Identification of discrete-time state-space nonlinear systems with stability guarantees

Vervaet, N.

Award date:
2021

MASTER'S THESIS

Identification of discrete-time state-space nonlinear systems with stability guarantees

DC 2021.054

Author: N. Vervaet
Student number: 0941902
Date: Eindhoven, September 16, 2021

DEPARTMENT OF MECHANICAL ENGINEERING

DYNAMICS & CONTROL

TU/e EINDHOVEN UNIVERSITY OF TECHNOLOGY

Supervisors:

prof. dr. ir. N. van de Wouw
dr. A.Y. Pogromskiy
ir. M.F. Shakib

Committee:

prof. dr. ir. N. van de Wouw
dr. A.Y. Pogromskiy
dr.ir. M. Schoukens *
ir. M.F. Shakib

** Control Systems group, Dept. of Electrical Eng. Eindhoven University of Technology*

Cover image: retrieved from [1].

Abstract

In this thesis, a system identification approach is developed for discrete-time state-space nonlinear models with a strong form of stability guarantees. The real world is complex, essentially nonlinear, and sometimes, modern day science still struggles with a qualitative understanding of the nonlinearities around us. Technologies that allow for an accurate explanation of observations are of great help to improve our understandings of these complex systems. Also, the next-generation high-tech and mechatronic systems come with extreme functionality and performance requirements, inducing ever-increasing model accuracy specifications. To keep on fulfilling these high expectations on system design and model-based control, it has become inevitable to exploit the nonlinear operation modes of real-world systems, raising the demand to go beyond linear system representations by modelling in the nonlinear domain.

Current works have shown that discrete-time block-oriented multiple-input multiple-output (MIMO) nonlinear state-space models are capable of accurately representing a large variety of real-world nonlinear systems. However, their implementations typically suffer from computational inefficiency due to the curse of dimensionality. Furthermore, the identified models cannot be safely used for other inputs than those seen during identification. Even a slight perturbation of the input signal can lead to a significantly different (possibly unbounded) model response.

Recent developments in stability analysis enable the assessment of the convergence property for such models by solving linear matrix inequality (LMI) conditions. This stability notion guarantees the existence of a unique, globally exponentially stable model response to periodic perturbations. Using a discrete-time, MIMO equivalent of the mixed-time-frequency (MTF) algorithm, we show that the steady-state model response and the gradient of the model response with respect to its parameters can be computed in an accurate and computationally efficient manner. These developments allow for the usage of both global and local optimisation methods to solve the system identification problem. Furthermore, by enforcing the identified model to be inside the set of convergent models that are characterized by the LMI conditions, we certify a stability property on the identified model. This property allows for reliable usage of model predictions, also for excitation signals other than those used during model training. The effectiveness and the benefits of the approach are validated on (i) experimental data from the well-known Silverbox nonlinear system identification benchmark and (ii) a simulation case study.

Keywords: Nonlinear system Identification; Data-driven modelling; Discrete-time; Lur'e-type models; Convergent systems; Incremental stability; Steady-state; Silverbox;

Acknowledgements

This master's thesis presents the results of my graduation project, which was carried out during the course of the 2020-2021 academic year within the Dynamics and Control research group of the Mechanical Engineering department at Eindhoven University of Technology. It is safe to say that last year has been *special* to me in many ways. My internship abroad was disrupted by the outbreak of the COVID-19 pandemic, and required me to get used to working remotely. I was still living in Germany when the first steps were taken in preparation for this graduation project. At that time, I committed myself to a year ahead full of hard work, with the intent to show and further explore the competences I acquired throughout my academic career. Because it is coming to an end soon, I want to thank the people that supported me through my time as a master's student.

In the first place, I want to thank my supervisor/mentor prof. dr. ir. N. Van de Wouw for his continuous support over the last three years. Unfortunately, the upcoming master's defence will only be the second time for us to meet in person, but still I am convinced that we made full use of our student-mentor relationship. I may not have walked the well-trodden road by voluntarily extending my studies by a year to enable my internship abroad, where I could contribute (within the core of my interests) to the development of automated driving technologies. Nathan, this is one out of many examples in which you provided me with the flexibility and freedom to make the best out of my master's program. Thank you for the pleasant conversations and your critical view on my research. Your insights kept on steering me in the right direction and your expertise in the analysis of nonlinear dynamical systems helped me to understand the technical details of both my internship and graduation project.

Secondly, I want to thank the whole team of university staff members that additionally guided me through the graduation project, consisting of dr. A.Y. Pogromskiy, ir. M.F. Shakib and dr.ir. M. Schoukens. Sasha, thank you for attending my monthly progress meetings as a supervisor and your efforts to try and make me understand all important technicalities during our fruitful discussions on stability theory. Fahim, thank you for acting as my daily coach and supervisor. Unquestionably, you took this role seriously, by actually guiding me from time-to-time on a daily basis (sometimes even outside office hours). I more than once told my relatives with the blink of an eye that you deserve a statue for playing this key role in my graduation project. I truly enjoyed our many (virtual) conversations in which your focus was really on the progression of my project in a way that definitely strengthened my scientific competences. It has been a pleasure working with you. Lastly, I want to thank Maarten for spending his time reading this thesis and acting as the graduation committee member from another research group.

Above all, I want to thank my family, friends and my girlfriend for their support and encouragement during the sometimes difficult periods. No matter whether I was working during weekends, holidays or even some exceptional night shifts, you have always supported me via inconspicuous, but important small-talks, lunch breaks or by taking me for a walk. Stan, Joris and Guus, the brainstorming sessions during our weekly virtual 'koffierondjes' were of great help as we shared and sometimes together solved our individual graduation project challenges of the week. Last but not least, Shannen, you probably underestimate your contributions to the completion of this thesis. Among other things, you have travelled forth and back between Hulst and Tilburg numerous times to see and support me, while I was only focussed on writing a thesis. You understood the necessity to shorten our only holidays of the year and it has been way too long since we spent some proper quality time together; everything with the goal in mind of successfully finishing this project. I am looking forward to our next steps, which lie just around the corner.

I want to finish with a word of thanks to everyone who read and/or reviewed this thesis and to whoever will.

Niels Vervaet
Eindhoven, September 2021

Contents

List of symbols

Roman symbols

Symbol	Description	Unit
A	System matrix	—
B	Nonlinearity input matrix	—
c	Hyper-parameter in Theorem 3.5	—
C	Nonlinearity output matrix	—
d	Process noise	—
D	External input to nonlinearity output feedthrough matrix	—
e	Output noise	—
F	Output matrix	—
f	Generic nonlinear dynamics	—
G	Nonlinearity output to system output feedthrough matrix	—
$\bar{G}$	Centroid in CRS	—
$\mathcal{G}_\Sigma$	Transfer function representation of Σ	—
H	External input to system output feedthrough matrix	—
J	Objective function	—
k	Time index	—
L	External input matrix	—
N	Measurement interval length	—
n_m	Number of subset agents in CRS	—
n_p	Population size in CRS	—
n_s	Number of successfull trial points in CRS	—
n_t	Number of trial points in CRS	—
n_θ	Decision variable dimension	—
n_u	Nonlinearity output dimension	—
n_x	Internal state dimension	—
n_y	Nonlinearity input dimension	—
n_w	Exogeneous excitation dimension	—
n_z	Output dimension	—
$\bar{P}_i$	Primary and secondary point in CRS for $i \in \{1, 2\}$, respectively	—
P	LMI decision variable in Theorems 3.5 to 3.7	—
R	Population in CRS	—
R_s	Subset in CRS	—
r_i	Subset agent in CRS for $i \in \{1, \ldots, m\}$	—
$\mathcal{S}_c$	Positively invariant set	—
S	LMI decision variable in Theorems 3.5 to 3.7	—
$\mathcal{S}$	Region of incremental stability	—
T	Period time	—
u	System input	—
$\mathcal{U}_{i\times j}$	Matrix of size $i \times j$ with elements from a uniform distribution on $[-1, 1]$	—
V	Lyapunov function	—
v	Measurement noise	—
$\mathcal{W}$	Input class	—
w	Measured excitation	—
$\mathcal{X}$	Convergence region	—
x	Internal state	—
z	System output	—

Greek symbols

Symbol	Description	Unit
α_i	Class $\mathcal{K}$- or $\mathcal{K}_\infty$-functions in Lemmas 3.1 and 3.3 for $i \in \{1, \ldots, 5\}$	—
β	Class $\mathcal{KL}$-functions in Definitions 3.1 and 3.3	—
γ	$\mathcal{H}_\infty$-norm upper-bound in Lemma 2.1	—
δ	Deadzone length	—
ϵ	Output error between model and (noisy) measurements	—
η	Convergence stopping criterium	—
λ	Scalar variable in the proof of Theorem 3.6	—
Ω	Sector condition upper bound	—
$\bar{\Omega}$	Incremental sector condition upper bound	—
$\tilde{\omega}$	DFT frequencies	Hz
ϕ	Solution of dynamics	—
Ψ	External nonlinearity input in a parameter sensitivity model	—
Ψ_i	Nonlinearity input and output loop scaling matrix for $i \in \{1, 2\}$	—
Ψ_3	Loop transformation matrix	—
Σ	LTI dynamical system	—
σ	Hyper-parameter in Theorem 3.5	—
σ_e	Standard deviation of Gaussian white noise	—
τ	Hyper-parameter in Theorem 3.6	—
τ^*	Scalar variable in Theorem 3.7	—
τ_1	Hyper-parameter in Theorem 3.5 for $i \in \{1, 2\}$	—
$\tau_{\max}$	Bisection interval upper-bound	—
$\tau_{\min}$	Bisection interval lower-bound	—
Θ	Decision variable space	—
θ	Decision variables	—
$\hat{\theta}_N$	Estimated parameter vector	—
θ_{init}	Estimated parameter vector	—
φ	Static nonlinearity	—

Subspaces

Symbol	Description	Unit
$\mathbb{C}$	The set of complex numbers on the entire complex plane	—
$\mathbb{C}_\bigcirc$	The set of complex numbers on the unit circle	—
$\mathbb{C}_\ominus$	The set of complex numbers inside the open unit disc	—
$\mathbb{C}_\otimes$	The set of complex numbers inside the closed unit disc	—
$\mathbb{C}_\oplus$	The set of complex numbers outside the closed unit disc	—
$\mathbb{C}^n$	The n-dimensional complex vector space	—
$\mathbb{C}^{n\times m}$	The set of complex matrices of size $n \times m$	—
$\mathbb{N}_0$	The set of natural numbers from 0 upwards	—
$\mathbb{N}$	The set of natural numbers from 1 upwards	—
$\mathbb{R}$	The set of real numbers	—
$\mathbb{R}_{<0}$	The set of negative real numbers	—
$\mathbb{R}_{\geq 0}$	The set of nonnegative real numbers	—
$\mathbb{R}_{\leq 0}$	The set of nonpositive real numbers	—
$\mathbb{R}_{>0}$	The set of positive real numbers	—
$\mathbb{R}^n$	The n-dimensional real vector space	—
$\mathbb{R}^{n\times m}$	The set of real matrices of size $n \times m$	—
$\mathbb{S}^n$	The set of real symmetric matrices of size $n \times n$	—
$\mathbb{Z}$	The set of integer numbers (including 0)	—

