Nonlinear dynamic systems: 
blind identification of block-oriented models, 
and instability under random inputs

Thesis submitted in the fulfillment of the requirements for the award of the degree of 
Doctor in de ingenieurswetenschappen (Doctor in Engineering) by

Laurent Vanbeylen

Private defense: April 5th, 2011 
Public defense: May 4th, 2011

Advisors: Prof. dr. ir. Rik Pintelon (Vrije Universiteit Brussel) 
Prof. dr. ir. Johan Schoukens (Vrije Universiteit Brussel)
Nonlinear dynamic systems:
blind identification of block-oriented models,
and instability under random inputs

Thesis submitted in the fulfillment of the requirements for the award of the degree of
Doctor in de ingenieurswetten (Doctor in Engineering) by
Laurent Vanbeylen

Private defense: April 5th, 2011
Public defense: May 4th, 2011

Chairman: Prof. dr. ir. Hugo Sol (Vrije Universiteit Brussel)
Vice-chairman: Prof. dr. ir. Herman Terryn (Vrije Universiteit Brussel)
Secretary: Prof. dr. ir. Gerd Vandersteen (Vrije Universiteit Brussel)

Advisors: Prof. dr. ir. Rik Pintelon (Vrije Universiteit Brussel)
Prof. dr. ir. Johan Schoukens (Vrije Universiteit Brussel)

Jury: Dr. ir. Adrian Wills (University of Newcastle, Australia)
Prof. dr. ir. Johan Suykens (Katholieke Universiteit Leuven)
Prof. dr. ir. Rodolphe Sepulchre (Université de Liège)
Prof. dr. ir. Patrick Guillaume (Vrije Universiteit Brussel)
Acknowledgements

Doing research and writing a Ph.D. is not a one-man performance. I wish to thank everyone who has contributed somehow to the accomplishment of this work during these 5.5 years.

Johan Schoukens, my initial (co)adviser, gave me the opportunity to start this Ph.D. The early phases of the research, on stability of nonlinear systems were very challenging. He provided me with inspiration and intuitive skills on the research topic, and continuously supported me. It was a pleasure to work with you. Thank you, Johan!

Rik Pintelon supported me throughout this Ph.D. I started collaborating with him, since Johan suggested a nice initial value algorithm to blindly identify a Wiener system. He initiated me with the advanced system identification tools, became my main adviser, was always willing to answer my questions, and proofread very carefully my writings. Your help was very valuable to me. Thank you, Rik!

Thanks to my both advisers, Rik and Johan, for keeping a critical attitude towards my results, for proofreading my papers and this thesis, and for giving me valuable comments.

I am very grateful to all the members of the jury, for willing to evaluate this thesis and to read it with care. Thank you!

Kurt Barbé learned me a lot with his knowledge on mathematics and statistics. It was a pleasure to have these interesting chats in our office. Your optimistic view on difficult challenges was really a good support. Thank you, Kurt!

Pieter de Groen is a professor at the mathematics department, with whom I got in touch via Kurt. He listened to my mathematical integration problem and taught me how to approximate integrals via the Laplace method. Thank you, Pieter!

With David Rijlaarsdam, staying at the department for several months, I had some interesting discussions on the stability problem. Thank you, David!

Thanks to all people of the ELEC-team (including visitors), for scientific support, and for the friendly atmosphere and nice time spent together in the koffiekot...

Last but not least, I wish to thank Anne, for her permanent love and care, our daughter Alison, which gave us already a lot of pleasure, my parents always willing to help at any time, my friendly sister and her boyfriend, my parents-in-law for taking care of Alison, and all other people of my family supporting me during my studies and Ph.D.

Laurent Vanbeylen
Foreword

How it all started: a motivation of the research topics

My research at ELEC, under the supervision of Johan Schoukens, started during my master thesis, on “reduction of unstable phenomena in acoustic systems in feedback”. That work focussed on the use of a frequency shifting element placed in between microphone and loudspeaker to reduce the Larsing effect (howling). Algorithms inspired on system identification methods were applied and the destabilizing effect of nonlinear distortions was briefly touched.

After my graduation, Johan proposed me to join the department (which I accepted, since my master thesis was a pleasant and instructive experience) and to work on unstable nonlinear systems. He had been in contact with Prof. Jan Swevers (dept. PMA, KULeuven), reporting about controllers for mechanical active suspension systems. Those systems seemed to work fine for (very) long periods of time, until very suddenly a controller failure (“blow-up”) took place. Johan started to think about the idea that the model-plant discrepancy in the nonlinear feedback loop was playing a crucial role: a controller is usually designed based on a priorly obtained plant model (working well together in theory). However, actually model errors are present, jeopardizing the stability guarantees of the control design methodology for the controlled plant. He proposed a method to analyze the stability of feedback systems in the presence of nonlinear distortions starting from a series of measurements (Schoukens et al., 2004). That was the original research problem, on which Johan proposed me to pursue research.

Two attempts to discover the relationship between the probability of blow-up and the input standard deviation, were based on the analysis of the measured output properties: power spectrum and histogram. The aim of the first attempt was to analyze the evolution of the poles of an ARMA model fitted to the output spectrum (it was expected that they would move towards the unstable region). The idea of the second attempt is that a static nonlinearity with a vertical asymptote (\( \lim_{u \to a} f(u) = \infty \), continuous for \( u < a \) and producing an unbounded output for \( u > a \)) and driven by Gaussian noise could allow one to calculate the probability of a blow-up. Although these two attempts did not solve the instability problem, their combination (both dynamic linear and static nonlinear parts) gave rise to the blind (Wiener and Hammerstein) identification setups, for which Johan suggested an initial value algorithm. I then got the support of Rik Pintelon on that topic, which I developed in parallel with the research on the “instability of nonlinear dynamic systems”, supported by Johan.
# Table of contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abbreviations</td>
<td>15</td>
</tr>
<tr>
<td>Symbols and notations</td>
<td>17</td>
</tr>
<tr>
<td>1 Introduction</td>
<td>23</td>
</tr>
<tr>
<td>1.1 Introduction: from observation to model</td>
<td>23</td>
</tr>
<tr>
<td>1.2 Part I: Blind system identification of Wiener and Hammerstein models</td>
<td>26</td>
</tr>
<tr>
<td>1.3 Part II: Analysis of instability under random inputs</td>
<td>26</td>
</tr>
<tr>
<td>1.4 Publications</td>
<td>28</td>
</tr>
<tr>
<td>I Maximum-likelihood approach to the blind identification of Wiener and Hammerstein nonlinear models</td>
<td>33</td>
</tr>
<tr>
<td>Introduction</td>
<td>35</td>
</tr>
<tr>
<td>2 Blind identification of Wiener systems</td>
<td>39</td>
</tr>
<tr>
<td>2.1 Chapter outline and contributions</td>
<td>39</td>
</tr>
<tr>
<td>2.2 Class of discrete-time Wiener systems considered</td>
<td>40</td>
</tr>
<tr>
<td>2.3 Parameterization of the Wiener system</td>
<td>41</td>
</tr>
<tr>
<td>2.4 Stochastic framework</td>
<td>41</td>
</tr>
<tr>
<td>2.5 The Gaussian maximum-likelihood estimator</td>
<td>42</td>
</tr>
<tr>
<td>2.5.1 The negative log-likelihood (NLL) function</td>
<td>42</td>
</tr>
<tr>
<td>2.5.2 Simplified cost function</td>
<td>42</td>
</tr>
<tr>
<td>2.5.3 Identifiability discussion</td>
<td>43</td>
</tr>
<tr>
<td>2.5.4 Frequency domain formulation</td>
<td>45</td>
</tr>
<tr>
<td>2.5.5 Asymptotic properties</td>
<td>46</td>
</tr>
<tr>
<td>2.5.6 Loss of consistency in the case of a non-Gaussian input</td>
<td>47</td>
</tr>
<tr>
<td>2.6 Generation of initial estimates</td>
<td>47</td>
</tr>
</tbody>
</table>
### Table of contents

2.6.1 Subproblem 1: monotonically increasing static nonlinearity driven by Gaussian noise .................................. 48  
2.6.2 Subproblem 2: LTI driven by white input noise ............ 50  
2.7 Minimization of the cost function ............................... 51  
2.7.1 Parameterization of the nonlinearity ....................... 51  
2.7.2 Gauss-Newton procedure .................................... 52  
2.7.3 Practical implementation of the constraint .................. 52  
2.8 Complex-valued case .......................................... 53  
2.9 The Cramér-Rao lower bound .................................. 53  
2.10 Simulation results ............................................. 54  
2.10.1 Setup: presentation of the example ....................... 54  
2.10.2 Graphical presentation of the results ..................... 55  
2.11 Practical considerations about real-life systems operating in continuous-time ............................................ 58  
2.12 Laboratory experiment ....................................... 59  
2.13 Conclusion .................................................... 60  
Appendices ......................................................... 63  
2.A Calculation of the log-likelihood function .................... 63  
2.B Asymptotic properties .......................................... 64  
2.C Loss of consistency for a non-Gaussian distributed input ........ 67  
2.D Calculation of the Jacobian matrix ............................ 67  
2.E Proof of Theorem 2.13 (Fisher information matrix) ........ 68  
2.F Practical calculation of the Fisher information matrix ...... 71  
2.G Law of large numbers applied to the computation of the Fisher information matrix ............................................. 72  
References .......................................................... 74  

3 **Blind identification of Hammerstein systems** ................................................................. 79  
3.1 Chapter outline and contributions .............................. 80  
3.2 Assumptions and choices ....................................... 80  
3.2.1 Class of discrete-time Hammerstein systems considered 80  
3.2.2 Parameterization ............................................ 81  
3.2.3 Stochastic framework ....................................... 82  
3.3 The Gaussian maximum-likelihood estimator ................ 82  
3.3.1 The negative log-likelihood function ...................... 82  
3.3.2 Identifiability .............................................. 83  
3.3.3 Result ...................................................... 84  
3.3.4 Asymptotic properties ..................................... 84  
3.4 Generation of initial estimates ................................. 85  
3.5 Minimization of the cost function ............................. 85  
3.5.1 Parameterization of the nonlinearity ...................... 86
Table of contents

3.5.2 Gauss-Newton procedure ........................................ 86
3.5.3 Practical implementation of the constraint .................. 86
3.6 The Cramér-Rao lower bound ...................................... 86
3.6.1 Theoretical expressions ......................................... 86
3.6.2 Practical calculation ............................................ 88
3.7 Simulation results .................................................. 88
3.7.1 Setup: presentation of the example ......................... 88
3.7.2 Results and validation .......................................... 88
3.8 Conclusions ........................................................ 90
Appendices .......................................................... 90
3.A Asymptotic properties .............................................. 91
3.B Fisher information matrix ......................................... 93
3.B.1 Case (a): $\theta_i \in \theta_L, \theta_j \in \theta_L$ ....................... 94
3.B.1.1 Preliminary calculations .................................. 94
3.B.1.2 Calculation of the summand $T_t$ of equation (3.20) .. 95
3.B.1.3 Actual calculation of the expected values and summations ............................................. 96
3.B.2 Case (b): $\theta_i \in \theta_L, \theta_j \in \theta_{NL}$ ....................... 100
3.B.2.1 Preliminary calculations of the derivatives .......... 100
3.B.2.2 Calculation of the summand $T_t$ of equation (3.20) 100
3.B.2.3 Actual calculation of the expected values and summations ............................................. 101
References .......................................................... 101

4 Impact of output measurement noise .......................... 103
4.1 Introduction ....................................................... 103
4.2 Theoretic analysis of the noise-induced bias .................. 104
4.3 Simulation results ................................................ 105
4.3.1 Output simulation error (Wiener) ......................... 105
4.3.2 Parametric bias error (Hammerstein) .................. 106
4.4 Blind maximum-likelihood problem with output measurement noise .................................................. 107
4.4.1 Very high-dimensional integrals, related problems and solutions .................................................. 107
4.4.2 Identifiability (Wiener) ..................................... 108
4.5 Conclusion ........................................................ 109
Appendices ........................................................ 109
4.A Identifiability proof .............................................. 109
4.A.1 Identifiability condition for cumulant-based estimation of the nonlinear coefficients and the noise variance .................................................. 109
4.A.2 Making sure that the polynomial $P(\theta_{NL})$ is not identically zero, via the evaluation of the determinant for a cubic function ................................................. 110
4.A.3 Proving the full-rank condition .................................................. 111
References ...................................................................................... 112

II Analysis of instability under random inputs 113

Introduction 115

5 Instability under random inputs: data-driven perspective 117
  5.1 Introduction ............................................................. 117
  5.1.1 Chapter’s contributions and outline .............................. 119
  5.2 Blow-up event: illustration and dependence on the input variance 119
  5.3 Escape times and their distribution .................................. 122
  5.4 Estimating the probability of blow-up from successive experiments 124
    5.4.1 Validation of the geometric distribution ..................... 125
    5.4.2 Two remarks .......................................................... 126
      5.4.2.1 On the use of a finite $c$-value ................................. 126
      5.4.2.2 On the “start-up transient” ..................................... 126
    5.4.3 Some more results for the example system .................... 126
    5.4.4 Additional remark on an alternative way of estimating the probability of escape ................................................. 128
  5.5 Input-output stability for stochastic processes ..................... 128
  5.6 Extreme-value based method .......................................... 129
    5.6.1 Introduction ............................................................ 129
    5.6.2 Extreme value statistics ............................................. 130
    5.6.3 Estimation of the GPD’s shape parameter $\xi$ .................. 132
    5.6.4 Impact of additive noise ............................................. 133
    5.6.5 Results of numerical experiments ............................... 134
    5.6.6 Hypothesis test ......................................................... 134
  5.7 Conclusion ................................................................. 137
Appendices .................................................................................... 137
  5.A Maximum-likelihood estimation of the probability of escape . 138
  5.B Fisher information matrix of the probability of escape ........ 139
References ...................................................................................... 141

6 Instability under random inputs: model-based perspective 143
  6.1 Introduction ............................................................. 144
    6.1.1 Motivation of this work .............................................. 144
    6.1.2 Example of locally stable, globally unstable dynamics .... 144
6.1.3 Towards a concrete problem setting ........................................ 145
6.1.4 Background information on classical notions of stability, and situation of this work ......................................................... 146
6.1.5 Randomly excited systems and related literature ...................... 149
6.1.6 Main contributions .............................................................. 150
6.1.7 Chapter’s organization .......................................................... 150
6.2 Assumptions ............................................................................ 151
6.2.1 Prior knowledge on the discrete-time state-space model 151
6.2.2 Stochastic framework ......................................................... 152
6.2.3 Possible extensions and generalizations .................................. 152
6.2.4 Consequence: bounded operation over a time interval of length $\tau$. ................................................................. 153
6.3 Estimation of the state’s probability distribution ......................... 154
6.3.1 Detailed specification of the integrand: both factors in (6.8) ....... 154
6.3.2 Performing the integration .................................................. 155
6.3.3 Interpretation ................................................................. 156
6.3.4 Putting theory into practice .................................................. 157
6.3.5 Additional remarks ............................................................ 157
6.4 Generation of initial estimates ................................................... 158
6.4.1 Usage of a neighbouring solution ...................................... 158
6.4.2 Extrapolation on the state-space grid .................................. 159
6.4.3 Solution in the linear case .................................................. 159
6.5 Accuracy of the pdf estimation .................................................. 159
6.6 Simulation results ..................................................................... 159
6.6.1 Nonlinear model equations .................................................. 163
6.6.2 Underlying linear model ..................................................... 167
6.6.3 Gathering information at other input standard deviation values ................................................................. 167
6.6.4 Nonlinear model with colored input power spectrum .............. 170
6.6.5 Nonlinear model with an initial state different from the origin, and a deterministic input component ....................... 171
6.6.6 Nonlinear model appended with nonlinear terms in the input ................................................................. 173
6.6.7 Comparison with the Fokker-Plank approach ......................... 175
6.7 Conclusions ............................................................................ 178
Appendices ..................................................................................... 180
6.A Integration approach ............................................................... 180
6.A.1 Goal and workflow .............................................................. 180
6.A.2 Elimination of the constraints .............................................. 181
6.A.3 Application of the Laplace integration .................................. 182
Table of contents

6.B Analytic solution in the linear white noise situation .......................... 183
   6.B.1 Statement of the theoretical result ........................................... 183
   6.B.2 Interpretation of the result .................................................... 184
   6.B.3 Proof ......................................................................... 185
6.C Approximation of the integral approximation error ....................... 186
References ....................................................................... 189

7 Conclusions .................................................................. 193
   7.1 General discussion ................................................................. 193
   7.2 List of essential contributions .................................................... 195
   7.3 Limitations of the proposed approaches and suggestions for further research .................................................... 196
      7.3.1 Blind identification of block-oriented nonlinear models . 196
      7.3.2 Instability under random inputs .............................................. 197

Global reference list ................................................. 201
## Abbreviations

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARMA</td>
<td>Autoregressive moving average</td>
</tr>
<tr>
<td>CLT</td>
<td>Central limit theorem</td>
</tr>
<tr>
<td>CRB</td>
<td>Cramér-Rao lower bound</td>
</tr>
<tr>
<td>DFT</td>
<td>Discrete Fourier Transform (throughout this thesis we use the unitary definition, viz. scaled by $\sqrt{N}$)</td>
</tr>
<tr>
<td>FFT</td>
<td>Fast Fourier Transform</td>
</tr>
<tr>
<td>FIM</td>
<td>Fisher information matrix</td>
</tr>
<tr>
<td>GPD</td>
<td>Generalized Pareto distribution</td>
</tr>
<tr>
<td>i.i.d.</td>
<td>Independent and identically distributed</td>
</tr>
<tr>
<td>IDFT</td>
<td>Inverse Discrete Fourier Transform (throughout this thesis we use the unitary definition, viz. scaled by $\sqrt{N}$)</td>
</tr>
<tr>
<td>ISS</td>
<td>Input to state stability</td>
</tr>
<tr>
<td>LIM</td>
<td>Laplace integration method</td>
</tr>
<tr>
<td>LLN</td>
<td>Law of large numbers</td>
</tr>
<tr>
<td>LTI</td>
<td>Linear time-invariant dynamic system</td>
</tr>
<tr>
<td>MLE</td>
<td>Maximum-likelihood estimator</td>
</tr>
<tr>
<td>MSE</td>
<td>Mean squared error</td>
</tr>
<tr>
<td>NLL</td>
<td>Negative log-likelihood function</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>---------------------------------</td>
</tr>
<tr>
<td>pdf</td>
<td>probability density function</td>
</tr>
<tr>
<td>rms</td>
<td>Root mean square value</td>
</tr>
<tr>
<td>ROA</td>
<td>Region of attraction</td>
</tr>
<tr>
<td>SNR</td>
<td>Signal to noise ratio</td>
</tr>
<tr>
<td>w.r.t.</td>
<td>with respect to</td>
</tr>
</tbody>
</table>
Symbols and notations

$(\bullet)'$ Except when explicitely mentioned, partial derivative w.r.t. the first function argument

$0^{n \times m}$ Zero matrix with $n$ rows and $m$ columns

$[a(k)]_{k=0...K-1}$ $K$-dimensional vector consisting of the respective entries $a(k)$ by letting the row index $k$ vary in the indicated range

$[a(k,l)]_{k,l=0...K-1}$ $K \times K$ matrix consisting of the respective entries $a(k,l)$ by letting the row and column indices $(k$ and $l)$ vary in the indicated range

$\alpha$ Significance level of a statistical test

$\approx$ Approximately equal to

$\bullet^H$ Hermitian matrix transposition (i.e. transpose and complex conjugate)

$\bullet^T$ Matrix transposition

$\bullet_0$ Subscript 0 often representing the (usually unknown) exact model or parameters.

$\bullet_{re}$ Stacks the real and imaginary parts of vector $\bullet$ onto each other

$\delta(\bullet)$ Multivariate Dirac delta distribution

$\delta_r$ Kronecker delta, taking the value 1 if $r = 0$, and the value 0 otherwise.

$\delta_\Sigma(x)$ Multivariate normal pdf with zero mean and covariance matrix $\Sigma$
\begin{itemize}
\item Symbol denoting an estimator
\item $\lambda$ Variance of the unmeasured signal $e(t)$
\item $\lambda_t$ Lagrange multipliers
\item $\mathbb{C}^{m \times n}$ Set of all $m \times n$ matrices with complex entries
\item $\mathbb{E}$ Mathematical expectation
\item $\mathbb{N}$ Set of natural numbers
\item $\mathbb{N}_0$ Set of natural numbers, zero element left out
\item $\mathbb{P}(A)$ Probability of event $A$
\item $\mathbb{R}$ Set of real numbers
\item $\mathbb{R}^{m \times n}$ Set of all $m \times n$ matrices with real entries
\item $\mathbb{R}_0$ Set of real numbers, zero element left out
\item $y_N = y$ Vector of (output) observations
\item $\mathcal{O}$ Big O notation: $f(x) = \mathcal{O}(f_1(x))$ as $x \to a \iff \lim_{x \to a} \frac{f(x)}{f_1(x)} \in \mathbb{R}$
\item $\mathcal{O}_p(\cdot)$ Big O in probability:
\[
\mathcal{O}_p(X_N) = a_N \text{ as } N \to \infty \iff \lim_{N \to \infty} \mathbb{P}\left(\left| \frac{X_N}{a_N} \right| > M \right) = 0
\]
\item $\text{cov}(\cdot)$ Covariance matrix associated with vector $\cdot$
\item $\text{cov}(X,Y)$ Covariance between the random variables $X$ and $Y$
\item $\text{cum}(X_1, X_2, \ldots, X_r)$ Joint cumulant of the random variables $X_1 \ldots X_r$ (see Brillinger, 1981, or Pintelon and Schoukens, 2001)
\item $\exp(\cdot)$ Naperian exponentiation of $\cdot$
\item $\mathbf{F}_i$ Fisher information matrix
\item $\log(\cdot)$ Naperian logarithm of $\cdot$
\item $\text{std}(\cdot)$ Standard deviation of the random variable $\cdot$
\item $\text{var}(\cdot)$ Variance of $\cdot$
\end{itemize}
Symbols and notations

\( \mu \) \hspace{1cm} \text{Variance of the additive output measurement noise}

\( \Phi^{-1} \) \hspace{1cm} \text{Standard normal quantile function}

\( \sigma \) \hspace{1cm} \text{Scale parameter of the GPD}

\( \sigma_X \) \hspace{1cm} \text{Time-domain standard deviation of the random variable } X, \text{ corresponding to the stationary signal } x(t)

\( \tau_i \) \hspace{1cm} \text{Escape time of the } i^{th} \text{ experiment}

\( \theta \) \hspace{1cm} \text{Parameter vector}

\( \theta_L \) \hspace{1cm} \text{Parameters associated with the linear time-invariant part of the model}

\( \theta_{NL} \) \hspace{1cm} \text{Parameters associated with the nonlinear part of the model}

\( \tilde{\tau}_i \) \hspace{1cm} \text{Escape time of the } i^{th} \text{ experiment, truncated to a maximal value } N + 1

\( \varepsilon \) \hspace{1cm} \text{Residual vector such that the cost function } V = \varepsilon^H \varepsilon

\( \xi \) \hspace{1cm} \text{Shape parameter of the GPD}

\( c, d \ (C, D) \) \hspace{1cm} \text{Numerator, denominator coefficients of the LTI noise model (in the } z\text{-domain)}

\( e(t) \) \hspace{1cm} \text{Model’s input signal}

\( e(t) \) \hspace{1cm} \text{Unobserved stochastic input to the model}

\( f(\bullet) \) \hspace{1cm} \text{(Static) nonlinear function}

\( F_X(x) \) \hspace{1cm} \text{Cumulative distribution function of the random variable } X

\( F_{\tau} \) \hspace{1cm} \text{DFT matrix converting a time domain vector into the frequency domain vector (consisting of the DFT coefficients of the original vector)}

\( g(\bullet) \) \hspace{1cm} \text{Static nonlinear function written from output to input}

\( G_{\xi, \sigma} \) \hspace{1cm} \text{The GPD’s cumulative distribution function}

\( g_N(y, \theta) \) \hspace{1cm} \text{Correction factor in the cost function definition depending on the derivatives of the static nonlinear function}
Symbols and notations

\( H(z) \) Discrete time noise filter

\( I_{n \times n} \) Identity matrix of size \( n \times n \)

\( I_A \) Indicator function associated with statement \( A \), viz. it takes the value 1 if \( A \) is true, and 0 otherwise

\( J \) Jacobian matrix

\( j \) Imaginary unit \( j = \sqrt{-1} \)

\( L(y, \ldots) \) Negative log-likelihood function

\( N \) Number of data samples

\( n_c, n_d \) Numerator and denominator orders of an LTI transfer function

\( p_e \) Probability of escape, probability of blow-up

\( p_X(x) \) Probability density function of the random variable \( X \) evaluated at its argument \( x \)

\( q \) Forward shift operator \( q x(t) = x(t + 1) \)

\( r_{uu}(t) \) Autocovariance function of \( u \)

\( t \) Discrete time symbol

\( T(z) \) Transient term in the input-output DFT relationship

\( u(t) \) Model’s intermediate signal

\( V(\bullet) \) Cost function

\( y(t) \) Model’s output signal

\( z \) The \( z \) variable (Z-transform)

\( z_k \) \( k \)th root of the number -1, viz, \( \exp(j2\pi k/N) \)

\( \det(\bullet) \) Determinant of matrix \( \bullet \)

\( \text{diag}(\bullet) \) If \( \bullet \) is a vector: diagonal matrix with respective diagonal entries taken form \( \bullet \)’s elements.

If \( \bullet \) is a square matrix: vector consisting of the respective diagonal entries of \( \bullet \).
<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Im</td>
<td>Imaginary part</td>
</tr>
<tr>
<td>Re</td>
<td>Real part</td>
</tr>
<tr>
<td>tr(\bullet)</td>
<td>Trace of the matrix \bullet</td>
</tr>
</tbody>
</table>
Chapter 1
Introduction

This introductory chapter starts from the concepts of nonlinear dynamic systems, modelling and system identification, and proceeds with the preliminary problem descriptions of blind identification of block-oriented models and instability under random inputs, two topics constituting the two parts of this thesis. This chapter also contains a publication list of all the publications accomplished during the Ph.D.

1.1 Introduction: from observation to model

Since humans are living in their universe, from the first day of their life, they have always been eager to learn from looking, listening and experimenting. The observation of physical phenomena, such as tree branches moving in the wind, the sun and shadow, a butterfly’s ability to control its position close to a flower, the water flowing in a river, the motion of the moon, ... has always been intriguing them. These phenomena have created a certain perception of how all this works. This led to the development of the famous physical laws of gravitation and classical mechanics (Isaac Newton), electromagnetism (James Maxwell), optics, ... All these branches of modern physics are based on mathematical representations (which have very often their roots in human perception of nature), called models \(^1\) and intend to represent the phenomena to a certain

\(^1\)A simple example of model for a resistor is e.g. given by the proportional law \(V = R.I\) between current and voltage. (In engineering applications, differential equations are more commonly used as models for dynamic systems.)
Chapter 1. Introduction

Figure 1.1: Dynamic systems: block structure representation and terminology

degree of accuracy. Nowadays, lots of modern appliances and industrial applications (e.g. a car, a chemical process, a bridge, an electronic filter, a robot, etc.) are being designed based on models, which allow predicting the devices’ response by mathematical calculations or computer computations, without the need to build them. Although scientists developed laws and models (ranging from simplistic to very intricate) for many physical phenomena, in practice, real-life applications (also called dynamic systems \(^2\) from now on) often tend to be much more involved and complex than expected. Up to now, many dynamic systems have been and are being modelled as linear time-invariant systems (whose mathematical framework is well-known and well-developed). This mathematical simplification allows one to interpret and understand the system’s response better and easier: the linear system’s response to a sum of excitation signals can be decomposed into terms corresponding to each individual excitation signal (this is known as the superposition principle for linear systems), and a doubling of the input excitation amplitude results in a doubling of the response (also called output) amplitude. Mathematically, just a single function, known as the impulse response, is enough to characterize a linear system completely.

Although linear models can be used in many application areas, in other situations, the system’s behaviour may not be approximated well enough by a linear model. To describe those systems in a satisfactory way, it is necessary to construct nonlinear models.

To understand linear and nonlinear behaviour, an old LP-player/audio

\(^2\)In Figure 1.1, a block schematic representation is shown and the usual terminology related to dynamic systems is introduced. A dynamic system is an abstract representation of a physical device which gives a response (also called output signal) to an excitation (also called input). E.g. a mechanical structure, such as an airplane wing, can have typically a force as input acting upon the system, and a displacement as output response. In general, real-life systems are said to be “dynamic”, which means that the response depends not only on the present input excitation, but also on its past. The system somehow “remembers” for some time what happened to it; in other words, it has a “memory”. Moreover, dynamics cause a system’s behaviour to be frequency-dependent: a sine wave (single frequency signal) input results in a frequency-dependent response. This may result in so-called “natural frequencies” or “resonances”, to which the system can be very sensitive. For a bridge or a swing, (very) large displacement amplitudes can be the result of a well-tuned excitation frequency with given amplitude (e.g. the legs of the child moved back and forth).
an amplifier/loudspeaker system can be used as an example. Its input is the signal written on the LP. The output is the sound that comes out of the device. When the volume control is set on a low or normal level, the music sounds “nicely”. The music sounds simply amplified when the knob is turned up in this range: this is linear behaviour. Nevertheless, the old device may alter the sound somehow; some frequencies (more precisely frequency bands) may be somewhat attenuated, the music may lack sharpness and sounds dull: that frequency-dependent behaviour is called dynamics. This is similar to what a digital equalizer does; it can change the frequency contents of the signal. When the volume control is set on a (very) high level, the music tends to sound very ugly: it underwent a nonlinear distortion. The music sounds differently than just a scaled version of the first situation with the volume knob in the low or normal range.

System identification is the scientific discipline that develops tools to construct models from (possibly noisy) measurement data, in a statistically sound manner. Besides their usage for simulation purposes (i.e. computation of the models’ response to new excitation signals), the resulting models are needed for control purposes too: i.e. to design control algorithms adapting the input excitation in such a way that the response follows a user-supplied signal automatically.

Compared to linear systems, nonlinear systems can have an extreme variety and complexity. The quote of the mathematician Stan Ulam “Using a term like nonlinear science is like referring to the bulk of zoology as the study of non-elephant animals.” illustrates this fact. Nowadays, excellent generic models and system identification methods are available for the well-understood linear systems. On the other hand, for nonlinear systems it is extremely hard to introduce a generic model structure valid for any nonlinear system. Since this thesis is centered on nonlinear dynamic systems, restrictions will be imposed on the kind of nonlinear behaviour considered (e.g. block-oriented nonlinear behaviour or nonlinear state-space behaviour). It concentrates on two topics, constituting its two main parts:

• the blind system identification of nonlinear (Wiener and Hammerstein) block-oriented models,

• the analysis of the instability of dynamic nonlinear systems under random inputs.

Both will be introduced in the upcoming sections.
1.2 Part I: Blind system identification of Wiener and Hammerstein models

Besides other nonlinear model structures, block-oriented models have shown to be useful in a number of application domains, such as physiology, biology, chemistry and radio-frequency applications. Block-oriented models are quite popular, since they mimic the physical nature of many devices. They are composed of interconnections between two types of building blocks: linear time-invariant dynamic blocks, and static, memoryless nonlinear blocks. The static nonlinear blocks are introducing the nonlinearities (viz. the departures from the linear behaviour) into the model structure. The restriction is that these blocks are not dynamic, but static: they do not have memory, and hence react instantaneously to their input. In the block structure, they can be viewed as nonlinear amplifiers.

The Wiener and Hammerstein models are two very simple nonlinear model structures from this class of block-oriented models. They simply consist of one block of each kind, both placed in a cascade connection (viz. the second block receives the output of the first block as an input). In the Wiener case, the linear dynamic operation precedes the static nonlinear operation, and vice-versa for the Hammerstein case.

The peculiarity in this work, is that, in contrast to classical system identification approaches, which make use of input and output observation data, the nonlinear Wiener and Hammerstein models are identified here just from output observations only (having no access to the intermediate signal either), assuming the unmeasured input to be described by white Gaussian noise. Therefore, the identification methods will be called blind.

A maximum-likelihood approach will be used to identify the parameters of both the linear and the nonlinear block. The model representation is done in discrete-time, which means that all signals in the model structure are assumed to be defined on a discrete time axis (the time variable only admits integer values).

1.3 Part II: Analysis of instability under random inputs

Besides nonlinear systems, random noise is another central notion in this thesis. Whereas white Gaussian noise was used to represent the unmeasured input to a Wiener or Hammerstein system in the first part, in the second part, the response of a nonlinear system to a random noise excitation is investigated. In practice, it is often unknown whether the response will be bounded or un-
bounded. In this second part, the behaviour of nonlinear systems to random excitations is studied, and an attempt is given to answer the very challenging question “Will the response be bounded or unbounded?” for a given system (or model) and a random input with certain properties. The problem will be viewed from two perspectives: a data-driven (measurement-based) perspective, in which no prior knowledge of the nonlinear system is used, and a model-driven perspective, in which a (discrete-time) state-space model is assumed to represent the system exactly. As expected, for linear systems, the theory is well-established; the boundedness of the response of a linear system to any bounded signal simply depends on the absolute summability of the impulse response. The local and global state-space (3) pictures (viz. trajectories) are identical to each other. Moreover, in the unstable case, unless good control actions (e.g. feedback from output to input) are undertaken, the system starting from rest gets a diverging response immediately from the early beginning of the experiment. A nonlinear system is all the way different. Its local and global state-space pictures can be very dissimilar. As a consequence, even a smooth nonlinear state-space representation (which can be linearized locally), can behave in a totally different manner as the region of operation is changing. Multiple equilibria, attracting/repelling limit cycles, chaotic behaviour, harmonics and subharmonics are features which nonlinear systems do not share with linear systems. When the input’s random properties change, e.g. the variance or the power spectral content, the region covered in the state space is modified. Hence the global state-space picture becomes important. With unforced (viz. zero-input) nonlinear systems, a stable attracting equilibrium is generally attractive only from a subdomain of the entire state-space, called the region of attraction. It may therefore leave the region of attraction not from the early beginning of the experiment, but at a random time instant under the action of the input. A possibility is that the state becomes unbounded. The aim of this part will be to quantify the risk of such an undesired response.

3The physical quantities related to the degrees of freedom of a physical system (or model) are called states (e.g. for a pendulum, this is the angle relative to the vertical and the angular velocity). Any value of the state(s) therefore corresponds to a valid physical situation of the system (e.g. a pendulum is located at a given angle and has given angular velocity). If no external force (or input) is applied, knowing the state(s) of a physical system at some point in time determines completely the states’ future evolution (however, please note that this requires perfect model knowledge). The state-space is the abstract space in which the states evolve over time. A state-space picture is a collection of trajectories representing the states’ evolution over time, allowing to visualise easily how any given state would evolve in the unforced situation.
1.4 Publications

A large part of this thesis is the result of material presented at international conferences and published in journals.

1. Journal publications:


2. Book chapter:


3. International conferences with proceedings:


4. Local conferences with abstract:


5. International workshops:


(e) Van Mulders, A., **L. Vanbeylen** and J. Schoukens (2009). Reduction of polynomial nonlinear state-space models by means of nonlinear similarity transforms. 18th ERNSI workshop on system identification, Vorau, Austria, September 30-October 2, 2009.

6. National workshops:


Part I

Maximum-likelihood approach to the blind identification of Wiener and Hammerstein nonlinear models
Introduction

In the context of block-oriented nonlinear modelling, the simplest nonlinear models built from two different building blocks are the Wiener and the Hammerstein model (see Figure 1.2). The former consists of a linear time-invariant dynamic part (viz. a linear filter) followed by a static (viz. memoryless) nonlinear part, which causes the response to deviate from a purely linear behaviour, and introduces the nonlinear distortion. The latter consists of the same two building blocks (parts), but in reversed order. The menu for this first part of the thesis is the identification of (discrete-time) Wiener and Hammerstein models from output measurements only (which is called blind identification), by following a maximum-likelihood approach. Thence, the main assumptions are the Gaussianity and whiteness of the unmeasured input signal and the invertibility of the model structure. In Chapters 2 and 3, we start with the additional assumption that no measurement noise comes into the output signal. Those two chapters are devoted to Wiener and Hammerstein systems respectively. In Chapter 4, the impact of the noise on these methods is studied, assuming that is not accounted for explicitly in the algorithm. The maximum-likelihood problem with noise will also be formulated, and the relevant problems will be discussed.

Blind nonlinear modelling

Despite their structural simplicity, Wiener and Hammerstein nonlinear model structures have been effective in many application areas, where linear modelling has failed, e.g., the chemical process industry (Eskinat et al., 1991; Kalafatis et al., 1995; Norquay et al., 1999; Zhu, 1999), microwave and radio frequency (RF) technology (Clark et al., 1998; Greblicki, 1996; Prakriya and Hatzinakos, 1997), seismology (Taleb et al., 2001), flood forecasting (Young, 2002), biology (Hunter and Korenberg, 1986), physiology and psychophysics (Nykamp and Ringach, 2002). They can also be used in model predictive control (Wang and
The Wiener and the Hammerstein model – consisting of the cascade connection of a linear time invariant dynamic system and a static nonlinearity, in this or in reverse order (see Fig. 1.2) – both fit inside the family of block-oriented nonlinear structures, which have been extensively studied over the past few decades, see e.g. Billings and Fakhouri (1982), Schoukens et al. (2003), Giri and Bai (2010).

Most of these identification approaches assume the availability of both input and output measurements of the system. However, in several real-world applications, such as operational modal analysis (Peeters and De Roeck, 2001), one often does not have access to the system input: in this case, blind identification of the nonlinear system becomes the only option.

Without measurements of the input, and with little prior knowledge and hypotheses about the system and the input, our goal is to identify the parameters of the nonlinear block-structured model. Blind identification (closely related to blind inversion and equalization) for linear systems has a long history with many theoretical results and applications, especially in telecommunication. Comprehensive overviews can be found in Abed-Meraim et al. (1997) and Tong and Perreau (1998). However, the present knowledge on blind identification of nonlinear systems is rather limited; Table 1.1 gives an overview and comparison of several recently proposed categories of blind identification methods for nonlinear systems.

Maximum-Likelihood Estimators (MLE’s) are proposed, assuming a white unobserved Gaussian input signal, a minimum-phase linear part, an invertible nonlinearity and errorless output observations. The absence of output measurement noise eases the analysis, as will be seen further on. In the context of RF technology, the methods developed in Chapters 2 and 3 can only take place in a (very) controlled environment (e.g. laboratory for characterization of amplifier responses, transmitter-based nonlinear processing) where additive output noise (as would appear after transmission) is negligible in a good approximation. The methods could e.g. be applied to identify the response of an
RF amplification chain when the input is an Orthogonal Frequency-Division Multiplexing (OFDM) communication signal (which tends to a Gaussian distribution, due to its multi-carrier nature, see Banelli, 2003). The major advantages of the maximum-likelihood approach are the consistency (convergence to the true value as the number of data tends to infinity), the asymptotic normality and the asymptotic efficiency (asymptotically the smallest variance) of the estimates. To the knowledge of the author, prior to this Ph.D., no MLE-based method was available for the blind identification of the Wiener and the Hammerstein systems.

Table 1.1: Comparison of several categories of blind identification methods for nonlinear systems

<table>
<thead>
<tr>
<th>Model: approach</th>
<th>Method’s assumptions &amp; properties</th>
</tr>
</thead>
</table>
| Volterra-Hammerstein (Kalouptsidis and Koukoulas, 2005): cumulant-based | - Input: Gaussian and white  
- Robust to low-order moving-average output noise  
- Parametric  
- Consistent |
| Hammerstein-Wiener (Bai, 2002) | - Input: zero-order-hold (piecewise constant)  
- Oversampled output  
- Invertible output nonlinearity admitting a polynomial representation  
- Not robust to output noise  
- Parametric  
- Consistency not analyzed |
- Polynomial nonlinearity  
- First linear system has minimum phase  
- Robust to circular symmetric output noise  
- Nonlinearity not identified  
- Nonparametric  
- Consistency not analyzed |
<table>
<thead>
<tr>
<th>Method (Authors, Year): Description</th>
<th>Characteristics</th>
</tr>
</thead>
</table>
| Wiener (Taleb et al., 2001): mutual information | - Input: non-Gaussian i.i.d.  
- Invertible nonlinearity and filter  
- Not robust to output noise  
- Quasi-nonparametric  
- Consistency not analyzed |
| Wiener (Wills and Ljung, 2010): maximum likelihood, expectation-maximization and particles | - Input: Gaussian and white  
- Robust to white Gaussian output noise  
- Parametric  
- Consistency not formally proven |
| Wiener & Hammerstein (Chapters 2 & 3 of this thesis): maximum likelihood | - Input: Gaussian and white  
- Invertible nonlinearity and filter  
- Not robust to output noise  
- Parametric  
- Consistent and efficient |
Chapter 2

Blind maximum-likelihood identification of Wiener systems

This chapter is about the identification of discrete-time Wiener systems from output measurements only (blind identification). Assuming that the unobserved input is white Gaussian noise, that the static nonlinearity is invertible, and that the output is observed without errors, a Gaussian maximum-likelihood estimator is constructed. Its asymptotic properties are analyzed and the Cramér-Rao lower bound is calculated, allowing to obtain uncertainty bounds. A two-step procedure for the generation of high-quality initial estimates is described; the iterative Gauss-Newton based optimization scheme, needed to find the estimate, is then started up from this point. The chapter includes the illustration of the method on a simulation example and measurement results.

