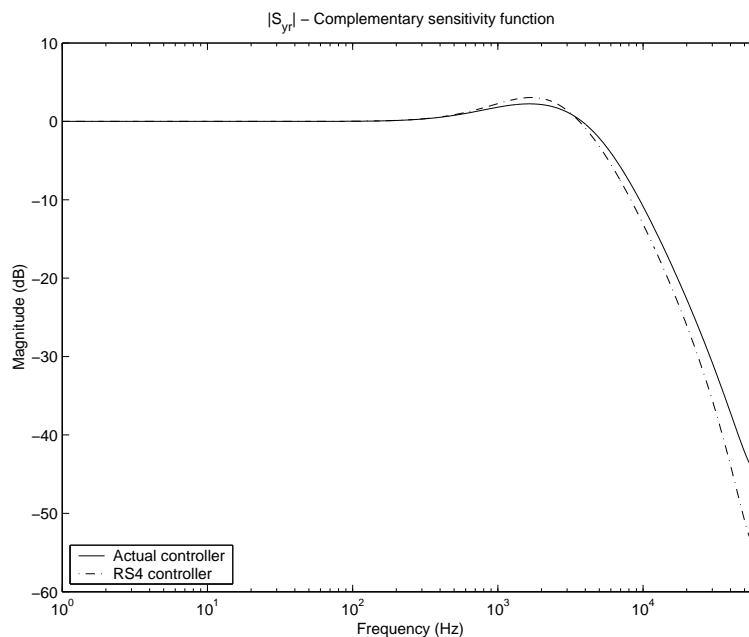


Figure 4.7: Complementary sensitivity function, focus loop.

Table 4.3: Comparison of the controller order reduction and generalized stability margin, $N = 1.5$, focus loop.

$K(q^{-1})$	n_R	n_S	\widehat{n}_R	\widehat{n}_S	$b(K, G)$	$b(\widehat{K}, \widehat{G})$
K_{act}	3	3	3	3	0.15563	
K_{RS3}	3	3	4	4	0.16476	0.16476
K_{RS4}	4	4	5	5	0.23351	0.23294

4.6.2 Focus Loop: Real-time Measurements

The experimental results are shown only for the 3rd order RS controller K_{RS3} and the 3rd order actual (standard) controller K_{act} because the 4th order controller is not implementable into a DSP controller structure in DVD-video player benchmark.

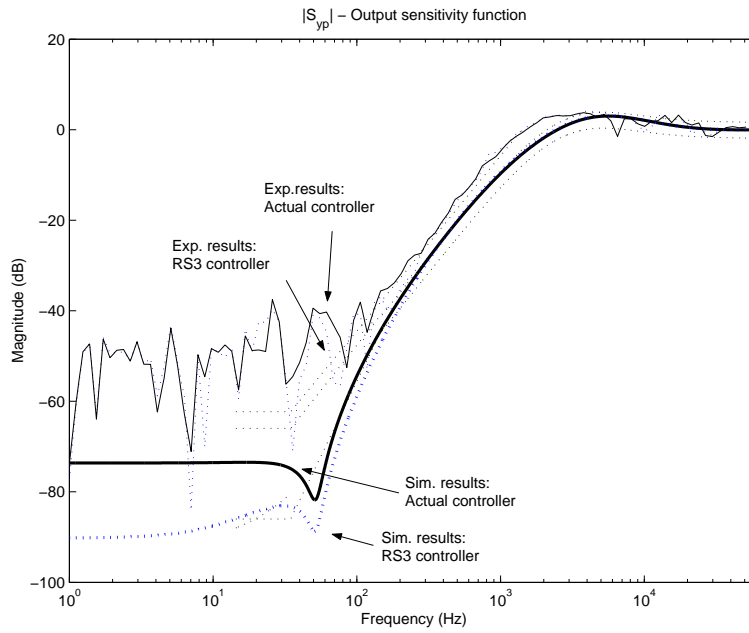
In fig. 4.8 the measured and the simulated output sensitivity function magnitudes $|S_{yp}|$, obtained for the 3rd order controller K_{RS3} are presented. Here, as term of comparison, we present also the measured and the simulated output sensitivity function magnitudes computed with the actual implemented controller K_{act} .

From this figure it can be seen that closed-loop measurements are very ill-conditioned at low frequencies since the high attenuation level needed at low frequencies to suppress external disturbances reduces the output Signal to Noise Ratio (SNR). This is true when both designed and actual controllers are used. This explains why, in fig. 4.8, the measured frequency responses of $|S_{yp}|$, obtained for both controllers, lie above the specification requirements up to 100 Hz, and matches with the simulated curves obtained with the same controllers only after this frequency.

Finally, in figs. 4.9 and 4.10 the measured Power Spectral Densities (PSD) of the focus error signal e_F , obtained for the 3rd order controller K_{RS3} and the actual controller

Table 4.4: Comparison of the various reduced controllers, $N = 1.5$, focus loop.

$K(q^{-1})$	$ S_{yp} _{\max}$ (dB)	$ S_{yp} _{14.4}$ (dB)	$ S_{yp} _{34.7}$ (dB)	$ S_{up} _{\max}$ (dB)	t_r (μs)	Overshot (%)	$\Delta\tau$ (μs)	$\Delta\phi$ (degree)
K_{act}	3.02	-73.5	-74.9	16.1	72.9	25	48.0	51.6
K_{RS3}	3.28	-80.2	-82.5	15.5	72.9	31	46.4	48.5
K_{RS4}	3.45	-88.0	-80.3	12.3	81.0	32	48.9	46.1
Spec.	3.85	-66.0	-66.0		122.5			

Figure 4.8: The measured and the simulated magnitude of the output sensitivity function obtained for the 3rd order controller K_{RS3} and the actual controller K_{act} , focus loop.

implemented in the current industrial solution K_{act} , are presented. Measurements have been acquired by using the worst-case disk (test disk having nominal vertical deviation at the disk outer edge $z_{d_max} = 0.5$ mm and nominal eccentricity of $x_{d_max} \approx 0$ μm) and for two different disk rotational frequencies of about $f_{rot} = 15$ Hz and $f_{rot} = 33$ Hz. From these figures it appear that the 3rd order designed controller K_{RS3} provides better level of periodic disturbance rejection than the actual implemented (standard) controller K_{act} , for these rotational frequencies.

These results also point out that the obtained improvements are still influenced by disk rotational frequency f_{rot} and its harmonics.

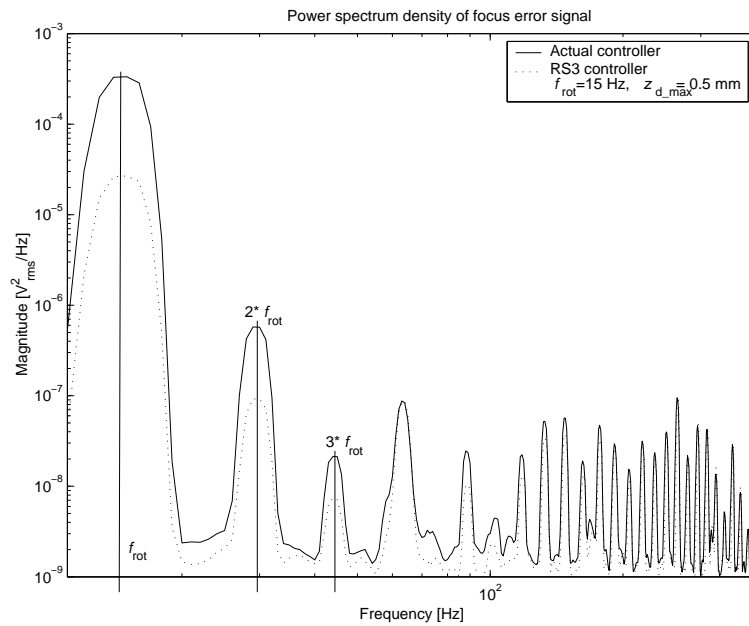


Figure 4.9: The measured power spectrum density of the focus error signal e_F for the 3rd order designed controller K_{RS3} and actual implemented controller K_{act} . The test disk has very small disk eccentricity x_{d_max} , but with high disk vertical deviation at the disk outer edge $z_{d_max} = 0.5$ mm, $f_{rot} = 15$ Hz.

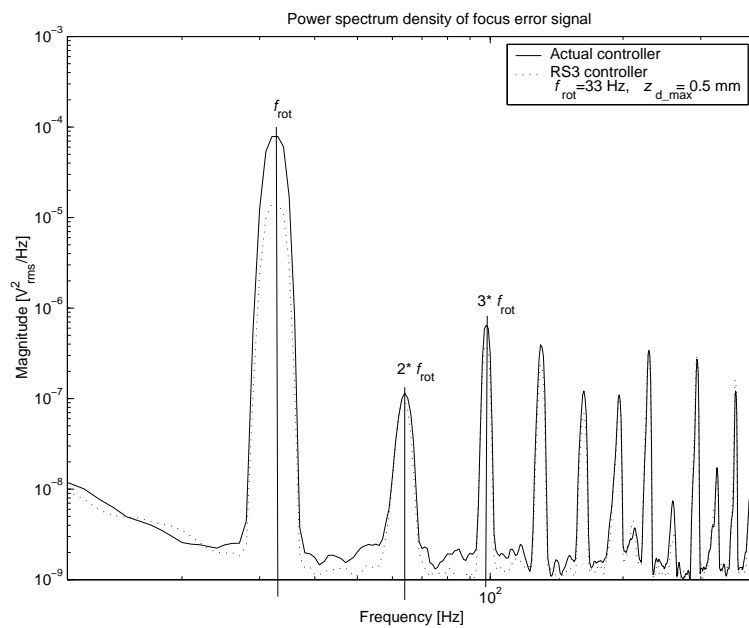


Figure 4.10: The same as in fig. 4.9 but for $f_{rot} = 33$ Hz.

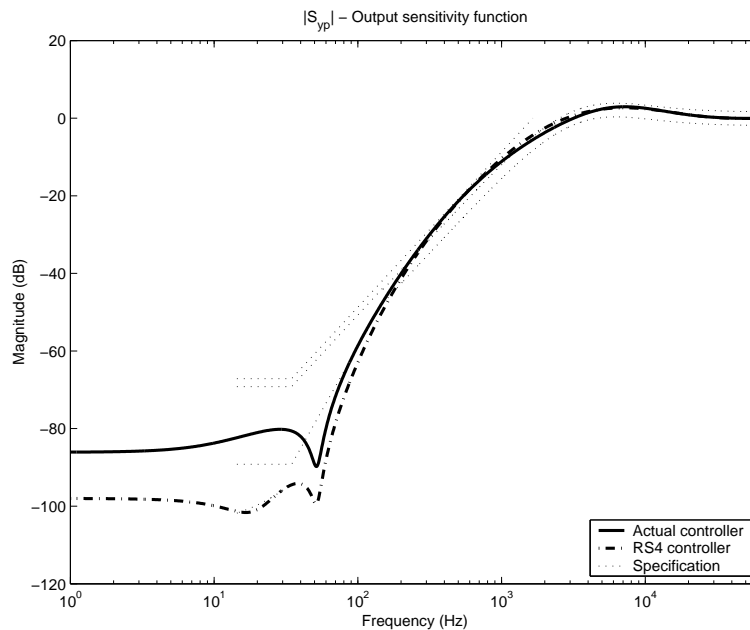


Figure 4.11: Output sensitivity function, radial loop.

4.6.3 Radial Loop: Simulation Experiments

The simulation results are shown for the designed RS controllers of the 2nd/3rd/4th order (K_{RS2} , K_{RS3} , K_{RS4}) and the 2nd order actual implemented (standard) controller K_{act} . In fact, actual controller K_{act} is the 2nd order controller because some constants are set to zero in the implemented structure which can be also used for the 3rd order controller implementation. Therefore the 4th order can not be implemented on the industrial benchmark.

The disturbance rejection is illustrated by the output sensitivity function modulus $|S_{yp}|$ in fig. 4.11 for the K_{RS4} and K_{act} controller. Notice that the perturbations suppression at $f_{rot} = 19$ Hz has been achieved by the fixed H_S polynomial choice, as explained in section 4.5.5.

Figs. 4.12 and 4.13 present the input sensitivity function modulus $|S_{up}|$ and complementary sensitivity function modulus $|S_{yr}|$ for these controllers, respectively. The lower peak in $|S_{yp}|$ and lower values of $|S_{up}|$ in high frequencies for the controller K_{RS4} than ones for actual controller K_{act} are seen in these figures. However, in case of the 4th order controller K_{RS4} , higher peak of $|S_{yr}|$ is a trade-off between disturbance rejection and robustness requirements.

To compare all designed controllers, the disturbance rejection of the 3rd order controller K_{RS3} and the 2nd order controller K_{RS2} together with K_{act} and K_{RS4} is shown by the output sensitivity function modulus $|S_{yp}|$ in fig. 4.14. Here, one can see the possibilities of each controllers to suppress the repetitive disturbances at low frequencies.

Another way how to compare the designed controllers is given in fig. 4.15 where the step responses of the closed-loop are shown. One can see that the rise time is satisfied in all cases but the overshoot values are different.

A good controller order reduction methodology and the generalized stability margin of the 2nd/3rd/4th order designed controllers (K_{RS2} , K_{RS3} , K_{RS4}) are shown in table 4.5,

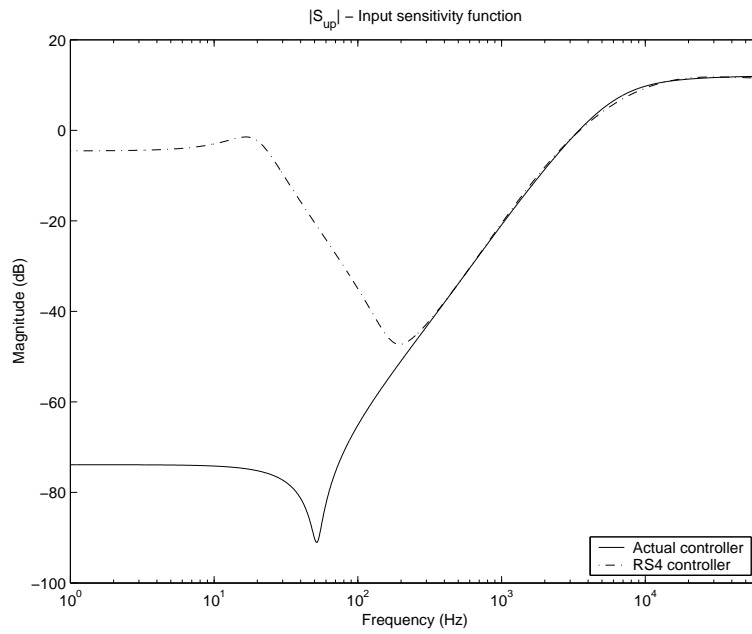


Figure 4.12: Input sensitivity function, radial loop.

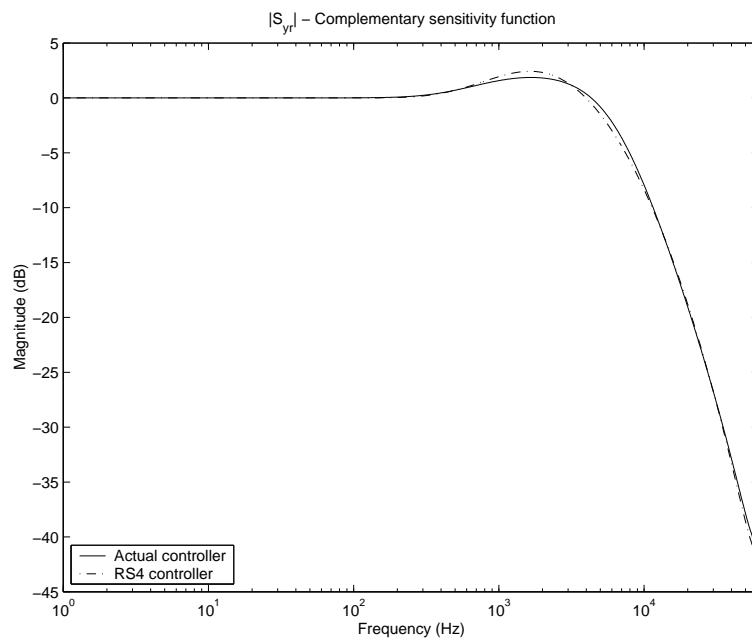


Figure 4.13: Complementary sensitivity function, radial loop.

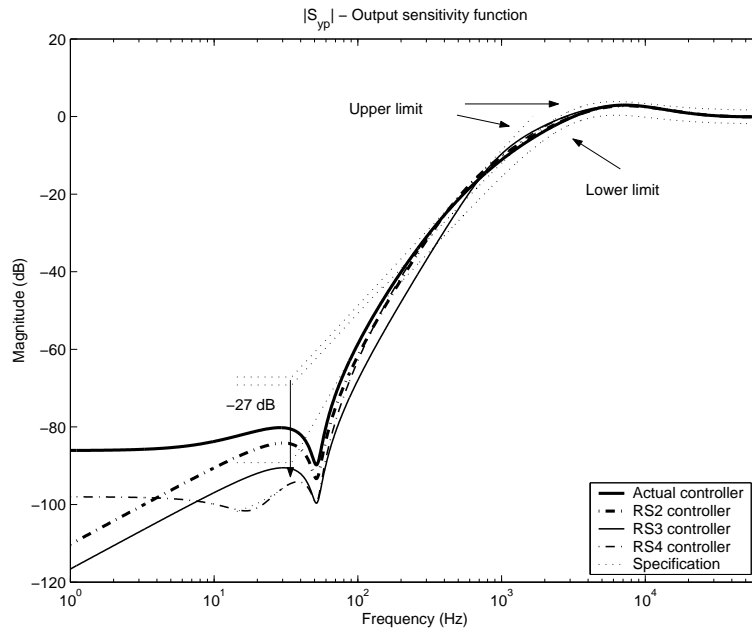


Figure 4.14: Summary: Output sensitivity function, radial loop.

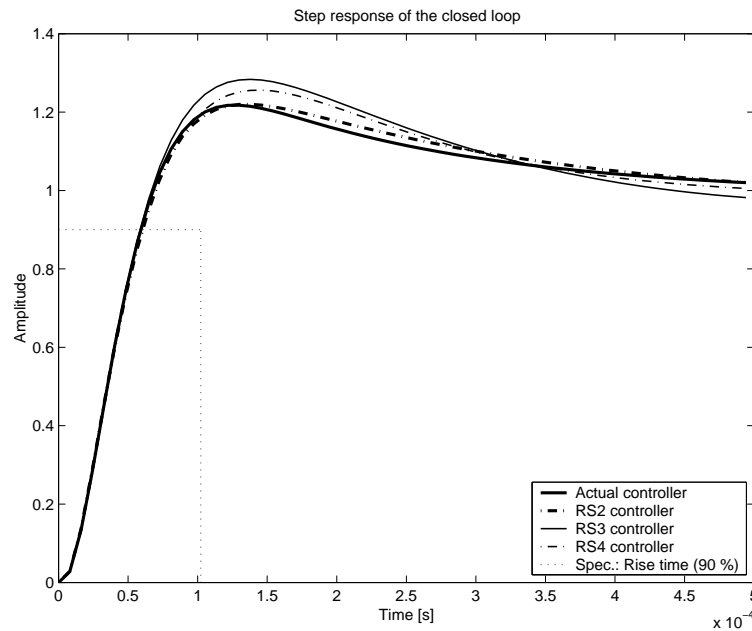


Figure 4.15: Summary: Step response of the closed-loop, radial loop.

Table 4.5: Comparison of the controller order reduction and generalized stability margin, $N = 1.5$, radial loop.

$K(q^{-1})$	n_R	n_S	\widehat{n}_R	\widehat{n}_S	$b(K, G)$	$b(\widehat{K}, \widehat{G})$
K_{act}	2	2	2	2	0.2127	
K_{RS2}	2	2	2	2	0.1984	
K_{RS3}	3	3	3	3	0.2156	
K_{RS4}	4	4	5	5	0.2253	0.2151

Table 4.6: Comparison of the various reduced controllers, $N = 1.5$, radial loop.

$K(q^{-1})$	$ S_{yp} _{max}$ (dB)	$ S_{yp} _{14.4}$ (dB)	$ S_{yp} _{34.7}$ (dB)	$ S_{up} _{max}$ (dB)	t_r (μs)	Overshot (%)	$\Delta\tau$ (μs)	$\Delta\phi$ (degree)
K_{act}	2.95	-82.3	-80.6	12.1	64.8	22	42.1	54.4
K_{RS2}	2.78	-87.8	-84.4	12.5	64.8	22	44.0	55.4
K_{RS3}	2.88	-94.1	-90.8	12.0	64.8	28	42.4	52.1
K_{RS4}	2.67	-101.3	-94.5	11.8	64.8	26	44.7	53.5
Spec.	3.86	-69.2	-69.2		102			

where the parameters of reduced controllers are n_R , n_S , $b(K, G)$ and the ones of non-reduced controllers are \widehat{n}_R , \widehat{n}_S , $b(\widehat{K}, \widehat{G})$.

Finally, the comparison of the various reduced controllers performances is given in table 4.6. One can see that the specification requirements are satisfied by all controllers. Moreover, the lower values of $|S_{yp}|_{14.4}$ and $|S_{up}|_{34.7}$ for the designed controllers than ones for actual controller lead to higher disturbance suppression at low frequencies which correspond to the varying disk rotational frequencies f_{rot} as in the focus control loop.

In table 4.6 the overshoot values of the all controllers are comparable to each other and the maximal values of output sensitivity function modulus $|S_{yp}|_{max}$ for the designed controllers are even smaller in contrast to actual controller.

Figs. 4.16, 4.17 and 4.18 illustrate the envelopes of the output sensitivity functions modulus $|S_{yp}|$ that have been calculated for the plant model set, given from the percentage variation of the actuator physical parameters, and the actual/2nd/3rd order designed controllers, respectively. This simple robustness analysis have been also used to show the envelopes of the closed-loop step responses, illustrated in figs. 4.19, 4.20 and 4.21. One can see that the stability templates and the desired performances are fulfilled in all cases.

4.6.4 Radial Loop: Real-time Measurements

The experimental results are shown for the 2nd/3rd order RS controller (K_{RS2} , K_{RS3}) and the 2nd order actual (standard) controller K_{act} because the 4th order RS controller is not implementable into a DSP controller structure in DVD-video player benchmark.

In figs. 4.22, 4.23 and 4.24 the frequency responses of the output sensitivity function magnitude are shown for the actual/2nd/3rd order controllers (K_{act} , K_{RS2} , K_{RS3}) on simulation using the actuator linear model and on experimentations using the industrial

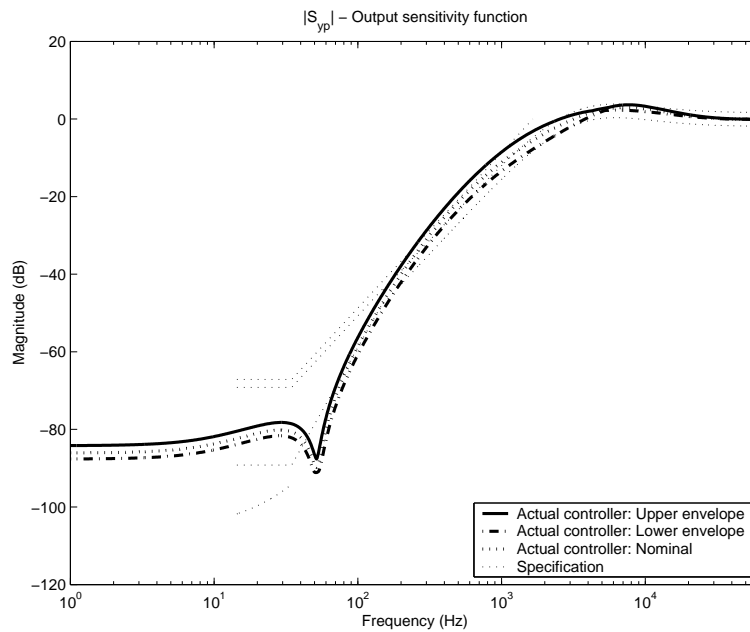


Figure 4.16: Envelopes of the output sensitivity functions, actual controller, radial loop.

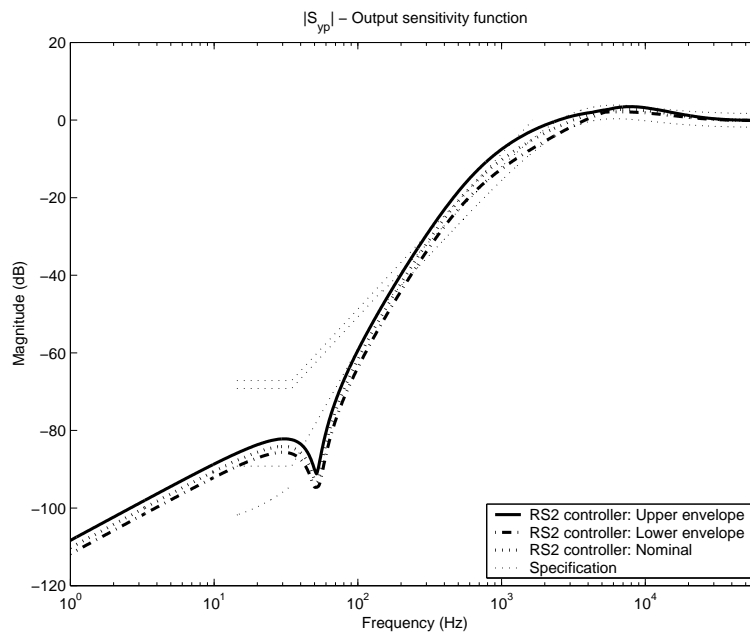


Figure 4.17: Envelopes of the output sensitivity functions, RS2 controller, radial loop.

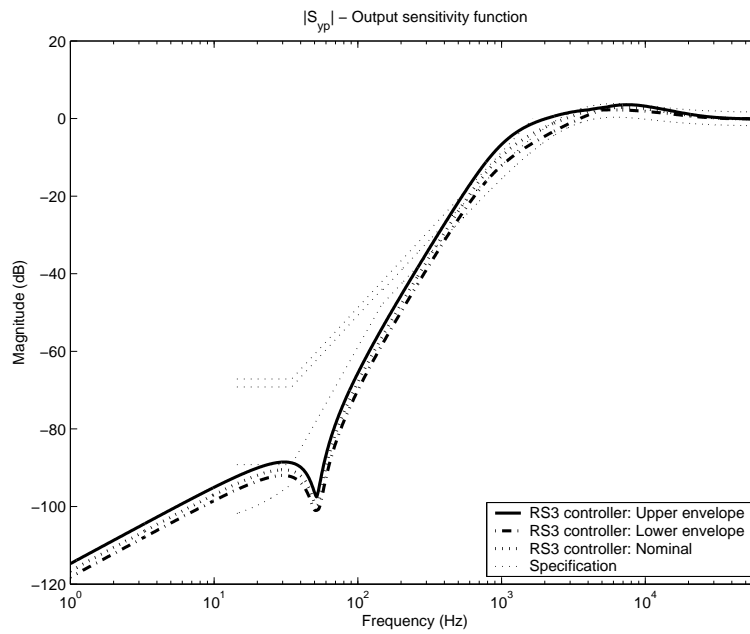


Figure 4.18: Envelopes of the output sensitivity functions, RS3 controller, radial loop.

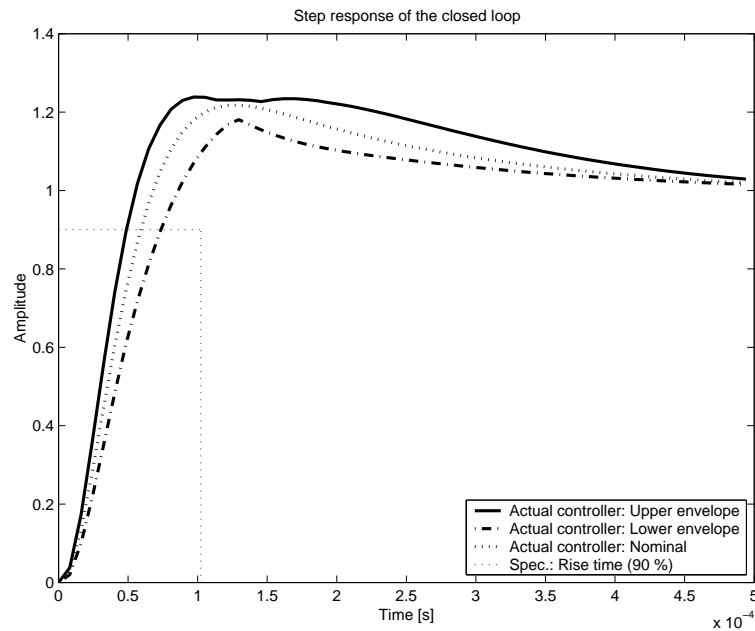


Figure 4.19: Envelopes of the radial closed-loop step responses, actual controller.

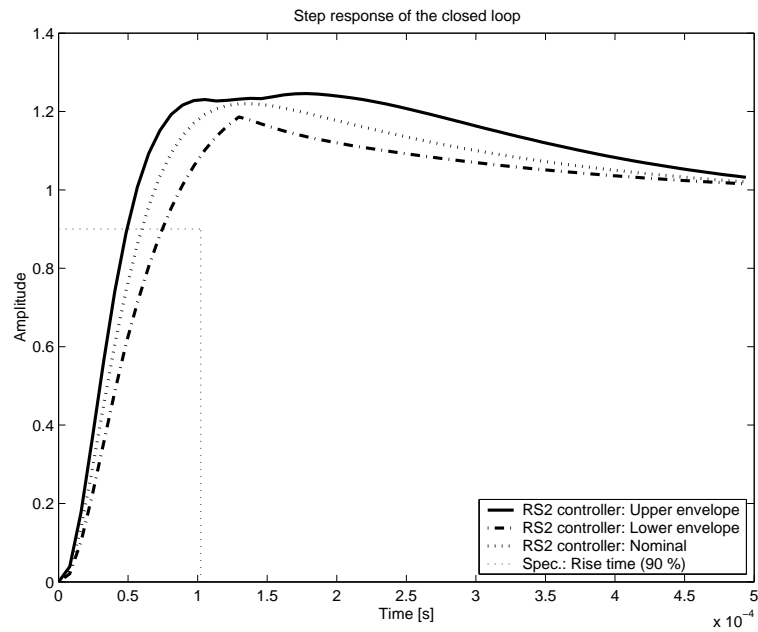


Figure 4.20: Envelopes of the radial closed-loop step responses, RS2 controller.

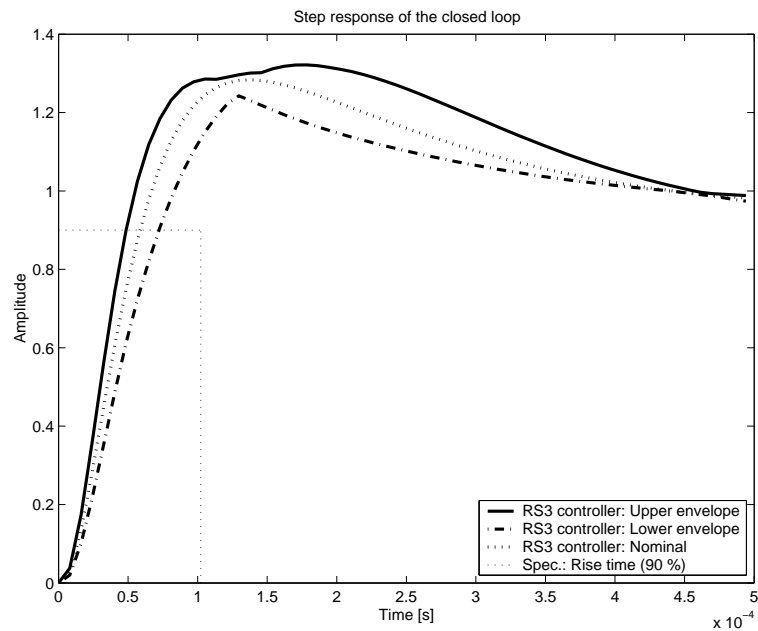


Figure 4.21: Envelopes of the radial closed-loop step responses, RS3 controller.

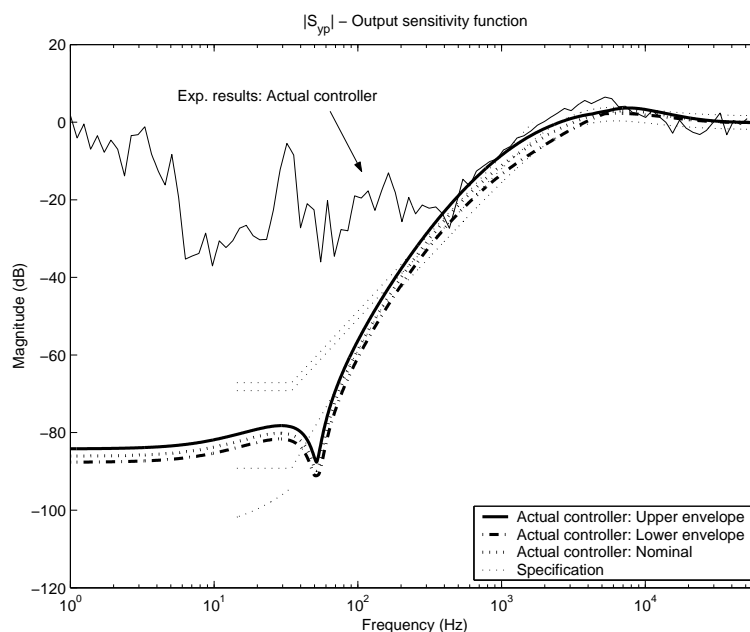


Figure 4.22: Magnitude of the output sensitivity function obtained from measurements and the envelopes of output sensitivity functions from simulation for K_{act} , radial loop.

benchmark. In this setup, since real-time measurements are only available under closed-loop conditions, the measurements are very ill-conditioned at low frequencies, like in the focus control loop.

This explains why, in figs. 4.22, 4.23 and 4.24, the measured frequency responses of $|S_{yp}|$, obtained for three controllers (K_{act} , K_{RS2} , K_{RS3}), lie above the specification requirements up to 500 Hz, and matches with the simulated envelope curves obtained with the same controllers only after this frequency. Nevertheless, one can see that the stability and performances conditions are satisfied for all controllers.

Finally, in figs. 4.25 and 4.26 the measured Power Spectral Densities (PSD) of the radial error signal e_R , obtained for the 2nd/3rd order designed controllers (K_{RS3} , K_{RS3}) and the actual controller implemented in the current industrial solution K_{act} , are presented. Measurements have been acquired by using the best-case disk (test disk having nominal vertical deviation at the disk outer edge $z_{d_max} \approx 0$ mm and nominal eccentricity of $x_{d_max} \approx 0$ μ m) and for two different disk rotational frequencies of about $f_{rot} = 15$ Hz and $f_{rot} = 33$ Hz.

From these figures it appear that the 2nd order designed controller K_{RS2} provides comparable level of periodic disturbance rejection as the actual implemented (standard) controller K_{act} , for these rotational frequencies. On the other hand the 3rd order designed controller K_{RS3} provides better level of periodic disturbance rejection than the actual implemented (standard) controller K_{act} , for these rotational frequencies. One can see that the disturbance rejection levels between disk rotational frequencies harmonics is trade off between the controller complexity and first harmonics disturbance rejection.

These results also point out, as in focus control loop, that the obtained improvements are still influenced by disk rotational frequency f_{rot} and its harmonics.

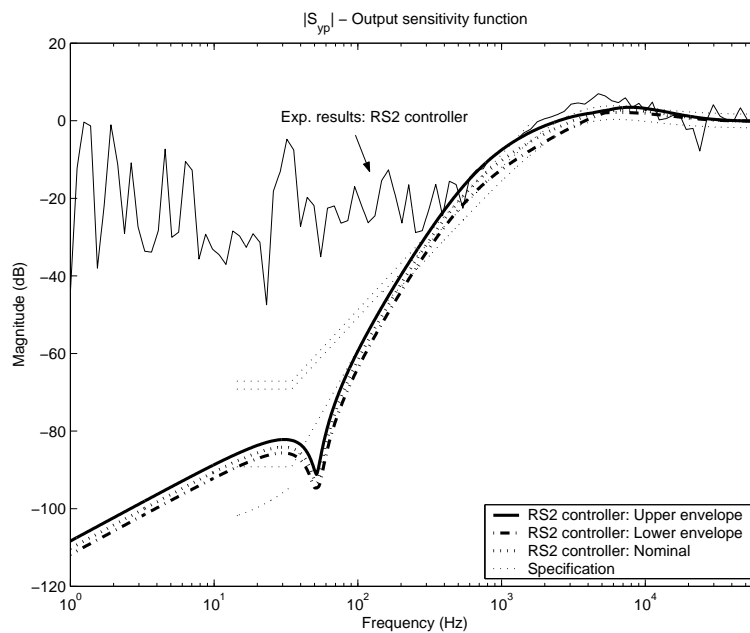


Figure 4.23: Magnitude of the output sensitivity function obtained from measurements and the envelopes of output sensitivity functions from simulation for K_{RS2} , radial loop.

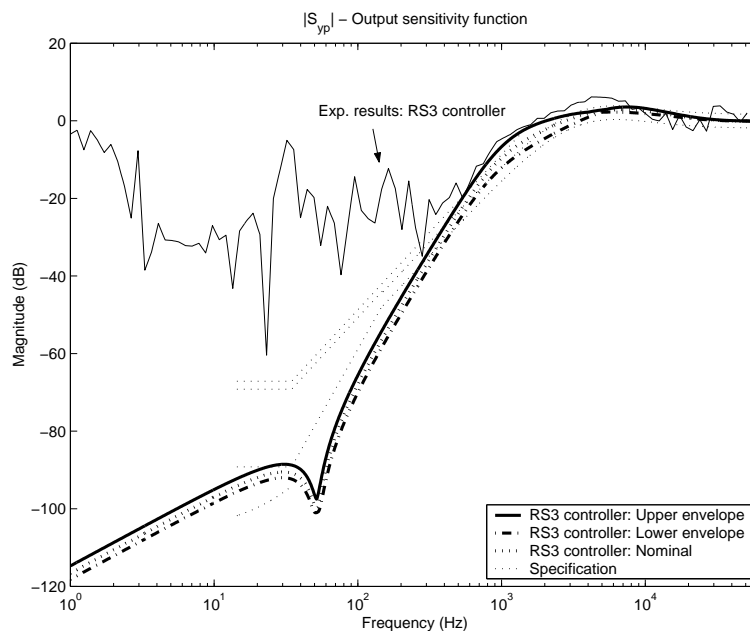


Figure 4.24: Magnitude of the output sensitivity function obtained from measurements and the envelopes of output sensitivity functions from simulation for K_{RS3} , radial loop.

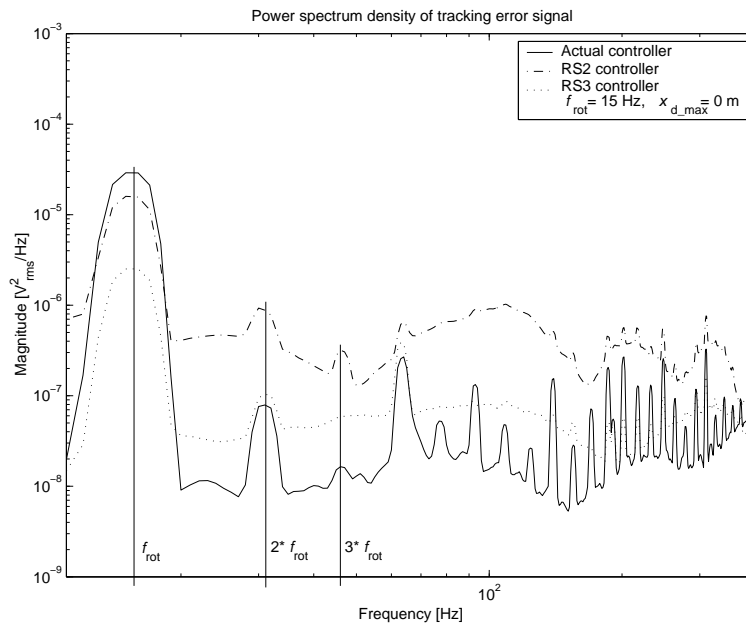


Figure 4.25: The measured power spectrum density of the radial error signal e_R for the 2nd/3rd order designed controllers (K_{RS2} , K_{RS3}) and actual implemented controller K_{act} . The tested disk has very small disk vertical deviation at the disk outer edge $z_{d,max} \approx 0 \mu\text{m}$ and very small disk eccentricity $x_{d,max} \approx 0 \mu\text{m}$, $f_{rot} = 15 \text{ Hz}$.

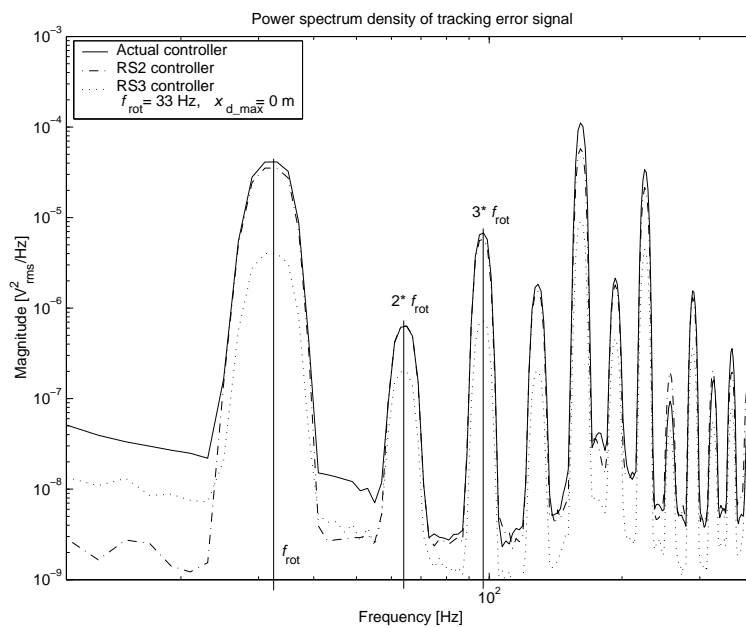


Figure 4.26: The same as in fig. 4.25 but for $f_{rot} = 33 \text{ Hz}$.

4.7 Conclusions

In this chapter we have presented a model-based control design methodology applied to an industrial DVD-video player, to enhance track following (in the focus and radial directions) and improve control performances with respect to different repetitive disturbances and manufacturing tolerances of optical pick-up units.

We have shown that repetitive disturbances are mainly due to the geometry of the disk. For these reasons, the simple models of the repetitive disturbances generation are given before control system design.

The combined pole placement/sensitivity function shaping control methodology, under industrial performance specifications, followed by the controller order reduction have been used to design a robust SISO controllers to minimize the laser spot position errors in the focus and radial control loop. Modifications on the output sensitivity template have been explained and directly fulfill in controllers design to obtain improvement in terms of repetitive disturbances. Control order reduction is performed to make the digital implementation feasible on the industrial benchmark.

An uncertainty model set, based on a parametric description, is considered to analyze how the variations of the plant physical parameters influence performances and robustness of the achieved solution. Simple robustness analysis point out that the achieved controllers remain stable for large uncertainty of the actuator physical parameters, and that performance specifications are met.