Notations

Symbol	Description	Unit
$A \prec 0$	Matrix A is negative definite ($A \prec B \implies A - B \prec 0$)	–
$A \preceq 0$	Matrix A is semi-negative definite ($A \preceq B \implies A - B \preceq 0$)	–
$A \succ 0$	Matrix A is positive definite ($A \succ B \implies A - B \succ 0$)	–
$A \succeq 0$	Matrix A is semi-positive definite ($A \succeq B \implies A - B \succeq 0$)	–
$\lambda(A)$	Set of all eigenvalues of matrix A	–
$\lambda_i(A)$	The i^{th} eigenvalue of matrix A	–
$\lambda_{\max}(A)$	Maximum eigenvalue of matrix A	–
$\lambda_{\min}(A)$	Minimum eigenvalue of matrix A	–
$\rho(A)$	Spectral radius of matrix A	–
$\operatorname{diag}(A)$	Main diagonal of matrix A.	–
$\operatorname{diag}(x)$	Diagonal matrix with x on the main diagonal.	–
$\operatorname{rank}(A)$	Rank of matrix A	–
$\|A\|_{p,q}$	The p, q-induced norm of matrix A	–
$\|A\|_1$	The row-sum norm of matrix A	–
$\|A\|_2$	The spectral norm of matrix A	–
$\|g\|_{\ell_2}$	The ℓ_2-norm on scalar valued sequence g	–
$\|g\|_{\ell_2^p}$	The ℓ_2^p-norm on vector valued sequence g	–
$\|\Sigma(\varkappa)\|_{\mathcal{H}_\infty}$	The $\mathcal{H}_\infty$ signal norm on a transfer function	–
$\|\Sigma\|_{\mathcal{L}_2}$	The induced $\mathcal{L}_2$ gain of a system Σ	–
$\|x\|_2$	The Euclidean norm of vector x	–
$\|x\|_P$	The P-weighted norm of vector x	–
$\star$	Symmetric block in a matrix	–
$\tilde{(\cdot)}$	Related to loop transformation and/or loop scaling	–
$\bar{(\cdot)}$	Steady-state	–
$\dot{(\cdot)}$	First-order time derivative	–
$\ddot{(\cdot)}$	Second-order time derivative	–
$(\cdot)^w$	Related to external input w	–
$(\cdot)^*$	Stopping criterium tolerance	–
$(\cdot)^\top$	Transpose	–
$(\cdot)^\dagger$	Conjugate/Hermitian transpose	–
$(\cdot)^e$	Subject to output noise	–
$\breve{(\cdot)}$	Related to the parameter sensitvity model	–
$(\cdot)_{\theta_i}$	Partial derivative with respect to θ_i	–
$(\cdot)^{\theta_i}$	Related to the parameter sensivity model of θ_i	–
$(\cdot)_0$	Related to the 'true' system	–
$(\cdot)^{-1}$	Inverse	–
$\mathcal{E}(P, c)$	Ellipsoid defined as $\left\{x \in \mathbb{R}^m \mid \|x\|_P < c\right\}$	–
$\widehat{\mathcal{F}}(\cdot)$	Discrete fourier transformation	–
$\mathcal{Z}(\cdot)$	$\varkappa$-transformation	–
$\sup_{x \in \mathcal{X}} f(x)$	Least upper bound on $f(x)$ for $x \in \mathcal{X}$	–
$\max_{x \in \mathcal{X}} f(x)$	Maximum of $f(x)$ for $x \in \mathcal{X}$	–
I_n	The identity matrix of order n	–
$\mathbb{1}_n$	The one vector in $\mathbb{R}^n$	–
$0_{m \times n}$	The zero matrix in $\mathbb{R}^{m \times n}$	–
0_n	The zero matrix in $\mathbb{R}^{n \times n}$	–
$\underline{0}_n$	The zero vector in $\mathbb{R}^n$	–
e_n^i	The basis vector $\begin{bmatrix}\underline{0}_{i-1}^\top & 1 & \underline{0}_{n-i}^\top\end{bmatrix}^\top$ in $\mathbb{R}^n$	–

Acronyms

BLA	best linear approximation
BMI	bilinear matrix inequality
CRS	controlled random search
DFT	Discrete Fourier Transform
DTFT	Discrete time Fourier transform
FDIDENT	Frequency Domain System Identification Toolbox
FRF	frequency response function
GAS	global asymptotic stability
GES	global exponential stability
GEUC	global exponential uniform convergence
IDFT	Inverse discrete Fourier transform
ISS	input-to-state stability
KYP	Kalman–Yakubovich–Popov
LMI	linear matrix inequality
LTI	linear time-invariant
MIMO	multiple-input multiple-output
MTF	mixed-time-frequency
NARX	nonlinear autoregressive exogenous
NL-LFR	nonlinear linear fractional representation
ODE	ordinary differential equation
PDE	partial differential equation
PEM	prediction error method
PISPO	periodic input the same period output
PoNCS	performance of nonlinear control systems
RMSE	Root Mean Squared Error
RNN	recurrent neural network
SEM	simulation error method
SIMO	single-input multiple-output
SISO	single-input single-output

Chapter 1

Introduction

This thesis is conducted on the topic of *Identification of discrete-time state-space nonlinear systems with stability guarantees*. The project background is discussed in Section 1.1. In Section 1.2, an overview of the state-of-the-art on this research area is given, where we summarise relevant existing results and open challenges. Section 1.3 addresses the research goal and planned contributions of this project. Finally, Section 1.4, summarises the outline for this report and Section 1.5 provides preliminaries on our used notations.

1.1 Background

The focus of this project is on the identification of a certain class of dynamical systems. In this section, describing the background of this project, the concept of a *dynamical system* is introduced first. Consecutively, the motivation for *modelling* such dynamical systems is given, followed by an introduction of the considered model class. Finally, it is explained how *system identification* contributes to dynamic modelling, specifically in the engineering research domain.

> "The totality is not, as it were, a mere heap, but the whole is something besides the parts."
>
> — Aristotle, *Metaphysics*, 350 B.C.E [2]

Dynamic systems

The net of science aims to cover the empirical universe by addressing the following two questions [3]: what is it made of and why does it work this way? A subset of this universe is typically studied as a complex system, interacting with its respective environment. Dealing with these complex systems has been an important discussion topic lately. On the one hand, the *analysis* philosophy tries to gain understanding on complex phenomena by breaking them down into their constituent parts. This inherently depromotes the relationships between components and, as was already suggested by Aristotle, a system may be more than the sum of its parts. On the other hand, the *synthesis* paradigm tries to describe an entity through the context of its relations and functioning within the whole. Note that both paradigms go hand in hand as no synthesis can be done without analysis and no analysis makes sense without verification via synthesis [4]. It goes without saying that these abstract systems can be related to phenomena in areas such as physics and mechanics, but also sociology and biology [5], [6].

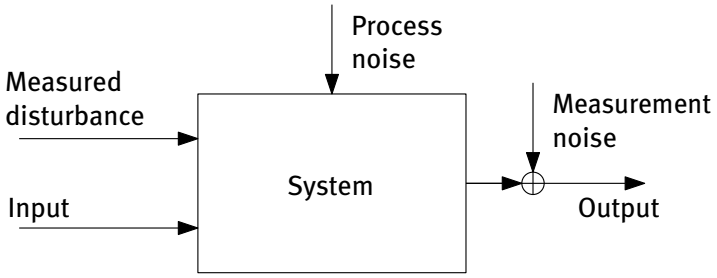


Figure 1.1: Schematic definition of a dynamic system.

Phenomena originating from either one of these research areas can be described by a mathematical system object in which variables of different kinds interact, producing observable signals while being subject to external stimuli, see Figure 1.1 [7]. Observable signals of interest are referred to as *outputs*, whereas manipulable stimuli are called *inputs*. The remaining external signals are *disturbances* and can either be directly measurable or unknown to the observer. With respect to the unknown disturbances, one distinguishes between measurement- and process-noise, respectively affecting only the outputs and also the system's internal signals. Systems commonly show behavioural changes over time according to a fixed rule, making them *dynamic systems*.

Dynamic models

When interacting with a system, it is of great use to have some concept, a *model*, of how its variables relate based on observed signals [7]. Such a mathematical model allows for (i) causal explanation, such as the health care industry trying to relate patient complaints to the disease causative agent, (ii) model output prediction for any input and (iii) description, which provides insights in principles that determine the system behaviour [8]. Note that a model never perfectly reflects reality, but requires only to be sufficiently accurate.

Consider Figure 1.2, in which the necessity to motivate the considered model class in this project becomes clear, given the large variety of existing model classes. It was mentioned before that the model is required to describe the future and/or history of a dynamic system, explaining the decisions in Step I and II. The systems under study are dynamic in the time domain and do not change their dynamics as function of a spatial coordinate (as is the case for distributed partial differential equation (PDE) models). As a result, lumped models are preferred in terms of an ordinary differential equation (ODE) in Step III. In this project, the derived model is expected to mimic systems that change their state only at a discrete set of points in time. Such systems commonly exist in research areas like biology and economics, see [9], [10], or as discretised versions of continuous-time models arising in physics or mechanics. At the same time, one should note that system outputs (even if they evolve in continuous-time) can typically only be measured at discrete time instances, raising the demand for models in terms of difference equations, which motivates the decision in Step IV. The model parametrisation does not change as function of time and model updates are assumed to be deterministic in Steps V and VI, implying that stochastic and time-varying difference equations are excluded.

One should understand why this project considers the class of *nonlinear* systems over its linear counterpart in Step VII, as indeed, the Polish-American mathematician Stanislaw Ulam once compared nonlinear science to studying the bulk of zoology as non-elephant animals [11]. Nevertheless, the vast majority of physically

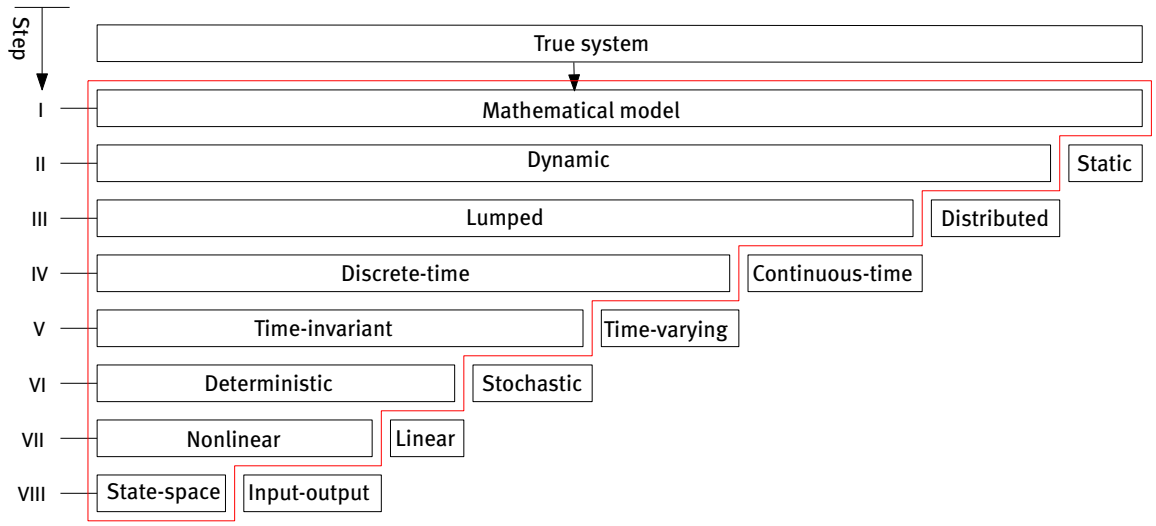


Figure 1.2: Definition of the model structure considered in this project (——).

relevant systems exhibits essential nonlinear dynamic behaviour. On the one hand, all (non-significant) nonlinear phenomena can be neglected while modelling in a linear framework. On the other hand, one can try to include nonlinearities in the model, increasing the model accuracy at the cost of model complexity. The latter is only recommended when a linear approximation fails the requirements, since nonlinear modelling has proven to be significantly more involved [12]–[14]. Among other things: (i) nonlinear models do not necessarily preserve stability under extrapolation outside the input domain that was seen during training, (ii) structural model errors are inevitable as it may be challenging to fit the true system nonlinearities in a certain nonlinear model class and (iii) process noise transforms nonlinearly such that it is no longer Gaussian, which results in undesirable statistical properties of identified model parameters [12].