2.1 Chapter outline and contributions

The Wiener model structure and the blind identification problem were introduced right before this chapter. This chapter’s contents is organized as follows: We will start with the problem setting and main assumptions (Section 2.2 ff.), and the Gaussian maximum-likelihood cost function and the asymptotic properties (Section 2.5). Next, we discuss the generation of high-quality initial
Chapter 2. Blind identification of Wiener systems

estimates via a two-step procedure (Section 2.6), the numerical optimization (Section 2.7), the calculation of uncertainty bounds via the Cramér-Rao Lower Bound (Section 2.9). Finally, we proceed with the simulation and measurement results (Sections 2.10 and 2.12).

The main contributions of this chapter are the following:

- generation of high-quality initial estimates for both the linear and the nonlinear part of a Wiener system (also applicable to Hammerstein systems);
- construction of the Gaussian MLE for Wiener systems, as an extension to classical ARMA modelling;
- analysis of the asymptotic properties of the MLE;
- calculation of the Cramér-Rao lower bound (Fisher information matrix);
- verification of the theoretical results by means of a simulation example;
- application of the MLE to a measurement example.

2.2 Class of discrete-time Wiener systems considered

Although, throughout the chapter, we assume all signals and parameters to be real-valued, with a few small modifications, it is possible to use the method for complex-valued variables as well (this may be necessary in Radio Frequency applications; see Section 2.8). Wiener systems are defined as a cascade of a linear time-invariant dynamic system (LTI) $H_0$ and a static (memoryless) nonlinearity (see Figure 2.1). In general, the LTI system can be characterized by its transfer function $H_0(z)$. The static nonlinearity can be characterized by its mapping function from $u$ (intermediate signal) to $y$ (output signal): $y = f_0(u)$. 
2.3 Parameterization of the Wiener system

**Assumption 2.1.** (the class of discrete-time Wiener systems considered)

1. $y = f_0(u)$ is a monotonic, bijective function and its derivative is bounded and non-zero over $\mathbb{R}$.
2. $H_0(z)$ is a causal, stable and inversely stable monic transfer function.

**Remark 2.2.** Assumption 2.1(1) is needed because we work in a blind framework: the value of $u$ should be recoverable from the observed output $y$. Under Assumption 2.1(2), the filter $H_0(z)$ has minimum phase and contains no pure delay. It allows for the computation of the input power spectrum from the power spectrum of the intermediate signal $u$.

### 2.3 Parameterization of the Wiener system

The proposed identification procedure is parametric, hence parameter vectors characterizing the linear and the nonlinear part, respectively denoted by $\theta_L$ and $\theta_{NL}$, are introduced. The transfer function corresponding to the LTI part becomes $H(z, \theta_L)$, parameterized as a rational form in $z^{-1}$,

$$H(z, \theta_L) = \frac{C(z, \theta_L)}{D(z, \theta_L)} = \frac{1 + \sum_{r=1}^{n_c} c_r z^{-r}}{1 + \sum_{r=1}^{n_d} d_r z^{-r}}$$  \hspace{1cm} (2.1)

with $\theta_L^T = [c_1, c_2, \ldots, c_{n_c}, d_1, d_2, \ldots, d_{n_d}]$, and the parameterization of the nonlinearity is in its inverse form: $u = g(y, \theta_{NL})$. Roughly stated, $g$ approximates the inverse of the (forward) nonlinear function $f$.

**Assumption 2.3.** (nonlinear parameterization)

1. $\theta_{NL}$ gives a unique representation of the nonlinearity, viz. $g(y, \theta_{NL}) = g(y, \tilde{\theta}_{NL}) \forall y \Leftrightarrow \theta_{NL} = \tilde{\theta}_{NL}$.
2. $g$ is twice continuously differentiable w.r.t. both function arguments.

Note that the inverse parameterization enables a direct inversion (or equalization) of the Wiener system by means of a Hammerstein system formed by $g$ in series with $H^{-1}$. This is useful for, e.g., calibration applications. Finally, we define a global parameter vector by stacking the parameters of both sub-blocks on top of each other:

$$\theta^T = [\theta_L^T, \theta_{NL}^T]$$  \hspace{1cm} (2.2)

### 2.4 Stochastic framework

**Assumption 2.4.** (stochastic framework)

1. The unknown, unobserved input $e(t)$ is zero mean, white Gaussian noise with unknown variance $\lambda_0 > 0$.
2. The output $y(t)$ is known exactly (i.e. observed without errors).
Note 2.5. In Assumption 2.4(1), note that the identification process does not need any knowledge about the variance $\lambda_0$ of the input. A value of the input variance, compatible with the estimated Wiener model, is estimated from the data during the identification process. Assumption 2.4(2) means that the output measurement signal-to-noise ratio equals infinity in theory, or, is “sufficiently high” in practice. Unfortunately, this is quite a restrictive, but unavoidable assumption of the method. The presence of noise will cause a bias, which will be small for high signal-to-noise ratios (we will come back to this in Chapter 4).

2.5 The Gaussian maximum-likelihood estimator

2.5.1 The negative log-likelihood (NLL) function

Theorem 2.6. Under Assumptions 2.1, 2.3 and 2.4, the conditional Gaussian negative log-likelihood (NLL) function of the observations

$$y_N^T = y^T = [y(0), \ldots, y(N - 1)]$$

(2.3)

given the model parameters $\theta$ and input variance $\lambda$, is:

$$L(y|\theta, \lambda) = \frac{N}{2} \log (2\pi \lambda) + \sum_{t=0}^{N-1} \frac{(H^{-1}(q, \theta_L)g(y(t), \theta_{NL}))^2}{2\lambda} - \sum_{t=0}^{N-1} \log |g'(y(t), \theta_{NL})|$$

(2.4)

with $q$ the forward shift operator ($qx(t) = x(t + 1)$; more information can be found in Ljung (1999) at pp. 24-25), and with $(\bullet)'$ the derivative w.r.t. $y$. Here, “conditional” means “given the initial conditions of the LTI part”. However, asymptotically ($N \rightarrow \infty$), the conditional MLE equals the true MLE.

Proof. The proof is based on the classical expressions of the log-likelihood for ARMA models and on the transformation formula of probability density functions through nonlinear mappings. More details can be found in Appendix 2.A on page 63.

2.5.2 Simplified cost function

By eliminating $\lambda$, the cost function (2.4) can be simplified to a sum-of-squares form. Indeed, $\lambda$ should minimize $L(y|\theta, \lambda)$. Hence, setting $\partial L(y|\theta, \lambda)/\partial \lambda = 0$,.
2.5. The Gaussian maximum-likelihood estimator

yields:

\[ \lambda = \frac{1}{N} \sum_{t=0}^{N-1} (H^{-1}(q, \theta_L)g(y(t), \theta_{NL}))^2 \]  

(2.5)

Substituting (2.5) in (2.4), and defining

\[ \varepsilon_{PE}(t, \theta) = H^{-1}(q, \theta_L)g(y(t), \theta_{NL}) \]  

(2.6)

yields a compressed form of the NLL:

\[ L(y|\theta) = \frac{N}{2} \log 2\pi + \frac{N}{2} \log \left( \frac{1}{N} \sum_{t=0}^{N-1} |\varepsilon_{PE}(t, \theta)|^2 \right) + \frac{N}{2} - \sum_{t=0}^{N-1} \log |g'(y(t), \theta_{NL})| \]  

(2.7)

It can easily be seen that \( \exp(2/N \cdot L(y|\theta)) \) is proportional to the following expression, which will be used as the definition of the cost function \( V(y, \theta) \):

\[ V(y, \theta) = g_N^2(y, \theta_{NL}) \sum_{t=0}^{N-1} |\varepsilon_{PE}(t, \theta)|^2 \]  

(2.8)

with \( g_N(y, \theta_{NL}) = \exp(-\frac{1}{N} \sum_{t=0}^{N-1} \log |g'(y(t), \theta_{NL})|) \) a scalar correction factor which becomes equal to one when no nonlinearity is present (\( u = g(y) = y \), hence \( g'(y) = 1 \)). Note that this correction factor is bounded under the assumption that the nonlinear function is monotonically increasing and differentiable.

Note 2.7. Rewriting (2.8) as

\[ V(y, \theta) = \sum_{t=0}^{N-1} |\varepsilon(t, \theta)|^2 \]  

(2.9)

with

\[ \varepsilon(t, \theta) = (H^{-1}(q, \theta_L)g(y(t), \theta_{NL})) g_N(y, \theta_{NL}) \]  

(2.10)

it is clear that the cost function (2.4) has indeed been reduced to a sum-of-squares form. Moreover, from (2.10), we notice that the maximum likelihood cost function uses \( \varepsilon_{PE}(t, \theta)g_N(y, \theta_{NL}) \) as residuals, whereas the classical prediction error framework would simply use \( \varepsilon_{PE}(t, \theta) \) (without the factor \( g_N(y, \theta_{NL}) \)) instead.

2.5.3 Identifiability discussion

In this paragraph, a set of identifiability conditions are introduced, because thus far the Wiener system was overparameterized. There is a scaling ambiguity,
due to the possibility to displace a scaling factor from the linear part to the nonlinearity, and vice versa. Figure 2.2 illustrates several setups which are not distinguishable from output measurements only (assume $a \in \mathbb{R}_0$).

Assumption 2.8. (identifiability conditions)

1. $H(z, \theta_L)$ has no common pole-zero pairs, and is causal, stable and inversely stable.

2. $H(z, \theta_L) = \frac{C(z, \theta_L)}{D(z, \theta_L)}$ is monic (this means that the constants in the $C$ and $D$ polynomials are 1).

3. $g'(y_1, \theta_{NL}) = 1$ for some $y_1 \in \mathbb{R}$.

Theorem 2.9. (uniqueness of the constrained Wiener model)

Suppose that a signal $y$ is exactly described as the output of a Wiener model subject to Assumptions 2.1, 2.3 and 2.4 and constraints given in Assumption 2.8. Then there exists no different parameter set that provides exactly the same output signal.

Proof. Suppose that there exists a different $(\tilde{\lambda}, \tilde{\theta}_L, \tilde{\theta}_{NL}) \neq (\lambda, \theta_L, \theta_{NL})$. From the Assumption 2.4(1) (Gaussianity), it follows that all the statistical properties of the intermediate signal are captured by its power spectrum. Since Assumption 2.8(1) excludes that $\theta_L$ and $\tilde{\theta}_L$ differ in the presence of all-pass sections or in extra pole-zero canceling pairs, it is found that $\tilde{\theta}_L = \theta_L$. Setting the ratio $\tilde{\lambda}/\lambda = \beta^2$, $\tilde{u} = \beta u$, from which $g(y, \tilde{\theta}_{NL}) = \beta g(y, \theta_{NL})$. Evaluating at $y_1$, and taking Assumption 2.8(3) into account, shows that $\beta = 1$. From Assumption 2.3(1), it follows that $\tilde{\theta}_{NL} = \theta_{NL}$. $\square$
2.5. The Gaussian maximum-likelihood estimator

As a consequence, and since both the model structure and the cost function (2.8) share the same parameter ambiguities, the constraints of Assumption 2.8 remove them all. Since the cost function has been written as a sum-of-squares, its minimizer

$$\hat{\theta} = \arg\min_{\theta} V(y, \theta)$$

(2.11)

can be calculated in a numerical stable way via the classical Gauss-Newton based iterative schemes (Fletcher, 1991). The most likely input variance $\hat{\lambda}$ corresponding to this parameter set $\hat{\theta}$ can be found by evaluating (2.5).

2.5.4 Frequency domain formulation

The time domain expression (2.8) of the cost function corresponds to the MLE given the initial conditions of the linear part $H$ (implicitly assumed in the equations). Since in practice, the initial conditions are unknown, they are estimated. We choose to do this in the frequency domain (1). Applying Parseval’s equality $\sum_{t=0}^{N-1} |x(t)|^2 = \sum_{k=0}^{N-1} |X_k|^2$, where $X_k$ is the (unitary) Discrete Fourier Transform (DFT) at frequency line index $k$ of the vector $x$, defined similarly as $y$ in (2.3),

$$X_k = \text{DFT}_k(x) = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} x(n) z_k^{-n}$$

(2.12)

with $z_k^{-1} = \exp(-j2\pi k/N)$ – to signal $\varepsilon(t, \theta)$ in equations (2.9-2.10), the cost function is rewritten. The DFT of $H^{-1}(q, \theta_L)g(y, \theta_{NL})$, appearing in the equations, equals

$$H^{-1}(z_k, \theta_L) (G_k(y, \theta_{NL}) - T(z_k, \theta_L))$$

(2.13)

with

$$G_k(y, \theta_{NL}) = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} g(y(n), \theta_{NL}) z_k^{-n}$$

(2.14)

and with $T(z_k, \theta_L)$ a transient term stemming from the exact relationship between the input and output DFT spectra $E_k$ and $U_k$ of the linear part (Pintelon et al., 1997):

$$U_k = H(z_k)E_k + T(z_k)$$

(2.15)

1The frequency and time domains are known to be asymptotically equivalent. However, the frequency domain is more convenient and faster to implement the linear filtering operations (the convolution in the time domain becomes a product in the frequency domain). Another advantage of frequency domain is that it allows identifying unstable linear systems without trouble.
with

\[ T(z) = \frac{I(z)}{D(z)} = \sum_{r=0}^{\max(n_e,n_d)-1} i_r z^{-r} \]

The transient \( T(z) \) is a rational form in \( z^{-1} \) with the same denominator coefficients \( d_r \) as the transfer function \( H \), and with numerator coefficients \( i_r \), depending on the initial and final conditions of the experiment. It decreases as \( O(N^{-1/2}) \) as \( N \to \infty \). The parameter vector \( \theta_L \) is extended with the coefficients of the numerator polynomial \( I(z) \).

To summarize, the frequency domain formulation of the maximum likelihood cost function is given by:

\[ V(y, \theta) = \sum_{k=0}^{N-1} |\varepsilon(k, \theta)|^2 \] (2.17)

\[ \varepsilon(k, \theta) = g_N(y, \theta_{NL}) H^{-1}(z_k, \theta_L) (G_k(y, \theta_{NL}) - T(z_k, \theta_L)) \] (2.18)

Remarks:

- It follows that the dependence of the cost function on the noise filter \( H \) does not make it possible to distinguish between a minimum-phase filter and the same filter where some of the zeroes are mirrored around the unit circle. This explains Assumption 2.1(2), where the zeroes were assumed to be inside the unit circle.

- Similar to Pintelon and Schoukens (2006), one could think of performing the identification in the frequency band of interest only. However, since this would throw away information about the nonlinearity, the sum in (2.17) always runs over the complete unit circle.

2.5.5 Asymptotic properties

Although the classical MLE is known to satisfy the asymptotic \((N \to \infty)\) properties (consistency, efficiency and asymptotic normality, see Caines, 1988, p. 317, theorem 3.2), it is here needed to study these properties explicitly, since, in contrast to the standard theory, the underlying noise process is not white, but filtered and nonlinearly distorted white noise. To study them, standard assumptions are first needed concerning the true model, the parametric model and the excitation.

Assumption 2.10. (consistency)

(1) The true model is within the proposed model set, that contains models that satisfy Assumptions 2.1, 2.3, 2.4 and 2.8.
(2) The scaled NLL (without parameter ambiguities) \( L(y_N|\theta, \lambda=1)/N \) has continuous second order derivatives w.r.t. \( \theta \) in a compact (i.e. closed and bounded, Kaplan, 1993, at section 2.2) set \( \Theta \) for any \( N \), infinity included. The compact set \( \Theta \) is constructed such that it contains its unique global minimum is an interior point (i.e. not on the boundaries, Kaplan, 1993, at section 2.2) of \( \Theta \).

(3) There exists an \( N_0 \) such that for any \( N \geq N_0 \), infinity included, the Hessian of the expected value of the scaled NLL (without parameter ambiguities) \( \mathbb{E}[L(y_N|\theta, \lambda=1)/N] \), subject to the constraints given as in Assumption 2.8, is regular (i.e. invertible) at its unique global minimizer in \( \Theta \).

**Assumption 2.11.** (nonlinear rates)

1. \( f_0(u) = \mathcal{O}(|u|^r_1) \) with \( r_1 \in \mathbb{R}_0^+ \) as \( |u| \to \infty \).
2. \( f_0'(u) = \mathcal{O}(|u|^r_2) \) with \( r_2 \in \mathbb{R}_0^+ \) as \( |u| \to \infty \).
3. \( g(y, \theta_{NL}) = \mathcal{O}(|y|^r_3) \) with \( r_3 \in \mathbb{R}_0^+ \) as \( |y| \to \infty \).
4. \( g'(y, \theta_{NL}), \frac{\partial g}{\partial \theta_{NL}}(y, \theta_{NL,0}) \) and \( \frac{\partial g'}{\partial \theta_{NL}}(y, \theta_{NL,0}) \) satisfy a similar power-law rate in \( y \). Herein \( \theta_{NL,0} \) contains the exact parameter values corresponding to \( \theta_{NL} \).

These rate constraints guarantee that the mixing versions of the strong law of large numbers and the central limit theorem (required in the proof the theorem below) apply.

**Theorem 2.12.** Under Assumptions 2.1, 2.3, 2.4, 2.8, 2.10 and 2.11, the MLE \( \hat{\theta} \) is consistent, has convergence rate \( \mathcal{O}(N^{-1/2}) \) in probability, is asymptotically efficient, and asymptotically normally distributed.

**Proof.** See Appendix 2.B on page 64.

2.5.6 Loss of consistency in the case of a non-Gaussian input

Since the MLE is constructed by assuming a white Gaussian input \( e(t) \), the consistency of the estimate \( \hat{\theta} \) can be lost if \( e(t) \) is non-Gaussian distributed. This effect is due to the nonlinearity, since in the linear case (Pintelon and Schoukens, 2006) the estimate is still consistent. A counterexample (in Appendix 2.C on page 67) illustrates this loss of consistency.

2.6 Generation of initial estimates

Since the MLE is defined as the minimizer of a non-quadratic-in-the-parameters cost function, the Newton-Gauss minimization may be affected by local minima.
Hence, there is a need for good initial estimates to start the iterative procedure from.

The properties that

- a linear system operating on a Gaussian input yields a Gaussian output;
- the samples from a white Gaussian signal are independent and identically distributed (i.i.d.);
- and, a static system operating on an i.i.d. input yields an i.i.d. output;

lead to the observation that, for the Wiener (and the Hammerstein) system driven by white Gaussian noise, the input to the LTI part is i.i.d. and the input to the static nonlinearity is Gaussian (see Fig. 1.2). Hence, the initial estimates of both the linear and the nonlinear block may be generated by a combination of the following two subproblems:

1. the estimation of the Gaussian input \( u_{SNL}(t) \) of a monotonically increasing static nonlinearity from its output signal \( y_{SNL}(t) \) (and the nonparametric and parametric estimation of the nonlinear function);
2. the estimation of the white input \( u_{LTI}(t) \) of an LTI system from its output signal \( y_{LTI}(t) \) (and the parametric estimation of the transfer function).

For a Wiener system, application of 1. to the (measured) output signal yields the intermediate signal (and the static nonlinearity \( f \)), from which 2. allows one to estimate the LTI \( H \). (For a Hammerstein, application of 2. to the (measured) output signal yields the intermediate signal (and the transfer function of the LTI \( H \)), from which 1. allows to estimate the static nonlinearity \( f \).)

### 2.6.1 Subproblem 1: monotonically increasing static nonlinearity driven by Gaussian noise

The nonlinear function \( g \) is monotonically increasing; as a consequence it preserves order relationships. For every pair \((\bar{u}, \bar{y})\) satisfying \( \bar{y} = g(\bar{u}) \), we have that

\[
F_{U_{SNL}}(\bar{u}) = P(U_{SNL} \leq \bar{u}) = P(Y_{SNL} \leq \bar{y}) = F_{Y_{SNL}}(\bar{y})
\] (2.19)

with \( U_{SNL} \) and \( Y_{SNL} \) the random processes corresponding respectively to the input \( u_{SNL}(t) \) and the output \( y_{SNL}(t) \) of the static nonlinearity, and \( F_{U_{SNL}} \) and \( F_{Y_{SNL}} \) the cumulative distribution functions of respectively \( U_{SNL} \) and \( Y_{SNL} \), and \( P(A) \) the probability of event \( A \). The following explains how (2.19), equating the corresponding quantiles, allows to estimate \( u_{SNL}(t) \). First, note that the (marginal) distribution of \( U_{SNL} \) is known (within an unknown variance
scaling factor). Second, we define the empirical distribution function of \( Y_{SNL} \) as

\[
\hat{F}_{Y_{SNL}}(y_{SNL}(t)) = \frac{1}{N} \sum_{r=0}^{N-1} I\{y_{SNL}(r) \leq y_{SNL}(t)\} - \frac{1}{2N}
\] (2.20)

with the finite sample correction \( \frac{1}{2N} \), and \( I\{A\} \) the indicator function associated with the statement \( A \), admitting the value 1 if \( A \) is true and 0 otherwise. The finite sample correction ensures that the expression (2.20) is contained in the interval \( \left[ \frac{1}{2N}, 1 - \frac{1}{2N} \right] \) and that the next step (2.21) results in a finite value. Finally, the input signal samples are reconstructed (within a scaling factor) as

\[
\hat{u}_{SNL}(t) = \Phi^{-1}\left( \hat{F}_{Y_{SNL}}(y_{SNL}(t)) \right)
\] (2.21)

with \( \Phi^{-1} \) the standard normal quantile function. The latter is the inverse function \( x = \Phi^{-1}(y) \) of the cumulative distribution function \( \Phi(x) = \mathbb{P}(X < x) = y \) of a standard normally distributed random variable \( X \). Plotting \( \hat{u}_{SNL}(t) \) and \( y_{SNL}(t) \) as function of each other gives a nonparametric estimate of the nonlinearity, which can be used to select a parameterization \( g(y_{SNL}, \theta_{NL}) \) for the static nonlinear function. Additive independent measurement noise would stretch the support of the output pdf compared to the noiseless case, due the convolution of pdf’s taking place. As a consequence, the nonlinearity would be affected by a bias.

The variances of each reconstructed input sample can be estimated, and next be used as weights in a weighted least squares procedure to obtain the parameter estimates \( \hat{\theta}_{NL}^{(0)} \):

\[
\hat{\theta}_{NL} = \arg\min_{\theta_{NL}} \sum_{t=0}^{N-1} w_t |\hat{u}_{SNL}(t) - g(y_{SNL}(t), \theta_{NL})|^2
\] (2.22)

The weights \( w_t \) are chosen as the inverses of the variances of the reconstructed \( \hat{u}_{SNL} \). These can be approximated (Taylor expansion) as (dropping the subscript \( SNL \)):

\[
\text{var } \hat{u}_{SNL} \approx (\partial \Phi^{-1}/\partial \hat{F}_Y)^2 \text{var } \hat{F}_Y(y)
\] (2.23)

Although the output observations are not independent, \( \hat{F}_Y(y) \) is treated as if it were binomially distributed (mean of the Bernoulli experiments \( Y \leq y \)), and a first-order approximation of its variance is estimated by \( \text{var } \hat{F}_Y(y) \approx \hat{F}_Y(y)(1 - \hat{F}_Y(y))/N \) \(^2\); finally \( \hat{F}_Y(y) \) is replaced by \( \hat{F}_Y(y) \). The optimization (2.22) can be done iteratively with a Gauss-Newton algorithm.

\(^2\)The variance can be obtained neglecting the covariances over different time instants, and
Chapter 2. Blind identification of Wiener systems

The following Matlab code illustrates an implemented version of the algorithm generating the nonparametric reconstruction and the first order variances:

```matlab
function [uHat, varU] = NonparNL(y)
% Nonparametric estimate of the nonlinearity with Gaussian input
% Uses cdf to compute corresponding u0 and y0 values
% assuming that U is standard normally distributed.
% y = measured output SNL as column
% uHat = estimated input SNL as column
% varU = variances of each reconstructed data point uHat as column
% plot(uHat,y,'.') produces the nonparametric estimate
N = length(y);
ysort = sort(y);
[yu, i] = unique(ysort);
Prob_levels = (cumsum([i(1); diff(i)])-0.5)/N; % empirical distribution
function of the sorted data
varP = Prob_levels.*(1-Prob_levels)/N;
% put in original order
[yunique, i_, j_] = unique(y);
% y0 = yunique(j_); % is automatically true
uHat = u0(j_);
varU = varU(j_);
```

2.6.2 Subproblem 2: LTI driven by white input noise

The second subproblem can be solved by applying classical ARMA algorithms to the output of the LTI \( y_{LTI}(t) \), either in the time domain (Ljung, 1999) or the frequency domain (noise model of Pintelon and Schoukens, 2006). This also coincides with the special case of a blind linear MLE of \( \theta_L \) (cost function (2.7) with \( g(x) = x \)). Let the estimate be denoted as \( \hat{\theta}_L^{(0)} \). After inverse filtering of

\[
\operatorname{var} \hat{\theta}_L^{(0)}(y) = \frac{1}{N^2} \sum_{r,s} \operatorname{cov} (I_{\{y(r) \leq y\}}, I_{\{y(s) \leq y\}})
\]

\[
= \frac{1}{N^2} \sum_r \operatorname{var} (I_{\{y(r) \leq y\}}) + \text{mixed terms being neglected}
\]

\[
\approx \frac{1}{N} \operatorname{var} (I_{\{y(r) \leq y\}})
\]

\[
= \frac{1}{N} \left( \mathbb{E} \left[ I_{\{y(r) \leq y\}}^2 \right] - \mathbb{E} \left[ I_{\{y(r) \leq y\}} \right]^2 \right) = \frac{1}{N} F_Y(y) (1 - F_Y(y))
\]

using ergodicity and stationarity of the random process.
2.7 Minimization of the cost function

Gauss-Newton based iterative algorithms (Fletcher, 1991) for optimizing sum-of-squares cost functions require use of the Jacobian matrix, defined as the partial derivatives of the residuals \( \varepsilon(k, \theta) \) (2.18) w.r.t. the parameters: 
\[
J_{ki}(y, \theta) = \frac{\partial \varepsilon(k, \theta)}{\partial \theta_i}.
\]
Its calculation is given in Appendix 2.D. In the general case, the practical implementation requires the calculation of the mixed second order derivatives \( \partial^2 g(y, \theta_{NL}) / (\partial y \partial \theta_i) \), which can be quite a tedious task. Therefore, we propose a simple parameterization for the nonlinearity.

2.7.1 Parameterization of the nonlinearity

To simplify the computations, a model that is linear-in-the-parameters is used to describe the nonlinearity \( g(y) \):
\[
u = g(y, \theta_{NL}) = f(y)^T \theta_{NL}
\] (2.24)
with \( \theta_{NL}^T = [\theta_{NL,1}, \ldots, \theta_{NL,M}] \) the parameter vector of the nonlinear block, and with \( f(y)^T = [f_1(y), \ldots, f_M(y)] \) the vector of basis functions. Moreover, this linear parameterization simplifies, considerably, the weighted least squares computation (2.22) in the initial estimates algorithm; the problem is linear-in-the-parameters, and an iterative procedure is avoided.

For a fixed vector of basis functions \( f(y) \), the parameters \( \theta_{NL} \) satisfy Assumption 2.3(1). Hence, the model becomes identifiable. It can be noted that a linear combination of the basis functions (actually, here we only use a fixed basis) would result in a same input-output behaviour if the parameters \( \theta_{NL} \) would be altered accordingly in a linear way.

Convenient choices for the basis functions \( f_m(y) \) are: (i) powers of \( y \) (polynomial expansion) (ii) properly scaled orthonormal polynomials, e.g. scaled Hermite polynomials which are orthonormal w.r.t. the Gaussian probability density function (with the same variance as the measured output sample variance):
\[
f_m(y) = \frac{1}{\sqrt{2^m m!}} H_m\left(\frac{y}{\sqrt{2} \sigma_y}\right)
\] (2.25)
This scaling of Hermite polynomials \( H_m(x) \) is needed, since classical Hermite polynomials are orthogonal w.r.t. \( \exp(-x^2) \). To avoid bad numerical condi-
tioning, the basis functions are divided by their standard deviations over time. In Appendix 2.D, explicit expressions for the Jacobian matrix are given for the parameterization (2.24).

### 2.7.2 Gauss-Newton procedure

The equations of \( J_{ki} \) make it possible to compute the Jacobian matrix from the data and the parameters, and to minimize the cost function \( V(y, \theta) \). This minimization is performed by using a Levenberg-Marquardt algorithm. It yields the most likely parameter vector \( \hat{\theta} \), given the output observations \( y \), if the global minimum is found.

### 2.7.3 Practical implementation of the constraint

In practice, for the model (2.24), the constraint on the nonlinearity is not implemented as stated in Assumption 2.8(3), but by putting the two-norm of the nonlinear parameter vector in the linear parameterization to unity:

\[
\sum_m \theta_{NL,m}^2 = \|\theta_{NL}\|^2_2 = 1 \tag{2.26}
\]

This norm-one constraint avoids fixing a particular coefficient to unity. The latter has the disadvantage that it could be detrimental to the numerical conditioning if the true value of the coefficient is zero. When performing the optimization, this constraint between the coefficients \( \theta_{NL} \) is taken into account at each step of the Gauss-Newton procedure, by using the pseudo-inverse of

\[
J_{re}(\theta) = \begin{bmatrix} \text{Re}(J(\theta)) \\ \text{Im}(J(\theta)) \end{bmatrix} \tag{2.27}
\]

for solving the normal equations, and then imposing (2.26) on the updated parameter vector.

Our approach is motivated as follows. Locally speaking, the steps are taken orthogonal to the degenerate directions of the cost function. Since the parameterization is linear in \( \theta_{NL} \), at each step the unit-norm constraint is easily implemented by dividing \( \theta_{NL} \) by its two-norm. This would not be possible for every general parameterization. More advanced normalizations of the global nonlinear gain (at the input-side of the nonlinearity) should then be used.

The unit-norm constraint could also be imposed via an explicit constraint during the optimization. However, this would require a constrained optimization method, which is more sophisticated than an unconstrained one.
2.8 Complex-valued case

As stated before, the method remains applicable in the case of complex signals and coefficients. This may be necessary in Radio Frequency applications. However, a few small modifications are needed and are listed below:

Assumption 2.1(1) becomes somewhat more complicated: \( f_0 \) should be a bijective, analytic and diffeomorphic (i.e. differentiable and inversely differentiable) function, with a Jacobian matrix (w.r.t. real and imaginary parts), whose determinant is bounded and different from zero (almost everywhere). The expression in the log-likelihood expression (2.4) (and resulting equations) should be interpreted as the determinant of the Jacobian matrix, which equals the magnitude squared of the complex derivative of \( g \) to \( y \). Moreover, Assumption 2.4(1) should be modified to a circular complex normally distributed \( e(t) \), i.e. real and imaginary parts are uncorrelated, have zero mean and have the same variance \( \lambda_0/2 \). Applying this and following the lines of the calculation of the NLL function, yields:

\[
L(y|\theta,\lambda) = N \log (\pi \lambda) + \sum_{t=0}^{N-1} \left| H^{-1}(q,\theta_L)g(y(t),\theta_{NL}) \right| \frac{N-1}{\lambda} \log |g'(y(t),\theta_{NL})| - 2 \sum_{t=0}^{N-1} \log |g'(y(t),\theta_{NL})| \tag{2.28}
\]

with (\( \bullet \)' denoting the complex derivative. Also, the algorithm that generates initial estimates for the nonlinear part still works in the complex-valued case: it suffices to apply the method to the real and imaginary parts separately \((3)\). This allows to reconstruct the intermediate signal up to a complex factor (amplitude and phase) \((4)\). Since we are in a blind setting, this complex factor is part of the identifiability issue and is solved by imposing the identifiability constraints given before.

2.9 The Cramér-Rao lower bound

Since we know from Section 2.5.5 that the estimator is asymptotically \((N \to \infty)\) efficient – which means the asymptotic covariance matrix equals the (CRB) – its calculation allows us to compute the uncertainties on the estimated model parameters. Note that the full model structure is overparameterized if no

\(3\) The vector consisting of the real and imaginary parts of \( y(t) \) follows a non-Gaussian distribution in 2 dimensions. Due to the analytic property of the static nonlinear function, a nonlinear “shrinking” of both the real and imaginary parts to the same variance is the only option to obtain a circular normal distribution.

\(4\) The phase is a consequence of the circular normal distribution, being invariant to a rotation. The amplitude factor is the consequence of the fact that the output is invariant to displacements of scaling factors in the model structure (see Section 2.5.3).
constraints are imposed, and that the NLL (2.4) remains invariant under the transformation
\[
\lambda \rightarrow 1, \, g \rightarrow g/\sqrt{\lambda}
\] (2.29)
which preserves the statistical properties of the output observations (cf. the equivalence between setup 1 and 3 in Fig. 2.2 with \(a = 1/\sqrt{\lambda}\)). From Pintelon et al. (1999), it follows that exactly the same result would have been obtained if another choice had been made to make the model structure identifiable. Although the obtained Fisher Information Matrix (FIM) depends on this choice, the uncertainty bounds on invariants, like the statistical properties of the observations or the pole/zero positions, calculated via a certain chosen CRB, are independent of the specific choice of the imposed constraints. Hence, without any loss of generality, the computations of the FIM – which is exactly the inverse of the CRB – may be performed by imposing \(\lambda = 1\), thus simplifying the calculations considerably. In practice, transformation (2.29) is applied before calculating the FIM with the expressions in the theorem given below.

**Theorem 2.13.** For the linear parameterization (2.24), and under the Assumptions 2.1, 2.3, 2.4, 2.8, 2.10 and 2.11, the general entry \(F_{i,j}\) of the asymptotic FIM for unity input variance equals
\[
\begin{cases}
2 \sum_{k=0}^{N-1} \text{Re}(H_k^{-1} \frac{\partial H_k}{\partial \theta_i}) \text{Re}(H_k^{-1} \frac{\partial H_k}{\partial \theta_j}) & \text{for } \theta_i \in \theta_L, \, \theta_j \in \theta_L \\
-2E \left[ \sum_{k=0}^{N-1} \text{Re}(H_k^{-1} \frac{\partial H_k}{\partial \theta_i}) \text{Re}(|H_k|^{-2} G_k \frac{\partial G_k}{\partial \theta_j}) \right] & \text{for } \theta_i \in \theta_L, \, \theta_j \in \theta_{NL} \\
E \left[ \sum_{k=0}^{N-1} \text{Re}(|H_k|^{-2} \frac{\partial G_k}{\partial \theta_i} \frac{\partial G_k}{\partial \theta_j}) \right] + N \text{E} \left[ g' - 2 \frac{\partial g'}{\partial \theta_i} \frac{\partial g'}{\partial \theta_j} \right] & \text{for } \theta_i \in \theta_{NL}, \, \theta_j \in \theta_{NL}
\end{cases}
\] (2.30)
with \(H_k\) is a short-hand for \(H(z_k, \theta_L)\), \(G_k\) defined at (2.13) and with \(g' = \frac{\partial g(y, \theta_{NL})}{\partial y}\).

**Proof and practical calculation.** See Appendices 2.E on page 68 and 2.F on page 71.

Notice that the FIM does not contain rows and columns corresponding to the transient coefficients, since these vanish asymptotically.

### 2.10 Simulation results

#### 2.10.1 Setup: presentation of the example

In this section, the results of the method applied to a simulation example are discussed. The setup consists of a fourth order filter \(H_0\) with transfer function
\[
\frac{1 + 0z^{-1} - 0.49z^{-2} + 0.01z^{-3} + 0.065z^{-4}}{1 + 0.3676z^{-1} + 0.88746z^{-2} + 0.52406z^{-3} + 0.55497z^{-4}}
\] (2.31)
2.10. Simulation results
cascaded with a fifth-order polynomial inverse nonlinearity: $g_0(y) = \theta_{NL,1} + \theta_{NL,2}y + \theta_{NL,3}y^3 + \theta_{NL,4}y^5$ (the basis functions $f_m$ are powers of $y$). Moreover, if $\theta_{NL,2}$, $\theta_{NL,3}$, and $\theta_{NL,4}$ all have the same sign, then the invertibility Assumption 2.1(1) is met (the derivative has the same sign everywhere). The particular choices of the coefficients are given by: $\theta_{NL}^T = [2, 1, 4, 2]$. Since $\hat{\theta}_{NL}$ only approximates $\theta_{NL}$ within an unknown scaling factor, the excitation variance is chosen to be $1$, such that the resulting estimates $\hat{\theta}_{NL}/\sqrt{\lambda}$ can be compared with the true parameters $\theta_{NL}$. The simulations consisted of 250 runs of a $N = 8192$-samples simulation with independent, zero mean, Gaussian noise excitation. After filtering $e(t)$ with $H_0(z)$ (zero initial conditions) to generate the intermediate signal $u(t)$, the corresponding output signal $y(t)$ is calculated numerically by solving the nonlinear equation $u = g_0(y)$ w.r.t. $y$ (Gauss-Newton). As already mentioned, the basis functions $f_m$ were scaled (normalization) by their standard deviations, to avoid numerical problems. After generation of the initial estimates, and joint MLE optimization (relative tolerance on cost function and relative tolerance on parameters are both $10^{-10}$, with a maximum of 80 iterations), the CRB was calculated. Finally, both the estimates and the Cramér-Rao matrix were denormalized, run per run (the standard deviations used as norm are data dependent).

2.10.2 Graphical presentation of the results
The transfer function $H_0$ and the nonlinear function $u = g_0(y)$ are depicted in Figure 2.3 and Figure 2.4 respectively. Moreover, in each figure, the properties of the mean estimate over $R = 250$ runs are shown: (i) its sample standard deviation (computed as $R^{-1/2}$ times the sample standard deviation over $R$ runs), (ii) the absolute value of its difference with the true values, and (iii) the 95% simultaneous confidence level of the difference. The latter is computed by applying the bootstrap method given in Tjärnström (2000) using 5000 multivariate normally distributed parameter vectors with mean $\theta_0$ and covariance matrix the estimated CRB. Then, the simultaneous confidence bounds of the absolute value of the difference between the true values and a random realization at 100 points (frequencies or $y$-values), are multiplied by $R^{-1/2}$ to obtain the result in the figures.

From these figures, it can be seen that the difference between the true values and its mean estimate, is (a) orders of magnitudes smaller than the true values, and (b) smaller than its 95% simultaneous confidence level. Therefore no systematic error can be detected, which confirms the consistency of the MLE. Moreover, from Figure 2.4, it is clear that the curves (i) and (iii) have a minimum at about the zero crossing of the nonlinearity, which is related to the fact that $u(t)$ is a zero mean signal, and most of the data are present around
Chapter 2. Blind identification of Wiener systems

Figure 2.3: Transfer function of the linear part of the Wiener system, and properties of the mean estimate: (sample) standard deviation, magnitude of the complex difference between the true and the mean estimate, and its 95% simultaneous confidence level.

Figure 2.4: Mapping function $u = g(y)$ of the nonlinear part of the Wiener system, and properties of the mean estimate: (sample) standard deviation (i), absolute value of the difference between the true and the mean estimate (ii), and its 95% simultaneous confidence level (iii).

zero.

In Figure 2.5 and Figure 2.6, the normalized histograms of the parameters are depicted, together with their true values, the normalized histogram of their initial estimates, and their probability densities estimated from the mean CRB.
2.10. Simulation results

Figure 2.5: Normalized histogram (grey bars) of the estimated parameters of the linear system (above: numerator coefficients, below: denominator coefficients), their true value (circle below), normalized histogram of initial estimates (white bars, coinciding with histogram of optimized values), probability densities according to mean CRB (bold).

Figure 2.6: Normalized histogram (grey bars) of the estimated parameters of the nonlinear system, their true value (circle below), normalized histogram of initial estimates (white bars), probability densities according to mean CRB (bold).

Note 2.14. Consistent with the MLE properties outlined earlier, the results clearly show:

(i) the excellent quality of the initial estimates, especially for the linear system (no significant improvements are achieved by the joint optimization in the given example)
Figure 2.7: Cross validation by comparing the spectrum of the intermediate
signal via the linear and via the nonlinear block (for a randomly selected run).
Notice that the model explains the observations very well.

(ii) the asymptotic normality of the estimates

(iii) the consistency (estimates are nicely gathered around the true pa-
rameter values)

(iv) the asymptotic efficiency of the MLE (asymptotic covariance matrix
equals CRB).

To cross-validate the results, the power spectrum of the intermediate signal,
calculated as $\sqrt{\lambda H}$ is compared with the spectrum of $g(y)$. Figure 2.7 represents
both spectra, and the standard deviations, computed as the diagonal elements
of the matrix $\text{cov}(\tilde{H}) \approx \frac{\partial \tilde{H}}{\partial \theta} \cdot \text{cov}(\theta) \cdot (\frac{\partial \tilde{H}}{\partial \theta})^H$ with $\tilde{H} = \sqrt{\lambda H}$.