Final comparison of the actual and designed controllers illustrates that new controllers provide better system performances and robustness than the actual controllers. It is clear from these results that we have obtained the limit of the controller order reduction for these imposed performances and implementation constraints.

This methodology can be applied to other families of DVD players.

5.1 Review

Three reasons are the main motivation of this thesis:

1. Develop and analyze the properties of a closed-loop identification method for state-space models with a given structure.
2. Find optical theory in order to develop physical models of the position error signals generation in DVD-player optical device.
3. Design advanced automatic control method in order to develop model-based control design for an industrial optical disk drive system.

5.2 Contributions of Presented Thesis

In the context of the structure outlined above, the following points have been elaborated in this work:

Part I - Identification

- Several novel recursive parameter adaptation algorithms have been presented to identify various state-space discrete-time plant models with a given structure (parametrization). They belong to the class of output-error algorithms, and can be interpreted as a recursive pseudolinear regression. The RLS-2 algorithm incorporates a numerical optimization component represented by computation of a matrix pseudoinverse. On the other hand, the IGM-1, IGM-2 and GM-2 algorithms, which all use the gradient technique to minimize an one-step quadratic criterion, have the advantage of a closed analytic form.
- In general, the RLS-2 algorithm possesses the fastest convergence. The strength of the IGM-1 and IGM-2 algorithms lies in robustness, *i.e.* in their insensitivity to the realization of the disturbance noise and to various initial settings (initial parameter vector estimate, adaptation gain settings). One can conclude from the verification that RLS-2 algorithm offers the best properties from these algorithms.
- Sufficient conditions for stability in a deterministic environment and convergence in a stochastic environment are always related to the actual state-space model and its parametrization. Only the particular case of parametrization, where the transformation function is linear towards the parameter vector, has been analyzed in more detail. The properties of the algorithms have been shown for this parametrization and they are related to a positive real condition on a sensitivity-type function. This

condition can be relaxed by data filtering or adding a proportional adaptation. It can be concluded that the canonical state-space representation is the best theoretical case.

- It has been shown that developed algorithms give on one hand a bias distribution which is not influenced by noise and, on the other hand, contain an implicit frequency weighting filter which is matched with a robust performance control criterion. These properties make these algorithms a suitable tool for control relevant identification.
- The new theoretical part of developed algorithms and the experiment for SISO and MIMO systems were discussed in chapter 2 and appendix B. The results of this study have been presented in [BBVH03] and [HBVBP04], respectively.

Part II - DVD player: Modelling and Control

Modelling the focus error, radial error and read-out signals generation (Appendix C)

- A new approach has been developed to model the DVD players properties. An analytical and a numerical model of the focus error signal generation have been developed, based on the astigmatic method and the opto-geometrical analysis.
- To estimate the unknown model parameters, a curve fitting method is applied, using measured data from an industrial DVD-video player.
- The performance of the analytical model is inferior in comparison with the numerical model but this model is important because it is a *complete analytical model*. This advantage is most important from the identification point of view where the non-measurable parameters (or hardly measurable ones) could be estimated much faster than in the case of the numerical model. Then, the analytical model is more useful in the focus closed loop of the DVD player. Moreover, any analytical model of the S-curve $e_F(\Delta z)$ significantly saves simulation time.
- Comparison with the real data, acquired from an industrial DVD-video player during start-up procedure, illustrates the quality of the analytical model.
- The theoretical part of the developed models and the experimental results were discussed in appendix C. The results of this study have been presented in [HBVSdF02], [HBVSd03] and [HVSd04].
- The model of the radial error signal generation has been presented in appendix D and the model of the read-out signal generation has been presented in appendix E. The presented models are useful for testing the DVD player functionality and future improvements of this system.

Control system design

- A new designed control system based on a combined pole placement/sensitivity function shaping methodology is proposed for construction of a low-complexity controller, being able to achieve an enhanced focus/radial following performance and repetitive disturbance rejection.
- Controller order reduction is performed to allow its practical implementation.
- An uncertainty model set, based on a parametric description, is considered to analyze how the variations of the plant physical parameters influence performance and robustness of the achieved solution. A simple robustness analysis demonstrates that the designed closed-loops remain stable for large uncertainty of the physical actuator parameters and the performance specifications are met.
- Final comparison of the existing and designed controllers illustrates that new controllers provide better system performance and robustness than the existing controllers. It is clear from these results that we have obtained the limit of the controller order reduction for the imposed performance and implementation constraints.
- This control design methodology can be applied to the other families of DVD players (DVD-ROM, DVD-Audio, DVD-Recordable) and a next generation, high-capacity, blue-laser DVD format.
- The control system design and the experiment results were shown in chapter 4. The results of this study have been published in [HBVFSd03a], [HBVFSd03b], [HVP03], [HBVF04] and [HVFS04].

This thesis fulfilled the goals traced out in the introduction. It can be considered that the chapters 2, C and 4 create the core of this thesis.

5.3 Perspectives

Part I - Identification

In this thesis, has been assumed the case of the parametrization, where the transformation function from state-space to the input-output transfer operator is linear towards the parameterization vector. Furthermore, the proposed algorithms have been extended for MIMO systems. Nevertheless, the parametrization function is given in a very general way. Therefore, a development of a methodology for building a convenient parametrization of state-space models can be of further interest for research.

Moreover, the estimation of physical parameters via identification of a discrete-time state-space model requires the nontrivial transformation from a continuous-time model to a discrete-time model. For this reason, more sophisticated optimization algorithms to minimize a non-linear criterion function of its parameters should be used.

Part II - DVD player: Modelling and Control

Modelling the focus error, radial error and read-out signals generation

The proposed models make a compromise between complexity and accuracy. Full analysis of propagation in terms of Gaussian beams can be carried out analytically to improve the accuracy of proposed models near the focus position.

Interesting aspects related to diffraction from the optical disk layer(s), like radial tracking to focus cross-talk and signal perturbation due to more than one information layer can be studied. Nevertheless, understanding of complicated diffraction theory and an access to a simulation environment are necessary.

Finally, possible higher order aberrations of the astigmatic lens and spherical lenses used in the optical pick-up unit and their optical tolerances can be included in more complicated models, based on diffraction theory. Consequently, a compromise between model complexity and accuracy has to be done because of increase of the simulation time-complexity.

Control system design

This thesis has proved that model-based control design applied to a physical system requires understanding:

1. The disturbances affecting the DVD systems.
2. The model uncertainty for evaluating the system robust performance.
3. The need of a systematic methodology useful to perform robust controllers design.

We have considered in this work the uncertainty of inaccurate modelling of an optical disk drive electro-mechanical actuator.

However, the system focus/radial performance is also determined by the limited accuracy of the implemented controller which operates in the feedback loop. For this reason, control design methodology should be extended in order to take also into account these restrictions.

Another problem, often encountered in consumer optical disk drives mass production is the variability of dynamical behavior, due to variation of environmental conditions and system aging. There is a clear need for models that accurately describe this variable dynamical behavior for a *large* number of systems in view of an enhanced robust focus/radial performance and disturbance rejection control design.

In this work, we have used a model uncertainty set based on technical specifications of one industrial DVD-video optical pick-up unit to analyze how the variations of the plant physical parameters influence the performance and the robustness of the achieved solution. However, the utility of this approach to achieve a *large* number of optical pick-up units demands a more detailed analysis of parametric uncertainty acting on DVD-video players, and it requires a thorough investigation of alternative modelling techniques.

Closed-loop Output-Error Identification Method for SISO Systems

A

A.1 Introduction

This chapter is devoted to the description of the closed-loop output-error identification method for SISO systems.

The aim of this chapter is to describe the closed-loop output-error identification recursive method and to present its properties. It makes more easier to understand the identifications algorithms for state-space systems that are presented in chapter 2.

In practice, many plants cannot be easily operated in open loop for carrying an experimentation protocol for identification (*e.g.* integrator behavior, drift). In some situations, a controller may already exist and there is no reason to open the loop when the object of the identification will be to get a better model for either design a new controller or re-tuning the existing controller.

In adaptive control, one also has to estimate the parameters of the plant model in closed-loop operation. In the context of adaptive control, the study of recursive plant model identification in closed-loop will emphasize on the one hand the importance of the filters to be used with open loop recursive identification algorithms, [LLM97; LK97b], and, on the other hand, the use of specific algorithms dedicated to recursive identification in closed-loop.

The problem of plant model identification in closed-loop is also of crucial importance in the context of the iterative combination of the identification in closed-loop and robust control re-design, see [Lju99; LLM97; LK97b]. This can be viewed as an adaptive control scheme in which data acquisition and parameter identification is done at sampling period, but the up-dating of the controller is done only on longer time horizons.

Figure A.1 illustrates the basis of the closed-loop identification iterative procedure (for the case of an RST digital controller). The upper part represents the true closed-loop system and the lower part represents the design system. The objective is to minimize the error between the two systems by using new data acquired in closed-loop operation. In adaptive control, the idea is first to improve the plant model estimation, and then to re-design the controller based on the new model. This sequence of operations is carried out one or several times. However, a key point is that the new plant model estimation should be done in order to reduce the error between the two systems. In fact, the objective of the plant model identification in closed-loop is to get a better predictor for the closed-loop via a better estimation of the plant model.

The problem of plant model identification in closed-loop can be viewed in two different ways:

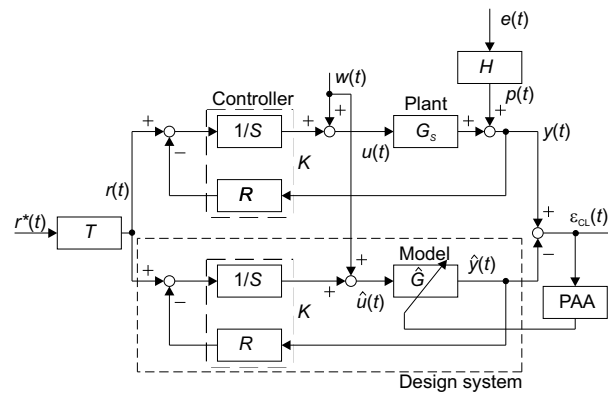


Figure A.1: Closed-loop identification scheme for SISO systems.

- *Model Reference Adaptive Control (MRAS) point of view.* The true closed-loop system corresponds to a reference model, and a parallel adjustable system having a feedback configuration is built up. This adjustable feedback system contains a fixed controller and an adjustable model of the plant.
- *Identification point of view.* The aim is to construct an adjustable predictor for the closed-loop system in terms of a known fixed controller and of an adjustable plant model \hat{G} .

The external excitation can either be applied on the reference $r^*(t)$ or added to the output of the controller $w(t)$, see fig. A.1. This does not change the structure of the algorithm; it just affects the external quantity added to the input of the plant and of the estimated plant model.

From the basic scheme in fig. A.1, the family of closed-loop output-error recursive algorithms can be derived, in which the parameters of the estimated model are driven by a Parameter Adaptation Algorithm (PAA) using the closed-loop prediction error ε_{CL} as adaptation error.

This chapter is organized as follows: The overview on the basic equations is in section A.2 and the family of the closed-loop out-put error algorithms is presented in section A.3.

Parameter adaptation algorithms are discussed in section A.4 and their properties are presented in section A.5. The stability of PAA in deterministic environment, parametric convergence analysis in a stochastic environment and frequency distribution of the asymptotic bias are discussed here.

Conclusions are presented in section A.6.

A.2 Basic Equations of Plant, Fixed Predictor and Output-Error

The objective is to estimate the parameters of the plant model defined by the transfer operator:

$$G_S(q^{-1}) = \frac{q^{-d}B(q^{-1})}{A(q^{-1})}, \quad (\text{A.1})$$

where:

$$B(q^{-1}) = b_1q^{-1} + \dots + b_{n_B}q^{-n_B} = q^{-1}B^*(q^{-1}), \quad (\text{A.2})$$

$$A(q^{-1}) = 1 + a_1q^{-1} + \dots + a_{n_A}q^{-n_A} = 1 + q^{-1}A^*(q^{-1}). \quad (\text{A.3})$$

In these formulae, we introduce the *forward shift operator* q which is defined by:

$$qu(k) = u(k+1),$$

and the *backward shift operator* q^{-1} which is defined by:

$$q^{-1}u(k) = u(k-1),$$

where T_s is the sampling period and $f_s = 1/T_s$ is the sampling frequency. One should note that q and q^{-1} operate in time domain, therefore they cannot be confused with \mathcal{Z} -transform operators z and z^{-1} .

The output of the **plant** can be expressed from (A.1) using (A.2) and (A.3):

$$\begin{aligned} A(q^{-1})[y(k+1) - p(k+1)] &= q^{-d}B(q^{-1})u(k+1) \\ y(k+1) &= -A(q^{-1})^*y(k) + B(q^{-1})^*u(k-d) + A(q^{-1})p(k+1), \end{aligned} \quad (\text{A.4})$$

where $u(k)$ is the plant input, $y(k)$ is the plant output and $p(k)$ is the output disturbance noise given by:

$$p(k) = H(q^{-1})e(k), \quad (\text{A.5})$$

where $e(k)$ is a zero-mean Gaussian white noise and $H(q^{-1})$ is the transfer operator of the noise model, as it is illustrated in fig. A.1. The equation (A.4) can be rewritten to the following form:

$$y(k+1) = \theta^T \varphi(k) + A(q^{-1})p(k+1), \quad (\text{A.6})$$

where θ^T denotes the true parameter vector:

$$\theta^T = [a_1 \dots a_{n_A} \quad b_1 \dots b_{n_B}], \quad (\text{A.7})$$

and $\varphi(k)$ the plant *regressor vector*:

$$\varphi(k) = [-y(k) \dots -y(k-n_A+1) \quad u(k-d) \dots u(k-n_B+1-d)]^T. \quad (\text{A.8})$$

The plant is operated in closed-loop with an RST digital controller which is given by three polynomials $R = R(q^{-1})$, $S = S(q^{-1})$ and $T = T(q^{-1})$. In this case, the plant input is expressed as follows:

$$u(k) = -\frac{R}{S}y(k) + w(k), \quad (\text{A.9})$$

where $w(k)$ is the equivalent external excitation added to the output of the controller either directly, either indirectly if the external excitation is applied on the reference $r^*(k)$, in which case $w(k) = \frac{T}{S}r^*(k)$.

If the external excitation $w(k)$ is added directly to the output of the controller, it will be filtered by the output sensitivity function $S_{yp}(q^{-1}) = \frac{y(t)}{p(t)} = \frac{1}{G_S K}$ which will enhance

the plant input signal in the frequency region where the output sensitivity function has its maximum. As a consequence, a more precise model will be obtained in the frequency regions which are critical for control design [HGdB96; LLM97; dH98; FL99].

For fixed values of the estimated parameters $\hat{\theta}^T = [\hat{a}_1 \dots \hat{a}_{n_A} \quad \hat{b}_1 \dots \hat{b}_{n_B}]$, the **predictor** of the closed-loop is described by:

$$\hat{y}(k+1) = -\hat{A}(q^{-1})^* \hat{y}(k) + \hat{B}(q^{-1})^* \hat{u}(k-d), \quad (\text{A.10})$$

where $\hat{u}(k)$ is the model input and $\hat{y}(k)$ is the model output, see fig. A.1. One can also write:

$$\hat{y}(k+1) = \hat{\theta}^T \phi(k, \hat{\theta}), \quad (\text{A.11})$$

where $\hat{\theta}^T$ denotes the estimated parameter vector:

$$\hat{\theta}^T = [\hat{a}_1 \dots \hat{a}_{n_A} \quad \hat{b}_1 \dots \hat{b}_{n_B}], \quad (\text{A.12})$$

and $\phi(k, \hat{\theta})$ is the *predictor regressor vector*:

$$\phi(k, \hat{\theta}) = [-\hat{y}(k) \dots -\hat{y}(k-n_A+1) \quad \hat{u}(k-d) \dots \hat{u}(k-n_B+1-d)]^T. \quad (\text{A.13})$$

The model input is given by:

$$\hat{u}(k) = -\frac{R}{S} \hat{y}(k) + w(k). \quad (\text{A.14})$$

One can compare the equations (A.4)–(A.9) describing the **plant** branch with the equations (A.10)–(A.14) describing the **model** branch. Finally, the closed-loop prediction **output-error** is defined as:

$$\varepsilon_{\text{CL}}(k+1) = y(k+1) - \hat{y}(k+1). \quad (\text{A.15})$$

It should be mentioned that, as clearly results from fig. A.1, for constant values of the estimated parameters the predictor regressor vector $\phi(k)$ depends only on the external excitation. Therefore, under the assumption that the external excitation $r^*(k)$ or $w(k)$ is independent from the stochastic disturbance $p(k)$, $\phi(k)$ and $p(k)$ are not correlated.

We may note that the previous predictions $\hat{y}(k-i)$ are used rather than the observed outputs in the *predictor regressor vector* $\phi(k)$. This is in accordance with fig. A.1, and motivates the name “output error method” or “model reference method”. As a consequence, the *predictor regressor vector* $\phi(k)$ does not directly depend on the observed variables $y(k)$ and hence not on any output disturbance noise $p(k)$ added to the measurements. This also means, as it is shown in section 2.4.1, that the convergence properties of parameter adaptation algorithm using (closed-loop) output error predictor is less sensitive to the properties of $p(k)$ than the usual recursive least square predictor (known also as *equation error*) algorithm, see [LS83; Lju99; LLM97].

Output error algorithm can also be interpreted as a pseudolinear regression. Since, in (A.13), the sequence $\hat{y}(k-i)$ is known then (A.11) is a pseudolinear regression for this model. To simplify notation, *backward shift operator* q^{-1} will be omitted. In the sequel $\hat{\theta}$ is also omitted in pseudo terms whenever there is no risk of confusion.

A.3 The Family of Closed-loop Output-Error Methods

One can easily prove the relationship:

$$\theta^T \varphi(k) = \theta^T \phi(k) - \left(A^* + \frac{q^{-d} B^* R}{S} \right) \varepsilon_{\text{CL}}(k), \quad (\text{A.16})$$

which enables to rewrite (A.6) as:

$$y(k+1) = \theta^T \phi(k) - \left(A^* + \frac{q^{-d} B^* R}{S} \right) \varepsilon_{\text{CL}}(k) + Ap(k+1). \quad (\text{A.17})$$

Subtracting (A.11) from (A.17), one obtains:

$$\varepsilon_{\text{CL}}(k+1) = (\theta - \hat{\theta})^T \phi(k) - \left(A^* + \frac{q^{-d} B^* R}{S} \right) \varepsilon_{\text{CL}}(k) + Ap(k+1). \quad (\text{A.18})$$

Furthermore, one can write:

$$1 + q^{-1} \left(A^* + \frac{q^{-d} B^* R}{S} \right) = \frac{AS + q^{-d} BR}{S} = \frac{P}{S}, \quad (\text{A.19})$$

where $P = AS + q^{-d} BR$ defines the poles of the true closed-loop system. From (A.19) it results that:

$$A^* + \frac{q^{-d} B^* R}{S} = q \left(\frac{P}{S} - 1 \right). \quad (\text{A.20})$$

Introducing (A.20) into (A.18), one finally gets:

$$\varepsilon_{\text{CL}}(k+1) = \frac{S}{P} (\theta - \hat{\theta})^T \phi(k) + \frac{AS}{P} p(k+1), \quad (\text{A.21})$$

and in the deterministic case ($p(k+1) \equiv 0$):

$$\varepsilon_{\text{CL}}(k+1) = \frac{S}{P} (\theta - \hat{\theta})^T \phi(k). \quad (\text{A.22})$$

The closed-loop prediction error written as (A.21) or (A.22) is the main player of the identification methods from the family of the closed-loop output-error (CLOE).

In the following subsections, different versions of the closed-loop prediction error ε_{CL} will be presented, in the case of adjustable predictors.

CLOE Method

Now replacing the fixed predictor \hat{y} of the closed-loop (A.11) by an adjustable predictor, one obtains *a priori* predicted output which can be computed:

$$\hat{y}^\circ(k+1) = \hat{y} [k+1 | \hat{\theta}(k)] = \hat{\theta}^T(k) \phi(k), \quad (\text{A.23})$$

and *a posteriori* predicted output which can be computed only after deriving $\hat{\theta}(k+1)$ as:

$$\hat{y}(k+1) = \hat{y} [k+1 | \hat{\theta}(k+1)] = \hat{\theta}^T(k+1) \phi(k). \quad (\text{A.24})$$

Defining the *a priori* prediction error which can be computed just after the acquisition of the new measurement $y(k+1)$ (before deriving $\hat{\theta}(k+1)$) as:

$$\varepsilon_{\text{CL}}^{\circ}(k+1) = y(k+1) - \hat{y}^{\circ}(k+1), \quad (\text{A.25})$$

and the *a posteriori* prediction error which can be computed only after deriving $\hat{\theta}(k+1)$ as:

$$\varepsilon_{\text{CL}}(k+1) = y(k+1) - \hat{y}(k+1), \quad (\text{A.26})$$

and using (A.17) and (A.24), the equation for the *a posteriori* prediction error becomes in the deterministic case:

$$\varepsilon_{\text{CL}}(k+1) = \frac{S}{P} \left[\theta - \hat{\theta}(k+1) \right]^{\text{T}} \phi(k). \quad (\text{A.27})$$

In the stochastic case, one obtain:

$$\varepsilon_{\text{CL}}(k+1) = \frac{S}{P} \left[\theta - \hat{\theta}(k+1) \right]^{\text{T}} \phi(k) + Ap(k+1). \quad (\text{A.28})$$

This is termed the *closed-loop output-error* method (CLOE).

F-CLOE and AF-CLOE Methods

However, (A.22) can also be rewritten as:

$$\varepsilon_{\text{CL}}(k+1) = \frac{S}{P} \frac{\hat{P}}{S} (\theta - \hat{\theta})^{\text{T}} \frac{S}{\hat{P}} \phi(k) = \frac{\hat{P}}{P} (\theta - \hat{\theta})^{\text{T}} \phi_{\text{f}}(k), \quad (\text{A.29})$$

where:

$$\phi_{\text{f}}(k) = \frac{S}{\hat{P}} \phi(k), \quad (\text{A.30})$$

$$\hat{P} = \hat{A}S + q^{-d}\hat{B}R. \quad (\text{A.31})$$

In (A.31) \hat{P} is an estimation of the true closed-loop poles based on an initial estimation of the plant model \hat{A} , \hat{B} (for example using an open-loop experiment). This formulation leads to the *filtered closed-loop output-error* method (F-CLOE) which uses the same adjustable predictor as CLOE (see (A.23) and (A.24)). An advantage of the F-CLOE method is that it gives the minimum value of the uncorrelation test in comparison with other methods, see [LLM97].

If one replaces $\hat{P}(q^{-1})$ by $\hat{P}(q^{-1}, k)$ (the current estimate of the closed-loop characteristic polynomial) in the data filter of the F-CLOE method, the *adaptive filtered closed-loop output-error* method (AF-CLOE) is obtained.

X-CLOE Method

Let's consider the following disturbance model now:

$$p(k+1) = H_e e(k+1) = \frac{C}{A} e(k+1), \quad (\text{A.32})$$

where $e(k+1)$ is a zero-mean Gaussian white noise, H_e is the transfer operator of the noise model and:

$$C(q^{-1}) = 1 + q^{-1}C^*(q^{-1}). \quad (\text{A.33})$$

An extended output-error prediction model can be defined:

$$\hat{y}(k+1) = -\hat{A}^*\hat{y}(k) + \hat{B}^*\hat{u}(k-d) + \hat{H}_e^* \frac{\varepsilon_{\text{CL}}(k)}{S}, \quad (\text{A.34})$$

$$\hat{y}(k+1) = \hat{\theta}^T \phi(k) + \hat{H}_e^* \frac{\varepsilon_{\text{CL}}(k)}{S} = \hat{\theta}_e^T \phi_e(k), \quad (\text{A.35})$$

where:

$$H_e^* = h_1 + h_2q^{-1} + \dots + h_{n_H}q^{-n_H+1} = C^*S - A^*S - q^{-d}B^*R, \quad (\text{A.36})$$

$$H_e = 1 + q^{-1}H_e^* = 1 + CS - P. \quad (\text{A.37})$$

The equation (A.17) for the plant output becomes in this case:

$$y(k+1) = \theta^T \phi(k) + H_e^* \frac{\varepsilon_{\text{CL}}(k)}{S} - C^* \varepsilon_{\text{CL}}(k) + Ce(k+1). \quad (\text{A.38})$$

Subtracting (A.35) from (A.38), one obtains the expression for the closed-loop prediction error:

$$\varepsilon_{\text{CL}}(k+1) = \frac{1}{C}(\theta_e - \hat{\theta}_e)^T \phi_e(k) + e(k+1), \quad (\text{A.39})$$

where the extended parameter vectors are defined:

$$\theta_e^T = [\theta^T \quad h_1 \quad \dots \quad h_{n_H}], \quad (\text{A.40})$$

$$\hat{\theta}_e^T = [\hat{\theta}^T \quad \hat{h}_1 \quad \dots \quad \hat{h}_{n_H}], \quad (\text{A.41})$$

and the extended predictor regressor vector:

$$\phi_e(k) = [\phi^T(k) \quad \varepsilon_{\text{CLf}}(k) \quad \dots \quad \varepsilon_{\text{CLf}}(k - n_H + 1)]^T, \quad (\text{A.42})$$

$$\varepsilon_{\text{CLf}}(k) = \frac{1}{S} \varepsilon_{\text{CL}}(k). \quad (\text{A.43})$$

The equation (A.39) clearly shows that for $\hat{\theta}_e = \theta_e$ the closed-loop prediction error tends asymptotically towards $e(k+1)$.

Replacing the fixed predictor (A.35) by an adjustable one, the *extended closed-loop output-error* method (X-CLOE) can be obtained. Thus, the *a priori* prediction error is given by:

$$\varepsilon_{\text{CL}}^\circ(k+1) = y(k+1) - \hat{\theta}_e^T(k) \phi_e(k) \quad (\text{A.44})$$

and the equation for the *a posteriori* prediction error becomes:

$$\varepsilon_{\text{CL}}(k+1) = y(k+1) - \hat{\theta}_e^T(k+1) \phi_e(k). \quad (\text{A.45})$$

G-CLOE Method

For the case of narrow-band disturbances, it is interesting to consider a disturbance model of the form:

$$p(k+1) = H_g e(k+1) = \frac{C}{DA} e(k+1), \quad (\text{A.46})$$

where $e(k+1)$ is a zero-mean Gaussian white noise, H_g is the transfer operator of the noise model and:

$$C(q^{-1}) = 1 + q^{-1}C^*(q^{-1}), \quad (\text{A.47})$$

$$D(q^{-1}) = 1 + q^{-1}D^*(q^{-1}) = 1 + d_1q^{-1} + \dots + d_{n_D}q^{-n_D+1}. \quad (\text{A.48})$$

The same procedure as for X-CLOE can be used for deriving a recursive parameter estimation method. Briefly, the asymptotically optimal predictor takes the form:

$$\hat{y}(k+1) = -\hat{A}^*\hat{y}(k) + \hat{B}^*\hat{u}(k-d) - \hat{D}^*\alpha_g(k) + \hat{H}_g^* \frac{\varepsilon_{\text{CL}}(k)}{S}, \quad (\text{A.49})$$

where:

$$\alpha_g(k) = \hat{A}\hat{y}(k) - \hat{B}\hat{u}(k-d), \quad (\text{A.50})$$

$$H_g^* = h_1 + h_2q^{-1} + \dots + h_{n_H}q^{-n_H+1} \quad (\text{A.51})$$

$$= C^*S - A^*S - q^{-d}B^*R - D^*AS - D^*q^{-d}BR, \quad (\text{A.52})$$

$$H_g = 1 + q^{-1}H_g^* = 1 + CS - DP. \quad (\text{A.53})$$

In the deterministic case, the closed-loop prediction error is given by:

$$\varepsilon_{\text{CL}}(k+1) = (\theta_{g0} - \hat{\theta}_g)^T \phi_g(k), \quad (\text{A.54})$$

where:

$$\theta_{g0}^T = [\theta^T \quad h_1 \quad \dots \quad h_{n_H} \quad d_1 \quad \dots \quad d_{n_D}], \quad (\text{A.55})$$

$$\hat{\theta}_g^T = [\hat{\theta}^T \quad \hat{h}_1 \quad \dots \quad \hat{h}_{n_H} \quad \hat{d}_1 \quad \dots \quad \hat{d}_{n_D}], \quad (\text{A.56})$$

$$\phi_g(k) = \left[\begin{array}{c} \phi^T(k) \quad \varepsilon_{\text{CLf}}(k) \quad \dots \quad \varepsilon_{\text{CLf}}(k - n_H + 1) \\ -\alpha_g(k) \quad \dots \quad -\alpha_g(k - n_D + 1) \end{array} \right]^T, \quad (\text{A.57})$$

$$H_{g0}^* = h_1 + h_2q^{-1} + \dots + h_{n_H}q^{-n_H+1} = -A^*S - B^*R, \quad (\text{A.58})$$

$$\varepsilon_{\text{CLf}}(k) = \frac{1}{S} \varepsilon_{\text{CL}}(k). \quad (\text{A.59})$$

Taking an adjustable predictor instead of the fixed one, these expressions lead to the *generalized closed-loop output-error* method (G-CLOE) which uses the following definition of the *a priori* prediction error:

$$\varepsilon_{\text{CL}}^{\circ}(k+1) = y(k+1) - \hat{\theta}_g^T(k) \phi_g(k) \quad (\text{A.60})$$

and the equation for the *a posteriori* prediction error becomes:

$$\varepsilon_{\text{CL}}(k+1) = y(k+1) - \hat{\theta}_g^T(k+1) \phi_g(k) \quad (\text{A.61})$$

See [LK97b] for more detailed information about the G-CLOE method. Finally, a summary of presented closed-loop output-error methods is shown in table A.1.

Table A.1: Summary of closed-loop output-error methods.

	Noise model $p(k)$	Parameter vector $\hat{\theta}$	Observation $\phi(k)$	Adaptation error $\varepsilon_{\text{CL}}(k+1)$
CLOE	$e(k)$	$\hat{\theta}$	$\phi(k)$	$y(k+1) - \hat{\theta}^T \phi(k)$
F-CLOE	$e(k)$	$\hat{\theta}$	$\phi_f(k)$	$y(k+1) - \hat{\theta}^T \phi(k)$
X-CLOE	$\frac{C}{A}e(k)$	$\hat{\theta}_e$	$\phi_e(k)$	$y(k+1) - \hat{\theta}_e^T \phi_e(k)$
G-CLOE	$\frac{C}{AD}e(k)$	$\hat{\theta}_g$	$\phi_g(k)$	$y(k+1) - \hat{\theta}_g^T \phi_g(k)$

A.4 Parameter Adaptation Algorithms

In this section, we will focus on the parameter adaptation algorithms that are used in the closed-loop identification scheme.

The key element for implementing the on-line estimation of the plant model parameters is the *parameter adaptation algorithm* (PAA) which drives the parameters of the adjustable predictor from the data acquired on the system at each sampling instant. This algorithm has a *recursive* structure, *i.e.* the new value of the estimated parameters is equal to the previous value plus a correcting term which depends on the most recent measurements.

Parameter adaptation algorithms generally have the following structure:

$$\begin{aligned} \begin{bmatrix} \textit{New estimated} \\ \textit{parameters} \\ \textit{(vector)} \end{bmatrix} &= \begin{bmatrix} \textit{Previous estimated} \\ \textit{parameters} \\ \textit{(vector)} \end{bmatrix} + \begin{bmatrix} \textit{Adaptation} \\ \textit{gain} \\ \textit{(matrix, scalar)} \end{bmatrix} \times \\ &\times \begin{bmatrix} \textit{Measurement} \\ \textit{function} \\ \textit{(vector)} \end{bmatrix} \times \begin{bmatrix} \textit{Prediction error} \\ \textit{function} \\ \textit{(scalar)} \end{bmatrix} \end{aligned}$$

This structure corresponds to the so-called *integral type adaptation algorithm*. The algorithm has memory, and therefore maintains the estimated value of the parameters when the correcting term becomes zero. The algorithm can be viewed as a discrete-time integrator fed by the correcting term at each instant.

The adaptation gain plays an important role in the performance of the parameter adaptation algorithm and it can be constant or time-varying. The measurement function is generally called the *observation vector*. The prediction error function is generally called the *adaptation error*.

In the case of the basic CLOE algorithm, the estimated parameter vector is denoted by $\hat{\theta}$, the observation vector corresponds to the predictor regressor vector $\phi(k)$ and the adaptation error corresponds to the closed-loop prediction error ε_{CL} . Accordingly, this notation is used in the following text. However, if a different closed-loop output-error algorithm (F-CLOE, AF-CLOE, X-CLOE or G-CLOE) is needed, the appropriate definition should be considered, as shown in table A.1.

Gradient Algorithm

The idea of the gradient parameter adaptation algorithm is based on the minimization of a quadratic criterion in terms of the *a priori* prediction error:

$$\min_{\hat{\theta}(k)} J(k+1) = \frac{1}{2} [\varepsilon_{\text{CL}}^{\circ}(k+1)]^2. \quad (\text{A.62})$$

A solution can be provided by the gradient technique. In order to minimize the value of the criterion, one moves in the opposite direction of the gradient of the criterion. Hence, the corresponding PAA has the form:

$$\hat{\theta}(k+1) = \hat{\theta}(k) - F \frac{\partial J(k+1)}{\partial \hat{\theta}(k)}, \quad (\text{A.63})$$

where F is the matrix adaptation gain and $\partial J(k+1)/\partial \hat{\theta}(k)$ is the gradient of the criterion at $k+1$ with respect to the estimated parameter vector at k . From (A.62) one obtains:

$$\frac{\partial J(k+1)}{\partial \hat{\theta}(k)} = \frac{\partial \varepsilon_{\text{CL}}^{\circ}(k+1)}{\partial \hat{\theta}(k)} \varepsilon_{\text{CL}}^{\circ}(k+1). \quad (\text{A.64})$$

In addition, the *a priori* prediction error is given by:

$$\varepsilon_{\text{CL}}^{\circ}(k+1) = y(k+1) - \hat{y}^{\circ}(k+1) = y(k+1) - \hat{\theta}^{\text{T}}(k)\phi(k), \quad (\text{A.65})$$

and therefore:

$$\frac{\partial \varepsilon_{\text{CL}}^{\circ}(k+1)}{\partial \hat{\theta}(k)} = -\phi(k) - \frac{\partial \phi(k)}{\partial \hat{\theta}(k)}^{\text{T}} \hat{\theta}(k), \quad (\text{A.66})$$

where $\partial \phi(k)/\partial \hat{\theta}(k)$ is the Jacobian matrix of the predictor regressor vector defined by:

$$\frac{\partial \phi(k)}{\partial \hat{\theta}(k)} = \begin{bmatrix} -\frac{\partial \hat{y}(k)}{\partial \hat{a}_1(k)} & \cdots & -\frac{\partial \hat{y}(k)}{\partial \hat{a}_{n_A}(k)} & -\frac{\partial \hat{y}(k)}{\partial \hat{b}_1(k)} & \cdots & -\frac{\partial \hat{y}(k)}{\partial \hat{b}_{n_B}(k)} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ -\frac{\partial \hat{y}(k-n_A+1)}{\partial \hat{a}_1(k)} & \cdots & -\frac{\partial \hat{y}(k-n_A+1)}{\partial \hat{a}_{n_A}(k)} & -\frac{\partial \hat{y}(k-n_A+1)}{\partial \hat{b}_1(k)} & \cdots & -\frac{\partial \hat{y}(k-n_A+1)}{\partial \hat{b}_{n_B}(k)} \\ \frac{\partial \hat{u}(k-d)}{\partial \hat{a}_1(k)} & \cdots & \frac{\partial \hat{u}(k-d)}{\partial \hat{a}_{n_A}(k)} & \frac{\partial \hat{u}(k-d)}{\partial \hat{b}_1(k)} & \cdots & \frac{\partial \hat{u}(k-d)}{\partial \hat{b}_{n_B}(k)} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial \hat{u}(k-n_B+1-d)}{\partial \hat{a}_1(k)} & \cdots & \frac{\partial \hat{u}(k-n_B+1-d)}{\partial \hat{a}_{n_A}(k)} & \frac{\partial \hat{u}(k-n_B+1-d)}{\partial \hat{b}_1(k)} & \cdots & \frac{\partial \hat{u}(k-n_B+1-d)}{\partial \hat{b}_{n_B}(k)} \end{bmatrix}. \quad (\text{A.67})$$

From (A.24) one can clearly see that $\hat{y}(k)$ is a function of $\hat{\theta}(k)$. Moreover, $\hat{u}(k-d)$ can also become a function of $\hat{\theta}(k)$ under certain circumstances (if d is zero and the relative order of R/S is zero—see (A.14)). Therefore, the Jacobian matrix $\partial \phi(k)/\partial \hat{\theta}(k)$ cannot be generally considered null in the closed-loop identification scheme.¹

Introducing (A.66) into (A.64), the parameter adaptation algorithm (A.63) becomes:

$$\hat{\theta}(k+1) = \hat{\theta}(k) + F \left(\phi(k) + \frac{\partial \phi(k)}{\partial \hat{\theta}(k)}^{\text{T}} \hat{\theta}(k) \right) \varepsilon_{\text{CL}}^{\circ}(k+1). \quad (\text{A.68})$$

There are two possible choices of the matrix adaptation gain F :

¹However, in the open-loop identification scheme the plant regressor vector $\varphi(k)$ is used, which obviously does not depend on the estimated parameter vector, see (A.8). Hence, the second term of the right hand side of (A.66) becomes $\partial \varphi(k)^{\text{T}}/\partial \hat{\theta}(k)$ which is null.

- $F = \alpha I$, where $\alpha > 0$ and I is the identity matrix,
- $F > 0$ (positive definite matrix).

The resulting algorithm (A.68) has an integral structure. Unfortunately, it presents instability risks. If the adaptation gain is large near the optimum, one can move away from this minimum instead of getting closer.

It is well known that this method is fairly inefficient, in particular when the iterates are getting close to the minimum. So-called Newton methods or quasi-Newton methods (see *e.g.* [LS83]) give a distinctly better results. In these variants the search direction is modified using the Hessian: $\frac{\partial^2 J(k+1, \hat{\theta}(k))}{\partial \hat{\theta}^2(k)}$, *i.e.* the second derivative matrix of criterion function $J(k+1, \hat{\theta}(k))$. These iteration algorithms give convergence in one step to the minimum of $J(k+1, \hat{\theta}(k))$, if this function is quadratic in θ . Therefore, close to the minimum, where a second-order approximation of $J(k+1, \hat{\theta}(k))$ describes well the least squares criterion $\sum_{i=0}^k \varepsilon_{\text{CL}}^2(i+1)$, the Newton method is very efficient. Far away from the minimum, the algorithm maybe inefficient or even diverge. Therefore, the Hessian is usually replaced by a guaranteed positive-definite approximation in order to secure a search direction that points “downhill”.

Our study, nevertheless, also uses the gradient technique to minimize a one-step quadratic criterion because the proposed algorithms take advantage of their closed analytical forms (see chapter 2).

Improved Gradient Algorithm

In order to assure the stability of the PAA for any value of the adaptation gain F , the same gradient approach is used but a different criterion is considered, using the *a posteriori* prediction error $\varepsilon_{\text{CL}}(k+1)$:

$$\min_{\hat{\theta}(k+1)} J(k+1) = \frac{1}{2} [\varepsilon_{\text{CL}}(k+1)]^2. \quad (\text{A.69})$$

The gradient of the criterion becomes:

$$\frac{\partial J(k+1)}{\partial \hat{\theta}(k+1)} = \frac{\partial \varepsilon_{\text{CL}}(k+1)}{\partial \hat{\theta}(k+1)} \varepsilon_{\text{CL}}(k+1), \quad (\text{A.70})$$

and since the *a posteriori* prediction error is given by:

$$\varepsilon_{\text{CL}}(k+1) = y(k+1) - \hat{y}(k+1) = y(k+1) - \hat{\theta}^T(k+1)\phi(k), \quad (\text{A.71})$$

one gets:

$$\frac{\partial \varepsilon_{\text{CL}}(k+1)}{\partial \hat{\theta}(k+1)} = -\phi(k). \quad (\text{A.72})$$

In this case, the Jacobian matrix $\partial \phi(k) / \partial \hat{\theta}(k+1)$ is null because all elements of the predictor regressor vector $\phi(k)$ are independent of the estimated parameter vector $\hat{\theta}(k+1)$.

Introducing (A.72) into (A.70), the parameter adaptation algorithm becomes:

$$\hat{\theta}(k+1) = \hat{\theta}(k) + F\phi(k)\varepsilon_{\text{CL}}(k+1). \quad (\text{A.73})$$

This algorithm depends on $\varepsilon_{\text{CL}}(k+1)$, which is a function of $\hat{\theta}(k+1)$. For implementing the algorithm, $\varepsilon_{\text{CL}}(k+1)$ must be expressed as a function of $\varepsilon_{\text{CL}}^{\circ}(k+1)$.

One can rewrite (A.71) as:

$$\varepsilon_{\text{CL}}(k+1) = y(k+1) - \hat{\theta}^{\text{T}}(k)\phi(k) - \left[\hat{\theta}(k+1) - \hat{\theta}(k) \right]^{\text{T}} \phi(k). \quad (\text{A.74})$$

The first two terms of the right hand side correspond to $\varepsilon_{\text{CL}}^{\circ}(k+1)$, and from (A.73) one obtains:

$$\hat{\theta}(k+1) - \hat{\theta}(k) = F\phi(k)\varepsilon_{\text{CL}}(k+1), \quad (\text{A.75})$$

which enables to rewrite (A.74) as:

$$\varepsilon_{\text{CL}}(k+1) = \varepsilon_{\text{CL}}^{\circ}(k+1) - \phi^{\text{T}}(k)F^{\text{T}}\phi(k)\varepsilon_{\text{CL}}(k+1), \quad (\text{A.76})$$

from which the desired relation between $\varepsilon_{\text{CL}}(k+1)$ and $\varepsilon_{\text{CL}}^{\circ}(k+1)$ is obtained:

$$\varepsilon_{\text{CL}}(k+1) = \frac{\varepsilon_{\text{CL}}^{\circ}(k+1)}{1 + \phi^{\text{T}}(k)F^{\text{T}}\phi(k)}, \quad (\text{A.77})$$

and the algorithm (A.73) becomes:

$$\boxed{\hat{\theta}(k+1) = \hat{\theta}(k) + \frac{F\phi(k)\varepsilon_{\text{CL}}^{\circ}(k+1)}{1 + \phi^{\text{T}}(k)F^{\text{T}}\phi(k)}}, \quad (\text{A.78})$$

which is a stable algorithm irrespective of the value of the matrix adaptation gain F (positive definite).

The division by $1 + \phi^{\text{T}}(k)F^{\text{T}}\phi(k)$ introduces a normalization which reduces the sensitivity of the algorithm with respect to F and $\phi(k)$.

Recursive Least Squares Algorithm

When using the gradient algorithm, one moves in the quickest decreasing direction of the criterion, with a step depending on F . However, the minimization of $\varepsilon_{\text{CL}}^2(k+1)$ at each step does not necessarily lead to the minimization of:

$$\sum_{i=0}^k \varepsilon_{\text{CL}}^2(i+1),$$

on a time horizon. If the adaptation gain F is not low enough, oscillations may occur around the minimum of the criterion (A.62) used in gradient algorithm. On the other hand, a high adaptation gain F is preferable in order to reach a satisfactory convergence speed at the beginning of iteration when the optimum of the gradient criterion (A.62) is far away. In fact, the least squares algorithm offers such a variation profile for the adaptation gain F .