It is not necessarily possible to translate one nonlinear model representation into another, in contrast to the linear time-invariant (LTI) case, where for example state-space and transfer-function representations are often interchanged [15]. Several nonlinear model classes are proposed, which can roughly be divided into state-space and input-output models [16]. The latter does not directly consider the internal behaviour of the modelled system, which, e.g., complicates the analysis of stability properties. Furthermore, the state-space model representation is more flexible than the well-known nonlinear autoregressive exogenous (NARX) input-output models, increasing the likelihood of an accurate data fit [12]. Besides that, most of the nonlinear control theory is based on state-space representations, which makes these models the intuitively preferred class and explains the decision in Step VIII [17]. In conclusion, the model class of interest, being nonlinear discrete-time state-space models is now fully motivated.

System identification

What remains is to find which model inside the set of candidate models defined by the model class fits the system under study best. To that end, one can take a white-box-, or first-principles modelling approach, reflecting simplifying assumptions together with well-known physical laws such as, e.g., mass balance, momentum balance or heat transfer relations. These models are typically time-consuming to derive and some physically meaningful parameters might still not be available to the user. In that case, additional system calibration measurements are to be performed such that the unknown parameters are found. Since experimentation is then required for model development anyway, the other approach would be to derive the model via an identification experiment, in which a configured input design is applied to the system and its outputs are measured with an appropriate sampling procedure [18]. The art of turning large input-output data sets that are obtained experimentally from dynamical systems into compact models described by only a few parameters is referred to as *system identification* [13], [14]. The procedure, as shown schematically in Figure 1.3, is characterised by four main components, being *data*, *model structure*, *estimation method* and *validation*.

The experimentally obtained system input-output data is the fundamental information source. It is important to make sure that the prediction domain that is relevant to the model is mostly incorporated in the experiment. If essential system features are not observed in the data, they are most likely not described by the identified model. The model structure defines the set of candidate models up to a parameter vector and ideally includes the true system dynamics. The estimation method should ensure that a specific parameter vector is selected that describes the observed data best, given the model structure. Typically, a cost function being nonlinear

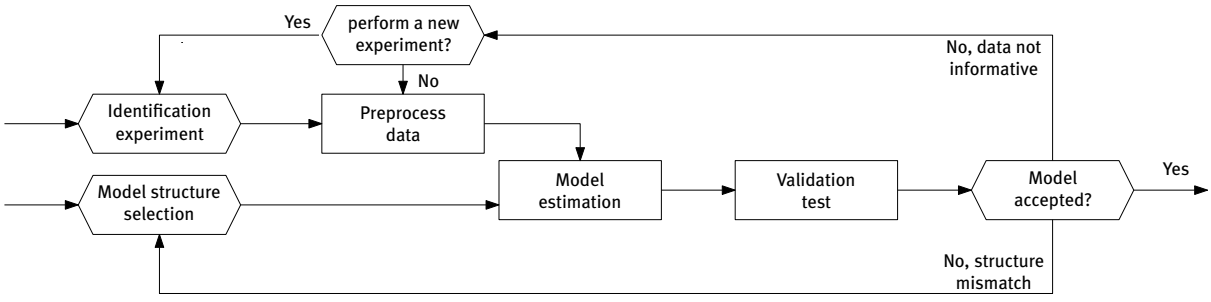


Figure 1.3: System identification [7]. User decisions are in hexagons and algorithms are in rectangles.

in the parameters has to be minimised in this step. Finally, the validation stage executes once a model is identified. In this stage, the model is validated by comparing its simulated output to a new data set that was not used during identification. Besides that, prior knowledge on the modelled system can be verified in this stage. If the validation test fails, we can alter properties of our model structure, e.g., an LTI model's state dimension, and/or the experiment data, before estimating a new model. This process iterates until a model passes the validation test according to user-defined criteria.

Engineering applications

So far, it has been described why dynamical systems are modelled and how system identification can be of great use to establish the actual model describing a system under study best. These results connect well to engineering challenges in dynamics and control, for which some motivating examples are listed below.

1. *Control of turbulent fluid systems* [19]: Imagine an aircraft wing positioned at a high angle of attack, causing a stalled, turbulent downstream flow. One may desire to control this dynamical system to re-laminarize the flow, which increases lift and decreases drag. Hereto, the nonlinear fluid dynamics are to be modelled, which are described by the well-known Navier-Stokes equation. Even though the nonlinear equations of motion are known, they are hard to solve and therefore not applicable for control purposes. It is desired to derive a simpler nonlinear model, for which system identification can be used.
2. *Damage analysis of large mechanical structures subjected to dynamic loading* [20]: Structural elements such as beams, cables and columns may be damaged under dynamic seismic, wind, or impact loading. To prevent structural failure, one tries to identify damage in an early stage via dynamic modelling. Large and complex real-life structures cannot be modelled accurately by first principles, which raises the demand for nonlinear system identification techniques.
3. *Modelling and control of mechanical ventilation* [21]: Respiratory modules assist patients to breathe. Even though the nonlinear dynamic models (including the patient's state-of-health) can be derived via first principles, the parameters describing the system are unknown and require offline calibration, which is a time-consuming task. One can think of implementing system identification techniques to online update the model, hereby minimising downtime of the control system and therefore increasing the treatment efficiency.

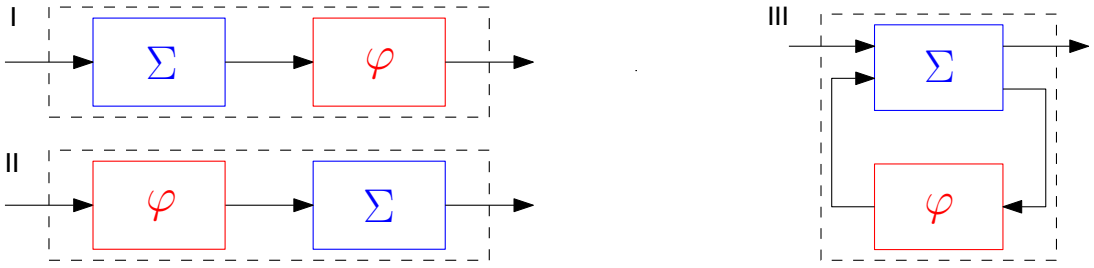
Some major challenges in these examples include (i) nonlinear dynamic behaviour, (ii) (possibly) unmodelled dynamics, (iii) high-dimensionality of the model parametrisation and (iv) limited measurements and actuations. Via system identification techniques, dominant patterns in complex systems can be described by relatively low-order models. As soon as the model is established, nonlinear control theory can be applied to steer the identified system, or retrieved insights potentially lead to improved system design. This project thereby contributes to bridging the gap between traditional dynamics and control techniques and recent developments in data-driven modelling. It does so by studying system identification of discrete-time nonlinear state-space models.

1.2 State of the art

The topic of system identification for discrete-time nonlinear state-space models has been introduced in Section 1.1. The current section summarises the achievements, limitations and open challenges of identification strategies proposed in literature.

Many different nonlinear system identification approaches have been proposed in literature, among which the family of block-oriented nonlinear models has been popular [22]. In these models, nonlinear dynamic behaviour is separated into LTI dynamics and static nonlinearities. These building blocks can be combined in different ways, of which some examples are shown in Figure 1.4. The popular Wiener and Hammerstein structures, respectively structure I and II in Figure 1.4a, can be interpreted as a linear plant interacting with a nonlinear output sensor or a nonlinear actuator input. These model structures showcase the advantages of block-oriented nonlinear models being relatively simple to understand and easy to use [23].

The Lur'e-type structure in Figure 1.4b includes, but is not limited to, the Wiener and Hammerstein structures.



(a) Wiener (I) and Hammerstein (II) type structures

(b) Lur'e (III) type structure

Figure 1.4: Block-oriented nonlinear models, interconnecting LTI models Σ and static nonlinearities φ [12].

Many engineering systems can be modelled as LTI with local static nonlinearities such as hardening/softening springs, backlash and elastic stops [24]. Therefore, this model structure is chosen as the sweet spot between complexity and representation capability of identified models [25]–[27]. System identification of discrete-time, Lur'e-type models is an active research area, see [23], [28]–[35]. Lur'e-type systems are also referred to as nonlinear-feedback and nonlinear linear fractional representation (NL-LFR) systems [22], [31].

The research in [36] studies system identification of *discrete-time* Lur'e-type models. A general multiple-input multiple-output (MIMO) static nonlinearity is allowed and the model complexity can be balanced by tuning the nonlinearity in- and output dimensions as well as the LTI model order. In an example, a single hidden layer feedforward neural network with a linear output layer is chosen as the memoryless nonlinear function, since it can be a universal approximator [37]. The identification procedure is split into two steps, starting with a best linear approximation (BLA) of the nonlinear system, followed by solving a nonlinear non-convex optimisation problem by a gradient-based routine. The model structure considered in [36] coincides with the model class considered in this project, since MIMO discrete-time Lur'e-type models include, but are not limited to, the class of nonlinear discrete-time state-space models. It should be noted that a thorough theoretical analysis on, e.g., stability of the identified models is lacking. Furthermore, there was no focus on computational efficiency of the system identification algorithm, which raises questions on the topic of applicability to problems with a large number of variables [38].

One of the major challenges in system identification is to enforce a form of stability on the identified model, see [39], [40]. Even though stable responses are observed for the true system, the responses of a model resulting from an identification routine may not necessarily obey this property. For LTI models, enforcing exponential stability is well understood [40]–[46]. However, the literature on the nonlinear counterpart with some guaranteed form of stability is scarce and the consequences are more pronounced and complex [47], [48]. It is well-known that nonlinear models can exhibit multiple stable solutions being attractive for different sets of initial conditions [27]. Moreover, depending on the applied excitation signal, the response of nonlinear models can become unstable. Consequently, even though the model explains the identification data set perfectly well, the model response to a new bounded excitation (that is only slightly different from those used during training) might be close to the true system response, but might as well be far off or even grow unbounded. Such models are not capable of accurately predicting the system response in scenarios other than the exact identification data set.