2.11 Practical considerations about real-life systems operating in continuous-time

Almost any real-life system operates in continuous-time, while the modelling is
done in discrete-time. The sampling of the continuous-time signals may hence
result in systematic errors. These can be reduced by following the advices listed
below:

- Reducing the aliasing errors due to the nonlinearity – which transports
2.12. Laboratory experiment

The excitation signal was originally generated in Matlab, using the “randn” function, producing white Gaussian noise. After the acquisition (32768 data points and sampling frequency 78125 Hz), the algorithm was given the output data (CH2) only. The basis functions for the inverse nonlinearity were carefully selected to approximate its true behaviour as good as possible (inverting the quadratic function, yields , hence and ), and the orders of the linear part were chosen to produce spectrally white residuals (orders / were retained). The identification results are shown in Fig. 8, 9 and 10. The reconstructed wave forms cannot be distinguished from Gaussians; this is a first validation.

Figure 8: Properties of the reconstructed input signal
Upper figure: histogram and black line representing the Gaussian probability density function with variance as computed from the model.
Bottom figure: power spectrum (computed from reconstructed data) and white line representing power spectrum as computed from the model.

Figure 2.8: Measurement setup.

signals’ energy contents over frequency \(^5\) – requires oversampling, i.e. to choose the Nyquist frequency (viz. half the sampling frequency) at least e.g. 10 times the bandwidth of the linear part. An anti-aliasing filter can always be used to suppress noise and other disturbances outside the measured frequency band.

- If a piecewise constant input is assumed, a discrete-time LTI will be obtained that corresponds to the step-invariant equivalent of the continuous LTI (not affecting the model order). The nonlinearity is static and hence remains unaffected.

- If the input is assumed to be band-limited continuous-time noise \(^6\) with a flat power spectrum in the bandwidth of the LTI, the discrete-time LTI will also capture the inter-sample behaviour of the input in that case. Note that this may result in a model order increase. As the sampling frequency grows (very) large, the model order will however become correct. Besides, the model will also only be valid, under the condition that the input properties during the estimation are in agreement with those holding while using the model later on.

- In practice, model order selection criteria such as the Akaike Information Criterion or the Minimal Description Length should be applied to balance well the trade-off between the model complexity and model fit.

2.12 Laboratory experiment

To show the potential usefulness of the approach in measurement science, a laboratory experiment has been elaborated, and the measurement results are

\(^5\)E.g. a cubic nonlinearity yields a response contributions up to three times the input’s highest frequency component.

\(^6\)Note that white continuous-time noise does not exist, since this would require an infinite variance.
Chapter 2. Blind identification of Wiener systems

reported. The setup, intended to emulate a nonlinear sensor, consisted of a Tektronix AM502 differential amplifier (cutoff frequency set to 3 kHz), connected to a static squaring electrical circuit based on Analog Devices’ Internally Trimmed Integrated Circuit Multiplier AD532KH. Although this nonlinearity is globally non-invertible, the adjustable output DC offset of the differential amplifier was set to obtain a positive voltage waveform before the squaring circuit. Consequently, the measurements are located on one branch of the squaring circuit, and thus the nonlinearity is invertible in this part. To avoid electrical loading between the subsystems, wideband $50\Omega$ buffers were inserted. The devices were excited by means of an Agilent HP E1445A Arbitrary Waveform Generator, and two synchronized Agilent HP E1430A Acquisition channels were used. Figure 2.8 depicts the whole measurement setup. The excitation signal was originally generated in Matlab, using the “randn” function, producing white Gaussian noise. After the acquisition (32768 data points at a sampling frequency of 78125 Hz), the algorithm was given the output data (CH2) only. The basis functions for the inverse nonlinearity were carefully selected to approximate its true behaviour as good as possible (inverting the quadratic function $y = \left(\frac{u-a}{b}\right)^2$, $y > 0$ yields $u = a + b\sqrt{y}$, hence $f_1(y) = 1$ and $f_2(y) = \sqrt{y}$), and the orders of the linear part were chosen to produce spectrally white residuals (orders $n_c/n_d = 3/3$ were retained). The identification results are shown in Figures 2.9 to 2.11. The reconstructed waveforms cannot be distinguished from Gaussians; this is a first validation.

Finally the obtained estimates were validated a second time by comparison with the results of a multisine measurement. The input variance, as well as all the other settings, was kept constant. A special odd multisine (almost normally distributed periodic signal with only odd harmonics excited) was applied and one period of ACQ CH1 and CH2 were measured (observe that this is not a blind experiment). The frequency response function (FRF) of the linear part was then calculated between the applied waveform (AWG) and CH1. Afterwards, a scatter plot of CH1 versus CH2 was used for comparison with the estimate of the static nonlinear function. This validation can be found in Figures 2.12 and 2.13.

2.13 Conclusion

In this chapter, a maximum-likelihood estimator for a blind identification of Wiener systems has been proposed. The main assumptions made are the following: the (unobserved) input signal is white, Gaussian noise, the output signal is measured exactly, and the nonlinearity is invertible with nonzero derivative. It has been shown that the MLE cost function can be reduced to a sum-of-squares
2.13. Conclusion

Figure 2.9: Properties of the reconstructed input signal $e$. Top: histogram and black line representing the Gaussian probability density function with variance $\hat{\lambda}$ as computed from the model. Bottom: power spectrum (computed from reconstructed data) and white line representing power spectrum $\hat{\lambda}H$ as computed from the model.

Figure 2.10: Properties of the reconstructed intermediate signal $u$. Top: histogram and black line representing the Gaussian probability density function with variance computed from the model. Bottom: power spectrum (computed from reconstructed data) and white line representing power spectrum $\hat{\lambda}H$ as computed from the model.
Chapter 2. Blind identification of Wiener systems

Figure 2.11: Estimated nonlinearity: comparison of the different estimates (almost coinciding).

Figure 2.12: Validation of the linear part $H$. Top: notice the excellent agreement of both FRFs (almost coinciding). Bottom: the grey line indicates the magnitude of the complex difference between FRFs, the dotted line represents the measurement variance of the multisine-measured FRF.
form, which is the classical prediction error cost function, multiplied by a correction factor due to the nonlinearity. The iterative Gauss-Newton-based methods can be used for calculating the estimates, and high quality initial estimates can be obtained via a two-step procedure. Finally, an estimate for the asymptotic covariance matrix of the parameters was calculated via the Cramér-Rao lower bound. A simulation example has been included showing that the results agree nicely with the theory, and measurement results are shown.

Appendices

2.A Calculation of the log-likelihood function

Let \(e\) and \(u\) be defined (similarly to \(y\)) by the \(N \times 1\) data vectors containing the unobserved input \(e(t)\) and the unobserved intermediate signal \(u(t)\). Up to the static nonlinearity, the model structure is exactly the classical ARMA model (also known as time-series model or spectral model) with a monic noise filter \(H(q)\). Therefore, as follows from equation (7.87) in Ljung (1999), the expression of the NLL of the intermediate signal, if it were known, is given by (within a constant):

\[
- \log p_u(u|\theta, \lambda) = \frac{N}{2} \log \lambda + \frac{1}{2\lambda} \sum_{t=0}^{N-1} (H^{-1}(q, \theta_L)u(t))^2
\]  

(2.32)
Chapter 2. Blind identification of Wiener systems

Herein, $p_u(u|\theta, \lambda)$ represents the probability density function (pdf) of the unobserved intermediate signal $u$.

Since $u(t) = g(y(t), \theta_{NL})$, the pdf’s of the output signal $p_y(y|\theta, \lambda)$ and of the intermediate signal $p_u(u|\theta, \lambda)$ are related by:

$$p_y(y|\theta, \lambda) = p_u(u|\theta, \lambda) |\partial u/\partial y|$$  \hspace{1cm} (2.33)

with $\partial u/\partial y$ the Jacobian matrix of partial derivatives, which – since $y(t)$ only depends on $u(t)$ – is given by:

$$\partial u/\partial y = \text{diag}(g'(y(0), \theta_{NL}), \ldots, g'(y(N - 1), \theta_{NL}))$$  \hspace{1cm} (2.34)

$$|\partial u/\partial y| = \prod_{t=0}^{N-1} |g'(y(t), \theta_{NL})|$$  \hspace{1cm} (2.35)

Combining (2.33) and (2.35) yields:

$$- \log p_y(y|\theta, \lambda) = - \log p_u(u|\theta, \lambda) - \sum_{t=0}^{N-1} \log |g'(y(t), \theta_{NL})|$$  \hspace{1cm} (2.36)

Plugging in (2.32), finally proves (2.4).

\[\square\]

2.B Asymptotic properties

Lemma 2.15. Under the consistency assumptions of Theorem 2.12, $y(t)$ is $\alpha$-mixing at exponential rate and cumulant mixing ($\gamma$) of order $\infty$.

Proof. From Assumption 2.1(2), the impulse response of the LTI decays at exponential rate. The input $e(t)$ is i.i.d. Therefore, the process $u(t)$ is $\beta$-mixing at exponential rate (see Fan and Yao, 2003, p.69, iii), and hence $\alpha$-mixing at the same rate (Bradley, 2005, eq. (1.11)). The process $y(t)$ is a memoryless function of $u(t)$, and has therefore the same $\alpha$-mixing properties (Fan and Yao, 2003, p.69, ii). Assumption 2.11(1) implies that $y(t)$ has bounded (absolute) moments of all orders; for, by continuity, the function $f_0(u)$ is bounded, the

\[\gamma \] $x_N(k)$ is said to be mixing of order $p$ if it satisfies an absolute summability condition of cumulants (Brillinger, 1981; Leonov and Shiryaev, 1959):

$$\max_{k_p} \lim_{N \to \infty} \sum_{k_1, \ldots, k_{p-1}=0}^{N-1} |\text{cum}(x_N(k_1), x_N(k_2), \ldots, x_N(k_p))| < \infty$$  \hspace{1cm} (2.37)

This condition implies that the span of dependence (up to the $p$th order order moments) over time is sufficiently small (see section 14.4.1 of Pintelon and Schoukens (2001)).
2.B. Asymptotic properties

The integral contribution of the pth \((p \in \mathbb{N}_0)\) order moment is finite over a finite interval, and the contribution in the tails are bounded by the moments of a Gaussian. Hence, \(y(t)\) is cumulant mixing of order \(\infty\) (Doukhan and León, 1989, proposition 2.2).

For the proof of Theorem 2.12, we follow the lines of section 14.13 of Pintelon and Schoukens (2001) applied to the NLL (2.4), and fill in the “gaps” — i.e., it only remains to guarantee the applicability of the Law of Large Numbers (LLN) and Central Limit Theorem (CLT) — for the actual model structure used. To avoid the model ambiguity problem in this reasoning, \(\lambda\) is fixed to unity (see also Assumptions 2.10(2) and 2.10(3)), allowing to treat \(\theta\) as an unconstrained parameter vector.

Note. The reader might wonder why \(\alpha\)-mixing (“strong mixing”) is used here in addition to cumulant mixing. The reason is that if the function \(g(\cdot, \theta_{NL})\) is polynomial, its inverse (being \(f_0(\cdot)\) at \(\theta_{NL,0}\)) cannot be polynomial. Cumulant mixing has properties for polynomial static nonlinearities, while \(\alpha\)-mixing is preserved over arbitrary static nonlinearities.

Consistency: As for any MLE with a finite-dimensional parameter-vector, and under Assumption 2.10(1), the expected NLL is minimal at the exact parameter values \(\theta_0\). Here comes a very short proof of that claim, from (Kendall and Stuart, 1979, section 18.10):

The geometric mean of any non-degenerate random variable can never exceed its arithmetic mean, hence for any \(\theta \neq \theta_0\)

\[
\mathbb{E} \left[ \log \frac{p_y(y|\theta, \lambda=1)}{p_y(y|\theta_0, \lambda=1)} \right] < \log \mathbb{E} \left[ \frac{p_y(y|\theta, \lambda=1)}{p_y(y|\theta_0, \lambda=1)} \right]
\]

The expected values are w.r.t. the data, assuming that the true model exists with parameter \(\theta_0\). Since \(\mathbb{E}[\cdot]\) weighs its operand with \(p_y(y|\theta_0, \lambda=1)\) and integrates out over the data, the expectation in the right-hand side becomes the integral of a valid pdf, which is 1. This proves the claim:

\[
\mathbb{E} \left[ -\log p_y(y|\theta_0, \lambda=1) \right] < \mathbb{E} \left[ -\log p_y(y|\theta, \lambda=1) \right]
\]

The NLL (2.4) \(L(y|\theta, \lambda=1)\) is a continuous function over a compact parameter set \(\Theta\).

- For the first sum, as in the above-mentioned lemma 2.15, \(y(t)\) is \(\alpha\)-mixing at exponential rate, hence so is \(g(y(t), \theta_{NL})\). The latter has bounded (absolute) moments of all orders (dominated in the tails by a power law in \(u(t)\)) and is therefore cumulant mixing of order \(\infty\) (Doukhan and León, 1989, proposition 2.2). A stable linear system \((H^{-1}(q, \theta_L))\) and a polynomial function \((x^2)\) preserve cumulant mixing of order \(\infty\) (Pintelon and Schoukens, 2001, lemmas 14.5 and 14.9).
The terms in the second sum are a memoryless function of \(y(t)\), and have therefore the same \(\alpha\)-mixing properties as \(y(t)\) \cite[p.69, ii]{FanYao2003}. Since \(g'(y(t), \theta_{NL})\) is nonzero and dominated in the tails by a power law in \(u(t)\), the terms have all (absolute) moments bounded. The terms are therefore cumulant mixing of order \(\infty\) \cite[proposition 2.2]{DoukhanLeon1989}.

The cumulant mixing LLN \cite[section 14.9, point 3]{PintelonSchoukens2001} now applies. This guarantees convergence to the expected NLL, minimized at \(\theta_0\). Therefore, consistency follows and the convergence rate is \(O_p(N^{-1/2})\) in probability.

**Asymptotic normality** requires a CLT (under mixing conditions) applicable to the stochastic terms of the derivative of the NLL (2.4) w.r.t. \(\theta\) evaluated at \(\theta_0\): as shown in section 14.13 of Pintelon and Schoukens (2001),

\[
\hat{\theta} - \theta_0 \to \delta_\theta = -E \left[ \frac{\partial^2 L}{\partial \theta^2} (y|\theta, \lambda=1) \right]^{-1} \frac{\partial L}{\partial \theta} (y|\theta_0, \lambda=1)
\]

(2.38)

in probability. When a general term of the first stochastic sum in (2.4) is derived w.r.t. \(\theta_L\) and \(\theta_{NL}\), and evaluated at \(\theta_0\), one finds

\[
2e(t) \cdot \frac{\partial H^{-1}(q, \theta_L)}{\partial \theta_L} u(t) \quad \text{and} \quad 2e(t) \cdot H^{-1}(q, \theta_L) \frac{\partial g'}{\partial \theta_{NL}} (y(t), \theta_{NL,0})
\]

(2.39)

respectively, expressions which, under Assumption 2.11, have bounded moments of all orders. When a general term of the second stochastic sum in (2.4) is derived w.r.t. \(\theta_{NL}\) and evaluated at \(\theta_0\), one finds:

\[
f'_0(u(t)) \cdot \frac{\partial g'}{\partial \theta_{NL}} (y(t), \theta_{NL,0})
\]

(2.40)

an expression which, under Assumption 2.11, has bounded moments of all orders. The signals left and right of the dot-symbols in (2.39) and (2.40) are jointly cumulant mixing sequences of order \(\infty\) (this follows from multivariable application of lemma 14.5 in Pintelon and Schoukens (2001), of p.69(ii) in Fan and Yao (2003), and of proposition 2.2 in Doukhan and León (1989)). Therefore, the (stationary) signals in (2.39) and (2.40) (i.e. the terms of the derivative of the NLL at \(\theta_0\)) are cumulant mixing sequences of order \(\infty\). Hence, the cumulant mixing CLT of Pintelon and Schoukens (2001) (section 14.10, point 4) applies, and normality of the MLE follows from the reasoning of section 14.13 of Pintelon and Schoukens (2001).

**Asymptotic efficiency:** By equation (14-85) of Pintelon and Schoukens (2001), valid for any MLE with a finite-dimensional parameter vector, the covariance of the (first order) random variable \(\delta_\theta\) (section 14.13 of Pintelon and Schoukens, 2001) added to \(\theta_0\) equals exactly the inverse of the FIM. \(\square\)
2.C Loss of consistency for a non-Gaussian distributed input

The considered effect will be shown by proving that the asymptotic expected cost function is no longer minimal at the exact parameter values \( \theta_0 \). That is (under continuous differentiability of \( V \) in a neighbourhood of \( \theta_0 \)):

\[
\frac{\partial}{\partial \theta} \lim_{N \to \infty} E \left[ \frac{V(y, \theta)}{N} \right] \Bigg|_{\theta_0} = \lim_{N \to \infty} E \left[ \frac{1}{N} \frac{\partial V(y, \theta)}{\partial \theta} \right] \Bigg|_{\theta_0} \neq 0 \quad (2.41)
\]

Consider the Wiener system given by \( H(q, \theta_L) = 1 \) and \( g(y, \theta_{NL}) = ay + by^3 \) with \( a, b \in \mathbb{R}_+^+ \), excited by a (non-Gaussian) input that results in a uniformly distributed output between -1 and 1. Differentiating (2.9) w.r.t. \( a \) yields:

\[
\frac{1}{N} \frac{\partial V}{\partial a} \Bigg|_{\theta_0} = 2 \sum_{t=0}^{N-1} \varepsilon_T(t, \theta_0) \frac{\partial \varepsilon_T(t, \theta_0)}{\partial a} = 2 \frac{g_N(y, \theta_0)}{N} \times (2.42)
\]

Applying the strong LLN, it is found that, as \( N \to \infty \), the derivative converges with probability one to

\[
2\exp(-2E \left[ \log (a + 3by(t)^2) \right]) \times \left( E \left[ ay(t)^2 + by(t)^4 \right] - E \left[ (ay(t) + by(t)^3)^2 \right] \right)
\]

The leading factor can never be zero, and is hence unimportant in this analysis. Working out the mathematical expectations, leads to (for the expression between brackets):

\[
\frac{a}{3} + \frac{b}{5} - \left( \frac{a^2}{3} + \frac{b^2}{7} + \frac{2ab}{5} \right) \frac{1}{\sqrt{3ab}} \tan \sqrt{\frac{3b}{a}} \quad (2.43)
\]

Since, generally speaking, this expression is different from zero, the consistency is lost. However, it is observed that, taking the limit for \( b \to 0 \), the value zero is retrieved. This confirms that, in the linear case, there is no consistency loss.

2.D Calculation of the Jacobian matrix

Here two different cases are considered: either \( \theta_i \) is one of the parameters of the linear part \( (\theta_i \in \theta_L) \), or it is one of the parameters of the nonlinear part \( (\theta_i \in \theta_{NL}) \).
If \( \theta_i \in \theta_L \) (except for transient polynomial coefficients), it can easily be seen that:

\[
J_{ki}(y, \theta) = \partial H^{-1}(z_k, \theta_L) / \partial \theta_i \cdot G_k(y, \theta_{NL}) \cdot g_N(y, \theta_{NL})
\]  

(2.44)

For the transient polynomial coefficients,

\[
J_{ki}(y, \theta) = -H^{-1}(z_k, \theta_L) \cdot \partial T(z_k, \theta_L) / \partial \theta_i \cdot g_N(y, \theta_{NL})
\]  

(2.45)

If \( \theta_i \in \theta_{NL} \), the computation is more complicated: in general,

\[
J_{ki}(y, \theta) = H^{-1}(z_k, \theta_L) \cdot \partial G_k(y, \theta_{NL}) / \partial \theta_i \cdot g_N(y, \theta_{NL})
\]  

(2.46)

(2.47)

In particular, if the linear parameterization (2.24) is used, then

\[
J_{ki}(y, \theta) = H^{-1}(z_k, \theta_L) \cdot DFT_k(f_i(y)) \cdot g_N(y, \theta_{NL})
\]  

(2.48)

In general, for a non-monic transfer function \( H \), similar to Pintelon and Schoukens (2006), the frequency domain NLL, corresponding to (2.4), evaluated at \( \lambda = 1 \),
2.E. Proof of Theorem 2.13 (Fisher information matrix)

is given by:

\[ \frac{N}{2} \log 2\pi + \frac{1}{2} \sum_{k=0}^{N-1} \left| H^{-1}(z_k, \theta_L) G_k(y, \theta_{NL}) \right|^2 - \sum_{t=0}^{N-1} \log |g'(y(t), \theta_{NL})| + \sum_{k=0}^{N-1} \log |H(z_k, \theta_L)| \]  

(2.50)

The arguments are now omitted for convenience:

\[ \frac{N}{2} \log 2\pi + \frac{1}{2} \sum_{k=0}^{N-1} \left| H^{-1}G \right|^2 - \sum_{t=0}^{N-1} \log |g'| + \sum_{k=0}^{N-1} \log |H| \]  

(2.51)

Using the following formulas, valid for any complex function \( z \), of a real vectorial argument \( \theta \), \( z : \mathbb{R}^{n_{\theta}} \to \mathbb{C} : \theta \to z(z(\theta)) \):

\[ \frac{\partial^2 \log |z|}{\partial \theta_i \partial \theta_j} = \text{Re}(z^{-1} \frac{\partial^2 z}{\partial \theta_i \partial \theta_j} - z^{-2} \frac{\partial z}{\partial \theta_i} \frac{\partial z}{\partial \theta_j}) \]  

(2.52)

\[ \frac{\partial^2 |z|^2}{\partial \theta_i \partial \theta_j} = 2\text{Re}(\frac{\partial z}{\partial \theta_i} \frac{\partial z}{\partial \theta_j} + \overline{z} \frac{\partial^2 z}{\partial \theta_i \partial \theta_j}) \text{ (Hint: use } |z|^2 = \overline{z}z), \]  

(2.53)

and the fact that \( \frac{\partial^2 g'}{\partial \theta_i \partial \theta_j} = \frac{\partial f'(g(t))}{\partial \theta_j} = 0 \) for parameterization (2.24) of the nonlinearity, we find:

\[ \frac{\partial^2 L(y \mid \theta, \lambda=1)}{\partial \theta_i \partial \theta_j} = \sum_{k=0}^{N-1} \text{Re}(\alpha_k + \beta_k + \gamma_k) + \sum_{t=0}^{N-1} g'^{-2} \frac{\partial g'}{\partial \theta_i} \frac{\partial g'}{\partial \theta_j} \]  

(2.54)

where \( \alpha_k, \beta_k \) and \( \gamma_k \) are defined as:

\[ \alpha_k = \partial H^{-1}G / \partial \theta_i \cdot \partial H^{-1}G / \partial \theta_j \]  

(2.55)

\[ \beta_k = H^{-1}G \cdot \partial^2 H^{-1}G / \partial \theta_i \partial \theta_j \]  

(2.56)

\[ \gamma_k = H^{-1} \cdot \partial^2 H / \partial \theta_i \partial \theta_j - H^{-2} \cdot \partial H / \partial \theta_i \cdot \partial H / \partial \theta_j \]  

(2.57)

\( \alpha_k + \beta_k \) corresponds to the second term in equation (2.51), while \( \gamma_k \) corresponds to the last term in equation (2.51). Using the following equations:

\[ \partial H^{-1}G / \partial \theta_j = -H^{-2} \cdot \partial H / \partial \theta_j \cdot G + H^{-1} \cdot \partial G / \partial \theta_j \]  

(2.58)
Chapter 2. Blind identification of Wiener systems

\[
\frac{\partial^2 H^{-1}G}{\partial \theta_i \partial \theta_j} = 2H^{-3}\frac{\partial H}{\partial \theta_i} \frac{\partial H}{\partial \theta_j} G - H^{-2} \frac{\partial^2 H}{\partial \theta_i \partial \theta_j} G - H^{-2} \frac{\partial H}{\partial \theta_i} \frac{\partial G}{\partial \theta_j} \\
- H^{-2} \frac{\partial H}{\partial \theta_i} \frac{\partial G}{\partial \theta_j} + H^{-1} \frac{\partial^2 G}{\partial \theta_i \partial \theta_j}
\]  (2.59)

yields:

\[
\alpha_k = |E|^2 |H|^{-2} \frac{\partial H}{\partial \theta_i} \frac{\partial H}{\partial \theta_j} - E |H|^{-2} \frac{\partial H}{\partial \theta_i} \frac{\partial G}{\partial \theta_j} \\
- E |H|^{-2} \frac{\partial G}{\partial \theta_i} \frac{\partial H}{\partial \theta_j} + |H|^{-2} \frac{\partial G}{\partial \theta_i} \frac{\partial G}{\partial \theta_j}
\]  (2.60)

\[
\beta_k = 2|E|^2 H^{-2} \frac{\partial H}{\partial \theta_i} \frac{\partial H}{\partial \theta_j} - |E|^2 H^{-1} \frac{\partial^2 H}{\partial \theta_i \partial \theta_j} \\
- EH^{-2} \frac{\partial H}{\partial \theta_i} \frac{\partial G}{\partial \theta_j} - EH^{-2} \frac{\partial H}{\partial \theta_i} \frac{\partial G}{\partial \theta_j}
\]  (2.61)

with \(\frac{\partial^2 G}{\partial \theta_i \partial \theta_j} = 0\), due to the linear parameterization of the nonlinearity (2.24). Special care should be taken when considering the expectations; notice that \(H^{-1}(z_k, \theta_L)\) and its derivatives are deterministic, while \(G(y, \theta_{NL})\) and \(\frac{\partial G}{\partial \theta_j}\) are random. In addition, \(\mathbb{E}[|E|^2] = \lambda = 1\), and \(\mathbb{E}[\text{Re}(\bullet)] = \text{Re}(\mathbb{E}[\bullet])\) are used.

\[
\mathbb{E}[\alpha_k] = |H|^{-2} \frac{\partial H}{\partial \theta_i} \frac{\partial H}{\partial \theta_j} + \mathbb{E} \left[ -\mathbb{E} |H|^{-2} \frac{\partial H}{\partial \theta_i} \frac{\partial G}{\partial \theta_j} \right] \\
- \mathbb{E} |H|^{-2} \frac{\partial G}{\partial \theta_i} \frac{\partial H}{\partial \theta_j} + |H|^{-2} \frac{\partial G}{\partial \theta_i} \frac{\partial G}{\partial \theta_j}
\]  (2.62)

\[
\mathbb{E}[\beta_k + \gamma_k] = H^{-2} \frac{\partial H}{\partial \theta_i} \frac{\partial H}{\partial \theta_j} + \mathbb{E} \left[ -\mathbb{E} H^{-2} \frac{\partial H}{\partial \theta_j} \frac{\partial G}{\partial \theta_i} - \mathbb{E} H^{-2} \frac{\partial H}{\partial \theta_j} \frac{\partial G}{\partial \theta_i} \right]
\]  (2.63)

Evaluating the above expressions in the different cases, yields:

\[
\theta_i, \theta_j \in \theta_L:\n \]

\[
F_{i,j} = \sum_{k=0}^{N-1} \text{Re}(|H|^{-2} \frac{\partial H}{\partial \theta_i} \frac{\partial H}{\partial \theta_j} + H^{-2} \frac{\partial H}{\partial \theta_i} \frac{\partial H}{\partial \theta_j})
\]  (2.64)
2. F. Practical calculation of the Fisher information matrix

θ_i \in \theta_L, \theta_j \in \theta_{NL}:

\begin{equation}
F_{i,j} = -\mathbb{E} \left[ \sum_{k=0}^{N-1} \text{Re}(\mathbb{E} |H|^2 \frac{\partial H}{\partial \theta_i} \frac{\partial G}{\partial \theta_j} + \mathbb{E} H^2 \frac{\partial H}{\partial \theta_i} \frac{\partial G}{\partial \theta_j}) \right] \quad (2.66)
\end{equation}

θ_i, \theta_j \in \theta_{NL}:

\begin{equation}
F_{i,j} = \mathbb{E} \left[ \sum_{k=0}^{N-1} \text{Re}(|H|^2 \frac{\partial G}{\partial \theta_i} \frac{\partial G}{\partial \theta_j}) \right] + \sum_{t=0}^{N-1} \mathbb{E} \left[ g^{-2} \frac{\partial g'}{\partial \theta_i} \frac{\partial g'}{\partial \theta_j} \right] \quad (2.67)
\end{equation}

Applying: (i) the property \( \forall a, b, c \in \mathbb{C} : \text{Re}(c(|a|^2b + a^2b)) = 2\text{Re}(ab)\text{Re}(ac) \) to (2.65) and (2.66), and (ii) the stationarity of \( y(t) \) (as follows from section 2.9 in Brillinger, 1981), which implies that the last expected value in (2.67) is \( t \)-independent, finally proves the theorem.

2. F. Practical calculation of the Fisher information matrix

Two main difficulties arise in these calculations: (i) calculation of the expected values of the nonlinear random expressions

\begin{equation}
\mathbb{E} \left[ G_k \frac{\partial G_k}{\partial \theta_j} \right], \mathbb{E} \left[ \frac{\partial G_k}{\partial \theta_i} \frac{\partial G_k}{\partial \theta_j} \right], \text{ and } \mathbb{E} \left[ g^{-2} \frac{\partial g'}{\partial \theta_i} \frac{\partial g'}{\partial \theta_j} \right] \quad (2.68)
\end{equation}

(ii) for LTI’s \( H \) with zeroes close to the unit circle, special care has to be taken to guarantee positive definiteness of the Fisher matrix. This is essentially due to the transient terms (leakage) when going from the time domain to the frequency domain.

Solution to these problems: (i) The elements of the Fisher matrix contain expected values of sums over \( k \) (see Theorem 2.12). Since the span of dependence of the frequency domain terms (in \( k \)) is sufficiently small (also known as mixing condition), a version of the LLN for the sums over \( k \) applies (see Pintelon and Schoukens, 2001, page 452, item 3). Therefore, the (partial) sum and its expected value are very close; hence a very good approximation of the sought result is found by simply omitting the expectation operator. Further explanations are provided in Appendix 2.G. The time domain term in (2.68) is computed by evaluating the following integral numerically:

\begin{equation}
\int_{-\infty}^{\infty} g^{-2} \frac{\partial g'}{\partial \theta_i} \frac{\partial g'}{\partial \theta_j} p_Y(y) dy \quad (2.69)
\end{equation}
with \( p_Y(y) \) the probability density function of \( y \). Since \( y(t) \) is a stationary process, it can be computed as:

\[
p_Y(y) = \frac{1}{\sqrt{2\pi\sigma_U^2}} \exp \left( -\frac{g(y)^2}{2\sigma_U^2} \right) g'(y)
\]

(2.70)

with \( \sigma_U^2 = \frac{1}{N} \sum_{k=0}^{N-1} |H(z_k)|^2 \) (consequence of Parseval’s theorem applied to \( u \)). In practice, the integral limits \( \pm\infty \) are replaced by the \( y \) values that correspond to \( \pm 5\sigma_U \). (ii) The second problem is due to the transient and leakage effects. These, in turn, are due to zeroes of \( H \) that are very close to the unit circle (this means that the transients of \( H^{-1} \) are not well damped out after \( N \) samples). The effect pops up in the expressions \( \frac{\partial \hat{G}_k}{\partial \theta_j} \) and \( \frac{\partial \hat{G}_k}{\partial \theta_i} \frac{\partial \hat{G}_k}{\partial \theta_j} \). This issue was solved by replacing these expressions by, respectively \( \frac{8}{3} \hat{G}_k \frac{\partial \hat{G}_k}{\partial \theta_j} \) and \( \frac{8}{3} \frac{\partial \hat{G}_k}{\partial \theta_i} \frac{\partial \hat{G}_k}{\partial \theta_j} \), where \( \hat{G}_k \) is the DFT of the Hanning windowed signal. The correction factor appears because, asymptotically the frequency lines become independent (Pintelon and Schoukens, 2001), and the Hanning window combines three neighbouring frequency lines with weights \( 1/4, 1/2, 1/4 \) (Harris, 1978), resulting in a variance scaled by \( (1/4)^2+(1/2)^2+(1/4)^2=3/8 \). However, in some cases, the Hanning window cannot suppress the leakage effects sufficiently, and the issue still remains. In which case, since the effect disappears with growing amounts of data, the following is suggested: (i) generate a larger amount of data \( y_{N_{CR}} \) (\( N_{CR} \) data points) using the estimated model structure, (ii) compute the CRB with that data set, and multiply with \( N_{CR}/N \), to correct for the different number of data points. Here the inverse proportionality of the Cramér-Rao covariance with the number of points \( N \) is exploited (Pintelon and Schoukens, 2001).

### 2.G Law of large numbers applied to the computation of the Fisher information matrix

For convenience, the version of the LLN used, it is recalled over here: If \( x_N(k), N = 1, 2, \ldots, \infty, \) is mixing of order 2 and depends on \( N \), with partial sum \( S(N) = \sum_{k=1}^{N} x_N(k) \), then the LLN

\[
\text{l.i.m.}_{N \to \infty} (S(N) - \mathbb{E}[S(N)]) / N = 0
\]

(2.71)

applies, and the convergence rate is given by

\[
S(N)/N = \mathbb{E}[S(N)] / N + \mathcal{O}_{\text{m.s.}}(N^{-1/2})
\]

(2.72)

(mean square convergence, implying convergence in probability) (Pintelon and Schoukens, 2001).
2.G. Law of large numbers applied to the computation of the Fisher information matrix

We claim that the result of the theorem above is applicable to the random frequency domain sums (containing \( G_k \) or its derivatives) in (2.30):

\[
\text{Re}(H_k^{-1} \frac{\partial H_k}{\partial \theta_i}) \text{Re}(|H_k|^{-2} \overline{G_k} \frac{\partial G_k}{\partial \theta_j}) \text{ or } \text{Re}(|H_k|^{-2} \overline{G_k} \frac{\partial G_k}{\partial \theta_i} \frac{\partial G_k}{\partial \theta_j}).
\] (2.73)

First notice that if a signal \( x(t) \) is (cumulant) mixing, the signal weighted by absolutely bounded nonrandom numbers \( \alpha(t) \) is also mixing of the same order. Due to Assumption 2.1(2), it holds that \( H_k, \) its derivatives, and its inverse \( H_k^{-1} \) are absolutely bounded. Hence, it suffices to check whether (2.37) holds \((p = 2)\) with \( x_N(k) \) equal to:

\[
\text{Re}(G_k \frac{\partial G_k}{\partial \theta_j}) \text{ or } \text{Re}(\overline{G_k} \frac{\partial G_k}{\partial \theta_i} \frac{\partial G_k}{\partial \theta_j}).
\] (2.74)

The computations are shown for the first case, as the second case yields similar results. First, using the definition (2.14), combined with (2.24), we obtain:

\[
\overline{G_k} \frac{\partial G_k}{\partial \theta_j} = N^{-1} \sum_{n_1, n_2 = 0}^{N-1} g(y(n_1), \theta_{NL}) f_j(y(n_2)) z_{n_1 - n_2}^{n_1 - n_2}
\] (2.75)

Applying the multilinearity to the cumulant in (2.37), yields:

\[
\text{cum} \left( \text{Re} \left( \overline{G_k} \frac{\partial G_k}{\partial \theta_j} \right), \text{Re} \left( \overline{G_l} \frac{\partial G_l}{\partial \theta_j} \right) \right) = N^{-2} \times \sum_{n_1, n_2, n_3, n_4 = 0}^{N-1} \cos \left( \frac{2\pi}{N} n_1 - n_2 \right) \cos \left( \frac{2\pi}{N} n_3 - n_4 \right) \times
\]

\[
\text{cum} \left( g(y(n_1), \theta_{NL}) f_j(y(n_2)), g(y(n_3), \theta_{NL}) f_j(y(n_4)) \right)
\] (2.76)

The absolute value can be bounded by:

\[
N^{-2} \sum_{n_1, n_2, n_3, n_4 = 0}^{N-1} |\text{cum} \left( g(y(n_1), \theta_{NL}) f_j(y(n_2)), g(y(n_3), \theta_{NL}) f_j(y(n_4)) \right)|
\] (2.77)

Now if the \( f_j \) are assumed to be polynomial functions of the cumulant mixing (of order \( \infty \)) process \( y(t) \) (Lemma 2.15 of Appendix 2.B), theorems 2.3.2 and 2.9.1 in Brillinger (1981), about cumulants of products and mixing of Volterra series, can be applied to compute the last cumulant and to show that the four-fold sum of these cumulants is an \( \mathcal{O}(N) \). If we now consider the left hand side
of equation (2.37) \((p = 2)\), it follows that it can be bounded by:

\[
\max_l \lim_{N \to \infty} N^{-2} \sum_{k=0}^{N-1} O(N) < c_1 < \infty
\]  

(2.78)

with \(c_1 \in \mathbb{R}_0^+\). This shows that the conditions for applying the LLN are fulfilled. □

References


2.G. Law of large numbers applied to the computation of the Fisher information matrix


Chapter 2. Blind identification of Wiener systems


Chapter 3

Blind maximum-likelihood identification of Hammerstein systems

Similar to Wiener systems, this chapter is about the blind identification of discrete-time Hammerstein systems. Assuming that the unobserved input is white Gaussian noise, that the static nonlinearity is invertible, and that the output is observed without errors, a Gaussian maximum-likelihood estimator is constructed. Its asymptotic properties are analyzed and the Cramér-Rao lower bound is calculated. In practice, the latter can be computed accurately without using the strong law of large numbers. A simulation example illustrating the method is included to support the theory.

Figure 3.1: Block diagram of a Hammerstein system. A static nonlinearity $f_0$ and a linear dynamic block $H_0$ are cascaded.
### 3.1 Chapter outline and contributions

The Hammerstein model structure (Figure 3.1) and the blind identification problem were already introduced in the introduction to part I of the thesis. Since the Wiener and the Hammerstein structure are similar (only the order of both blocks is different), and the natural assumptions introduced in the Wiener case (see previous chapter) are reused, this chapter’s material and outline is very close to the previous one. It was attempted to avoid repetitions, while keeping the chapter easily readable on its own.

The main contributions of this chapter are the following:

- presentation of the Gaussian MLE for Hammerstein systems;
- calculation of the Cramér-Rao lower bound (Fisher information matrix);
- verification of the theoretical results by means of a simulation example.

### 3.2 Assumptions and choices

The assumptions and choices are very close to those of the Wiener case. For the sake of brevity, for certain similar technical details, we refer to the chapter on Wiener systems.

#### 3.2.1 Class of discrete-time Hammerstein systems considered

Hammerstein systems are defined as a cascade of a static (memoryless) non-linearity and a linear time-invariant dynamic system (LTI) $H_0$ (Fig. 3.1). In general, the static nonlinearity can be characterized by its mapping function from $e$ (input signal) to $u$ (intermediate signal): $u = f_0(e)$. The LTI system can be characterized by its transfer function $H_0(z)$.

**Assumption 3.1.** (the class of discrete-time Hammerstein systems considered)

1. $u = f_0(e)$ is a monotonic, bijective function and its derivative is bounded and non-zero over $\mathbb{R}$.
2. $H_0(z)$ is a causal, stable and inversely stable monic transfer function.

**Note 3.2.** The bijectivity is needed because we work in a blind framework. Under Assumption 3.1(2), the filter $H_0(z)$ has minimum phase and contains no pure delay. It allows for the computation of (the power spectrum of) the intermediate signal $u$ from (the power spectrum of) the output signal $y$. Combined with Assumption 3.1(1), the input signal $e$ can be recovered from the observed output $y$.\
3.2. Assumptions and choices

Figure 3.2: An unloaded resistor-diode network as an example of a static nonlinearity. $e$ is the voltage applied at the left terminals. $u$ is the voltage across the diode. The static nonlinear relationship between both can be modelled by a parameter vector $\theta_{NL} = [\alpha \quad \beta \quad \gamma]^T \in \mathbb{R}^3$.

3.2.2 Parameterization

The parameter vectors characterizing the linear and the nonlinear part are respectively denoted by $\theta_L$ and $\theta_{NL}$. The nonlinear function is parameterized inversely: $e = g(u, \theta_{NL})$. The choice of this function could be based on physical knowledge: e.g. an electrical resistor-diode circuit (as represented in Fig. 3.2), can be modelled by an equation of the form:

$$e = g(u, [\alpha \quad \beta \quad \gamma]^T) = \alpha u + \beta (\exp(\gamma u) - 1) \quad (3.1)$$

On the other hand, the LTI is parameterized in the numerator and denominator coefficients of its transfer function $H(z, \theta_L)$:

$$H(z, \theta_L) = \frac{C(z, \theta_L)}{D(z, \theta_L)} = \frac{1 + \sum_{r=1}^{n_c} c_r z^{-r}}{1 + \sum_{r=1}^{n_d} d_r z^{-r}} \quad (3.2)$$

with $\theta_L^T = [c_1, c_2, \ldots, c_{n_c}, d_1, d_2, \ldots, d_{n_d}]$.

Assumption 3.3. (nonlinear parameterization)

(1) $\theta_{NL}$ gives a unique representation of the nonlinearity, viz. $g(u, \theta_{NL}) = g(u, \tilde{\theta}_{NL}), \forall u \Leftrightarrow \theta_{NL} = \tilde{\theta}_{NL}$.

(2) $g$ is twice continuously differentiable w.r.t. both function arguments.

Note that the inverse parameterization enables a direct inversion (or equalization) of the Hammerstein system by means of a Wiener system formed by $H^{-1}$ in series with $g$. (This could be useful for, e.g., calibration applications.) Finally, we define a global parameter vector by stacking the parameters of both subblocks onto each other:

$$\theta^T = [\theta_L^T, \theta_{NL}^T] \quad (3.3)$$
3.2.3 Stochastic framework

Assumption 3.4. (stochastic framework)

1. The unknown, unobserved input \( e(t) \) is zero mean, white Gaussian noise with unknown variance \( \lambda_0 > 0 \).