The aim is to find a recursive adaptation algorithm which minimizes the *least squares* (LS) criterion:

$$\min_{\hat{\theta}(k)} J(k) = \sum_{i=1}^k \left[y(i) - \hat{\theta}^T(k) \phi(i-1) \right]^2. \quad (\text{A.79})$$

The term $\hat{\theta}^T(k) \phi(i-1)$ corresponds to $\hat{y}(i) = \hat{y}[i, \hat{\theta}(k)]$, the prediction of the output at instant i ($i \leq k$) based on the parameter estimate at instant k obtained using k measurements.

The value of $\hat{\theta}(k)$, which minimizes the criterion (A.79), is obtained by seeking the value that cancels $\partial J(k) / \partial \hat{\theta}(k)$:

$$\frac{\partial J(k)}{\partial \hat{\theta}(k)} = -2 \sum_{i=1}^k \left[y(i) - \hat{\theta}^T(k) \phi(i-1) \right] \phi(i-1) = 0. \quad (\text{A.80})$$

From (A.80), taking into account the relationship:

$$\left[\hat{\theta}^T(k) \phi(i-1) \right] \phi(i-1) = \phi(i-1) \phi^T(i-1) \hat{\theta}(k), \quad (\text{A.81})$$

one obtains:

$$\left[\sum_{i=1}^k \phi(i-1) \phi^T(i-1) \right] \hat{\theta}(k) = \sum_{i=1}^k y(i) \phi(i-1), \quad (\text{A.82})$$

and therefore, the estimation algorithm is:

$$\hat{\theta}(k) = \left[\sum_{i=1}^k \phi(i-1) \phi^T(i-1) \right]^{-1} \sum_{i=1}^k y(i) \phi(i-1) = F(k) \sum_{i=1}^k y(i) \phi(i-1), \quad (\text{A.83})$$

in which:

$$F(k)^{-1} = \sum_{i=1}^k \phi(i-1) \phi^T(i-1). \quad (\text{A.84})$$

It is assumed that the matrix $\sum_{i=1}^k \phi(i-1) \phi^T(i-1)$ is invertible, which corresponds to the so-called *excitation condition*. See [LLM97] for explanation.

This estimation algorithm does not have a recursive form. To obtain a recursive algorithm, the estimation of $\hat{\theta}(k+1)$ is considered:

$$\hat{\theta}(k+1) = F(k+1) \sum_{i=1}^{k+1} y(i) \phi(i-1), \quad (\text{A.85})$$

where:

$$F(k+1)^{-1} = \sum_{i=1}^{k+1} \phi(i-1) \phi^T(i-1) = F(k)^{-1} + \phi(k) \phi^T(k). \quad (\text{A.86})$$

From (A.85) one gets:

$$\hat{\theta}(k+1) = F(k+1) \left[\sum_{i=1}^k y(i) \phi(i-1) + y(k+1) \phi(k) \right]. \quad (\text{A.87})$$

Taking into account (A.83), one can rewrite (A.87) as:

$$\hat{\theta}(k+1) = F(k+1) \left[F(k)^{-1} \hat{\theta}(k) + y(k+1) \phi(k) \right]. \quad (\text{A.88})$$

From (A.86), after multiplying both sides by $\hat{\theta}(k)$, the following relation is obtained:

$$F(k)^{-1} \hat{\theta}(k) = F(k+1)^{-1} \hat{\theta}(k) - \phi(k) \phi^T(k) \hat{\theta}(k), \quad (\text{A.89})$$

and (A.88) becomes:

$$\hat{\theta}(k+1) = F(k+1) \left\{ F(k+1)^{-1} \hat{\theta}(k) + \phi(k) \left[y(k+1) - \hat{\theta}^T(k) \phi(k) \right] \right\}. \quad (\text{A.90})$$

Using the expression (A.65) for the *a priori* prediction error $\varepsilon_{\text{CL}}^{\circ}(k+1)$, the result is:

$$\hat{\theta}(k+1) = \hat{\theta}(k) + F(k+1) \phi(k) \varepsilon_{\text{CL}}^{\circ}(k+1). \quad (\text{A.91})$$

The adaptation algorithm (A.91) has a recursive form similar to the gradient algorithm, except that the matrix adaptation gain $F(k+1)$ is time-varying now. It corrects the gradient direction and the step length automatically. A recursive formula for $F(k+1)$ remains to be given from the recursive formula for $F(k+1)^{-1}$ given in (A.86). This is obtained using the following *matrix inversion lemma*.

Matrix Inversion Lemma: Let F be an $n \times n$ dimensional nonsingular matrix, R an $m \times m$ dimensional nonsingular matrix and H an $n \times m$ dimensional matrix of maximum rank. Then the following identity holds:

$$(F^{-1} + HR^{-1}H^T)^{-1} = F - FH(R + H^T FH)^{-1}H^T F. \quad (\text{A.92})$$

In our case, one selects $H = \phi(k)$, $R = 1$, $F = F(k)$ and the following recursive formula is obtained:

$$F(k+1) = F(k) - \frac{F(k) \phi(k) \phi^T(k) F(k)}{1 + \phi^T(k) F(k) \phi(k)}. \quad (\text{A.93})$$

Putting together the different equations, a basic formulation of the recursive least squares (RLS) parameter adaptation algorithm can be given:

$$\hat{\theta}(k+1) = \hat{\theta}(k) + F(k+1) \phi(k) \varepsilon_{\text{CL}}^{\circ}(k+1), \quad (\text{A.94})$$

$$F(k+1) = F(k) - \frac{F(k) \phi(k) \phi^T(k) F(k)}{1 + \phi^T(k) F(k) \phi(k)}, \quad (\text{A.95})$$

$$\varepsilon_{\text{CL}}^{\circ}(k+1) = y(k+1) - \hat{\theta}^T(k) \phi(k). \quad (\text{A.96})$$

An equivalent form of this algorithm is obtained by introducing the expression of $F(k+1)$ given by (A.95) in (A.94) and using equations (A.65) and (A.71). Then the equivalent form of the parameter adaptation algorithm for the recursive least squares is obtained:

$$\boxed{\begin{aligned} \hat{\theta}(k+1) &= \hat{\theta}(k) + F(k) \phi(k) \varepsilon_{\text{CL}}^{\circ}(k+1), \\ F(k+1)^{-1} &= F(k)^{-1} + \phi(k) \phi^T(k), \\ F(k+1) &= F(k) - \frac{F(k) \phi(k) \phi^T(k) F(k)}{1 + \phi^T(k) F(k) \phi(k)}, \\ \varepsilon_{\text{CL}}^{\circ}(k+1) &= \frac{y(k+1) - \hat{\theta}^T(k) \phi(k)}{1 + \phi^T(k) F(k) \phi(k)}. \end{aligned}} \quad (\text{A.97})$$

For the recursive least squares algorithm to be exactly equivalent in sense of criterion minimization to the non-recursive least squares algorithm, it would have to be started from a first estimation obtained at instant $k_0 = \dim \phi(k)$, since $F(k)^{-1}$ given by (A.84) becomes nonsingular for $k = k_0$. In practice, the algorithm is started up at $k = 0$ by choosing:

$$F(0) = \frac{1}{\delta} I = G_{\text{ini}} I, \quad (\text{A.98})$$

in which the typical value is $\delta = 0.001$ ($G_{\text{ini}} = 1000$). It can be observed in the expression of $F(k+1)$ given by (A.86) that the influence of this initial error decreases with the time. In this case one minimizes the following criterion:

$$\min_{\hat{\theta}(k)} J(k) = \sum_{i=1}^k \left[y(i) - \hat{\theta}^T(k) \phi(i-1) \right]^2 + \left[\theta - \hat{\theta}(0) \right]^T F(0)^{-1} \left[\theta - \hat{\theta}(0) \right]. \quad (\text{A.99})$$

Choice of the Adaptation Gain F

The RLS algorithm is an algorithm with a decreasing gain. This can be clearly seen if the estimation of a single parameter is considered. In this case, $F(k)$ and $\phi(k)$ are scalars, and (A.95) becomes:

$$F(k+1) = \frac{F(k)}{1 + \phi^2(k) F(k)} \leq F(k). \quad (\text{A.100})$$

In fact, the RLS algorithm gives less and less weight to the new prediction errors, and thus to the new measurements. Consequently, this type of variation of the adaptation gain is not suitable for the estimation of time-varying parameters, and therefore other variation profiles must be considered in this case.

The recursive formula for the inverse of the adaptation gain $F(k+1)^{-1}$ given by (A.86) can be generalized by introducing two weighting sequences $\lambda_1(k)$ and $\lambda_2(k)$:

$$F(k+1)^{-1} = \lambda_1(k) F(k)^{-1} + \lambda_2(k) \phi(k) \phi^T(k), \quad (\text{A.101})$$

$$0 < \lambda_1(k) \leq 1, \quad 0 \leq \lambda_2(k) < 2, \quad F(0) > 0.$$

Note that $\lambda_1(k)$ and $\lambda_2(k)$ have the opposite effect. $\lambda_1(k) < 1$ tends to increase the adaptation gain while $\lambda_2(k) > 0$ tends to decrease the adaptation gain.

Using the matrix inversion lemma (A.92), one obtains:

$$F(k+1) = \frac{1}{\lambda_1(k)} \left[F(k) - \frac{F(k) \phi(k) \phi^T(k) F(k)}{\frac{\lambda_1(k)}{\lambda_2(k)} + \phi^T(k) F(k) \phi(k)} \right]. \quad (\text{A.102})$$

Each choice of the sequences $\lambda_1(k)$ and $\lambda_2(k)$ corresponds to a variation profile of the adaptation gain. Several possible choices are presented now:

A1. *Decreasing (vanishing) gain.* In this case:

$$\lambda_1(k) = \lambda_1 = 1, \quad \lambda_2(k) = \lambda_2 = 1, \quad (\text{A.103})$$

and $F(k+1)$ is given by (A.95), which leads to a decreasing adaptation gain. This type of profile is suitable for the estimation of the parameters of stationary systems and it represents the particular case of the previously presented “standard” RLS algorithm.

A2. *Constant forgetting factor.* In this case:

$$\lambda_1(k) = \lambda_1, \quad 0 < \lambda_1 < 1, \quad \lambda_2(k) = \lambda_2 = 1, \quad (\text{A.104})$$

where the typical λ_1 values are:

$$\lambda_1 = 0.95 \text{ to } 0.99.$$

The criterion to be minimized will be:

$$\min_{\hat{\theta}(k)} J(k) = \sum_{i=1}^k \lambda_1^{(k-i)} \left[y(i) - \hat{\theta}^T(k) \phi(i-1) \right]^2. \quad (\text{A.105})$$

The value $\lambda_1(k) < 1$ introduces increasingly weaker weighting of the old data. The maximum weight is given to the most recent prediction error. This is why λ_1 is known as the *forgetting factor*. The use of a constant forgetting factor without monitoring the maximum value of $F(k)$ causes problems in adaptive control if the $\phi(k)\phi^T(k)$ sequence becomes null in the average (*i.e.* steady state is reached). In such a case, the adaptation gain tends to infinity:

$$\begin{aligned} F(k+i)^{-1} &= (\lambda_1)^i F(k)^{-1}, \\ F(k+i) &= (\lambda_1)^{-i} F(k), \\ \lim_{i \rightarrow \infty} (\lambda_1)^{-i} &= \infty \quad \text{for } \lambda_1 < 1. \end{aligned}$$

This type of profile is suitable for the estimation of the parameters of slowly time-varying systems.

A3. *Variable forgetting factor.* In this case:

$$\lambda_2(k) = \lambda_2 = 1, \quad (\text{A.106})$$

and the forgetting factor $\lambda_1(k)$ is given by:

$$\lambda_1(k) = \lambda_0 \lambda_1(k-1) + 1 - \lambda_0, \quad 0 < \lambda_0 < 1, \quad (\text{A.107})$$

where the typical values are:

$$\lambda_1(0) = 0.95 \text{ to } 0.99, \quad \lambda_0 = 0.5 \text{ to } 0.99.$$

This type of profile is recommended for the model identification of stationary systems, since it avoids a too fast decrease of the adaptation gain, which generally results in acceleration of the convergence.

Other types of $\lambda_1(k)$ evolution can be considered. For example:

$$\lambda_1(k) = 1 - \beta \frac{[\varepsilon_{\text{CL}}^{\circ}(k)]^2}{1 + \phi^T(k)F(k)\phi(k)}, \quad \beta > 0. \quad (\text{A.108})$$

This forgetting factor tends towards 1 when the prediction error tends towards zero. On the other hand, when a change occurs in the system parameters, the prediction error increases which leads to a forgetting factor smaller than 1 to assure a good adaptation capability.

A4. *Constant trace.* In this case, $\lambda_1(k)$ and $\lambda_2(k)$ are automatically chosen at each step in order to assure a constant trace of the gain matrix (constant sum of the diagonal elements):

$$\text{tr } F(k+1) = \text{tr } F(k) = \text{tr } F(0) = d_0 G_{\text{ini}}, \quad (\text{A.109})$$

in which d_0 is the number of parameters and G_{ini} determines the initial matrix adaptation gain:

$$F(0) = \begin{bmatrix} G_{\text{ini}} & 0 \\ & \ddots \\ 0 & G_{\text{ini}} \end{bmatrix}, \quad (\text{A.110})$$

where the typical G_{ini} values are:

$$G_{\text{ini}} = 0.1 \text{ to } 4.$$

Using this technique, there is a movement in the optimal direction of the RLS algorithm at each step, but the gain is maintained approximately constant. The values of $\lambda_1(k)$ and $\lambda_2(k)$ are determined from the equation:

$$\text{tr } F(k+1) = \frac{1}{\lambda_1(k)} \text{tr} \left[F(k) - \frac{F(k)\phi(k)\phi^T(k)F(k)}{\gamma(k) + \phi^T(k)F(k)\phi(k)} \right], \quad (\text{A.111})$$

fixing the ratio $\gamma(k) = \lambda_1(k)/\lambda_2(k)$. This relationship is a direct consequence of (A.102).

This type of profile is suitable for the model identification of systems with time-varying parameters and for adaptive control with non-vanishing adaptation.

A5. *Constant gain (gradient algorithm).* In this case:

$$\lambda_1(k) = \lambda_1 = 1, \quad \lambda_2(k) = \lambda_2 = 0 \quad (\text{A.112})$$

and thus from (A.102), one obtains:

$$F(k+1) = F(k) = F(0). \quad (\text{A.113})$$

The improved gradient adaptation algorithm given by (A.73) or (A.78) is then obtained.

This algorithm can be used for the identification and adaptive control of stationary or time-varying systems with few parameters (≤ 3) and in presence of a reduced noise level.

This type of adaptation gain results in performance that are inferior to those provided by the A1, A2, A3, A4 profiles, but it is simpler to implement.

Choice of the Initial Matrix Adaptation Gain $F(0)$

The initial matrix adaptation gain $F(0)$ is usually chosen as a diagonal matrix of the form given by (A.98) and, respectively, (A.110).

In the absence of initial information upon the parameters to be estimated, a high initial gain G_{ini} is chosen. A typical value is $G_{\text{ini}} = 1000$ but higher values can be chosen.

If an initial parameters estimation is available, a low initial gain is chosen. In general, in this case $G_{\text{ini}} < 1$.

A General Structure of PAA

For all methods the parameter adaptation algorithm (PAA) has the general form (integral type), [LK97b]:

$$\begin{aligned}
 \hat{\theta}(k+1) &= \hat{\theta}(k) + F(k)\phi(k)\varepsilon_{\text{CL}}(k+1), \\
 F(k+1)^{-1} &= \lambda_1(k)F(k)^{-1} + \lambda_2(k)\phi(k)\phi^{\text{T}}(k), \\
 0 < \lambda_1(k) \leq 1, \quad 0 \leq \lambda_2(k) < 2, \\
 F(0) > 0, \quad F(k)^{-1} > \alpha F(0)^{-1}, \quad 0 < \alpha < \infty, \\
 F(k+1) &= \frac{1}{\lambda_1(k)} \left[F(k) - \frac{F(k)\phi(k)\phi^{\text{T}}(k)F(k)}{\lambda_1(k) + \phi^{\text{T}}(k)F(k)\phi(k)} \right], \\
 \varepsilon_{\text{CL}}(k+1) &= \frac{\varepsilon_{\text{CL}}^{\circ}(k+1)}{1 + \phi^{\text{T}}(k)F(k)\phi(k)},
 \end{aligned} \tag{A.114}$$

where $\varepsilon_{\text{CL}}^{\circ}(k+1) = f_1(\hat{\theta}(k), \hat{\theta}(k-1), \dots, y(k+1), \varepsilon_{\text{CL}}(k), \varepsilon_{\text{CL}}(k-1), \dots)$ is the *a priori* adaptation error, $\varepsilon_{\text{CL}}(k+1) = f_2(\hat{\theta}(k+1), \hat{\theta}(k), \dots, y(k+1), \varepsilon_{\text{CL}}(k), \varepsilon_{\text{CL}}(k-1), \dots)$ is the *a posteriori* adaptation error and $\phi(k)$ is the observation vector.

For each recursive identification algorithm (CLOE, F-CLOE, AF-CLOE, X-CLOE, G-CLOE), the variables θ , ϕ , $\varepsilon_{\text{CL}}^{\circ}(k+1)$ have a specific expression. Note that the sequences $\lambda_1(k)$ and $\lambda_2(k)$ have different laws for the time evolution of the adaptation gain $F(k)$. For convergence analysis in the stochastic environment, it is assumed that a PAA with decreasing gain is used (*i.e.* $\lambda_1(k) \equiv 1$, $\lambda_2(k) > 0$ or $\lambda_2(k) = \lambda_2$).

A.5 Properties of Parameter Adaptation Algorithms

In this section, only the properties of CLOE algorithm are summarized briefly. The properties of the other algorithms (F-CLOE, AF-CLOE, X-CLOE and G-CLOE) are given in the same way in [LK97b].

Stability of PAA in a Deterministic Environment

Equivalent Feedback Representation of PAA

In the case of recursive least squares or of the improved gradient algorithm, the following *a posteriori* predictor has been used, see (A.24):

$$\hat{y}(k+1) = \hat{y} \left[k+1 | \hat{\theta}(k+1) \right] = \hat{\theta}^{\text{T}}(k+1)\phi(k). \tag{A.115}$$

The PAA has the following form, see (A.114):

$$\hat{\theta}(k+1) = \hat{\theta}(k) + F(k)\phi(k)\varepsilon_{\text{CL}}(k+1). \tag{A.116}$$

The parameter error is defined as:

$$\tilde{\theta}(k) = \hat{\theta}(k) - \theta. \tag{A.117}$$

Subtracting θ in both sides of (A.116) and taking into account (A.117), one obtains:

$$\tilde{\theta}(k+1) = \tilde{\theta}(k) + F(k)\phi(k)\varepsilon_{\text{CL}}(k+1). \tag{A.118}$$

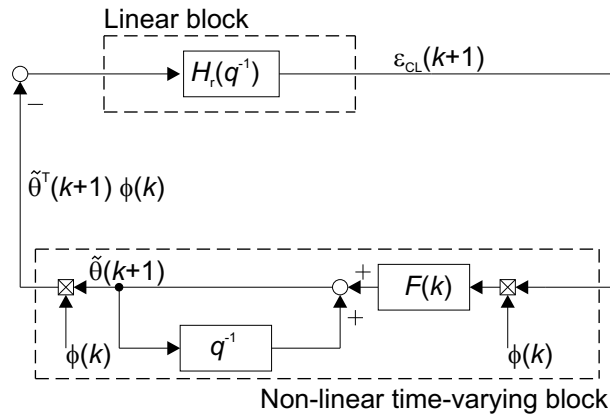


Figure A.2: Equivalent feedback representation of PAA associated to the output-error predictor (OE plant model).

From the definition of the *a posteriori* prediction error $\varepsilon_{\text{CL}}(k+1)$ given by (A.26) and taking into account (A.6) for zero output disturbance noise $p(k) \equiv 0$ (deterministic case) and (A.117), one gets:

$$\varepsilon_{\text{CL}}(k+1) = -H_r(q^{-1})\tilde{\theta}^{\text{T}}(k+1)\phi(k), \quad (\text{A.119})$$

where $H_r(q^{-1})$ is a transfer operator given from the closed-loop predictor. The operator $H_r(q^{-1})$ will be specified more precisely for CLOE algorithm in the following (see (A.122) and (A.125)). Using (A.118), one can write:

$$\tilde{\theta}^{\text{T}}(k+1)\phi(k) = \tilde{\theta}^{\text{T}}(k)\phi(k) - \phi^{\text{T}}(k)F(k)\phi(k)\varepsilon_{\text{CL}}(k+1). \quad (\text{A.120})$$

Equation (A.118) through (A.120) defines an equivalent feedback system represented in fig. A.2. Equation (A.119) defines a linear block with constant parameters on the feedforward path, whose input is $-\tilde{\theta}^{\text{T}}(k+1)\phi(k)$. This block is characterized by a transfer function $H_r(q^{-1})$ in the case of output-error predictor (OE plant model). Equations (A.118) and (A.120) define a non-linear time-varying block in the feedback path.

The equivalent feedback representation associated to PAA is always formed by two blocks: The linear time-invariant block and the non-linear time-varying block. Passivity (hyperstability) properties or Lyapunov functions are well suited for the stability analysis of such feedback system and the same results can be obtained with both approaches.

We shall next present the results of the stability of PAA for CLOE algorithm using passivity approach. Detailed analysis can be found in [LLM97].

Stability of PAA Using the Equivalent Feedback Representation

Associated with the PAA (A.114) one considers the class of adaptive systems for which *a posteriori* adaptation error satisfies:

$$\varepsilon_{\text{CL}}(k+1) = H_r(q^{-1}) \left[\theta - \hat{\theta}(k+1) \right]^{\text{T}} \phi(k), \quad (\text{A.121})$$

see (A.27), where:

$$H_r(q^{-1}) = \frac{H_1(q^{-1})}{H_2(q^{-1})}, \quad (\text{A.122})$$

with:

$$H_i(q^{-1}) = 1 + \sum_{i=1}^{n_j} h_i^j q^{-i}; \quad j = 1, 2 \quad (\text{A.123})$$

and θ is a fixed value of the unknown parameter vector.

The equation for the *a posteriori* prediction error, in the deterministic case ($p(k) \equiv 0$), is:

$$\varepsilon_{\text{CL}}(k+1) = \frac{S(q^{-1})}{P(q^{-1})} [\theta - \hat{\theta}(k+1)]^T \phi(k). \quad (\text{A.124})$$

and consequently one can define $H_r(q^{-1})$ as:

$$H_r(q^{-1}) = \frac{S(q^{-1})}{P(q^{-1})}. \quad (\text{A.125})$$

Exploiting the input-output properties of the equivalent feedback and feedforward block from fig. A.2, one has the following general result for CLOE algorithm, see [LK97b]:

Theorem 3 *Assuming that the closed-loop system (operated in closed-loop with a RST digital controller) is stable, the recursive parameter estimation algorithm given by (A.114) assures*

$$\begin{aligned} \lim_{k \rightarrow \infty} \varepsilon_{\text{CL}}(k+1) &= 0, \\ \lim_{k \rightarrow \infty} \varepsilon_{\text{CL}}^\circ(k+1) &= 0, \\ \|\phi(k)\| &< C, \quad 0 < C < \infty, \quad \forall k \end{aligned}$$

for all initial conditions $\hat{\theta}(0)$, $\varepsilon_{\text{CL}}^\circ(0)$ and $\phi(0)$ if

$$H'_r(z^{-1}) = H_r(z^{-1}) - \frac{\lambda_2}{2} = \frac{S(z^{-1})}{P(z^{-1})} - \frac{\lambda_2}{2} \quad (\text{A.126})$$

is strictly positive real (SPR) transfer function, where: $\max_k (\lambda_2(k)) \leq \lambda_2 < 2$ and $P(z^{-1})$ represents the characteristic polynomial (closed-loop poles).

The transformed equivalent feedback associated to the Theorem 3 is shown in fig. A.3. The equivalent feedback path has a local negative feedback with a gain $\max_k \frac{\lambda_2(k)}{2} \leq \frac{\lambda_2}{2}$, which makes the transformed feedback path passive. However, in order that the whole feedback system remains unchanged, a gain $-\frac{\lambda_2}{2}$ should be added in parallel to the feedforward path and, therefore, the equivalent transformed linear path will be characterized by the transfer function $H'_r(q^{-1}) = H_r(q^{-1}) - \frac{\lambda_2}{2}$.

Parametric Convergence Analysis in a Stochastic Environment

One of the objectives of closed-loop identification is to obtain asymptotic unbiased estimates in the presence of noise on the plant output. We shall use for this analysis an approach based on an associated *Ordinary Differential Equation* (ODE), see [Lju77], and a specific result for a class of parameter-estimation algorithms [DL80].

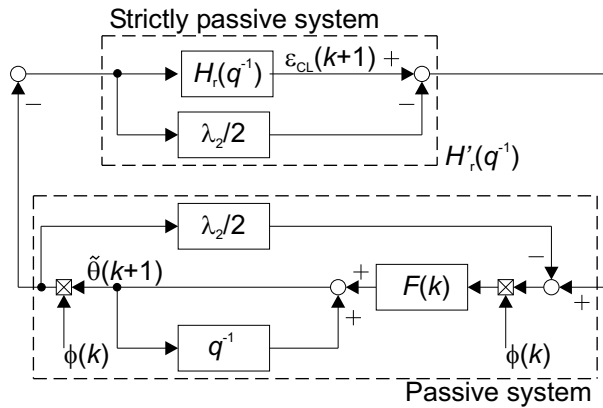


Figure A.3: Transformed equivalent feedback systems associated to the PAA with time-varying gain for the output-error predictor (OE plant model).

The equation of the *a posteriori* prediction error in the presence of noise is:

$$\varepsilon_{\text{CL}}(k+1) = \frac{S(q^{-1})}{P(q^{-1})} [\theta - \hat{\theta}(k+1)]^T \phi(k) + \frac{A(q^{-1})S(q^{-1})}{P(q^{-1})} p(k+1). \quad (\text{A.127})$$

One has the following result, see Theorem 4.1 in [LK97a]:

Theorem 4 Consider the recursive parameter estimation algorithm given by (A.114) with $\lambda_1(k) = 1$. Define

$$\phi(k, \hat{\theta}) \triangleq \phi(k)|_{\hat{\theta}(k)=\hat{\theta}=\text{const.}}, \quad (\text{A.128})$$

$$\varepsilon_{\text{CL}}(k+1, \hat{\theta}) \triangleq \varepsilon_{\text{CL}}(k+1)|_{\hat{\theta}(k)=\hat{\theta}=\text{const.}}, \quad (\text{A.129})$$

$$D_s \triangleq \{\hat{\theta} : \hat{A}(z^{-1})S(z^{-1}) + z^{-d}B(z^{-1})R(z^{-1}) = 0 \Rightarrow |z| < 1\} \quad (\text{A.130})$$

Assume that $\hat{\theta}(k)$ generated by the algorithm belongs infinitely often to the domain D_s for which the stationary process $\phi(k, \hat{\theta})$ and $\varepsilon_{\text{CL}}(k+1, \hat{\theta})$ can be defined.

Assume that $p(k)$ is a zero-mean stochastic process with finite moments independent of the reference sequence $r^*(k)$.

If

$$H'_r(z^{-1}) = H_r(z^{-1}) - \frac{\lambda_2}{2} = \frac{S(z^{-1})}{P(z^{-1})} - \frac{\lambda_2}{2} \quad (\text{A.131})$$

is strictly positive real (SPR) transfer function, where: $\max_t (\lambda_2(t)) \leq \lambda_2 < 2$ then

$$\text{Prob} \left\{ \lim_{k \rightarrow \infty} \hat{\theta}(k) \in D_c \right\} = 1,$$

where $D_c = \{\hat{\theta} : (\phi^T(k, \hat{\theta}))(\theta - \hat{\theta}) = 0\}$. If, furthermore, $\phi^T(k, \hat{\theta})(\theta - \hat{\theta}) = 0$ has a unique solution (richness condition) then the condition on $H'_r(z^{-1})$ (A.131) to be strictly positive real implies that

$$\text{Prob} \left\{ \lim_{k \rightarrow \infty} \hat{\theta}(k+1) = \theta \right\} = 1.$$

This theorem is establish here because it is useful for the parametric convergence analysis in a stochastic environment of the state-space identification methods in chapter 2.

Frequency Distribution of the Asymptotic Bias

The asymptotic expression for the bias distribution of model estimate plays an important role in the analysis of different identification methods, particularly when the true system does not belong to the model set. In this section the bias distributions of the estimates for CLOE algorithm is presented using the framework considered in [Lju99].

Consider the discrete-time SISO system with the additive disturbance $p(k)$:

$$\begin{aligned} y(k) &= G_S(q^{-1})u(k) + p(k), \\ p(k) &= H(q^{-1})e(k), \end{aligned}$$

where $y(k)$ is the output signal, $u(k)$ is the input signal and $e(k)$ is a zero-mean Gaussian white noise signal. $G_S(q^{-1})$ and $H(q^{-1})$ are the transfer operators of the plant and noise model, respectively.

Consider the excitation signal $w(k)$ is added to the plant input. For the case of the **closed-loop** methods, the formula for the estimated parameter vector when number of data $N \rightarrow \infty$ is given by:

$$\hat{\theta}^* = \arg \min_{\theta} \int_{-\pi}^{\pi} |S_{yp}|^2 \left[|G_S - \hat{G}|^2 |\hat{S}_{yp}|^2 \phi_w + |H|^2 \phi_e \right] d\omega, \quad (\text{A.132})$$

where $S_{yp}(\omega)$ corresponds to the true output sensitivity transfer function between the output disturbance p and the plant output y , $\hat{S}_{yp}(\omega)$ is the estimated output sensitivity transfer function, $\phi_w(\omega)$ corresponds to the spectral density of the external excitation signal w and the $\phi_e(\omega)$ is the spectral density of the noise e . This formula shows that the noise does not affect the parameter estimation.

On the other hand, for the case of **open-loop** identification (direct method) $e(k)$ and $u(k)$ are independent, the formula for the estimated parameter vector when number of data $N \rightarrow \infty$ is given by:

$$\hat{\theta}^* = \arg \min_{\theta} \int_{-\pi}^{\pi} |\hat{H}^{-1}|^2 \left[|G_S - \hat{G}|^2 \phi_u + |H - \hat{H}|^2 \phi_e \right] d\omega, \quad (\text{A.133})$$

where \hat{H} is the estimate of noise model transfer operator H and $\phi_u(\omega)$ corresponds to the spectral density of the input u . Comparing this equation (A.133) with (A.132), one observes that in the closed-loop operation the bias of the plant model is related to the bias of the noise model. It means that contrary to the open-loop case when the noise model and plant model are independently parametrized, incorrect estimation of the noise model leads to the bias estimation of the plant model. It can be shown that this bias is proportional to the noise variance and $H - \hat{H}$, [Lju93].

For more details see [LK97b], [LK97a] and [LLM97].

A.6 Conclusions

In this chapter, we have presented the parameter adaptation algorithms using closed-loop output-error identification predictor. We have examined their properties in deterministic and stochastic environment.

We can emphasize the following properties:

The recursive algorithms for plant model identification in closed-loop operation are efficient tools either for improving open loop identified models or for re-design and re-tuning of existing controllers, based on the improved models.

Identification in closed-loop allows the identification of models for plants which can hardly be operated in open loop (presence of integrators, open loop unstable, drift, *etc.*).

Recursive identification in closed-loop is based on the use of an adaptive predictor for the closed-loop which is re-parameterized in terms of the model to be identified. The estimated parameters asymptotically minimize a criterion of the closed-loop predictor error.

As for the case of identification in open loop, there is no single algorithm which gives the best identified model, but a family of algorithms are available.

The properties of the PAA (stability, convergence, frequency distribution of the asymptotic bias) in the presence of stochastic disturbances depend upon the structure of the adjustable predictor and the way in which the observation vector and adaptation error is generated.

State-space Identification Experiments

B

B.1 Simulation Experiments for SISO Systems

Simulation Experiment 3

The following state-space representation of the test plant \mathcal{S} is considered:

$$\begin{aligned} A_0(\theta) &= \left[\begin{array}{c|c} -\theta_1 & -\theta_2 \\ \hline 1 & 0 \end{array} \right], & B_0(\theta) &= \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \\ C_0(\theta) &= \left[\begin{array}{cc} 2.9573 & 2.9490 \end{array} \right], \end{aligned}$$

for the parameter vector defined by:

$$\theta = [\theta_1, \theta_2]^T.$$

This is the first experiment in which the number of estimated parameters is smaller than the number of transfer function coefficients ($2n > d_0$). It means that some knowledge about identified systems constants have been included to the estimated model. Two constants remains to be estimated in this case.

Since the state-space matrices are given in a canonical controller form, one can immediately determine the vector of transfer function coefficients $\Gamma(\theta)$ and its Jacobian matrix $\Gamma'_\theta(\theta)$:

$$\Gamma(\theta) = \begin{bmatrix} \theta_1 \\ \theta_2 \\ 2.9573 \\ 2.9490 \end{bmatrix}, \quad \Gamma'_\theta(\theta) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

This model structure is also linear in its inputs but nonlinear in its parameters since the right-hand side of closed-loop predictor (2.70) involves the product of parameters by model outputs that depend on $\hat{\theta}$. Identifiability of this model is satisfied (Laplace transform approach [BÅ70]). The true values of the estimated parameters are straightforward:

$$\theta = \begin{bmatrix} -1.9887 \\ 0.9915 \end{bmatrix} \Rightarrow \Gamma(\theta) = \begin{bmatrix} -1.9887 \\ 0.9915 \\ 2.9573 \\ 2.9490 \end{bmatrix}.$$

Table B.1 gives an overview of the simulation settings. One can see that the same constant of 100 is selected as the scalar adaptation gain for the IGM-1 and IGM-2 algorithm.

The simulation results are summarized in tables B.2 and B.3. Figure B.1 shows that the convergence speed of the IGM-1 and IGM-2 algorithms are the best. The RLS-2

Table B.1: Settings of the SISO experiment 3.

Initial parameter vector estimate	$\hat{\theta}(0) = [0, 0]^T$
Standard deviation of white noise	$\sigma_e = 0.00001$
RLS-2 initial matrix adaptation gain	$F(0) = 10000I$
RLS-2 variation profile sequence $\lambda_1(k)$	$\lambda_1(k) \equiv 0.98$
RLS-2 variation profile sequence $\lambda_2(k)$	$\lambda_2(k) \equiv 1$
IGM-1, IGM-2 scalar adaptation gain	$\alpha_1 = \alpha_2 = 100$
GM-2 scalar adaptation gain	$\alpha = 0.04$
Number of simulation steps	$k_{\text{end}} = 2047$
Number of simulation runs	$N = 100$

Table B.2: Results of the SISO experiment 3.

	$\bar{\theta}[\times 10^{-4}]$	$\bar{\sigma}[\times 10^{-4}]$
RLS-2	1.3895	± 0.3771
IGM-1	4.5635	± 0.7992
IGM-2	4.0345	± 0.8429
GM-2	7599	± 0.0716

algorithm suffers from a lower convergence speed while the GM-2 algorithm has the worst convergence speed irrespective of the scalar adaptation gain settings.

Lower mean value of the parametric error $\bar{\theta}$ and standard deviation of the estimated parameter vector $\bar{\sigma}$ for the RLS-2 algorithm in table B.2 in comparison with the IGM-1 and IGM-2 algorithms are caused by the time-varying adaptation gain matrix for the RLS-2 algorithm.

Figures B.2 and B.3 illustrate the evolution of the individual parameters in the case of the RLS-2 and IGM-2 algorithms, respectively. One can see that the IGM-2 algorithm requires only 80 steps in comparison 1000 steps for the RLS-2 algorithm in order to come near the true values.

Table B.3: Results of the SISO experiment 3, mean values and standard deviations of the individual parameters.

	j	1	2
RLS-2	$\bar{\theta}_j$	-1.9886	0.9914
	$\sigma_j[\times 10^{-3}]$	± 0.0343	± 0.0410
IGM-2	$\bar{\theta}_j$	-1.9891	0.9919
	$\sigma_j[\times 10^{-3}]$	± 0.0964	± 0.0721

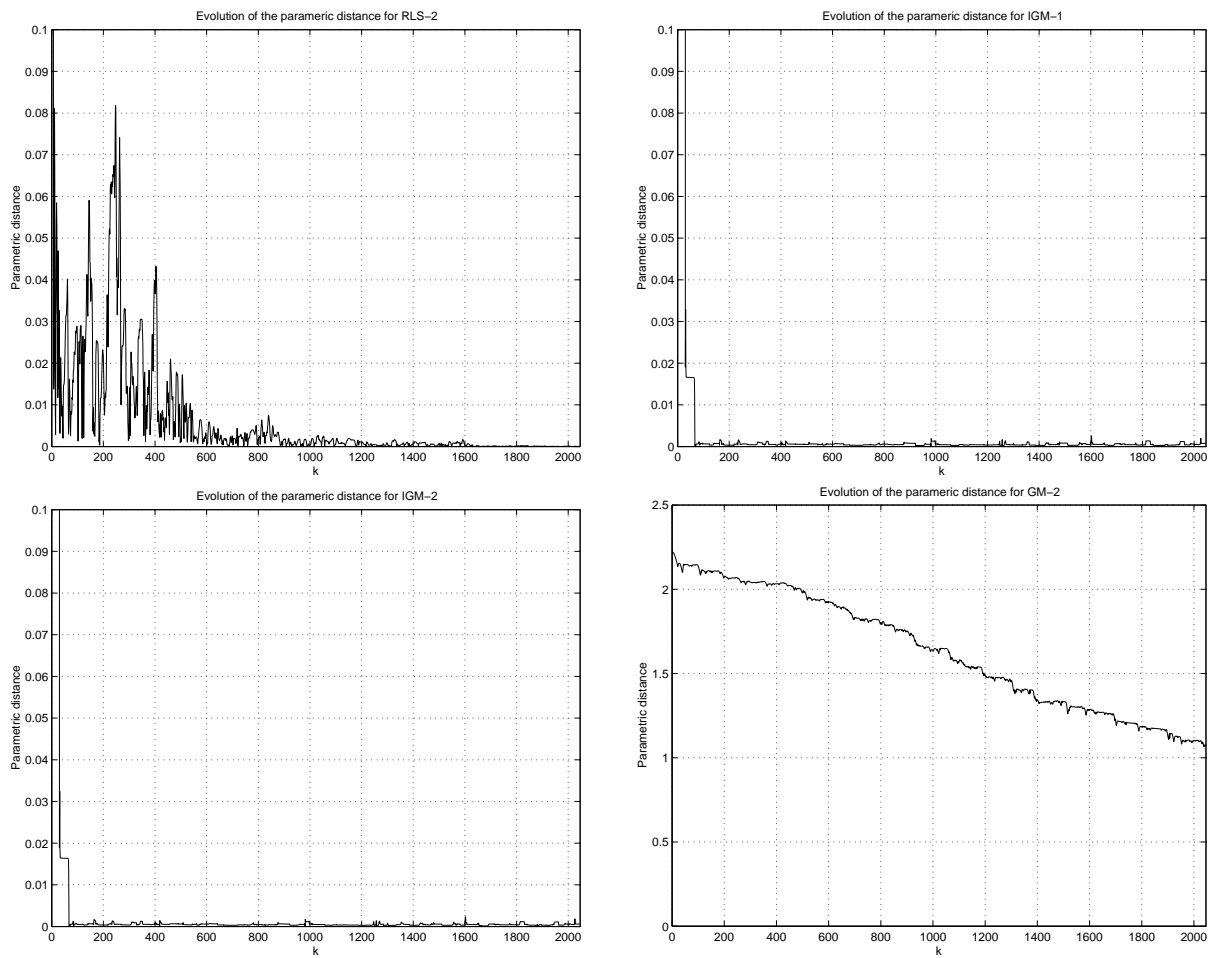


Figure B.1: Evolution of the parametric distance for one simulation run (SISO experiment 3).

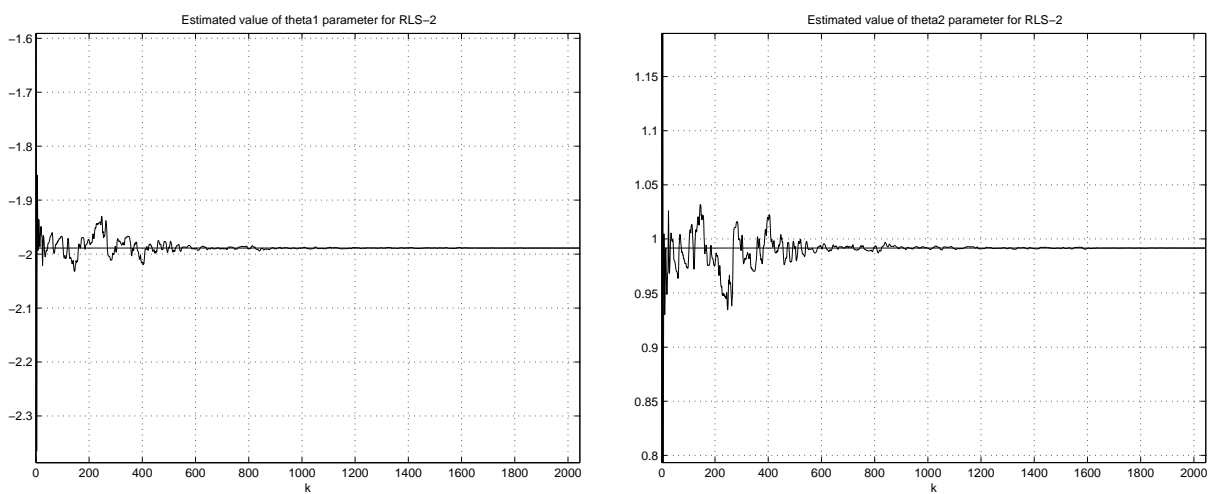


Figure B.2: Evolution of the estimated parameters for one simulation run (RLS-2, SISO experiment 3).

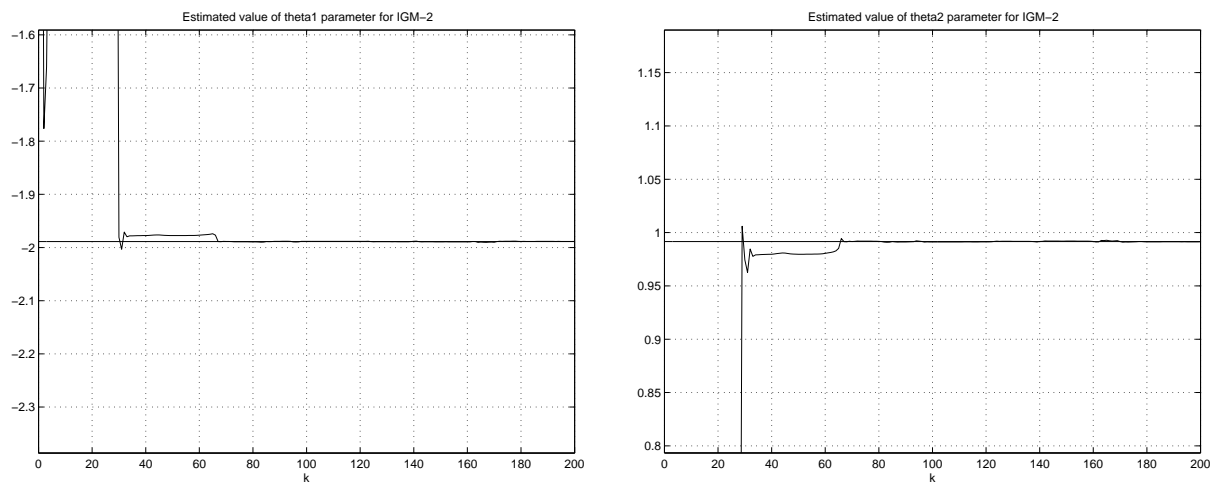


Figure B.3: Evolution of the estimated parameters for one simulation run (IGM-2, SISO experiment 3).