Another important challenge is related to the computational efficiency of algorithms used for nonlinear system identification, since these algorithms are reported as extremely time-consuming, especially for problems with a large number of decision variables [38], [49]. Most approaches minimise a non-convex cost function using gradient-based optimisation routines. These algorithms require the computation of model responses to evaluate the cost function. Moreover, additional model response computations are required for numerical estimation of the cost function gradient (and Hessian) with respect to the model parameters [13], [14]. Next to improvements on the hardware side to speed-up computations, one can think of data-parallelism, in which shorter, but an increased number of model response simulations should be computed, increasing performance by parallel computing [50]. Nevertheless, it is still an open challenge to minimise the computational load of a single simulation step during identification.

In [13], [14], [51], system identification of *continuous-time* Lur'e-type models is discussed, while addressing

the previously mentioned challenges on stability guarantees and computational efficiency. This model structure comes with relatively easy-to-check, sufficient circle-criterion-like conditions for global exponential convergence, which is a strong form of model stability. In particular, this property guarantees the existence and uniqueness of a globally exponentially stable steady-state solution [52]. Exponentially convergent systems obey the property that the periodic steady-state model response is of the same period time as its periodic excitation [52]. This property was exploited in the development of a computationally efficient algorithm to compute steady-state responses to evaluate the cost function and its gradient with respect to the model parameters. These steady-state model responses are compared to steady-state measurements to quantify model performance. Thereby, it is no longer required to estimate the initial system state. Also, one can benefit computationally from implementing the efficient methods to compute the unique steady-state model response. In system identification of linear models, it is common practice to process steady-state model responses in frequency domain methods, which boils down to estimation of a transfer function or a frequency response function in parametric and non-parametric approaches, respectively [53]. In the nonlinear system identification field, steady-state model response are also exploited, e.g., when BLA methods are deployed as a starting point in nonlinear model estimation, as shown in [13], [14], [51]. The system identification problem is typically solved via a two-step algorithm. An initialisation step is performed, which is not limited to the BLA (see, e.g., [23]), but can also be performed by deployment of global optimisation routines. In the second step, a nonlinear non-convex optimisation problem is solved via a gradient-based routine. To that end, a Lur'e-type parameter sensitivity model approach is presented to calculate the exact objective function gradient with respect to the model parameters. These works, hence, solve the previously mentioned major challenges in nonlinear system identification. These works are limited to the continuous-time case, which do not fall into the discrete-time nonlinear state space model class considered in this project. Furthermore, only single-input single-output (SISO) nonlinearities are considered, which greatly reduces the representation capabilities of identified models.

Other nonlinear system identification approaches that address our stability challenges are modelling in a recurrent neural network (RNN) structure. [54] discusses nonlinear system identification of black-box discrete-time Lur'e-type models, viewed as RNNs, in which parameters have no physical interpretation. These models have been shown to exhibit excellent expressive power, but lack stability or robustness guarantees. To that end, methods that identify stable RNNs via a convex parametrisation are proposed in [55], [56]. Next to global exponential stability (GES) of all solutions, an upper bound to the Lipschitz constant of the learnt network is formulated, such that a sensitivity of the model output with respect to a small input perturbation (in the infinity norm) can be quantified. This allows for robustness analysis of the identified model and a direct link to robust control theory. Note that the stability conditions in the proposed methods are conservative in the sense that they are only sufficient. These kind of models can generally not be used in combination with white-box modelling, which require a physical understanding of the parameters, such as [13], [14].

Open challenges

The presented review of the state-of-the-art on the topic of system identification of nonlinear discrete-time state-space models identifies a set of open challenges that are to be addressed in this project. In [36], a flexible Lur'e-type discrete-time model structure has been proposed for the purpose of system identification. Furthermore, [13], [14] proposes a methodology for a less flexible continuous-time model structure that deals with major challenges in system identification, being stability guarantees and computational efficiency of the estimation algorithm. These challenges have not been resolved in literature for the flexible discrete-time case as of yet. This insight initiates the following open challenges: How to guarantee stability for flexible discrete-time models with the same structure as [36]? How to compute steady-state solutions of stable models inside this class computationally efficiently? How can these features be combined into a system identification algorithm? These open challenges were taken as a starting point for this project's research goal formulation.

1.3 Research goals

Based on the state-of-the-art in Section 1.2 and especially the remaining open challenges, the foremost objective for this project is to develop a methodology for system identification of the flexible model structure

for discrete time systems, proposed in [36]. Inspired by the continuous-time contributions in [13], [14], this project aims to include stability guarantees to the identified model and benefit from computationally efficient model response calculation methods. In short, the main goal of this research is defined as:

"*Develop an experimentally validated, computationally efficient system identification methodology for flexible discrete-time nonlinear state-space models with stability guarantees.*"

From this goal, a number of sub-goals can be identified towards an implementation of the main goal:

- Identification of a *flexible* discrete-time nonlinear state-space model is guaranteed by the Lur'e-type model structure. This structure is extremely flexible if, e.g., deep feedforward neural networks are chosen to characterize the nonlinear part of this model class as was shown in [36].
- An equivalent to the steady-state solution computation algorithm that was used in [13], [14] is to be developed in this project. Once computational advantages of the mixed-time-frequency (MTF) algorithm for discrete-time systems are numerically investigated and confirmed, a contraction analysis of the algorithm towards the steady-state model response to periodic excitation should be proven, similar to [57].
- A method is to be developed to calculate the exact gradient of the objective function (the mismatch between model and system outputs) with respect to the model parameters efficiently, as was done for the continuous-time case in [13], [14].
- Stability guarantees of the identified model are required, defined in terms of the global exponential convergence property for nonlinear models, see [13], [14]. This property is selected because the identified model is desired to have, for each (periodic) excitation, a unique (independent from the initial state conditions) and GES steady-state (periodic) model response that is not highly sensitive to small input perturbations in the infinity norm of the excitation. Additionally, conditions that are computationally cheap to check are desired, such that the identification algorithm progression is not limited by frequent, computationally costly stability checks.
- The identification methodology is required to be validated using experimentally obtained benchmark datasets [58], [59]. The performance of our algorithms can then be compared quantitatively to state-of-the-art methods in literature. Also, a simulation case study is to be performed for validation purposes.

1.4 Report outline

The outline of the remainder of this report is as follows:

Chapter 2 provides an overview of the Lur'e-type class of discrete-time dynamical systems that are of interest to this project. It presents important assumptions and properties of these dynamics as preliminaries before the focus shifts to answering the research questions.

Chapter 3 explains the stability guarantees that are required for the Lur'e-type dynamics to be a candidate model. First we elaborate on the challenges in stability analysis of nonlinear systems and how/why global exponential uniform convergence (GEUC) is a desired system property to overcome these challenges. Thereafter, we propose numerically tractable conditions to verify this property as a stability guarantee for Lur'e-type systems.

Chapter 4 combines the information of its preceding chapters into a system identification algorithm for Lur'e-type models with stability guarantees. Here, we formalise our model class and explain how we solve the system identification problem, which is formulated as a standard constrained optimisation problem. We focus on a computational efficient implementation by proposing methods to calculate steady-state model responses fast and accurately in order to evaluate the objective function and its gradient.

Chapter 5 validates the system identification algorithm by presenting system identification results on Benchmark datasets, as well as a numeric case-study.

Chapter 6 presents the conclusions of this thesis, together with recommendations for future research.

1.5 Preliminaries on notation

Definition 1.1 (Signal norms and normed signal spaces [60, p. 94-97])

- *For* real vectors $x \in \mathbb{R}^n$*, one can define:*
  - *the Euclidean norm* $\|\cdot\|_2$ *of* x *by*
  $$\|x\|_2^2 = x^\top x. \tag{1.1}$$
  - *the* P*-weighted norm* $\|\cdot\|_P$ *of* x *for* $0 \prec P \in \mathbb{S}^n$ *by*
  $$\|x\|_P^2 = x^\top P x. \tag{1.2}$$
- *For* scalar valued sequences $g : \mathbb{N} \to \mathbb{R}$*, the Banach space* $\ell_2(N)$ *of all* N*-periodic sequences having a finite* ℓ_2*-norm* $\|\cdot\|_{\ell_2}$ *is defined, such that*
$$\|g\|_{\ell_2}^2 = \textstyle\sum_{i=1}^n |g(i)|^2 < \infty. \tag{1.3}$$
- *[61, p. 3]: For* vector valued sequences $g : \mathbb{N} \to \mathbb{R}^n$ *(with* g_i *being the* i*-th sequence in* g*), the Banach space* $\ell_2^n(N)$ *of all* N*-periodic sequences having a finite* ℓ_2^n*-norm* $\|\cdot\|_{\ell_2^n}$ *is defined, such that*
$$\|g\|_{\ell_2^n}^2 = \textstyle\sum_{i=1}^n \|g_i\|_{\ell_2}^2 < \infty. \tag{1.4}$$
- *For* vector valued complex functions $f : \mathbb{C}_\oplus \to \mathbb{C}^n$ *(with* f_i *being the* i*-th sequence in* f*), one can define the Hardy space* $\mathcal{H}_2^n$ *of functions for which all* f_i *are analytic on , as well as continuous and square-summable on every circle in up to the unit disc. This function space only appears in Definition 1.3.*

Definition 1.2 (Matrix and system norms [27, p. 648])

- *A matrix* $A \in \mathbb{R}^{m\times n}$ *defines a bounded linear mapping between vectors* $x \in \mathbb{R}^n$*,* $y \in \mathbb{R}^m$*, such that* $y = Ax$*. The induced* p, q*-norm of matrix* A *reads (for* $p \geq 1$*, and* $q \leq \infty$*) as*
$$\|A\|_{p,q} = \sup_{x\neq 0} \frac{\|Ax\|_q}{\|x\|_p}. \tag{1.5}$$
  Some special cases are $\|A\|_1 = \|A\|_{1,1}$*,* $\|A\|_2 = \|A\|_{2,2}$ *and* $\|A\|_\infty = \|A\|_{\infty,\infty}$*, respectively given by*
$$\|A\|_1 = \max_j \sum_{i=1}^m |a_{ij}|, \qquad \|A\|_2 = \sqrt{\rho\left(A^\top A\right)}, \qquad \|A\|_\infty = \max_i \sum_{j=1}^n |a_{ij}|.$$
- *A system* Σ *defines a bounded linear mapping between sequences* $u \in \mathcal{U}, y \in \mathcal{Y}$*, such that* $y = \Sigma \circ u$*. One can introduce the normed space of linear systems* $\mathcal{L}(\mathcal{U}, \mathcal{Y})$ *with induced norm*
$$\|\Sigma\|_{\mathcal{L}(\mathcal{U},\mathcal{Y})} := \sup_{u\in\mathcal{U}\setminus\{0\}} \frac{\|\Sigma \circ u\|_{\mathcal{Y}}}{\|u\|_{\mathcal{U}}}. \tag{1.6}$$
  A special cases that is exploited in this project is $\|\Sigma\|_{\mathcal{L}_2} = \|\Sigma\|_{\mathcal{L}(\ell_2^{n_u}, \ell_2^{n_y})}$*, which is referred to as the system's induced* $\mathcal{L}_2$*-gain.*

Definition 1.3 ($\varkappa$-transform of vector-valued sequences)
The bilateral $\varkappa$-transform maps a n-dimensional complex vectors valued sequence $x : \mathbb{Z} \to \mathbb{C}^n$ (with x_i being the i-th sequence in x) into a n-dimensional complex vector valued function *$X \in \mathcal{H}_2^n$ (with X_i being the i-th sequence in X). This transform is denoted by the linear operator $\mathcal{Z} : \mathbb{C}^n \to \mathcal{H}_2^n$ and defined for each component x_i as*

$$X_i(\varkappa) = (\mathcal{Z} \circ x_i(n))(\varkappa) = \sum_{n=-\infty}^{\infty} x_i(n)\varkappa^{-n} \qquad \forall\, \varkappa \in \mathbb{C}. \tag{1.7}$$