2. The output \( y(t) \) is known exactly (i.e. observed without errors).

Note 3.5. In Assumption 3.4(1), note that the identification process does not need any knowledge about the variance \( \lambda_0 \) of the input. A value of the input variance, compatible with the estimated Hammerstein model, is estimated from the data during the identification process. Assumption 3.4(2) means that the output measurement Signal-to-Noise Ratio (SNR) equals infinity in theory, or, is “sufficiently high” in practice. The impact of noise on the estimator will be treated in the next chapter.

3.3 The Gaussian maximum-likelihood estimator

3.3.1 The negative log-likelihood function

Theorem 3.6. Under Assumptions 3.1 and 3.4, the conditional Gaussian negative log-likelihood function of the observations

\[
y_T^N = y^T = [y(0), \ldots, y(N-1)]
\]

given the model parameters \( \theta \) and input variance \( \lambda \), is:

\[
L(y|\theta,\lambda) = \frac{N}{2} \log 2\pi + \frac{N}{2} \log \lambda + \frac{N-1}{2} \sum_{t=0}^{N-1} \left( \frac{g(H^{-1}(q,\theta_L)y(t),\theta_{NL})}{2\lambda} \right)^2
\]

\[
- \sum_{t=0}^{N-1} \log |g'(H^{-1}(q,\theta_L)y(t),\theta_{NL})|
\]

with \( q \) the forward shift operator, with \((...)')'\ denoting the first order partial derivative w.r.t. the first argument of the function, viz. \( g'(x_1,x_2) = \frac{\partial g}{\partial x_1}(x_1,x_2) \).

Here, “conditional” means “given the initial conditions of the LTI part". However, asymptotically \((N \to \infty)\), the conditional MLE equals the true MLE.

Proof. Follow the same lines as in Appendix 2.A on page 63 from Chapter 2. The proof is based on the classical expressions of the log-likelihood for ARMA models and on the transformation formula of probability density functions through nonlinear mappings. \(\square\)
3.3. The Gaussian maximum-likelihood estimator

As in the Wiener case, it is possible to eliminate the input variance $\lambda$ analytically from the log-likelihood cost function, by setting $\partial L(y|\theta,\lambda)/\partial \lambda = 0$, and solving for $\lambda$. This yields the following result:

$$
\lambda = \frac{1}{N} \sum_{t=0}^{N-1} \left( g(H^{-1}(q, \theta_L)y(t), \theta_{NL}) \right)^2
$$

(3.6)

and the likelihood-based cost function boils down to the sum-of-squares expression

$$
V(y, \theta) = g_N^2(y, \theta) \sum_{t=0}^{N-1} \left( g(H^{-1}(q, \theta_L)y(t), \theta_{NL}) \right)^2
$$

(3.7)

with

$$
g_N(y, \theta) = \exp \left( -\frac{1}{N} \sum_{t=0}^{N-1} \log |g'(H^{-1}(q, \theta_L)y(t), \theta_{NL})| \right)
$$

(3.8)

Stated differently, the cost function $V(y, \theta)$ can be rewritten as a sum of squares of residuals $\varepsilon(t, \theta)$:

$$
\varepsilon(t, \theta) = g(H^{-1}(q, \theta_L)y(t), \theta_{NL}) g_N(y, \theta)
$$

(3.9)

The sum-of-squares property is useful, since it allows to use efficient Gauss-Newton based routines for the optimization.

At this point, it is also interesting to note that the first two differences arise in contrast to the Wiener case: first, the factor $g_N$ no longer only depends on the data and the parameter estimates of the nonlinear part; also $\theta_L$ pops up in the equations, and $g_N$ becomes a function of the full parameter vector $\theta$. This results in slightly more complicated expressions for the Jacobian matrix. Second, in contrast to the Wiener case, we prefer to stay in the time domain to perform the calculations, since it is the natural domain in which the reconstructed input signal $g(H^{-1}(q, \theta_L)y(t), \theta_{NL})$ is found (a static nonlinearity is most easy to calculate in the time domain).

### 3.3.2 Identifiability

As it is the case for the blind identification of Wiener systems, the actual parameterization of the Hammerstein is not identifiable as such. It can be seen that e.g. the displacement of a gain factor from the input variance to the linear or nonlinear block would result in the same output signal (similar to Figure 2.2 on page 44). This motivates the need of introducing a set of identifiability conditions.
Chapter 3. Blind identification of Hammerstein systems

Assumption 3.7. (identifiability conditions)

1. $H_0(z, \theta_L)$ has no common pole-zero pairs.
2. $H_0(z, \theta_L) = \frac{C_0(z, \theta_L)}{D_0(z, \theta_L)}$ is monic (this means that the constants in the $C_0$ and $D_0$ polynomials are 1).
3. $g'_0(u_1, \theta_{NL}) = 1$ for some $u_1 \in \mathbb{R}$.

Note 3.8. The above assumptions intend to normalize the scaling factors associated with both model sub-blocks, such that $\lambda$ can be identified and the aforementioned scaling ambiguity has been solved. $\lambda$ then remains the only degree of freedom of the model to control the output signal amplitude.

3.3.3 Result

As a consequence of the preceding paragraphs, the following holds:

Result: Under Assumptions 3.1 and 3.4, the maximum likelihood estimator $\hat{\theta}$ of the system parameters minimizes the cost function (3.7), subject to the constraints given in Assumption 3.7.

Since the cost function has been written as a sum-of-squares, its minimizer

$$\hat{\theta} = \text{argmin}_{\theta} V(y, \theta)$$

(3.10)

can be calculated in a numerically stable way via the classical Gauss-Newton based iterative schemes (Fletcher, 1991). The most likely input variance $\hat{\lambda}$ corresponding to this parameter set $\hat{\theta}$ can be found by evaluating (3.6).

3.3.4 Asymptotic properties

As in the previous chapter on the Wiener case (see section 2.5.5 on page 46), it is also not obvious that the MLE enjoys the usual asymptotic ($N \to \infty$) properties in the Hammerstein case, since the classical i.i.d. noise assumption (Caines, 1988, p. 317, theorem 3.2) does not apply here. Therefore, the asymptotic properties of the estimator have to be studied explicitly. Firstly, this requires standard assumptions about the true model, the parametric model and the excitation.

Assumption 3.9. (consistency)

1. The true model is within the proposed model set, that contains models that satisfy Assumptions 3.1, 3.3, 3.4 and 3.7.
2. The normalized cost function $V(y_N, \theta)/N$ has continuous second order derivatives w.r.t. $\theta$ in a compact (i.e. closed and bounded) set $\Theta_r$ for any $N$, infinity included. The compact set $\Theta_r$ is constructed such that it contains a...
unique global minimum of \( V(y_N, \theta)/N \), which is an interior point (i.e. not on the boundaries) of \( \Theta_r \).

(3) There exists an \( N_0 \) such that for any \( N \geq N_0 \), infinity included, the Hessian of the expected value of the normalized cost function \( V(\theta) = E[V(y_N, \theta)/N] \), subject to the constraints given in Assumption 3.7, is regular (i.e. invertible) at its unique global minimizer in \( \Theta_r \).

**Assumption 3.10.** (nonlinear rates)

1. \( f_0(e) = O(|e|^{r_1}) \) with \( r_1 \in \mathbb{R}^+_0 \) as \( |e| \to \infty \).
2. \( f'_0(e) = O(|e|^{r_2}) \) with \( r_2 \in \mathbb{R}^+_0 \) as \( |e| \to \infty \).
3. \( g(u, \theta_{NL}) = O(|u|^{r_3}) \) with \( r_3 \in \mathbb{R}^+_0 \) as \( |u| \to \infty \).
4. \( g'(u, \theta_{NL}), g''(u, \theta_{NL,0}), \frac{\partial g}{\partial \theta_{NL}}(u, \theta_{NL,0}) \) and \( \frac{\partial g'}{\partial \theta_{NL}}(u, \theta_{NL,0}) \) satisfy a similar power-law rate in \( u \). Herein \( \theta_{NL,0} \) refers to the exact parameter values corresponding to \( \theta_{NL} \).

**Theorem 3.11.** Under Assumptions 3.1, 3.3, 3.4, 3.7, 3.9 and 3.10, the MLE \( \hat{\theta} \) is consistent, has convergence rate \( O(N^{-1/2}) \) in probability, is asymptotically efficient, and asymptotically normally distributed.

**Proof.** See Appendix 3.A on page 91.

### 3.4 Generation of initial estimates

As already mentioned, an iterative algorithm can be used to find the MLE. However, the generation of initial estimates is still necessary. This is done in two steps: in a first step, the LTI part \( H \) is estimated. This allows one to calculate an estimate of the intermediate signal \( u(t) \) (Fig. 3.1), from which – in a second step – a nonparametric estimate of the nonlinearity \( f \) is generated via the increasing property of the nonlinear function and the input Gaussianity. The procedure was explained in full detail in chapter 2, at section 2.6 on page 47. It can be noted that Assumption 3.10(1) guarantees that the moments of the intermediate signal exist, and therefore allows the use of ARMA modelling in the first step.

### 3.5 Minimization of the cost function

Gauss-Newton based iterative algorithms for optimizing sum-of-squares cost functions require use of the Jacobian matrix, defined as the partial derivatives of the residuals \( \varepsilon(t, \theta) \) (3.9) w.r.t. the parameters: \( J_t(y, \theta) = \partial \varepsilon(t, \theta)/\partial \theta_t \). Its calculation is straightforward, and, hence, left out of this chapter. In the general case, the practical implementation requires the calculation of all mixed
second order derivatives $\partial^2 g(u, \theta_{NL}) / (\partial u \partial \theta_i)$, which can be quite a tedious task. Therefore, we propose a simple parameterization for the nonlinearity.

3.5.1 Parameterization of the nonlinearity

As in the Wiener case, to simplify the computations, a linear-in-the-parameters representation is used to describe the nonlinearity $g(y)$:

$$ e = g(u, \theta_{NL}) = f(u)^T \theta_{NL} $$

(3.11)

with $\theta_{NL}^T = [\theta_{NL,1}, \ldots, \theta_{NL,M}]$ the parameter vector of the nonlinear block, and with $f(u)^T = [f_1(u), \ldots, f_M(u)]$ the vector of basis functions. Moreover, this linear parameterization simplifies, considerably, the weighted least squares computation in the initial estimates algorithm; the problem is linear-in-the-parameters, and an iterative procedure is avoided.

The same remarks apply for the identifiability and the options for the selection of the basis $f(u)$ (see Section 2.7.1 on page 51).

3.5.2 Gauss-Newton procedure

The equations of $J_{ti}$ make it possible to compute the Jacobian matrix from the data and the parameters, and to minimize the cost function $V(y, \theta)$. This minimization is performed by using a Levenberg-Marquardt algorithm. It yields the most likely parameter vector $\hat{\theta}$, given the output observations $y$, if the global minimum is found.

3.5.3 Practical implementation of the constraint

In practice, as in the Wiener case, for the model (3.11), the constraint on the nonlinearity is not implemented as stated in Assumption 3.7(3), but by setting the two-norm of the nonlinear parameter vector in the linear parameterization to unity:

$$ \sum_m \theta_{NL,m}^2 = \|\theta_{NL}\|_2^2 = 1 $$

(3.12)

This alternative constraint has the same properties as described in chapter 2.

3.6 The Cramér-Rao lower bound

3.6.1 Theoretical expressions

Since it is known from subsection 3.3.4 that the estimator is asymptotically $(N \rightarrow \infty)$ efficient – which means the asymptotic covariance matrix equals the
3.6. The Cramér-Rao lower bound

Cramér-Rao lower bound (CRB) – its calculation allows us to compute the uncertainties on the estimated model parameters.

Exactly as in the Wiener case, the overparameterization issue can be dealt with by rescaling the nonlinear parameters in such a way that the new input variance becomes unity, and leaving out the parameter $\lambda$ from the Fisher Information Matrix (FIM) calculations (cf. full motivation in chapter 2, section 2.9 on page 53): in practice, $\hat{\theta}_{NL}$ is replaced by $\hat{\theta}_{NL}/\sqrt{\lambda}$ before calculating the FIM with the expressions in the theorem given below.

**Theorem 3.12.** For the Gaussian MLE given in (3.10), and under the Assumptions 3.1, 3.3, 3.4, 3.7, 3.9 and 3.10, and for polynomial nonlinearities of the form (3.11), asymptotically, the elements of the FIM for unit input variance are given by:

(a) for $\theta_i \in \theta_L, \theta_j \in \theta_L$,

$$F_{i,j} = \mathbb{E} \left[ (g')^2 + \frac{(g'')^2}{g'} \right] \times$$

$$\left( N\mu_U H_{0}^{-2} \frac{\partial H_0}{\partial \theta_i} \frac{\partial H_0}{\partial \theta_j} + 2\sigma_U^2 \sum_{k=0}^{N-1} \text{Re} \left( H_k^{-1} \frac{\partial H_k}{\partial \theta_i} \right) \text{Re} \left( H_k^{-1} \frac{\partial H_k}{\partial \theta_j} \right) \right)$$

(b) for $\theta_i \in \theta_L, \theta_j \in \theta_{NL}$,

$$F_{i,j} = -N\mu_U H_{0}^{-1} \frac{\partial H_0}{\partial \theta_i} \cdot \mathbb{E} \left[ f_j g' + f_j' g'' / (g')^2 \right]$$

(c) for $\theta_i \in \theta_{NL}, \theta_j \in \theta_{NL}$,

$$F_{i,j} = N \cdot \mathbb{E} \left[ f_i f_j + f_i' f_j' / (g')^2 \right]$$

Each of these symbols is a compact notation where, for notational simplicity, the dependencies on the parameters have been omitted. In these expressions, $\mu_U = \mathbb{E}[U], \sigma_U^2 = \mathbb{E}[U - \mu_U]^2, H_k = H(\exp(j2\pi k/N), \theta_L)$, and $U$ is the random variable corresponding to the stationary signal $u(t)$.

**Proof.** See Appendix 3.B on page 93.

**Note 3.13.** First, it is interesting to note that – in contrast to the Wiener case – the obtained CRB matrix can be calculated without having recourse to the strong law of large numbers to approximate the mathematical expectations. Since the expectation operators do not contain frequency domain expressions like $H$, no leakage effects can occur either. Second, it is seen that, in the case of an antisymmetric nonlinearity (viz. $f(-e) = -f(e)$), $\mu_U$ is zero, and, therefore, the statistical coupling between the linear and nonlinear parts disappears.
3.6.2 Practical calculation

In practice, the mathematical expectations are calculated by numerical integration of the considered expressions weighted by the pdf of $U$ denoted by $p_U(u)$. The integration limits are chosen as the values of $u$ which correspond with $e = \pm 5\sqrt{\hat{\lambda}}$. The pdf $p_U(u)$ is given by:

$$p_U(u) = p_E(g(u, \hat{\theta}_{NL})) \left| g'(u, \hat{\theta}_{NL}) \right|$$ \hspace{1cm} (3.13)

with

$$p_E(e) = \frac{1}{\sqrt{2\pi\hat{\lambda}}} \exp\left( -\frac{e^2}{2\hat{\lambda}} \right)$$ \hspace{1cm} (3.14)

The values of $\mu_U$ and $\sigma^2_U$ are calculated similarly by numerical integration.

3.7 Simulation results

3.7.1 Setup: presentation of the example

In this section, the results of the method applied to a simulation example are discussed. The underlying Hammerstein system used in the simulations was given by a fifth-order polynomial inverse nonlinearity with powers of $u$ as basis functions $f_m(u)$:

$$g_0(u) = \theta_{NL,1} + \theta_{NL,2}u + \theta_{NL,3}u^3 + \theta_{NL,4}u^5 = 2 + u + 4u^3 + 2u^5$$ \hspace{1cm} (3.15)

and a fourth order filter with transfer function

$$H_0(z^{-1}) = \frac{1 + 0z^{-1} - 0.49z^{-2} + 0.01z^{-3} + 0.065z^{-4}}{1 + 0.3676z^{-1} + 0.88746z^{-2} + 0.52406z^{-3} + 0.55497z^{-4}}$$ \hspace{1cm} (3.16)

Subsection 3.7.2 presents the results of Monte-Carlo simulations in absence of output noise (Assumption 3.4(2)). The behaviour of the estimator in the presence of noise is examined in the next chapter.

3.7.2 Results and validation

A 500-run Monte-Carlo simulation with $N = 8192$ points each was performed, and the resulting parameter estimates were analyzed. It was first noticed that, as expected, the MLE improves (i.e. lowers) the variances of the estimated

\footnote{Here, the symbols $U$ and $E$ refer respectively to the random variables associated with the stationary signals $u(t)$ and $e(t)$.}
3.7. Simulation results

parameters, compared to the initial estimates. Also some statistical hypothesis tests (at significance level $\alpha = 5\%$) were performed on the parameters estimated from the Monte-Carlo data:

- Kolmogorov-Smirnov: could not detect significant deviations from normality (12 out of 12 parameters).
- Deviation of the mean from the true value: for 11 out of the 12 parameters, no significant deviations could be detected.
- Sample variance vs. Cramér-Rao lower bound: could not detect significant (variance) deviations on any of the 12 components of the parameter vector. The preceding facts are shown graphically for the parameter $\theta_{NL, 3}$ in Figure 3.3.

The other parameters give similar results. For this simulation example, the maximal ratio between the standard deviations of initial and final estimates is 1.86. This illustrates the high quality of the initial estimates.

Finally, the cross-validation of the results is performed (after the MLE optimization) by calculation, for a randomly selected realization, of:

- the pdf of the intermediate signal $p_U(u)$ based on the estimated nonlinear parameters $\hat{\theta}_{NL}$ and the estimated input variance $\hat{\lambda}$, as given in (3.13)-(3.14),
and, the normalized histogram of the data sequence $\hat{u}$ obtained by inverse filtering of the output with the estimated parameters of the linear part $\hat{\theta}_L$.

Both are shown in Figure 3.4, from which the excellent agreement between both can be seen.

This cross-validation provides a practical way to the user to check whether the Hammerstein model with white Gaussian input is appropriate to model the signal provided to the algorithm.

### 3.8 Conclusions

Following the lines of the Wiener case, a maximum-likelihood procedure for a blind identification of Hammerstein systems has been handled. As in the Wiener case, the cost function can be rewritten as a sum of squares. A two-step algorithm for generating high-quality initial estimates has been presented. Furthermore, expressions of the Fisher information matrix have been provided, allowing one to generate uncertainty bounds for the parameters or any parameter-related quantity. The simulation experiments illustrate the theoretical results of the estimator.
3.A. Asymptotic properties

Appendices

3.A Asymptotic properties

Note. Although the tools for the analysis and proofs are equal to the Wiener case, the reasoning itself has to be repeated, since the way in which they are combined is somewhat different.

Lemma 3.14. Under the consistency assumptions of Theorem 3.11, \( y(t) \) is \( \alpha \)-mixing at exponential rate and cumulant mixing of order \( \infty \).

Proof. From the input’s i.i.d. property, its Gaussianity and the smoothness of the static nonlinear function (Assumption 3.1(1)), it follows that the process \( u(t) \) is a continuous i.i.d. random variable and has a pdf. From Assumption 3.1(2), the impulse response of the LTI decays at exponential rate. Therefore, the process \( y(t) \) is \( \beta \)-mixing at exponential rate (see Fan and Yao, 2003, p.69, iii), and hence \( \alpha \)-mixing at the same rate (see Bradley, 2005, eq. (1.11)). Assumption 3.10(1) implies that \( u(t) \) has bounded (absolute) moments of all orders; for, by continuity, the function \( f_0(u) \) is bounded, the integral contribution of the \( p \)th \((p \in \mathbb{N}_0)\) order moment is finite over a finite interval, and the contribution in the tails are bounded by the moments of a Gaussian. Thanks to the stability of the LTI, the output process \( y(t) \) has also bounded (absolute) moments of all orders. Hence, \( y(t) \) is cumulant mixing of order \( \infty \) (see Doukhan and León, 1989, proposition 2.2).

For the proof of Theorem 3.11, we follow the lines of section 14.13 of Pintelon and Schoukens (2001) applied to the NLL (3.5), and fill in the ‘gaps’ – i.e. it only remains to guarantee the applicability of the Law of Large Numbers (LLN) and Central Limit Theorem (CLT) – for the actual model structure used. To avoid the model ambiguity problem in this reasoning, \( \lambda \) is fixed to unity (see also Assumptions 3.9(2) and 3.9(3)), allowing to treat \( \theta \) as an unconstrained parameter vector.

Note. The reader might wonder why \( \alpha \)-mixing (“strong mixing”) is used here in addition to cumulant mixing. The reason is that if the function \( g(\bullet, \theta_{NL}) \) is polynomial, its inverse (being \( f_0(\bullet) \) at \( \theta_{NL,0} \)) cannot be polynomial. Cumulant mixing has properties for polynomial static nonlinearities, while \( \alpha \)-mixing is preserved over arbitrary static nonlinearities.

Consistency: As for any MLE with a finite-dimensional parameter-vector, and under Assumption 3.9(1), the expected NLL is minimal at the exact parameter values \( \theta_0 \) (Kendall and Stuart, 1979, section 18.10); see also Appendix 2.B on page 64 for a short proof. The NLL is a continuous function over a compact parameter set.
• For the first sum in (3.5), knowing that stable linear systems and polynomial nonlinearities preserve cumulant mixing of order $\infty$ (Pintelon and Schoukens, 2001, lemmas 14.5 and 14.9), it follows that the expression $g(H^{-1}(q, \theta_L)y(t), \theta_{NL})^2$ is cumulant mixing of order $\infty$.

• The terms in the second sum of (3.5) are a memoryless function of $H^{-1}(q, \theta_L)y(t)$, which is $\alpha$-mixing at exponential rate. Since $g'(\bullet, \theta_{NL})$ is nonzero and dominated in the tails by a power law in its first argument (denoted by $\bullet$), the terms have all (absolute) moments bounded. The terms are therefore both $\alpha$-mixing at exponential rate and cumulant mixing of order $\infty$ (see Doukhan and León, 1989, proposition 2.2).

The cumulant mixing LLN (Pintelon and Schoukens, 2001, section 14.9, point 3) now applies. This guarantees convergence to the expected NLL, minimized at $\theta_0$. Therefore, consistency follows and the convergence rate is $O(N^{-1/2})$ in probability.

*Asymptotic normality* requires a CLT (under mixing conditions) applicable to the stochastic terms of the derivative of the NLL (3.5) w.r.t. $\theta$ evaluated at $\theta_0$: (see (2.38), or section 14.13 of Pintelon and Schoukens, 2001). When a general term of the first stochastic sum in (3.5) is derived w.r.t. $\theta_L$ and $\theta_{NL}$, and evaluated at $\theta_0$, one finds

$$2e(t) \cdot g'(u(t), \theta_{NL,0}) \cdot \frac{\partial H^{-1}(q, \theta_L)}{\partial \theta_L} y(t) \quad \text{and} \quad 2e(t) \cdot \frac{\partial g}{\partial \theta_{NL}}(u(t), \theta_{NL,0})$$

(3.17)

respectively, expressions which, under Assumption 3.10, have bounded moments of all orders. When a general term of the second stochastic sum in (3.5) is derived w.r.t. $\theta_L$ and $\theta_{NL}$ and evaluated at $\theta_0$, one finds

$$f'(e(t)) \cdot g''(u(t), \theta_{NL,0}) \cdot \frac{\partial H^{-1}(q, \theta_L)}{\partial \theta_L} y(t) \quad \text{and} \quad f'_0(e(t)) \cdot \frac{\partial g'}{\partial \theta_{NL}}(u(t), \theta_{NL,0})$$

(3.18)

respectively, both of which, under Assumption 3.10, have bounded moments of all orders. The signals constituting the factors of (3.17) and (3.18) are jointly cumulant mixing sequences of order $\infty$ (this follows from multivariable application of lemma 14.5 in Pintelon and Schoukens (2001), of p.69(ii) in Fan and Yao (2003), and of proposition 2.2 in Doukhan and León (1989)). Therefore, the (stationary) signals in (3.17) and (3.18) (i.e. the terms of the derivative of the NLL at $\theta_0$) are cumulant mixing sequences of order $\infty$. Hence, the cumulant mixing CLT of Pintelon and Schoukens (2001) (section 14.10, point 4) applies, and normality of the MLE follows from the reasoning of section 14.13 of Pintelon and Schoukens (2001).
Asymptotic efficiency: By equation (14-85) of Pintelon and Schoukens (2001), valid for any MLE with a finite-dimensional parameter vector, the covariance of the (first order) random variable \( \delta_\theta \) (section 14.13 of Pintelon and Schoukens, 2001) added to \( \theta_0 \) equals exactly the inverse of the FIM.

3.B Fisher information matrix

The FIM is defined as the mathematical expectation of the Hessian of the negative log-likelihood function w.r.t. the parameters, evaluated at the exact parameter values. First \( \lambda \) is set to unity in (3.5), and the second derivative is calculated. It is easily found that

\[
\frac{\partial^2 L(y|\theta, 1)}{\partial \theta_i \partial \theta_j} = \sum_{t=0}^{N-1} T_t
\]

with

\[
T_t = \frac{\partial g_t}{\partial \theta_i} \frac{\partial g_t}{\partial \theta_j} + g_t \frac{\partial^2 g_t}{\partial \theta_i \partial \theta_j} - g_t^{-1} \frac{\partial^2 g_t'}{\partial \theta_i \partial \theta_j} + g_t^{-2} \frac{\partial g_t}{\partial \theta_i} \frac{\partial g_t}{\partial \theta_j}
\]

and with \( g_t = g(H^{-1}(q, \theta_L)y(t), \theta_{NL}) \).

From this equation and parameterization (3.11), the FIM in case (c) follows immediately. Cases (a) and (b) are found after some very lengthy – but not so complicated – calculations, which are the contents of the remaining part of this Appendix. We start with a lemma which will be useful during the calculations.

**Lemma 3.15.** For any smooth increasing functions \( g(u) \) and \( h(u) \), satisfying a power-law rate of increase similar to Assumption 3.10,

\[
E[h(u)g(u)] = E\left[h'(u)\frac{g(u)}{g'(u)}\right]
\]

for \( U = f(E) \) with \( E \) standard normally distributed, viz.

\[
p_E(e) = (2\pi)^{-1/2} \exp\left(-\frac{e^2}{2}\right)
\]

Proof. The left hand side is given by the following integral (the integration interval \((-\infty, \infty)\) is dropped for notational conciseness):

\[
\int h(u)g(u)p_U(u)du = \int h(u)g(u)g'(u)p_E(g(u))du = (2\pi)^{-1/2} \int h(u)g(u)g'(u)\exp\left(-\frac{g(u)^2}{2}\right)du
\]
Integration by parts yields

\[
= -(2\pi)^{-1/2} h(u) \exp \frac{-g(u)^2}{2} \bigg|_{-\infty}^{\infty} + (2\pi)^{-1/2} \int h'(u) g'(u) \exp \frac{-g(u)^2}{2} du
\]  

(3.25)

Since the first term equals zero, it is easily seen that the right hand side \( E \left[ \frac{h'(u)}{g'(u)} \right] \) has been found. Notice that the division by \( g'(u) \) is allowed since it differs from zero.

\subsection*{3.B.1 Case (a): \( \theta_i \in \theta_L, \theta_j \in \theta_L \)}

\subsubsection*{3.B.1.1 Preliminary calculations}

Here expressions for \( g = g_t = g(H^{-1}(q, \theta_L)y(t), \theta_{NL}) \) and its derivatives are calculated. Since we want to derive w.r.t. \( \theta_L \), we may assume that \( \theta_{NL} \) equals its true value, but we need to express explicitly the quantities as function of the transfer function \( H \) and the observations:

\[
g_t = g(u_t, \theta_{NL}) = g \left( \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} H_k^{-1} Y_k \exp \frac{j2\pi kt}{N}, \theta_{NL} \right) \]  

(3.26)

Herein we used the well-known asymptotic input/output relationship for linear systems, combined with the IDFT (Inverse Discrete Fourier Transformation) definition:

\[
u_t = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} H_k^{-1} Y_k \exp \frac{j2\pi kt}{N} \]  

(3.27)

with \( H_k = H(\exp j2\pi k/N, \theta_L) \), and \( Y_k = \frac{1}{\sqrt{N}} \sum_{t=0}^{N-1} y_t \exp -j2\pi kt \).

In the following, the arguments will often be omitted for notational conciseness. (Let \( \xi^{(r)} \) denote the \( r \)th partial derivative of the function \( \xi(x, y, \ldots) \) w.r.t. its first argument; \( \xi^{(r)} = \frac{\partial^r \xi(x, y, \ldots)}{\partial x^r} \).) Whenever omitted, the summation indices run from 0 to \( N - 1 \).

\[
\frac{\partial g_t^{(r)}}{\partial \theta_i} = \frac{\partial g_t^{(r)}}{\partial u_t} \frac{\partial u_t}{\partial \theta_i} = g_t^{(r+1)} \left( \frac{1}{\sqrt{N}} \sum_k H_k^{-2} \frac{\partial H_k}{\partial \theta_i} Y_k \exp \frac{j2\pi kt}{N} \right) \]  

(3.28)
\[ \frac{\partial^2 g_t^{(r)}}{\partial \theta_i \partial \theta_j} = \frac{\partial}{\partial \theta_j} \frac{\partial g_t^{(r)}}{\partial \theta_i} \tag{3.29} \]

\[ = \frac{g_t^{(r+1)}}{\partial \theta_j} \left( -\frac{1}{\sqrt{N}} \sum_k H_k^{-2} \partial H_k \frac{\partial H_l}{\partial \theta_i} Y_k \exp \frac{j2\pi kt}{N} \right) \tag{3.30} \]

\[ + g_t^{(r+1)} \left( -\frac{1}{\sqrt{N}} \sum_k \frac{\partial}{\partial \theta_j} \left( H_k^{-2} \frac{\partial H_k}{\partial \theta_i} \right) Y_k \exp \frac{j2\pi kt}{N} \right) \]

\[ = \frac{g_t^{(r+2)}}{N} \sum_{k,l} H_k^{-2} H_l^{-2} \frac{\partial H_k}{\partial \theta_i} \frac{\partial H_l}{\partial \theta_j} Y_k Y_l \exp \frac{j2\pi (k - l) t}{N} \tag{3.31} \]

\[ + \frac{g_t^{(r+1)}}{\sqrt{N}} \sum_k H_k^{-2} \left( 2H_k^{-1} \frac{\partial H_k}{\partial \theta_i} \frac{\partial H_k}{\partial \theta_j} - \frac{\partial^2 H_k}{\partial \theta_i \partial \theta_j} \right) Y_k \exp \frac{j2\pi kt}{N} \]

Introducing

\[ \alpha_t = \frac{1}{N} \sum_{k,l} H_k^{-2} H_l^{-2} \frac{\partial H_k}{\partial \theta_i} \frac{\partial H_l}{\partial \theta_j} Y_k Y_l \exp \frac{j2\pi (k - l) t}{N} \tag{3.32} \]

and

\[ \beta_t = \frac{1}{\sqrt{N}} \sum_k H_k^{-2} \left( 2H_k^{-1} \frac{\partial H_k}{\partial \theta_i} \frac{\partial H_k}{\partial \theta_j} - \frac{\partial^2 H_k}{\partial \theta_i \partial \theta_j} \right) Y_k \exp \frac{j2\pi kt}{N} \tag{3.33} \]

we find the more compact expression:

\[ \frac{\partial^2 g_t^{(r)}}{\partial \theta_i \partial \theta_j} = g_t^{(r+2)} \alpha_t + g_t^{(r+1)} \beta_t \tag{3.34} \]

### 3.B.1.2 Calculation of the summand \( T_t \) of equation (3.20)

\[ \frac{\partial g_t}{\partial \theta_i} \frac{\partial g_t}{\partial \theta_j} = g_t \frac{1}{N} \sum_{k,l} H_k^{-2} H_l^{-2} \frac{\partial H_k}{\partial \theta_i} \frac{\partial H_l}{\partial \theta_j} Y_k Y_l \exp \frac{j2\pi (k - l) t}{N} \tag{3.35} \]

\[ = g_t^2 \alpha_t \tag{3.36} \]

\[ g_t \frac{\partial^2 g_t}{\partial \theta_i \partial \theta_j} = g_t g_t' \alpha_t + g_t g_t' \beta_t \tag{3.37} \]

\[ - g_t^{-1} \frac{\partial^2 g_t}{\partial \theta_i \partial \theta_j} = -g_t^{-1} g_t^{(3)} \alpha_t - g_t^{-1} g_t'' \beta_t \tag{3.38} \]
\[ g_t^{-2} \frac{\partial^2 g_t'}{\partial \theta_i \partial \theta_j} = g_t^{-2} g_t'' \times \frac{1}{N} \sum_{k,l} H_k^{-2} \overline{H_l}^{-2} \frac{\partial H_k}{\partial \theta_i} \frac{\partial \overline{H_l}}{\partial \theta_j} Y_k \overline{Y_l} \exp \left( \frac{j2\pi (k-l) t}{N} \right) \] (3.39)
\[ = g_t^{-2} g_t'' \alpha_t \] (3.40)

Putting all this together, and rewriting, yields:
\[ T_t = \eta_t \alpha_t + \zeta_t \beta_t \] (3.41)

with
\[ \eta_t = g_t'^2 + g_t g_t'' - g_t'^{-1} g_t^{(3)} + g_t^{-2} g_t''^2 \] (3.42)

and
\[ \zeta_t = g_t g_t' - g_t'^{-1} g_t'' \] (3.43)

### 3.B.1.3 Actual calculation of the expected values and summations

Notice that no more derivatives need to be calculated. Therefore, at this point, the parameters may be equated to their true values. Since the statistical properties of \( u_t \) (this signal is identically, independently and equally distributed) are a lot easier than those of \( Y_k \), using the DFT, we go back to expressions in terms of \( u_t \). So, \( \mathbb{E} \left[ \sum_t T_t \right] \) is written out and rearranged by substitution of
\[ Y_k = H_k \frac{1}{\sqrt{N}} \sum_p u_p \exp \left( -\frac{j2\pi kp}{N} \right) \] (3.44)

in \( \alpha_t \) and \( \beta_t \):
\[ \mathbb{E} \left[ \sum_t T_t \right] = \frac{1}{N^2} \sum_{k,l} H_k^{-1} \overline{H_l}^{-1} \frac{\partial H_k}{\partial \theta_i} \frac{\partial \overline{H_l}}{\partial \theta_j} A_{kl} \] (3.45)
\[ + \frac{1}{N} \sum_k I_k \sum_{t,p} \exp \left( \frac{j2\pi k (t-p)}{N} \right) \mathbb{E} \left[ \zeta_t u_p \right] \]

with
\[ A_{kl} = \sum_t \exp \left( \frac{j2\pi (k-l) t}{N} \right) \sum_{p,q} \exp \left( \frac{j2\pi (lq - kp)}{N} \right) \mathbb{E} \left[ \eta_t u_p u_q \right] \] (3.46)
\[ I_k = H_k^{-1} \left( 2H_k^{-1} \frac{\partial H_k}{\partial \theta_i} \frac{\partial H_k}{\partial \theta_j} - \frac{\partial^2 H_k}{\partial \theta_i \partial \theta_j} \right) \] (3.47)
Since $\eta_t$ and $\zeta_t$ can be seen as functions of $u_t$, and all $u_t$ are independent random variables, we can distinguish the following cases:

$$
\mathbb{E}[\eta_t u_p u_q] = \begin{cases} 
\mathbb{E}[\eta_t] \mathbb{E}[u_p] \mathbb{E}[u_q] = \mathbb{E}[\eta] \mu_U^2 & t \neq p \neq q \\
\mathbb{E}[\eta_t] \mathbb{E}[u_p^2] = \mathbb{E}[\eta] (\mu_U^2 + \sigma_U^2) & t \neq p = q \\
\mathbb{E}[\eta_t u_t] \mathbb{E}[u_q] = \mathbb{E}[\eta u] \mu_U & t = p \neq q \\
\mathbb{E}[\eta_t u_t] \mathbb{E}[u_p] = \mathbb{E}[\eta u] \mu_U & t = q \neq p \\
\mathbb{E}[\eta^2] & t = p = q 
\end{cases} \tag{3.48}
$$

with $\mu_U = \mathbb{E}[U]$ and $\sigma_U^2 = \mathbb{E}[(U - \mu_U)^2]$. Notice that those moments are time independent as follows from the i.i.d. property of $U$.

Thanks to Lemma 3.15, it can now be seen that $\mathbb{E}[\zeta] = 0$, hence, the second term of (3.45) can be written as:

$$
\frac{1}{N} \sum_k I_k \sum_{t,p} \exp \left( \frac{j2\pi k(t-p)}{N} \right) \mathbb{E}[\zeta_t u_t] \delta_{t-p} = \sum_k I_k \mathbb{E}[\zeta u] \tag{3.50}
$$

with the $\delta$ represents a Kronecker delta:

$$
\delta_r = \begin{cases} 
0 & r \neq 0 \\
1 & r = 0 
\end{cases} \tag{3.51}
$$

Using ‘note b’ of the errata of Pintelon and Schoukens (2001), about “sums over the unit circle of functions that are analytic outside the unit circle and zero at infinity”, $\sum_k I_k$ asymptotically vanishes to zero as an $O(N|\lambda_{\text{max}}|^N)$ with $\lambda_{\text{max}} = \max(|z_1|, |z_2|, \ldots, |z_{nc}|, |p_1|, \ldots, |p_{nd}|)$ and with $z_r$ and $p_r$ the zeros and poles of $H$ respectively.