B.2 Simulation Experiments for MIMO Systems

In the case of MIMO systems, the experiments are performed in the closed-loop identification scheme (fig. 2.2) using the RLS-2 and IGM-2 algorithm. One should note that, independently of the criterion taken, the simulation results are controller-dependent. A MIMO test plant \mathcal{S} is operated by a digital linear quadratic (LQ) controller \mathcal{K} .

The statistic criteria (2.249), (2.250), (2.251), (2.252) have been used to evaluate the results.

The following discrete-time state-space representation of the test plant \mathcal{S} , which represents a flying insect model [DSS03], is considered:

$$A_{\mathcal{S}} = \begin{bmatrix} 0.9923 & 0.5807 \cdot 10^{-3} & 0.1472 \cdot 10^{-3} & 3.1350 \cdot 10^{-8} \\ 0.1229 \cdot 10^{-3} & 0.9900 & 4.5050 \cdot 10^{-8} & 0.1471 \cdot 10^{-3} \\ -102 & 7.6770 & 0.9600 & 0.6065 \cdot 10^{-3} \\ 1.616 & -131.8 & 0.6338 \cdot 10^{-3} & 0.9579 \end{bmatrix},$$

$$B_{\mathcal{S}} = \begin{bmatrix} -1.015 \cdot 10^{-6} & 0.5643 \cdot 10^{-3} \\ 1.122 \cdot 10^{-3} & -1.133 \cdot 10^{-3} \\ -0.0119 & 7.472 \\ 14.85 & -14.97 \end{bmatrix},$$

$$C_{\mathcal{S}} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

The matrices $\{A_{\mathcal{S}}, B_{\mathcal{S}}, C_{\mathcal{S}}\}$, that describes two inputs and four outputs MIMO system, can be obtained directly from micromechanical flying insect morphological parameters such as mass, moment of inertia, center of mass, *etc.* if a continuous-time nonlinear system is approximated by a linear discrete-time system, [DSS03].

The output feedback LQ controller $u(k) = -Ky(k)$ has been designed to minimize the following quadratic cost function:

$$J = \lim_{N \rightarrow \infty} E \left(\sum_{k=1}^N x(k)^T Q x(k) + u(k)^T R u(k) \right),$$

where $Q \geq 0$ and $R \geq 0$ are the weighting matrices to reflect the trade-off between regulation performance and control effort, and the diagonal entries in the weighting matrices are iteratively tuned to ensure a good transient response without saturating the control inputs. The final choice of the weighting matrices Q and R for the controller are $Q = \text{diag}(10, 20, 1, 1)$ and $R = \text{diag}(1, 2)$. The sampling period is $T_s = 0.15$ ms.

In order to test the proposed state-space identification algorithms, various state-space representations $\{A_0(\theta), B_0(\theta), C_0(\theta)\}$ of the plant \mathcal{S} operated by the controller \mathcal{K} are identified in the following simulation experiments.

Two pseudo-random binary sequences (PRBS), generated by a 8-bit register, added to the controller outputs vector $u(k)$ are taken as excitation signals vector $w(k)$. Their value alternates between $+0.02$ and -0.02 and their clock frequency is equal to $f_s/2$.

Four zero-mean Gaussian white noises of standard deviation $\sigma_{e1}, \sigma_{e2}, \sigma_{e3}, \sigma_{e4}$ are used as the output disturbance signals vector $e(k) = [e_1(k), e_2(k), e_3(k), e_4(k)]^T$ and a noise filter $H = \text{diag}[1, 1, 1, 1]$ has been considered.

Simulation Experiment 4

The following state-space representation of the test plant \mathcal{S} is considered:

$$\begin{aligned}
 A_0(\theta) &= \left[\begin{array}{c|c|c|c} 0.9923 & 0.5807 \cdot 10^{-3} & 0.1472 \cdot 10^{-3} & 3.1350 \cdot 10^{-8} \\ \hline 0.1229 \cdot 10^{-3} & 0.9900 & 4.5050 \cdot 10^{-8} & 0.1471 \cdot 10^{-3} \\ \hline -102 & 7.6770 & 0.9600 & 0.6065 \cdot 10^{-3} \\ \hline 1.616 & -131.8 & 0.6338 \cdot 10^{-3} & 0.9579 \end{array} \right], \\
 B_0(\theta) &= \left[\begin{array}{c|c} -1.015 \cdot 10^{-6} & 0.5643 \cdot 10^{-3} \\ \hline 1.122 \cdot 10^{-3} & -1.133 \cdot 10^{-3} \\ \hline \theta_1 & \theta_2 \\ \hline \theta_3 & \theta_4 \end{array} \right], \\
 C_0(\theta) &= \left[\begin{array}{c|c|c|c} 1 & 0 & 0 & 0 \\ \hline 0 & 1 & 0 & 0 \\ \hline 0 & 0 & 1 & 0 \\ \hline 0 & 0 & 0 & 1 \end{array} \right].
 \end{aligned}$$

for the parameter vector defined by:

$$\theta = [\theta_1, \theta_2, \theta_3, \theta_4]^T.$$

One can see that the parameter vector is composed of four parameters which define the unknown parameters in matrix $B_0(\theta)$. The state-space representation used for identification $\{A_0(\theta), B_0(\theta), C_0(\theta)\}$ corresponds directly to the discrete-time state-space physical description. Moreover, this form has also canonical observer form.

The vector of transfer function coefficient $\Gamma(\theta)$ and its Jacobian matrix can be defined as follows:

$$\Gamma(\theta) = \begin{bmatrix} -3.9003 \\ 5.7385 \\ -3.7744 \\ 9.3648 \cdot 10^{-1} \\ -1.0151 \cdot 10^{-6} \\ 3.6033 \cdot 10^{-6} + 1.4718 \cdot 10^{-4}\theta_1 + 3.1347 \cdot 10^{-8}\theta_3 \\ -2.8674 \cdot 10^{-6} + 1.1354 \cdot 10^{-7}\theta_3 - 2.8670 \cdot 10^{-4}\theta_1 \\ 3.1940 \cdot 10^{-7} + 2.5609 \cdot 10^{-8}\theta_3 + 1.4244 \cdot 10^{-4}\theta_1 \\ 5.6429 \cdot 10^{-4} \\ -1.6416 \cdot 10^{-3} + 1.4718 \cdot 10^{-4}\theta_2 + 3.1347 \cdot 10^{-8}\theta_4 \\ 1.6014 \cdot 10^{-3} + 1.1354 \cdot 10^{-7}\theta_4 - 2.8670 \cdot 10^{-4}\theta_2 \\ -5.2365 \cdot 10^{-4} + 2.5609 \cdot 10^{-8}\theta_4 + 1.4244 \cdot 10^{-4}\theta_2 \\ 1.1216 \cdot 10^{-3} \\ -3.2642 \cdot 10^{-3} + 1.4707 \cdot 10^{-4}\theta_3 + 4.5046 \cdot 10^{-8}\theta_1 \\ 3.1830 \cdot 10^{-3} + 2.3461 \cdot 10^{-8}\theta_1 - 2.8714 \cdot 10^{-4}\theta_3 \\ -1.0397 \cdot 10^{-3} - 3.2032 \cdot 10^{-8}\theta_1 + 1.4232 \cdot 10^{-4}\theta_3 \\ -1.1306 \cdot 10^{-3} \\ 3.2904 \cdot 10^{-3} + 1.4707 \cdot 10^{-4}\theta_4 + 4.5046 \cdot 10^{-8}\theta_2 \\ -3.2085 \cdot 10^{-3} + 2.3461 \cdot 10^{-8}\theta_2 - 2.8714 \cdot 10^{-4}\theta_4 \\ 1.0479 \cdot 10^{-3} - 3.2032 \cdot 10^{-8}\theta_2 + 1.4232 \cdot 10^{-4}\theta_4 \\ \theta_1 \\ 8.7144 \cdot 10^{-3} - 2.9403\theta_1 + 6.0646 \cdot 10^{-4}\theta_3 \\ -1.7151 - 7.6295 \cdot 10^{-5}\theta_3 + 2.9007\theta_1 \\ 8.4384 \cdot 10^{-3} - 5.3016 \cdot 10^{-4}\theta_3 - 9.6032 \cdot 10^{-1}\theta_1 \\ \theta_2 \\ -6.6265 \cdot 10^{-2} - 2.9403\theta_2 + 6.0646 \cdot 10^{-4}\theta_4 \\ 1.2926 \cdot 10^{-1} - 7.6295 \cdot 10^{-5}\theta_4 + 2.9007\theta_2 \\ -6.4135 \cdot 10^{-2} - 5.3016 \cdot 10^{-4}\theta_4 - 9.6032 \cdot 10^{-1}\theta_2 \\ \theta_3 \\ -1.4786 \cdot 10^{-1} + 6.3384 \cdot 10^{-4}\theta_1 - 2.9424\theta_3 \\ 2.8869 \cdot 10^{-1} + 2.9005\theta_3 - 1.0246 \cdot 10^{-3}\theta_1 \\ -1.4309 \cdot 10^{-1} - 9.5802 \cdot 10^{-1}\theta_3 + 3.9072 \cdot 10^{-4}\theta_1 \\ \theta_4 \\ 1.4996 \cdot 10^{-1} + 6.3384 \cdot 10^{-4}\theta_2 - 2.9424\theta_4 \\ -2.9282 \cdot 10^{-1} + 2.9005\theta_4 - 1.0246 \cdot 10^{-3}\theta_2 \\ 1.4514 \cdot 10^{-1} - 9.5802 \cdot 10^{-1}\theta_4 + 3.9072 \cdot 10^{-4}\theta_2 \end{bmatrix},$$

$$\Gamma'_\theta(\theta) = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1.4718 \cdot 10^{-4} & 0 & 3.1347 \cdot 10^{-8} & 0 \\ -2.8670 \cdot 10^{-4} & 0 & 1.1354 \cdot 10^{-7} & 0 \\ 1.4244 \cdot 10^{-4} & 0 & 2.5609 \cdot 10^{-8} & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1.4718 \cdot 10^{-4} & 0 & 3.1347 \cdot 10^{-8} \\ 0 & -2.8670 \cdot 10^{-4} & 0 & 1.1354 \cdot 10^{-7} \\ 0 & 1.4244 \cdot 10^{-4} & 0 & 2.5609 \cdot 10^{-8} \\ 0 & 0 & 0 & 0 \\ 4.5046 \cdot 10^{-8} & 0 & 1.4707 \cdot 10^{-4} & 0 \\ 2.3461 \cdot 10^{-8} & 0 & -2.8714 \cdot 10^{-4} & 0 \\ -3.2032 \cdot 10^{-8} & 0 & 1.4232 \cdot 10^{-4} & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 4.5046 \cdot 10^{-8} & 0 & 1.4707 \cdot 10^{-4} \\ 0 & 2.3461 \cdot 10^{-8} & 0 & -2.8714 \cdot 10^{-4} \\ 0 & -3.2032 \cdot 10^{-8} & 0 & 1.4232 \cdot 10^{-4} \\ 1 & 0 & 0 & 0 \\ -2.9403 & 0 & 6.0646 \cdot 10^{-4} & 0 \\ 2.9007 & 0 & -7.6295 \cdot 10^{-5} & 0 \\ -9.6032 \cdot 10^{-1} & 0 & -5.3016 \cdot 10^{-4} & 0 \\ 0 & 1 & 0 & 0 \\ 0 & -2.9403 & 0 & 6.0646 \cdot 10^{-4} \\ 0 & 2.9007 & 0 & -7.6295 \cdot 10^{-5} \\ 0 & -9.6032 \cdot 10^{-1} & 0 & -5.3016 \cdot 10^{-4} \\ 0 & 0 & 1 & 0 \\ 6.3384 \cdot 10^{-4} & 0 & -2.9424 & 0 \\ -1.0246 \cdot 10^{-3} & 0 & 2.9005 & 0 \\ 3.9072 \cdot 10^{-4} & 0 & -9.5802 \cdot 10^{-1} & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 6.3384 \cdot 10^{-4} & 0 & -2.9424 \\ 0 & -1.0246 \cdot 10^{-3} & 0 & 2.9005 \\ 0 & 3.9072 \cdot 10^{-4} & 0 & -9.5802 \cdot 10^{-1} \end{bmatrix}.$$

The true values of the estimated parameters are straightforward:

$$\theta = \begin{bmatrix} -1.1898 \cdot 10^{-2} \\ 7.4716 \\ 14.849 \\ -14.967 \end{bmatrix}$$

Table B.4 gives an overview of the simulation settings. There are two cases (A, B) which have been studied using RLS-2 algorithm. The difference comes from two different initial matrix adaptation gains $F(0)$. Initial matrix adaptation gain $F(0)$ is equal to $0.1I$ in the first case (A) while it is equal to $2I$ in the second case (B).

Table B.4: Settings of the MIMO experiment 4.

Standard deviation of white noise $e_1(k)$	$\sigma_{e1} = 0.000001$
Standard deviation of white noise $e_2(k)$	$\sigma_{e2} = 0.000001$
Standard deviation of white noise $e_3(k)$	$\sigma_{e3} = 0.000001$
Standard deviation of white noise $e_4(k)$	$\sigma_{e4} = 0.000001$
Initial parameter vector estimate (A, B)	$\hat{\theta}(0) = [-0.01, 7.4, 14, -14]^T$
RLS-2 initial matrix adaptation gain (A)	$F(0) = 0.1I$
RLS-2 initial matrix adaptation gain (B)	$F(0) = 2I$
RLS-2 variation profile sequence $\lambda_1(k)$	$\lambda_1(k) \equiv 0.995$
RLS-2 variation profile sequence $\lambda_2(k)$	$\lambda_2(k) \equiv 1$
Initial parameter vector estimate	$\hat{\theta}(0) = [0, 0, 0, 0]^T$
IGM-2 scalar adaptation gain	$\alpha_2 = 100$
Number of simulation steps	$k_{\text{end}} = 255$
Number of simulation runs	$N = 100$

Table B.5: Results of the MIMO experiment 4.

	$\bar{\theta}[\times 10^{-4}]$	$\bar{\sigma}[\times 10^{-4}]$
RLS-2 (A)	–	–
RLS-2 (B)	–	–
IGM-2	1.4892	± 19.8312

One can also see that a constant forgetting factor of 0.995 is selected as an adaptation gain variation profile in both cases (A, B) for the RLS-2 algorithm.

The simulation results are summarized in tables B.5 and B.6. The evolution of the parametric distance for RLS-2 algorithm in both cases (A, B) is shown in fig. B.4. One can see that RLS-2 does not converge for given number of simulation step irrespective of the initial matrix adaptation gain settings. Better estimate is provided by the IGM-2 algorithm.

Figures B.5 and B.7 illustrate the evolution of the individual parameters in the case of the RLS-2 and IGM-2 algorithms, respectively. One can see that the IGM-2 algorithm requires about 200 steps to come near the all true values.

Table B.6: Results of the MIMO experiment 4, mean values and standard deviations of the individual parameters.

	j	1	2	3	4
IGM-2	$\bar{\theta}_j$	-0.01181	7.4715	14.8487	-14.9673
	$\sigma_j[\times 10^{-3}]$	± 1.1816	± 2.0679	± 1.7008	± 2.9821

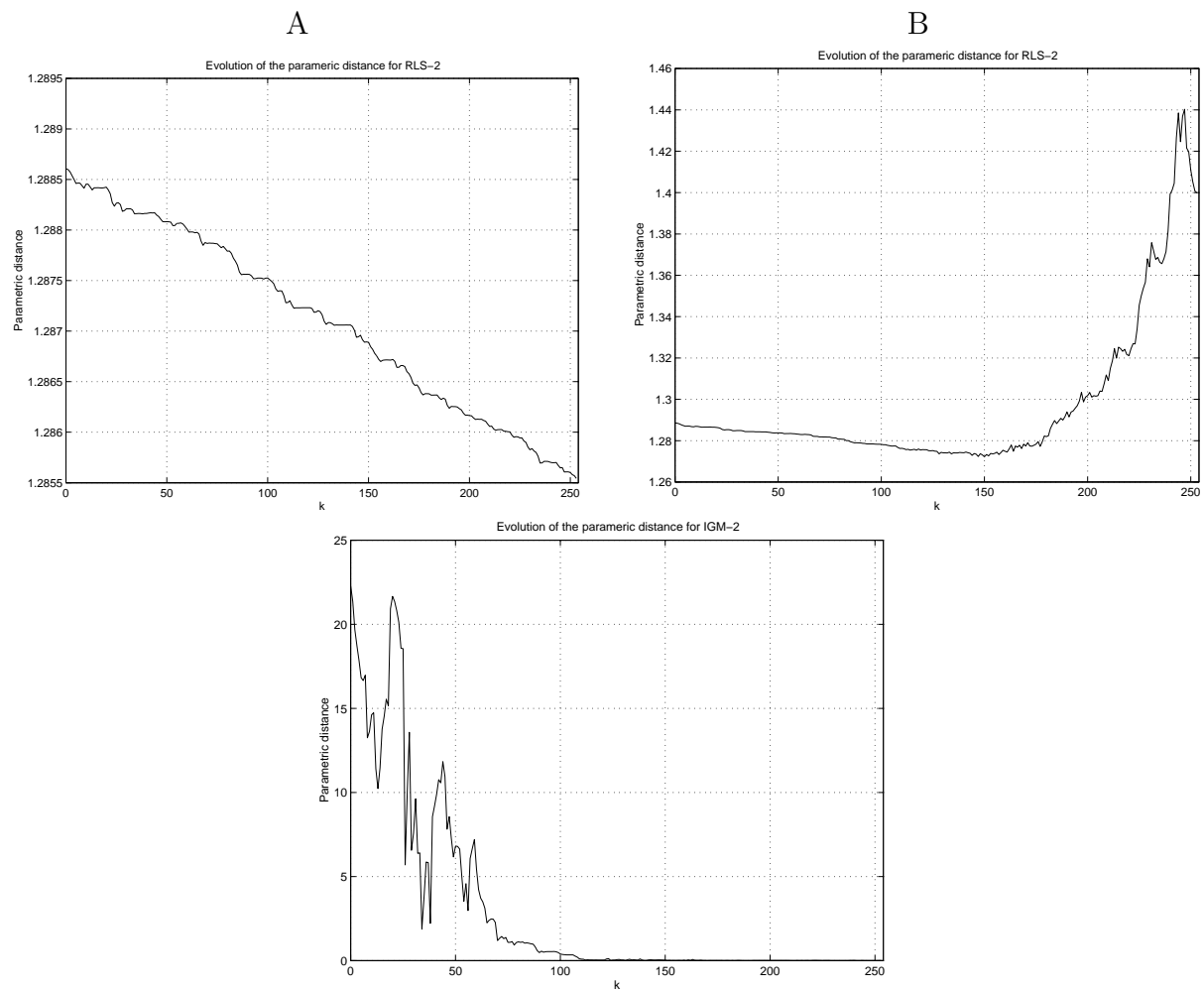


Figure B.4: Evolution of the parametric distance for one simulation run (MIMO experiment 4). Case A: $F(0) = 0.1I$. Case B: $F(0) = 2I$.

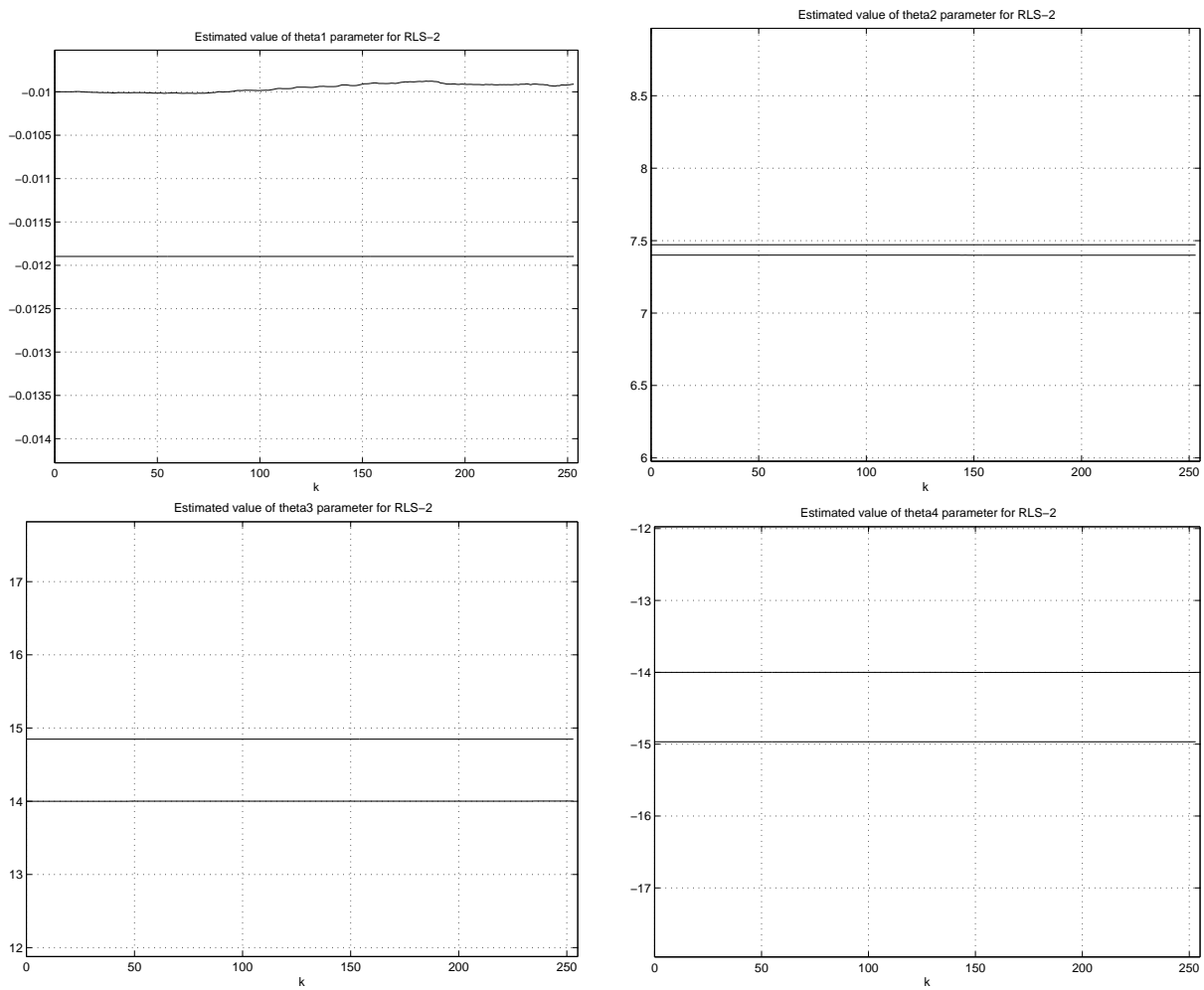


Figure B.5: Evolution of the estimated parameters for one simulation run (RLS-2, SISO experiment 4). Case A: $F(0) = 0.1I$.

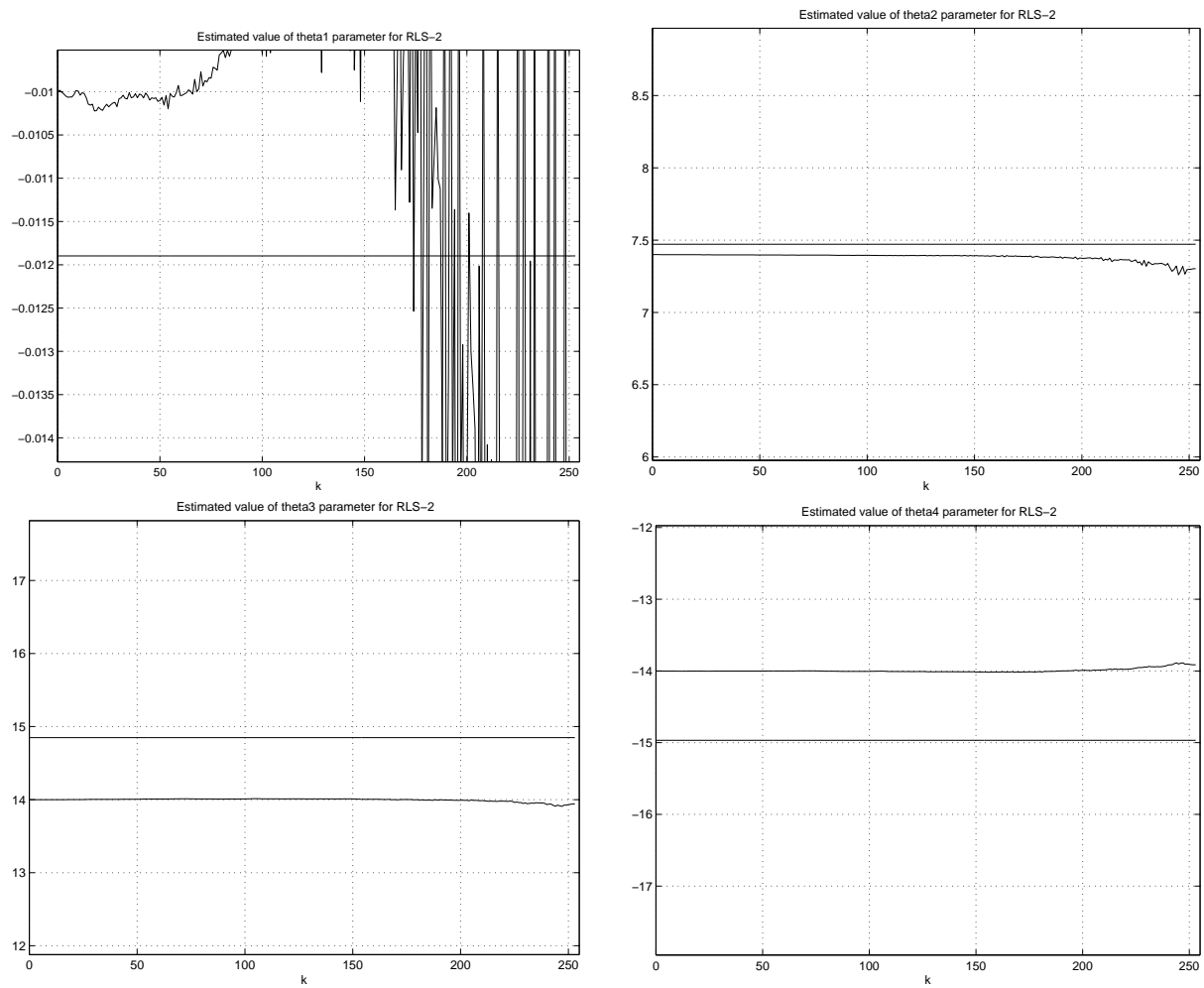


Figure B.6: Evolution of the estimated parameters for one simulation run (RLS-2, MIMO experiment 4). Case B: $F(0) = 2I$.

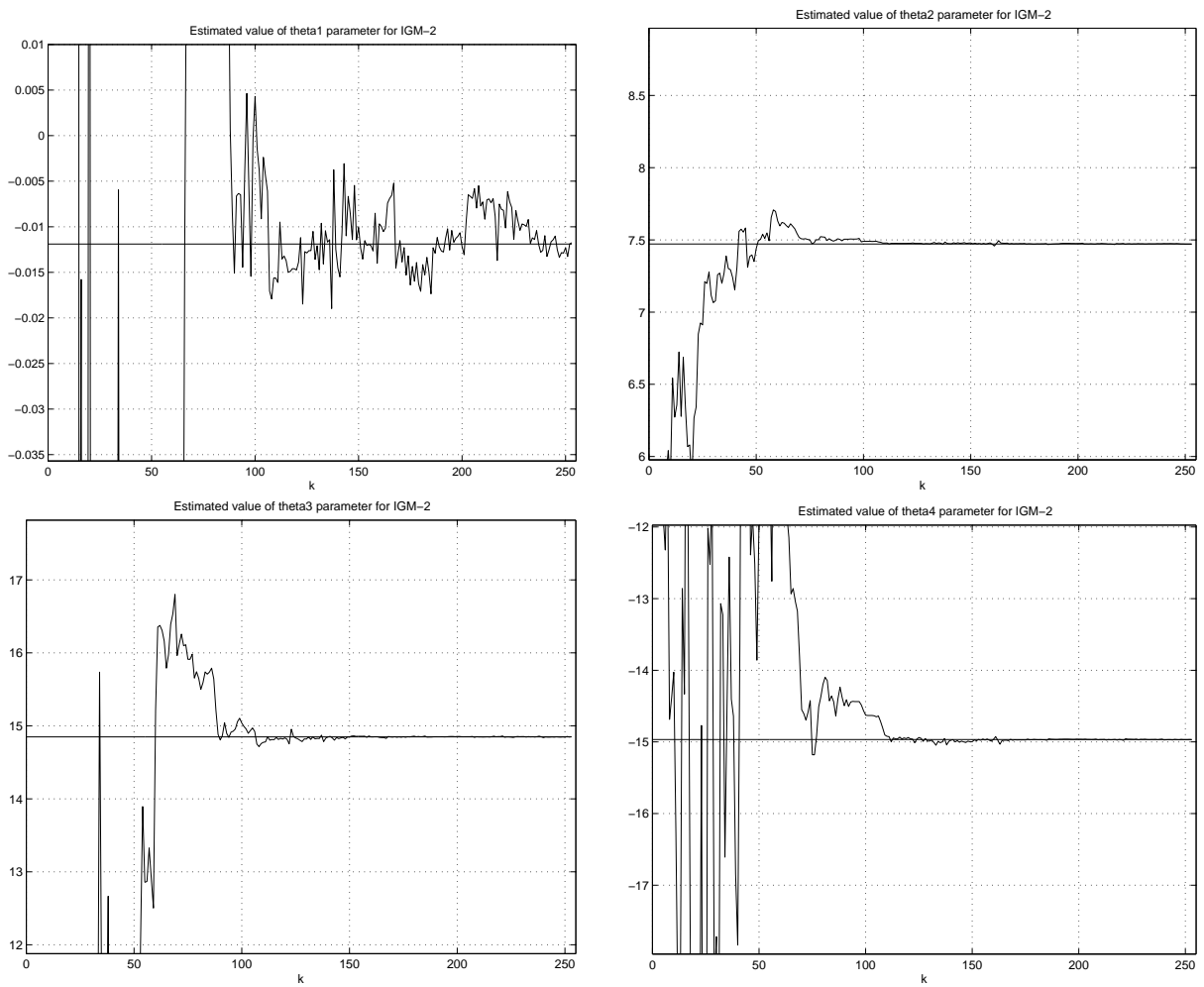


Figure B.7: Evolution of the estimated parameters for one simulation run (IGM-2, MIMO experiment 4).

Table B.7: Settings of the MIMO experiment 5.

Standard deviation of white noise $e_1(k)$	$\sigma_{e1} = 0.000001$
Standard deviation of white noise $e_2(k)$	$\sigma_{e2} = 0.000001$
Standard deviation of white noise $e_3(k)$	$\sigma_{e3} = 0.000001$
Standard deviation of white noise $e_4(k)$	$\sigma_{e4} = 0.000001$
Initial parameter vector estimate	$\hat{\theta}(0) = [0]^T$
IGM-2 scalar adaptation gain	$\alpha_2 = 1.8$
Number of simulation steps	$k_{\text{end}} = 255$
Number of simulation runs	$N = 100$

Simulation Experiment 5

The following state-space representation of the test plant \mathcal{S} is considered:

$$\begin{aligned}
 A_0(\theta) &= \begin{bmatrix} 0.9923 & 0.5807 \cdot 10^{-3} & 0.1472 \cdot 10^{-3} & 3.1350 \cdot 10^{-8} \\ 0.1229 \cdot 10^{-3} & 0.9900 & 4.5050 \cdot 10^{-8} & 0.1471 \cdot 10^{-3} \\ -102 & 7.6770 & \theta_1 & 0.6065 \cdot 10^{-3} \\ 1.616 & -131.8 & 0.6338 \cdot 10^{-3} & 0.9579 \end{bmatrix}, \\
 B_0(\theta) &= \begin{bmatrix} -1.015 \cdot 10^{-6} & 0.5643 \cdot 10^{-3} \\ 1.122 \cdot 10^{-3} & -1.133 \cdot 10^{-3} \\ -0.0119 & 7.472 \\ 14.85 & -14.97 \end{bmatrix}, \\
 C_0(\theta) &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.
 \end{aligned}$$

for the parameter vector defined by:

$$\theta = [\theta_1]^T.$$

One can see that the parameter vector is composed of one parameter which define one unknown diagonal parameter in matrix $A_0(\theta)$. The same state-space representation (canonical observer form) as in simulation experiment 4 has been used. The true value of the estimated parameter is straightforward:

$$\theta = [0.9600]$$

Table B.7 gives an overview of the simulation settings.

The simulation results are summarized in tables B.8 and B.9. The RLS-2 does not converge irrespective of the initial matrix adaptation gain setting and therefore only the results of the IGM-2 algorithm are presented. The evolution of the parametric distance for IGM-2 algorithm is shown in fig. B.8 while fig. B.7 illustrates the evolution of the individual parameter for the IGM-2 algorithm. One can see that the IGM-2 algorithm requires about 100 steps to come near the all true values.

Table B.8: Results of the MIMO experiment 5.

	$\bar{\theta}[\times 10^{-4}]$	$\bar{\sigma}[\times 10^{-4}]$
RLS-2	—	—
IGM-2	2.7441	± 38.211

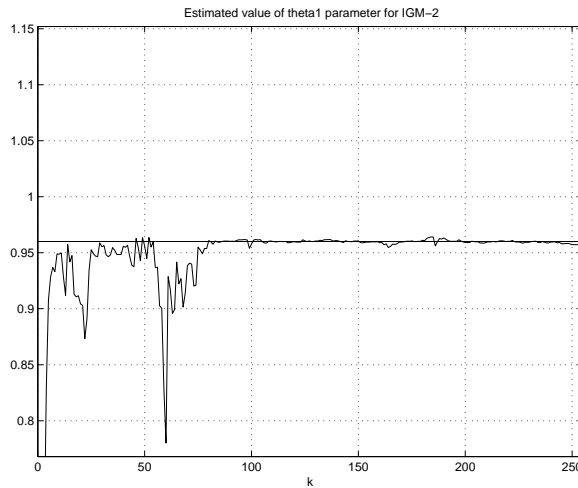


Figure B.8: Evolution of the parametric distance for one simulation run (MIMO experiment 5).

Table B.9: Results of the MIMO experiment 5, mean value and standard deviation of the individual parameter.

	j	1
IGM-2	$\bar{\hat{\theta}}_j$	0.9603
	$\sigma_j[\times 10^{-3}]$	± 3.8211

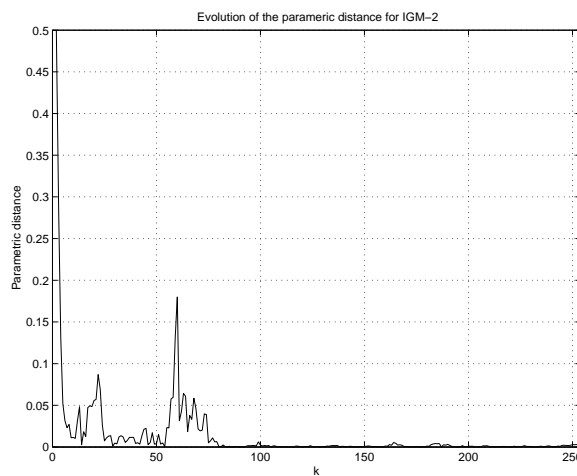


Figure B.9: Evolution of the estimated parameters for one simulation run (IGM-2, MIMO experiment 5).

Table B.10: Settings of the MIMO experiment 6.

Standard deviation of white noise $e_1(k)$ (A)	$\sigma_{e1} = 0$
Standard deviation of white noise $e_2(k)$ (A)	$\sigma_{e2} = 0$
Standard deviation of white noise $e_3(k)$ (A)	$\sigma_{e3} = 0$
Standard deviation of white noise $e_4(k)$ (A)	$\sigma_{e4} = 0$
Standard deviation of white noise $e_1(k)$ (B)	$\sigma_{e1} = 0.000001$
Standard deviation of white noise $e_2(k)$ (B)	$\sigma_{e2} = 0.000001$
Standard deviation of white noise $e_3(k)$ (B)	$\sigma_{e3} = 0.000001$
Standard deviation of white noise $e_4(k)$ (B)	$\sigma_{e4} = 0.000001$
Initial parameter vector estimate	$\hat{\theta}(0) = [0]^T$
IGM-2 scalar adaptation gain	$\alpha_2 = 0.3$
Number of simulation steps	$k_{\text{end}} = 255$
Number of simulation runs	$N = 100$

Simulation Experiment 6

The following state-space representation of the test plant \mathcal{S} is considered:

$$\begin{aligned}
 A_0(\theta) &= \begin{bmatrix} 0.9923 & 0.5807 \cdot 10^{-3} & 0.1472 \cdot 10^{-3} & 3.1350 \cdot 10^{-8} \\ 0.1229 \cdot 10^{-3} & 0.9900 & 4.5050 \cdot 10^{-8} & 0.1471 \cdot 10^{-3} \\ -102 & 7.6770 & 0.9600 & \theta_1 \\ 1.616 & -131.8 & 0.6338 \cdot 10^{-3} & 0.9579 \end{bmatrix}, \\
 B_0(\theta) &= \begin{bmatrix} -1.015 \cdot 10^{-6} & 0.5643 \cdot 10^{-3} \\ 1.122 \cdot 10^{-3} & -1.133 \cdot 10^{-3} \\ -0.0119 & 7.472 \\ 14.85 & -14.97 \end{bmatrix}, \\
 C_0(\theta) &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.
 \end{aligned}$$

for the parameter vector defined by:

$$\theta = [\theta_1]^T.$$

One can see that the parameter vector is composed of one parameter which define one unknown non-diagonal parameter in matrix $A_0(\theta)$. The same state-space representation (canonical observer form) as in simulation experiment 4 has been used. The true values of the estimated parameters are straightforward:

$$\theta = [0.6065 \cdot 10^{-3}]$$

Table B.10 gives an overview of the simulation settings. There are two cases (A, B) which have been studied using IGM-2 algorithm. The difference comes from two different output disturbances $e(k)$. The standard deviations of white noises σ_{ei} are zero in the first case (A) while they are equal to 0.000001 in the second case (B).

The simulation results are summarized in tables B.11 and B.12. The RLS-2 does not converge irrespective of the initial matrix adaptation gain setting and therefore only the

Table B.11: Results of the MIMO experiment 6.

	$\bar{\theta}[\times 10^{-4}]$	$\bar{\sigma}[\times 10^{-4}]$
RLS-2	—	—
IGM-2 (A)	0.0237	$\pm 1.1986 \cdot 10^{-14}$
IGM-2 (B)	0.0776	± 1.9046

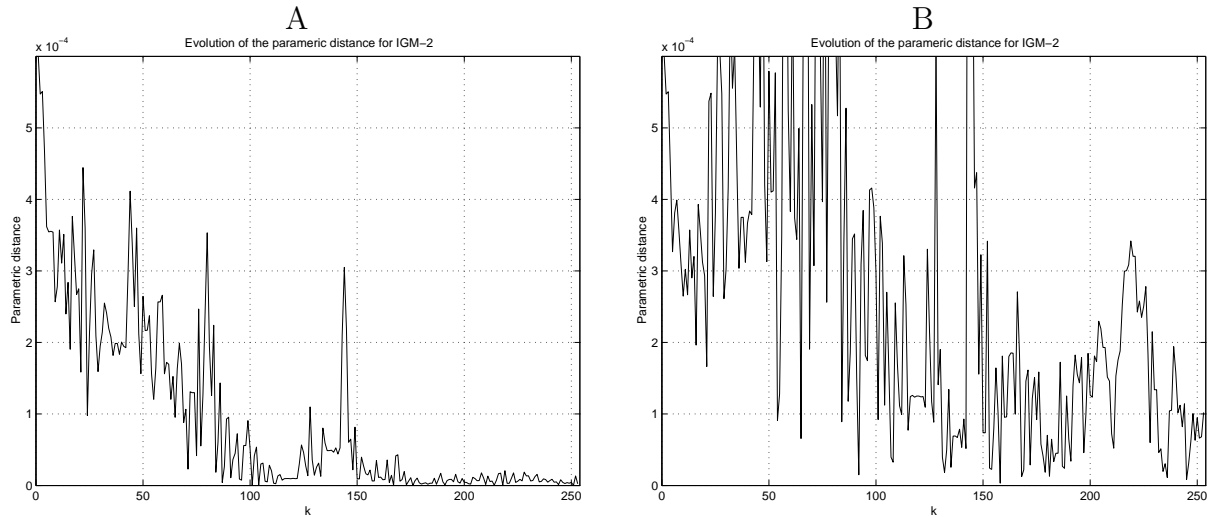


Figure B.10: Evolution of the parametric distance for one simulation run (MIMO experiment 6). Case A: $\sigma_{ei} = 0$. Case B: $\sigma_{ei} = 0.000001$.

results of the IGM-2 algorithm are presented. The evolution of the parametric distance for IGM-2 algorithm is shown in fig. B.10 while fig. B.11 illustrates the evolution of the individual parameter for the IGM-2 algorithm. One can see that the IGM-2 algorithm is very noise sensitive and in additional, there is also small bias if the zero output noise disturbances have been used (case A).

Table B.12: Results of the MIMO experiment 6, mean value and standard deviation of the individual parameter.

	j	1
IGM-2(A)	$\hat{\theta}_j$ $\sigma_j[\times 10^{-3}]$	$0.6040 \cdot 10^{-3}$ $\pm 0.1986 \cdot 10^{-14}$
IGM-2(B)	$\bar{\theta}_j$ $\sigma_j[\times 10^{-3}]$	$0.5986 \cdot 10^{-3}$ ± 0.1904

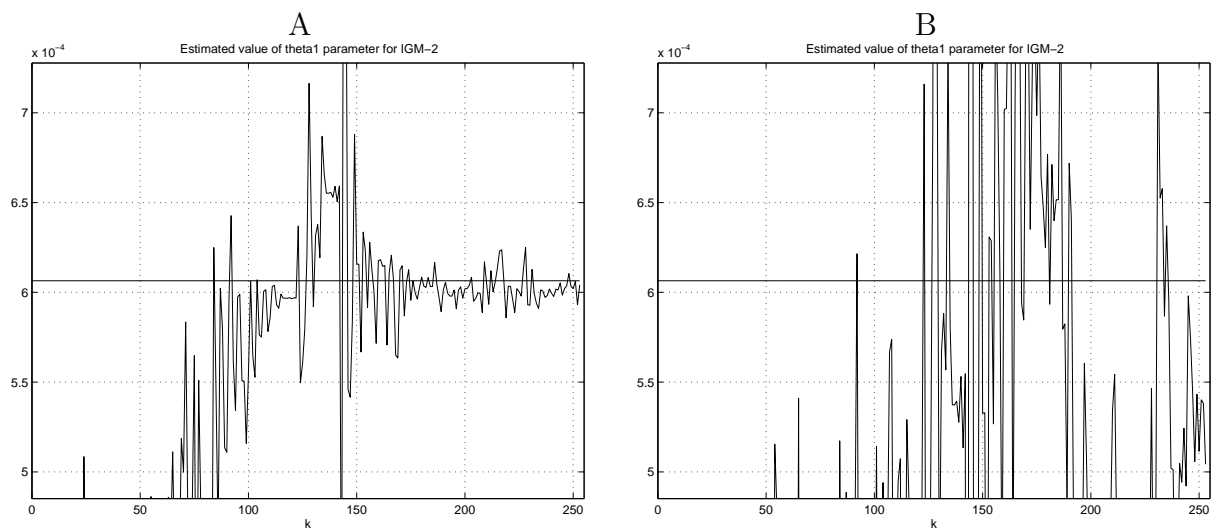


Figure B.11: Evolution of the estimated parameter for one simulation run (IGM-2, MIMO experiment 6). Case A: $\sigma_{ei} = 0$. Case B: $\sigma_{ei} = 0.000001$.