The inverse $\varkappa$-transform is denoted by the linear operator $\mathcal{Z}^{-1} : \mathcal{H}_2^n \to \mathbb{C}^n$ and reads for stable $X_i(\varkappa)$, i.e., all poles of $X_i(\varkappa)$ are inside the open unit disc, as

$$x_i(n) = \mathcal{Z}^{-1} \circ X_i(\varkappa) = \frac{1}{2\pi} \int_{-\pi}^{\pi} X_i\left(e^{j\omega}\right) e^{j\omega n} d\omega \qquad \forall\, n \in \mathbb{Z}. \tag{1.8}$$

Definition 1.4 (N-point (Inverse) discrete Fourier transform of vector-valued sequences)

The N-point Discrete Fourier Transform (DFT) maps from x, a periodic sequence of N complex vectors of dimension n (with $x_i \in \mathbb{C}^n$ being the i-th vector in x) into X, another sequence of N complex vectors of dimension n (with $X_i \in \mathbb{C}^n$ being the i-th vector in X). This transform is denoted by the linear operator $\widehat{\mathcal{F}} : \mathbb{C}^{N\times n} \to \mathbb{C}^{N\times n}$ and defined for the component x_i as

$$X_i = \widehat{\mathcal{F}} \circ x_i, \qquad \text{where} \qquad X_i(k) = \sum_{n=0}^{N} x_i(n) e^{-j\tilde{\omega}_i(k)n} \qquad \forall k \in \{0, 1, \ldots, N-1\}. \tag{1.9}$$

We recognize the discrete set of so-called normalised DFT frequencies as $\tilde{\omega}_i$, where $\tilde{\omega}_i(k) = 2\pi k/N$. An inverse procedure, which is referred to as Inverse discrete Fourier transform (IDFT), is denoted by the linear operator $\widehat{\mathcal{F}}^{-1} : \mathbb{C}^{N\times n} \to \mathbb{C}^{N\times n}$ and defined as

$$x_i = \widehat{\mathcal{F}}^{-1} \circ X_i, \qquad \text{where} \qquad x_i(k) = \frac{1}{N} \sum_{n=0}^{N-1} X_i(n) e^{j\,\tilde{\omega}_i(k)n} \qquad \forall k \in \{0, 1, \ldots, N-1\}. \tag{1.10}$$

Definition 1.5 (Comparison functions [27, def. 4.2 & 4.3])

- *A continuous function $\alpha : [0, a) \to [0, \infty)$ is said to belong to class $\mathcal{K}$ if:*
  1. *it is strictly increasing.*
  2. *$\alpha(0) = 0$.*
- *A continuous function $\alpha : [0, \infty) \to [0, \infty)$ is said to belong to class $\mathcal{K}_\infty$ if*
  1. *it belongs to class $\mathcal{K}$.*
  2. *$\lim_{r\to\infty} \alpha(r) = \infty$.*
- *A continuous function $\beta : [0, a) \times [0, \infty) \to [0, \infty)$ belongs to class $\mathcal{KL}$ if*
  1. *for each fixed r:*
     (a) *the mapping $\beta(r, s)$ is decreasing with respect to s.*
     (b) *$\lim_{s\to\infty} \beta(r, s) = 0$.*
  2. *for each fixed s: the mapping $\beta(r, s)$ belongs to class $\mathcal{K}$ with respect to r.*

Chapter 2

Lur'e-type models

The purpose of this chapter is to elaborate on important preliminaries for Lur'e-type models. To that end, the Lur'e-type structure is introduced first, after which the remainder of this chapter is structured as follows. Section 2.1 discusses assumptions and instrumental properties of the linear block. The nonlinear block is treated in Section 2.2 in a similar fashion. Section 2.3 discusses the problem of non-unique parametrisation of the Lur'e-type structure, after which a summary of this chapter follows in Section 2.4.

Let us consider a discrete-time Lur'e-type structure, expressed into the following form:

$$\Sigma : \begin{cases} x(k+1) = Ax(k) + Bu(k) + Lw(k), \quad x(0) = x_0, & \text{(2.1a)} \\ y(k) = Cx(k) + Dw(k), & \text{(2.1b)} \\ z(k) = Fx(k) + Gu(k) + Hw(k), & \text{(2.1c)} \end{cases}$$

$$\varphi : \{ \quad u(k) = \varphi(y(k), k) \quad . \qquad \text{(2.1d)}$$

In this set of equations, one finds $x(k) \in \mathbb{R}^{n_x}$, $z(k) \in \mathbb{R}^{n_z}$ and $y(k) \in \mathbb{R}^{n_y}$ as the evaluation at time-instant $k \in \mathbb{Z}$ of state x, measured output z and unmeasured nonlinearity input y, respectively. Moreover, $u(k) \in \mathbb{R}^{n_u}$ and $w(k) \in \mathbb{R}^{n_w}$ are recognised as the evaluations of the nonlinearity output u and measured periodic external excitation w at this moment in time. Periodicity of discrete time signals is defined as follows SISO and SISO.

Definition 2.1 (Discrete periodic signals [62, p. 121])
A discrete-time signal $x : \mathbb{Z} \to \mathbb{R}^n$ *is* periodic *if there exists a non-zero integer* $p \in \mathbb{Z}$ *such that*

$$x(n+p) = x(n), \qquad \forall\, n \in \mathbb{Z}. \tag{2.2}$$

The Lur'e-type structure (2.1) can be split into (i) a LTI block Σ having a representation according to Equations (2.1a) to (2.1c) and (ii) a nonlinear block φ following Equation (4.17d). These dynamics can then be interpreted as the feedback interconnection that is shown in Figure 2.1. Throughout this work, a discrete-time Lur'e-type system can be referred to via either Equation (2.1) or the tuple (Σ, φ).

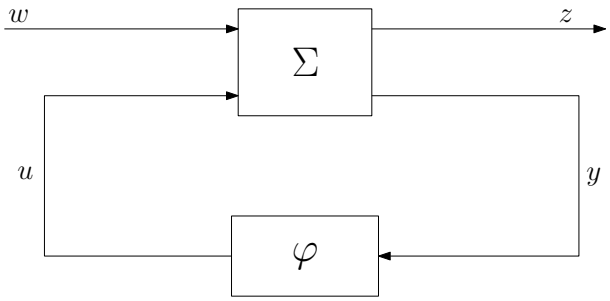


Figure 2.1: Lur'e-type system interconnection.

2.1 Linear block

One recognizes the state-space representation of a finite-dimensional MIMO LTI causal discrete-time dynamical system Σ in Equations (2.1a) to (2.1c). The state equation (2.1a) is parametrised by state matrix $A \in \mathbb{R}^{n_x \times n_x}$, together with input matrices $B \in \mathbb{R}^{n_x \times n_u}$ and $L \in \mathbb{R}^{n_x \times n_w}$. The output equations (2.1b) and (2.1c) are characterised by output matrices $C \in \mathbb{R}^{n_y \times n_x}$ and $F \in \mathbb{R}^{n_z \times n_x}$, as well as feedthrough matrices $D \in \mathbb{R}^{n_y \times n_w}$, $G \in \mathbb{R}^{n_z \times n_u}$ and $H \in \mathbb{R}^{n_z \times n_w}$. Given an initial condition $x_0 \in \mathbb{R}^{n_x}$, this state-space representation can be used to fully describe the behaviour of Σ under the time-varying excitation w and nonlinear feedback u. This section starts by discussing assumptions on linear block Σ in the Lur'e-type structure.

Assumption 2.1 (Linear block)
Throughout this thesis, the LTI block Σ respects certain properties, which can be listed as follows:

1. *System Σ is represented via a state-space form to which one can refer to as $(A_\Sigma, B_\Sigma, C_\Sigma, D_\Sigma)$. This representation is defined in terms of the parametrisation of Equation* (2.1) *via*

$$A_\Sigma = A, \qquad B_\Sigma = \begin{bmatrix} B & L \end{bmatrix}, \qquad C_\Sigma = \begin{bmatrix} C \\ F \end{bmatrix}, \qquad D_\Sigma = \begin{bmatrix} 0_{n_y} & D \\ G & H \end{bmatrix}. \tag{2.3}$$

2. *$(A_\Sigma, B_\Sigma, C_\Sigma, D_\Sigma)$ is a minimal realization of Σ, such that the system is described by the minimum number of state variables.*
3. *The nonlinearity input y and nonlinearity output u are of equal dimension, i.e., $n_y = n_u$.*

The feedback interconnection in Figure 2.1 reveals a clear nonlinear algebraic loop in the Lur'e-type structure. This direct feedthrough from nonlinearity output u into nonlinearity input y is excluded from the considered Lur'e-type structure. An alternative method to deal with this challenge is mentioned in [23] and assumes a unit sample delay in either the forward or the backward path. Finally, note that the second item in the list of assumptions on the linear block can be guaranteed via the controllability and observability properties verified by the well-known Kalman test [63, Th. 8.9].

All important assumptions on Σ are now explained, so in the next paragraph the focus lies on key properties of these LTI dynamics. In particular, we aim to analyse solutions in the frequency domain. Let us once more consider the state-space representation of LTI system Σ in Equations (2.1a) to (2.1c). It is well known that the system output responses of these dynamics can be written as

$$y(k) = \underbrace{CA^k x_0}_{\text{I}} + \underbrace{\sum_{i=1}^{k} CA^{k-i-1}Bu(i) + \sum_{i=1}^{k} CA^{k-i-1}Lw(i) + Dw(k)}_{\text{II}}, \tag{2.4a}$$

$$z(k) = \underbrace{FA^k x_0}_{\text{III}} + \underbrace{\sum_{i=1}^{k} FA^{k-i-1}Bu(i) + \sum_{i=1}^{k} FA^{k-i-1}Lw(i) + Gu(k) + Hw(k)}_{\text{IV}}. \tag{2.4b}$$

In these equalities, terms I and III are recognised as the so-called *free* response contributions to y and z, respectively. Furthermore, terms II and IV denote the *forced* response contributions to the respective system outputs. For Schur stable LTI systems, i.e., $\lambda(A) \subset \mathbb{C}_\ominus$, it is known that free responses decay asymptotically to zero. Because the solutions 'forget' their initial conditions, only the forced solution in Equation (2.4) remains forward in time. Each solution then converges to this so-called *steady-state* solution, which is unique and exists on the entire time axis $k \in \mathbb{Z}$ without growing unbounded, as is formalised in Property 2.1.