Therefore, only the first term of (3.45) survives. Now we focus on the part indicated with the brace below in $A_{kl}$ (3.46), and write out expressions for it in the cases of equation (3.48):

$$
\sum_{p,q} \exp \left( \frac{j2\pi (lq - kp)}{N} \right) \mathbb{E}[\eta_t u_p u_q] = \begin{cases} 
s_1 & t \neq p \neq q \\
s_2 & t \neq p = q \\
s_3 & t = p \neq q \\
s_4 & t = q \neq p \\
s_5 & t = p = q 
\end{cases} \tag{3.52}
$$
Chapter 3. Blind identification of Hammerstein systems

with

\[ s_1 = \mathbb{E} [\eta] \mu_U^2 \sum_{p \neq t} \exp \left( \frac{-j2\pi kp}{N} \right) \sum_{q \neq p \neq t} \exp \left( \frac{j2\pi lq}{N} \right) \] (3.53)

\[ = \mathbb{E} [\eta] \mu_U^2 \sum_{p \neq t} \exp \left( \frac{-j2\pi kp}{N} \right) \left( N\delta_l - \exp \left( \frac{j2\pi lt}{N} \right) - \exp \left( \frac{j2\pi lp}{N} \right) \right) \] (3.54)

\[ = \mathbb{E} [\eta] \mu_U^2 \left[ \left( N\delta_l - \exp \left( \frac{j2\pi lt}{N} \right) \right) \left( N\delta_k - \exp \left( \frac{-j2\pi kt}{N} \right) \right) \right. \] (3.55)

\[ \left. - N\delta_{l-k} + \exp \left( \frac{j2\pi (l-k) t}{N} \right) \right] \]

\[ = \mathbb{E} [\eta] \mu_U^2 \left[ N^2 \delta_k \delta_l - N\delta_l \exp \left( \frac{-j2\pi kt}{N} \right) - N\delta_k \exp \left( \frac{j2\pi lt}{N} \right) \right. \] (3.56)

\[ + 2 \exp \left( \frac{j2\pi (l-k) t}{N} \right) - N\delta_{l-k} \left. \right] \]

\[ s_2 = \mathbb{E} [\eta] \left( \mu_U^2 + \sigma_U^2 \right) \sum_{p \neq t} \exp \left( \frac{j2\pi (l-k) p}{N} \right) \] (3.57)

\[ = \mathbb{E} [\eta] \left( \mu_U^2 + \sigma_U^2 \right) \left( N\delta_{l-k} - \exp \left( \frac{j2\pi (l-k) t}{N} \right) \right) \] (3.58)

\[ s_3 = \mathbb{E} [\eta u] \mu_U \exp \left( \frac{-j2\pi kt}{N} \right) \sum_{p \neq t} \exp \left( \frac{j2\pi lp}{N} \right) \] (3.59)

\[ = \mathbb{E} [\eta u] \mu_U \exp \left( \frac{-j2\pi kt}{N} \right) \left( N\delta_l - \exp \left( \frac{j2\pi lt}{N} \right) \right) \] (3.60)

\[ s_4 = \mathbb{E} [\eta u] \mu_U \exp \left( \frac{j2\pi lt}{N} \right) \sum_{p \neq t} \exp \left( \frac{-j2\pi kp}{N} \right) \] (3.61)

\[ = \mathbb{E} [\eta u] \mu_U \exp \left( \frac{j2\pi lt}{N} \right) \left( N\delta_k - \exp \left( \frac{-j2\pi kt}{N} \right) \right) \] (3.62)

\[ s_5 = \mathbb{E} [\eta u^2] \exp \left( \frac{j2\pi (l-k) t}{N} \right) \] (3.63)
Now the expression $A_{kl}$ (3.46) is computed:

$$
\sum_t \exp \frac{j2\pi (k - l) t}{N} \sum_{p,q} \exp \frac{j2\pi (lq - kp)}{N} \mathbb{E} [\eta_t u_p u_q] = \begin{cases} 
S_1 & t \neq p \neq q \\
S_2 & t \neq p = q \\
S_3 & t = p \neq q \\
S_4 & t = q \neq p \\
S_5 & t = p = q 
\end{cases}
$$

(3.64)

with

$$S_1 = \mathbb{E} [\eta] \mu_U^2 \sum_t \left( N^2 \delta_k \delta_l - N \delta_l \exp \frac{-j2\pi lt}{N} \right)$$

$$-N \delta_k \exp \frac{j2\pi kt}{N} + 2 - N \delta_{l-k} \right)$$

$$= \mathbb{E} [\eta] \mu_U^2 \left( N^3 \delta_k \delta_l - N^2 \delta_l - N^2 \delta_k + 2N - N^2 \delta_{l-k} \right)$$

(3.65)

$$S_2 = \mathbb{E} [\eta] \left( \mu_U^2 + \sigma_U^2 \right) \sum_t \left( N \delta_{l-k} - 1 \right)$$

$$= \mathbb{E} [\eta] \left( \mu_U^2 + \sigma_U^2 \right) \left( N^2 \delta_{l-k} - N \right)$$

(3.66)

$$S_3 = \mathbb{E} [\eta \mu_U] \sum_t \exp \frac{-j2\pi lt}{N} \left( N \delta_l - \exp \frac{j2\pi lt}{N} \right)$$

$$= \mathbb{E} [\eta \mu_U] \left( N^2 \delta_l - N \right)$$

(3.67)

$$S_4 = \mathbb{E} [\eta \mu_U] \sum_t \exp \frac{j2\pi kt}{N} \left( N \delta_k - \exp \frac{-j2\pi kt}{N} \right)$$

$$= \mathbb{E} [\eta \mu_U] \left( N^2 \delta_k - N \right)$$

(3.68)

$$S_5 = \mathbb{E} [\eta^2] \left( N \delta_l \delta_k \right) \left( \mu_U^2 + \sigma_U^2 \right)$$

(3.69)

All this together gives:

$$\mathbb{E} [\eta] \mu_U^2 \left( N^3 \delta_k \delta_l - N^2 \delta_l - N^2 \delta_k + N \right) + \mathbb{E} [\eta] \sigma_U^2 \left( N^2 \delta_{l-k} - N \right)$$

$$+ \mathbb{E} [\eta u] \mu_U \left( N^2 \delta_l + N^2 \delta_k - 2N \right) + \mathbb{E} [\eta^2] \left( N \right)$$

(3.70)
Finally, depending on the \(\delta\)'s, this leads to double sums (no \(\delta\)), single sums (\(\delta_{l-k}, \delta_k\) and \(\delta_l\)) and a single term (\(\delta_k \delta_l\)) in (3.45). Notice that all the double sums are proportional to \(\left(\sum_k H_k^{-1} \frac{\partial H_k}{\partial \theta_i}\right) \left(\sum_l H_l^{-1} \frac{\partial H_l}{\partial \theta_j}\right)\), and that it concerns sums over the unit circle, similarly to what was done above. Hence these terms will vanish to zero as \(O\left(N |\lambda_{max}|^2 N\right)\). Likewise, the terms proportional to \(\delta_k\) or \(\delta_l\) vanish as \(O\left(N |\lambda_{max}| N\right)\). Therefore, the remaining terms are given by the reals:

\[
E[\eta] \mu_U^2 \sum_k \text{Re} \left( H_k^{-1} \frac{\partial H_k}{\partial \theta_i} \right) \text{Re} \left( H_k^{-1} \frac{\partial H_k}{\partial \theta_j} \right)
\]

and a vanishing term

\[
- E[\eta] \sigma_U^2 \text{Re} \left( \sum_k H_k^{-2} \frac{\partial H_k}{\partial \theta_i} \frac{\partial H_k}{\partial \theta_j} \right) = O\left(N |\lambda_{max}|^N\right)
\]  

(3.76)

The last step is to apply once more Lemma 3.15. This yields \(\mathbb{E}[\eta] = \mathbb{E}[g^2 + g''^2]\), and it is seen that the \((\theta_L, \theta_L)\) components of the Fisher matrix have been found as given in Theorem 3.12, part (a).

### 3.B.2 Case (b): \(\theta_i \in \theta_L, \theta_j \in \theta_{NL}\)

#### 3.B.2.1 Preliminary calculations of the derivatives

\[
T_t = \frac{\partial g_t}{\partial \theta_i} f_{j,t} + g_t \frac{\partial f_{j,t}}{\partial \theta_i} - g_t^{-1} \frac{\partial f^{*,t}_{j,t}}{\partial \theta_i} + g_t^{-2} \frac{\partial g_t}{\partial \theta_i} f_{j,t}
\]

(3.77)

with \(f_{j,t}\) the basis function corresponding to the coefficient \(\theta_j\) evaluated at time instant \(t\). Similarly as for case (a), we find

\[
\frac{\partial f^{(r)}_{j,t}}{\partial \theta_i} = f^{(r+1)}_{j,t} \left(-\frac{1}{\sqrt{N}} \sum_k H_k^{-2} \frac{\partial H_k}{\partial \theta_i} Y_k \exp \frac{j2\pi kt}{N}\right)
\]

(3.78)

and (3.28) remains valid for \(\frac{\partial g_t^{(r)}}{\partial \theta_i}\).

#### 3.B.2.2 Calculation of the summand \(T_t\) of equation (3.20)

Now, (3.77) becomes:

\[
T_t = \psi_t \frac{1}{\sqrt{N}} \sum_k H_k^{-2} \frac{\partial H_k}{\partial \theta_i} Y_k \exp \frac{j2\pi kt}{N}
\]

(3.79)
with

$$\psi_t = -g'_tf_{j,t} - g_tf'_{j,t} + g'^{-1}f''_{j,t} - g'^{-2}g''f'_{j,t}$$

(3.80)

3.B.2.3 Actual calculation of the expected values and summations

Again, we use (3.44), and rearranging, this leads to:

$$\mathbb{E} \left[ \sum T_t \right] = \frac{1}{N} \sum_k H_k^{-1} \frac{\partial H_k}{\partial \theta_i} \left( \sum_t \exp \frac{j2\pi kt}{N} \sum_p \exp \frac{-j2\pi kp}{N} \mathbb{E} [\psi_t u_p] \right)$$

(3.81)

Using

$$\mathbb{E} [\psi_t u_p] = \begin{cases} \mathbb{E} [\psi_t] \mathbb{E} [u_p] = \mathbb{E} [\psi] \mu_U & t \neq p \\ \mathbb{E} [\psi u] & t = p \end{cases}$$

(3.82)

the expression indicated with a brace below becomes:

$$\mathbb{E} [\psi] \mu_U \left( N\delta_k - \exp \frac{-j2\pi kt}{N} \right) + \mathbb{E} [\psi u] \exp \frac{-j2\pi kt}{N}$$

(3.83)

and the expression indicated with a brace above becomes:

$$\mathbb{E} [\psi] \mu_U \left( N^2\delta_k - N \right) + \mathbb{E} [\psi u] N$$

(3.84)

Finally, we find:

$$\mathbb{E} \left[ \sum T_t \right] = N\mathbb{E} [\psi] \mu_U H_0^{-1} \frac{\partial H_0}{\partial \theta_i} + O \left( N |\lambda_{max}|^N \right)$$

(3.85)

The $O ()$-term comes from the sums over the unit circle, and asymptotically, it vanishes. Observing that, using Lemma 3.15, $\mathbb{E} [\psi] = -\mathbb{E} [g'f + g'^{-2}g''f']$, proves part (b) of the theorem.

References


Chapter 4

Impact of output measurement noise

The estimators discussed in the two previous chapters are based on the assumption that no noise is present in the output measurements. The goal of this short chapter is to study of the effects that can arise if this assumption is violated, i.e. if the methods are simply applied to data corrupted by output measurement noise (which is also much more realistic). The asymptotical bias is derived theoretically for high output Signal-to-Noise-Ratio’s (SNR’s), and shown to be, in a first-order approximation, proportional to the noise variance. Therefore, there exists an output SNR beyond which the bias is insignificant compared to the estimation variability. Hence, the impact of additive output noise is small in high SNR situations. The effect is illustrated with simulation examples. This chapter ends with a discussion on the blind maximum-likelihood problem with output measurement noise.

4.1 Introduction

In the previous chapters, the observations were assumed to be errorless. However, as a consequence of output measurement noise, a bias will pop up in the estimates. In Section 4.2, it will be shown that, asymptotically, this bias has the property to be – in a first order approximation – proportional to the variance of the additional unmodelled noise source. In a practical situation, if one has an idea of the measurement noise level, by using the derived expressions,
Chapter 4. Impact of output measurement noise

one could obtain an order of magnitude for the bias. In Section 4.3, the effect will be illustrated by means of simulation examples. Finally, in Section 4.4, the integration problems involved with the extension of the model structure are reported and an identifiability proof is provided.

4.2 Theoretic analysis of the noise-induced bias

In this analysis, the noise \( n_y \) is assumed to be independent of \( e(t) \), additive, zero mean and white with variance \( \mu \): \( y_m = y + n_y \). Remember that the MLE is given by the minimum of a cost function \( V(y, \theta) \) (subject to the appropriate identifiability constraints) with a noise-free output \( y \). Minimizing \( V(\theta) = \mathbb{E}[V(y + n_y, \theta)/N] \) instead of \( V(\theta) = \mathbb{E}[V(y, \theta)/N] \), will obviously have an effect of the position of the minimum and hence cause a systematic deviation (or bias). Let \( \theta_0 \) and \( \tilde{\theta} \) be the respective minimizers of \( V(\theta) \) and \( V(\theta) \):

\[
V'(\theta_0) = 0 \quad \text{and} \quad \tilde{V}'(\tilde{\theta}) = 0 \quad \text{with} \quad V'(\theta) = \frac{\partial V(\theta)}{\partial \theta} \quad (4.1)
\]

First, \( V_1(\theta) \) is defined such that:

\[
\tilde{V}(\theta) = V(\theta) + V_1(\theta) \quad (4.2)
\]

Second, \( V'(\theta) \) is expanded linearly about \( \theta_0 \):

\[
V'^T(\theta) \approx V'^T(\theta_0) + V''(\theta_0)(\theta - \theta_0) \quad (4.3)
\]

with \( V''(\theta) = \frac{\partial^2 V(\theta)}{\partial \theta^2} \). Collecting (4.1)-(4.3) yields:

\[
0 = \tilde{V}'(\tilde{\theta}) \approx V''(\theta_0)(\tilde{\theta} - \theta_0) + V'^T_1(\tilde{\theta}) \quad (4.4)
\]

This shows that the bias \( \tilde{\theta} - \theta_0 \) is approximately given by:

\[
\tilde{\theta} - \theta_0 \approx -V''(\theta_0)^{-1}V'^T_1(\tilde{\theta}) \quad (4.5)
\]

The Hessian \( V''(\theta_0) \) does not depend on the noise \( n_y \); therefore, \( V''(\theta_0)^{-1} = \mathcal{O}(\mu^0) \). However, \( V'^T_1(\tilde{\theta}) \) depends on the noise \( n_y \):

\[
V'^T_1(\theta) = \tilde{V}'(\tilde{\theta}) - V'^T(\theta) = \mathbb{E}[V'^T(y + n_y, \theta)/N - V'^T(y, \theta)/N] \quad (4.6)
\]

Making a quadratic expansion around \( y \) yields (The notation is symbolic, since \( V'^T_1(\theta) \) is a vector and its Hessian w.r.t. \( y \) is not a matrix. Nevertheless, we may reason as follows on each component of \( V'^T_1(\theta) \).

\[
V'^T(\theta) \approx \mathbb{E} \left[ \frac{1}{N} \frac{\partial V'^T(y, \theta)}{\partial y} n_y \right] + \mathbb{E} \left[ \frac{1}{2N} n_y^T \frac{\partial^2 V'^T(y, \theta)}{\partial y^2} n_y \right] \quad (4.7)
\]
Now, the independence of $n_y$ and $e(t)$ is used. The second term is the mathematical expectation of a random quadratic form with a matrix with $n_y$-independent entries, and can be handled similarly to Mathai and Provost (1992), p. 50. Therefore,

$$V_1'(\theta) \approx \mathbb{E} \left[ \frac{1}{N} \frac{\partial V'^T(y,\theta)}{\partial y} \right] \mathbb{E} [n_y] + \frac{1}{2} \text{tr} \left( \mathbb{E} [n_y n_y^T] \mathbb{E} \left[ \frac{1}{N} \frac{\partial^2 V'^T(y,\theta)}{\partial y^2} \right] \right)$$

$$= \frac{1}{2} \text{tr} \left( \mu \mathbb{E} \left[ \frac{1}{N} \frac{\partial^2 V'^T(y,\theta)}{\partial y^2} \right] \right) = O(\mu) \quad (4.8)$$

Plugging these results in (4.5), it is found that in general – i.e. if the diagonal elements of $\mathbb{E} \left[ \frac{\partial^2 V'^T(y,\theta)}{\partial y^2} \right]$ do not sum to zero for at least one component of the parameter vector – there is a bias, which behaves as an $O(\mu)$. An estimate of the bias can be found by evaluating:

$$\hat{\theta} - \theta_0 \approx -\frac{\mu}{2} \mathbb{E} \left[ \frac{1}{N} V''(y,\hat{\theta}) \right]^{-1} \text{tr} \left( \mathbb{E} \left[ \frac{1}{N} \frac{\partial^2 V'^T(y,\hat{\theta})}{\partial y^2} \right] \right) \quad (4.9)$$

### 4.3 Simulation results

The bias is illustrated on:

- a Wiener simulation example, focusing on the output simulation error;
- and, a Hammerstein example, focusing on the parametric bias error, and
- the above results are verified numerically.

#### 4.3.1 Output simulation error (Wiener)

The same system as presented in chapter 2 was used (also 8192 points per run) to study the impact of additive output noise on the simulation error of the model, for several output SNR conditions. At each SNR, the system was identified 50 times, and these 50 parameter vectors were averaged. Finally, for a fixed input signal, the root mean squared value (rms) of the difference between the output of the estimated model and the true system’s output was calculated relatively to the rms value of the latter. Figure 4.1 shows that the influence of the output noise on the simulation error is small (order of magnitude of a few percent) in high SNR cases. It is seen that the deviation between the full line and the horizontal dash-dotted line (obtained in a noiseless situation) is insignificant when the SNR exceeds the value of approximately 45 dB. The reason is that the variability of the estimator dominates over the systematic errors (bias) induced by the output noise. Of course, one should still be careful since this numerical value is valid for this particular simulation example.
Chapter 4. Impact of output measurement noise

4.3.2 Parametric bias error (Hammerstein)

This part deals with the analysis of the MLE when the available data sequence is a noise-corrupted version of the true model output. We restrict ourselves to the Hammerstein case. The method is applied as if the data was noiseless. White Gaussian noise $n_y$ was added to the noiseless output $y$: $y_m = y + n_y$. A Monte-Carlo analysis was carried out at 8 different output SNR’s with $R = 100$ runs (of $N = 8192$ points each) per SNR. At each SNR, the Mean Squared Error (MSE) between estimated and true parameter values was calculated as follows:

$$\text{MSE}_{\theta} = \frac{1}{R} \sum_{r=1}^{R} \left\| \hat{\theta}^{[r]} - \theta_0 \right\|^2_2$$  \hspace{1cm} (4.10)

with $\hat{\theta}^{[r]} \in \mathbb{R}^{n_{\theta}}$ the estimated parameter vector at run number $r$, and $\theta_0 \in \mathbb{R}^{n_{\theta}}$ denotes the exact parameter vector. The MSE was divided by $\left\| \theta_0 \right\|^2_2$ to obtain the relative MSE shown in Figure 4.2. From the asymptotic properties in the noiseless case, the MLE is normally distributed with mean $\theta_0$ and covariance CRB $\in \mathbb{R}^{n_{\theta} \times n_{\theta}}$, which can be calculated from the FIM. Let the eigenvalues of the CRB matrix be given by $\lambda_i$ $(i = 1 \ldots n_{\theta})$. Then, provided $R$ is large enough, $\text{MSE}_{\theta}$ (4.10) tends in distribution to a normally distributed variable:

$$\text{MSE}_{\theta} \rightarrow N(\mu_{\text{MSE}}, \sigma_{\text{MSE}}^2)$$  \hspace{1cm} (4.11)

with $\mu_{\text{MSE}} = \sum_{i=1}^{n_{\theta}} \lambda_i$ and $\sigma_{\text{MSE}}^2 = \frac{2}{R} \sum_{i=1}^{n_{\theta}} \lambda_i^2$; result based on (Mathai and Provost, 1992, p. 28, section 3.1a), and $\chi^2_R \rightarrow N(R, 2R)$.

Figure 4.1: Relative simulation error as function of the output SNR.
4.4 Blind maximum-likelihood problem with output measurement noise

It is well known that the MSE contains a variance and a (squared) bias term. In the noiseless case, no bias is present; hence, the MSE should coincide with $\mu_{\text{MSE}}$. From Fig. 4.2, it is also observed that for SNR’s exceeding 40dB, the curve stays in between the confidence bounds. Therefore, no significant bias occurs at these SNR values. It is concluded that the impact of noise is hardly visible at sufficiently high SNR’s.

4.4 Blind maximum-likelihood problem with output measurement noise

4.4.1 Very high-dimensional integrals, related problems and solutions

The previous sections show that the noise can be neglected in high SNR situations. However, the bias may be unacceptably large if this is not the case. Therefore, it may be needed to extend the model in such a way to incorporate the measurement noise source explicitly, e.g. via an additive, zero-mean white, Gaussian noise source (the variance of which has to be estimated as well), denoted $w(t)$, independent of the input and added to the noiseless model output,
denoted $v(t)$. The measurements are then given by $y(t) = v(t) + w(t)$. The nasty consequence is that the likelihood pops up as a multidimensional convolution of probability density functions, which is the result of the summation of independent random variables; since it is a likelihood, the number of integration dimensions is as high as the number of data $N$. Clearly, neither the classical analytical nor numerical integration methods can yield a result which is usable in practice. The author has been working on the analysis and implementation of the multi-integral (and the related optimization scheme). It turns out that, rewritten in an exponential form, the integral can be approximated via the Laplace integration method (de Bruijn, 1970); a quadratic approximant (involving the Hessian) of the exponent’s argument is then constructed at its maximizer and an approximation to the integral results via the multivariate normal density function’s integral. The approximation qualities, such as integration error and bias were under study, while Wills and Ljung (2010) came up with an elegant approach, based on the Expectation-Maximization method and particles, to solve the Wiener identification problem without needing quadratic approximations nor high-dimensional integrals (the approximation made by these authors is related to the finite number of particles to approximate the probability density functions). Besides, the paper by Shun and McCullagh (1995) reports that the unmodified Laplace approximation is reliable provided that the integration dimension, which is here the number of data points $N$, tends to infinity at speed given by the third power of the parameter in the exponent’s argument, which is here $\mu^{-1}$ (where $\mu$ represents the new noise source’s variance). This implies that the estimator obtained via the Laplace approximation would have good properties for $\mu \rightarrow 0$ at fixed $N$, and for $N \rightarrow \infty$ at fixed $\mu$; the combination of both may result in unreliable results, since the speed of convergence then also plays a role.

4.4.2 Identifiability (Wiener)

In the context of blind identification of the Wiener system in a noisy environment, the identifiability of the model parameters may be a relevant issue, which, to the knowledge of the author, has not yet been investigated theoretically. In Appendix 4.A on the facing page, it is shown that, the nonlinear function being a fully parameterized polynomial, all model parameters and the noise variance are jointly locally identifiable with probability 1. The proof is based on cumulants (Leonov and Shiryaev, 1959).
4.5 Conclusion

In this short chapter, it has been shown that the maximum-likelihood estimators for blind identification of Wiener and Hammerstein systems from chapters 2 and 3 are subject to a bias when applied to noisy data. In a theoretical analysis, it was demonstrated that additive measurement noise in the output data introduces a bias, in a first-order approximation, proportional to the variance of that additive, unmodelled noise source. As a consequence, if it turns out that the bias does not exceed the variability level (e.g. twice the square root of diagonal entries of the CRB matrix), it can be concluded that it is insignificant in that situation. The simulations support this result, and show that this bias is insignificant beyond a certain Signal-to-Noise Ratio (40 dB in the example). The blind maximum-likelihood problem in the presence of output noise has been outlined, and the associated multi-integral problem has been discussed. Finally, an identifiability proof has been given for polynomial Wiener systems in the presence of noise.

Appendices

4.A Identifiability proof

If the number of data tends to infinity, the output cumulants (i.e., coefficients of the log-characteristic function, see Pintelon and Schoukens, 2001) of any order may be considered to be known. In the following, it will be shown that $\theta_{NL}$ and the noise variance $\mu$ can both be recovered locally in a unique way by equating the numerical cumulant values to their theoretic model-based expressions for almost any $\varsigma^T = [\theta_{NL}^{T} \mu]$. Once this is done, the distribution of the noiseless output $v(t)$ is known (from the cumulants and $\mu$), from which $\theta_L$ and $\theta_{NL}$ can be identified (see Theorem 2.9 on page 44).

4.A.1 Identifiability condition for cumulant-based estimation of the nonlinear coefficients and the noise variance

Identifiability problems may arise at some $\varsigma_1$ if the Jacobian matrix (with infinitely many rows)
Chapter 4. Impact of output measurement noise

\[ J_\zeta = \frac{\partial}{\partial \zeta} \begin{bmatrix} \text{cum}(y(t)) \\
\text{cum}(y(t), y(t)) \\
\vdots \\
\text{cum}(y(t), \ldots, y(t)) \\
p \\
\vdots \end{bmatrix} \] (4.12)

has linearly dependent columns at \( \zeta_1 \). If we can make sure that the determinant of the top square part \( J_\square \) of \( J_\zeta \) is nonzero, local identifiability is obtained. Using independence of the input and measurement noise sources and Gaussianity of \( w(t) \), and the multilinearity and symmetry of the cumulants in any of its arguments, (4.12) becomes:

\[
\frac{\partial}{\partial \zeta} \text{cum}(y, \ldots, y) = \frac{\partial}{\partial \zeta} \left[ \text{cum}(v, \ldots, v) + \mu \delta_{p-2} \right] \\
= p \text{cum} \left( \frac{\partial}{\partial \zeta}, v, \ldots, v \right) + \frac{\partial \mu}{\partial \zeta} \delta_{p-2}
\]

in which each \( v \) and each \( y \) is evaluated at the same time instant \( t \). From this and the linear parameterization of \( v \) in \( \theta_{NL} \), it follows that \( \det J_\zeta \) results in a polynomial expression \( P(\theta_{NL}) \). Viz. non-identifiability can only happen in this lower-dimensional space of roots of \( P(\theta_{NL}) = 0 \), unless \( P(\theta_{NL}) \) is exactly a zero polynomial, in which case \( \zeta \) would be nowhere identifiable. The rest of this Appendix is devoted to proving that \( P(\theta_{NL}) \) is not the zero polynomial.

4.A.2 Making sure that the polynomial \( P(\theta_{NL}) \) is not identically zero, via the evaluation of the determinant for a cubic function

This is done by evaluation of \( P(\theta_{NL}^\triangle) \) (or the corresponding determinant) with \( \theta_{NL}^\triangle \) the parameter vector corresponding to the cubic nonlinearity \( v(t) = \left( \frac{u(t)}{\text{std}(u(t))} \right)^3 \). Since the \( f_i(u) \) form a basis by assumption, the span of \( \{f_i(u)\}_i \) equals the span of \( \{u^1\}_i \); hence \( J_\zeta \) is row equivalent with the matrix \( J_{\text{eq}} \) with \((p, l + 1)\)th entry given by
\[ \kappa_l^{(p)} = \text{cum}(x^l, x^3, \ldots, x^3) \]  
(4.13)

for \( x \) standard normally distributed, to be extended with a zero column with a 1 on position 2 from the derivative w.r.t. \( \mu \). The cumulant \( \kappa_l^{(p)} \) is zero if its actual degree \( l + 3(p - 1) \) is odd (all odd order central moments vanish for a normal distribution); this happens when the parity of \( l \) and \( p \) is the same.

Writing out a part of \( J_{eq} \) yields a checkerboard pattern:

\[
J_{eq} = \begin{bmatrix}
\kappa_0^{(1)} & 0 & \kappa_2^{(1)} & 0 & \ldots & 0 \\
0 & \kappa_1^{(2)} & 0 & \kappa_3^{(2)} & \ldots & 1 \\
\kappa_0^{(3)} & 0 & \kappa_2^{(3)} & 0 & \ldots & 0 \\
0 & \kappa_1^{(4)} & 0 & \kappa_3^{(4)} & \ldots & 0 \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots 
\end{bmatrix}
\]  
(4.14)

If the parity of \( l \) and \( p \) is different, \( \kappa_l^{(p)} \) can be found by counting of indecomposable partitions in pairs of the table consisting of \( l \) entries on the top row and \( 3 \) entries on each of the \( p - 1 \) other rows (Leonov and Shiryaev, 1959) results in \( (l + 3(p - 1) - 1)!! =
\[
\sum_{r=1}^{p/2} (3p - 6r - 1)!! \left( \begin{array}{c}
p - 1 \\
2r - 1 
\end{array} \right) \kappa_l^{(2r)} \quad \text{odd } l \text{ & even } p
\]
\[
\sum_{r=1}^{(p+1)/2} (3p - 6r + 2)!! \left( \begin{array}{c}
p - 1 \\
2r - 2 
\end{array} \right) \kappa_l^{(2r-1)} \quad \text{even } l \text{ & odd } p
\]  
(4.15)

It can be remarked that \( \kappa_0^{(p)} = \delta_{p-1} \).

### 4.A.3 Proving the full-rank condition

Combining the linear dependency condition on \( J_{eq} \) with (4.15), the full rank condition of \( J_{eq} \) is easily seen to be equivalent to a full rank condition of the matrices \( \tilde{J}^{\text{odd}} \) defined as \([(6i + 2j - 9)!!]_{i,j} \), and \( \tilde{J}^{\text{even}} \) defined as \([(6i + 2j - 5)!!]_{i,j} \), with the column \([(2i - 1) (6i - 7)!!] \), appended on the right, with \( \tilde{J}^{\text{odd}} \) and \( \tilde{J}^{\text{even}} \) corresponding to the odd and even numbered rows of \( J_{eq} \), and with the convention that the indices \( i \) and \( j \) start from 1. \( \tilde{J}^{\text{odd}} \) is seen to have the property
\[
\frac{\tilde{J}_{i,j+1}^{\text{odd}}}{\psi_{j+1}^{\text{odd}}} = \frac{\tilde{J}_{i,j}^{\text{odd}}}{\psi_{j}^{\text{odd}}} (\alpha_i^{\text{odd}} + \beta_j^{\text{odd}})
\]  
(4.16)
with \( \alpha_i^{\text{odd}} = 6i \), \( \beta_{j+1}^{\text{odd}} = 2j - 7 \) and \( \psi_j^{\text{odd}} = 1 \) (used more generally in the sequel). It can be shown with elementary matrix operations that the general \( n \times n \) matrix possessing this property has the following determinant:

\[
\det \left[ \phi_i \psi_j \prod_{r=2}^{j} (\alpha_i + \beta_r) \right]_{i,j}^{n \times n} = \prod_{i=1}^{n} (\phi_i \psi_i) \prod_{r=1}^{n-1} \left( \prod_{j=1}^{n-r} (\alpha_{j+r} - \alpha_r) \right)
\]

As a consequence, this kind of matrices has full rank unless some of the \( \phi_i \) or \( \psi_i \) are zero, or at least two \( \alpha_i \) coincide with each other. The latter is clearly not the case for the matrix \( \tilde{J}^{\text{odd}} \), hence it has full rank. After shifting circularly the matrix elements of \( \tilde{J}^{\text{even}} \) over 1 entry to the right, the resulting matrix can be seen to obey (4.16) with \( \alpha_i^{\text{even}} = 6i - 5 \) and \( \beta_{j+1}^{\text{even}} = 2j - 2\delta_{j-1} \) and with nonzero \( \phi_i^{\text{even}} = (2i - 1)(6i - 7)!! \) and \( \psi_i^{\text{even}} = 3 - 2\delta_{i-1} \). Hence, \( \tilde{J}^{\text{even}} \) has also full rank. This concludes the proof. \( \square \)

**References**


Part II

Analysis of instability under random inputs
Introduction

This second part of the thesis concentrates on a different topic in the field of nonlinear dynamic systems, namely their behaviour subject to random input excitations. Simply stated, stability means that any bounded input results in a bounded output. A stationary input acting on a nonlinear system may result in a response which is at first seemingly bounded, possibly for long periods of time, but actually is not: the response may "blow up" (i.e. grow very large) suddenly, without warning, at some random time instant (depending on the input signal’s realization). In practice, for controlled nonlinear plants, this means that the control over the plant is lost, the output can saturate and, if it is not broken, the setup has to be restarted manually again. These unstable effects result in a loss of efficiency, a waste of time and money for the company. On the other hand, obtaining a system with a guaranteed stable and bounded behaviour typically requires large design efforts.

The general objective can be stated as follows: “an attempt to produce quantitative, statistically sound measures, about how frequent a blow-up can occur”. A zero frequency (viz. blow-ups occur with probability zero) would be ideal, but this is not easy to achieve in practice. Indeed, the average frequency has a practical impact: it may be acceptable to use the system anyway, without modification, if the average time before blow-up (also called escape time further on) is known to be long enough. The respective aims of the two chapters composing this second part will be:

- to investigate the (non-immediate) blow-up phenomenon, which can never be encountered with linear systems (1);

- to propose a method that allows to approximate the probability of blow-

---

1This point is motivated as follows. Consider a causal LTI system driven by zero-mean white noise with time-domain variance $\sigma^2_u$, starting from rest at time $t = 0$; it has an output given by

$$ y(t) = \sum_{k=0}^{t} h(k) u(t-k) $$

115
up for a given model subject to a Gaussian input distribution.

The first chapter will explore the phenomenon rather intuitively, starting from input-output measurements, and proceeding with a statistical analysis.

In the second, more formally written chapter, a state-space interpretation will be given, and a method will be set up theoretically to approximate the probability of blow-up on a finite time horizon under some conditions.

with impulse response samples represented as $h(k)$. The output’s variance is given by

$$E[y^2(t)] = \sum_{k=0}^{t} \sum_{l=0}^{t} h(k) h(l) E[u(t-k)u(t-l)]$$

$$= \sum_{k=0}^{t} \sum_{l=0}^{t} h(k) h(l) \sigma_u^2 \delta_{k-l}$$

$$= \sigma_u^2 \sum_{k=0}^{t} h^2(k) \leq E[y^2(\tilde{t})] \text{ for } t < \tilde{t}$$

From which it is seen that this quantity is non-decreasing with time:

- either converging to a limit value as $t \to \infty$, resulting in a second-order stationary response (this behaviour is guaranteed for stable LTI systems by absolute summability of the impulse response);

- either diverging towards infinity, resulting in a loss of stationarity. In this case, an immediate blow-up occurs.
Chapter 5

Instability under random inputs: data-driven perspective

In this chapter, the blow-up phenomenon (also termed “escape”) – occurring with nonlinear dynamic systems driven by stationary random excitations – and its properties, are studied from a statistical point of view, i.e. starting from measurement data. After illustrating the problem first via a simulation example, the escape time’s properties are explored. The probability of escape is introduced as the parameter of the geometrical distribution characterizing the escape time. Next, some issues with a candidate stability definition are addressed. The potential of extreme value statistics to assess the output power’s boundedness is presented. The results are illustrated via Monte-Carlo simulation experiments.

5.1 Introduction

This chapter shows, starting from a very simple example, that, from an experimental point of view (i.e. in a given time record), a nonlinear system may behave “as if” it were stable at certain input amplitudes (e.g. low standard deviation values), and reveal true unstable behaviour at other amplitudes. However, the system is unstable in the classical sense, since there exist input realizations which result in an unbounded output. The aim of this work is...
to provide quantitative measures of the extent of the system’s unstable nature, refining the classical instability concept.

This kind of instabilities may e.g. occur in applications such as sigma-delta modulators (Lota et al., 2007) and accurate electronic amplifiers (de Jong and Meijer, 1997).

The experimental stability analysis of nonlinear feedback systems has already been handled mainly in two publications in the control literature, both relying on the small gain theorem (Zames, 1966; Khalil, 2002), and concerned with the estimation of the gain of the unmodelled nonlinear dynamics. These papers use the loop gain as a quantitative measure:

- In Schoukens et al. (2002, 2004), the first ideas are presented, and, the simplification is made that the maximum gain is replaced by the gain estimate calculated from the available data.

- In Mosskull et al. (2003), the results of the former are somewhat refined, allowing to estimate the maximum gain by an iterative procedure (called power iterations) consisting of a (possibly high) number of experiments.

As mentioned in Mosskull et al. (2003), another drawback of both methods is that the small-gain criterion is conservative; i.e. even if the loop gain exceeds the value $1$, the closed loop system may still be stable. It can be noted that a weaker gain definition has been introduced by Chen et al. (2007), resulting in a necessary and sufficient small-gain criterion. The above-mentioned methods do not apply to more general nonlinear systems not satisfying a nonlinear feedback form $^{(1)}$.

Besides, the related topic of stationary response of nonlinear systems driven by random inputs has been the subject of the interesting work by van de Wouw (1999) and Brouwers (2006), but the former does not consider nonlinear unstable effects and the latter uses a white-box approach in which the governing system equations are known.

The presentation of theoretical stability notions is deferred to the next chapter (in which a model-based point of view will be adopted), since this chapter emphasizes the experimental point of view, in which no stability information is available.

The context of the problem considered here is the following situation occurring in real-life control systems. To control a plant, one usually proceeds in two steps: a model is first constructed (via classical system identification

---

$^{(1)}$Nonlinear feedback systems are defined as the systems which can be rewritten as a feedback interconnection of two dynamic (or static) (non)linear single-input-single-output systems: the output of the feed-forward branch is fed back via the feedback branch to the input in an additive way, see Figure 5.1.
methods), and, next, a (feedback) controller is designed based on the obtained model (via classical control design techniques). In such a way, the feedback system, consisting of controller and plant model, is guaranteed to operate stably. However, the real-life system, consisting of controller and real-life plant, has no stability guarantees anymore, do to model errors taking place in the first step: a bounded input could generate an unbounded output (a “blow-up”). The aim of this chapter is to characterize the stability in a statistically sound way, starting from measurements of the controlled real-life plant.

5.1.1 Chapter’s contributions and outline

The remainder of this chapter is organized as follows. The reader will first be familiarized with the blow-up phenomenon via a simulation example (Section 5.2). The main contributions of this chapter will then be addressed in the sections to follow:

Section 5.3 Analysis and numerical validation of the statistical distribution of the escape times;

Section 5.4 Maximum-likelihood estimation (MLE) of a (record-length independent) probability of escape from successive experiments;

Section 5.5 Intermezzo: a few thoughts on stability for stochastic inputs;

Section 5.6 Introduction of an extreme-value-based quantitative measure revealing the boundedness of the output power, and its estimation from a single input-output data record; formulation of a statistical hypothesis test that draws a decision on the boundedness of the output power and verification with numerical Monte-Carlo experiments.

5.2 Blow-up event: illustration and dependence on the input variance

To give the reader some insight into the problem, a simple example is presented. Consider the nonlinear feedback system shown in Figure 5.1, built out of 2 nonlinear blocks $f_{FF}(q,e)$ and $f_{FB}(q,y)$ present in the feed-forward and the feedback branch respectively. Herein, $q$ is the forward shift operator $x(t - 1) = q^{-1}x(t)$ that introduces the dynamics. The nonlinear feedback system is assumed to be excited with a stationary random input signal $u(t)$. The
Chapter 5. Instability under random inputs: data-driven perspective

Figure 5.1: Block schematic representation of a nonlinear feedback system. It consists of two interconnected single-input-single-output systems, called feed-forward and feedback.

system consists of an LTI system in the feed-forward branch, and a simple static nonlinearity in the feedback branch:

\[ f_{\text{FF}}(q, e(t)) = q^{-1} - q^{-3} \frac{e}{2(1 + 0.9q^{-1})} \quad \text{and} \quad f_{\text{FB}}(q, y(t)) = -y^2 \] (5.1)

It could be noted that there are no specific reasons to choose for this specific example. The phenomena to be studied would also occur with different system orders and nonlinear degrees.

The feedback system is driven by zero-mean Gaussian noise with a fixed power spectrum (here: white Gaussian noise filtered with a fourth order digital low-pass Butterworth filter with cutoff frequency at 0.1 relative to the sampling frequency), with a (total, time-domain) standard deviation \( \sigma_U \). The response (output \( y(t) \)) of the nonlinear feedback system to a given realization of the driving noise is shown in Figure 5.2. From this figure, it is seen that, although the system’s behaviour seems stable at first (time indices up to 90000), the output signal very suddenly blows up without warning \( ^2 \). Moreover, as expected from the random property of the input, the escape time is also random (realization-dependent). Of course, in many practical situations, either saturation or oscillations can occur, and no real blow-up can happen. In these cases, the escape event could be seen as the undesirable saturation or oscillation of the system. The aim is to estimate the probability that an escape occurs, using input-output measurement data.

From Monte-Carlo simulations (1000 runs at each standard deviation), it is seen that, for this system, the risk of a blow-up increases with the standard deviation of the excitation signal.

\(^2\)The blow-up might be expected to be the result of a (very) large spike in the input signal. However, this would be too simplistic, due to the dynamics of the system: a specific sequence of – not necessarily large – input samples can also drive the system to a blow-up. This dynamic view has been the starting point of the different perspective on blow-up as described in chapter 6, where a state-space interpretation is central.
5.2. Blow-up event: illustration and dependence on the input variance

Figure 5.2: Possible behaviour of the nonlinear feedback system for $\sigma_U = 1.54$. The time instant of blow-up (90522) is called escape time.

Figure 5.3: Fraction of simulations with blow-up observed during a simulation of length $N$, vs. input standard deviation $\sigma_U$. Each $\sigma_U$-value corresponds to $N_{\text{sim}} = 1000$ Monte-Carlo runs with simulation length $N = 100000$. 
Figure 5.3 reveals that the system behaves “as if” it were stable below $\sigma_U = 1.1$ (no realizations with blow-up have been observed), and behaves “strongly unstable” beyond $\sigma_U = 1.7$ (almost all simulations of length $N = 100000$ become unstable). In between, there is a transition region where the operation of the system gradually shifts from an “almost safe” to a “very unsafe” mode.

While $\sigma_U$ is increased, it was also observed that the average escape time decreases, i.e. the blow-up tends to occur earlier. Please note that an increase of the simulation length, yields an increase of the fraction of simulations experiencing a blow-up. Those two facts will be confirmed in the next subsection via the theoretical escape-time distribution. A record-length independent measure will be introduced as well.

### 5.3 Escape times and their distribution

Escape times ($^3$) – defined as the time instants where the output leaves some compact domain $(-c, c)$ with $c > 0$ a high, unlikely value of the output – are now further investigated. As a consequence of the stochastic property of the input, the escape time inherits this randomness. Its statistical description could be seen as a memoryless ($^4$) waiting process, i.e. a series of independent Bernoulli experiments (at each time instant) waiting for an outcome 1. The outcome 0 is associated with the event $|y(t)| < c$, and has a probability $q_e = 1-p_e$; the outcome 1 is associated with the event $|y(t)| \geq c$ and has a probability $p_e$. This leads in a natural way to a geometrically distributed escape time with parameter $p_e$. Its probability mass function is given by:

$$P(\text{escape time} = k > 0) = q_e^{k-1}p_e$$

(indeed, $k-1$ zero-outcomes followed by a single 1-outcome are needed to obtain an escape at time $k$). Notice that the support is chosen to exclude $k = 0$, viz. $\{1, 2, 3, \ldots\}$. This interpretation is in agreement with Freidlin’s results (Freidlin and Wentzell, 1998) about perturbations of nonlinear systems, where (in the continuous-time case) it is found that, under input excitations with weak long-range dependency, the first exit time of the state trajectories out of some compact domain are approximately exponentially distributed (which is the continuous memoryless counterpart of the geometric distribution). Note

---

$^3$The words escape and blow-up will be used interchangeably.

$^4$For a Gaussian input, the memoryless assumption can be motivated via the Wold decomposition theorem, which states that there exist an i.i.d. signal driving the nonlinear system through an LTI system.
that the probability of a $N$-samples-long simulation without blow-up equals
\[ P(\text{simulation without blow-up}) = P(\text{escape time} > N) = P(\forall k \in \{1, \ldots, N\} \text{ no escape occurs}) = (1 - p_e)^N \]  
(5.3)

from which the record-length dependency is clear. Hence, the fraction of simulations without blow-up vanishes exponentially as the record length increases.

The parameter $p_e$ depends on the nonlinear system itself, and on the statistical properties of the input signal (among which its standard deviation). \(^5\) The symbol $p_e$ will be termed “probability of escape”.

The expected escape time follows directly from the geometric distribution (Stuart and Ord, 1987, sections 5.16 and 5.17):
\[ E[\text{escape time}] = p_e^{-1} \]  
(5.4)

and hence it is indeed verified that blow-ups tend to occur earlier as $p_e$ increases. Note that (5.4) is expressed in samples (it is dimensionless; the time variable is assumed to be discrete).