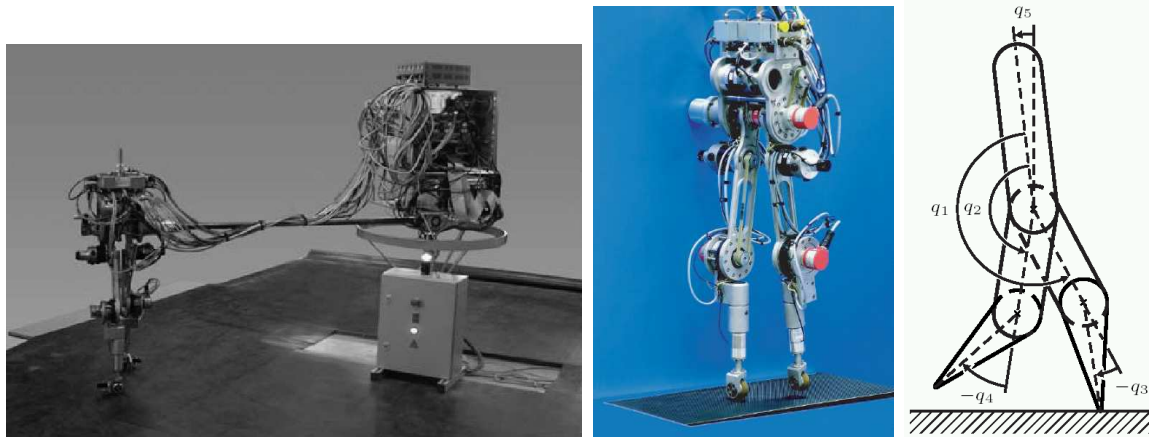


Figure B.12: The biped prototype RABBIT's experimental setup. Schematics of the prototype RABBIT with measurement conventions are illustrated on the right side.

B.3 Real-time Experiment for SISO System

The bipedal robot (RABBIT), see fig. B.12, constructed jointly by several research laboratories by the French project *Commande de Robots à Pattes* of the CNRS—GdR Automatique [Buc], [CAA⁺03], consists of five-link.

RABBIT weights approximately 32 kg and is 1.43 m tall. Its five links are connected by revolute joints that form two symmetric legs and a torso, see fig. B.12. Actuators supply torque between each of the four internal joints: One at each knee and one between the torso and each femur. All actuators are identical and capable of producing a peak torque of 150 Nm, though, for safety reasons, the peak torque was limited to 70 Nm during the experiments reported here. To prevent motions in the frontal plane, RABBIT was constructed with a boom attached at the hip, see fig. B.12. RABBIT has no feet and no means of supplying actuation between the stance leg end and the ground.

To obtain configuration information, encoders are located at each internal joint giving the robot's shape, and between the boom and hip giving the robot's orientation with respect to a world frame. Binary contact switches located at the leg ends are used to detect whether or not a leg is in contact with the walking surface.

The experimental measurement was performed for one knee of the RABBIT, the angle $q_3(t)$ has been chosen as the controlled system output, see fig. B.12. In the actual implementation a decoupled PD controller with friction compensation was used, see [CAA⁺03]. The PD based feedback was chosen over sliding mode, or finite-time converging controller because of its robustness to noise and uncertainty. The controller was implemented on the dSpace DS1103 system running with a sampling period $T_s = 1.5$ ms, *i.e.* a sampling frequency $f_s = 666,6$ Hz.

The identified system \mathcal{S} can be described by the following differential equations of Direct Current (DC) motor:

$$\begin{aligned} J_s q_3(t) &= \tau_e = k_1 i_a(t), \\ v_\omega(t) &= k_2 \dot{q}_3(t), \\ i_a(t) &= \frac{u_a(t) - k_2 \dot{q}_3(t)}{R_a}. \end{aligned}$$

where $u_a(t)$ denotes the armature voltage, $q_3(t)$ the output angle, J_s the inertia of the shaft, τ_e the electrical torque, $i_a(t)$ the armature current, $v_\omega(t)$ the angular velocity, R_a the armature resistance, k_1, k_2 constants. Thus, the model input $u(t) = u_a(t)$ and the model output $y(t) = q_3(t)$.

This model can be converted to state-space form by introducing:

$$\begin{aligned}x_1(t) &= q_3(t), \\x_2(t) &= \dot{q}_3(t).\end{aligned}$$

The model can be written as

$$\begin{aligned}\frac{d}{dt} \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix} &= \begin{bmatrix} 0 & 1 \\ 0 & -\frac{k_1 k_2}{R_a} \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} + \begin{bmatrix} 0 \\ \frac{k_1}{R_a} \end{bmatrix} u(t), \\ y(t) &= \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}.\end{aligned}$$

The continuous-time transfer function can be written as:

$$\mathcal{S}: \quad G_S(s) = \frac{\frac{1}{k_2}}{s\left(\frac{R J_s}{k_1 k_2} s + 1\right)}. \quad (\text{B.1})$$

This linear model corresponds to the second order discrete-time model:

$$\mathcal{S}: \quad G_S(q^{-1}) = \frac{b_1 q^{-1} + b_2 q^{-2}}{1 + a_1 q^{-1} + a_2 q^{-2}}. \quad (\text{B.2})$$

A pseudo-random binary sequence (PRBS), generated by a 10-bit register, added to the controller output u is taken as excitation signal w . Its value alternates between +20 Nm and -20 Nm and its clock frequency is equal to $f_s/20$.

The measured data have been acquired from the closed-loop operation and they have been used to identify unknown parameters of discrete-time model (B.2). Only the non-recursive identification algorithm has been used to identify three parametric models of the second order, namely the Output-Error model (OE), AutoRegresive with eXogenous variable model (ARX) and AutoRegresive with eXogenous variable and Moving Average noise model (ARMAX). A generalized model structure of SISO models is given by, [Lju99]:

$$A(q^{-1})y(k) = \frac{B(q^{-1})}{F(q^{-1})}u(k) + \frac{C(q^{-1})}{D(q^{-1})}e(k), \quad (\text{B.3})$$

where $A(q^{-1}), B(q^{-1}), C(q^{-1}), D(q^{-1})$ and $F(q^{-1})$ are polynomials representing the model input-output transfer operator.

The identified models obtained by using the developed state-space identification algorithms and the basic CLOE algorithm for the black-box model are not usable in this case because these algorithms need higher sampling period of the control algorithm. It is difficult to increase this sampling period because of RABBIT complexity. Nevertheless, the identified OE, ARX and ARMAX models are usable for the following controller redesign and therefore they are presented in the thesis.

To evaluate the results the following criteria have been used: The comparison of the measured and simulated model output is given by the best fits in table B.13.

Table B.13: Statistical parameters of measured and simulated model output.

Model structure name	Best fits
ARX221	73.1
ARMAX2221	72.6
OE221	64.6

The autocorrelation of residuals for output and crosscorrelation for input and output residuals of identified models are illustrated in fig. B.13 while their frequency responses are shown in fig. B.14. It can be concluded that the Output-Error model (OE) is the best identified model while the AutoRegressive with eXogenous variable model (ARX) has the worst performance. AutoRegressive with eXogenous variable and Moving Average noise model (ARMAX) is very close to the Output-Error model.

Numerical values of corresponding models and their variations are summarized as follows:

ARX221:

$$A(q) = 1 - 1.861(\pm 0.002895)q^{-1} + 0.8654(\pm 0.002898)q^{-2}$$

$$B(q) = -2.706 \times 10^{-7}(\pm 1.078 \times 10^{-7})q^{-1} + 2.695 \times 10^{-6}(\pm 1.14 \times 10^{-7})q^{-2}$$

ARMAX2221:

$$A(q) = 1 - 1.908(\pm 0.001186)q^{-1} + 0.9131(\pm 0.001184)q^{-2}$$

$$B(q) = -5.24 \times 10^{-7}(\pm 6.667 \times 10^{-8})q^{-1} + 2.544 \times 10^{-6}(\pm 7.099 \times 10^{-8})q^{-2}$$

$$C(q) = 1 - 0.8708(\pm 0.009175)q^{-1} + 0.308(\pm 0.009096)q^{-2}$$

OE221:

$$B(q) = -7.577 \times 10^{-7}(\pm 6.78 \times 10^{-7})q^{-1} + 2.688 \times 10^{-6}(\pm 7.286 \times 10^{-7})q^{-2}$$

$$F(q) = 1 - 1.91(\pm 0.005384)q^{-1} + 0.9151(\pm 0.005315)q^{-2}$$

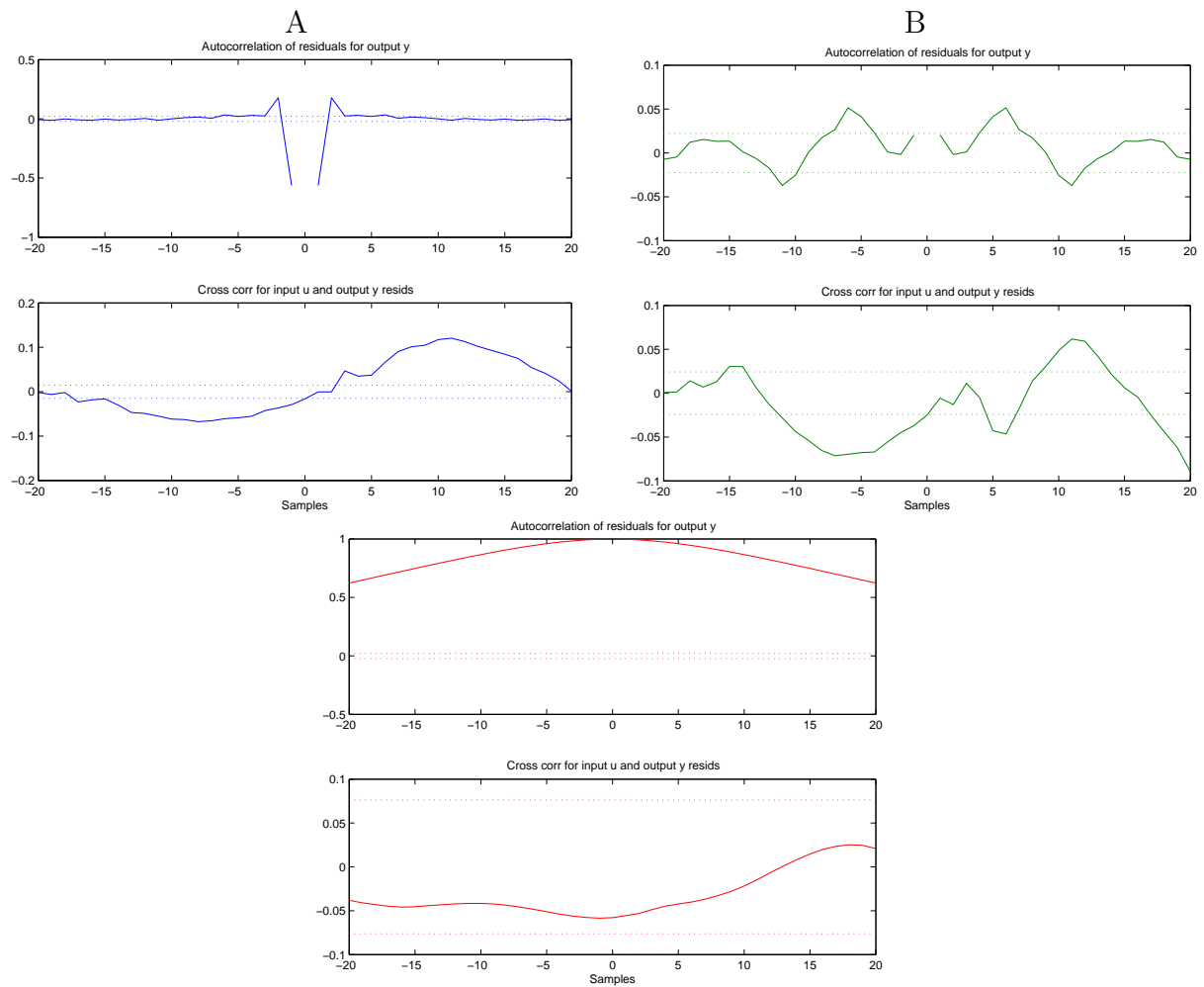


Figure B.13: Autocorrelation of residuals for output and crosscorrelation for input and output residuals (SISO experiment 7). Case A: ARX221. Case B: ARMAX2221. Case C: OE221.

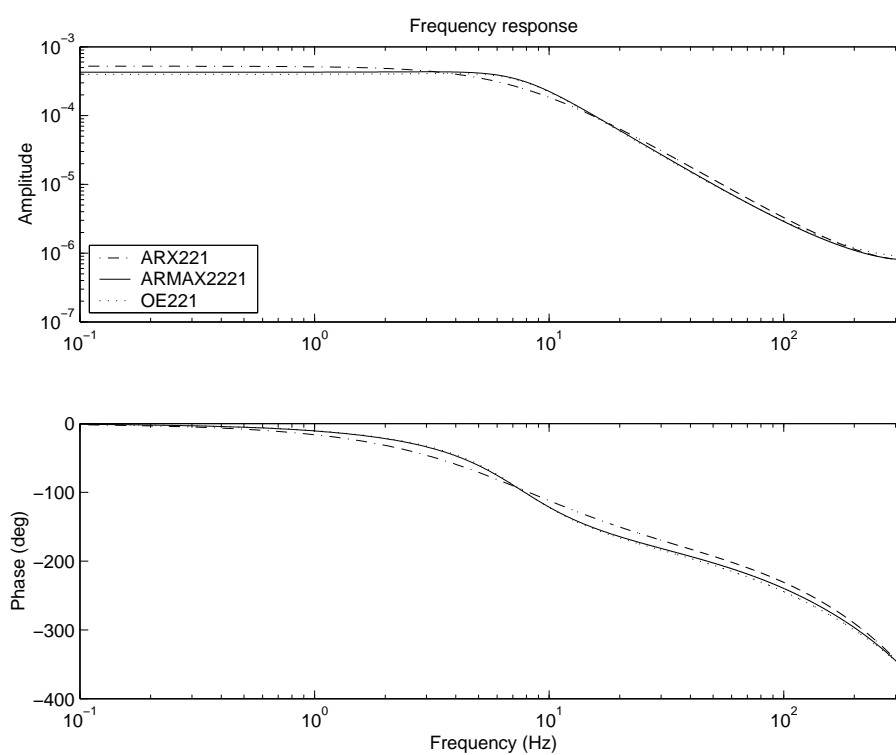


Figure B.14: Frequency response of identified models.

Modelling of the Focus Error Signal Generation

C

C.1 Introduction

As discussed in section 3.3.3, this chapter concerns modelling the non-linear characteristics of a photodetector (so called the S-curve) in a DVD player, using optical theory.

The photodetector is used in the DVD player to generate the focus error signal which in turn has to be controlled at a minimum level. An analytical and a numerical photodetector models are developed here, based on the astigmatic method and on opto-geometrical analysis. The influence of model parameters on the focus error signal is discussed. To estimate the unknown model parameters a curve fitting method is applied by using measured data from an industrial DVD-video player developed in STMicroelectronics laboratories. Model quality is illustrated by a comparison with the real focus error signal acquired during the start-up procedure (from open loop to closed loop).

The chapter is organized as follows: In section C.2, a brief state of the art is presented. Section C.3 describes briefly the principle of the astigmatic method in comparison with the section 3.3.3. S-curve modelling is presented in section C.4. The detailed S-curve analysis is given in section C.5. Section C.6 explains the unknown parameters estimation and section C.7 shows the possible applications of photodetector characteristic. Finally, conclusions are presented in section C.8.

C.2 State of the Art

In a DVD player, the quality of the sound and image is obtained with the help of well-tuned control loops to maintain the laser beam position on the desired track.

The focus positioning is one of the most important control loops which is designed to minimize the focus error of the laser beam. The focus error signal is generated by a photodetector. The present chapter treats the focus error signal generation by the photodetector, namely the modelling of the photodetector non-linear characteristics. This characteristic is useful firstly during the start-up procedure (when the data layer is searched for) and secondly during normal playing mode to tune the controller gain for the desired closed-loop bandwidth of the focus loop.

Despite the fact many references exist about the photodetector construction and functionality, (*e.g.* like [BBH⁺85], [Isa85], [Poh95], [Bra98] and [Sta98]), there is no reference, from our knowledge, about models of the photodetector characteristic.

In case of DVD players, there are not so many types of photodetectors like for CD, and the most widely used method is based on the astigmatism principle proposed by [BLPIC76]. In this case, there is no model available also to describe the photodetector characteristics.

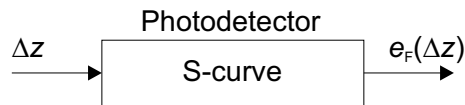


Figure C.1: Block-scheme of the focus error generation model.

To present the photodetector characteristics briefly, recall that the photodetector is composed by four photodiodes which are used to measure the vertical spot position error Δz from the voltages acquired from its photodiodes. The photodetector provides a non-linear bipolar characteristic (the *S-curve*) between the input vertical spot position error Δz and output focus error signal $e_F(\Delta z)$, as shown in fig. C.1.

In practice, a very rough approximation of the S-curve by a linear function is used, *i.e.* $e_F(\Delta z) = k \cdot \Delta z$, where k is the constant value. Nevertheless, this linear approximation is valid only for small changes of the vertical spot position error Δz and this is the case for the normal playing mode only but not for the start-up procedure. The development of a more complex non-linear model (analytical or numerical) of the S-curve involves a large amount of optical theory of the astigmatic principle.

The principle of the astigmatic method is often explained schematically, as in [NO92], [BBH⁺85], [Poh95], [Sin88], [Sta98] and [Wil94]. The focus error generation modelling has been discussed in more detail in [CGLL84] and [Man87] where the analysis based on the diffraction theory and the full Gaussian optics considerations has been shown.

In this chapter, the proposed models are then more complex than classically used linear model described above with one parameter k (the slope of the S-curve). The analytical model has four parameters and the numerical one has five parameters. More complex models (with much more parameters) have been developed by using the diffraction theory in order to take into account crosstalk effects, see *e.g.* [Hop79], [MCW⁺96], [Mil98], [MU99], [UAM⁺00] and [UM01].

The model complexity, however, is an important issue because the unknown model parameters have to be estimated. The proposed models in this chapter make a compromise between complexity and accuracy.

A curve fitting procedure, based on the measured data obtained on an industrial DVD-video player developed by STMicroelectronics, is used to estimate of the unknown model parameters.

Possible applications of the presented model are, for example, in the focus processing procedure or even for new photodetector design.

C.3 Principle of the Astigmatic Method

As presented in section 3.3.3, the focus error signal $e_F(\Delta z)$ which has to be minimized by the control system is obtained when some asymmetry element (the cylindrical lens (12) in our case) is present in the optical path (10, 7, 5, 4, 3, 12, 13) of the return beams (11), see fig. 3.2.

The principle of the astigmatic method is simply illustrated in fig. 3.8 and it is based upon an optical aberration, called *astigmatism*. Fig. 3.8 is a zoom of fig. 3.2 and it illustrates the simplified model of the return optical path from the focusing area (15) to the detection area (16). An astigmatic image is rotated with a respect to its optical axis

z and a focus error signal $e_F(\Delta z)$ can be extracted if a special arrangement of the four photodiodes A, B, C, D is used. The focus error signal $e_F(\Delta z)$ has been already defined by (3.8).

The vertical spot position error $\Delta z = z_{\text{obj}} - z_{\text{ref}}$ is defined as the difference between the objective lens position z_{obj} from the disk information layer and its reference value z_{ref} , see fig. 3.13. Using optical theory, one can verify that $z_{\text{ref}} \equiv f'_{\text{obj}}$ and $f'_{\text{obj}} = -f_{\text{obj}}$, where f_{obj} is the objective lens focal length.

Since this chapter involves the optical theory and $z_{\text{ref}} \equiv f'_{\text{obj}}$, the vertical spot position error Δz is noted by $\Delta z = z_{\text{obj}} - f'_{\text{obj}}$ in this chapter.

C.4 Photodetector Model by Astigmatic Method

The modelling problems and their solutions have been presented briefly in [HBVSdF02] and [HBVSd03] where only some ideas have been introduced. A detailed model description based on the opto-geometrical analysis is given here. Moreover, the presented numerical model is more general.

An important range of the vertical spot position error Δz from “in focus position” could be $\pm 150 \mu\text{m}$ according to a DVD standard. We will, however, deal with much smaller range where the S-curve can be found out. We concentrate to the range $|\Delta z| = |z_{\text{obj}} - |f_{\text{obj}}|| \in \langle 0, 20 \rangle \mu\text{m}$ because it is fully sufficient for the practical use.

Before going too far in treating the optics theory, the sign convention has to be defined. We have considered usual sign convention where the light rays pass through the optical system from the left side to right side, and assumed this direction as positive.

As presented in section 3.3.1, the laser beams pass through the lenses in the optical pick-up unit. These lenses can be described by thin spherical lenses and a diverging cylindrical lens. Therefore, the description of their behavior in air is discussed in the following sections C.4.1 and C.4.2.

C.4.1 Thin Lens

A thin lens can be described by the Lensmaker’s equation, see *e.g.* [BW87]:

$$\frac{1}{z'_i} - \frac{1}{z_o} = \frac{1}{f'} = -\frac{1}{f}, \quad (\text{C.1})$$

where z_o is the objective distance, z'_i is the image distance, f is the objective focal length and f' is the image focal length of the lens. Fig. C.2 shows thin lens and its object, image distance and focal lengths. Equation (C.1) is valid if $n = n'$ where n is refractive index of objective space and n' is refractive index of image space. It means that $f = -f'$ for lens in air. Theory usually assumes thin lens in air but if it is not then $f'/f = -n'/n$ has to be used.

An objective space defines the space in front of lens where the light source is placed, while an image space is defined as the space where the light source image could be observed, [BW87]. For the thin spherical lens, the image of the parallel rays with the optical axis z is so called a focal point that is placed in the focal distance behind of lens, as illustrated in fig. C.3. Nevertheless, for the cylindrical lens, the image is not one focal point but so called a focal line, as it will be presented in section C.4.2.

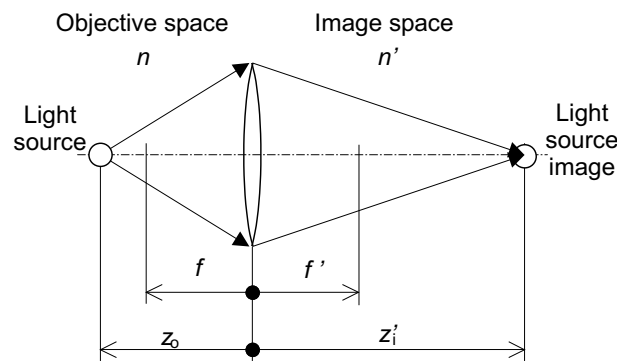


Figure C.2: The thin lens and its Lensmaker's equation.

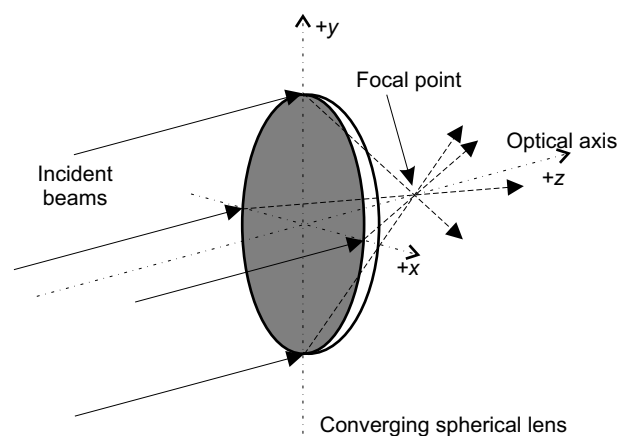


Figure C.3: The converging spherical lens with coordinates system. If incident beams are parallel to the optical axis z than the focal point is placed in lens focal length behind of the spherical lens.

If we assume the light point-source in object distance then we can observe circular shape on the screen in the image distance (if the screen is perpendicular to the optical axis). If the screen is placed just in image focal length f' from thin lens then only one point is displayed on the screen. This projection appears as one point that is why is called stigmatic projection. Moreover, the approximation of light flux distribution on screen can be assumed as a uniform.

C.4.2 Diverging Cylindrical Lens

A few accessible references exist that describe the behavior of cylindrical lens, see [Mah98], [BW87]. Therefore we have to imagine the function of the cylindrical lens in three dimensional space in order to trace two focal lines creation. The behavior of diverging cylindrical lens is presented because this type is used in OPUs.

Following mathematical description requires some coordinate system that is presented in fig. C.4.

If we assume incident beam of parallel rays with optical axis z then the refraction is in one direction. But only rectangular spot shapes are displayed on the screen that is

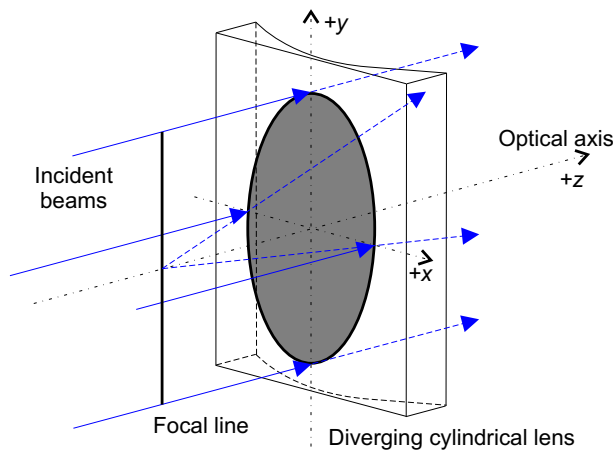


Figure C.4: The diverging cylindrical lens with coordinates system. If we assume parallel incident beams with optical axis z than apparent focal line is placed in focal length in front of the cylindrical lens.

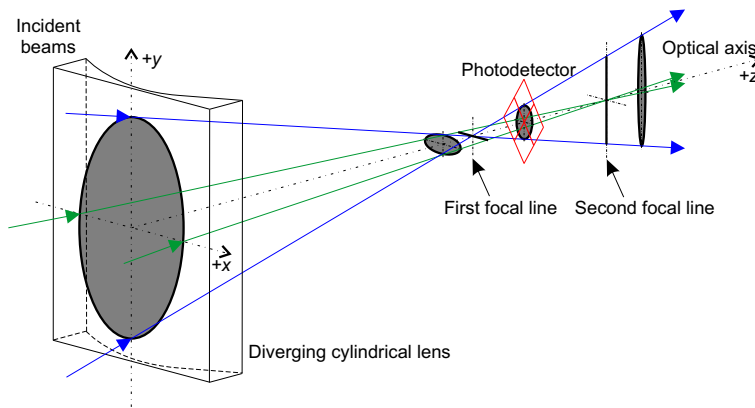


Figure C.5: The marginal rays of the cylindrical diverging lens if light source (disk) is in focus.

placed perpendicular to optical axis because passing rays are divergent. Apparent image of the focal line is placed in front of the cylindrical lens. Fig. C.4 shows some marginal laser beams and apparent focal line in this case. The focal line is parallel to axis y too.

On the other hand, if we assume non-parallel incident rays with optical axis z that have to be *convergent enough* then the situation becomes even more complicated. Fig. C.5 shows this situation when the disk is in focus, $\Delta z = 0$. The detector position, spot size and spot shape are also presented here.

In addition, if the laser beam, whose cross section is a true circle, enters the cylindrical lens, it will be focused in the transverse direction only. As a result, the laser beam passed through the lens will be elliptical in cross section. The eccentricity of ellipse will vary with the distance from the lens or with the angle of the incident rays falling on the cylindrical lens. This later case is a key idea of the usage of the cylindrical lens to form the light spot intensity distribution on the photodetector. Thus, the vertical spot position error

Δz can be measured from four voltages V_A, V_B, V_C, V_D which measure the light intensity distribution on the photodetector, as explained in section 3.3.1.

C.4.3 Optical System Separation

Knowledge of the four voltages V_A, V_B, V_C, V_D generation must be explained to model a non-linear characteristic, given by (3.8). Since they measure the laser spot energy falling on the four quadrant photodetector, the laser spot shape creation and its energy calculation are presented beforehand.

In addition, to calculate spot shape and spot energy on the photodetector, some optical theory has to be used, because the reflected laser beam from the disk (carrying the vertical spot position error information Δz) is passing through the few lenses in optical pick-up unit.

The main idea of model creation is expressed by the following sentences: Since the whole optical system contains cylindrical lens (causing the astigmatism), we can assume that *the whole optical system is separated into two subsystems with lenses, which are easier to describe*. The separation of the whole optical system into the two orthogonal planes yz and xz allows to use the theory of a system formed by two centered thin lenses, [BW87]. Figs. C.6 and C.7 show the separated optical system with thin lens approximation at planes yz and xz where the introduced variables have the following meaning:

- f_{obj} : objective focal length of the objective lens,
- z_{obj} : distance between the disk and the objective lens,
- d_3 : distance between the objective lens and the collimator,
- d_4 : distance between the collimator and the cylindrical lens,
- z_{det} : distance between the collimator and the photodetector,
- z_1 : distance between the collimator and the second focal line,
- z_2 : distance between the collimator and the first focal line,
- f_{L1} : distance between the collimator and the first focal line if disk is in focus,
- f_{L2} : distance between the collimator and the second focal line if disk is in focus.

C.4.4 Position of the Focal Lines

The lenses are specified by its objective space, image space and focal distance, [BW87]. For the thin lens, the image of the parallel rays with the optical axis z at objective space is one focal point at focal distance of the image space. Nevertheless, for the cylindrical lens, the image is not one focal point but so called focal line. Situation is more complicated for the cylindrical lens if the incident rays are converging or diverging because then there are two distances z_1, z_2 in the image space where the focal line could be found, (see figs. C.6 and C.7). Therefore, the focal lines are defined as light lines that can be observed on the screen in the image space of the cylindrical lens. More information about cylindrical lenses can be found in [BW87].

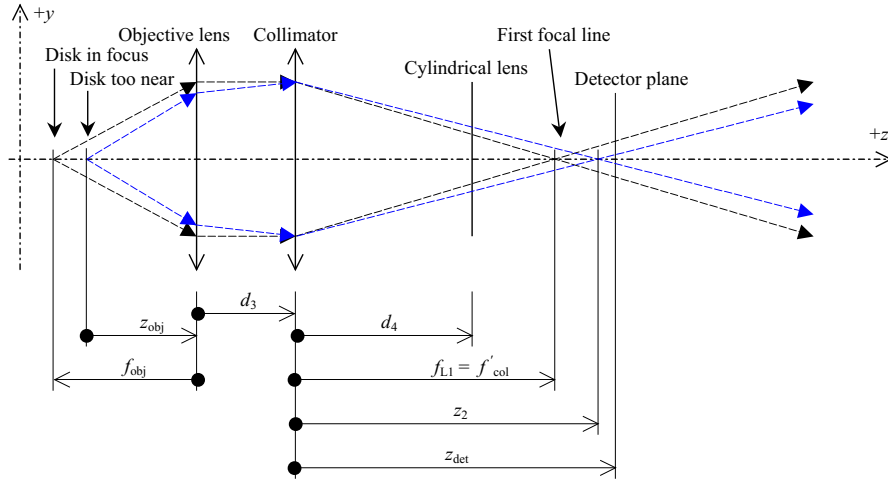


Figure C.6: Arrangement in the yz plane. Diverging cylindrical lens is approximated by air.

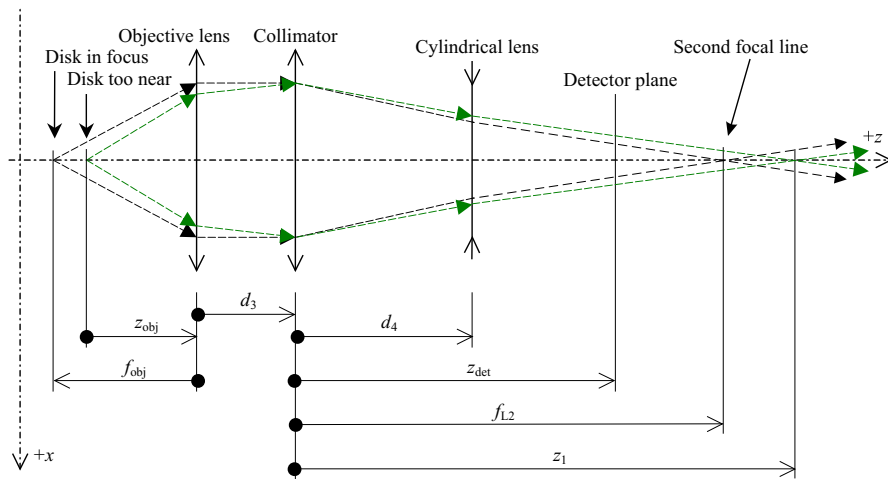


Figure C.7: Arrangement in the xz plane. Diverging cylindrical lens is approximated by diverging thin lens.

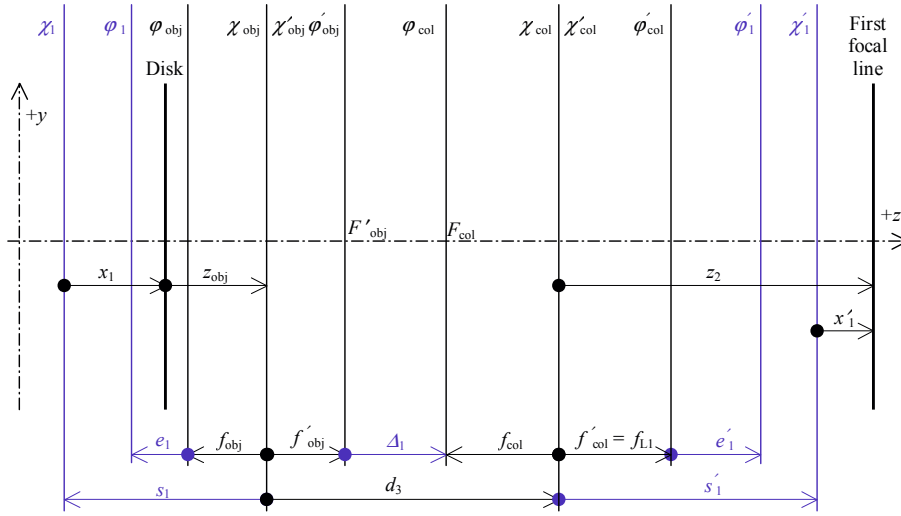


Figure C.8: System formed by two centered thin lens in the yz plane.

Knowledge of two focal lines position $z_1(z_{obj})$, $z_2(z_{obj})$ is the first condition to obtain the S-curve model because they establish the laser spot size on the photodetector, see figs. C.6 and C.7.

Firstly, a dependence between distance z_2 and z_{obj} is derived in the yz plane. Fig. C.8 shows a solution in the yz plane where the notation has the following meaning:

- F'_{obj}, F_{col} : focal points,
- χ_{obj}, χ'_{obj} : unit planes of the objective lens,
- χ_{col}, χ'_{col} : unit planes of the collimator,
- χ_1, χ'_1 : unit planes of the “resulting” lens,
- $\varphi_{obj}, \varphi'_{obj}$: focal planes of the objective lens,
- $\varphi_{col}, \varphi'_{col}$: focal planes of the collimator,
- φ_1, φ'_1 : focal planes of the “resulting” lens,
- Δ_1 : optical interval,
- e_1 : distance of the “resulting” focal plane φ_1 from objective focal plane φ_{obj} of the first lens,
- e'_1 : distance of the “resulting” focal plane φ'_1 from image focal plane φ'_{col} of the second lens,
- s_1 : distance of the “resulting” first unit (principal) plane χ_1 from the first unit plane χ_{obj} of the first lens,
- s'_1 : distance of the “resulting” second unit (principal) plane χ'_1 from last unit plane χ'_{col} of the second lens,

- x_1, x'_1 : objective and image distance of the “resulting” lens,
- f'_1 : “resulting” objective focal length in yz plane.

Let the objective lens focal length $f'_{\text{obj}} = -f_{\text{obj}}$ and the collimator focal length $f_{\text{col}} = -f'_{\text{col}} = f_{L1}$ then the model parameters are given by the following relationships:

$$\Delta_1 = d_3 - f'_{\text{obj}} + f_{\text{col}} = d_3 + f_{\text{obj}} - f'_{\text{col}}, \quad (\text{C.2})$$

$$s_1 = \frac{d_3 f'_{\text{obj}}}{f'_{\text{obj}} + f'_{\text{col}} - d_3}, \quad (\text{C.3})$$

$$s'_1 = -\frac{d_3 f'_{\text{col}}}{f'_{\text{obj}} + f'_{\text{col}} - d_3}, \quad (\text{C.4})$$

$$e_1 = -\frac{f'_{\text{obj}}{}^2}{\Delta_1}, \quad (\text{C.5})$$

$$e'_1 = \frac{f'_{\text{col}}{}^2}{\Delta_1}, \quad (\text{C.6})$$

$$f'_1 = -\frac{f'_{\text{obj}} f'_{\text{col}}}{\Delta_1}. \quad (\text{C.7})$$

Using the Lensmaker’s equation (C.1), one can write:

$$\frac{1}{-s'_1 + z_2} - \frac{1}{-s_1 - z_{\text{obj}}} = \frac{1}{f'_1}. \quad (\text{C.8})$$

Then the dependence between distances z_2 and z_{obj} can be obtained from (C.8) as follows:

$$\boxed{z_2 = \frac{1}{\frac{1}{f'_1} + \frac{1}{-s_1 - z_{\text{obj}}}} + s'_1}. \quad (\text{C.9})$$

The analytical solution in the xz plane (a dependence between distance z_1 and z_{obj}) is more complicated (due to the cylindrical lens approximation), nevertheless, it is very similar to the solution in the yz . Since the solution in the yz plane approximates the cylindrical lenses by air, only one transformation from two lenses to one lens has been used. As there are three lenses in the xz plane, therefore the transformation from two lenses to one lens has been used twice. The first transformation, however, is the same as in the yz plane and thus the results of transformation in the yz plane have been applied. Fig. C.9 shows the solution in the xz plane where the remaining variables have the following meaning:

- f_{cyl} : objective focal length of the cylindrical lens,
- f'_{cyl} : image focal length of the cylindrical lens,
- F'_{cyl}, F'_1 : focal points,
- v_2 : distance between two lenses,
- Δ_2 : optical interval,

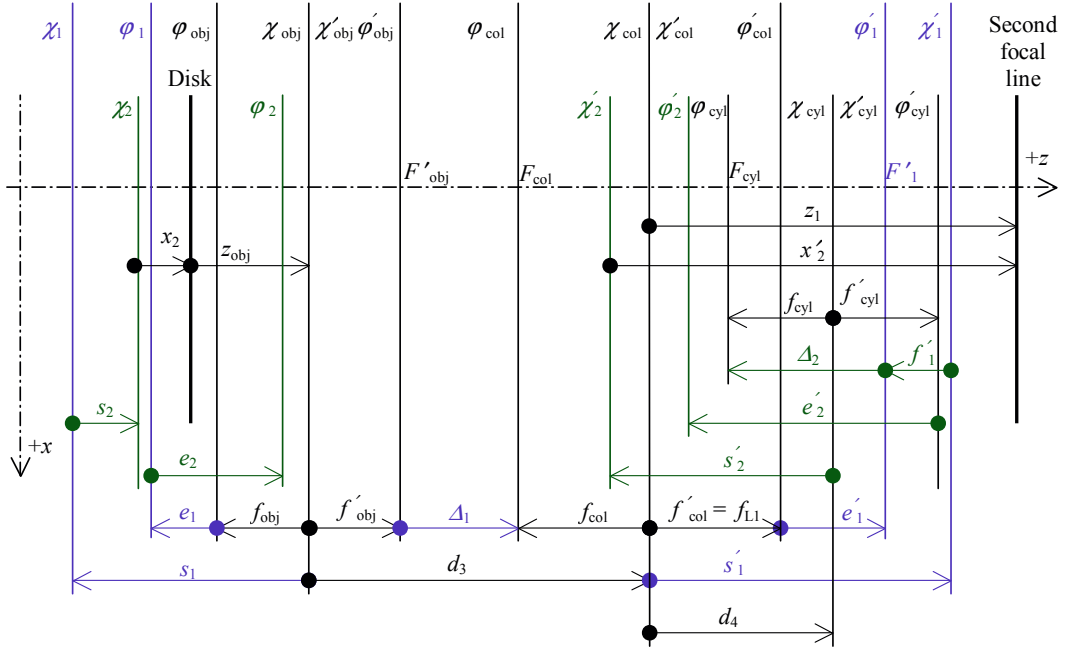


Figure C.9: System formed by two centered thin lenses in the xz plane.

- $\chi_{cyl}, \chi'_{cyl}, \chi_2, \chi'_2$: unit (principal) planes of lenses,
- $\varphi_{cyl}, \varphi'_{cyl}, \varphi_2, \varphi'_2$: focal planes of lenses,
- e_2 : distance of the “resulting” focal plane φ_2 from the objective focal plane φ_1 of the first lens,
- e'_2 : distance of the “resulting” focal plane φ'_2 from image focal plane φ'_{cyl} of the second lens,
- s_2 : distance of the “resulting” first unit (principal) plane χ_2 from the first unit plane χ_1 of the first lens,
- s'_2 : distance of the “resulting” second unit (principal) plane χ'_2 from last unit plane χ'_{cyl} of the second lens,
- x_2, x'_2 : objective and image distance of the “resulting” lens approximation,
- f'_2 : “resulting” objective focal length in xz plane.