Property 2.1 *Consider a LTI dynamical system Σ according to Equations (2.1a) to (2.1c) having a state-space representation $(A_\Sigma,\, B_\Sigma,\, C_\Sigma,\, D_\Sigma)$. The dynamics are subjected to T-periodic bounded inputs w, u. If the state matrix A_Σ is Schur stable, then there exists a unique T-periodic globally exponentially stable steady-state solution $\left(\bar{x} \in \ell_2^{n_x},\, \bar{y} \in \ell_2^{n_y},\, \bar{z} \in \ell_2^{n_z}\right)$.*

Note that the macron diacritic was chosen to denote steady-state signals. So-called *time-domain steady-state operators* are defined to evaluate the T-periodic steady-state outputs of Σ in time-domain while being subjected to T-periodic inputs for all $m \in \{1, \ldots, T\}$ according to

$$\bar{y}^w = \hat{\mathcal{F}}_{yw} \circ w, \qquad \text{where} \qquad \left(\hat{\mathcal{F}}_{yw} \circ w\right)(m) = \textstyle\sum_{i=-\infty}^{m} CA^{m-i-1}Lw(i) + Dw(m), \tag{2.5a}$$

$$\bar{y}^u = \hat{\mathcal{F}}_{yu} \circ u, \qquad \text{where} \qquad \left(\hat{\mathcal{F}}_{yu} \circ u\right)(m) = \textstyle\sum_{i=-\infty}^{m} CA^{m-i-1}Bu(i), \tag{2.5b}$$

$$\bar{z}^w = \hat{\mathcal{F}}_{zw} \circ w, \qquad \text{where} \qquad \left(\hat{\mathcal{F}}_{zw} \circ w\right)(m) = \textstyle\sum_{i=-\infty}^{m} FA^{m-i-1}Lw(i) + Hw(m), \tag{2.5c}$$

$$\bar{z}^u = \hat{\mathcal{F}}_{zu} \circ u, \qquad \text{where} \qquad \left(\hat{\mathcal{F}}_{zu} \circ u\right)(m) = \textstyle\sum_{i=-\infty}^{m} FA^{m-i-1}Bu(i) + Gu(m). \tag{2.5d}$$

Also, due to the superposition principle of LTI dynamics [63, p. 57], it holds

$$\bar{y} = \bar{y}^w + \bar{y}^u, \tag{2.6}$$

$$\bar{z} = \bar{z}^w + \bar{z}^u. \tag{2.7}$$

These definitions are used in Appendix A.11, where a contraction mapping property is proven that is used in Chapter 4. According to [64, Ch. 9], it is generally preferred to avoid calculating the convolution sums in Equation (2.5), because (i) the operation is mathematically difficult to deal with, e.g., suppose we are given an output sequence and system description, how can we then easily reconstruct the inputs sequence? and (ii) evaluating this operation requires many algebraic actions, which is computationally unattractive. To overcome these issues, fundamental results in linear systems theory often make use of analysis techniques in the frequency domain, where the convolution sum operation reduces to a single multiplication. There also exist computationally cheap methods to transform signals forth and back between the time- and frequency-domain and, therefore, it is computationally attractive to compute model responses in the frequency domain. In the following, one first introduces Σ's transfer function representation for which an upper-bound to its $\mathcal{H}_\infty$-norm is derived. Thereafter, Σ's frequency response function is treated, from which linear steady-state operators that map periodic system inputs to periodic steady-state system outputs can be derived.

Transfer function representation

Next to the state-space formulation, there exist numerous representations of Σ, such as the difference equation, impulse-response representation and transfer function representation. This last form is characterised by the *transfer matrix* $\mathcal{G}_\Sigma(\varkappa)$ for the complex variable $\varkappa \in \mathbb{C}$, which exactly relates the $\varkappa$-transform (as per Definition 1.3) of periodic system inputs to the $\varkappa$-transform of steady-state system outputs. To distinguish between the respective input- and output channels of Σ, the transfer matrix can be written as

$$\begin{pmatrix} \bar{Y}(\varkappa) \\ \bar{Z}(\varkappa) \end{pmatrix} = \underbrace{\left(\begin{array}{c|c} \mathcal{G}_{\Sigma_{yw}}(\varkappa) & \mathcal{G}_{\Sigma_{yu}}(\varkappa) \\ \hline \mathcal{G}_{\Sigma_{zw}}(\varkappa) & \mathcal{G}_{\Sigma_{zu}}(\varkappa) \end{array}\right)}_{\mathcal{G}_\Sigma(z)} \begin{pmatrix} W(\varkappa) \\ U(\varkappa) \end{pmatrix}, \tag{2.8}$$

in which $\varkappa$-transforms of signals are denoted by capital letters and $\varkappa$ arguments. This expression allows for a characterization of the transfer matrix in terms of the parametrisation of Σ via

$$\mathcal{G}_{\Sigma_{yw}}(\varkappa) = C(\varkappa I_{n_x} - A)^{-1}L + D, \tag{2.9a}$$

$$\mathcal{G}_{\Sigma_{yu}}(\varkappa) = C(\varkappa I_{n_x} - A)^{-1}B, \tag{2.9b}$$

$$\mathcal{G}_{\Sigma_{zw}}(\varkappa) = F(\varkappa I_{n_x} - A)^{-1}L + H, \tag{2.9c}$$

$$\mathcal{G}_{\Sigma_{yu}}(\varkappa) = F(\varkappa I_{n_x} - A)^{-1}B + G. \tag{2.9d}$$

Note that $\mathcal{G}_\Sigma(\varkappa)$ is unique, whereas Σ may have infinitely many state-space realizations. Also, $\mathcal{G}_\Sigma(\varkappa)$ is known to be proper and rational, since it would not have a state-space representation otherwise.

Frequency response function

The *frequency response function* of system Σ is defined as the evaluation of the transfer matrix $\mathcal{G}_\Sigma(\varkappa)$ on a uniformly sampled subset of $\mathbb{C}_\bigcirc$. This operator is denoted by $\mathcal{G}_\Sigma(e^{j\tilde{\omega}})$ for so-called normalised DFT frequencies $\tilde{\omega}$, where $\tilde{\omega}(k) = 2\pi k/N$ for all $k \in \{0, 1, \ldots, N-1\}$. If a periodic signal with measurement interval N is analysed, whose length is an integer multiple of the excitation period, then it is well known that the N-point DFT (as per Definition 1.4) of steady-state LTI system outputs and the N-point DFT of harmonic inputs are exactly related via the frequency response function $\mathcal{G}_\Sigma(e^{j\tilde{\omega}})$ [65].

Because of this exact relation, one can introduce so-called *frequency-domain steady-state operators* for the LTI system Σ, which map from periodic excitations to periodic steady-state outputs. These operators involve element-wise multiplication with the frequency response function and are defined for all $m \in \{1, \ldots, N\}$ as

$$\bar{Y}^w = \hat{\mathcal{F}}_{YW} \circ W, \qquad \text{where} \qquad \left(\hat{\mathcal{F}}_{YW} \circ W\right)(m) = \mathcal{G}_{\Sigma_{yw}}(e^{j\tilde{\omega}(m-1)})\, W(m), \tag{2.10a}$$

$$\bar{Y}^u = \hat{\mathcal{F}}_{YU} \circ U, \qquad \text{where} \qquad \left(\hat{\mathcal{F}}_{YU} \circ U\right)(m) = \mathcal{G}_{\Sigma_{yu}}(e^{j\tilde{\omega}(m-1)})\, U(m)\,, \tag{2.10b}$$

$$\bar{Z}^w = \hat{\mathcal{F}}_{ZW} \circ W, \qquad \text{where} \qquad \left(\hat{\mathcal{F}}_{ZW} \circ W\right)(m) = \mathcal{G}_{\Sigma_{zw}}(e^{j\tilde{\omega}(m-1)})\, W(m), \tag{2.10c}$$

$$\bar{Z}^u = \hat{\mathcal{F}}_{ZU} \circ U, \qquad \text{where} \qquad \left(\hat{\mathcal{F}}_{ZU} \circ U\right)(m) = \mathcal{G}_{\Sigma_{zu}}(e^{j\tilde{\omega}(m-1)})\, U(m). \tag{2.10d}$$

Here, e.g., $\bar{Y}^w \in \mathbb{C}^{N\times n_y}$ is recognised as the N-point DFT of steady-state output $\bar{y}^w$ driven by w. This concludes the discussion on LTI system properties and analysis techniques. All important assumptions on the linear block are listed, whereafter it is shown that steady-state solutions of linear systems can be solved efficiently in the frequency domain. Finally, it is shown that the induced $\mathcal{L}_2$-gain on the linear block can be found by solving a linear matrix inequality (LMI) optimisation problem, which is used in Chapters 3 and 4. An instrumental property of the transfer matrix representation reads as follows.

Property 2.2 ([66], [67, Th. 11.3]) *Consider a stable discrete-time LTI system Σ_{yu}, subject to input u and generating output y^u. Suppose that the system can be represented via the transfer matrix $\mathcal{G}_{\Sigma_{yu}}(\varkappa) : \ell_2^{n_u} \to \ell_2^{n_y}$. Then, by Definition 1.2, it holds*

$$\left\|\bar{Y}^u\right\|_{\ell_2^{n_y}} = \left\|(\mathcal{G}_{\Sigma_{yu}} U)\right\|_{\ell_2^{n_y}} \leq \|\Sigma_{yu}\|_{\mathcal{L}_2} \|U\|_{\ell_2^{n_u}} = \left\|\mathcal{G}_{\Sigma_{yu}}(\varkappa)\right\|_{\mathcal{H}_\infty} \|U\|_{\ell_2^{n_u}}. \tag{2.11}$$

In this property, one recognise the equivalence between a LTI system's induced $\mathcal{L}_2$-gain and the $\mathcal{H}_\infty$-norm of its transfer matrix [67, Th. 11.3]. Crucially, via this equivalence, it is possible to express a system property into a signal norm that can be evaluated by solving an LMI optimisation problem for which efficient solution methods exist. The well-known discrete-time Bounded-Real lemma is formalised below.

Lemma 2.1 (Discrete-time Bounded-Real Lemma [68, Lem. 5.1])
Consider a discrete-time LTI system Σ whose transfer matrix $\mathcal{G}_\Sigma(\varkappa)$ admits a state-space representation $(A,\ B,\ C,\ D)$, where $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n_x\times n_u}$, $C \in \mathbb{R}^{n_y\times n_x}$, $D \in \mathbb{R}^{n_y\times n_u}$. Then, the inequality $\|\mathcal{G}_\Sigma(\varkappa)\|_{\mathcal{H}_\infty} < \gamma$ holds if and only if there exist $0 \prec P \in \mathbb{S}^{n_x}$ and $\gamma \in \mathbb{R}_{>0}$, such that

$$\begin{bmatrix} A^\top PA - P & A^\top PB & C^\top \\ \star & B^\top PB - \gamma I_p & D^\top \\ \star & \star & -\gamma I_p \end{bmatrix} \prec 0. \tag{2.12}$$

The $\mathcal{H}_\infty$ norm of $\mathcal{G}_\Sigma(\varkappa)$ is the minimum value of $\gamma \in \mathbb{R}_{>0}$ that satisfies the above condition. If $(A,\ B,\ C,\ D)$ is a minimal realization, then the matrix inequalities can be non-strict.

Proof. See [68, Lem. 5.1]. □

The LMI optimisation problem in Lemma 2.1 is applied multiple times throughout this project. On the one hand, in Chapter 3 we show that once a Lur'e-type structure satisfies a certain stability notion, then this

implies an upper-bound to the $\mathcal{H}_\infty$-norm of its transfer matrix. On the other hand, this upper-bound is applied together with Property 2.2 in the contraction analysis of an iterative algorithm in Chapter 4.

2.2 Nonlinear block

This section treats the most important assumptions that were made on the nonlinear block φ. These assumptions partly define the model class and impose important properties on the nonlinear block as explained in the remainder of this section. The notions of (incremental) sector conditions are introduced first in Definitions 2.2 and 2.3, respectively. Consecutively, the assumptions on φ are listed in Assumption 2.2.