To verify the validity of the geometric distribution, a set of 1000 numerical Monte-Carlo runs has been performed for a fixed simulation length of $N = 100000$ samples, at two different input standard deviation values $\sigma_U$: $\sigma_U = 1.80$ and $\sigma_U = 1.53$. The escape times for $c = 1000$ are recorded, and, for $\sigma_U = 1.80$, both the normalized histogram and the geometric probability distribution function (pdf) are shown in Figure 5.4. The $p_e$-value was estimated using an MLE described later in this chapter.

From this figure, it is seen that the recorded escape times (represented by $\tau_i$) have a distribution which is very close to the geometric one (this visual agreement has also been confirmed objectively via a statistical test, see section 5.4.1 on page 125). For the second case ($\sigma_U = 1.53$), the value of $p_e$ is lower, and the simulation length $N$ was still kept fixed. Hence, for a certain fraction of the simulations, the output signal $y(t)$ does not escape from the domain $(-c, c)$ at any time instant from $t = 1 \ldots N$. Therefore, in these cases, the escape time would exceed the value $N$, but its actual value is not available. To be able to take this data into account, the escape times are truncated to a maximal value $N + 1$:
\[ \tilde{\tau}_i = \min(\tau_i, N + 1) = \begin{cases} \tau_i & \tau_i \leq N \\ N + 1 & \tau_i \geq N + 1 \end{cases} \]  
(5.5)

\(^5\) Although, strictly speaking, $p_e$ also depends on $c$, it will not be critically dependent on it for $c$ admitting high, unlikely values of the output signal $y(t)$. This is motivated as follows: if the output leaves the domain $|y(t)| < c$, it will be highly unlikely that it comes back. This dependency on $c$ will hence be ignored.
Chapter 5. Instability under random inputs: data-driven perspective

Those truncated escape times are always known at the end of the simulation. This truncation explains the large peak in the histogram represented in Figure 5.5. In the next section, the MLE to estimate $p_e$ is outlined.

5.4 Estimating the probability of blow-up from successive experiments

Starting from the (truncated) geometrically distributed escape times, an MLE $\hat{p}_e$ was constructed, that allows to estimate $p_e$ from a series of $N_{\text{sim}}$ independent simulations with length $N$. For each simulation $i$, $i = 1 \ldots N_{\text{sim}}$, the truncated escape time $\tilde{\tau}_i$ is recorded. Let $N_1$ and $N_2 = N_{\text{sim}} - N_1$ denote respectively the number of simulations for which an escape occurred ($\tilde{\tau}_i = \tau_i \leq N$) and for which no escape is observed ($\tilde{\tau}_i = N + 1$). Then it can very easily be shown (see Appendix 5.A on page 138) that the MLE of $p_e$, is given by

$$\hat{p}_e = \left( \frac{1}{N_1} \sum_{i=1}^{N_{\text{sim}}} I_{\{\tau_i \leq N\}} \tau_i + \frac{NN_2}{N_1} \right)^{-1} \quad (5.6)$$

where $I_{\{A\}}$ is an indicator function admitting the value 1 when $A$ is true and 0 otherwise. Besides, it is seen that the MLE automatically selects how the inverted average escape time (see (5.4)) has to be corrected via the second
5.4. Estimating the probability of blow-up from successive experiments

Figure 5.5: Distribution of the escape times for $\sigma_U = 1.53$, $\hat{p}_e = 1.43 \cdot 10^{-5}$

term to take the truncation (5.5) into account. Due to the truncation taking place, this MLE is a nonstandard variant of the classical MLE for a geometric distribution, and was not available prior to this work.

5.4.1 Validation of the geometric distribution

To validate the interpretation of geometrically distributed escape times, Kolmogorov-Smirnov goodness-of-fit tests (Kendall and Stuart, 1979) were performed. At neither of the two $\sigma_U$ values, the null hypothesis of geometrically distributed escape times could be rejected at significance level $\alpha = 5\%$. These tests thus indicate that the deviation between the observed distribution of the escape times and the postulated geometric distribution is not statistically significant.

The preceding paragraphs suggest a practical and powerful way to estimate the parameter $p_e$ of the geometric distribution (which is nothing else than the probability at each time step that the output leaves the domain $|y(t)| < c$), starting from a set of $N_{\text{sim}}$ simulations (or experiments) of length $N$. It can even handle cases where $N$ is low compared with the average escape time $1/p_e$ (at the cost of a somewhat higher variance). The asymptotic variance is given by the Cramér-Rao lower bound

$$\lim_{N_{\text{sim}} \to \infty} N_{\text{sim}} \text{var}\hat{p}_e = p_e^2 (1 - p_e) \left( 1 - (1 - p_e)^N \right)^{-1}$$ (5.7)
(see Appendix 5.B on page 139 for a proof). This expression has been validated via numerical Monte-Carlo simulations. From the asymptotic variance expression, it can be seen that, in order to produce an accurate estimate $\hat{p}_e$, the method requires high values of $N_{\text{sim}}$ and $N$. This means a large number of successive experiments with a large number of data points. Hence this may be time-consuming.

5.4.2 Two remarks

5.4.2.1 On the use of a finite $c$-value

The reader might wonder whether it could be interesting to consider the limit for $c$ tending to infinity. In many cases, the answer is no. This is because, unlike their continuous-time counterparts, discrete-time nonlinear systems cannot have finite escape times for $c = \infty$, as long as the global nonlinear dynamics maps finite arguments to finite function values (viz. no singularities may be present in the mapping from past inputs and outputs to present output). This means that, in these cases, even if there is a blow-up of the output signal, the individual values of this signal $y(t)$ will admit finite values at every time instant $t \in \mathbb{N}$. Hence, if the limit $c \to \infty$ was taken, no escape could occur at finite time instants.

5.4.2.2 On the “start-up transient”

To be precise, the actual escape times satisfy a “shifted” geometrical distribution instead of the standard geometrical distribution as mentioned before. This means that in practice no blow-ups are observed at the very first samples of the simulation. This has to do with the fact that the simulations start from zero initial conditions (the system starts from rest), and the system’s finite bandwidth (the time constant is several samples long). As a consequence, it takes the output some time to reach the value $\pm c$ from zero.

5.4.3 Some more results for the example system

As mentioned before, the extent of unstable behaviour of nonlinear dynamic systems can be very dependent on the nature of the input signal, e.g. its amplitude. This can now easily be seen from Figure 5.6, representing the estimated probability of escape of the system of section 5.2 versus the standard deviation $\sigma_U$ of the filtered Gaussian input signal. For this estimation, MLE (5.6) was used. Similar results have been obtained for nonlinear feedback systems with $f_{FB}(q, y) = -y^3$ and $f_{FB}(q, y) = -y^5$. 
5.4. Estimating the probability of blow-up from successive experiments

Figure 5.6: Probability of escape $p_e$ vs. standard deviation $\sigma_U$ (bold line, logarithmic axis on left); probability of an escape during a simulation of length $N$ (dashed line, linear axis on right). Each $\sigma_U$-value corresponds to $N_{\text{sim}} = 1000$ Monte-Carlo runs with simulation length $N = 100000$. 
5.4.4 Additional remark on an alternative way of estimating the probability of escape

It can be noted that the following alternative to the MLE can be proposed for the estimation of the probability of escape $p_e$: via the estimation of the fraction of simulations without blow-up and finally by solving (5.3) to $p_e$. To be more specific, this means, defining $P$ (simulation without blow-up) = $P_N$ that

$$\hat{P}_N = \frac{N_2}{N_{sim}} = \frac{1}{N_{sim}} \sum_{i=1}^{N_{sim}} I_{\{\tau_i > N\}}$$

$$\tilde{p}_e = 1 - \hat{P}_N^{1/N}$$

The question arising at this point is which of both estimates $\tilde{p}_e$ or $\hat{p}_e$ has to be preferred.

Due to the MLE’s consistency and efficiency, it follows that the variance of this alternative method would be asymptotically (6) at least as large as the MLE’s variance, and that nothing can be gained asymptotically (a loss in efficiency can even be expected). Therefore, asymptotically the MLE $\hat{p}_e$ has to be preferred.

5.5 Input-output stability for stochastic processes

Classically, the input-output stability concept is defined for deterministic signals (Schoukens et al., 2004; Zames, 1966): a system is said to be stable if every bounded input signal results in a bounded output (the norm most widely used is the “max” or $l_{\infty}$ norm). For random excitations (i.e. stochastic signals), stochastic stability notions can be used. E.g. stability in probability and stability in terms of moments have been defined for stochastic differential equations (Arnold, 1974; Has’minskii, 1980). Two thoughts for a (related) stochastic stability notion – a bounded output power (second-order moment) for any input power in a given range – are mentioned below:

- Even if no blow-up is observed at low input amplitudes, this does not imply that this may never occur at these levels. Suppose though that there is a significant risk of blow-ups at higher input amplitudes and that the input pdf’s support is infinite (e.g. a Gaussian input). Simulations running infinitely long, will all experience blow-ups with probability one, both at low and at high amplitude levels: indeed, there exist input realizations causing blow-ups, hence there is always at least a tiny risk of

---

6 Asymptotically here means that the number of simulations $N_{sim}$ tends to infinity.
5.6 Extreme-value based method

5.6.1 Introduction

Although the previous method (section 5.4) could be useful, the careful reader might argue that some information is lost, making the method suboptimal. Indeed, the actual data sequences as such are not used; only the (truncated) observed escape times are used. The forthcoming ideas will be applicable to a single input-output data record. Therefore, at most a single escape could be present in the observations. It is not possible to estimate the probability of escape in this case, which would be a big extrapolation issue\(^7\). The drawback is that another measure of the extent of instability has to be used instead, and is only indirectly related to the probability of instability.

The fundamental idea here is to make the plausible assumption of stationarity of the response on the portion of time before blow-up (the practical issue is the lack of ergodicity on the whole time axis; after blow-up, one cannot collect informative data on the stationary part anymore). Since the behaviour is very dependent on the input, it is assumed to be a member of the family of inputs with a given distribution and a given finite standard deviation (i.e. finite power). The existence of the output power in the stationary segment will be analyzed and quantified in the sequel. First notice that the only way to let the power – that is the integral \(E\left[X^2\right] = \int_{-\infty}^{\infty} x^2 p_X(x) dx\) – of a random variable \(X\) with probability density \(p_X(x)\) diverge, is via “heavy” tails (i.e. a too slow\(^{7}\)This extrapolation issue will also be discussed later on in the general conclusions chapter.

generating that scenario, also at low levels (since the pdf’s support is infinite). If the output variance can be unbounded at some input amplitudes, it cannot be bounded at other input amplitudes.

- A second issue is the non-stationarity: a time-invariant’s system’s response is only stationary if it is a nonlinear moving average system (only depending on a finite number of previous inputs, and not on past outputs, see Brillinger, 1981), or if it is a stable system. As already mentioned, classically speaking the systems considered are unstable. Therefore, non-stationarity can occur as a consequence of unstable operation and the output signal’s pdf and statistical moments become time dependent – inherently, the instability is a “transient” phenomenon (not damping out); speaking about the output’s variance as such does not make sense. Note that this lack of stationarity has also to do with the absence of the fading memory (Boyd and Chua, 1985) and convergence (Pavlov et al., 2007) properties, which would imply that the system tends to “forget” its past.
Chapter 5. Instability under random inputs: data-driven perspective

Figure 5.7: For this stochastic time series, the exceedances of the threshold $\eta = 0.5$ are indicated by means of encircled dots.

decrease of $p_X(x)$). The frequency and amplitude distribution of outliers or spikes present in the output signal will hence be crucial.

5.6.2 Extreme value statistics

This motivates a recourse to extreme-value statistics (Coles, 2001; Embrechts et al., 1997), which is a well-known discipline already applied in lots of financial applications (insurance) and in the statistical analysis of environmental phenomena (such as waves, wind speed, temperature, floods, earthquakes, and rainfall). To be more specific, the family of the generalized Pareto distributions (GPD, to be defined somewhat further on) is very powerful, since it is known that the tails of a very wide class of distributions (e.g. normal, exponential, lognormal, gamma, beta, uniform, log-gamma, Cauchy, ...) tend to GPD’s (Coles, 2001; Embrechts et al., 1997). The practical conclusion is a stochastic variable’s tail can usually be approximated by a GPD. This can be achieved by fitting the GDP model parameters to the excesses over a sufficiently high threshold of the data. As shown in Figure 5.7, the exceedances are the outcomes of the stochastic variable (say, $X$) which exceed a specified threshold $\eta$ (i.e. $X | X > \eta$), and the threshold excesses are defined as those values minus $\eta$ (i.e. $X - \eta | X > \eta$).

The GPD is usually defined by two parameters: a scale parameter $\sigma > 0$ and a shape parameter $\xi$. Its cumulative distribution function $G_{\xi, \sigma}$ is defined
5.6. Extreme-value based method

Figure 5.8: Probability density function of the Generalized Pareto Distribution

by:

\[ G_{\xi,\sigma}(x) = \begin{cases} 
1 - (1 + \xi x / \sigma)^{-1/\xi} & \xi \neq 0 \\
1 - \exp\left(-\frac{x}{\sigma}\right) & \xi = 0
\end{cases} \] (5.8)

defined on \{x: x > 0 \text{ and } 1 + \xi x / \sigma > 0\}. Figure 5.8, representing the corresponding probability density function for several values of \( \xi \), shows that \( \sigma \) indeed controls the scaling of the GPD variable (notice the labeling of the horizontal axis), and, that \( \xi \) determines the “tail behaviour” of the distribution; i.e. \( \xi < 0 \) corresponds to a tail with finite support \( x \in [0, -\frac{\sigma}{\xi}] \), \( \xi = 0 \) corresponds to an exponential decaying tail with infinite support, and the case \( \xi > 0 \) corresponds to a slower-than-exponential (polynomial) decaying tail with infinite support (“heavy tails”).

For \( \xi > 0 \), and for \( x \gg \sigma / \xi \), this leads to:

\[ \mathbb{P}(X > x) = (1 + \xi x / \sigma)^{-1/\xi} \approx (\xi x / \sigma)^{-1/\xi} \] (5.9)

The exponent \(-1/\xi\) indicates that the shape parameter \( \xi \) dictates the rate of decay of the tail of the distribution. The higher \( \xi \), the heavier the probabilistic mass present in the tails. Therefore, there exists some critical value of \( \xi \) (which turns out to be \( 1/2 \)), where the second moment stops to exist. It can be shown (Embrechts et al., 1997) that:

\[ \mathbb{E}[X^2] = 2\sigma^2 (\xi - 1)^{-1} (2\xi - 1)^{-1} \quad \xi < 1/2 \] (5.10)

The whole idea of the method relies on estimating the value of the shape parameter \( \xi \) of the output signal, given that the input power is bounded. The
closer $\xi$ approaches the value $1/2$, the higher the output power. If $\xi \geq 1/2$, the output power does not exist. The shape parameter $\xi$ is therefore believed to be a measure for the unstable nature of the system.

### 5.6.3 Estimation of the GPD’s shape parameter $\xi$

Before estimating $\xi$, in a first step, the nonlinear system is split up in two contributions: one contribution being the output of a stable model (to be estimated from the input-output data), and an (unmodelled) remaining part, defined as the residuals of the stable model (viz. the difference between the measured and observed output) (see Figure 5.9). The unmodelled part is possibly unstable.

The stable model is assumed to have an output $\hat{y}(t)$ with bounded power. If the model approximates the behaviour of the system, the residuals $y_r(t)$ will be small in some sense, e.g. in least-squares sense. Then a finite power condition on $y_r(t)$ is pursued. If the user has no idea of a stable (nonlinear) model structure that might approximate the system, the best-linear-approximation framework allows one to easily do the linear identification in the presence of nonlinearities (Pintelon and Schoukens, 2001; Pintelon et al., 2010) (8). The splitting of the output in two parts has two advantages:

- Already a (possibly large) part of the power which is automatically bounded is taken away, due to the stability and bounded power assumption of the model;

- If the stable model is the best linear approximation, $y_r(t)$ will reveal the hidden (unmodelled) nonlinear effects occurring in the system; it contains more information about the potential instability than $y(t)$. This has also to do with the Gaussianizing effect of a linear system (if sufficient dynamics are present); the non-Gaussian part contains the essential nonlinear information.

Both effects ease the estimation process of the shape parameter $\xi$.

---

8Note that a stability constraint has to be used in addition to the standard framework.
After estimation of a stable model, \( |y_r(t)| \) is calculated. The absolute value is taken since this allows one to apply the extreme-value method to both the upper and lower tail in one step. Then, as classically in extreme-value statistics, a threshold level \( \eta \) is chosen (e.g. the 90\% percentile of the data \( |y_r(t)| \)) \(^9\). Next, those \( |y_r(t)| \)-values are retained, which are exceeding \( \eta \) and sufficiently separate in time (further away than the dominant time constant of the stable model). This process is known as data-declustering (Coles, 2001). It is very important, since not taking care of the dependencies, invalidates the likelihood expressions \(^{10}\). Finally, neglecting the (very) small dependencies between the retained data samples, and, after removing the threshold level \( \eta \), the values of the GPD-parameters \( \sigma \) and \( \xi \) are estimated by a standard MLE (Coles, 2001; Embrechts \textit{et al.}, 1997). It is interesting to know that extreme value theory indicates that the scale parameter does not depend on the chosen threshold level \( \eta \), as long as the approximation of the distribution of the exceedances of \( \eta \) by a GPD is valid. For the example given in section 5.2, the Kolmogorov-Smirnov goodness-of-fit test (Kendall and Stuart, 1979) could not distinguish significant (\( \alpha = 5\% \)) model errors between the data and the estimated GPD.

It can be noted that other extreme-value methods exist (e.g. based on block maxima, order statistics, point processes), each with their advantages and drawbacks (Coles, 2001; Embrechts \textit{et al.}, 1997). The choice for a threshold-based approach allows easily to make a connection with the boundedness of the second order moment. Although extreme value statistics were used as well in Schoukens \textit{et al.} (2004), the method outlined here makes use of the GPD to model the tails, while Schoukens \textit{et al.} (2004) use the method of block-maxima (leading to the “Generalized Extreme Value” distribution).

### 5.6.4 Impact of additive noise

Generally speaking, additive input and output noise might have an impact on the estimated shape parameter \( \hat{\xi} \). Nevertheless, if the noise contribution is independent of the noiseless input and output, and if it has a finite power, theoretically, this does not influence the results about the (un)boundedness of the output power. Hence, if these conditions are satisfied, the methodology remains valid.

\(^9\)Note that a trade-off has to be made: a low \( \eta \)-value increases the fraction of data-point retained (which reduces the variance), while a high \( \eta \)-value is needed to reduce the bias (the GPD requires a high threshold).

\(^{10}\)Considering an i.i.d. random process over a time window, the threshold exceedances would be uniformly distributed over time in that window. But, practical random processes are not iid, and therefore threshold exceedances tend to cluster together (Coles, 2001).
5.6.5 Results of numerical experiments

Finally, some results obtained on the system introduced in section 5.2 are reported. To this end, at 5 different standard deviation values $\sigma_U$, $\xi$ was estimated repeatedly (Monte-Carlo with 40000 simulations without blow-up \(^{11}\) of 10000 data points each) and the normalized histograms of the obtained $\hat{\xi}$ values are shown below in Figure 5.10.

It is seen that, on average, the shape parameter $\xi$ increases with increasing input standard deviation values $\sigma_U$. Since from Figure 5.6 it is known that also the probability of escape $p_e$ is an increasing function of $\sigma_U$, this demonstrates that $\xi$ makes sense as a quantitative measure for the unstable behaviour of nonlinear dynamic systems. The precise relationship between $\xi$ and $p_e$ is believed to be system-dependent.

5.6.6 Hypothesis test

Now a statistical hypothesis test can be drawn about the boundedness of the output power. Doing so, results in an automatic decision scheme looking for

\(^{11}\)This means that the simulations in which a blow-up was observed were discarded, since after blow-up, estimation of $\xi$ is useless.
5.6. Extreme-value based method

statistical arguments (in the data) either in favor of, or against the boundedness of the output power.

The null hypothesis $H_0$ (resp. alternative hypothesis $H_a$) states that the output has unbounded power (resp. bounded). From the preceding, this can be translated in terms of the shape parameter $\xi$: 

\[
\begin{align*}
H_0 : & \quad \xi = 0.5 \\
H_a : & \quad \xi < 0.5
\end{align*}
\]

This is a one-sided hypothesis test. From the properties of the MLE of the estimated $\hat{\xi}$, it follows that, under the null hypothesis, $\xi$ is normal with mean 0.5 and standard deviation $1.5/\sqrt{N_{\text{exc}}}$ (with $N_{\text{exc}}$ the number of excesses which were used for the GPD fit). Hence, if the estimated $\hat{\xi}$ exceeds the critical value $\xi_c = 0.5 - \Phi^{-1}(1 - \alpha) \cdot 1.5/\sqrt{N_{\text{exc}}}$ (5.11)

(where $\Phi^{-1}(1 - \alpha)$ the $1 - \alpha$ quantile of the standard normal distribution, $\alpha$ is the significance level, typically 5%), $H_0$ cannot be rejected and it is concluded that the system might still behave unstably at the input variance $\sigma_U^2$. Otherwise, the null hypothesis is rejected, and one can conclude, at significance level $\alpha$, at input variance $\sigma_U^2$, that the system’s output power in the stationary part is bounded. E.g. if the number of retained excesses $N_{\text{exc}} = 150$, and the significance level $\alpha = 5\%$ is chosen, one obtains a critical value $\xi_c$ of about 0.3. The corresponding power curve (probability of rejecting $H_0$ as function of the true value of $\xi$) is shown in Figure 5.11.

From (5.11), it can be seen that, the more data is used, the closer $\xi_c$ comes to 0.5, and the lower the risk of type II errors (and the higher the distinguishing power of the test). Notice that, compared to Schoukens et al. (2004), where the risk for an unstable behaviour is estimated based on a very long data set, this method uses typically less data and makes a statistical statement on the boundedness of the output’s power. This work hence complements the paper by Schoukens et al. (2004), which allows one to estimate the maximal nonlinear power gain over a given time horizon (cf. the comparison in Table 5.1). However, one must be cautious with the interpretation of the result: the rejection of the null hypothesis allows a “hard” decision on the boundedness of the output’s power, whereas in the case of non-rejection, the correct statement is to say that there is no strong statistical argument in favor of the bounded-power operation. In the latter case, the system might operate close to the theoretical border $\xi = 0.5$. The preceding discussion motivates the choice of the alternative hypothesis as $\xi < 0.5$ (and not as $\xi > 0.5$); since it allows to make “hard” decisions on the boundedness of the output’s power, allowing to validate the bounded-power operation of nonlinear systems.
### Table 5.1: Comparison between Schoukens et al. (2004) and this work.

<table>
<thead>
<tr>
<th></th>
<th>Schoukens et al. (2004)</th>
<th>This work</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Gain(s)</td>
<td>block maxima method applied to nonlinear power gain(s)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>at a given experiment length (nothing is said)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>power gain(s)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>estimates the extreme values of the nonlinear power gain(s)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>needs (very) long experiments</td>
</tr>
<tr>
<td></td>
<td></td>
<td>can be conservative (small gain theorem)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>apply to any time-invariant nonlinear system</td>
</tr>
<tr>
<td></td>
<td></td>
<td>only applies to nonlinear feedback systems</td>
</tr>
<tr>
<td></td>
<td></td>
<td>bounded support function input</td>
</tr>
<tr>
<td></td>
<td></td>
<td>critical value (CDF) shape parameter = 0.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>critical value nonlinear gain = 1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>power is defined as an (amplitude) distribution property</td>
</tr>
<tr>
<td></td>
<td></td>
<td>power is defined over time</td>
</tr>
<tr>
<td></td>
<td></td>
<td>also considers an input variation to output</td>
</tr>
<tr>
<td></td>
<td></td>
<td>validation gain definition</td>
</tr>
<tr>
<td></td>
<td></td>
<td>model can improve the results</td>
</tr>
<tr>
<td></td>
<td></td>
<td>availability of an approximating nonlinear model</td>
</tr>
<tr>
<td></td>
<td></td>
<td>needs nonlinear model</td>
</tr>
<tr>
<td></td>
<td></td>
<td>data-based, robust to noise</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
5.7. Conclusion

In this chapter, using a simple example, the instability phenomenon occurring with nonlinear systems driven by random excitations has been treated statistically. The probability of an undesirable blow-up in a simulation generally depends (among others) on the input variance and the simulation length. The escape has been interpreted statistically as a memoryless waiting process over time, leading in a natural way to a geometrical distribution for the escape time, with a parameter termed “probability of escape”. This parameter is simulation-length independent and inversely proportional to the average escape time. A MLE has been set-up to estimate this parameter from multiple experiments, in which the escape time is recorded. After an intermezzo highlighting the lack of bounded operation at any input variance on the infinite time-scale, and the non-stationarity of the response, extreme-value theoretic tools (the Generalized Pareto Distribution) have been used to relate the output signal’s boundedness to its tails’ heaviness (determined by the shape parameter). The theory is illustrated throughout via numerical Monte-Carlo experiments.

Figure 5.11: Power curve of the above statistical test
Appendices

5.A Maximum-likelihood estimation of the probability of escape

The likelihood is the probability density function (pdf) of the data evaluated at the observations (here the truncated escape times \(\tilde{\tau}_i\)) and viewed as function of the parameters (here there is a single parameter \(p_e\)). Here, the pdf of one escape time is a truncated geometric distribution:

\[
p_{\tilde{\tau}_i}(x_i) = \begin{cases} 
(1 - p_e)^{x_i-1} p_e & x_i \leq N \\
(1 - p_e)^N & x_i > N 
\end{cases}
\]  

Since each simulation is independent, the multivariate pdf of \(N_{\text{sim}}\) repeated experiments is given by a product of these univariate pdf's:

\[
p_{\tilde{\tau}_1, \ldots, \tilde{\tau}_{N_{\text{sim}}}}(x_1, \ldots, x_{N_{\text{sim}}}) = \prod_{i=1}^{N_{\text{sim}}} p_{\tilde{\tau}_i}(x_i)
\]

Evaluating at the data \(\tilde{\tau}_i\), and taking the logarithm, yields

\[
\log p_{\tilde{\tau}_1, \ldots, \tilde{\tau}_{N_{\text{sim}}}}(\tilde{\tau}_1, \ldots, \tilde{\tau}_{N_{\text{sim}}}) = \sum_{i=1}^{N_{\text{sim}}} \log p_{\tilde{\tau}_i}(\tilde{\tau}_i)
\]  

(5.13)

with, according to (5.12) and the indicator function \(I_A\) (taking the value zero if \(A\) is false and the value 1 if \(A\) is true)

\[
\log p_{\tilde{\tau}_i}(\tilde{\tau}_i) = I_{\{\tilde{\tau}_i \leq N\}} \log \left( (1 - p_e)^{\tilde{\tau}_i-1} p_e \right) + I_{\{\tilde{\tau}_i > N\}} \log \left( (1 - p_e)^N \right)
\]

The negated left-hand side of (5.13) is termed negative log-likelihood and denoted by the symbol \(\mathbf{L}\). Now the terms are rearranged in a standard way.

\[
-L(\tilde{\tau}_1, \ldots, \tilde{\tau}_{N_{\text{sim}}}|p_e) = \sum_{i=1}^{N_{\text{sim}}} \left[ I_{\{\tilde{\tau}_i \leq N\}} \log \left( (1 - p_e)^{\tilde{\tau}_i-1} p_e \right) 
\right.
\]

\[
+ I_{\{\tilde{\tau}_i > N\}} \log \left( (1 - p_e)^N \right) \right]
\]

(5.12)
5.B. Fisher information matrix of the probability of escape

Setting

\[ N_1 = \sum_{i=1}^{N_{\text{sim}}} I_{\{\tilde{\tau}_i \leq N\}} = \# \text{ experiments with a blow-up observed} \]

\[ N_2 = \sum_{i=1}^{N_{\text{sim}}} I_{\{\tilde{\tau}_i > N\}} = \# \text{ experiments without blow-up observed} \]

results into

\[-L(\tilde{\tau}_1, \ldots, \tilde{\tau}_{N_{\text{sim}}} | p_e) = \left( \sum_{i=1}^{N_{\text{sim}}} I_{\{\tilde{\tau}_i \leq N\}} \tilde{\tau}_i - N_1 + N_2 N \right) \log (1 - p_e) + N_1 \log p_e \]

The maximizer of the likelihood is found by solving the equation \( \frac{\partial L(\tilde{\tau}_1, \ldots, \tilde{\tau}_{N_{\text{sim}}} | p_e)}{\partial p_e} = 0 \). This yields

\[ 0 = \left( \sum_{i=1}^{N_{\text{sim}}} I_{\{\tilde{\tau}_i \leq N\}} \tilde{\tau}_i - N_1 + N_2 N \right) \frac{-1}{1 - p_e} + \frac{N_1}{p_e} \]

\[ = \left( \sum_{i=1}^{N_{\text{sim}}} I_{\{\tau_i \leq N\}} \tau_i - N_1 + N_2 N \right) \frac{-1}{1 - p_e} + \frac{N_1}{p_e} \]

(since the event \( \{\tilde{\tau}_i \leq N\} \) is true iff the event \( \{\tau_i \leq N\} \) is true, and, in that case, \( \tau_i = \tilde{\tau}_i \)) and the MLE expression of \( \hat{p}_e \) follows:

\[ \hat{p}_e = \left( \frac{1}{N_1} \sum_{i=1}^{N_{\text{sim}}} I_{\{\tau_i \leq N\}} \tau_i + \frac{N N_2}{N_1} \right)^{-1} \quad (5.14) \]

This concludes the proof. \( \square \)

5.B Fisher information matrix of the probability of escape

The expected value of the negative log-likelihood’s second derivative (Hessian) w.r.t. \( p_e \) is known as the Fisher information matrix, denoted \( F_i \); the inverse of which equals the asymptotic covariance. This Hessian is given by:

\[ \left( \sum_{i=1}^{N_{\text{sim}}} I_{\{\tilde{\tau}_i \leq N\}} \tilde{\tau}_i - N_1 + N_2 N \right) \frac{1}{(1 - p_e)^2} + \frac{N_1}{p_e^2} \]
Chapter 5. Instability under random inputs: data-driven perspective

Taking the mathematical expectation according to the geometric distribution yields (due to the fact that each escape time has the same distribution, each term of the sum has the same expected value)

\[
F_i = \frac{1}{(1 - p_e)^2} \left( N_{\text{sim}} \mathbb{E} \left[ I_{\{\tilde{\tau}_i \leq N\}} \tilde{\tau}_i \right] - \mathbb{E} [N_1] + N \mathbb{E} [N_2] \right) + \frac{1}{p_e^2} \mathbb{E} [N_1]
\]  \hspace{1cm} (5.15)

The three expected values are now calculated, and will next be substituted into that expression

\[
\mathbb{E} \left[ I_{\{\tilde{\tau}_i \leq N\}} \tilde{\tau}_i \right] = \sum_{k=1}^{N} k (1 - p_e)^{k-1} p_e
\]

\[
= \frac{p_e}{1 - p_e} \sum_{k=0}^{N} k (1 - p_e)^k
\]

applying

\[
\sum_{k=1}^{N} k q^k = - N q^{N+1} \frac{1 - q}{1 - q} + q \frac{1 - q^N}{(1 - q)^2},
\]

which is a direct consequence of formula (0.113) of Gradshteyn and Ryzhik (2000), one obtains

\[
\mathbb{E} \left[ I_{\{\tilde{\tau}_i \leq N\}} \tilde{\tau}_i \right] = \frac{p_e}{1 - p_e} \left[ - N (1 - p_e)^{N+1} \frac{1 - p_e}{p_e} + \frac{(1 - p_e) \left( 1 - (1 - p_e)^N \right)}{p_e^2} \right]
\]

\[
= - N (1 - p_e)^N + p_e^{-1} \left( 1 - (1 - p_e)^N \right)
\]

\[
\mathbb{E} \left[ \frac{N_1}{N_{\text{sim}}} \right] = 1 - \mathbb{P} \text{ (simulation without blow-up)}
\]

\[
= 1 - (1 - p_e)^N
\]

\[
\mathbb{E} \left[ \frac{N_2}{N_{\text{sim}}} \right] = \mathbb{P} \text{ (simulation without blow-up)}
\]

\[
= (1 - p_e)^N
\]

Defining \( P_N = (1 - p_e)^N \), and rescaling the left and right hand sides, we find

\[
N_{\text{sim}} \mathbb{E} \left[ I_{\{\tilde{\tau}_i \leq N\}} \tilde{\tau}_i \right] = N_{\text{sim}} (-N P_N + p_e^{-1} (1 - P_N))
\]

\[
\mathbb{E} [N_1] = N_{\text{sim}} (1 - P_N)
\]

\[
N \mathbb{E} [N_2] = N_{\text{sim}} N P_N
\]
5.B. Fisher information matrix of the probability of escape

Plugging the above into (5.15), results in

\[
F_i = \frac{N_{sim}}{(1-p_e)^2} \left[-NP_N + p_e^{-1} (1 - P_N) - (1 - P_N) + NP_N \right] + \frac{N_{sim}}{p_e^2} (1 - P_N) \\
= N_{sim} \left(p_e^{-1} - 1 \right) \left(1 - P_N \right) \\
= N_{sim} p_e^{-2} (1 - p_e)^{-1} \left(1 - (1 - p_e)^N \right)
\]

This concludes the proof. □

References


Chapter 6

Instability under random inputs: model-based perspective

In this chapter, the behaviour of nonlinear dynamic systems driven by stationary random excitations is studied from a model-based perspective – i.e. starting from a perfect knowledge of the system under study and its driving random input – over a finite time interval (a burst excitation is assumed). For a given discrete-time nonlinear state-space model operating in the neighbourhood of a stable equilibrium, a “blow-up” is seen as the event of escaping out of a region of attraction. Based on Laplace integration, a method is outlined to approximate a future state’s probability density function (pdf) at low excitation amplitudes. Inspection of this pdf can reveal additional insights into the complex behaviour of an abstract state-space model, compared with the simulation approach. The probability of staying inside the region of attraction (viz. obtaining a bounded operation subject to an input active in a finite time interval) can be obtained by integration of this pdf. The state pdf estimation is illustrated with numerical Monte-Carlo simulation experiments.
Chapter 6. Instability under random inputs: model-based perspective

6.1 Introduction

6.1.1 Motivation of this work

Since all real-life systems are to some extent nonlinear, linear models can fail to represent their behaviour in a way satisfactory to the user. A large class of nonlinear plants can be very well approximated via the family of polynomial state-space models (with a state evolution of polynomial form in the state and input). It has good black-box approximation capabilities, and was successfully applied in practice (Paduart et al., 2010). The model can be constructed from input-output measurements of the system under study, via a least-squares data-fitting approach. The problem is that no stability guarantees are given for the resulting nonlinear state-space model; only the underlying linear dynamics are easy to analyze. It is therefore desirable to investigate (automatically) how a given state-space model behaves under a random input with predefined nature when no stability information is given. Moreover, an unstable system’s state can stay within a given bounded region for very long periods of time and with a high probability (possibly not arbitrary close to one). In this chapter, a method is proposed that allows one to approximate the probability distribution of a future state, given the state space equation, the initial state and the nature of the random input, consisting of a Gaussian term with known power spectrum and (if desired) a known deterministic term. Besides the randomly excited black box nonlinear model structures, possible applications include the analysis of the start-up behaviour of nonlinear oscillators, the investigation of physical (white-box) models, and controlled nonlinear plants with random noise on top of the deterministic reference signal.

6.1.2 Example of locally stable, globally unstable dynamics

As follows from the previous chapter, for unstable systems, the time interval over which a bounded response is observed depends both on the system itself and on the stochastic input’s properties. Nonlinear models’ responses (starting from rest) can, e.g., be bounded (this behaviour can appear “stable”) for long periods of time at low input amplitudes (variance for stochastic inputs), and at high amplitudes only remain bounded for short periods of time, revealing the true unstable nature of the system. For the amplitude levels in between, there can be a transition, with a decrease of the typical time over which the response remains bounded as function of the input variance. This kind of effect can be expected with state space systems with a state transition function of the form
6.1. Introduction

Figure 6.1: Unbounded response with interpretation based on a Region Of Attraction (boundary shown in black) in the state space. Based on Khalil (2002), exercise 1.22 (4).

\[ f(x, u) = Ax + Bu + o\left(\begin{bmatrix} x \\ u \end{bmatrix}\right) \]

(where \( u \) is the input and \( x \) is the state) with eigenvalues of the \( A \) matrix lying in the open unit disk, i.e., nonlinear systems having asymptotically stable underlying linear dynamics. If both the input and the state vector are small, the \( o() \)-term vanishes; as a result, the origin is a locally stable equilibrium of the unforced system (autonomous, \( u = 0 \)). The autonomous system’s response is a given deterministic state trajectory. In this sense, a nonzero random input causes state fluctuations. Inside a so-called Region Of Attraction (ROA) around the equilibrium point, the state is attracted towards the equilibrium in the unforced (i.e., zero input) situation. Outside, the nonlinear terms can become very active and cause the state to be repelled away from the origin. Assuming that the input enters linearly into the state equation, Figure 6.1 shows an example where, most likely, the operation will be bounded at low excitation amplitudes, and unbounded at high amplitudes.

6.1.3 Towards a concrete problem setting

In this context, the practical goal is to quantify the “risk” for the state to leave a certain domain \( \mathcal{D} \) in a given period of time, say, of \( \tau \) samples, for a specified random excitation. In this work, we restrict ourselves to the case of a known model equation, initial state and colored Gaussian input with known power spectrum. To achieve this goal, we propose in this chapter to calculate the
probability density function (pdf) of the state at time $\tau$. If the pdf is integrated over a domain $D$ chosen as the ROA – see e.g. Zečević and Šiljak (2010) and Balint et al. (2006) and references therein for methods estimating the ROA in resp. the continuous- and the discrete-time setting – the probability is obtained for the state to be asymptotically attracted to the equilibrium after a burst of $\tau$ noise samples (followed by a zero input). This will be called bounded operation over a time interval of length $\tau$ (to be defined in the sequel).

The reader may comment that the state pdf at time $\tau$ can also be estimated from Monte-Carlo simulations of the system. One advantage of the method to be presented, is that it obtains a much better accuracy in low pdf regions, where a Monte-Carlo simulation hardly ever turns up. A second advantage is that the results for a range of standard deviations can be obtained, at almost no additional cost.

6.1.4 Background information on classical notions of stability, and situation of this work

This section aims at situating our problem setting w.r.t. the well-established theory on stability of nonlinear dynamic systems. A large number of notions and definitions have been introduced; a thorough elaboration of all of them is out of the scope of this thesis. However, one can distinguish between only a few main philosophical stability notions:

- Lyapunov stability,
- input-output stability,
- input-to-state stability,
- fading memory and convergence,
- passivity and dissipativity.

After stressing their main ideas (which is done here only in the continuous-time setting for simplicity), the present work will be situated.

Lyapunov stability dates back to 1892 (An English translation was republished for the centennial: Lyapunov, 1992). This classical stability notion was originally formulated for an unforced (autonomous) state-space system $\dot{x} = f(x)$. Since no external inputs are assumed to affect the system, it is also referred to as “internal” stability. It studies the stability of solutions $x(t, x_0)$ of the differential state equation w.r.t. perturbations of the initial state $x_0$, in particular for equilibria (i.e. constant solutions). E.g. asymptotic stability expresses that a state close to an equilibrium would converge towards that point.
as time evolves towards infinity. The system is assumed to be finite-dimensional
and the absence of input somehow isolates it from its environment.

Input-output stability is a more recent concept, originating from the work
by Sandberg (1965) and Zames (1966). In essence, it studies the stability of a
nonlinear system w.r.t. perturbations in the input. Since, through the input,
the system is “open” to influences coming from its environment, it is also known
as external stability. The system is viewed as an operator mapping an input
signal to an output signal (both members of well-defined Banach-signal-spaces),
for fixed initial conditions. A system is input-output stable, if inputs which
are bounded in some sense, produce bounded outputs. Then, a nonlinear gain
exists that relates the input norm to the output norm. It is a natural framework
for cascade and feedback connections of systems (e.g. small-gain theorem), and
does not require the system to be written in a finite-dimensional state-space
form.

The two above-mentioned stability concepts are interrelated and comple-
ment each other; under some conditions, there exist implications between both
notions of stability (Willems, 1971; Vidyasagar, 1986; Fromion et al., 1996).
However, the “region of attraction” (ROA) concept is not (well) formulated in
the input-output stability framework, while the latter allows to obtain more
powerful and general results (Willems, 1971). Nevertheless, the necessity for
state-space systems to specify the set of initial states from which small-signal
input-output stability holds, implicitly presumes that something similar to the
ROA is contained in that framework. The survey paper of Vidyasagar (1986)
provides an excellent introduction to both Lyapunov and input-output stabil-
ity, discussing their definitions, properties and interrelations in very accessible
manner.