The unknown focal length f'_{cyl} variable can be calculated from the following condition which defines the disk in focus situation:

$$\text{if } (z_{obj} = -f_{obj}) \text{ then } (z_1 = f_{L2}). \quad (\text{C.10})$$

Let $f'_{cyl} = -f_{cyl}$ then the quantities of the optical system can be expressed by these formulas:

$$\Delta_2 = -e'_1 - f'_{col} + d_4 + f_{cyl} = -e'_1 - f'_{col} + d_4 - f'_{cyl}, \quad (\text{C.11})$$

$$v_2 = \Delta_2 + f'_1 - f_{cyl} = -e'_1 - f'_{col} + d_4 + f'_1, \quad (\text{C.12})$$

$$s_2 = \frac{v_2 f'_1}{f'_1 + f'_{\text{cyl}} - v_2}, \quad (\text{C.13})$$

$$s'_2 = -\frac{v_2 f'_{\text{cyl}}}{f'_1 + f'_{\text{cyl}} - v_2}, \quad (\text{C.14})$$

$$e_2 = -\frac{f'^2_1}{\Delta_2}, \quad (\text{C.15})$$

$$e'_2 = \frac{f'^2_{\text{cyl}}}{\Delta_2}, \quad (\text{C.16})$$

$$f'_2 = -\frac{f'_1 f'_{\text{cyl}}}{\Delta_2}. \quad (\text{C.17})$$

In this case the Lensmaker's equation is:

$$\frac{1}{x'_2} - \frac{1}{x_2} = \frac{1}{f'_2}. \quad (\text{C.18})$$

Using the substitution for x'_2 , x_2 in (C.18) which is clear from the geometry in fig. C.9, one can write:

$$\frac{1}{-s'_2 - d_4 + z_1} - \frac{1}{-s_2 - s_1 - z_{\text{obj}}} = \frac{1}{f'_2}. \quad (\text{C.19})$$

Finally, the dependence between distance z_1 and z_{obj} is given from (C.19), thus

$$\boxed{z_1 = \frac{1}{\frac{1}{f'_2} + \frac{1}{-s_2 - s_1 - z_{\text{obj}}}} + s'_2 + d_4.} \quad (\text{C.20})$$

One can note that the variables s_2 , s'_2 and f'_2 depend on the unknown variable f'_{cyl} in (C.20), see (C.13), (C.14) and (C.17). The condition (C.10) has been used in Lensmaker's equation (C.19) of the xz plane to derive unknown variable f'_{cyl} as follows:

$$\frac{1}{-s'_2 - d_4 + f_{L2}} - \frac{1}{-s_2 - s_1 - f'_{\text{obj}}} - \frac{1}{f'_2} = 0. \quad (\text{C.21})$$

It is possible to obtain an explicit solution for quantity f'_{cyl} from the relationship (C.21). The deduction is complicated but the results are not so complex. The approximation by the diverging thin lens is used in the xz plane and therefore the focal length of diverging thin lens f'_{cyl} is negative. This means that $f'_{\text{cyl}} < 0$, which is a vital property for deriving the proper root of (C.21). Thus, f'_{cyl} is given by:

$$\boxed{f'_{\text{cyl}} = \frac{f_{L1} f_{L2} - f_{L1} d_4 - f_{L2} d_4 + d_4^2}{f_{L1} - f_{L2}}.} \quad (\text{C.22})$$

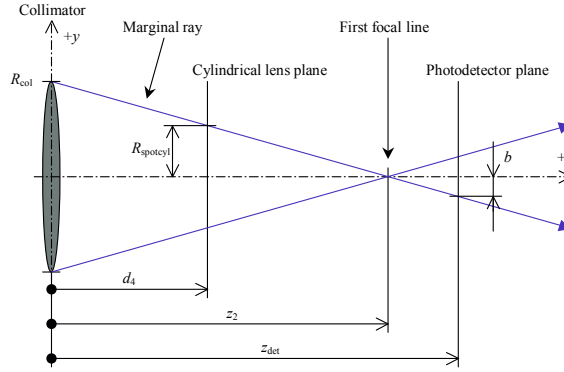


Figure C.10: Eduction of R_{spotcyl} and half-axis b in the yz plane.

C.4.5 Spot Size

Light spot shape determination on the photodetector is the next step to obtain the S-curve model. It is usually approximated by elliptical shape that is shown in fig. C.12, therefore the aim of this section is to define the dependence between the half-axis sizes of the elliptic laser spot a , b and the focal line distances z_1 , z_2 , respectively.

Firstly, the half-axis size b (a dependence between half-axis b and z_2) is solved in the yz plane. The laser spot radius in the cylindrical lens plane R_{spotcyl} is determined from triangle similarity. Fig. C.10 shows this situation in the yz plane where the cylindrical diverging lens is approximated by air. Out of the geometry illustrated in fig. C.10, one can obtain:

$$\frac{R_{\text{col}} - R_{\text{spotcyl}}}{d_4} = \frac{R_{\text{col}}}{z_2}, \quad (\text{C.23})$$

where the collimator and the objective lens radius are $R_{\text{col}} = R_{\text{obj}} = \frac{f'_{\text{obj}} NA}{\sqrt{1-NA^2}}$ and NA is the objective lens numerical aperture. Then, using (C.23), the laser spot radius in the cylindrical lens plane R_{spotcyl} is given by:

$$R_{\text{spotcyl}} = \frac{R_{\text{col}}(z_2 - d_4)}{z_2}. \quad (\text{C.24})$$

The laser spot half-axis size b is given by:

$$b = \left| R_{\text{spotcyl}} \frac{z_{\text{det}} - z_2}{z_2 - d_4} \right|, \quad (\text{C.25})$$

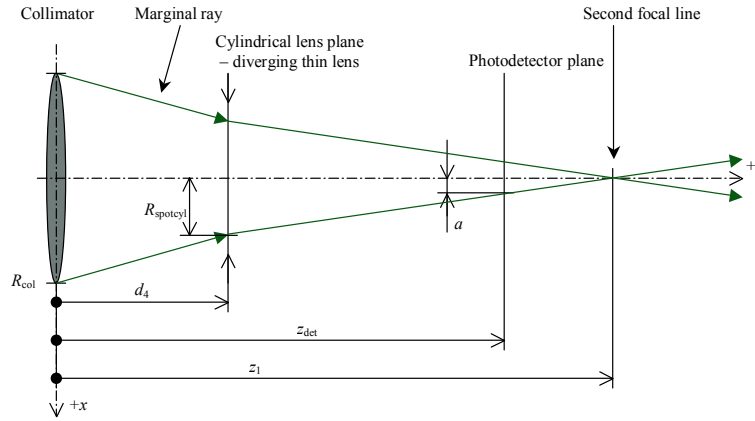
where triangle similarity is used (see fig. C.10) and the absolute value expresses the non-negativity of the laser spot half-axis size b .

The laser spot half-axis size a (a dependence between a and z_2) can be obtained in a similar way from the triangles, see fig. C.11. Hence,

$$\frac{R_{\text{spotcyl}}}{z_1 - d_4} = \frac{a}{z_{\text{det}} - z_1} \quad (\text{C.26})$$

and therefore

$$a = \left| R_{\text{spotcyl}} \frac{z_{\text{det}} - z_1}{z_1 - d_4} \right|, \quad (\text{C.27})$$


 Figure C.11: Eduction of half-axis a in the xz plane.

where absolute value is used because the laser spot half axis is never negative.

The distance between the collimator and the photodetector z_{det} , needed for half-axis calculation, is given by the following condition: *If the disk is in focus, then the laser spot half-axes a_{foc} , b_{foc} are equal to each other and distances of two focal lines from the collimator z_1 , z_2 are f_{L1} and f_{L2} , respectively. Then*

$$a_{\text{foc}} = b_{\text{foc}}, \quad z_1 = f_{L2}, \quad z_2 = f_{L1}. \quad (\text{C.28})$$

Using triangle similarity, see figs. C.10 and C.11, (C.28) has been rewritten as follows:

$$-R_{\text{spotcylfoc}} \frac{z_{\text{det}} - d_4}{f_{L1}} + R_{\text{spotcylfoc}} = -R_{\text{spotcylfoc}} \frac{z_{\text{det}} - d_4}{f_{L2}} - R_{\text{spotcylfoc}}, \quad (\text{C.29})$$

where $R_{\text{spotcylfoc}}$ is a light spot radius in the cylindrical lens plane if the disk is in focus. Using (C.29), one can obtain the distance between collimator and the photodetector:

$$z_{\text{det}} = \frac{-f_{L1}(d_4 - 2f_{L2}) + d_4 f_{L2}}{f_{L2} - 2d_4 + f_{L1}}. \quad (\text{C.30})$$

Remark 11 *If the disk is in focus, then the laser spot radius $R_{\text{spotcylfoc}}$ is expressed from (C.24) and (C.30) by relationship:*

$$R_{\text{spotcylfoc}} = \frac{R_{\text{col}}(f_{L1} - d_4)}{f_{L1}}. \quad (\text{C.31})$$

Remark 12 *If the disk is in focus, the spot is circular and the laser spot half-axes a_{foc} , b_{foc} on the photodetector, are computed from the following equations. They are defined from figs. C.10 and C.11, so:*

$$\frac{R_{\text{spotcylfoc}}}{f_{L1} - d_4} = \frac{a_{\text{foc}}}{z_{\text{det}} - f_{L1}}, \quad \frac{R_{\text{spotcylfoc}}}{f_{L2} - d_4} = \frac{a_{\text{foc}}}{f_{L2} - z_{\text{det}}} \quad (\text{C.32})$$

and therefore:

$$a_{\text{foc}} = b_{\text{foc}} = \frac{R_{\text{spotcylfoc}}(f_{L1} + f_{L2} - 2z_{\text{det}})}{f_{L2} - f_{L1}}. \quad (\text{C.33})$$

C.4.6 Light Intensity Distribution

The amount of light energy, captured by the photodetector from the laser spot on the photodetector, must be calculated to complete S-curve model creation. Since it is difficult to define light intensity distribution on the photodetector analytically, rough approximations have been used in [HBVSdF02], such as *uniform* and “*paraboloid*” light intensity distribution (in the case of analytical models) or *Gaussian* intensity distribution (in case of numerical model) rather than more complicated light distribution described by Rayleigh-Sommerfeld and Kirchhoff diffraction integrals. From the last research, see [HBVSdF02], uniform approximation is better than paraboloid approximation. An explanation of the models that use uniform and Gaussian light intensity distribution is presented here because the analytical model is useful for its calculation time saving while the numerical model is more accurate.

If uniform intensity distribution approximation is used, then the whole energy (voltage), that can be measured on the *infinite dimension photodetector* placed in the detector plane, is given by:

$$V_{\text{WholeU}} = 4 \int_0^a \int_0^{b\sqrt{1-\frac{x^2}{a^2}}} c_U dx dy = \pi abc_U, \quad (\text{C.34})$$

where c_U includes all proportional constants, such as

1. the relation between real light intensity distribution and its uniform approximation and
2. photodetector sensitivity and other proportional parameters in the light beam path.

Due to the “energy (light flux) conservation law”, the light flux coming out from the cylindrical lens has to be equal to the light flux which falls on the whole infinite photodetector plane. Therefore the variable c_U is not constant for different values of half-axes a , b while V_{WholeU} is always constant.

If Gaussian intensity distribution approximation is used, then the laser spot half-axes a , b on the distribution function has been defined. It is useful to introduce a variable H_R (height ratio) on normalized Gaussian intensity distribution which defines the quantity of intensity for half-axes a , b on normalized Gaussian intensity distribution, $H_R \in (0, 1)$. For example, if $H_R = 0.2$, then the half-axes a , b are given at 0.2 height on the graph of normalized Gaussian intensity distribution. Then, the whole energy on *infinite dimension photodetector* is given by:

$$V_{\text{WholeG}} = 4 \int_0^\infty \int_0^\infty \left(c_G \exp \left(-\frac{x^2}{a_{1G}^2} - \frac{y^2}{b_{1G}^2} \right) \right) dx dy, \quad (\text{C.35})$$

where a_{1G} , b_{1G} take the forms

$$a_{1G} = \frac{a}{\sqrt{-\log(H_R)}}, \quad (\text{C.36})$$

$$b_{1G} = \frac{b}{\sqrt{-\log(H_R)}}, \quad (\text{C.37})$$

where c_G is proportional constant with the same meaning as c_U in the uniform intensity distribution case.

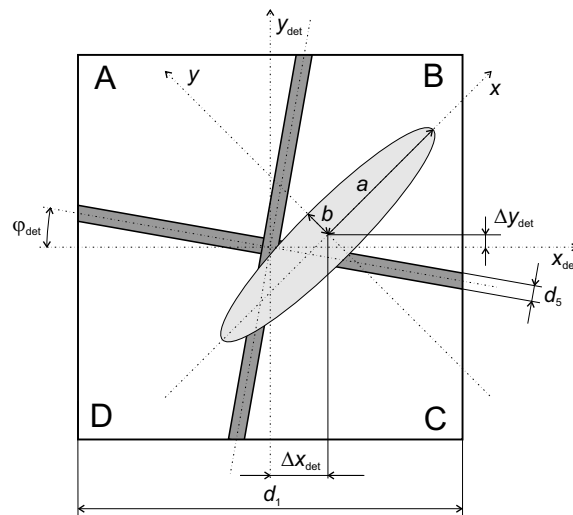


Figure C.12: The geometrical arrangement of the photodetector.

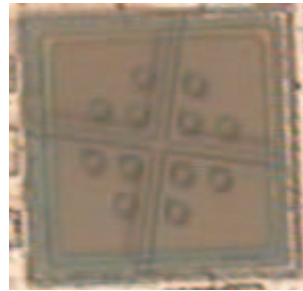


Figure C.13: A photo of the real photodetector arrangement.

Thus, both proportional constants, c_U and c_G , can be calculated from (C.34) and (C.35) with assumption of constant V_{WholeU} and V_{WholeG} , respectively.

Fig. C.12 shows a geometrical arrangement of the photodetector, normally used in presented S-curve modelling where the introduced variables have the following meaning:

- $x_{\text{det}}, y_{\text{det}}$: photodetector coordinates,
- d_1 : photodetector size,
- d_5 : gap between photodiodes,
- φ_{det} : angle of photodetector rotation,
- Δx_{det} : displacement in x_{det} direction,
- Δy_{det} : displacement in y_{det} direction.

Fig. C.13 illustrates a photo of the photodetector used in DVD-video players. The measured voltage from the photodiode A is given by:

$$V_A = \iint_{\Omega_A} c_U dx dy. \quad (\text{C.38})$$

in case of the analytical model, where Ω_A is the photodiode area A and c_U is given from (C.34). In case of the numerical model, the voltage V_A has the form

$$V_A = \iint_{\Omega_A} \left(c_G \exp \left(-\frac{x^2}{a_{1G}^2} - \frac{y^2}{b_{1G}^2} \right) \right) dx dy, \quad (C.39)$$

where c_G is given from (C.35). The other voltages V_B , V_C and V_D of the analytical/numerical model are defined analogically.

To calculate the voltage V_{WholeU} from the optical pick-up unit and DVD disk specifications, the following notation has been introduced:

- P_{LD} : laser power,
- η_{LD2disk} : effectiveness from the laser diode to the disk,
- η_{disk} : disk reflectivity,
- η_{disk2PD} : effectiveness from the disk to the photodetector,
- a_{focMax} : maximum value of the half axis a on the photodetector,
- s_{det} : photodetector sensitivity.

If we assume that the spot size is smaller than the photodetector lateral size d_1 , the laser spot power on the photodetector is given by:

$$P_{\text{PD}} = P_{\text{LD}} \cdot \eta_{\text{LD2disk}} \cdot \eta_{\text{disk}} \cdot \eta_{\text{disk2PD}}. \quad (C.40)$$

Photodetector effectiveness is:

$$\eta_{\text{PD}} = \frac{a_{\text{focMax}}}{\left(a_{\text{focMax}} - \frac{d_5}{\sqrt{2}} \right)} \quad (C.41)$$

and the power on the whole photodetector is defined by:

$$P_{\text{PDWholeU}} = P_{\text{PD}} \cdot \eta_{\text{PD}}. \quad (C.42)$$

Finally, the whole voltage in the photodetector plane is given by:

$$V_{\text{WholeU}} = s_{\text{det}} \cdot P_{\text{PDWholeU}}. \quad (C.43)$$

If we assume a truncated Gaussian intensity profile, the V_{WholeG} calculation is analog to the V_{WholeU} calculation presented above.

Remark 13 *If the gaps between photodiodes A, B, C, D are ignored (i.e. $d_5 = 0$) then V_{WholeU} directly equals to the maximum value of the focus error voltage e_F on the S-curve.*

Remark 14 *If there are not such complete specifications to calculate V_{WholeU} then this quantity V_{WholeU} could be estimated with other unknown or hardly measurable variables from the measured S-curve.*

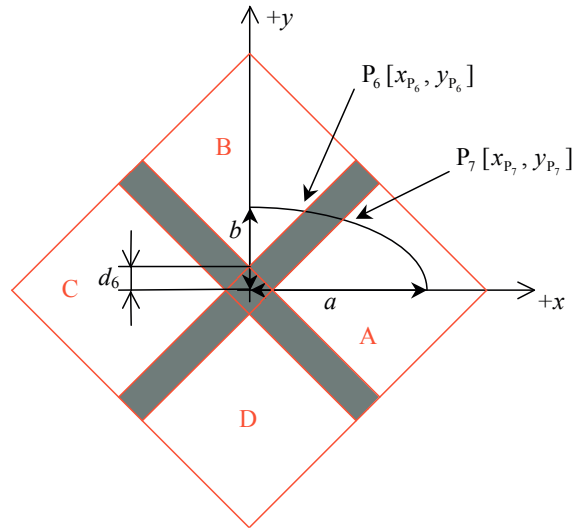


Figure C.14: Special position of the photodetector.

The analytical model: solution in special case

Let the photodetector position be centered exactly with the optical axis z , *i.e.* $\Delta x_{\text{det}} = 0$, $\Delta y_{\text{det}} = 0$, and let the angle of the detector rotation equal the zero, *i.e.* $\varphi_{\text{det}} = 0^\circ$ as it is shown in fig. C.14.

The intersections $P_6 [x_{P_6}, y_{P_6}]$ and $P_7 [x_{P_7}, y_{P_7}]$, (see fig. C.14), have been analytically calculated from the set of equations defining geometrical photodetector properties and elliptic laser size. The general solution of intersections is given by:

$$x_{P_6} = \frac{(-a d_6 + b \sqrt{a^2 - d_6^2 + b^2}) a}{a^2 + b^2}, \quad (\text{C.44})$$

$$x_{P_7} = \frac{(a d_6 + b \sqrt{a^2 - d_6^2 + b^2}) a}{a^2 + b^2}, \quad (\text{C.45})$$

$$d_6 = \frac{d_5}{\sqrt{2}}. \quad (\text{C.46})$$

The general solution of output voltage V_A is given by:

$$\begin{aligned} V_A &= 2 \int_{d_6}^{x_{P_7}} \int_0^{x-d_6} c_U dx dy + 2 \int_{x_{P_7}}^a \int_0^{b \sqrt{1 - \frac{x^2}{a^2}}} c_U dx dy = \\ &= c_U (x_{P_7}^2 - d_6^2) - 2c_U d_6 (x_{P_7} - d_6) - \\ &\quad - \frac{c_U b \left(-a^2 \pi + 2x_{P_7} \sqrt{a^2 - x_{P_7}^2} + 2a^2 \arcsin\left(\frac{x_{P_7}}{a}\right) \right)}{2a}, \end{aligned} \quad (\text{C.47})$$

$$\begin{aligned} V_B &= 2 \int_0^{x_{P_6}} \int_{x+d_6}^{b \sqrt{1 - \frac{x^2}{a^2}}} c_U dx dy = \\ &= \frac{\left(b x_{P_6} \sqrt{a^2 - x_{P_6}^2} + b a^2 \arcsin\left(\frac{x_{P_6}}{a}\right) - x_{P_6}^2 a - 2d_6 x_{P_6} a \right) c_U}{a}, \end{aligned} \quad (\text{C.48})$$

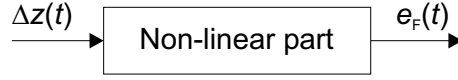


Figure C.15: Block-scheme of the focus error generation model.

Table C.1: Unknown model parameters vector θ and known model parameters vector p for the analytical/numerical model.

Item	Analytical model	Numerical model
θ	$f_{L1}, f_{L2}, d_4, V_{\text{WholeU}}$	$f_{L1}, f_{L2}, d_4, V_{\text{WholeG}}, H_R$
p	$f_{\text{obj}}, NA, d_1, d_3, d_5$	$\Delta x_{\text{det}}, \Delta y_{\text{det}}, \varphi_{\text{det}}, \Delta z_{\text{pp}}$

$$V_C = V_A, \quad (\text{C.49})$$

$$V_D = V_B, \quad (\text{C.50})$$

where c_U , (see (C.34)), is

$$c_U = \frac{V_{\text{WholeU}}}{\pi ab}. \quad (\text{C.51})$$

The solution introduced is useable if the laser spot, defined by half-axes a , b , is inside the photodetector area. Other cases of analytical solutions are even more complicated and are not presented here.

Limited photodetector size d_1 is the reason why the focus error signal is close to zero for large vertical spot position errors Δx .

C.4.7 Modelling Summary

The model is described by a non-linear function. Output of analytical/numerical model is the focus error signal e_F . It is given by:

$$e_F(\Delta z, \theta, p, t) = \left(V_A(\Delta z, \theta, p, t) + V_C(\Delta z, \theta, p, t) \right) - \left(V_B(\Delta z, \theta, p, t) + V_D(\Delta z, \theta, p, t) \right), \quad (\text{C.52})$$

where V_A , V_B , V_C , V_D are voltages from photodiodes A, B, C, D of the photodetector, respectively. Expression (C.52) where the focus error signal e_F is expressed as a function of the vertical spot position error Δz and of known/unknown model parameters p and θ , corresponds to (3.8). The unknown/known model parameters are summarized in table C.1.

Finally, a block-scheme of the focus error signal generation model is shown in fig. C.15.

C.5 Influence of Different Model Parameters

Before discussing the results of the simulation, the terms must be defined. A typical S-curve is presented in fig. 3.7. The area between peaks of the S-curve is called the

Table C.2: Nominal parameters of system and models.

Item	Specification	Label
Focal length	-3.05 mm	f_{obj}
Numerical aperture	0.60	NA
Photodetector size	100 μm	d_1
Distance between objective lens and collimator	5 mm	d_3
Distance between collimator and cylindrical lens	22.168 mm	d_4
Gap between photodiodes	0 μm	d_5
Distance between collimator and the first focal line	26.099 mm	f_{L1}
Distance between collimator and the second focal line	26.581 mm	f_{L2}
Displacement in x_{det} direction	0 μm	Δx_{det}
Displacement in y_{det} direction	0 μm	Δy_{det}
Angle of photodetector rotation	0°	φ_{det}
Whole energy uniform intensity distribution	1	V_{WholeU}
Whole energy Gaussian intensity distribution	1	V_{WholeG}
Gauss ratio	0.4	H_{R}
Peak to peak value of focus error characteristic $e_{\text{F}}(\Delta z)$	6 μm	Δz_{pp}

lock-on range. The physical distance corresponding to this area represents the extent of physical displacement that the system is able to respond to and still remain locked. An *acquisition range* is the distance from focus in which $e_{\text{F}}(\Delta z)$ is large enough to allow the focus system to acquire focus, *i.e.* the range between the points S_2 and S_8 in fig. 3.9. *The final characteristic of interest is its shape.* A very asymmetric S-curve would lead to problems because in this case the focusing system needs to be electronically adjusted to set the focus error $e_{\text{F}}(\Delta z)$ equal to zero if the disk is in focus.

For demonstration purposes, an optical system specification is given in table C.2. The following discussion describes the effects of varying one of the above variables on the model $\hat{e}_{\text{F}}(\Delta z, \theta)$, (C.52) that uses the Gaussian distribution. In the other words, the influence of model parameters $\theta = [d_1, \Delta x_{\text{det}}, \Delta y_{\text{det}}, d_5]$ is explained below. Note that the graphs of the simulated S-curves $\hat{e}_{\text{F}}(\Delta z, \theta)$ are all normalized with respect to peak to peak value, so that parameter influence can be compared.

Fig. C.16 shows the influence of varying detector size d_1 . As presented, the *lock-on range* remains unchanged but the *acquisition range* increases with increasing detector size. Limited photodetector size d_1 is the reason why only part of light energy is captured by photodetector and the focus error signal is close to zero for large vertical spot position errors Δx .

Figs. C.17–C.19 show the analysis of the quadrature detector displacement $\Delta x_{\text{det}}, \Delta y_{\text{det}}$ along two orthogonal directions $x_{\text{det}}, y_{\text{det}}$. Note that in figs. C.17 and C.18, the displacement is along the axes of the quadrant separation while the displacement is along the direction at 45° to the axes $x_{\text{det}}, y_{\text{det}}$ in fig. C.19.

Figs. C.17 and C.18 illustrate that the detector shift in one direction (x_{det} or y_{det}) leads only to the lower $\hat{e}_{\text{F}}(\Delta z, \theta)$ peak to peak value and lower sensitivity (*i.e.*, lower slope through the S-curve origin). They show that the influence of detector displacement in the x_{det} and the y_{det} direction on the $\hat{e}_{\text{F}}(\Delta z, \theta)$ is the same. Note that because of the

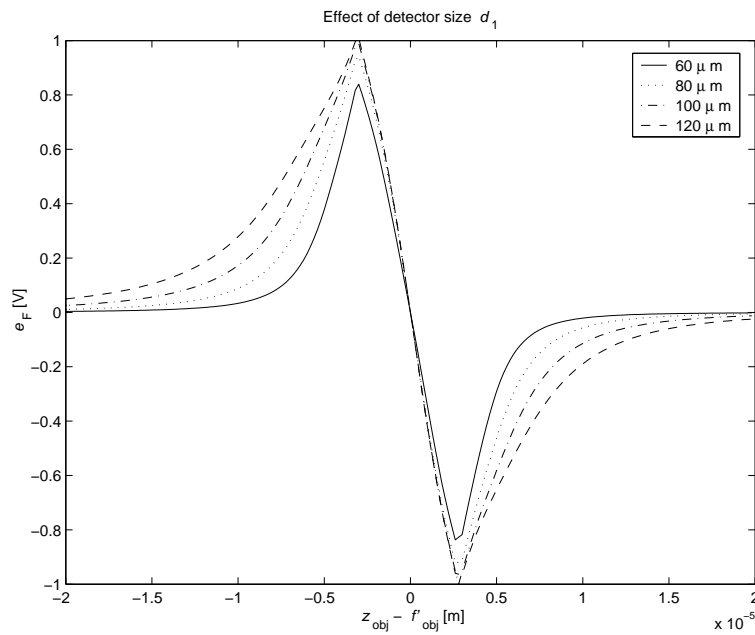


Figure C.16: Effect of photodetector size d_1 .

symmetry of the optical system a displacement in the positive direction x_{det} , (y_{det}) is equivalent to a displacement in the negative direction x_{det} , (y_{det}).

Nevertheless, a non-zero displacement in both directions could cause a strong shift of the S-curve as it is shown in fig. C.19 where the $\Delta x_{\text{det}} = \Delta y_{\text{det}}$. Focus error characteristic $\hat{e}_F(\Delta z, \theta)$ becomes very asymmetrical and its sensitivity quickly decreases. The photodetector displacement analysis is very important because it explains the shift of the S-curve during normal DVD playing mode. This shift is caused by the disk warping, so called a vertical deviation, and places a large repetitive error into focus closed-loop.

Fig. C.20 shows the effect of gap size d_5 between photodiodes A, B, C, D. It is clear that sharp peaks on the S-curve are caused by the zero size of gaps between the photodiodes. As is shown, the *lock-on range* is not changed but sensitivity decreases with the increasing gap size.

Fig. C.21 presents the influence of varying angle of dividing line φ_{det} between photodetector parts. If a rotation is smaller than 30° , the $\hat{e}_F(\Delta z, \theta)$ is not strongly affected. The *lock-on range* is unchanged but the *acquisition range* increases and sensitivity decreases slightly with the increasing angle.

Fig. C.22 shows the effect of varying Gaussian intensity distribution ratio H_R . In general, as ratio H_R increases, the *acquisition range* increases, nevertheless, photodetector sensitivity and its *lock-on range* do not undergo radical changes.

C.6 Estimation of Unknown Model Parameters

Real S-curves of DVD players are measured during a focus checking procedure. An example of the measured characteristic $e_F(t)$ of the real optical pick-up unit is shown in fig. 3.7.

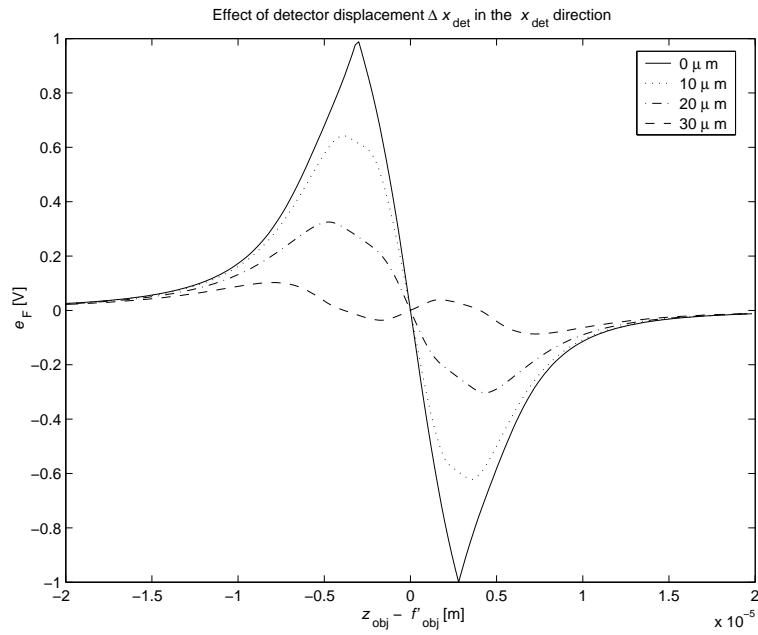


Figure C.17: Effect of photodetector displacement Δx_{det} .

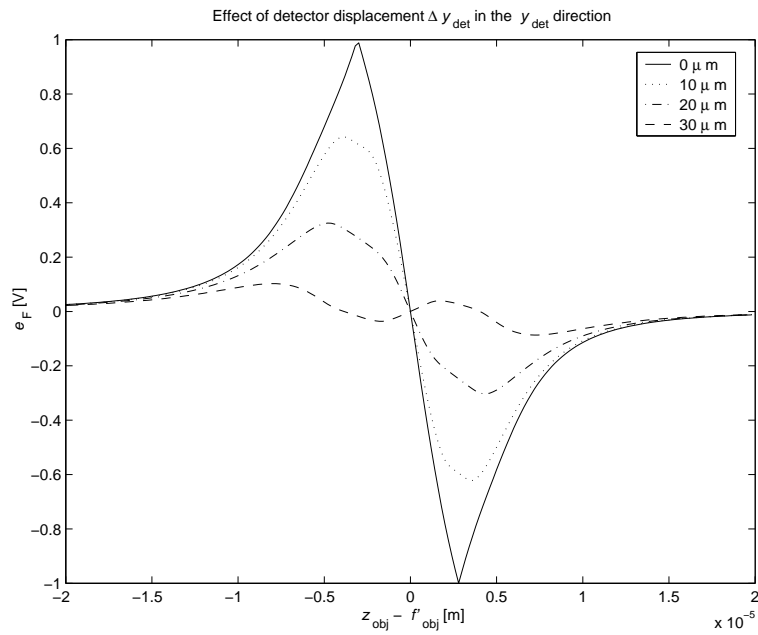


Figure C.18: Effect of photodetector displacement Δy_{det} .

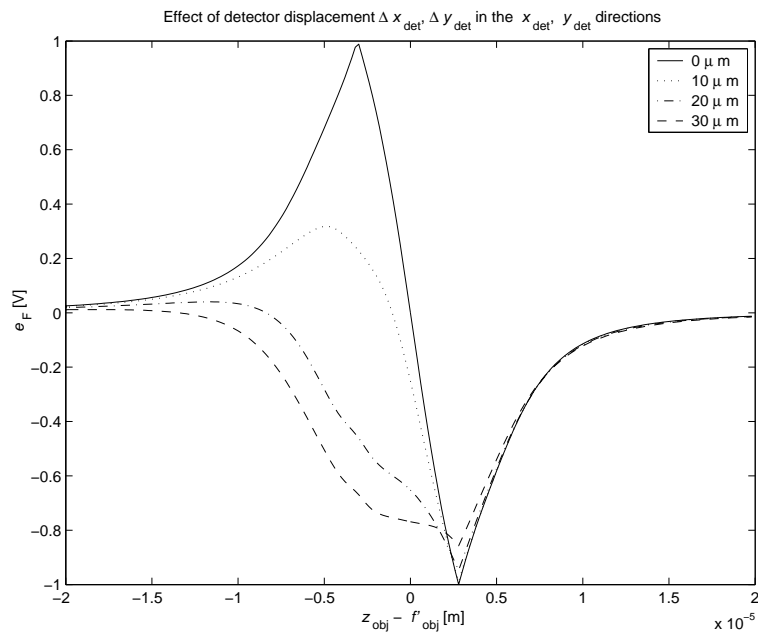


Figure C.19: Effect of photodetector displacement $\Delta x_{\text{det}}, \Delta y_{\text{det}}$; $\Delta x_{\text{det}} = \Delta y_{\text{det}}$.

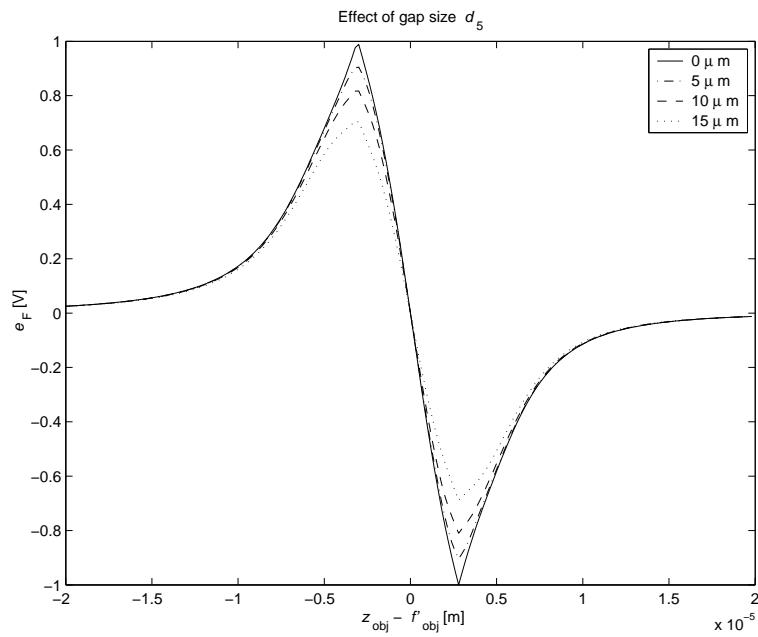


Figure C.20: Effect of photodetector gap size d_5 .

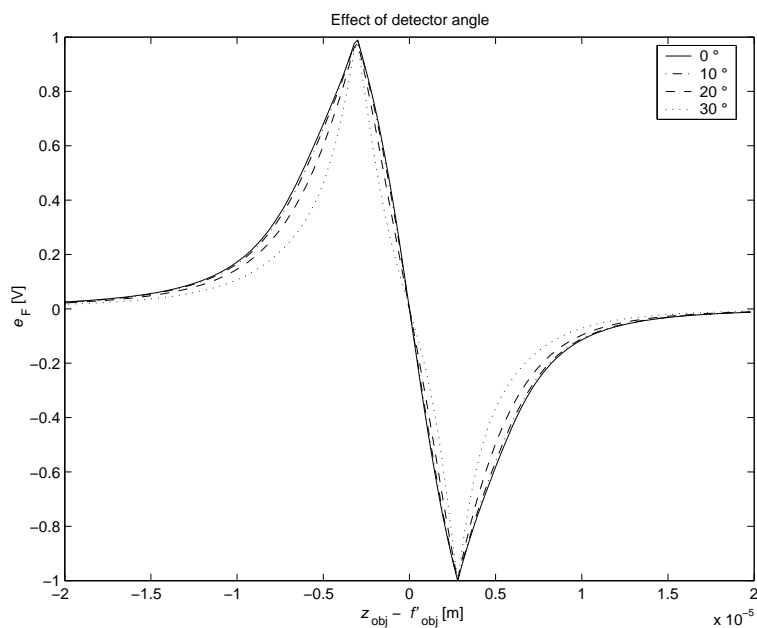


Figure C.21: Effect of photodetector angle φ_{det} .

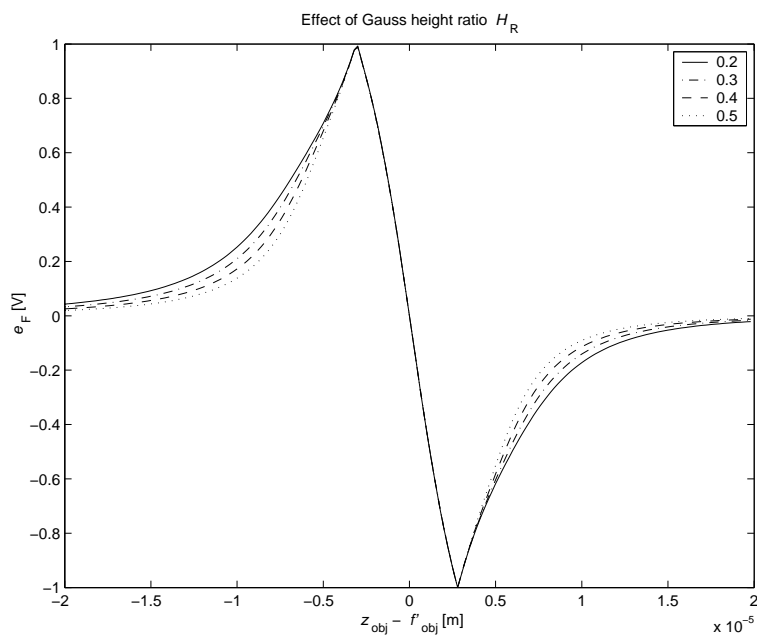


Figure C.22: Effect of Gauss height ratio H_R .

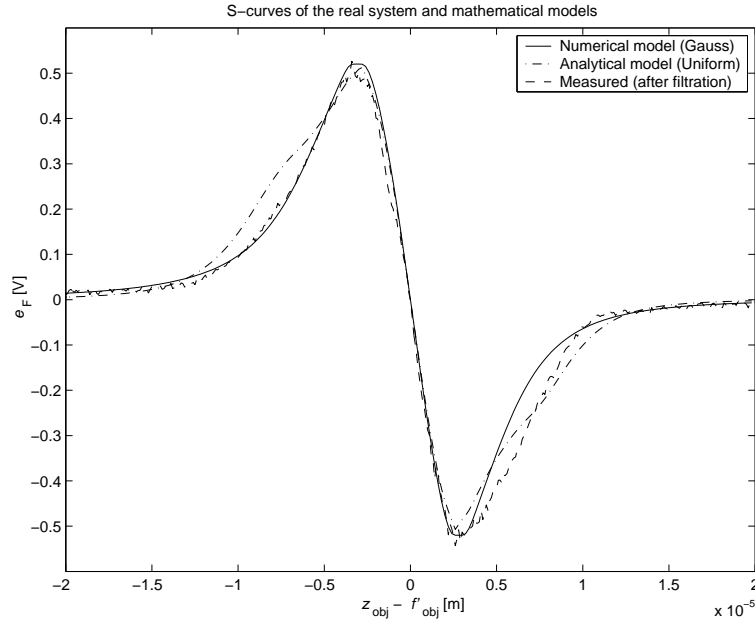


Figure C.23: S-curves of the real system and mathematical models.

Some preprocessing procedures have to be performed on the measured data to obtain the focus error characteristic $e_F(\Delta z)$ from $e_F(t)$. Following procedures are described below:

Filtering: The causal moving average filter has been used to filter the real data.

Calibration: The peak to peak values of the filtered focus error characteristic $e_F(t)$ have been found to normalize the filtered focus error characteristic $e_F(t)$ from time domain (t) to the vertical spot position error domain (Δz).

Interpolation: The calibrated focus error characteristic $e_F(\Delta z)$ was resampled by linear interpolation to obtain smaller number of the samples on the S-curve. Re-sampling accelerates greatly the next curve fitting procedure.

Fig. C.23 shows the focus error characteristic $e_F(\Delta z)$ after the preprocessing procedures described above (dashed line). This characteristic is referred to as the “nominal” focus error characteristic in the next curve fitting procedure.

The analytical and numerical model have been created in section C.4. The analytical model uses uniform intensity distribution of the laser beam on the photodetector while the numerical model uses Gaussian intensity distribution. Since few important optical parameters, such as f_{L1} , f_{L2} , d_4 , V_{WholeU} and V_{WholeG} , H_R are difficult to measure in the real system, their values have been identified by the curve fitting procedure.

The curve fitting method is based on a weighted nonlinear least-squares criterion:

$$J(\theta) = \sum_{\Delta z = \Delta z_{\text{ini}}}^{\Delta z_{\text{fin}}} w [e_F(\Delta z) - \hat{e}_F(\Delta z, \theta)]^2, \quad (\text{C.53})$$

where the highest weight, given by a vector w , has been set on the *lock-on range* of the S-curve and smaller weights on the remaining parts. $e_F(\Delta z)$ stands for the measured

Table C.3: Statistical parameters of the curve-fitting procedure.

Model	$J(\theta)_{\min}$	$\bar{\sigma}$	$\bar{\nu}$
Analytical	0.2262	± 0.03067	0.00094
Numerical	0.1615	± 0.02821	0.00076

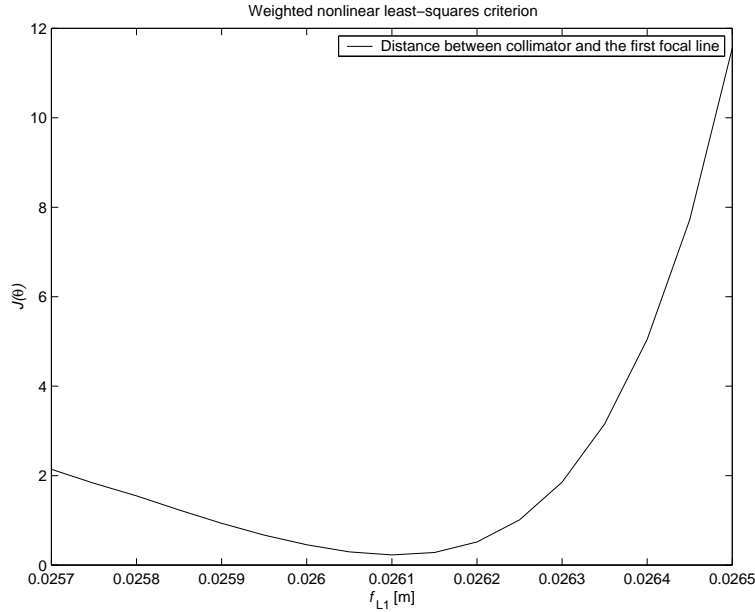


Figure C.24: Weighted nonlinear least-squares criterion for the analytical model.

error signal from the real system (“nominal” characteristic) while $\hat{e}_F(\Delta z, \theta)$ is the error signal calculated by the model. The unknown parameters vector θ is different for the analytical and numerical model as it is shown in table C.1.

The gradient searching method has been applied to minimize the criterion (C.53). The criterion minimum value $J(\theta)_{\min}$, the standard deviation $\bar{\sigma}(e_F(\Delta z) - \hat{e}_F(\Delta z, \theta))$ and the variance $\bar{\nu}(e_F(\Delta z) - \hat{e}_F(\Delta z, \theta))$ of the estimated models have been used to evaluate both models. These statistical parameters are presented in table C.3.