Definition 2.2 (Sector condition [27, Def. 6.2])
A memoryless function $\varphi : \mathbb{R}^{n_y} \times \mathbb{R}_{>0} \to \mathbb{R}^{n_y}$ is said to be sector bounded within sector bounds $[K_1,\ K_2]$, for real symmetric $K_1,\ K_2 \in \mathbb{S}^{n_y}$ with $K_2 - K_1 \succ 0$, if this function satisfies:

$$[\varphi(y,t) - K_1 y]^\top [\varphi(y,t) - K_2 y] \leq 0, \tag{2.13}$$

for all $y \in \mathbb{R}^{n_y}$ and $t \in \mathbb{R}_{>0}$.

Definition 2.3 (Incremental sector condition [69, Def. 4])
A memoryless function $\varphi : \mathbb{R}^{n_y} \times \mathbb{R}_{>0} \to \mathbb{R}^{n_y}$ is said to be incrementally sector bounded within sector bounds $[S_1, S_2]$, for real symmetric $S_1,\ S_2 \in \mathbb{S}^{n_y}$ with $S_2 - S_1 \succ 0$, if this function satisfies

$$\left[\varphi(y^a,t) - \varphi(y^b,t) - S_1(y^a - y^b)\right]^\top \left[\varphi(y^a,t) - \varphi(y^b,t) - S_2(y^a - y^b)\right] \leq 0, \tag{2.14}$$

for all $y^a, y^b \in \mathbb{R}^{n_y}$ and $t \in \mathbb{R}_{>0}$.

Assumption 2.2 (Nonlinear block assumptions)

1. *Nonlinear block φ is memoryless, which ensures that the nonlinearity output at any instant of time is determined uniquely by its input at that time instant. Hence, it does not depend on the input history [27, p. 19]. This property is also referred to as zero-memory or static.*
2. *Nonlinear block φ is decentralized in the sense that for any time instant k, its i^{th} output $\varphi_i(y_i(k),\ k)$ depends solely on its respective i^{th} input $y_i(k)$ [27, p. 228]. Therefore, one can write*

   $$\varphi(y(k),\ k) = \begin{bmatrix}\varphi_1(y_1(k),\ k) & \varphi_2(y_2(k),\ k) & \dots & \varphi_p(y_p(k),\ k)\end{bmatrix}^\top, \tag{2.15}$$

   which clearly shows the decoupled input-output behaviour of the nonlinearity.
3. *The nonlinearity φ is* sector bounded *as per Definition 2.2 within known sector bounds $[0_{n_y}, \Omega]$. This assumption is verified if*

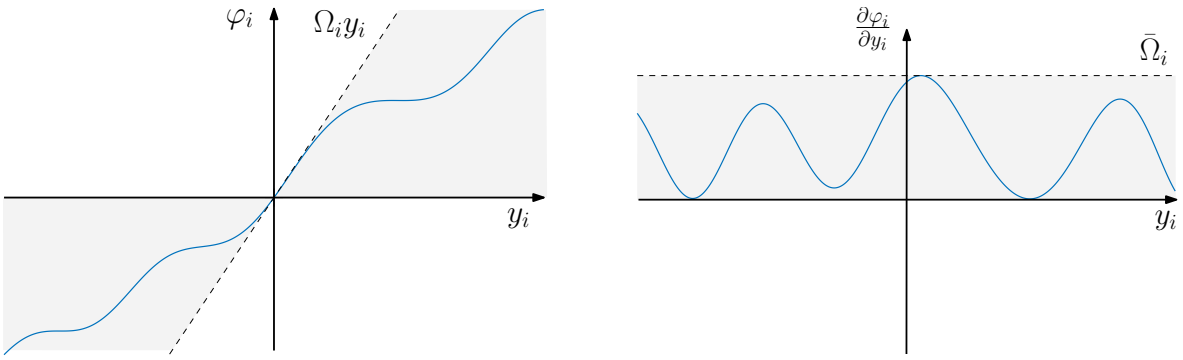
   $$s_\Omega(R,y) := \varphi(y,k)^\top R(\varphi(y,k) - \Omega y) \leq 0 \quad \forall y \in \mathbb{R}^{n_y},\ k \in \mathbb{Z}, \tag{2.16}$$

   for any diagonal $0 \prec R \in \mathbb{S}^{n_y}$. The diagonal matrix $0 \prec \Omega \in \mathbb{S}^{n_y}$ with $\operatorname{diag}\left(\left\{\Omega_1, \dots, \Omega_{p_y}\right\}\right) = \Omega$ and $\Omega_i > 0,\ \forall i \in \{1, \dots, p_y\}$ fully characterizes this condition.
4. *The nonlinearity φ is* incrementally sector bounded *as per Definition 2.3 within known incremental sector bounds $[0_{n_y}, \bar{\Omega}]$. This assumption is verified if*

   $$s_{\bar{\Omega}}(R, y^a, y^b) := \left(z^a - z^b\right)^\top R\left(z^a - z^b - \bar{\Omega}(y^a - y^b)\right) \leq 0 \quad \forall y^a,\ y^b \in \mathbb{R}^{n_y},\ k \in \mathbb{Z}, \tag{2.17}$$

   for any diagonal $0 \prec R \in \mathbb{S}^{n_y}$ and where $z^i = \varphi(y^i, k)$. The diagonal matrix $0 \prec \bar{\Omega} \in \mathbb{S}^{n_y}$ with $\operatorname{diag}\left(\left\{\bar{\Omega}_1, \dots, \bar{\Omega}_{p_y}\right\}\right) = \bar{\Omega}$ and $\bar{\Omega}_i > 0,\ \forall i \in \{1, \dots, n_y\}$ fully characterizes this condition.

The third and fourth item in the list of assumptions are from now on referred to as the *cone bounded sector constraints* on nonlinearity φ. In the general MIMO case ($p_y > 1$), it is hard to find and interpret the cone bounded sector constraints on a certain nonlinearity φ. However, if we recall that φ is decentralized in the scope of this project, then we can think of these constraints as a set of SISO cone bounded constraints on the individual nonlinearity outputs $\varphi_i(y_{i,k},\, k)$. Fortunately, it is relatively easy to find tight bounds on the (incremental) sector of a SISO nonlinear function, which can be understood for an example nonlinearity via the visual interpretations in Figure 2.2.



(a) Sector bounded within $[0, \Omega_i]$.

(b) Incrementally sector bounded within $[0, \bar{\Omega}_i]$.

Figure 2.2: Cone bounded sector constraints on $\varphi_i(y_i,\, k)$.

The nonlinearity output φ_i is clearly constrained by a sector condition to the grey area in Figure 2.2a, whose boundary is characterized by $\Omega_i > 0$. The sector clearly imposes y_i and φ_i to be of equal sign, which is also referred to as $\varphi_i(y_i)$ lying in the first and third quadrant. Note that this sector condition imposes each nonlinearity output φ_i to pass through the origin, such that

$$\varphi_i(0, k) = 0 \qquad \forall i \in \{1, \ldots, n_y\},\ k \in \mathbb{Z}. \tag{2.18}$$

Besides that, the nonlinearity slope $\frac{\partial \varphi_i}{\partial y_i}$ is clearly constrained by a incremental sector condition to the grey area in Figure 2.2b, whose boundary is characterized by $\bar{\Omega}_i$. Here we can see that the incremental sector condition defines φ_i as *monotonic* since it is non-decreasing on its entire domain. Finally, observe that multiplying a SISO (incremental) sector condition by a nonnegative constant can never make the function violate its constraint. This fact allowed for the inclusion of weight matrix R in Equations (2.16) and (2.17).

2.3 Non-unique model parametrisation

A well-known issue of block-oriented nonlinear models, including the model class (2.1), is related to the uniqueness of the model parametrization [23]. Many different models inside this model class result in the same input-output behaviour even though they are parametrised differently. This phenomenon is undesired in the scope of system identification, as is illustrated later on in Chapter 5. In the following section, loop scaling, loop transformation and similarity transformations are identified as sources of non-unique parametrisations inside the model class. Also, a normalised, symmetric Lur'e-type model form is presented, which is used in Chapters 3 and 4 for stability analysis and system identification purposes, respectively.

Loop scaling

Loop scaling can be thought of as a linear scaling of the nonlinearity output and input channels, respectively, u and y. To this end, we can introduce the diagonal matrix $\Psi_1 \in \mathbb{S}^{n_u}$ to scale the nonlinearity output channel according to $u = \Psi_1 \tilde{u}$. Furthermore, the diagonal matrix $\Psi_2 \in \mathbb{S}^{n_y}$ scales the nonlinearity input channel via $y = \Psi_2^{-1}\tilde{y}$. Loop scaling is shown schematically as the block diagram in Figure 2.3.

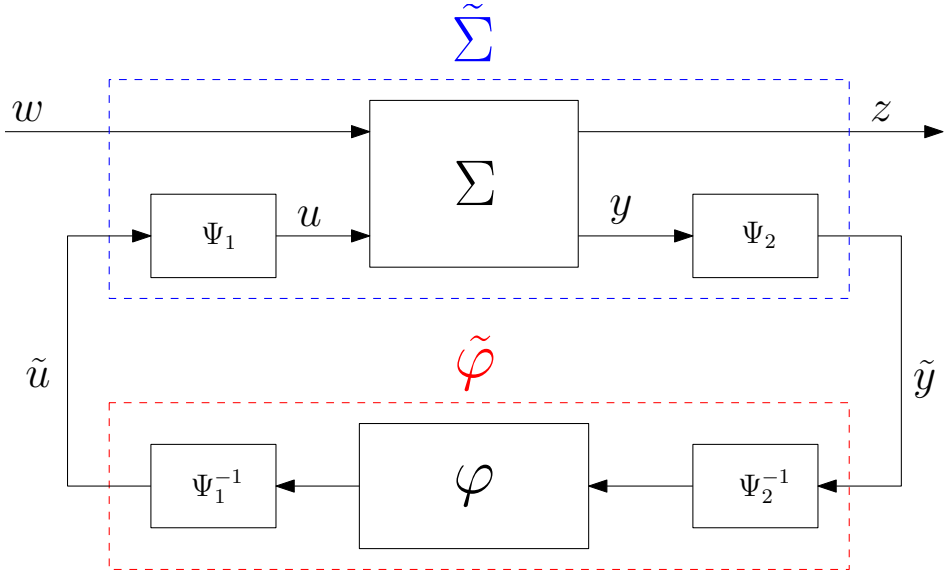


Figure 2.3: Loop scaling interconnection.

One can see that loop scaling gives rise to an updated Lur'e-type parametrisation $(\tilde{\Sigma},\, \tilde{\varphi})$, which is different from the original model $(\Sigma,\, \varphi)$, but still inside the model class and producing an equivalent input-output behaviour. This Lur'e-type model satisfies the properties that are explained in Lemma 2.2.