Input-to-state stability (ISS) has been introduced by Sontag (1989), who
aimed at developing a new notion which is consistent with both Lyapunov and
input-output stability theories. A state-space perspective is being adopted:
\[ \dot{x} = f(x, u) \]. A system is said to be ISS if the state’s magnitude is bounded
by a term coming from the input, plus a term vanishing over time depending
on the initial state. The ISS property is equivalent to the combination of
asymptotic stability of the equilibrium of the unforced system (obtained by
forcing \( u = 0 \)), and a property indicating that, as time tends towards infinity,
the system forgets its initial state and is then bounded only by a term dependent
on the input. An input signal converging to zero automatically leads to a state
converging to zero. An important motivation for the work by Sontag is that the
stability of the unforced system does not imply a nice behaviour w.r.t. inputs
(i.e. no input-output or input-state stability). A system is said to be ISS if the
state can be bounded by a nonlinear gain in the input signal magnitude plus a
term vanishing over time depending on the initial state. ISS enjoys also some
properties for interconnected systems (cascade and feedback). A large number of variants of ISS have been studied, but these fall out of the scope of this discussion. Sontag (2008) contains an accessible introduction to and survey of the ISS concept.

In the nonlinear case, most notions of stability have a local (or small-signal) and global version, while, in the linear case, all of them should coincide. The local versions indicate that the stability property may be lost if the initial state is too far from the equilibrium or a too large input is applied. A good textbook reviewing the three notions of stability already discussed, is Khalil (2002).

Next to these three stability notions, the “fading memory”, “convergence”, “passivity”, and “dissipativity” concepts are stronger and imply stability in a classical sense. **Fading memory** and **convergence** are quite related to each other, and to the idea of ISS, that the initial conditions (state or input signal) cannot affect the long-term future behaviour. Therefore, they exclude the presence of chaotic behaviour. “Fading memory” (Boyd and Chua, 1985; Sandberg, 2002) is also an input-output operator notion. It implies that input signals far in the past have relatively little influence on the present output. This notion is related to a unique steady state, and allows to approximate the system arbitrarily closely as a Volterra series. “Convergence” dates back to Russian mathematics in the 1960’s, but revived recently by the work of Pavlov *et al.* (2007). It allows to obtain a coordinate-independent notion (in state-space), leading to a unique bounded globally asymptotic solution (steady state), independent on the initial state. “Incremental stability” (Angeli, 2002) fits into the same idea: the state trajectories of a system, starting from a given initial state and driven by a given input, can tend to the state trajectories of the same system, starting from a neighbouring initial state and driven by an input close to that input. **Passivity** is a strong form of stability that expresses that the system’s internal behaviour tends to loose energy over time. Among others, Willems (1972) developed the theory for **dissipativity**, which is a generalization of passivity. Both “passivity” and “dissipativity” are notions developed in a state-space context.

How do the systems considered in this thesis fit into these classical theories? First, in a Lyapunov sense, these systems have a (uniformly) asymptotically stable equilibrium. This stability is local (or regional), not global. Therefore, if no input is applied, the state is only attracted towards that equilibrium, if the initial state is located in a bounded domain around it (ROA). The initial state is crucial for the ROA-concept. Viewed from an input-output perspective, the “open” system (with external input) is not $L_\infty$-stable (bounded-input-bounded-output), since only inputs below some maximal amplitude would generate a bounded output. This would moreover be the case only if the initial state is contained in some domain (close to the equilibrium). Therefore, the systems
considered in this thesis are small-signal-input-output stable, if the initial state is contained in a given bounded domain. From an ISS perspective, a local version of ISS holds for the systems considered. No fading memory, convergence, dissipativity or similar properties would hold either on a global or large-input level.

### 6.1.5 Randomly excited systems and related literature

The classical stability notions just outlined, assume a bounded input. However, due to its unboundedness, a random input with an infinite (amplitude) support does not fit into these classical frameworks.

Various definitions for stochastic stability have been introduced for stochastic differential equations (Arnold, 1974; Has’minskii, 1980). ISS has also a stochastic counterpart in continuous-time, requiring Itô stochastic calculus, in which the state upper bound holds with probability arbitrary close to one (stability in probability), see e.g. Liu et al. (2008).

The monograph by Freidlin and Wentzell (1998) gives a quite theoretical and mathematically deep treatment of the problem of continuous-time dynamical systems subject to random perturbations (requiring advanced functional calculus). In a paper by Roy and Nauman (1995), a second order continuous-time equation forced with a single sine wave is analyzed numerically. The latter refers to a few publications on the effect of white Gaussian perturbations to the system’s response. Interesting notions as escape probability, mean exit time and most probable exit path (originally introduced by Freidlin and Wentzell (1998)) are reviewed therein. Besides, the optimal path concept is studied for systems driven by Gaussian noise in Dykman and Smelyanskiy (1998), focusing on the prehistory rather than the future. Deterministic and random state changes between multiple basins of attraction were analyzed specifically for semiconductor ring lasers in Gelens (2010).

It can be noted that for continuous-time systems with white Gaussian noise entering the state equation in an affine way, the Fokker-Planck partial differential equation can be written down immediately from the state equation. We will come back to this approach at the very end of this chapter. However, the method to be outlined here also applies to more general state equations with an arbitrary function of the states and inputs. On the other hand, in contrast to the proposed method, the Fokker-Planck approach does not linearize the state equation.

The related concept “stochastic resonance” (Gammaitoni et al., 1998) is often studied for bistable systems (two stable equilibria) excited by the superposition of a small-amplitude (usually periodic) input component and a random noise process. As the input noise level is increased, the response signal-to-noise
ratio increases towards a maximum (this is the point of maximal “cooperation” between the noise and the deterministic part of the input to cross the barrier), and then decreases. In this work, the focus is on one stable equilibrium (with ROA), and, loosely speaking, a “stable equilibrium at infinity” (since states starting at very large amplitudes get attracted towards infinity). The relative size of the ROA of both equilibria is infinite, which is not the case in classical stochastic resonance. The probability to jump back from the ROA of the point at infinity towards the finite ROA is really tiny in the types of systems considered here. Due to this property, stochastic resonance is expected not to take place here.

In this chapter, the model is discrete in time, the input fluctuations on top of the deterministic input are assumed to be colored and Gaussian. We focus on the calculation of the pdf of $x(\tau)$, without needing to simulate the model’s response extensively and explicitly. Doing so, it can e.g. be seen which regions of the state space are very unlikely to be visited on the time scale $\tau$. Since instability is inherently a transient phenomenon, and initially the state can stay inside a ROA for long periods of time, it is therefore interesting to consider the state’s distribution after a finite time in the future. Moreover, due to the instability, the state response to a stationary input (without deterministic term) is not stationary. In the locally stable situation, up to the initial transient and before the leaving the ROA, it just appears as a stationary random sequence.

6.1.6 Main contributions

The main contributions of this chapter are the following:

- (theoretical) calculation of a future state’s pdf as a multivariate integral involving Dirac delta distributions;
- reformulation of the dominant contributions to the integral as a constrained optimization problem, providing an interpretation of its solution, and initial estimates for the iterative procedure;
- evaluation via a variant of the Laplace Integration Method (LIM);
- (theoretical) analysis of the integral approximation error;
- verification of the theoretical results by means of a Monte-Carlo simulation example.

6.1.7 Chapter’s organization

This chapter is organized as follows. In Section 6.2, the assumptions are formalized. In Section 6.3 (the core part), the state’s pdf approximation algorithm
6.2 Assumptions

In this Section, the assumptions are formalized, and a few possible variants, extensions and generalizations to the problem setting are mentioned.

6.2.1 Prior knowledge on the discrete-time state-space model

**Assumption 6.1.** (the class of discrete-time state-space models considered)

*In the state evolution equation*

\[ x(t + 1) = f(x(t), u(t)) \]  \hspace{1cm} (6.1)

*with* \( x(t) \in \mathbb{R}^n \) (\( n \in \mathbb{N}_0 \) is called the model order), \( u(t) \in \mathbb{R} \), the function \( f(\bullet, \bullet) \) is a known twice continuously differentiable function w.r.t. both arguments.

**Assumption 6.2.** (initial state)

*The initial state* \( x(0) \) *is known.*

To allow for an interpretation as a probability of unbounded operation, a ROA is assumed to exist and to be compact. Therefore, it is also assumed that the origin is an asymptotically stable and isolated equilibrium point of (6.1) for \( u(t) = 0 \).

Note that the procedure also applies for non-state-affine and non-input-affine state equations.

If the next assumption is also made, the output will nowhere (no state/input-combination) be allowed to become infinite or undefined.

**Assumption 6.3.** (state,input)-to-output mapping

*The function* \( y = g(x, u) \), *from state and input to output, is a function defined for each* \( x \in \mathbb{R}^n \), *and each* \( u \in \mathbb{R} \) *such that* \( y \in \mathbb{R} \), *viz. the domain is* \( \mathbb{R}^n \times \mathbb{R} \).
6.2.2 Stochastic framework

Assumption 6.4. (stochastic framework)

The input $u(t)$ consists of the superposition

$$u(t) = u_0(t) + n_u(t)$$  \hspace{1cm} (6.2)

of a known deterministic term $u_0(t)$ (defined on the time interval $t \in [0, \tau - 1]$) and a zero-mean, stationary, colored Gaussian noise term $n_u(t)$ with known power spectrum $S_{uu}(\omega)$, and with $\omega$ the (normalized, discrete-time) angular frequency. Translated into the frequency domain, this means that the $k$th frequency component (normalized Discrete Fourier Transform $F_k \frac{\tau}{\sqrt{\tau}}$, as defined below, and denoted DFT for short) of $u(t)$, $U_k$, is circular complex normally distributed (Picinbono, 1993)

$$U_k \sim N_C(U_{0,k}, \sigma^2_k)$$ \hspace{1cm} (6.3)

with $U_{0,k}$ the DFT of $u_0(t)$, $\sigma^2_k = S_{uu}(\omega_k) > 0$ the (total complex) variance, and $\omega_k$ the angular frequency at frequency line $k$.

Throughout, the (total) time domain variance will be denoted as $\sigma^2_U$.

Please notice that, by this assumption, the input signal is automatically real (and hence never infinite) with probability one on $t \in [0, \tau - 1]$. As a second remark, it can be said that a zero input power spectrum is not allowed, since it can create problems further on when it appears in the denominator.

6.2.3 Possible extensions and generalizations

The proposed stochastic framework can be extended without any additional technical difficulty:

- In a production context, model parameters can be considered as random variables changing from product to product (but not over time). One could be interested in $x(\tau)$’s probability distribution as a result of Gaussian distributed model parameters (with specified mean and covariance), stochastically independent of the driving input.

- An extension to Assumption 6.2 is to assume the initial state $x(0)$ to be Gaussian distributed, with a given mean and covariance. This is a (Bayesian) prior on the initial state $x(0)$.

Both extensions can also be combined with each other. For clarity, we stick to the basic situation and leave the extensions for future research.

Moreover, further possible generalizations are the following:
6.2. Assumptions

- Introduction of explicit time dependence in the state evolution equation.
- Usage of a state evolution equation in implicit form

\[ f^{\text{impl}}(x(t), x(t+1), u(t)) = 0 \]

(subject to additional constraints making sure that the solution is well defined).
- Extension towards a multivariate input signal.
- Extension towards a non-Gaussian input signal (this can be achieved quite easily by replacing \( u(t) \) by a static nonlinear function of \( u(t) \) in the state evolution equation); the restriction is that this function is smooth, so that this idea cannot be used for discontinuous input pdfs.

Please note that most of these generalizations do not require a separate analysis, and can be dealt with in the way described in the sequel.

6.2.4 Consequence: bounded operation over a time interval of length \( \tau \)

Assuming (i) that the ROA of the origin exists in the unforced situation, viz.

\[
\text{ROA} = \left\{ x(0) \in \mathbb{R}^n \mid \lim_{t \to \infty} x(t) = 0 \right\}
\]

with \( x(t+1) = f(x(t), 0) \)

and (ii) that the input is a burst of length \( \tau \), viz. \( u(t) = 0, t > \tau \), the probability to obtain a state finally converging towards the (isolated) asymptotic equilibrium point at the origin under all above-mentioned assumptions equals the state’s pdf at time \( \tau \) integrated over the ROA:

\[
\int_{\text{ROA}} p_{x(\tau)}(x)dx \tag{6.4}
\]

Note that the state may escape the ROA during the first \( \tau - 1 \) samples, as long as it comes down into it before the time interval has come to an end. At the same time, a bounded state guarantees a bounded output, since the output mapping is well-behaved.

In the unforced situation, initial states not contained into the ROA can – besides drift away to infinity – also be attracted to other equilibria (fixed points), limit cycles, periodic solutions or chaotic “strange” attractors. If all these phenomena are considered as undesired behaviour, the probability for undesired behaviour is the complement of the above-defined probability (6.4).
6.3 Estimation of the state’s probability distribution

This section contains the core material of the chapter, which is the estimation of the state’s probability distribution at time instant \( \tau \in \mathbb{N}_0 \) under the above assumptions. The probability density function (pdf) \( p_{x(\tau)}(x(\tau)) \) of the state \( x(\tau) \) can be calculated as the integrated form of a higher-dimensional joint pdf based on all states and inputs:

\[
p_{x(\tau)}(x(\tau)) = \int_{-\infty}^{+\infty} p_{x(\tau),u}(x, x(\tau), u) \, dx \, du \tag{6.5}
\]

with \( x \) and \( u \) the vector of intermediate states and intermediate inputs respectively, defined as

\[
x^T = \begin{bmatrix} x(1)^T & \ldots & x(\tau - 1)^T \end{bmatrix} \in \mathbb{R}^{(\tau-1)n} \tag{6.6}
\]

and

\[
u^T = \begin{bmatrix} u(0) & \ldots & u(\tau - 1) \end{bmatrix} \in \mathbb{R}^\tau. \tag{6.7}
\]

The integrand can be rewritten as

\[
p_{x(\tau),u}(x, x(\tau), u) = p_u(u) p_{x(\tau) | u}(x, x(\tau), u) \tag{6.8}
\]

with \( p_u(u) \) a known Gaussian pdf based on Assumption 6.4, and \( p_{x(\tau) | u} \) the conditional distribution of the states given the input sequence. Since in any case, the initial state is always known, the states are deterministic in \( p_{x(\tau) | u} \), which becomes a multivariate Dirac delta distribution imposing each state evolution equation (6.1) at \( t = 0 \ldots \tau - 1 \).

6.3.1 Detailed specification of the integrand: both factors in (6.8)

The first factor is given by

\[
p_u(u) = (\det 2\pi \Sigma_u)^{-1/2} \exp \left( -\frac{1}{2} (u - u_0)^T \Sigma_u^{-1} (u - u_0) \right) \tag{6.9}
\]

with deterministic part \( u_0^T = [ u_0(0) \ldots u_0(\tau - 1) ] \). From the stationarity of \( n_u(t) \), the variance-covariance matrix \( \Sigma_u \in \mathbb{R}^{\tau \times \tau} \) is a Toeplitz matrix with as signal the autocovariance of \( n_u(t) \), denoted as \( r_{uu} \), viz.

\[
\Sigma_{u,k,l} = \text{cov} (n_u(t), n_u(t + l - k)) = r_{uu}(l - k) \tag{6.10}
\]
This formulation can be transformed to the frequency domain as follows. From Gray (2005), it is known that a Toeplitz matrix has asymptotically (in the sense \( \tau \to \infty \)) the same properties as a circulant matrix whose signal (first row) is \( r_{uu} \). As a result, it follows that

\[
\Sigma \approx F^{-1}_\tau \text{diag}(F_\tau r_{uu})F_\tau \tag{6.11}
\]

with the autocovariance vector denoted as \( r_{uu} = [r_{uu}(t)]_{t=0...\tau-1} \), \( F_\tau = [\exp(-j\frac{2\pi}{\tau} kl)]_{k,l=0...\tau-1} \in \mathbb{C}^{\tau \times \tau} \) denoting the DFT matrix and “diag” the operator placing the vector elements on the respective diagonal positions (note that this result is similar to the so-called Whittle approximation as used in time-series analysis). Due to the Wiener-Khinchin theorem, we may replace \( F_\tau r_{uu} \), interpreted as the DFT of the autocovariance \( r_{uu} \), by \( S_{uu} = [S_{uu}(\omega_k)]_{k=0...\tau-1} = [\sigma_k^2]_{k=0...\tau-1} \). Using the unitarity of \( \frac{F_\tau}{\sqrt{\tau}} \), it follows that

\[
(u - u_0)^T \Sigma_u^{-1} (u - u_0) \approx \varepsilon^H \varepsilon \tag{6.12}
\]

with

\[
\varepsilon = \text{diag} \left( S_{uu}^{-1/2} \right) \frac{F_\tau}{\sqrt{\tau}} (u - u_0) \tag{6.13}
\]

This means that the signal \( u - u_0 \) has simply to be DFT-transformed and scaled frequency by frequency to obtain \( \varepsilon \), which can be interpreted as “driving” white noise.

The second factor is given by

\[
p_{x,x(\tau)|u} = \prod_{t=0}^{\tau-1} \delta (f(x(t), u(t)) - x(t+1)) \tag{6.14}
\]

with \( \delta(x) \) the multivariate Dirac delta distribution. Since a Dirac delta distribution only has nonzero contributions at the places where its argument is nonzero, one observes that this second factor introduces a constraint on the integration variables (states and inputs).

### 6.3.2 Performing the integration

The careful reader could wonder how the high-dimensional integration containing Dirac distributions in (6.5) can be performed. Let us first observe that (6.5) only has contributions at the places where the argument of each Dirac delta is zero, and consists of at least as many integration variables \((\tau - 1) n + \tau\) as constraints implied by these Dirac delta’s \((n \tau)\) if \( \tau \geq n \), which is the nontrivial case – otherwise the integral would be zero. If the constraints are nowhere satisfied (this situation corresponds to a non-controllability of the model), then the integral vanishes too.
We refer to Appendix 6.A on page 180, in which it is precisely set out how to proceed with the (approximate) integration of the integrand with a Dirac delta distribution of the form

\[ I = \int_{-\infty}^{+\infty} e^{-\gamma V(\eta)} \delta (\Psi (\eta)) \, d\eta \]  

(6.15)
making use of Laplace’s integration method (de Bruijn, 1970). For the reader not interested to dig into all technical details, in this section we restrict ourselves to mentioning the following: local linear and quadratic Taylor approximants are formed to Dirac’s operand \( \Psi (\eta) \) (corresponding to \( f(x(t), u(t)) - x(t + 1) \) in vector form) and the exponent \( V(\eta) \) (corresponding to the quadratic form (6.12), up to a scalar factor denoted \( \gamma \)) respectively. Then the integral is approximated by using the approximants instead of the functions themselves. Notice that the scaling of \( \gamma \) and \( V(\eta) \) in the left hand side of \( \gamma V(\eta) = (u - u_0)^T \Sigma_u^{-1} (u - u_0) \) is fixed by viewing \( \gamma \) as the inverse of the input variance \( \sigma^2_U \)

(6.16)
The local expansions can be motivated as follows. The integrand is, in the subspace satisfying Dirac’s constraint, dominant at the minimizer of \( V(\eta) \), denoted \( \eta^* \). A decrease to zero of the input variance, corresponding to an increase to infinity of the parameter \( \gamma \) in (6.15), tends to reduce the region of “activity” of the integrand to a small neighbourhood (on the constraint subspace) of the constrained minimizer of \( V(\eta) \). As a consequence, the integral approximation error tends to zero as the input variance tends to zero. Moreover, in this specific situation (integral (6.5)), due to the quadratic property of \( V(\eta) \), this error only comes from the local linearization of the nonlinear function in \( \Psi (\eta) \).

6.3.3 Interpretation

Interesting to report is the interpretation of the constrained minimizer \( \eta^* \)

\[ \eta^* = \arg\min_{\eta \atop \Psi(\eta) = 0} V(\eta) \]  

(6.17)
the neighbourhood of which is contributing the most to the integral representing \( x(\tau) \)’s pdf. The variable \( \eta \) corresponds to the integration variables, which are the concatenation of the states \( x \) and inputs \( u \). The constrained optimization problem, translated in these terms, reads

\[ \arg\min_{x,u \atop x(t + 1) = f(x(t), u(t))} (u - u_0)^T \Sigma_u^{-1} (u - u_0) \]  

(6.18)
Due to the fulfillment of the constraints, $\eta^*$ coincides with a state-input combination that satisfies the state equation at each time instant, and the state trajectory obtained starts at the initial state $x(0)$ and it ends at the argument $x(\tau)$ of the pdf $p_{x(\tau)}(x(\tau))$. Due to the minimization of the quadratic form, we conclude, in the Gaussian case, that the normalized input energy (normalized with the specified input power spectrum) of $\eta^*$ has to be minimal. In general, even if the pdf is not Gaussian, the input signal of $\eta^*$ can be interpreted as the most likely one (maximization of $p_u(u)$) that drives the state of $\eta^*$ on a trajectory from $x(0)$ to the point of evaluation of $p_{x(\tau)}$.

### 6.3.4 Putting theory into practice

Application of the results of Appendix 6.A to the estimation of (6.5) involves $n_\eta = (\tau - 1)n + \tau$ integration dimensions (related to the concatenation of $x$ and $u$), and $n_\Psi = n\tau$ constraint dimensions (related to the state evolution at each time instant). According to the method, the integration variables are to be split up in two parts, $\eta_1$ and $\eta_2$, the first of which is of dimension $n_\Psi$. The choice is made to let $\eta_1$ correspond to $x(1), \ldots, x(\tau - 1), u(0), \ldots, u(n - 1)$ and $\eta_2$ to $u(n), \ldots, u(\tau - 1)$. As a consequence, $J_1^*$ and $J_2^*$ consist of the respective columns of $\Psi$’s Jacobian, which is very sparse (note that by construction we obtain a full rank $J_1^*$ matrix, which is needed in Appendix 6.A). The constrained minimization (6.18) can be achieved via, e.g., Lagrange-Newton algorithms, and involves a Fast-Fourier-Transform (FFT) based implementation combined with sparse matrices. Not forgetting the first factor of (6.9), the approximation of the pdf’s logarithm based on the result of Appendix 6.A,

$$
\log \tilde{I} = -\log |\det J_1^*| - \gamma V(\eta^*) + \frac{\Delta n}{2} \log 2\pi - \frac{1}{2} \log |\det \gamma H_W| 
$$

(in which the symbols are defined in the Appendix), can be rewritten as:

$$
\log p_{x(\tau)}(x(\tau)) \approx -\frac{n}{2} \log 2\pi - \frac{1}{2} \log |\det \Sigma_u| - \log |\det J_1^*| - \frac{1}{2} (u - u_0)^T \Sigma_u^{-1} (u - u_0) - \frac{1}{2} \log |\det \gamma H_W| 
$$

### 6.3.5 Additional remarks

1. The last term of (6.19) can be computed more efficiently than via brute-force computation: $H_W$ is written as $M^T M$, with

$$
M = \begin{bmatrix}
0^{n(\tau-1)\times n(\tau-1)} & 0^{n(\tau-1)\times \tau} \\
0^{\tau\times n(\tau-1)} & \Sigma_u^{-1/2}
\end{bmatrix}
\begin{bmatrix}
-I_{\Delta n\times \Delta n} \\
Q^*
\end{bmatrix}
$$
$$\Sigma_u^{-\frac{1}{2}} = \text{diag} \left( S_u^{-\frac{1}{2}} \right) \frac{F}{\sqrt{\tau}}, \quad Q^* = J_1^*-1 J_2^* \quad \text{and} \quad \Delta n = \tau - n, \quad M \text{ is decomposed as } Q_M R_M, \quad \text{and as a consequence}$$

$$-\frac{1}{2} \log |\text{det} \gamma H_W| = -\frac{\Delta n}{2} \log \gamma - \log |\text{det} R_M|$$

$$= -\frac{\Delta n}{2} \log \gamma - \sum_r \log |R_{M,r,r}|$$

2. The pdf at another input variance can be obtained with almost no additional computational effort (the optimization problem, and so its solution, remain the same, just a slight modification of the constants in the pdf’s expression is needed at the places where a $\gamma$ or covariance matrix is present).

3. It can be noted that the intermediate states (for $1 \leq t \leq \tau - 1$) can be eliminated directly from the problem, both in the Delta-containing integral and in the constrained optimization problem. The result should be equivalent. Nevertheless, our present implementation does not use this elimination, but this large-size problem has the advantage to be sparse.

### 6.4 Generation of initial estimates for the constrained nonlinear optimization problem

Till here, the constrained optimization problem was formulated, without addressing the question of how to generate good initial estimates. Since the constraints are nonlinear, the iterative optimization procedure may end up in local optima. The closer the initial estimate is to the desired global optimum, the less likely the iterations will converge to spurious local optima. In this section, a few options allowing one to generate these initial estimates are discussed.

#### 6.4.1 Usage of a neighbouring solution

Since the state’s pdf is likely to be evaluated over a region, a grid can be used in the state space domain. At each grid point, the optimization has to be rerun. Advantage can be taken of smoothness of the solution w.r.t. the problem’s parameters: the solution $\eta^*$ of the closest neighbouring point can be used as an initial guess. This may also be a way to save computation time, reducing the number of iterations.
6.4.2 Extrapolation on the state-space grid

Another option is to use a (linear) extrapolation of the solution from neighbouring grid points. In a 2-dimensional state space, 3 non-collinear points define a plane, which can be used to estimate a 4th grid point. In an \( n \)-dimensional state space, \( n + 1 \) linear independent points define a hyperplane, which can be used to obtain an initial guess for the solution at a grid point nearby.

6.4.3 Solution in the linear case

A third option is to take advantage from the fact that the linear case can be completely worked out analytically. The solution can, in that case, be calculated in one step using linear matrix algebra. The initial estimates consist of neglecting the nonlinear terms and using the analytic solution to start the optimization from. The technical details can be found in Appendix 6.B on page 183, together with an interpretation adding insight into the problem.

6.5 Accuracy of the pdf estimation

One may wonder how well the proposed LIM-based method approximates the exact pdf. A complete theoretical analysis is provided in Appendix 6.C on page 186, with an analytic expression of the first order approximation of the relative error (for \( \sigma_U \to 0 \), i.e. \( \gamma \to \infty \)). Clearly, as expected, the approximation quality of the linear Taylor approximant to the state-space equations determines the approximation quality of the pdf. In the linear case, the LIM will introduce no integration error. In the general case, the magnitude of the second derivatives determines the accuracy.

6.6 Simulation results

The present implementation is based on the Matlab’s nonlinear constrained optimization routine \texttt{fmincon}, and uses an interior point algorithm (iteration step computed via conjugated gradients, Hessian calculated via finite differences, analytically calculated gradients of cost and constraint functions, and a constraint tolerance of \( 10^{-13} \)). The method is illustrated by means of 2-dimensional examples, since this allows graphical representations of the pdf.

The following continuous-time state equation, formulated in polar coordinates \((\rho, \phi)\), is considered:

\[
\dot{\rho} = -\rho \left( 1 - \rho^2 \right) \quad (6.21)
\]

\[
\dot{\phi} = 1 \quad (6.22)
\]

\[
\dot{\rho} = -\rho \left( 1 - \rho^2 \right)
\]

\[
\dot{\phi} = 1
\]
Figure 6.2: Evolution of the radial coordinate: the sign of $\dot{\rho}$ (changing at $\rho = 1$), determines the attraction to, or repulsion from the equilibrium.

It has as ROA the region $\rho < 1$ located inside the unit circle (all initial states inside get attracted to the origin, while those starting outside get repelled, towards infinity). This is easily seen from the sign of the function $-\rho (1 - \rho^2)$ changing at $\rho = 1$, and the decoupling of both coordinates (see Figure 6.2). The phase portrait (collection of unforced trajectories) is shown in Figure 6.3. An unstable limit cycle is obtained.

Writing in Cartesian coordinates

$$x = \begin{bmatrix} \rho \cos \phi \\ \rho \sin \phi \end{bmatrix},$$

and adding a linear input-term, one obtains

$$\dot{x} = -x (1 - x^T x) + \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} x + Bu$$

Inspired by this continuous-time system, the discrete-time system given by Euler discretization of the former is introduced:

$$f(x, u) = x + T_s \left( -x (1 - x^T x) + \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} x + Bu \right)$$

with $T_s = 0.01$. We let the method estimate the pdf for $\tau = 300$, a Gaussian input with time-domain standard deviation $\sigma_U = 3$ (white power spectrum), $x(0) = \begin{bmatrix} 0 & 0 \end{bmatrix}^T$, $u_0(t) = 0 (\forall t = 0 \ldots \tau - 1)$, $B = \begin{bmatrix} 1 & 1 \end{bmatrix}^T$. The resulting
Figure 6.3: Phase portrait of the autonomous system under consideration (dashed lines). The states rotate counterclockwise. The ROA is the region located inside the unit circle (bold curve). Initial states inside the unit disk get attracted towards the origin, while initial states starting outside tend towards infinity.
Chapter 6. Instability under random inputs: model-based perspective

Figure 6.4: Trajectory corresponding to $\eta^*$ (the most likely path from $x(0)$ to $x(\tau)$), needed for the calculation of $p_{x(\tau)}$ at $x(\tau)$. The ROA is indicated in grey, and the dashed lines represent the unforced trajectories.

state- and input-trajectories corresponding to $\eta^*$ for $x(\tau) = \begin{bmatrix} -0.8621 & 0 \end{bmatrix}^T$ are shown in Figures 6.4 and 6.5 respectively.

A grid over the ROA (which is here the unit disk), parameterized in polar coordinates, was chosen with 30 linearly spaced $\rho$-values and 30 linearly spaced $\phi$-values. Each grid point requires a constrained optimization.

In the next subsections, the results will be shown:

- for the described nonlinear model at $\sigma_U = 3$;
- for the underlying linear model at $\sigma_U = 3$;
- for both of the above, extending to several $\sigma_U$-values, and comparing with Monte-Carlo simulations;
- for the nonlinear model with colored input power spectrum at $\sigma_U = 3$;
- for the nonlinear model with an initial state different from the origin, and a deterministic input component at $\sigma_U = 3$;
- for the nonlinear model extended with additional nonlinear terms in the input at $\sigma_U = 3$. 
6.6. Simulation results

Figure 6.5: Most likely input corresponding to $\eta^*$ and the trajectory of Figure 6.4, needed for the calculation of $p_{x(\tau)}$ at $x(\tau)$.

As a last section, we will briefly compare with the Fokker-Planck approach to obtain the state’s pdf.

The results at the fixed value $\sigma_U = 3$, to be discussed in the upcoming subsections, are summarized in Table 6.1 on the following page.

6.6.1 Nonlinear model equations

Figures 6.6, 6.7 and 6.8 show respectively the pdf, its (naperian) logarithm, and the location of the state $x(\tau)$ for 1500 Monte-Carlo runs (using a Gaussian input satisfying the assumptions) shown on top of the pdf logarithm’s contour plot. The pdf shows how the states $x(\tau)$ are distributed at time $\tau$. The distribution of the random Monte-Carlo realizations of the states is in good agreement with the estimated pdf. From the last two figures, it is seen that the pdf deviates from a Gaussian distribution and the contour lines are not elliptical (the pdf’s logarithm clearly deviates from a paraboloid). This indicates that, although a linearization took place in the method, the method still grasps some nonlinear behaviour. The method allows one to predict which regions are less likely locations of the state $x(\tau)$, without requiring extensive simulation efforts as would be the case with a purely Monte-Carlo approach. Based on this mesh, a 2-dimensional Simpson integration of the pdf over the ROA returns as prob-
Chapter 6. Instability under random inputs: model-based perspective

\[ P(x(\tau)) \in \text{ROA} \]

Integration of approximated Monte-Carlo (95\% confidence) under linear model with colored input power spectrum:
- 99.97\% ± 0.3\%
- 99.93\% ± 0.13\%

Table 6.1: Summarizing comparison between the proposed method and Monte-Carlo simulation results.

<table>
<thead>
<tr>
<th>and ( f = \begin{bmatrix} 1 &amp; 0.5 \ 0 &amp; 1 \end{bmatrix} )</th>
<th>Nonlinear terms in the input</th>
<th>non-Linear model appended with nonlinear terms in the input</th>
</tr>
</thead>
<tbody>
<tr>
<td>92.0% ± 0.24%</td>
<td>94.3% ± 0.4%</td>
<td>94.3% ± 0.4%</td>
</tr>
<tr>
<td>53.6% ± 4.4%</td>
<td>57.0% ± 0.4%</td>
<td>57.0% ± 0.4%</td>
</tr>
<tr>
<td>16.6% ± 3.3%</td>
<td>18.4% ± 3.3%</td>
<td>18.4% ± 3.3%</td>
</tr>
<tr>
<td>96.9% ± 0.3%</td>
<td>99.9% ± 0.13%</td>
<td>99.9% ± 0.13%</td>
</tr>
<tr>
<td>98.2% ± 0.84%</td>
<td>99.9% ± 0.3%</td>
<td>99.9% ± 0.3%</td>
</tr>
</tbody>
</table>

\( x(\tau) = 0.5 \cos(\pi \tau t) \) and initial conditions \( x(0) = 0 \), \( u(t) = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \).
6.6. Simulation results

Figure 6.6: Estimated pdf $p_{x(\tau)}$.

Figure 6.7: (Naperian) logarithm of the estimated pdf $p_{x(\tau)}$. 
Figure 6.8: Contour lines of the logarithm of $p_{x(\tau)}$, with a scatter plot of random Monte-Carlo values of the state $x(\tau)$ indicated with crosses. Notice the excellent agreement between both. The major axis of the approximately elliptical contour lines around the origin is slightly rotated compared with the diagonal axis along which the input enters into the state ($B = \begin{bmatrix} 1 & 1 \end{bmatrix}^T$); this is due to the natural rotation in the phase plane induced by $\dot{\phi} = 1$. 
ability of bounded operation over a time interval of length $\tau$ ($\mathbb{P}(x(\tau) \in \text{ROA})$) the value 98.53% (the value of 98.63% was obtained at a two times coarser grid, indicating that the order of magnitude of the integration error is 0.1%). This can be compared with the value 97.27% (with estimated standard deviation of 0.42%) obtained from the fraction of Monte-Carlo simulations resulting in a state $x(\tau)$ in the ROA. The 95% confidence-interval does not contain the estimated value, indicating a small mismatch between both values. This has probably to do with the violation of the assumption that the input should have a small standard deviation. Anyway, the method clearly succeeds in providing an order of magnitude of the risk of leaving the ROA, which is the complement of the above-mentioned probability values: 1.47\%±0.1\% against 2.73\%±0.84\% (2\sigma-confidence).

In addition, it can be mentioned that the whole procedure on this mesh (run on a PC with a 2.0 GHz Intel Xeon processor and 4.0 GB RAM), takes less than 8 minutes of computation time.

### 6.6.2 Underlying linear model

These results have been compared with what would be obtained by using the underlying linear model (which is the linearization at the operating point zero). The linear counterparts to Figures 6.6, 6.7 and 6.8 are respectively Figures 6.9, 6.10 and 6.11. Notice the pdf’s Gaussian shape, which has also been confirmed by fitting a paraboloid on the logarithm (both surfaces coincide). The resulting probabilities $\mathbb{P}(x(\tau) \in \text{ROA})$ are given by: 99.97\% for the estimated probability via the method and 99.93\%±0.13\% (2\sigma-confidence) for the Monte-Carlo simulation. A severe underestimation of the risk (of a factor 100: 0.03\% estimated instead of 2.73\%) hence takes place when a simple linear analysis is performed.

### 6.6.3 Gathering information at other input standard deviation values

Another advantage of the method is that it allows one to estimate the probability of stable operation also at other input standard deviation values at once, without needing to solve additional constrained optimization problems (just the constants in the pdf expressions are changing). Figure 6.12 on page 170 shows the estimated probabilities of stable operation, both in the nonlinear case and in the linear case, and compared with the results of 1500 Monte-Carlo simulations of length $\tau$. The limitation of this method is as follows: at small $\sigma_U$-values a discrepancy occurs between the true probabilities (very close to 1) and the estimated one (cf. the part of the figure corresponding to $\sigma_U < 2$).
Figure 6.9: Estimated pdf $p_{x(\tau)}$ (for the underlying linear system), coinciding with a Gaussian pdf (not visible in the graph).

Figure 6.10: (Naperian) logarithm of the estimated pdf $p_{x(\tau)}$ (for the underlying linear system), coinciding with a paraboloid (not visible in the graph).
Figure 6.11: Contour lines of the logarithm of $p_{x(\tau)}$ (for the underlying linear system), with a scatter plot of random Monte-Carlo values of the state $x(\tau)$ (for the underlying linear system) indicated with crosses. Notice the excellent agreement between both, and the typical elliptical contour shape of Gaussian pdfs.
Chapter 6. Instability under random inputs: model-based perspective

Figure 6.12: Estimated probabilities of bounded operation over a time interval of length $\tau$, both in the nonlinear case and in the linear case. Notice that the part of the graph below $\sigma_U = 2$ is not reliable, since the integration grid has not been modified, while the pdf concentrates more around the origin (serious integration errors are taking place). Notice the underestimation of the risk of leaving the ROA when the underlying linear model is used. The results are compared with the fraction of 1500 Monte-Carlo runs of length $\tau$ resulting in a bounded behaviour (grey line; the 95% confidence interval is shown with a dashed grey line).

This is due to the integration grid, which is not well matched to the fast variations of the pdf around the origin at low $\sigma_U$-values. In the region $2 < \sigma_U < 4.5$, a reasonable approximation is obtained (see Figure 6.13 on the facing page).

6.6.4 Nonlinear model with colored input power spectrum

The results for a colored Gaussian input with power spectrum corresponding to the linear filter $\frac{1}{1-0.8z^{-1}}$, renormalized to a time domain standard deviation of $\sigma_U = 3$, are shown in Figures 6.14, 6.15 and 6.16. The corresponding probability estimates $P(x(\tau) \in \text{ROA})$, shown in Table 6.1 on page 164, can be remarked to be much lower. This can be explained intuitively as follows: the input varies more slowly in time, and hence, whenever the state is pushed
6.6. Simulation results

![Graph showing probability of bounded operation over a time interval of length \( \tau \) vs. \( \sigma_U \).](image)

Figure 6.13: Detail of Figure 6.12.

in the wrong direction, this is lasting longer, leading to an increased risk to leave the ROA. This can be compared with the white noise situation, in which the fast variations of the excitation are averaged out in the system itself. Due to the slow fluctuations of the input, larger deviations from the origin create a “flattened” and more “stretched” pdf (as can be seen from the smaller pdf values in Figure 6.14) compared with the white noise situation (see Figure 6.6). These larger deviations cause an increase of the level of nonlinearities, and hence of the second derivatives, which are a measure of the errors on the estimated pdf. This explains the discrepancy between the estimated probability and the fraction of bounded Monte-Carlo simulations in Table 6.1.

6.6.5 Nonlinear model with an initial state different from the origin, and a deterministic input component

The same is carried out for the situation with an initial state different from the equilibrium point at the origin:

\[
x (0) = \begin{bmatrix} 0.5 \\ 0 \end{bmatrix}^T
\]

and a non-zero deterministic input component

\[
u_0 (t) = \frac{1}{2} \cos \left( \frac{\pi t}{\tau} \right)
\]
Chapter 6. Instability under random inputs: model-based perspective

Figure 6.14: Estimated pdf $p_x(\tau)$ (colored input power spectrum).

Figure 6.15: (Naperian) logarithm of the estimated pdf $p_x(\tau)$ (colored input power spectrum).
6.6. Simulation results

Figure 6.16: Contour lines of the logarithm of \( p_x(\tau) \) (colored input power spectrum), with a scatter plot of random Monte-Carlo values of the state \( x(\tau) \) indicated with crosses.

This time, the integration grid is centered around the nominal state at time \( t = \tau \). The result is shown at Figures 6.17, 6.18, and 6.19. Figure 6.19 shows also the deterministic, nominal trajectory. Note that the nonlinearity causes the mode of the pdf to depart from the nominal state at time \( t = \tau \). As can be seen from the corresponding row of Table 6.1 on page 164, the estimated probability falls within the 95%-confidence interval.

6.6.6 Nonlinear model appended with nonlinear terms in the input

In this last example, the nonlinear state equation is appended with nonlinear terms of the form \( u^2 \) and \( ux \). More precisely, the following model equation is being considered:

\[
f(x, u) = x + T_s \left( -x (1 - x^T x) + \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} x + Bu + B_2 u^2 + \kappa xu \right)
\]

with \( B_2 = 10^{-2} \begin{bmatrix} 1 & -2 \end{bmatrix}^T \) and \( \kappa = 1 \). A 50-times-50 grid in radial and angular directions is being used in this situation. The results are shown in
Chapter 6. Instability under random inputs: model-based perspective

Figure 6.17: Estimated pdf $p_{x(\tau)}$ (non-zero initial state and non-zero deterministic input component).

Figure 6.18: (Naperian) logarithm of the estimated pdf $p_{x(\tau)}$ (non-zero initial state and non-zero deterministic input component).
6.6. Simulation results

Figure 6.19: Contour lines of the logarithm of $p_x(\tau)$ (non-zero initial state and non-zero deterministic input component), with a scatter plot of random Monte-Carlo values of the state $x(\tau)$ indicated with crosses. The nominal trajectory (input purely deterministic) is shown in green.

Figures 6.20, 6.21, and 6.22. At the states where the constraints could not reasonably be satisfied (no feasible solution found: error on constraints larger than $10^{-9}$; this has to do with a lack of controllability from the initial state to that location), the pdf is manually set to zero (this explains the gaps in the log-pdf mesh in Figure 6.21). The resulting probability values agree up to a few percent (see Table 6.1 on page 164, this may be due to luck), although there is a disagreement between the pdf and the distribution of the random realizations in Figure 6.22. The higher level of the second order derivatives compared with the original situation (while $\sigma_U$ is kept unchanged), may explain an increase in integration error.