Figs. C.25 and C.24 illustrate the criterion function $J(\theta)$ map line for the analytical model and two different quantities f_{L1} , V_{WholeU} of the estimated vector θ . It is clear that the criterion is convex with respect to f_{L1} , V_{WholeU} .

The curve fitting procedure, described above, has been applied to obtain the focus error characteristics of two created mathematical models $\hat{e}_F(\Delta z, \theta)$ that are shown in fig. C.23. They are indicated by the solid and dot-dashed lines. The “nominal” focus error characteristic $e_F(\Delta z)$ obtained from preprocessing is indicated by the dashed line there.

The nominal parameters of real system are in table C.4 while the identified parameters of both models are presented C.5. Fig. C.23 and table C.3 show compatibility of the numerical model (Gaussian intensity distribution) with the real system. The analytical

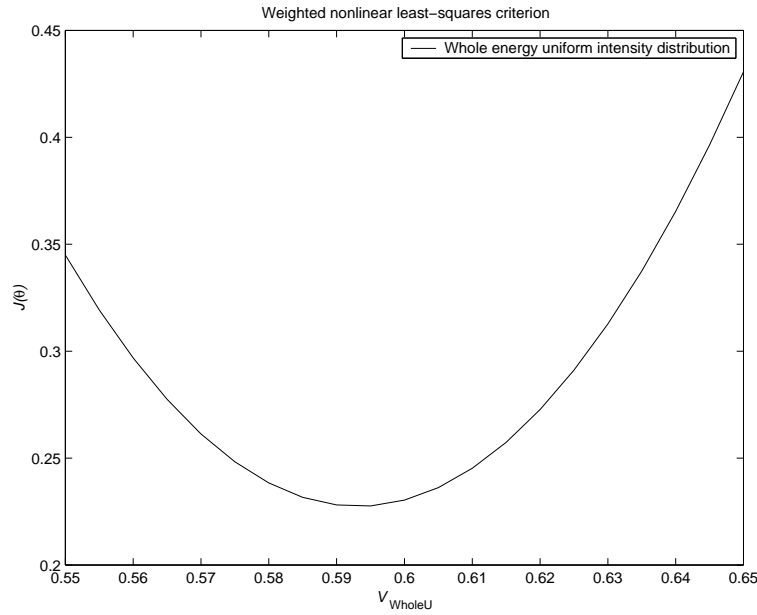


Figure C.25: Weighted nonlinear least-squares criterion for the analytical model.

model has two advantages. Firstly, despite the inferior performance of analytical model, this model is *completely analytical*.

C.7 Possible Applications of Photodetector Characteristics

C.7.1 Simulation of Start-up Procedure

The first application of the created models deals with the start-up procedure where data layer is searched. This focus processing procedure to automatically close the focus loop when the focus point is reached in the S-curve *lock-on range* is illustrated in fig. C.26.

Table C.4: Nominal parameters of real system.

Item	Specification	Label
Focal length	-3.05 mm	f_{obj}
Numerical aperture	0.60	NA
Photodetector size	100 μm	d_1
Distance between objective lens and collimator	5 mm	d_3
Gap between photodiodes	5 μm	d_5
Displacement in x_{det} direction	0 μm	Δx_{det}
Displacement in y_{det} direction	0 μm	Δy_{det}
Angle of photodetector rotation	10°	φ_{det}
Peak to peak value of focus error characteristic $e_F(\Delta z)$	6 μm	Δz_{pp}

Table C.5: Identified parameters.

Item	Value	Label
Distance between collimator and the first focal line	26.099 mm	f_{L1}
Distance between collimator and the second focal line	26.581 mm	f_{L2}
Distance between collimator and cylindrical lens	22.168 mm	d_4
Whole energy uniform intensity distribution	0.59	V_{WholeU}
Whole energy Gaussian intensity distribution	0.52	V_{WholeG}
Gauss ratio	0.4	H_R

Here, the $K_F(q^{-1})$ denotes the digital controller transfer function whose digital output is converted to analog signal by the Digital to Analog (D/A) convertor, A_{Fdri} is the power driver gain, $G_{\text{Fact}}(s)$ is the continuous-time transfer function of the actuator, z_d is the disturbance of the turning disk in vertical direction, $z_{\text{ref}} \equiv f'_{\text{obj}}$ is the objective lens reference value, $I(\Delta z)$ is the light intensity distribution of the return beams and $G_F(s)$ is the continuous-time transfer function of the plant.

The focus search function is performed as follows:

1. The switch is at position 1 and the linearly rising signal from the ramp generator causes the laser beam to approach the disk data layer.
2. At the moment when the focusing is reached, the switch moves to position 2 and the control loop is closed.

Figs. C.27 and C.28 present the focus error signal from simulation/real data measuring of the focus processing procedure to automatically close the focus loop. The simulated focus error signal $e_F(t)$, see fig. C.27, has been obtained from realistic DVD simulator using the developed models. Its block diagram is shown in fig. C.26. The disturbance z_d was set to zero during simulation experiment, *i.e.* $z_d = 0$.

This procedure is currently used in DVD players and a good coincidence can be observed between simulated and measured focus error signals $e_F(t)$. It is clear that the common gain approximation of the S-curve, widely used for controller design, does not allow testing of this start-up procedure that is in fact always executed at least at the beginning of DVD/CD disk playing.

C.7.2 Feedback Focus Control Loop

The second application of the models is the same as in linear gain approximation, (in *lock-on range*), for focus control loop, like in fig. C.26. Nevertheless, the designed models are more realistic. Moreover, they are also usable in the small/long jump procedure while the linear gain approximation cannot be used here.

C.7.3 New Photodetector Characteristics Design

The third application of models could be in new photodetector design, since the presented modelling describes which parameters influence the whole S-curve shape.

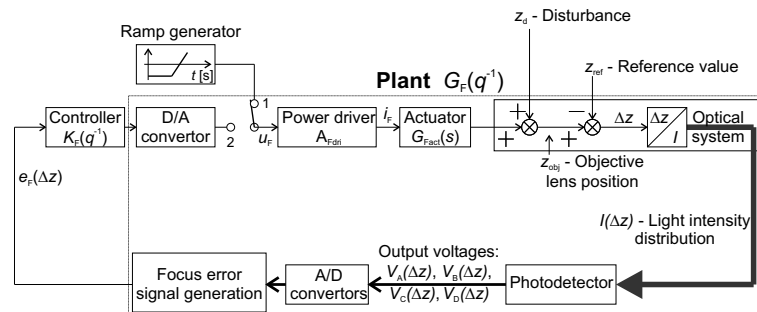


Figure C.26: Block diagram of the focus control loop, focus processing procedure to automatically close the focus loop.

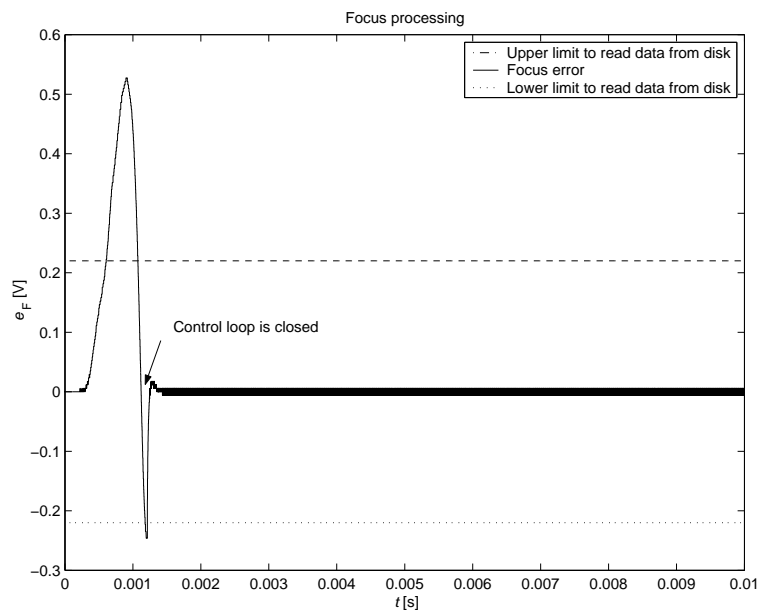


Figure C.27: Focus processing procedure to automatically close the focus loop obtained from simulation where the created analytical model of the S-curve has been included.

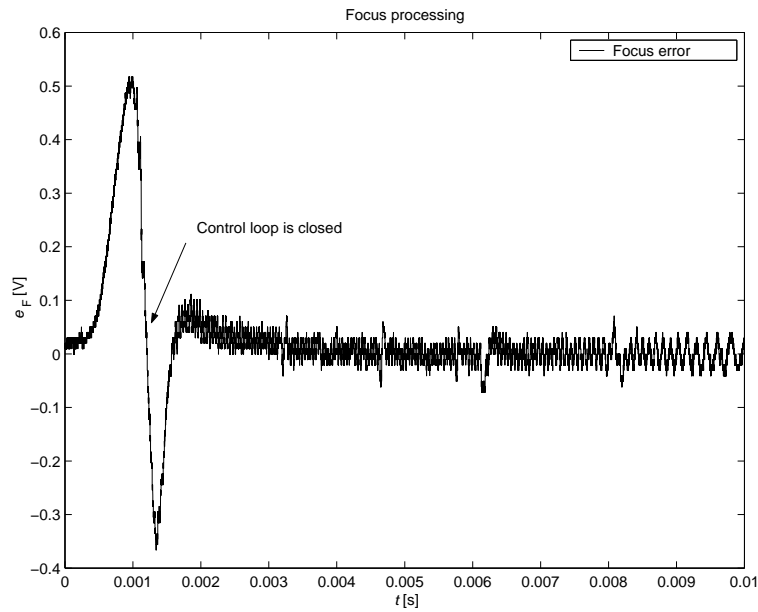


Figure C.28: Focus processing procedure to automatically close the focus loop obtained by measurement on a real DVD-video player from STMicroelectronics.

C.8 Conclusions

An analytical and a numerical models of the non-linear characteristic of a photodetector in a DVD player have been developed.

In addition, a detailed analysis of model parameter variations with respect to the performance and reliability of such models has been shown. The estimation of hardly measurable parameters is described and the results are compared. Comparison with real data, acquired from STMicroelectronics industrial DVD-video player, illustrates the model quality.

To conclude, we believe that the modelling results are advantageous for future improvements of this system in photodetector manufacturing and control system design.

Radial Error Signal Generation

D

There exist many methods for generating the radial error signal e_R from the radial spot position error Δx , for CD and DVD systems, as presented in [BBH⁺85], [Poh95] and [Sta98]. The most common strategies are usually known as 3-beam method, radial push-pull detection, 3-beam push-pull, and radial wobble method.

For the DVD player, a new method for radial error signal generation has been developed, since the smaller track pitch q_{real} and the increased storage capacity have required to enhance the accuracy of the error detection strategy, while minimizing the alignment effort for the optics.

Nowadays the Differential Phase Detection (DPD) method, that corresponds to the known Differential Time Detection (DTD) method, is the one widely used in DVD systems and there are two versions:

1. A first variant of the DTD principle is based on the time-delay differences of the signals from photodiodes A, B, C, D. One possibility to generate the radial error signal e_R , is to calculate the sums of the signals from diagonal pairs of photodiodes ($V_A + V_C$) and ($V_B + V_D$) and then measure the time difference between the rising and falling edges of that two signal sums. This method is usually called DTD2-method.
2. A second variant is to measure the time-delay differences between adjacent pairs of photodiodes B, C and D, A respectively. This will give two intermediate radial error signals, which are then added to form the signal for the filtering. This signal is filtered, by a first-order low pass filter with the cutting-off frequency above 30 kHz, to give the final radial error signal e_R , [ECM01]. In the case of DVD systems, this second variant of DTD method is used to generate the radial error signal. The common denomination is DTD4-method.

The advantage of the second variant is that the amplitude of the radial error signal e_R does not depend on the playback speed. This is because the rate of the edges on the four photodiodes signals is proportional to the playback speed; the values of the time-delay differences between those edges, however, is inversely proportional to the playback speed.

An example of radial error signal generation, used for DVD players, is given in fig. D.1, where the point 2 corresponds to the “spot on-track” situation, besides points 1 and 3 indicate that the focused laser spot is not perfectly centered on the scanned track (Track 0 in fig. D.1). q_{real} denotes the real distance between two consecutive tracks, *i.e.* the *track pitch*.

In this work we consider the second variant because it is the one used on the industrial benchmark. The measured radial error signal $e_R(t)$, in the time-domain, acquired during simply opening the radial control loop without jump back, is shown in fig. D.2.

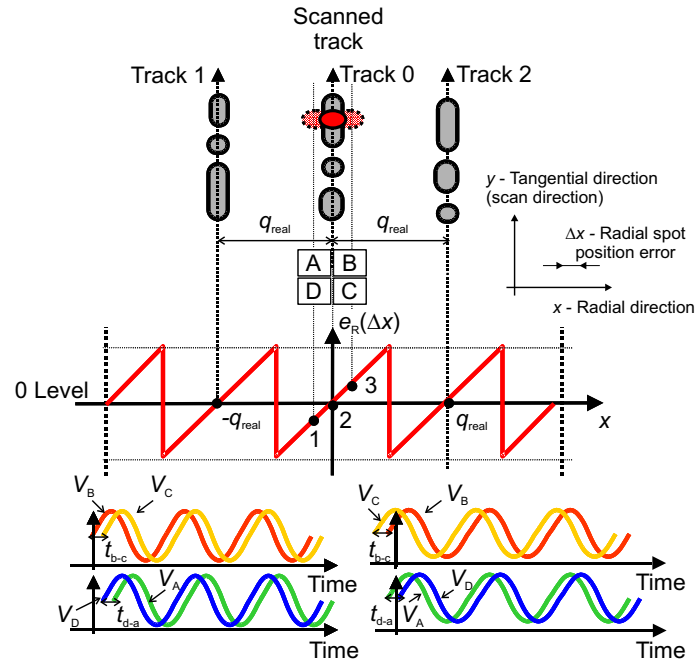


Figure D.1: An example of DTD radial error signal generation used in DVD players (DTD4-method).

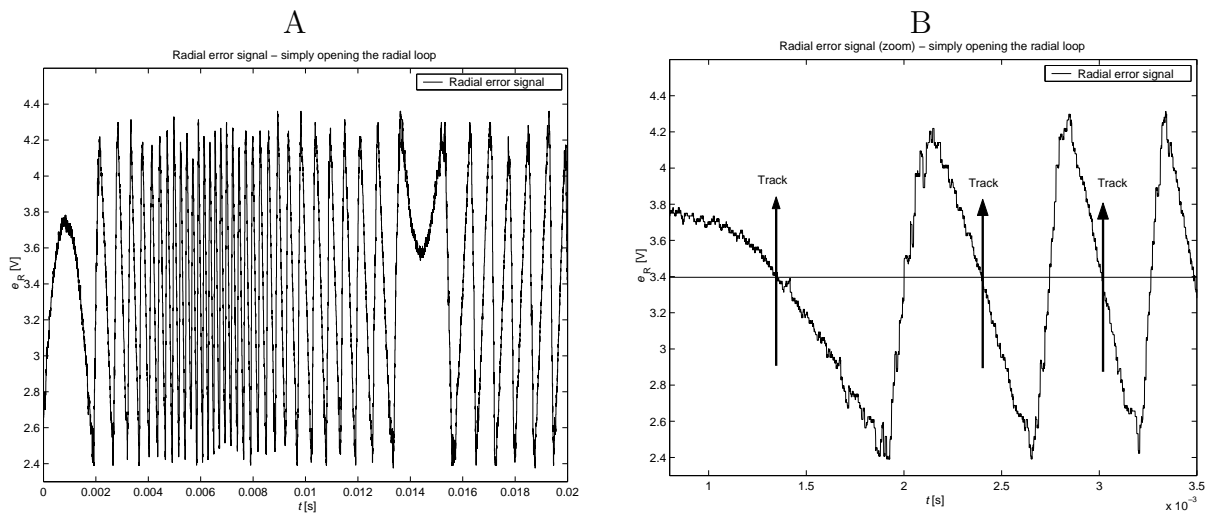


Figure D.2: Case A: An example of radial error signal, in the time-domain, from an industrial DVD-video player—simply opening the radial control loop without jump back, $f_{rot} = 33\text{ Hz}$. Case B: Zoom.

D.1 Modelling of the Radial Error Signal Generation

To understand principles that define image generation in optical systems, it is necessary to have a clear understanding of how the light field propagates through the optical system. This requires the knowledge of scalar diffraction theory and Fourier imaging.

In this section we will not go further in describing the wave nature of light, and how it can be used to derive the propagation of the light field between two arbitrary planes. In fact, a complete and exhaustive analysis of these phenomena is presented in many literature, *e.g.* [BW87], [Bra65], [Pap68], [Goo68], [Sta86], where a description of Huygens' principle as a solution to the scalar wave equation, and Kirchhoff's formula are given, to develop an expression describing the far-field diffraction pattern for a propagating optical field. Our aim is to briefly expose a simple mathematical analysis useful for understanding how the radial error signal is generated in current DVD-video players.

In the literature, there exist some work treating the radial error signal modelling in CD and DVD players, as in [BBH⁺85], [PvKH02], [Poh95] and [Sta98]. Nevertheless, the Differential Phase Detection (DPD) method which corresponds to the Differential Time Detection (DTD) method, widely used in DVD players, is explain more precisely only in few works, see [Bra98].

An easy picture for the understanding of the DPD or DTD method is obtained when the information in the tangential direction (scan direction) is replaced by a single frequency, so that, together with the radial periodicity, a purely two-dimensional grating is present on the disk.

Fig. D.3 shows a typical far field pattern generated by an optical disk with a regular diffracting structure on it. Although such a regular pattern is more representative for the former analog video disks, the *harmonic* analysis with a periodic pattern also gives the correct trends for digital CDs and DVDs.

In fig. D.3 the various quasi-Direct Current (DC) and Alternating Current (AC) signals are depicted that are present in the far-field on a quadrant detector (quadrants "a", "b", "c" and "d"). In this case, the quadrant detector is placed to the plane of the objective lens and its diameter is equal to the objective lens diameter. This means that the presented approximation does not contain the influence of the other optical elements because the light intensity is measured by four photodiodes A, B, C, D in reality, see 3.2.

The centers of the diffracted orders are off-set by a distance $\pm x_0 = \pm \lambda / (q_{\text{real}} NA)$ and $\pm y_0 = \pm \lambda / (p_{\text{real}} NA)$ in the x and y directions in fig. D.3. NA is the numerical aperture of the objective lens and λ is the laser beam wavelength. The detection region is the inner part of the zeroth order (heavy circle). x and y denote the radial and scanned directions, respectively.

Each modulation term depends on the diffraction direction. Therefore the *radially* diffracted orders show a varying intensity level in an overlapping region with the zeroth order that is proportional to $\cos(\psi \pm \phi_r)$. The reference phase $\psi = \psi_{(+1,0)} - \psi_{(0,0)}$ between the first and zeroth orders is determined by the disk structure; it varies from $\pi/2$ for very shallow structures to π if the phase depth $\Delta\phi$ of the optical effects attains the value of π .

The quantity $\phi_r = 2\pi\Delta x/q_{\text{real}}$ is proportional to the radial error Δx , where $q_{\text{real}} = 0.74 \mu\text{m}$ is the track pitch. The sign of ϕ_r depends on the order number ± 1 of the diffracted order.

Similarly, the overlapping regions along the *tangential* direction (scan direction) show an intensity variation proportional to $\cos(\psi \pm 2\pi f_{\text{pit}} t)$, due to the scanning of the tan-

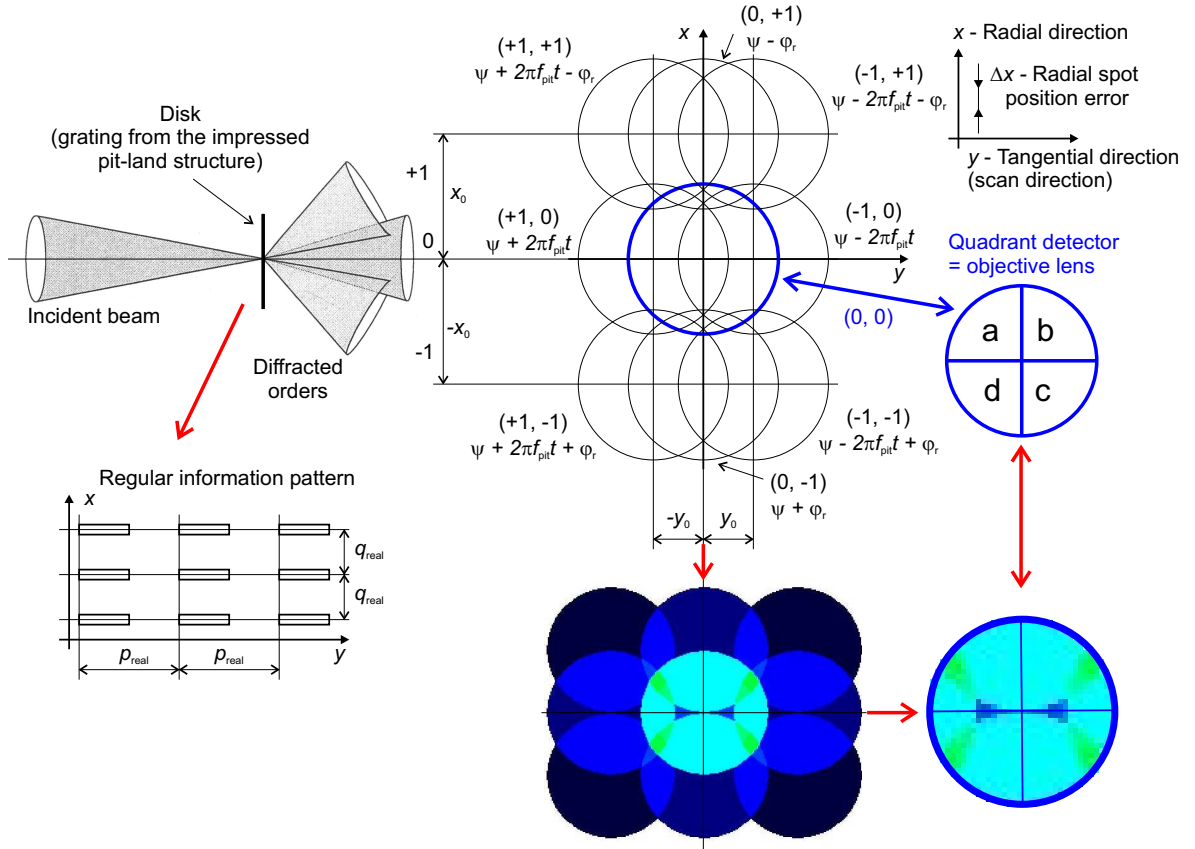


Figure D.3: Schematic drawing illustrating the far field pattern generated by an optical disk with a regular information pattern.

gentially stored periodic pattern at the uniform speed v_a . Here, the quantity:

$$f_{\text{pit}} = \frac{v_a N}{p_{\text{real}}} \quad (\text{D.1})$$

denotes a temporal frequency generated by scanning the spiral track with periodic information, N is an over-speed factor and p_{real} is a real distance between periodic pits on a track. Obliquely diffracted orders show a mixed phase contribution and carry interesting information for deriving the radial error signal e_R .

The expression for the high-frequency content of the signal in quadrant “a” of the objective lens has the following expression:

$$\begin{aligned} S_a(t) &= W_{(0,0;+1,0)} \cos(\psi + 2\pi f_{\text{pit}} t) & (1) \\ &+ W_{(0,+1;+1,+1)} \cos(2\pi f_{\text{pit}} t) & (2) \\ &+ W_{(0,0;+1,+1)} \cos(\psi + 2\pi f_{\text{pit}} t - \phi_r) & (3) \\ &+ W_{(0,+1;+1,0)} \cos(2\pi f_{\text{pit}} t - \phi_r), & (4) \end{aligned} \quad (\text{D.2})$$

where the weighting factors W are proportional to the product of the diffracted orders amplitudes and to the size of the overlapping regions on the quadrant detector. Similar expressions are obtained for the signals from the quadrants “b”, “c” and “d” with the appropriate algebraic signs in the term $(2\pi f_{\text{pit}} t)$ and phase term ϕ_r .

The Hopkins theory, presented in [Hop79], is used to compute the phase terms ψ in (D.2). These values are evaluated to be $\psi_{(0,0)} = 8^\circ$ while $\psi_{(+1,0)} = \psi_{(0,+1)} = \psi_{(+1,+1)} = 160^\circ$. These values very well apply to both the pit structure on a CD and on a DVD. One can see that the main contributions to the high-frequency signal come from the components 1) and 3) of (D.2).

Hence, for each photodetector quadrant (see fig. D.3), we can write:

$$\begin{aligned} S_a(t, \phi_r) &= \cos(2\pi f_{\text{pit}}t + \psi) + \alpha_3 \cos(2\pi f_{\text{pit}}t - \phi_r + \psi), \\ S_b(t, \phi_r) &= \cos(2\pi f_{\text{pit}}t - \psi) + \alpha_3 \cos(2\pi f_{\text{pit}}t + \phi_r - \psi), \\ S_c(t, \phi_r) &= \cos(2\pi f_{\text{pit}}t - \psi) + \alpha_3 \cos(2\pi f_{\text{pit}}t - \phi_r - \psi), \\ S_d(t, \phi_r) &= \cos(2\pi f_{\text{pit}}t + \psi) + \alpha_3 \cos(2\pi f_{\text{pit}}t + \phi_r + \psi), \end{aligned} \quad (\text{D.3})$$

where α_3 is a factor less than unity, *i.e.* $\alpha_3 < 1$, that accounts for the relatively small contribution of the diagonal orders with respect to the tangential orders. The consequence of this simplification is that the final model can not be used for the high frequency read-out signal generation modelling (only phase's elements are taken into account in (D.3) from (D.2)).

The standard DTD-signal is derived by firstly summing the light intensities on a diagonal pair of detector, then detecting the signals relative time shift. It is also possible to compare the phase difference between signals delivered by each single detector quadrant.

Thus, we can summarize the possible set of DTD methods as follows:

1. $\Delta t_2(t) = t_{a+c} - t_{b+d}$: The time difference between the sum of signals from quadrants "a+c" and the sum of signals from quadrants "b+d" is taken. This is the standard DTD-method based on the difference of diagonal quadrant signals, and it is called DTD2-method.
2. $\Delta t_4(t) = t_{b-c} + t_{d-a}$: This difference requires four independent high-frequency detectors to be computed. When the laser spot is on the track, each difference signal "d-a" and "b-c" is zero. This method is usually called DTD4-method.
3. $\Delta t_{4a}(t) = t_{b-a} + t_{d-c}$: This signal is comparable to the previous one. The time differences "b-a" and "d-c" are not zero when the laser spot is on the track, since these differences depend among other factors on the optical pit depth. The common name is DTD4a-method.

It is proved that the DTD4-method remains unaffected by de-focusing less than the DTD2-method [Bra98], as it is shown in fig. D.4. In particular, the slope of radial error e_R remains unaltered by de-focusing effects near the zero spot position error Δx . Therefore, the DTD4 method is the one commonly used in DVD players for radial error signal generation.

After the time differences $\Delta t_4(t)$ being generated, a IIR low-pass filter is used to integrate it to give the final radial error signal $e_R(t)$. This filter can be approximated by a simple RC low-pass filter, as follows:

$$\Delta t_4(t) = \frac{1}{K} \left(e_R(t) + \frac{1}{2\pi f} \frac{de_R(t)}{dt} \right), \quad (\text{D.4})$$

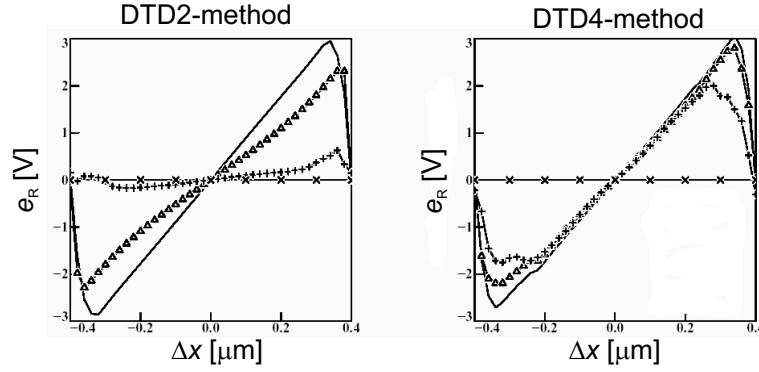


Figure D.4: The radial error signals generated by DTD2 and DTD4 method with different amounts of de-focus. Draw line: $\Delta z = -0.903 \mu\text{m}$, triangles: $\Delta z = 0 \mu\text{m}$, crosses: $\Delta z = +0.903 \mu\text{m}$.

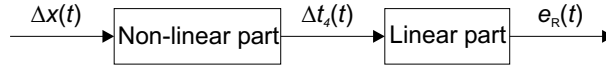


Figure D.5: Block-scheme of the radial error generation model.

where f is the filter cut-off frequency and K is a gain of the digital signals processing which gives the total gain value from the measured time differences t_{b-c} and t_{d-a} to the radial error signal e_R .

With this assumptions, the radial error signal generation model can be split-up in two different parts: The first part which is non-linear and the second part, considered as linear. In fig. D.5 a block-scheme of the radial error generation model is shown.

The model output is the radial error signal $e_R(t)$, which is given by the following expression:

$$\Delta t_4(\Delta x, \theta, \varrho, t) = \frac{1}{K} \left(e_R(\Delta x, \theta, \varrho, t) + \frac{1}{2\pi f} \frac{de_R(\Delta x, \theta, \varrho, t)}{dt} \right), \quad (\text{D.5})$$

where the model input $\Delta x(t)$, the unknown parameters vector θ and the known parameters vector ϱ are emphasized.

Equation (D.5) describes the linear part of fig. D.5. Its non-linear part, which gives the dependency of $\Delta t_4(t)$ from the radial spot position error $\Delta x(t)$ is described by:

$$\Delta t_4(\Delta x, \theta, \varrho, t) = \left(S_b(\Delta x, \theta, \varrho, t) - S_c(\Delta x, \theta, \varrho, t) \right) + \left(S_d(\Delta x, \theta, \varrho, t) - S_a(\Delta x, \theta, \varrho, t) \right), \quad (\text{D.6})$$

where S_a , S_b , S_c , S_d (see (D.3)) represent the phase contributions of the detector quadrants ‘‘a’’, ‘‘b’’, ‘‘c’’ and ‘‘d’’, respectively. Hence, the model is characterized by the following parameters:

- The input track displacement $\Delta x(t)$ along the disk radial direction.
- The radial error $e_R(\Delta x, \theta, \varrho, t)$.

- The unknown parameters vector $\theta = [\alpha_3, K]$.
- The known parameters vector $\varrho = [f_{\text{obj}}, NA, \psi_{00}, \psi_{10}, \lambda, q_{\text{real}}, T_{\text{pit}}, p_{\text{real}}, v_a, N, f]$.

Finally, one can verify that the radial spot position error $\Delta x(t)$ appears as $\phi_r = 2\pi\Delta x/q_{\text{real}}$ in (D.3).

Summary: Estimation of unknown parameters of the radial error signal generation model from measured data is hard task since:

- Linear models can be computed only by measuring the radial S-curves in open-loop (radial control loop opened, focus control loop closed). In this condition it is possible to correctly focus the laser spot with respect to the disk layer, but the radial spot position error Δx is affected by unknown periodical signal caused by disk displacement on the spindle, as it will be explained in section 4.3.2. This perturbation makes the laser cross tracks inward and outward with respect to the center of the track, resulting in not very reliable time acquisitions.
- Characteristics of the periodic perturbation entering at the plant output are not known exactly. Particularly, it is hard to know accurately the values of its phase and amplitude when radial S-curves measurements start. This makes difficult to estimate unknown parameters vector $\theta = [\alpha_3, K]$ of non-linear mathematical model.
- Sensor noise and quantization effects introduced by the D/A converters, at the output of the control chain, lead to very ill conditioned measurements which amplify inaccuracy of model parameters estimation algorithms.

The curve fitting procedure has been applied on the measured data acquired during the radial control loop opening, see fig. D.2, in order to estimate unknown model parameters, as illustrated in fig. D.6. The obtained results seems to be right but the estimated model parameters are not valid for the other tracks intersections because they are depending on the unknown periodical signal caused by disk displacement on the spindle.

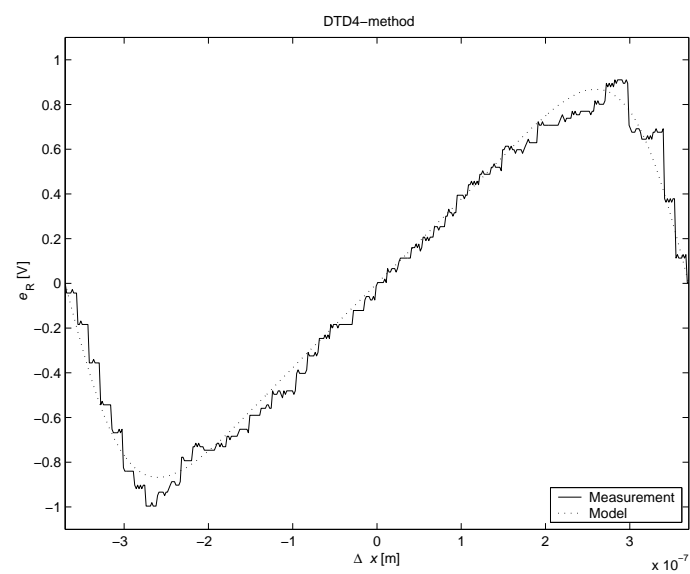


Figure D.6: An example of the radial error signal generation modelling.

The binary data encoded on a DVD surface can be retrieved by summing-up the light intensity retrieved by the four photodiodes A, B, C and D. In such a way, a High-Frequency signal (HF), modulated by the disk relief structure, is derived.

The so called *eye pattern* is presented in fig. E.1, and it is obtained by superposing slices of the HF signal, synchronized with the Phase Locked Loop (PLL) clock frequency, several times on an oscilloscope screen. Since the light reflected by pits is strongly affected by destructive interference, the HF signal reaches minimum and maximum values in correspondence of pit and land structures, respectively as can be seen in fig. E.1.

The symbols impressed on the disk surface as sequence of pits and lands, are converted into retrievable data, by using the Eight to Sixteen modulation, which belongs to the class of Run-Length Limited (RLL) codes, described in [Imm91] and [Sta98].

RLL codes are characterized by constraints in the symbol coding, where the minimum and the maximum number of identical symbols following each other is limited. In addition, from RLL codes it is also possible to retrieve the whole system clock frequency (self-clocking sequence). For the DVD system, the minimum and maximum run-lengths are set equal to 3 and 14, respectively [ECM01].

The minimum and the maximum fundamental frequencies of the modulation pattern can be computed by considering the channel bit rate, that is the frequency at which the binary sequence is coded, the run-length k , and the over-speed factor N , as indicated by:

$$f_{kT_{\text{pit}}} = \frac{f_{\text{ch}}}{2k} N, \quad (\text{E.1})$$

where $f_{\text{ch}} = 26.16$ Mbit/s is the channel bit rate computed for $N = 1$, and $k = 3, \dots, 14$. For $N = 1.5$, as on the industrial benchmark, we obtain $1.40 \text{ MHz} \leq f_{kT_{\text{pit}}} \leq 6.54 \text{ MHz}$.

From fig. E.1 some relevant parameter, characterizing the HF signal, can be distinguished:

- I_{14} and I_3 are the modulation peak to peak values, generated by the largest and the shortest lengths of pit or land, respectively when a $14T_{\text{pit}}$ or a $3T_{\text{pit}}$ symbol is read on the disk.
- I_{14H} and I_{3H} are the highest reflectivity amplitudes of the HF signal, generated by the largest and the shortest lengths of pit or land respectively when a $14T_{\text{pit}}$ or a $3T_{\text{pit}}$ symbol is read on the disk.
- I_{14L} and I_{3L} are the lowest reflectivity amplitudes of the HF signal, generated by the largest and the shortest lengths of pit or land respectively when a $14T_{\text{pit}}$ or a $3T_{\text{pit}}$ symbol is read on the disk.
- The zero level is the no reflection level without disk, and it is also called *dark level*.

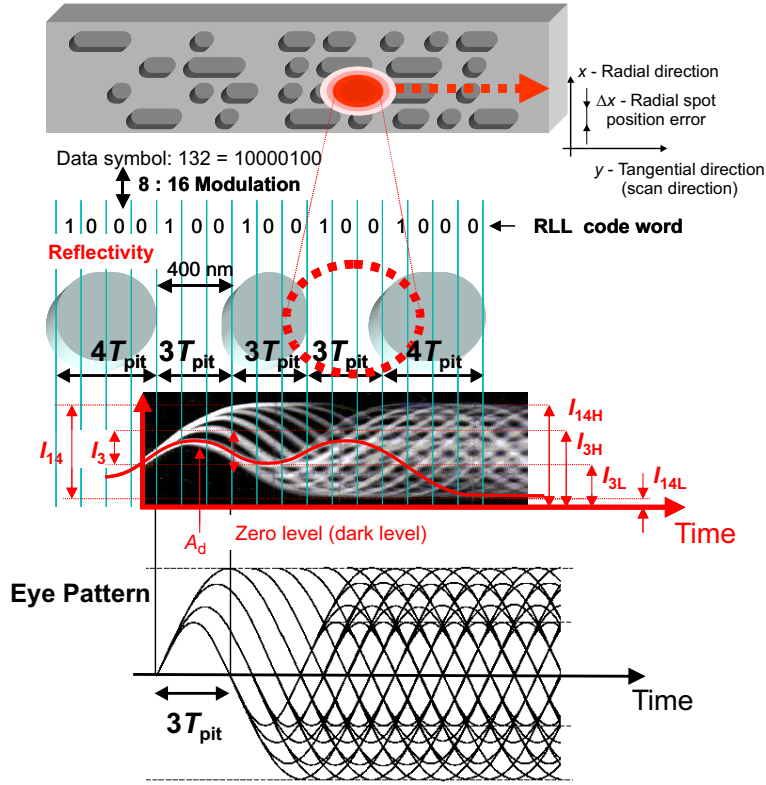


Figure E.1: Example of HF generation for a pit/land impressed structure.

The above mentioned parameters shall satisfy the following relations [ECM01]:

$$\frac{I_{14}}{I_{14H}} = 0.60 \quad \text{minimal}, \quad (\text{E.2})$$

$$\frac{I_3}{I_{14}} = 0.15 \div 0.20 \quad \text{minimal} \quad (\text{E.3})$$

and

$$\frac{(I_{14H\max} - I_{14H\min})}{I_{14H\max}} = 0.33 \quad \text{maximal}. \quad (\text{E.4})$$

Finally, a detection level A_d is applied to the HF signal, to let the system correctly recover the digital information. This level should satisfy the so called *asymmetry condition*, defined in [ECM01]:

$$-0.05 \leq \frac{(I_{14H} + I_{14L}) - (I_{3H} + I_{3L})}{2(I_{14H} - I_{14L})} \leq 0.15. \quad (\text{E.5})$$

Equations (E.2), (E.3), (E.4), and (E.5) represent conditions that should be satisfied, to confirm that the disk conforms to the standards fixed in [ECM01].

It is interesting to notice that the pit/land structure impressed on a DVD allows to retrieve, at the same time, the recorded data (used for audio and video reproduction), and the servo signals (needed to control the spot position during playback).

This is why an accurate description of data read-out, as well as error signal generation mechanisms, is fundamental for control design purposes.

E.1 Modelling of the Read-out Signal Generation

The method used in a DVD system for data read-out is that of the scanning microscope [BBH⁺85], [Bra98].

The principle is based on the fact that sequence of pits and lands forms along the disk a two-dimensional diffraction grating, as shown in fig. 3.3 and fig. 3.4. This grating splits the incident light into multiple diffraction orders, as shown in fig. D.3. Data read-out is accomplished by capturing, with photodiodes, the amount of light which goes through the objective lens, *i.e.* $HF(t) = V_A(t) + V_B(t) + V_C(t) + V_D(t)$. Therefore one can approximate the read-out data signal generation by sum of all contributions captured by the quadrant detector in fig. D.3, *i.e.* $HF(t) = S_a(t) + S_b(t) + S_c(t) + S_d(t)$. This detection method is typically used in all DVD and CD systems and it is usually called Central Aperture (CA) detection method.

It can be shown [BBH⁺85] that the intensity of the light received by the quadrant detector can be approximated by:

$$I_{\text{det}}(t) \approx 2A_0^2 \left[1 + \left(\frac{A_1}{A_0} \right)^2 + 2 \left(\frac{A_1}{A_0} \right) \cdot MTF(y_0) \cdot \cos \psi_{10} \cos \left(\frac{2\pi N v_a t}{p_{\text{real}}} \right) \right], \quad (\text{E.6})$$

where A_0 and A_1 are the amplitudes of the zeroth and first diffracted order respectively, ψ_{10} is the phase shift existing between them, and MTF is the Modulation Transfer Function (MTF) defined by:

$$MTF(y_0) = \frac{2}{\pi} \arccos \left(\frac{y_0}{2} \right) - \frac{y_0}{\pi} \sqrt{1 - \left(\frac{y_0}{2} \right)^2}. \quad (\text{E.7})$$

In the case of CA detection, $MTF(y_0)$ is basically a measure of the cumulated overlapping areas depicted in fig. D.3 and depends on the diffracted order off-set distance y_0 . One can also define the spatial frequency along the track as follows:

$$f_{\text{sp}} = y_0 \frac{NA}{\lambda} = \frac{1}{p_{\text{real}}}. \quad (\text{E.8})$$

MTF characterizes the low-pass filtering process underground by the time-modulated third term from (E.6). Finally, it is important to mention that possible aberrations in the optical system will not affect the MTF but do introduce phase distortions in the read-out signal [Goo68].

There exists very useful analysis of the read-out signal generation that has been presented in [Hop79]. It is based on the approximation of the disk structure by the regular information pattern described in section D. It allows to model the different aberrations of the focused laser beam and to explain the crosstalks phenomena. Nevertheless, this analysis can not be used to the focus error signal generation, neither to the radial error signal generation because it does not describe the four signals on the quadrant detector “a”, “b”, “c”, “d” in fig. D.3. It only allows to approximate the whole light intensity received by the quadrant detector (objective lens) analytically.

The second method that can be used to read-out signal generation modelling is based on the direct numerical solution of diffraction, see *e.g.* [Sta86], [Nut97]. In this case, a general disk information pattern can be used for a complex analysis. This methodology is complicated and numerically very time consuming in comparison to the analytical solution

by Hopkins's analysis. This method offers to generate the radial error signal, analyze the crosstalks phenomena, and obtain more accurate S-curve characteristic near the focused spot position, (*i.e.* $|\Delta z| < z_{\max}$, where z_{\max} is the focused depth, $z_{\max} = 0.903 \mu\text{m}$ for DVD).

Many software products are available to deal with the diffraction theory. Moreover, the modelling of read-out signal generation is not required for the control system design in the focus and radial loop. Therefore, the modelling of read-out signal generation is not more analyzed in this research.

Glossary

Acronyms and Abbreviations

A, B, C, D	four photodiodes of the photodetector
AC	Alternating Current
A/D	Analog to Digital converter
AF-CLOE	Adaptive Filtered-Closed-Loop Output-Error identification method
ARMAX	AutoRegresive with eXogenous variable and Moving Average noise model
ARX	AutoRegresive with eXogenous variable model
BW	Band-Width
CA	Central Aperture detection method
CD	Compact Disk
CD-ROM	Compact Disk-Read Only Memory
CIRC	Cross Cross Interleaved Reed-Solomon Code
CL	Closed-Loop
CLOE	Closed-Loop Output-Error identification method
CLV	Constant Linear Velocity
D/A	Digital to Analog converter
DC	Direct Current
DL	Dual Layer disk
DPD	Differential Phase Detection method
DTD	Differential Time Detection method
DTD2	Differential Time Detection method 2
DTD4	Differential Time Detection method 4
DSP	Digital Signal Processing (or Processor)
DVD	Digital Versatile Disk
DVD-ROM	Digital Versatile Disk-Read Only Memory
DVD-5	single-layer single-side Digital Versatile Disk
ECC	Error Correction Code
EFM	Eight to Fourteen modulation
F-CLOE	Filtered-Closed-Loop Output-Error identification method
FIR	Finite Impulse Response
FFT	Fast Fourier Transformation
FWHM	Full Width at Half Maximum
G-CLOE	Generalized-Closed-Loop Output-Error identification method
GM	Gradient Method
GM-2	Gradient Method 2 for state-space system
HF	High Frequency signal
IGM	Improved Gradient Method
IGM-1	Improved Gradient Method 1 for state-space system
IGM-2	Improved Gradient Method 2 for state-space system

IIR	Infinite Impulse Response
LMI	Linear Matrix Inequality
LQ	Linear Quadratic controller
LS	Least Square method
LTI	Linear Time Invariant system
MIMO	Multiple-Input Multiple-Output system
MTF	Modulation Transfer Function
MRAS	Model Reference Adaptive System
ODE	Ordinary Different Equation method
OE	Output-Error model
OL	Open-Loop
OPU	Optical Pick-up Unit
PAA	Parameter Adaptation Algorithm
PD	Proportional and Derivative controller
PID	Proportional, Integral and Derivative controller
PSD	Power Spectral Density
PRBS	Pseudo-Random Binary Sequence
RLL	Run-Length Limited code
RLS	Recursive Least Square method
RLS-2	Recursive Least Square method 2 for state-space system
RS	R-S controller (one degree of freedom digital controller)
RS2	R-S controller of the second order, $n_R = 2, n_S = 2$
RS3	R-S controller of the third order, $n_R = 3, n_S = 3$
RS4	R-S controller of the fourth order, $n_R = 4, n_S = 4$
RSPC	Reed Solomon Product Code
RST	R-S-T controller (a two degree of freedom digital controller)
SISO	Single-Input Single-Output system
SNR	Signal to Noise Ratio
SPR	Strictly Positive Real transfer function
TITO	Two-Input Two-Output system
X-CLOE	eXtended-Closed-Loop Output-Error identification method
SL	Single Layer disk
μC	micro Controller

Symbols

a_{\max}	maximum acceleration of the end point x_{\max} on data zone, in m/s^2
$A_{F\text{dri}}$	focus power driver gain, in V/V
$A_{F\text{dri}1}$	first focus power driver gain, in V/V
$A_{F\text{dri}2}$	second focus power driver gain, in V/V
$A_{F\text{opt}}$	optical gain in focus loop, in V/m
$A_{R\text{dri}}$	radial power driver gain, in V/V
$A_{R\text{dri}1}$	first radial power driver gain, in V/V
$A_{R\text{dri}2}$	second radial power driver gain, in V/V
$A_{R\text{opt}}$	optical gain in radial loop, in V/m
d	pure time delay
d_a	aperture diameter of the the objective lens, in m

d_{Airy}	Airy disk diameter, in m
$d_{\text{Airy(FWHM)}}$	Airy disk diameter (Full Width at Half Maximum), in m
$d_{e^{-2}}$	diameter of the laser spot for the Gaussian intensity distribution, in m
d_{Mreal}	real mechanical depth, in m
D_{F}	focus actuator dumping constant of surrounding medium, in Ns/m
D_{R}	radial actuator dumping constant of surrounding medium, in Ns/m
f	frequency, in Hz
f^{BW}	desired closed-loop band-width, in Hz
f_{D}	the slowest vibration frequency of the plant model, in Hz
f_{F}	frequency of complex pole in characteristic polynomial, in Hz
f_{Fn}	natural undamped frequency of the second order focus system, in Hz
f_{cr}	cross-over frequency of the compensated open-loop, in Hz
f_{LL1}	lead break frequency of reference lead-lag controller $K_{\text{ref}}(s)$, in Hz
f_{LL2}	lag break frequency of reference lead-lag controller $K_{\text{ref}}(s)$, in Hz
f_{obj}	(objective) focal length of the objective lens, in m
f'_{obj}	image focal length of the objective lens, in m
f_{rot}	disk rotational frequency, in Hz
$f_{\text{rot_max}}$	maximal disk rotational frequency during playback, in Hz
$f_{\text{rot_min}}$	minimal disk rotational frequency during playback, in Hz
f_{s}	sampling frequency used for digital computation in the servo DSP, in Hz
f_{Rn}	natural undamped frequency of the second order radial system, in Hz
f_{S}	frequency of complex pole in polynomial H_{S} , in Hz
g_{LL}	actual lead-lag controller gain
K	proportional gain, in general
K_{Fs}	focus actuator elastic constant of a spring, in N/m
K_{FDC}	focus actuator DC sensitivity
K_{Fe}	focus actuator back efm constant, in Vs/m
K_{Ff}	focus actuator force constant, in N/A
K_{Rs}	radial actuator elastic constant of a spring, in N/m
K_{RDC}	radial actuator DC sensitivity
K_{Re}	radial actuator back efm constant, in Vs/m
K_{Rf}	radial actuator force constant, in N/A
j	$\sqrt{-1}$
L_{F}	focus coil inductance, in H
L_{land}	land length, in m
L_{pit}	pit length, in m
L_{R}	radial coil inductance, in H
M_{F}	focus actuator moving mass, in kg
M_{R}	radial actuator moving mass, in kg
n_{A}	order of the plant numerator $G(q^{-1})$
n_{B}	order of the plant denominator $G(q^{-1})$
n_{r}	refractive index of objective space
n_{p}	refractive index of the transparent substrate
n_{R}	order of the RS controller numerator $R(q^{-1})$
\widehat{n}_{R}	order of the RS controller numerator $R(q^{-1})$ after controller order reduction
n_{S}	order of the RS controller denominator $S(q^{-1})$

\widehat{n}_S	order of the RS controller denominator $S(q^{-1})$ after controller order reduction
N	over-speed factor of the disk scanning velocity
NA	numerical aperture of the objective lens, in m
q, q^{-1}	forward and backward shift operators
q_{real}	track pitch, in m
$p_{\text{LL1}}, p_{\text{LL2}}$	actual lead-lag controller poles
R_{F}	focus coil resistance, in Ω
R_{obj}	objective lens radius, in m
R_{R}	radial coil resistance, in Ω
s	Laplace complex variable
S_{low}	minimum output sensitivity function magnitude for disk manufacturing at the low rotational frequencies
$ S_{\text{yp}} _{\text{max}}$	maximum of output sensitivity function magnitude, in dB
$ S_{\text{yp}} _{14.4}$	output sensitivity magnitude at $f_{\text{rot}} = 14.4$ Hz, in dB
$ S_{\text{yp}} _{34.7}$	output sensitivity magnitude at $f_{\text{rot}} = 34.7$ Hz, in dB
$ S_{\text{up}} _{\text{max}}$	maximum of input sensitivity function magnitude, in dB
t	time, in s
t_{r}	rise time of the system closed-loop response, in s
T_{d}	derivation-time constant, in s
T_{i}	integration-time constant, in s
T_{pit}	channel bit length, in m
T_{s}	sampling period used for digital computation in the servo DSP, in s
v_{a}	constant linear velocity during playback, in m/s
v_{x}	desired velocity of the laser spot in radial direction x , in m/s
x	direction of the orthogonal coordinate system
$x_{\text{d_high}}$	maximal radial deviation for high rotational frequencies given for disk manufacturing, in m
$x_{\text{d_low}}$	maximal radial deviation for low rotational frequencies given for disk manufacturing, in m
$x_{\text{d_max}}$	disk eccentricity, <i>i.e.</i> the maximal radial disk displacement on the spindle, in m
x_{max}	maximal radius of data zone, in m
x_{min}	minimal radius of data zone, in m
y	direction of the orthogonal coordinate system
z	direction of the orthogonal coordinate system
$z_{\text{d_high}}$	maximal vertical deviation for high rotational frequencies given for disk manufacturing, in m
$z_{\text{d_low}}$	maximal vertical deviation for low rotational frequencies given for disk manufacturing, in m
$z_{\text{d_max}}$	maximal vertical deviation at the maximal data zone radius, in m
$z_{\text{d_min}}$	minimal vertical deviation at the minimal data zone radius, in m
$z_{\text{LL1}}, z_{\text{LL2}}$	actual lead-lag controller zeros
α	multiple coefficient
γ	width of the pit, in m
γ_{F}	real pole in characteristic polynomial
γ_{R}	real pole in polynomial H_{R}

$\Delta\phi$	phase margin, in $^\circ$
ΔM	modulus margin, in dB
$\Delta\tau$	delay margin, in s
Δx_{\max}	maximal radial spot position error to avoid loosing the data read-out signal during playback, in m
Δz_{\max}	maximal vertical spot position error (<i>i.e.</i> the objective lens focus depth) to avoid loosing the data read-out signal during playback, in m
λ	wavelength of the laser beam in air, in m
ξ_D	damping ratio of the slowest plant model vibration frequency
ξ_F	damping ratio of complex pole in characteristic polynomial
ξ_S	damping ratio of complex pole in polynomial H_S
ϕ_{\max}	opening angle of the objective lens, in $^\circ$
φ_F	initial face of periodical disturbance model in the focus loop, in rad
φ_R	initial face of periodical disturbance model in the radial loop, in rad
ω	angular velocity, in rad/s
ω_{cr}	cross-over angular velocity of the compensated open-loop, in rad/s
ω_{rot}	disk angular velocity, in Hz

Notational conventions

$A(q^{-1})$	plant transfer operator numerator, as function of time operator
$A_0(\theta)$	matrix of discrete-time state-space model
A_c	matrix of state-space model in controller form
A_{ob}	matrix of state-space model in observer form
$B(q^{-1})$	plant transfer operator denominator, as function of time operator
$B_0(\theta)$	matrix of discrete-time state-space model
B_c	matrix of state-space model in controller form
B_{ob}	matrix of state-space model in observer form
$b(t)$	measurement noise, as function of time, in general
$b(K, G)$	generalized stability margin, in general
$\widehat{b(K, G)}$	generalized stability margin after controller order reduction, in general
$C_0(\theta)$	matrix of discrete-time state-space model
$D_0(\theta)$	matrix of discrete-time state-space model
$D_\theta^i(k)$	parametric distance
\mathcal{D}	domain of admissible parameters related to the model set
$e(t)$	zero-mean Gaussian white noise, as function of time, in general
$e_F(t)$	focus error signal, as function of time, in V
$e_F(\Delta z)$	focus error signal, as function of Δz , in V
$e_R(t)$	radial error signal, as function of time, in V
$e_R(\Delta x)$	radial error signal, as function of Δx , in V
$F(k)$	matrix adaptation gain
$F_0(\theta_c)$	matrix of continuous-time state-space model
$G(s), G(q^{-1})$	transfer operator of the model, in general
$\hat{G}(q^{-1})$	transfer operator of estimated model, in general
$G_F(s), G_F(q^{-1})$	focus loop plant transfer function
$G_{Fact}(s)$	focus actuator transfer functions, in Laplace domain

$G_R(s), G_R(q^{-1})$	radial loop plant transfer function
$G_{\text{Ract}}(s)$	radial actuator transfer functions, in Laplace domain
$G_S(q^{-1})$	transfer operator of the true plant (system)
$G_0(\theta_c)$	matrix of continuous-time state-space model
$H(q^{-1})$	transfer operator of the noise model
$H_s(s)$	open-loop compensated transfer function, in Laplace domain
$H_{\text{real}}(s)$	real open-loop transfer function, in Laplace domain
$H_f(q^{-1})$	transfer operator
$H'_f(q^{-1})$	transfer operator
$H_R(q^{-1})$	fixed part of the RS controller transfer function numerator
$H_S(q^{-1})$	fixed part of the RS controller transfer function denominator
$\hat{H}(q^{-1})$	estimated transfer operator of the noise model
$i_F(t)$	electrical current through the focus actuator coil, as function of time, in A
$i_R(t)$	electrical current through the radial actuator coil, as function of time, in A
$I(r)$	light intensity distribution, as function of the normalized focus spot radius r
$I(\Delta x)$	light intensity distribution, as function of Δx
$I(\Delta z)$	light intensity distribution, as function of Δz
$J_n(x)$	Bessel function
$J(k)$	criterion function
$K(s), K(q^{-1})$	controller transfer function, in general
$K_{\text{act}}(q^{-1})$	actual controller applied into the industrial benchmark, as function of time operator
$K_F(s), K_F(q^{-1})$	focus loop controller transfer function
$K_{\text{ref}}(s)$	reference lead-lag controller, in Laplace domain, in general
$K_R(s), K_R(q^{-1})$	radial loop controller transfer function
$K_{\text{RS2}}(q^{-1})$	RS controller of the 2th order, as function of time operator
$K_{\text{RS3}}(q^{-1})$	RS controller of the 3th order, as function of time operator
$K_{\text{RS4}}(q^{-1})$	RS controller of the 4th order, as function of time operator
$\mathcal{L}(\cdot)$	scalar-valued function
M	matrix of scalar elements, in general
\mathcal{M}	model structure
$P(q^{-1})$	characteristic polynomial of the closed-loop, as function of time operator
$R(q^{-1})$	control transfer function numerator, as function of time operator
$R'(q^{-1})$	part of control transfer function numerator as the solution of the Bezout equation, as function of time operator
$S(q^{-1})$	control transfer function denominator, as function of time operator
$S'(q^{-1})$	part of control transfer function denominator as the solution of the Bezout equation, as function of time operator
$S_{ij}(K, G)(s)$	sensitivity function, in Laplace domain, in general
$S_{\text{yp}}(e^{-j\omega})$	output sensitivity function, as function of frequency
$S_{\text{up}}(e^{-j\omega})$	input sensitivity function, as function of frequency
$S_{\text{yu}}(e^{-j\omega})$	input-output transfer operator, as function of frequency, in general
$S_{\text{yw}}(e^{-j\omega})$	output sensitivity function with respect to an input disturbance,

	as function of frequency
$S_{yr}(e^{-j\omega})$	complementary sensitivity function, as function of frequency
$S_{yb}(e^{-j\omega})$	noise sensitivity function, as function of frequency
T	similarity matrix
$\mathbf{T}(K, G)$	transfer matrix, in general
$p(t)$	output disturbance signal, as function of time, in general
\mathcal{P}	family of plant models
$r(t)$	reference signal, as function of time, in general
$u(t)$	controller output signal, as function of time, in general
$u_F(t)$	focus controller output voltage, as function of time, in V
$u_R(t)$	radial controller output voltage, as function of time, in V
$V_A(t)$	output voltage from the photodiode A, as function of time, in V
$V_B(t)$	output voltage from the photodiode B, as function of time, in V
$V_C(t)$	output voltage from the photodiode C, as function of time, in V
$V_D(t)$	output voltage from the photodiode D, as function of time, in V
$v(k)$	process noise
$x(t), x(k)$	state-vector, in general
$w(t)$	external excitation signal, as function of time, in general
$w_F(t)$	external excitation signal for the focus closed-loop, as function of time, in V
$w_R(t)$	external excitation signal for the radial closed-loop, as function of time, in V
$x_d(x_{d_max}, t)$	radial disk displacement on the spindle, as function of time, in m
$x_{obj}(t)$	objective lens position in radial direction x , as function of time, in m
$x_{ref}(t)$	radial loop reference signal, as function of time, in m
$y(t)$	plant output signal, as function of time, in general
$\hat{y}^\circ(k)$	<i>a priori</i> predicted output
$\hat{y}(k)$	<i>a posteriori</i> predicted output
$y_F(t)$	focus loop plant output signal, as function of time, in V
$y_R(t)$	radial loop plant output signal, as function of time, in V
$z_{obj}(t)$	objective lens position in focus direction z , as function of time, in m
$z_{ref}(t)$	focus loop reference signal, as function of time, in m
$z_d(x_{ref}, x_d, t)$	disk vertical deviation, as function of time, in m
$z_{dA}(x_{ref}, x_d, t)$	disk vertical deviation amplitude, as function of time, in m
$\Gamma(\theta)$	transformational function
$\Delta x(t)$	radial (tracking) spot position error, as function of time, in m
$\Delta z(t)$	vertical (focus) spot position error, as function of time, in m
$\varepsilon_{CL}^\circ(k)$	<i>a priori</i> prediction error of the closed-loop
$\varepsilon_{CL}(k)$	<i>a posteriori</i> prediction error of the closed-loop
θ	parameter vector of estimated model
θ_c	parameter vector of continuous-time model
$\hat{\theta}(k)$	estimated parameter vector
$\tilde{\theta}(k)$	parameter error
$\bar{\theta}$	mean value
$\lambda_1(k)$	forgetting factor

$\lambda_2(k)$	weighting factor
σ	standard deviation
$\phi(k)$	predictor regressor vector
$\phi_e(\omega)$	spectral density of the noise
$\phi_u(\omega)$	spectral density of the input
$\phi_w(\omega)$	spectral density of the external excitation signal
$\phi_{\varepsilon_{\text{CL}}}(e^{j\omega}, \hat{\theta})$	spectral density of the adaptation error
$\varphi(k)$	plant regressor vector

Bibliography

- [Abr01] D. Abramovitch. Magnetic and optical disk control: Parallels and contrast. In *Proceedings of the 2001 American Control Conference (ACC 2001)*, pages 421–428, Arlington, Virginia, USA, 2001.
- [BÅ70] R. Bellman and K. J. Åström. On structural identifiability. *Mathematical Biosciences*, 7:329–339, 1970.
- [BBH⁺85] G. Bouwhuis, J. Braat, A. Huijser, J. Pasman, G. van Rosmalen, and K. S. Immink. *Principles of Optical Disc Systems*. Adam Hilger Ltd, Bristol and Boston, 1985.
- [BBVH03] M. Bezděk, A. Besançon-Voda, and B. Hnilička. A novel algorithm for output-error identification of MIMO state-space models. In *Proceedings of the 3rd Computational Engineering in Systems Applications multiconference (CESA 2003)*, pages 1–8, Lille, France, 2003.
- [BDCS01] S. Bittanti, F. Dell’Orto, A. Di Carlo, and S. M. Savaresi. Radial tracking in high-speed DVD players. In *Proceedings of the 40th IEEE Conference on Decision and Control (CDC 2001)*, pages 4705–4710, Orlando, Florida, USA, 2001.
- [Bez01] M. Bezděk. Closed-loop identification of state-space representations. Master’s thesis, Brno University of Technology, 2001.
- [Bez04] M. Bezděk. Closed-loop identification of state-space representations. *Internal report of LAG, ENSIEG, INPG*, (AP04–49), 2004.
- [BLPIC76] C. Bricot, J. C. Lehureau, C. Puech, and F. le Carvenec. Optical readout of video disc. *IEEE Transactions on Consumer Electronics*, CE-22:304, 1976.
- [Bra65] R. Bracewell. *The Fourier Transform and Its Applications*. McGraw-Hill, Inc., New York, 1965.
- [Bra98] J. Braat. Differential time detection for radial tracking of optical disks. *Applied Optics*, Jaargang 37(29):6973–6982, 1998.
- [BS98] S. Bittanti and S. M. Savaresi. Safe estimate of sinusoidal signals for control applications. In *Proceedings of the 38th IEEE Conference on Decision and Control (CDC 1998)*, pages 2827–2832, Phoenix, Arizona, USA, 1998.
- [Buc] G. Buche. “ROBEA Home Page”. <http://www-lag.ensieg.inpg.fr/PRC-Bipedes/English/index.php>.

- [BW87] M. Born and E. Wolf. *Principles of Optics*. Pergamon, New York, 6th edition, 1987.
- [CAA⁺03] C. Chevallereau, G. Abba, Y. Aoustin, F. Plestan, E. R. Westervelt, C. Canudas de Wit, and J. W. Grizzle. RABBIT: A testbed for advanced control theory. *IEEE Control Systems Magazine*, 23(5):57–79, 2003.
- [CGLL84] D. K. Cohen, W. H. Gee, M. Ludeke, and J. Lewkowicz. Automatic focus control: the astigmatic lens approach. *Applied Optics*, 23(4):565–570, 1984.
- [CP01] B. Chen and A. Petropulu. Frequency domain blind Multiple-Input Multiple-Output system identification based on second and higher order statistics. *IEEE Transactions on Signal Processing*, 49(8):1677–1688, 2001.
- [dC98] R. de Callafon. *Feedback oriented identification for enhanced and robust control: A fractional approach to a wafer stage*. PhD thesis, Technical University of Delft, 1998.
- [dCdH96] R. de Callafon and P. M. J. Van den Hof. FREQID-frequency domain identification toolbox for use with MATLAB. *Selected Topics in Identification Modelling and Control*, 9:129–134, 1996.
- [Det01] M. Dettori. *LMI techniques for control with application to a Compact Disc player mechanism*. PhD thesis, Technical University of Delft, 2001.
- [dH98] P. M. J. Van den Hof. Closed-loop issues in system identification. *Annual Reviews in Control*, 22:173–186, 1998.
- [dHdC96] P. M. J. Van den Hof and R. de Callafon. Multivariable closed-loop identification: from indirect identification to dual-Youla parametrization. In *Proceedings of the 35th Conference on Decision and Control (CDC 1995)*, pages 1397–1402, Kobe, Japan, 1996.
- [dHS95] P. M. J. Van den Hof and R. P. J. Schrama. Identification and control closed-loop issues. *Automatica*, 31(12):1751–1770, 1995.
- [DL80] L. Dugard and I. D. Landau. Recursive output error identification algorithms. *Automatica*, 16:443–462, 1980.
- [DPS98] M. Dettori, V. Prodanovic, and C. W. Scherer. Mixed objectives MIMO control design for a Compact Disc player. In *Proceedings of the 1998 American Control Conference (ACC 1998)*, pages 1284–1288, Philadelphia, Pennsylvania, USA, 1998.
- [DS99] M. Dettori and C. W. Scherer. Digital implementation of a mixed objectives MIMO controller for a Compact Disc player using multiprocessor system. In *Proceedings of the 1999 American Control Conference (ACC 1999)*, pages 3630–3634, San Diego, California, USA, 1999.

- [DS00] M. Dettori and T. R. Stribos. Performance-robustness trade-off in the control of a CD player using mixed objectives design. In *Proceedings of the 39th IEEE Conference on Decision and Control (CDC 2000)*, pages 3112–3117, Sydney, Australia, 2000.
- [DS02] M. Dettori and C. W. Scherer. MIMO control design for a Compact Disc player with multiple norm specifications. *IEEE Transactions on Control Systems Technology*, 10(5):635–545, 2002.
- [DSdH95] H. Dotsch, H. T. Smakman, and P. M. J. Van den Hof. Adaptive repetitive control of a Compact Disc mechanism. In *Proceedings of the 34th Conference on Decision and Control (CDC 1995)*, pages 1720–1725, New Orleans, Louisiana, 1995.
- [DSS03] X. Deng, L. Schenato, and S. S. Sastry. Model identification and attitude control for a micromechanical flying insect including thorax and sensor models. In *Proceedings of the 2003 IEEE International Conference on Robotics and Automation (ICRA 2003)*, pages 1152–1157, Taipei, Taiwan, 2003.
- [ECM96] ECMA. Standard ECMA-130: Data interchange on read-only 120 mm optical data disks (CD-ROM). Technical report, ECMA, 1996.
- [ECM01] ECMA. Standard ECMA-267: 120 mm DVD-read-only disk. Technical report, ECMA, 2001.
- [Fil03] G. Filardi. *Robust control design strategies applied to a DVD-video player*. PhD thesis, Université Joseph Fourier-Grenoble, 2003.
- [FL99] U. Forssell and L. Ljung. Closed-loop identification revisited. *Automatica*, 35(7):1215–1241, 1999.
- [FSBVSd03] G. Filardi, O. Sename, A. Besançon-Voda, and H.-J. Schröder. Robust H_∞ control of a DVD drive under parametric uncertainties. In *Proceedings of the 17th European Control Conference (ECC 2003)*, pages 1–7, Cambridge, U.K., 2003.
- [GdH01] M. Gilson and P. M. J. Van den Hof. On the relation between a bias-eliminated least-squares BELS and an IV estimator in closed-loop identification. *Automatica*, 37(10):1593–1600, 2001.
- [Goo68] J. W. Goodman. *Introduction to Fourier Optics*. McGraw-Hill, Inc., New York, 1968.
- [HBVBP04] B. Hnilička, A. Besançon-Voda, M. Bezděk, and P. Pivoňka. Identification en boucle fermée des modèles représentation d'état physique. In *Conférence Internationale Francophone d'Automatique (CIFA 2002)*, page Accepted, Douz, Tunis, 2004. (in English).
- [HBVF04] B. Hnilička, A. Besançon-Voda, and G. Filardi. *Advances in Automatic Control*, chapter Control of DVD players; focus and tracking control loop, pages 101–128. Kluwer Academic Publishers, Boston, 2004.

- [HBVFSd03a] B. Hnilička, A. Besançon-Voda, G. Filardi, and H.-J. Schröder. Pole placement/sensitivity function shaping and controller order reduction in DVD players (Focus control loop). In *Proceedings of the 17th European Control Conference (ECC 2003)*, pages 1–6, Cambridge, U.K., 2003.
- [HBVFSd03b] B. Hnilička, A. Besançon-Voda, G. Filardi, and H.-J. Schröder. Pole placement/sensitivity function shaping and controller order reduction in DVD players (Tracking control loop). In *Proceedings of the 3rd Computational Engineering in Systems Applications multiconference (CESA 2003)*, pages 1–8, Lille, France, 2003.
- [HBVSd03] B. Hnilička, A. Besançon-Voda, and H.-J. Schröder. Modelling the focus error signal generation in a DVD player. In *Proceedings of the 3rd Computational Engineering in Systems Applications multiconference (CESA 2003)*, pages 1–8, Lille, France, 2003.
- [HBVSdF02] B. Hnilička, A. Besançon-Voda, H.-J. Schröder, and G. Filardi. Modelling the focus error characteristic of a DVD player. In *Proceedings of the 2002 IEEE International Conference on Control Applications (CCA 2002)*, pages 629–630, Glasgow, Scotland, U.K., 2002.
- [HGdB96] H. Hjalmarsson, M. Gevers, and F. de Bruyne. For model-based control design, closed-loop identification gives better performance. *Automatica*, 12(7):1659–1673, 1996.
- [Hop79] H. H. Hopkins. Diffraction theory of laser read-out systems for optical video discs. *Journal of the Optical Society of America*, 69(1):4–24, 1979.
- [HVFS04] B. Hnilička, A. Voda, G. Filardi, and O. Sename. Low order robust control of a DVD player. *IEEE Transactions on Control Systems Technology*, page Submitted, 2004.
- [HVP03] B. Hnilička, A. Voda, and P. Pivoňka. Further results on the DVD player robust control (tracking control loop). In *Proceedings of 9th Conference and Competition Student Electrical Engineering, Information and Communication Technologies (STUDENT EEICT 2003)*, pages 244–248, Brno, Czech Republic, 2003.
- [HVSd04] B. Hnilička, A. Voda, and H.-J. Schröder. Modelling the characteristics of a photodetector in a DVD player. *Sensors and Actuators A*, page Submitted, 2004.
- [Imm91] K. A. Schouhamer Immink. *Coding Techniques for Digital Recorders*. Prentice-Hall, Inc., New York, 1991.
- [Isa85] J. Isaloilović. *Videodisc and Optical Memory Systems*. Prentice-Hall, Inc., New Jersey, 1985.
- [Kai80] T. Kailath. *Linear Systems*. Prentice-Hall, Inc., New Jersey, 1980.

- [KON95] T. Katayama, M. Ogawa, and M. Nagasawa. High-precision tracking control system for digital video disk players. *IEEE Transactions on Consumer Electronics*, 41(2):313–321, 1995.
- [KYT⁺03] D. Koide, H. Yanagisava, H. Tokumaru, H. Okuda, K. Ohishi, and Y. Hayakawa. Feed-forward tracking servo system for high-data-rate optical recording. *Japanese Journal of Applied Physics*, 42(2B):939–945, 2003.
- [Lan80] I. D. Landau. An extension of a stability theorem applicable to adaptive control. *IEEE Transactions on Automatic Control*, AC-25(4):814–817, 1980.
- [Lee98] C. H. Lee. *Robust repetitive control and application to a CD player*. PhD thesis, Cambridge University Engineering, 1998.
- [Lju77] L. Ljung. On positive real transfer function and the convergence of some recursive schemes. *IEEE Transactions on Automatic Control*, AC-22:539–551, 1977.
- [Lju79] L. Ljung. Asymptotic behavior of the extended kalman filter as a parameter estimator for linear systems. *IEEE Transactions on Automatic Control*, AC-24(1):36–50, 1979.
- [Lju93] L. Ljung. Information contents in identification data from closed-loop operation. In *Proceedings of the 32nd Conference on Decision and Control (CDC 1993)*, pages 2248–2252, San Antonio, California, USA, 1993.
- [Lju99] L. Ljung. *System Identification: Theory for the User*. Prentice-Hall, Inc., New Jersey, 2nd edition, 1999.
- [LK97a] I. D. Landau and A. Karimi. An output error recursive algorithm for unbiased identification in closed loop. *Automatica*, 33(5):933–938, 1997.
- [LK97b] I. D. Landau and A. Karimi. Recursive algorithms for identification in closed loop: a unified approach and evaluation. *Automatica*, 33(8):1499–1523, 1997.
- [LKK02] S.-Y. Lee, J.-D. Kim, and S. Kim. Critical and flutter speeds of optical disks. *Microsystem Technologies*, 8:206–211, 2002.
- [LL99] J. Langer and I. D. Landau. Combined pole placement/sensitivity function shaping method using convex optimization criteria. *Automatica*, 35(6):1111–1120, 1999.
- [LLM97] I. D. Landau, R. Lozano, and M. M’Saad. *Adaptive Control*. Springer, London, 1997.
- [LS79] I. D. Landau and H. M. Silvera. A stability theorem with applications to adaptive control. *IEEE Transactions on Automatic Control*, AC-24(2):305–311, 1979.

- [LS83] L. Ljung and T. Söderström. *Theory and Practice of Recursive Identification*. MIT Press, Cambridge, 1983.
- [Mah98] V. N. Mahajan. *Optical Imaging and Aberrations: Part I, Ray geometrical optics*. SPIE, Washington, 1998.
- [Man87] M. Mansuripur. Analysis of astigmatic focusing and push-pull tracking error signals in magneto-optical disk systems. *Applied Optics*, 26(18):3981–3986, 1987.
- [McK95] T. McKelvey. *Identification of state-space models from time and frequency data*. PhD thesis, Linköping University, 1995.
- [MCW⁺96] T. D. Milster, Z. Chen, E. P. Walker, M. T. Tuell, and E. C. Gage. Optical data storage readout with quadrant pupil detection. *Applied Optics*, 35(14):2471–2476, 1996.
- [Mel94] H. Melgaard. *Identification of physical models*. PhD thesis, Technical University of Denmark, 1994.
- [Mil98] T. D. Milster. New way to describe diffraction from optical disks. *Applied Optics*, 37(29):6878–6883, 1998.
- [MLC98] J.-H. Moon, M.-N. Lee, and M. J. Chung. Repetitive control for the track-following servo system of an optical disk drive. *IEEE Transactions on Control Systems Technology*, 6(5):663–670, 1998.
- [MU99] T. D. Milster and R. S. Upton. Fundamental principles of crosstalk pattern in optical data storage. *Japanese Journal of Applied Physics, Part 1*, 38(3B):1608–1613, 1999.
- [NFP00] S. Nakamura, G. Fasol, and S. Pearton. *The Blue Laser Diode*. Springer, New York, 2nd edition, 2000.
- [NO92] H. Nakajima and H. Ogawa. *Compact Disc Technology*. Ohmsha, Tokyo, 1992.
- [Nut97] P. W. Nutter. *Image formation in the scanning optical microscope*. PhD thesis, The University of Manchester, 1997.
- [OM96] P. Van Overschee and B. De Moor. *Subspace Identification for Linear Systems. Theory, Implementation, Applications*. Kluwer Academic Publishers, Dordrecht, 1996.
- [OM97] P. Van Overschee and B. De Moor. Closed-loop subspace identification. In *Proceedings of the 36th IEEE Conference on Decision and Control (CDC 1997)*, pages 1848–1853, San Diego, California, USA, 1997.
- [OS97] A. Oppenheim and R. Schaffer. *Discrete-Time signal Processing*. Prentice-Hall, Inc., New Jersey, 1997.

- [Pap68] A. Papoulis. *Systems and Transforms with Applications in Optics*. McGraw-Hill, Inc., New York, 1968.
- [PEC00] Ltd Pioneer Electric Co. *Optical DVD Pick-up Technical Specifications*. Pioneer Electric Co., Optical Disc Division, 2000.
- [Piv02] P. Pivoňka. *Digital Control (Číslicová řídicí technika)*. VUT, Brno, 2002. (in Czech).
- [Poh95] K. C. Pohlmann. *Principles of Digital Audio*. McGraw-Hill, Inc., 3rd edition, 1995.
- [Pop72] V. M. Popov. Invariant description of linear, time-invariant controllable systems. *SIAM Journal on Control*, 15(2):252–264, 1972.
- [PvKH02] P. Pivoňka, K. Švancara, P. Krupanský, and B. Hnilička. Přímá implementace výpočetně náročných algoritmů. *Teorie a praxe automatizace*, pages 33–27, 2002. (in Czech).
- [Roc82] R. T. Rockafellar. *Convex Analysis*. Princeton University Press, New Jersey, 1982.
- [SEC01] Ltd Sanyo Electric Co. *Optical DVD Pick-up Technical Specifications*. Sanyo Electric Co., Ltd Multimedia company Optical Device Division, 2001.
- [SGSB94] M. Steinbuch, P. Van Groos, G. Schootstra, and O. H. Bosgra. Multi-variable control of a Compact Disc player using DSPs. In *Proceedings of the 1994 American Control Conference (ACC 1994)*, pages 2434–2438, Baltimore, Maryland, USA, 1994.
- [Sin88] G. T. Sincerbox. Miniature optics for optical recording. *Critical Reviews of Optical Science and Technology, SPIE Proceedings, Gradient-Index Optics and Miniature Optics*, 935:63–76, 1988.
- [SP96] S. Skogestad and I. Postlethwaite. *Multivariable Feedback Control: Analysis and Design*. John Wiley & Sons Ltd, Chichester, 1996.
- [SS89] T. Söderström and P. Stoica. *System Identification*. Prentice-Hall International Ltd, London, 1989.
- [SSB96] M. Steinbuch, G. Schoostr, and O. H. Bosgra. *Robust Control of a Compact Disk Mechanism*. CRC Press-IEEE Press, 1996.
- [Sta86] J. J. Stamnes. *Waves in Focal Regions: Propagation, Diffraction and Focusing of Light, Sound and Water Waves*. Adam Hilger Ltd, Bristol and Boston, 1986.
- [Sta98] S. G. Stan. *The CD-ROM Drive – A Brief System Description*. Kluwer Academic Publishers, Boston, 1998.

- [Sta99] S. G. Stan. *Optimization of the CD-ROM system towards higher data throughputs*. PhD thesis, Technische Universiteit Eindhoven, 1999.
- [Ste00] M. Steinbuch. Repetitive control for systems with uncertain period-time, with application to a compact disc drive. In *Proceedings of the 1st IFAC conference on Mechatronic Systems*, pages 409–414, Darmstadt, Germany, 2000.
- [SW70] J. Stoier and C. Witzgall. *Convexity and Optimization in Finite Dimensions*. Springer, Berlin, 1970.
- [TCPG02] J.-C. Trigeassou, F. J. Carrillo, T. Poinot, and O. Grospeaud. Convergence des algorithmes à erreur de sortie hors-ligne et récursifs. *Journal Européen des Systèmes Automatisés (JESA)*, 36(3):397–415, 2002. (in French).
- [UAM⁺00] R. S. Upton, F. Akhavan, T. D. Milster, M. Schweisguth, J. K. Ervin W. L. Bletscher, and A. M. Nichol. Electronic crosstalk cancellation with a quadrant cell detector. *Japanese Journal of Applied Physics, Part 1*, 39(2B):837–839, 2000.
- [UM01] R. S. Upton and T. D. Milster. Detector patterns from optical disks. *Optical Engineering*, 40(6):1030–1044, 2001.
- [Ver93] M. Verhaegen. Application of a subspace model identification technique to identify LTI systems operating in closed-loop. *Automatica*, 29(4):1027–1040, 1993.
- [Vin93] G. Vinnicombe. Frequency domain uncertainty and the graph topology. *IEEE Transactions on Automatic Control*, AC-38(9):1371–1383, 1993.
- [VSA⁺01] E. Vidal, J. Stoustrup, P. Andersen, T. S. Pedersen, and H. F. Mikkelsen. Open and closed loop parametric system identification in Compact Disk players. In *Proceedings of the 2001 American Control Conference (ACC 2001)*, pages 25–27, Arlington, Virginia, USA, 2001.
- [VSA⁺02] E. Vidal, J. Stoustrup, P. Andersen, T. S. Pedersen, and H. F. Mikkelsen. Deterministic method for obtaining nominal and uncertainty models of CD drivers. In *Proceedings of the 2002 IEEE International Conference on Control Applications (CCA 2002)*, pages 637–642, Glasgow, Scotland, U.K., 2002.
- [VSA⁺03] E. Vidal, J. Stoustrup, P. Andersen, T. S. Pedersen, and H. F. Mikkelsen. Parametric uncertainty with perturbations restricted to be real on 12 CD mechanisms. In *Proceedings of the 2003 American Control Conference (ACC 2003)*, pages –, Denver, Colorado, USA, 2003.
- [VV02] V. Verdult and M. Verhaegen. Subspace identification of multivariable linear parameter-varying systems. *Automatica*, 38(5):805–814, 2002.
- [Wil94] E. W. Williams. *The CD-ROM Optical Disc Recording Systems*. Oxford, New York, 1994.

- [WP97] E. Walter and L. Prozanto. *Identification of Parametric Models from Experimental Data*. Springer, London, 1997.
- [WY02] J. F. Whidborne and J. B. Yang. Multiobjective design of low complexity digital controllers. In *Proceedings of the 2002 IEEE International Conference on Computer Aided Control System Design (CACSD 2002)*, pages 27–32, Glasgow, Scotland, U.K., 2002.
- [YCC⁺02] K. Yang, Y. Choi, W. K. Chung, I. H. Suh, and S. R. Oh. Robust tracking control of optical disk drive systems using error based disturbance observer and its performance measure. In *Proceedings of the 2002 American Control Conference (ACC 2002)*, pages 1395–1400, Anchorage, Alaska, USA, 2002.
- [Zhu01] Y. C. Zhu. *Multivariable System Identification for Process Control*. Elsevier Science Ltd, Oxford, 2001.
- [ZKS02] Y. Zhou, D. Kostic, and M. Steinbuch. Estimator-based sliding mode control of an optical disc drive under shock and vibration. In *Proceedings of the 2002 IEEE International Conference on Control Applications (CCA 2002)*, pages 631–636, Glasgow, Scotland, U.K., 2002.

Curriculum Vitae

Personal:

Name: Bohumil Hnilička
Born: 11 April 1977 in Šumperk, Czech Republic
Marital status: single

Education:

1991–1995 Secondary School of Electrotechnical Engineering in Mohelnice, (Czech Republic)
1995–2000 Master degree in automation and control at Brno University of Technology, Faculty of Electrical Engineering and Computer Science, Department of Automatic Control and Instrumentation, (Czech Republic). Master Thesis: Adaptive control of the heat systems
2000–now Ph.D. student at the *Systemes Avancés de Régulation et Commande* Group, *Laboratoire d'Automatique de Grenoble-INPG/ENSIEG*, (France) and at Brno University of Technology, Faculty of Electrical Engineering and Computer Science, Department of Automatic Control and Instrumentation, (Czech Republic).

Published Papers

- [1] M. Bezděk, A. Besançon-Voda, and B. Hnilička. A novel algorithm for output-error identification of MIMO state-space models. In *Proceedings of the 3rd Computational Engineering in Systems Applications multiconference (CESA 2003)*, pages 1–8, Lille, France, 2003.
- [2] B. Hnilička, A. Besançon-Voda, M. Bezděk, and P. Pivoňka. Identification en boucle fermée des modèles représentation d'état physique. In *Conférence Internationale Francophone d'Automatique (CIFA 2002)*, page Accepted, Douz, Tunis, 2004. (in English).

- [3] B. Hnilička, A. Besançon-Voda, and G. Filardi. *Advances in Automatic Control*, chapter Control of DVD players; focus and tracking control loop, pages 101–128. Kluwer Academic Publishers, Boston, 2004.
- [4] B. Hnilička, A. Besançon-Voda, G. Filardi, and H.-J. Schröder. Pole placement/sensitivity function shaping and controller order reduction in DVD players (Focus control loop). In *Proceedings of the 17th European Control Conference (ECC 2003)*, pages 1–6, Cambridge, U.K., 2003.
- [5] B. Hnilička, A. Besançon-Voda, G. Filardi, and H.-J. Schröder. Pole placement/sensitivity function shaping and controller order reduction in DVD players (Tracking control loop). In *Proceedings of the 3rd Computational Engineering in Systems Applications multiconference (CESA 2003)*, pages 1–8, Lille, France, 2003.
- [6] B. Hnilička, A. Besançon-Voda, and P. Pivoňka. Further results on the DVD player robust control (tracking control loop). In *Proceedings of 9th Conference and Competition Student Electrical Engineering, Information and Communication Technologies (STUDENT EEICT 2003)*, pages 244–248, Brno, Czech Republic, 2003.
- [7] B. Hnilička, A. Besançon-Voda, and H.-J. Schröder. Modelling the focus error signal generation in a DVD player. In *Proceedings of the 3rd Computational Engineering in Systems Applications multiconference (CESA 2003)*, pages 1–8, Lille, France, 2003.
- [8] B. Hnilička, A. Besançon-Voda, H.-J. Schröder, and G. Filardi. Modelling the focus error characteristic of a DVD player. In *Proceedings of the 2002 IEEE International Conference on Control Applications (CCA 2002)*, pages 629–630, Glasgow, Scotland, U.K., 2002.
- [9] B. Hnilička and P. Pivoňka. Implementation of complex controller's algorithms from the program MATLAB-Simulink to program logic controllers. In *Proceedings of the Colloquium MATLAB 2000*, pages 318–321, Prague, Czech Republic, 2000. (in Czech).
- [10] B. Hnilička and P. Pivoňka. Implementation of heterogeneous algorithms from MATLAB-Simulink into PLC. In *Proceedings of the International 9th Zittau Fuzzy Colloquium*, pages 151–156, Zittau, Germany, 2001.
- [11] P. Pivoňka, K. Švancara, P. Krupanský, and B. Hnilička. Přímá implementace výpočetně náročných algoritmů. *Teorie a praxe automatizace*, pages 33–27, 2002. (in Czech).