6.6.7 Comparison with the Fokker-Plank approach

The present results could be compared with the solution of the so-called Fokker-Plank equation, which dictates how a particle (or state) probability distribution evolves over time, when it is subjected to an input-affine stochastic differential equation. In the scalar case, for the following stochastic state equation

$$\dot{x} = f(x) + \sigma(x)u$$

(6.23)
Chapter 6. Instability under random inputs: model-based perspective

Figure 6.20: Estimated pdf $p_{x(\tau)}$ (additional nonlinear terms in the input).

Figure 6.21: (Naperian) logarithm of the estimated pdf $p_{x(\tau)}$ (additional nonlinear terms in the input).
6.6. Simulation results

Figure 6.22: Contour lines of the logarithm of $p_x(\tau)$ (additional nonlinear terms in the input), with a scatter plot of random Monte-Carlo values of the state $x(\tau)$ indicated with crosses.
with \( u \) unit-variance continuous-time white Gaussian noise, the Fokker-Planck equation (Risken, 1996)

\[
\frac{\partial p(x,t)}{\partial t} = -\frac{\partial (f(x)p(x,t))}{\partial x} + \frac{1}{2} \frac{\partial^2 (\sigma(x)p(x,t))}{\partial x^2}
\]  

(6.24)

is the parabolic partial differential equation in the state’s pdf \( p(x,t) \) as function of time. It can be easily verified that the stationary solution of (6.24) (which becomes an ordinary differential equation, since the pdf then satisfies \( \frac{\partial p(x,t)}{\partial t} = 0 \)) is given by

\[
p(x) = \frac{c}{\sigma^2(x)} \exp \left( \int_0^x \frac{2f(y)}{\sigma^2(y)} dy \right)
\]  

(6.25)

Herein \( c \) is a normalization constant scaling the pdf to unit integral. This solution is known to be the single positive solution.

Now, the original continuous-time system (6.21) – from which the simulated discrete-time system is originating – is considered in the white noise case. We restrict ourselves to the one-dimensional equation of the radial coordinate subject to additive (state-independent) Gaussian noise. This yields:

\[
p(x) = \frac{c}{\sigma^2} \exp \left( \int_0^x \frac{2}{\sigma^2} \left( -\frac{x^2}{2} + \frac{x^4}{4} \right) dy \right)
\]  

The function \( \exp \left( -\frac{2}{\sigma^2} \left( \frac{x^2}{2} - \frac{x^4}{4} \right) \right) \) is shown in Figure 6.23 for \( \sigma = 1 \). As a result, the stationary solution is unbounded as \( x \to \infty \), and its integral diverges. This is a degenerate solution. However, since the function \( \exp \left( -\frac{2}{\sigma^2} \left( \frac{x^2}{2} - \frac{x^4}{4} \right) \right) \) is bounded in the ROA (\( |x| < 1 \)), as expected, in stationary regime (at time infinity), the state is located outside the ROA with probability 1. The (transient) solution of the full Fokker-Planck partial differential equation at any finite time would not be degenerate (it can be obtained via extensive numerical computations), and should be in agreement with the previous results (one-dimensional version of Figure 6.6). More precisely, using the Dirac initial condition for the state \( p(x,0) = \delta(x-x(0)) \), the comparison should be done at the corresponding time instant \( \tau T_s = 3 \).

6.7 Conclusions

In this chapter, given the input’s deterministic part, the power spectrum of the Gaussian part, the analytic form of the state evolution equation, and the initial
6.7. Conclusions

Figure 6.23: Stationary solution (non-normalized) to the Fokker-Planck equation for a similar one-dimensional system driven by white Gaussian noise.

state, the pdf of a future state was calculated as an integral and using Bayes’ rule. If the input variance is small, the (very high dimensional) integral can be well approximated at any point of the state space via the Laplace integration method. This, in turn, requires the solution of a constrained optimization problem, which can be interpreted as the most likely scenario (trajectory) driving the state from its initial position to the pdf’s argument. The constrained optimization involves sparse matrices. A few options for the generation of initial estimates to the nonlinearly constrained optimization problem have been presented. The problem setting (model equation, input noise and initial state) can be extended and generalized in different ways. A theoretical analysis of the integration error shows that it essentially depends on the magnitude of second derivatives of the state equation. If the input is a finite-time burst, one obtains the probability of getting asymptotically attracted towards the origin via the integral of the pdf over the region of attraction of the unforced system. The proposed method provides more insight and is more efficient to estimate the pdf than a brute-force simulation approach, especially in the regions where the pdf is very low, since this could require a huge number of simulations to get an accurate result. Another advantage compared with the Monte-Carlo approach, is that the pdf computations at other input standard deviations can be performed at once, without needing to solve additional constrained optimization problems. Simulation results are shown to illustrate and support the theory. Although the estimated probability of bounded operation is not always per-
fectly correct, due to the fact that the input standard deviation is not (very) small as assumed in the theory, a correct order of magnitude is obtained, outperforming a classical analysis of the underlying linear dynamics, which would result in a severe underestimation of the risk. The optimization of the choice of the integration method and of the density of the integration grid, is left for future research.

Appendices

6.A Integration approach

6.A.1 Goal and workflow

Let us consider the more general problem of approximating the integral

$$I = \int_{-\infty}^{+\infty} e^{-\gamma V(\eta)} \delta (\Psi (\eta)) d\eta$$

(6.26)

for $\gamma(> 0) \to \infty$, with $\eta \in \mathbb{R}^{n_{\eta}}$, $V : \mathbb{R}^{n_{\eta}} \to \mathbb{R}$, $\Psi : \mathbb{R}^{n_{\eta}} \to \mathbb{R}^{n_{\Psi}}$ and with $n_{\eta} \geq n_{\Psi}$, the aim being to set up its first order approximation of a series expansion in $\gamma$ around the value $\gamma = \infty$. Since the integrand is only “active” (viz. nonzero) at the solution set of $\Psi (\eta)$, and $\gamma$ is large, the dominant contributions to $I$ are coming from the neighborhood of the minimizer of $V$ (maximizer of $\exp (-\gamma V)$) on this solution set. Let us introduce

$$\eta^* = \arg\min_{\eta} V (\eta) \quad \text{s.t. } \Psi (\eta) = 0$$

(6.27)

As $\gamma$ increases towards infinity, the size of the neighborhood of $\eta^*$ contributing most to $I$ decreases. Therefore, a good approximation can still be obtained by integrating $\exp (-\gamma V (\eta))$ on the tangent subspace to $\Psi (\eta) = 0$ at $\eta^*$. This is schematically illustrated for $n_{\eta} = 2$ and $n_{\Psi} = 1$ in Figure 6.24.

To do so, a linearized version of $\Psi (\eta)$ has to be constructed in order to obtain an approximate linear relationship among the integration variable’s components. This will allow us to eliminate the $\delta$-distribution and proceed with the Laplace integration method. The idea is then to approximate the new (unconstrained) integrand as a multivariate Gaussian pdf (whose integral is known analytically) by performing a quadratic approximation of its exponent.

\footnote{The parameter $\gamma$ tends to infinity here. How large $\gamma$ should be in practice, has to be determined based on the integration approximation error (which is the series’ second term, i.e. the first order correction to the approximation), to be studied in Appendix 6.C on page 186.}
6.A. Integration approach

Figure 6.24: Since $\gamma$ is large, the immediate neighborhood of $\eta^*$ on the solution set of $\Psi(\eta)$ (represented by a solid line) contributes most to the integral. The contour lines of $V(\eta)$ are shown in grey. A tractable approximation of the integral will be obtained via a linearization of $\Psi(\eta)$ at $\eta^*$; the integration will take actually place along the hyperplane (represented by a dashed line) defined as the solution set of the Taylor approximant $\tilde{\Psi}(\eta)$ to $\Psi(\eta)$.

Applied to the pdf integral $p_{x(\tau)}$ (6.5), loosely spoken, the procedure implicitly imposes a compatibility constraint between the inputs and the states (forming the integration variables together). Weighting with the function $e^{-\gamma V(\eta)}$, corresponds to the likelihood that a given trajectory occurs. The behaviour in the neighbourhood of the most likely trajectory (coinciding with $\eta^*$) is going to determine $x(\tau)$’s pdf $p_{x(\tau)}$.

6.A.2 Elimination of the constraints

Mathematically speaking, a linear Taylor approximation of $\Psi(\eta)$ is constructed at $\eta^*$

$$
\tilde{\Psi}(\eta) = \frac{\partial \Psi}{\partial \eta} \bigg|_{\eta^*} (\eta - \eta^*) = \frac{\partial \Psi}{\partial \eta} \bigg|_{\eta^*} \bar{\eta},
$$

(6.28)

(with $\bar{\eta} = \eta - \eta^*$) and $\Psi(\eta)$ is replaced by $\tilde{\Psi}(\eta)$ in (6.26). Now $\eta$ is partitioned as $\begin{bmatrix} \eta_1 \\ \eta_2 \end{bmatrix}$ (and similarly for $\bar{\eta}$ and $\eta^*$), with $\eta_1 \in \mathbb{R}^{n*}$, the scalar function $V(\eta) = V(\eta_1; \eta_2)$, and $\Psi$’s Jacobian accordingly $\frac{\partial \Psi}{\partial \eta} \bigg|_{\eta^*} = \begin{bmatrix} J_1^* & J_2^* \end{bmatrix}$, with
$J_1^* \in \mathbb{R}^{n_\phi \times n_\phi}$ and $J_2^* \in \mathbb{R}^{n_\phi \times \Delta n}$ (with $\Delta n = n_\eta - n_\phi$). This yields

$$\tilde{\Psi} (\eta) = J_1^* \bar{\eta}_1 + J_2^* \bar{\eta}_2$$  \hfill (6.29)

Applying the property of multivariate delta distributions

$$\delta (A x + b) = |\det A|^{-1} \delta (x + A^{-1} b)$$

for a regular matrix $A$ (which is a generalization of the univariate property

$$\delta (a x) = |a|^{-1} \delta (x)$$

applied with $A = J_1^*$, yields

$$\delta \left( \tilde{\Psi} (\eta) \right) = \delta (J_1^* \bar{\eta}_1 + J_2^* \bar{\eta}_2) = |\det J_1^*|^{-1} \delta (\bar{\eta}_1 + Q^* \bar{\eta}_2)$$  \hfill (6.30)

with $Q^* = J_1^{-1} J_2^*$. After changing the integration variable to $\bar{\eta}$, we can go back to $I$, and apply the $\delta$-sifting property over the variable $\bar{\eta}_1$. This yields as approximation of $I$, in which the integration dimensionality has been reduced ($\bar{\eta}_2$ is the only free variable remaining):

$$\tilde{I} = |\det J_1^*|^{-1} \int_{-\infty}^{+\infty} e^{-\gamma W(\bar{\eta}_2)} d\bar{\eta}_2$$  \hfill (6.31)

with

$$W (\bar{\eta}_2) = \left. V (\eta_1; \eta_2) \right|_{\bar{\eta}_1 + Q^* \bar{\eta}_2 = 0} = \left. V (\eta_1^* - Q^* \bar{\eta}_2; \eta_2^* + \bar{\eta}_2) \right|_{\bar{\eta}_1 + Q^* \bar{\eta}_2 = 0}$$

6.A.3 Application of the Laplace integration

Now we are ready to proceed with the standard Laplace integration via a quadratic Taylor approximant $\tilde{W} (\bar{\eta}_2)$ of $W (\bar{\eta}_2)$ at $\bar{\eta}_2 = 0$:

$$\tilde{W} (\bar{\eta}_2) = V (\eta^*) + \frac{1}{2} \bar{\eta}_2^T H_W \bar{\eta}_2$$  \hfill (6.32)

with a short calculation revealing the expression of the Hessian

$$H_W = \frac{\partial^2 W}{\partial \bar{\eta}_2^2} = \left[ \begin{array}{c} -Q^* \\ I_{\Delta n \times \Delta n} \end{array} \right]^T \frac{\partial^2 V}{\partial \eta^2} \left[ \begin{array}{c} -Q^* \\ I_{\Delta n \times \Delta n} \end{array} \right]$$

(to be evaluated at $\eta = \eta^*$). Note that the linear term vanishes per construction of $\eta^*$ (it is a maximum of $V (\eta)$ along the constraint, which has been imposed in the previous subsection). Replacing $W$ by $\tilde{W}$ (6.32) in (6.31), and working out the integral over the multivariate normal density leads to:
6.B Analytic solution in the linear white noise situation

\[ \tilde{I} = |\text{det } J_1^*|^{-1} \int_{-\infty}^{+\infty} e^{-\gamma(V(\eta^* + \frac{1}{2}\eta_2^T H W \eta_2)}} d\tilde{\eta}_2 \]

\[ = |\text{det } J_1^*|^{-1} e^{-\gamma V(\eta^*)} \int_{-\infty}^{+\infty} e^{-\frac{1}{2} \eta_2^T (\gamma H W) \eta_2} d\tilde{\eta}_2 \]

where the integral is nothing but the normalization constant of a multivariate normal distribution with covariance matrix \((\gamma H W)^{-1}\), hence

\[ \tilde{I} = |\text{det } J_1^*|^{-1} e^{-\gamma V(\eta^*)} \left| \text{det } (2\pi(\gamma H W)^{-1}) \right|^{1/2} \]

Taking the logarithm yields

\[ \log \tilde{I} = -\log |\text{det } J_1^*| - \gamma V(\eta^*) + \frac{\Delta n}{2} \log 2\pi - \frac{1}{2} \log |\text{det } \gamma H W| \quad (6.33) \]

which concludes the calculation. \(\square\)

6.B Analytic solution in the linear white noise situation

In this appendix, the theoretical result is first stated, next interpreted and finally proven.

6.B.1 Statement of the theoretical result

Defining the nominal state at time \(\tau\) as the one which results from the deterministic inputs applied to the initial state,

\[ x_{\text{nom}}(\tau) = A^T x(0) + \sum_{r=0}^{\tau-1} A^{r-1} B u_0(\tau - r) \quad (6.34) \]

it will be shown in subsection 6.B.3 that the optimal input samples (the solution trajectory with the smallest value of \(V(\eta)\)) are given by:

\[ u(t) = u_0(t) - \frac{1}{2} B^T (A^T)^{\tau-1-t} \rho \quad (6.35) \]

where \(\rho\) solves

\[ -\frac{1}{2} X \rho = x(\tau) - x_{\text{nom}}(\tau) \quad (6.36) \]
and herein $X$ is the matrix defined as

$$X = \sum_{r=0}^{\tau-1} A^r B B^T (A^T)^r$$

(6.37)

and also satisfying the discrete Lyapunov equation

$$A X A^T - X = A^\tau B B^T (A^T)^\tau - B B^T$$

(6.38)

### 6.B.2 Interpretation of the result

The above can be interpreted as follows: the right-hand side of (6.36) expresses the mismatch between the position $x(\tau)$ and the “natural” position $x_{\text{nom}}(\tau)$ resulting from applying the forcing input $u_0(t)$ starting from $x(0)$. The larger this mismatch, the larger the resulting $\rho$, the latter being a measure for deviations between the nominal input $u_0(t)$ and the solution’s optimal input. Under stability, $A$’s eigenvalues (the poles) have magnitudes smaller than 1. Therefore, the right-hand side of (6.38) reduces to its second term as $\tau$ grows large. Under instability, at least one pole has magnitude larger than 1, and the first term tends to dominate over the second one, $X$ grows large. A short calculation, using (6.35) and (6.37), reveals

$$(u - u_0)^T (u - u_0) = \frac{1}{4} \rho^T X \rho = \left( \frac{X \rho}{2} \right)^T X^{-1} \frac{X \rho}{2}$$

combining with (6.36), this can also be rewritten as

$$(x(\tau) - x_{\text{nom}}(\tau))^T X^{-1} (x(\tau) - x_{\text{nom}}(\tau))$$

Since the other terms in the approximation of the pdf’s logarithm $\log p_x(\tau)(x(\tau))$ (6.20) do not contain $x(\tau)$, it equals (up to a constant term)

$$-\frac{1}{2} (x(\tau) - x_{\text{nom}}(\tau))^T (\sigma^2_{U} X)^{-1} (x(\tau) - x_{\text{nom}}(\tau))$$

Therefore, as expected, the state $x(\tau)$ is found to be normally distributed with mean $\mu_x(\tau) = x_{\text{nom}}(\tau)$ (default, nominal trajectory) and covariance matrix $\Sigma_x(\tau) = \sigma^2_U X$ (the larger the input variance, the larger the state deviations; in an unstable situation, $X$ also grows very large). Note that, for a stable system, taking the covariance of both sides of the state equation $x(t+1) = A x(t) + B u(t)$ and expressing stationarity for large $\tau$ (viz. $\text{cov} x(t) = \text{cov} x(t+1) = \Sigma_x(\tau)$) yields

$$\Sigma_x(\tau) = A \Sigma_x(\tau) A^T + B \sigma^2_U B^T$$

which is in agreement with (6.38), since the first term in (6.38) vanishes.
6.B. Analytic solution in the linear white noise situation

6.B.3 Proof

Using the Lagrangian, the optimization problem can be reframed as:

\[ \mathcal{L} = (u - u_0)^T (u - u_0) + \sum_{t=0}^{\tau-1} \lambda_t^T (Ax(t) + Bu(t) - x(t+1)) \]

with \( \lambda_t \in \mathbb{R}^n \) representing the Lagrange multipliers related to the time instant \( t \). Transposition of the derivatives w.r.t. \( u(t) \), \( x(t+1) \) and \( \lambda_t \) yields respectively:

\[ 2 (u(t) - u_0(t)) + B^T \lambda_t = 0 \quad t = 0 \ldots \tau - 1 \quad (6.39) \]
\[ A^T \lambda_{t+1} - \lambda_t = 0 \quad t = 0 \ldots \tau - 2 \quad (6.40) \]
\[ Ax(t) + Bu(t) - x(t+1) = 0 \quad t = 0 \ldots \tau - 1 \quad (6.41) \]

Setting \( \lambda_{\tau-1} = \rho \), \( \lambda_t \) can be eliminated from (6.39) using (6.40) as

\[ \lambda_t = (A^T)^{\tau-1-t} \rho \]

and (6.35) follows. Combining into (6.41), this results into

\[ Ax(t) + Bu_0(t) - \frac{1}{2} BB^T (A^T)^{\tau-1-t} \rho = x(t+1) \]

Therefore, applying this multiple times, the intermediate states \( x(t), t = 1 \ldots \tau - 1 \) can be eliminated:

\[ x(\tau) = Ax(\tau - 1) + Bu_0(\tau - 1) - \frac{1}{2} BB^T \rho \]
\[ = A^2 x(\tau - 2) + ABu_0(\tau - 2) - \frac{1}{2} ABB^T A^T \rho \]
\[ + Bu_0(\tau - 1) - \frac{1}{2} BB^T \rho \]
\[ \vdots \]
\[ = A^\tau x(0) + \sum_{r=0}^{\tau-1} A^{r-1} Bu_0(\tau - r) - \frac{1}{2} \sum_{r=0}^{\tau-1} A^{r} BB^T (A^T)^r \rho \]

Taking (6.37) and (6.34) into account, yields (6.36). Note that \( X \) (6.37) is easily seen to verify the discrete Lyapunov equation (6.38), which is more efficient to use for practical implementation purposes than the summation of (6.37). \( \square \)
6.C Approximation of the integral approximation error

To obtain an approximation of the integration error induced by the modified Laplace integration method (LIM), the difference between the integral (6.26) and its approximation (in which the linear approximant \( \tilde{\Psi} (\eta) \) to \( \Psi (\eta) \) is used) is considered:

\[
\Delta I = \int_{-\infty}^{+\infty} e^{-\gamma V(\eta)} \left[ \delta (\Psi (\eta)) - \delta (\tilde{\Psi} (\eta)) \right] d\eta
\]  

(6.42)

We will express the Dirac distributions as the limit of normal densities – in the univariate case, one has

\[
\delta (x - x_0) = \lim_{\sigma^2 \to 0} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-x_0)^2}{2\sigma^2}}
\]

The expression inside the square brackets of 6.42 will now be rewritten in terms of the nascent delta function arising from the limit of the multivariate normal pdf

\[
\delta_{\Sigma} (x) = |\det (2\pi\Sigma)|^{-1/2} \exp \left( -\frac{1}{2} x^T \Sigma^{-1} x \right)
\]

with a covariance matrix determinant tending to zero (here, \( \Sigma = \epsilon I \) will be used with \( \epsilon \to 0 \)). Next, the difference is approximated by an infinitesimal variation.

\[
\delta (\Psi (\eta)) - \delta (\tilde{\Psi} (\eta)) = \lim_{\epsilon \to 0} \delta_{\epsilon I} (\Psi (\eta)) - \delta_{\epsilon I} (\tilde{\Psi} (\eta)) \\
\approx \lim_{\epsilon \to 0} \delta'_{\epsilon I} (\tilde{\Psi} (\eta)) \Delta \Psi (\eta)
\]  

(6.43)

where \( \Delta \Psi (\eta) = \Psi (\eta) - \tilde{\Psi} (\eta) \), and where \((\bullet)'\) denotes differentiation w.r.t. the first \( n_{\Psi} \) arguments (note that the dimensions of the subscript matrix used are \( n_{\Psi} \times n_{\Psi} \)) and \( I \) represents an identity matrix. Now, from the chain rule

\[
\left( \delta_{\epsilon I} (\tilde{\Psi} (\eta)) \right)' = \delta'_{\epsilon I} (\tilde{\Psi} (\eta)) \tilde{\Psi}' (\eta), \text{ and } \tilde{\Psi}' (\eta) = J_1^* (\eta_1 - \eta_1^*) + J_2^* (\eta_2 - \eta_2^*),
\]

\( \tilde{\Psi}' (\eta) = J_1^* \) (both resulting from (6.29)), one gets for the first factor composing the limit:

\[
\delta'_{\epsilon I} (\tilde{\Psi} (\eta)) = \left( \delta_{\epsilon I} (\tilde{\Psi} (\eta)) \right)' J_1^{* -1}
\]  

(6.44)
Herein,
\[
\delta_{\epsilon I} \left( \tilde{\Psi} (\eta) \right) = |\det (2\pi\epsilon I)|^{-1/2} \exp \left( -\frac{1}{2\epsilon} \tilde{\Psi}(\eta)^T \tilde{\Psi}(\eta) \right)
\]
\[
= |\det (2\pi\epsilon I)|^{-1/2} \times \exp \left( -\frac{1}{2} \tilde{\Psi}(\eta)^T J_1^* - T \left( \epsilon (J_1^* T J_1^*)^{-1} \right) J_1^* - 1 \tilde{\Psi}(\eta) \right)
\]
\[
= |\det J_1^*|^{-1/2} \delta_{\epsilon \{J_1^* T J_1^*\}^{-1}} \left( J_1^* - 1 \tilde{\Psi}(\eta) \right)
\]

Defining the scalar function
\[
d_{\epsilon}(\eta) = |\det J_1^*|^{-1} \delta_{\epsilon \{J_1^* T J_1^*\}^{-1}} (\eta_1 - \eta_1^* + Q^* (\eta_2 - \eta_2^*))
\]
one obtains for (6.44)
\[
\delta'_{\epsilon I} \left( \tilde{\Psi} (\eta) \right) = (d_{\epsilon}(\eta))^T J_1^* - 1
\]

Hence, the expression inside the square brackets of (6.42) becomes approximately
\[
\lim_{\epsilon \to 0} (d_{\epsilon}(\eta))^T J_1^* - 1 \Delta \Psi (\eta)
\]

Defining \( \chi(\eta) = e^{-\gamma V(\eta)} J_1^* - 1 \Delta \Psi (\eta) \in \mathbb{R}^{n_2} \), a multivariate partial integration can be carried out over \( \eta \)'s vector part \( \eta_1 \) (and viewing \( \eta_2 \) provisionally as parameters):
\[
\Delta I \approx \lim_{\epsilon \to 0} \int_{-\infty}^{+\infty} (d_{\epsilon}(\eta))^T \chi(\eta) d\eta
\]
\[
= \lim_{\epsilon \to 0} \int_{-\infty}^{+\infty} \text{tr} \left[ \chi(\eta) (d_{\epsilon}(\eta))^T \right] d\eta
\]
\[
= \lim_{\epsilon \to 0} \int_{-\infty}^{+\infty} \text{tr} \left[ (d_{\epsilon}(\eta) \chi(\eta))^T - d_{\epsilon}(\eta) (\chi(\eta))^T \right] d\eta
\]

The first term vanishes since \( \lim_{\epsilon \to 0} d_{\epsilon}(\eta) \chi(\eta) = 0 \). The function \( d_{\epsilon}(\eta) \) now tends to a Dirac delta distribution up to the scaling factor \( |\det J_1^*|^{-1} \).

\[
\Delta I = -\lim_{\epsilon \to 0} \int_{-\infty}^{+\infty} d_{\epsilon}(\eta) \text{tr} \left[ (\chi(\eta))^T \right] d\eta_1 d\eta_2
\]
\[
= -|\det J_1^*|^{-1} \int_{-\infty}^{+\infty} \delta(\eta_1 - \eta_1^* + Q^* (\eta_2 - \eta_2^*)) \text{tr} \left[ (\chi(\eta))^T \right] d\eta_1 d\eta_2
\]
\[
= -|\det J_1^*|^{-1} \int_{-\infty}^{+\infty} \text{tr} \left[ (\chi(\eta))^T \right] \eta_1^* - Q^* (\eta_2 - \eta_2^*) d\eta_2
\]
The integrand can now be worked out, we have

\[(\chi(\eta))' = J_1^{* -1} e^{-\gamma V(\eta)} (\Delta \Psi' (\eta) - \Delta \Psi (\eta) \gamma V' (\eta))\]

and since \(\gamma \to \infty\), the first term of the difference can be neglected, and the trace of the remaining second term becomes

\[
\text{tr } [(\chi(\eta))'] = -\text{tr } [J_1^{* -1} e^{-\gamma V(\eta)} \Delta \Psi (\eta) \gamma V' (\eta)]
\]

with \(w(\eta)^T = V'(\eta) J_1^{* -1}\), and

\[
\Delta I \approx \gamma |\det J_1^{*}|^{-1} \int_{-\infty}^{+\infty} e^{-\gamma V(\eta_1^* - Q^*(\eta_2 - \eta_2^*); \eta_2)} \times \left[ w(\eta)^T (\Delta \Psi (\eta)) \right]_{\eta_1^* - Q^*(\eta_2 - \eta_2^*)} \, d\eta_2
\]

This integral can be approximated via the LIM; this requires a quadratic Taylor approximant of \(V(\eta_1^* - Q^*(\eta_2 - \eta_2^*); \eta_2)\) around its minimizer, and a Taylor approximant of the right integrand factor at the same point. This minimizer is exactly \(\eta_2^*\), since (6.27) also implies

\[\eta_2^* = \arg\min_{\eta_2} V(\eta) \quad \text{s.t. } \tilde{\Psi}(\eta) = 0\]

(The reader should notice the slight difference in the constraint compared with (6.27); this can be shown via the Lagrange equations) and this coincides with the constraint \(\eta_1 = \eta_1^* - Q^*(\eta_2 - \eta_2^*)\). By construction of the function \(\tilde{\Psi}(\eta)\) (which is a perfect representation of the linear components of \(\Psi(\eta)\)), each component \(l\) of the difference function \(\Delta \Psi(\eta)\) can be represented with a Taylor series with a first term which is quadratic:

\[\frac{1}{2} (\eta_2 - \eta_2^*)^T \left. \frac{\partial^2 \Psi_l}{\partial \eta_2^2} \right|_{\eta_2^*} (\eta_2 - \eta_2^*)\]

Hence,

\[w(\eta)^T (\Delta \Psi (\eta)) = \frac{1}{2} (\eta_2 - \eta_2^*)^T \sum_{l=1}^{n_{\Psi}} w_l(\eta) \left. \frac{\partial^2 \Psi_l}{\partial \eta_2^2} \right|_{\eta_2^*} (\eta_2 - \eta_2^*)\]

wherein a Taylor expansion of \(w_l(\eta)\) up to degree 1 only has contributions coming from the constant term, due to central 3rd order moments of the Gaussian pdf popping up. Then, after translating the integration variables as in
6.C. Approximation of the integral approximation error

Appendix 6.A, and using (6.32), the LIM can finally be applied

\[ \Delta I \approx \frac{\gamma}{2} \left| \det J_1^* \right|^{-1} \int_{-\infty}^{+\infty} e^{-\gamma \left( V(\eta^*) + \frac{1}{2} \eta_2^T H W \eta_2 \right)} \eta_2^T N (\eta^*) \eta_2 d\eta_2 \]

\[ = \frac{\gamma}{2} \left| \det J_1^* \right|^{-1} e^{-\gamma V(\eta^*)} \left| \det 2\pi (\gamma H_W)^{-1} \right|^{1/2} \times \operatorname{tr} \left[ (\gamma H_W)^{-1} N (\eta^*) \right] \]

\[ = \frac{1}{2} \operatorname{tr} \left[ H_W^{-1} N (\eta^*) \right] \tilde{I} \]

An approximation of the relative integration error \( \Delta I/\tilde{I} \) is therefore given by the expression \( \frac{1}{2} \operatorname{tr} \left[ H_W^{-1} N (\eta^*) \right] \); as expected, it depends essentially on the relative size of the constraints’ second derivative w.r.t. their first derivative. □

References


Chapter 7

Conclusions

This final chapter concludes the thesis with a general discussion, a list of contributions, followed by a number of thoughts, remarks, restrictions and open issues associated with this work.

7.1 General discussion

In this thesis two topics, composing the two main parts, in the field of nonlinear dynamic systems have been dealt with:

- blind identification of nonlinear block-oriented models;
- instability of nonlinear systems driven by random inputs.

In part I, using a maximum-likelihood approach, discrete-time Wiener and Hammerstein models – block-oriented model structures consisting of a linear time-invariant block and a static nonlinearity in this or reversed order – are identified from output data only. The model input, not available for measurement, is assumed to be white Gaussian noise, and the model structure is assumed to be invertible. The input variance parameter could be eliminated from the negative log-likelihood cost function, resulting in a sum-of-squares cost function. The latter can be efficiently minimized via Newton-Gauss methods, and can be viewed as the classical prediction-error cost function up to a correction factor coming from the nonlinearity. That iterative joint optimization algorithm is in practice preceded by a two-step procedure for the generation of high-quality initial estimates of both the linear and the nonlinear model parts.
A low output measurement noise level (coinciding to a high SNR condition) was assumed. The effects of additive noise were shown to be a bias, which is – in a first order approximation – proportional to the noise variance; therefore, this bias can be neglected if the estimator’s variance is high (e.g. when the number of data is not very high). The maximum-likelihood’s asymptotical properties – viz. the consistency, efficiency, rate of convergence and asymptotic normality of the estimates, as the number of data tend to infinity – were established using cumulant- and $\alpha$-mixing statistical tools. Since an estimator without confidence bounds is useless, the Cramér-Rao lower bound was calculated to generate confidence bounds of the estimated model parameters or parameter-related quantities. The methods and their properties for Wiener and Hammerstein systems have been profiled via Monte-Carlo simulation experiments and real-life measurements. The problems related to the generalized blind identification problem in the presence of additive measurement noise have been discussed, and an identifiability proof was given in the Wiener case.

In part II, instability with nonlinear dynamical systems under random inputs has been studied from essentially two points of view: a data-based (statistical) and a model-based perspective:

- The first perspective has revealed that the unstable behaviour of nonlinear systems subject to random inputs depends on the stochastic properties of the input (among which the input standard deviation) and the nature of the system. The output of a nonlinear dynamical system can be bounded for very long periods of time until a blow-up suddenly occurs. This blow-up (or escape) event has been viewed as a memoryless waiting process over time, resulting in geometrically distributed escape times, which can be characterized by a “probability of escape” parameter. This parameter is simulation-length independent, in contrast to the probability of blow-up during a finite-length simulation. A maximum-likelihood method has been outlined to estimate it from multiple finite-length experiments, in which the escape times are recorded. Strictly speaking, if blow-ups can occur at certain standard deviations, even when no blow-ups are observed at a given input standard deviation over a finite period of time, there is always at least a tiny risk of blow-up at any standard deviation. Instability is inherently a transient phenomenon, and therefore the response is clearly non-stationary on an infinite time scale. Extreme-value-statistical tools (the generalized Pareto distribution) have been used to establish the boundedness of the output power, which is possible from a single input-output data record, assuming stationarity in the pre-blow-up phase.

- The second, model-based perspective has been focusing on a state-space interpretation to the blow-up phenomenon: assuming a locally stable
equilibrium, the input can cause the state to jump out of its region of attraction. Starting from a known discrete-time state-space description, a known power spectrum of the Gaussian input and a known initial state, the state’s pdf after a burst of finite duration is approximated at low input amplitudes using the Laplace integration method. The approximation error was studied theoretically. The method allows one to estimate the probability of a bounded response via integration over the region of attraction, at several input standard deviation values at once, which is a huge advantage over Monte-Carlo methods. The results have been illustrated via simulation experiments. The probability obtained via numerical integration of the state’s pdf can provide an order of magnitude of the actual probability, although it is not perfect due to the small input standard deviation assumption; on the other hand, the simulations show that a linearization at the equilibrium point would result in a dramatic underestimation of the risk of unbounded operation.

7.2 List of essential contributions

Contributions of part I: Blind system identification of Wiener and Hammerstein models

- generation of high-quality initial estimates for both the linear and the nonlinear part of a Wiener system (also applicable to Hammerstein systems);
- construction of the Gaussian MLE for Wiener and Hammerstein systems;
- analysis of the asymptotic properties of the MLE;
- calculation of the Cramér-Rao lower bound (Fisher information matrix) for Wiener and Hammerstein systems;
- study of the bias of the methods not taking the output measurement noise explicitly into account;
- identifiability proof for Wiener systems with output measurement noise;
- illustration of the theoretical results via simulations and a measurement example.

Contributions of part II: Analysis of instability under random inputs

- maximum-likelihood estimation of the probability of escape based on multiple experiments;
• extreme-value-based approach to assess whether the output variance is bounded;

• approximation of the probability of bounded operation of a nonlinear state-space model under a Gaussian burst excitation, via the estimation of a future state’s probability density; a theoretical error analysis has been carried out;

• illustration of the results via simulation examples.

7.3 Limitations of the proposed approaches and suggestions for further research

7.3.1 Blind identification of block-oriented nonlinear models

In the first part, the limitations and open problems are essentially caused by the gap between theory and practice, due to the assumptions.

The most constraining assumptions are the whiteness and Gaussianity of the input, and the model structure assumption itself. Not satisfying one of these assumptions in general introduces a bias. Hardly any real-life nonlinear device satisfies the block-oriented assumption: this imposes prior knowledge, which is not always available.

Moreover, besides the verification of the intermediate signal’s properties, there is no possibility to validate the model on a fresh data set since the input cannot be measured.

The parameterization of the static nonlinear part could be done in the forward manner. This is easier for simulation starting from a given input signal, but comes at the expense of an additional inversion of the nonlinearity at each iteration step.

Other model structures, e.g. the block-oriented Hammerstein-Wiener, Wiener-Hammerstein, parallel structures, nonlinear state-space models, could also be used to approximate a nonlinear devices’ behaviour with unmeasured input. Multiple-input multiple-output and continuous-time generalizations of the blind identification methods could be investigated. The theoretical asymptotic maximum-likelihood properties should also be checked for each model and parameterization (in particular this requires more knowledge about $\alpha$-mixing).
7.3. Limitations of the proposed approaches and suggestions for further research

7.3.2 Instability under random inputs

The following thoughts and open issues on this challenging topic are enumerated below:

Chapter 5: Data-driven perspective

- Without prior knowledge, obtaining (in)stability information from experimental data is very complicated, especially if blow-ups occur very infrequently. This makes the problem somehow a short-data problem. Moreover, if a blow-up occurs during the measurement, the useful data set is bounded in time (stops at blow-up). In a single data record, at most one blow-up can occur. The estimation of the probability of blow-ups from this single observation is an ill-posed problem.

- For practical systems with unknown stability information, it is however often known that the probability of escape is relatively low (otherwise it would not make sense to be willing to use the considered system at all). Therefore, the desired information is related to rare events.

- Extreme values can play an important role, but since these are extremes, by definition, this keeps their number low. Lowering the threshold value comes at the expense of introducing a bias into the estimated extreme value parameters. Optimizing the choice of the threshold to the data set is an open issue.

- The estimation of extreme value parameters for non-i.i.d. data sequences is suboptimal in its present implementation: the samples in between spikes (and even the spikes themselves) are discarded if they are not sufficiently separate in time. At this moment, there is still a lot of research in that domain (see the journal “Extremes”, Springer). Using the so-called extremal index (related to the average inter-exceedance time) may be a way to take care of the dependency of the samples over time. Besides, the theory of point processes can also possibly provide inspiration to model dependent extreme events, without wasting data-samples containing precious information.

- The problem setting is inherently also an extrapolation problem. As mentioned in the main text, for input distributions with infinite support, the risk of a blow-up is believed to be present at any input standard deviation. Therefore, detecting the (extent of) instability at low levels will very likely depend on the system’s (very-)large-amplitude behaviour in the spikes and on a never-seen blow-up event. Figure 7.1 illustrates the
extrapolation issue, via a particle moving in a one-dimensional potential field, experiencing a force given by the negative derivative of the potential. However, assuming that the measurements have been performed in the central part, the system’s behaviour on a large-amplitude scale cannot be determined reliably. Estimation of stability information based on (finite-amplitude) measurements is a similar problem.

• The stationarity during the time window before blow-up may be an invalid assumption; to the author’s knowledge, testing for stationarity is a very difficult task, since the changes may occur in higher-order moments, which are not easy to estimate. Note that, in practice, non-stationary effects may also arise due to spurious measured periodic or aperiodic perturbations (e.g. power supplies, mains frequency, electromagnetic interference), time variations, or simply drifts of operating point of electronic circuitry causing a drift in the input and/or output signals. Also the system’s initial transient can be an obvious source of non-stationarity.

• The stationarity is probably a more important criterion than the boundedness of the output variance. This can be illustrated via a continuous static nonlinear function, say \( e^{x^3} \), driven by stationary white Gaussian noise, which results in a stationary and bounded output, but with an unbounded variance.
7.3. Limitations of the proposed approaches and suggestions for further research

Chapter 6: Model-based perspective

- Obtaining a validated nonlinear model is very difficult, since it requires one to find a good model structure, a good parameterization, a good initial estimate, a good model complexity and (not the least!) good measurements. Therefore, the application of the method of chapter 6 to a system without prior knowledge, just starting from measurements, does not guarantee that similar effects would take place in the real-life setup.

- It might be interesting to take a different look at systems driven by random inputs: those can also be seen as systems without input, but with randomly fluctuating parameters. There is some literature available and ongoing research in the field of “stochastic nonlinear systems”.

- The method may be of limited use as the input amplitude increases, due to the linearization w.r.t. the input and the states. As shown theoretically and in the simulations, the estimated pdf then becomes inaccurate. Using higher-order approximation, or reducing the linearization to fewer variables, may improve the accuracy. The choices of the density of the integration grid and of the numerical integration method are other subjects of improvement.

- It was assumed that the region of attraction is compact. This may not be the case in practice, e.g. it may extend towards infinity in a given direction. The (numerical) integration may be complicated and may require a large number of pdf evaluations. However, this does not pose a serious problem, since integration over a reasonable compact subdomain, would result in an overestimation of the probability of an unbounded behaviour. Another option is to reparameterize the numerical integral (via a transformation of variables) such that the new variables would vary in a finite interval.

- Taking the interplay between stability and stationarity into account, it may be interesting to investigate under which conditions the pdf of the state tends to a stationary pdf as the time window parameter \( \tau \) tends to infinity.
Global reference list


Nonlinear dynamic systems:
blind identification of block-oriented models, and instability under random inputs

Laurent Vanbeylen

Nonlinear dynamic systems are everywhere present in our daily life: a microchip, a loudspeaker, a robot, a car, an airplane, a chemical plant, a bridge, ... These applications reveal both dynamic and nonlinear behaviour. Dynamic systems have a memory and their response varies with frequency. “Nonlinear” means that the response does not scale (linearly) with the input amplitude. To provide insight, simulate, predict, design, optimize and control these nonlinear dynamic systems, mathematical models describing their complex behaviour are required. System identification is the theory that aims to build accurate models from measurement data. Till now, system identification has been mainly focusing on linear modelling, but this approach becomes inaccurate when applied to nonlinear systems. Hence, nonlinear models are needed.

The first part of the thesis concentrates on the identification of certain types of nonlinear systems (e.g. Wiener and Hammerstein systems). Usually, in system identification, the system’s input and output are both measured. But in some applications, one has no access to the input, e.g. the wind acting on a bridge or building, or the unknown stock market input. In the cases where only output data are available, blind identification becomes the only option. Due to this fact, blind identification is more involved than the classical identification theory. In this work, the theoretical properties and also the impact of measurement noise disturbances are analyzed.

In a second part, the focus is put on the (in)stability of nonlinear dynamical systems. A system is said to be stable if the response to a bounded input is also bounded. In practice, the stability or instability of a given physical system or model is often unknown. For linear systems the theory is well-established. In this work, the aim is to construct tools for (automated) retrieval of stability information of nonlinear systems (or models), assuming that the input is random.