Lemma 2.2 (Loop scaling)
Consider a Lur'e-type system $(\Sigma,\, \varphi)$ according to Equation (2.1). Let Σ and φ verify Assumptions 2.1 and 2.2, respectively. If we introduce the diagonal and invertible nonlinearity input- and output scaling matrices $\Psi_1 \in \mathbb{S}^{n_u}$ and $\Psi_2 \in \mathbb{S}^{n_y}$, such that $u = \Psi_1\tilde{u}$ and $y = \Psi_2^{-1}\tilde{y}$, then we find an updated Lur'e-type model $(\tilde{\Sigma}, \tilde{\varphi})$, for which:

- *The linear block $\tilde{\Sigma}$ verifies Assumption 2.1 for the state-space form $(A_{\tilde{\Sigma}},\, B_{\tilde{\Sigma}},\, C_{\tilde{\Sigma}},\, D_{\tilde{\Sigma}})$, where*

$$\begin{aligned} A_{\tilde{\Sigma}} &= A, & B_{\tilde{\Sigma}} &= \begin{bmatrix} B\Psi_1 & L \end{bmatrix}, \\ C_{\tilde{\Sigma}} &= \begin{bmatrix} \Psi_2 C \\ F \end{bmatrix}, & D_{\tilde{\Sigma}} &= \begin{bmatrix} 0_{n_y} & \Psi_2 D \\ G\Psi_1 & H \end{bmatrix}. \end{aligned} \tag{2.19}$$

- *The nonlinear block $\tilde{\varphi}$ is defined as*

$$\tilde{\varphi}(\tilde{y}) := \Psi_1^{-1}\varphi(\Psi_2^{-1}\tilde{y}) \tag{2.20}$$

and verifies Assumption 2.2 with updated bounds on (i) the sector condition: $[0_{n_y}, \Psi_1^{-1}\Omega\Psi_2^{-1}]$ and (ii) the incremental sector condition: $[0_{n_y}, \Psi_1^{-1}\bar{\Omega}\Psi_2^{-1}]$.

Proof. See Appendix A.1 □

Loop transformation

Via a loop transformation, one can add or subtract a purely linear term to the nonlinear block of the Lur'e-type model structure, while compensating for this in the linear block parametrisation [27, p. 255]. This transformation is established if the nonlinearity output u is replaced by $\tilde{u} = u - \Psi_3 y$ given a diagonal loop-transformation matrix $\Psi_3 \in \mathbb{S}^{n_u}$. The loop transformation is shown schematically in Figure 2.4.

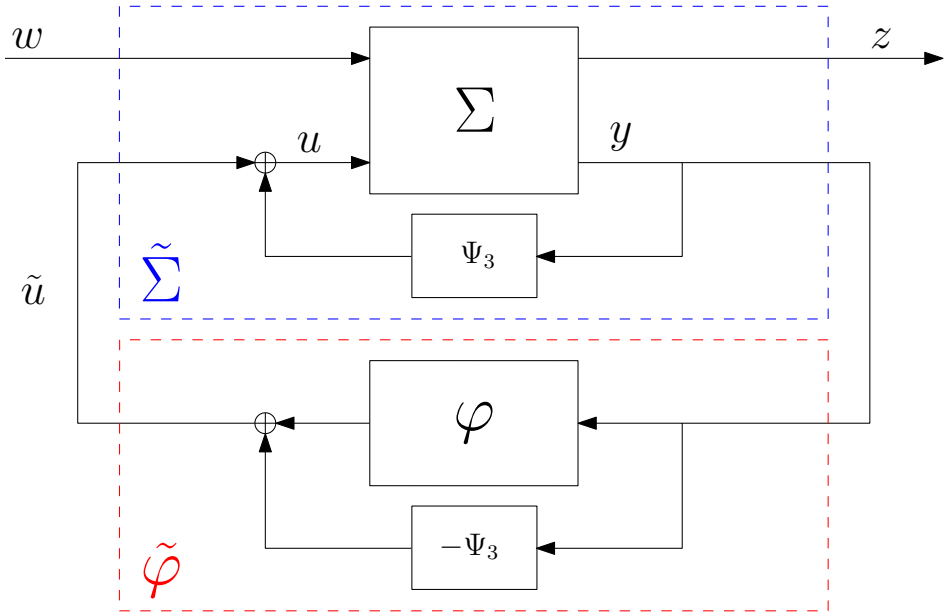


Figure 2.4: Loop transformation interconnection.

One can see that loop transformation gives rise to an updated Lur'e-type parametrisation $(\tilde{\Sigma},\ \tilde{\varphi})$, which is different from the original model $(\Sigma,\ \varphi)$, but still inside the model class and producing an equivalent input output behaviour. The properties of $(\tilde{\Sigma},\ \tilde{\varphi})$ follow the results in Lemma 2.3.

Lemma 2.3 (Loop transformation)
Consider a Lur'e-type system $(\Sigma,\ \varphi)$ according to Equation (2.1). Let Σ and φ verify Assumptions 2.1 and 2.2, respectively. If one introduces the diagonal loop-transformation matrix $\Psi_3 \in \mathbb{S}^{n_y}$ such that $u = \tilde{u} + \Psi_3 y$, then, an updated Lur'e-type model $(\tilde{\Sigma}, \tilde{\varphi})$ is retrieved, for which:

- *The linear block $\tilde{\Sigma}$ verifies Assumption 2.1 for an updated state-space form $(A_{\tilde{\Sigma}},\ B_{\tilde{\Sigma}},\ C_{\tilde{\Sigma}},\ D_{\tilde{\Sigma}})$, where*

$$\begin{aligned} A_{\tilde{\Sigma}} &= A + B\Psi_3 C, & B_{\tilde{\Sigma}} &= \begin{bmatrix} B & (L + B\Psi_3 D) \end{bmatrix}, \\ C_{\tilde{\Sigma}} &= \begin{bmatrix} C \\ F + G\Psi_3 C \end{bmatrix}, & D_{\tilde{\Sigma}} &= \begin{bmatrix} 0_{n_y} & D \\ G & H + B\Psi_3 D \end{bmatrix}. \end{aligned} \tag{2.21}$$

- *The nonlinear block $\tilde{\varphi}$ is defined as*

$$\tilde{\varphi}(y, k) := \varphi(y) - \Psi_3 y \tag{2.22}$$

and verifies Assumption 2.2 with updated bounds on (i) the sector condition: $[-\Psi_3, \Omega - \Psi_3]$ and (ii) the incremental sector condition: $[-\Psi_3, \bar{\Omega} - \Psi_3]$.

Proof. See Appendix A.2 □

A normalised, symmetric Lur'e-type system

The loop scaling and loop transformation methods in Lemmas 2.2 and 2.3, respectively, can be used to cast any Lur'e-type model (2.1) into a form with an equivalent input-output behaviour, whose nonlinearity is incrementally sector bounded within known incremental sector bounds $[-\frac{1}{2}I_{n_y}, \frac{1}{2}I_{n_y}]$. The loop scaling and loop transformation operations on (Σ, φ) are shown schematically as the block diagram in Figure 2.5.

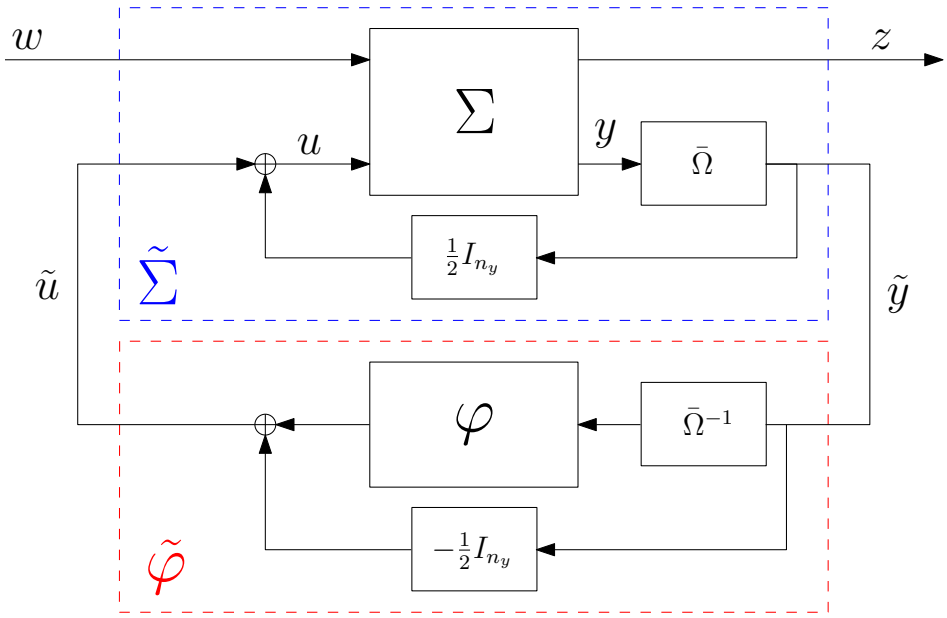


Figure 2.5: Transformation towards a normalised, symmetric Lur'e-type model.

The properties of this transformation, as formalised in Lemma 2.4, are applied multiple times in this project; e.g. Section 3.5 and Theorem 4.2 extensively use the normalised, symmetric form of a Lur'e-type system.

Lemma 2.4 (Normalized symmetric Lur'e-type structures)
Consider a Lur'e-type system (Σ, φ) according to Equation (2.1). Let Σ and φ verify Assumptions 2.1 and 2.2, respectively. Let us perform a loop scaling as per Lemma 2.2 with $\Psi_1 = I_{n_y}$ and $\Psi_2 = \bar{\Omega}$, followed by a loop transformation as per Lemma 2.3 of $\Psi_3 = \frac{1}{2}I_{n_y}$. One then finds an updated Lur'e-type model $(\tilde{\Sigma}, \tilde{\varphi})$ for which:

- *The linear block $\tilde{\Sigma}$ verifies Assumption 2.1 and admits a state-space form $(A_{\tilde{\Sigma}}, B_{\tilde{\Sigma}}, C_{\tilde{\Sigma}}, D_{\tilde{\Sigma}})$, where*

$$\begin{aligned} A_{\tilde{\Sigma}} &= A + \frac{1}{2}B\bar{\Omega}C, & B_{\tilde{\Sigma}} &= \begin{bmatrix} B & L + \frac{1}{2}B\bar{\Omega}D \end{bmatrix}, \\ C_{\tilde{\Sigma}} &= \begin{bmatrix} \bar{\Omega}C \\ F + \frac{1}{2}G\bar{\Omega}C \end{bmatrix}, & D_{\tilde{\Sigma}} &= \begin{bmatrix} 0_{n_y} & \bar{\Omega}D \\ G & H + \frac{1}{2}B\bar{\Omega}D \end{bmatrix}. \end{aligned} \tag{2.23}$$

- *The nonlinear block $\tilde{\varphi}$ is defined as*

$$\tilde{\varphi}(\tilde{y}, k) := \varphi(\bar{\Omega}^{-1}\tilde{y}, k) - \frac{1}{2}y \tag{2.24}$$

and verifies Assumption 2.2 with updated bounds on (i) the sector condition: $[-\frac{1}{2}I_{n_y}, \bar{\Omega}^{-1}\Omega - \frac{1}{2}I_{n_y}]$ and (ii) the normalised symmetric *incremental sector condition: $[-\frac{1}{2}I_{n_y}, \frac{1}{2}I_{n_y}]$.*

Linear block similarity transformation

The final source of non-unique Lur'e-type model parametrisation that is discussed in this work addresses *similarity transformations* of the linear block Σ. It is well known from linear system theory, that any one-to-one linear transformation of state variables does not alter the input-output behavior of a LTI system and, hence, induces non-unique parametrisations of Σ [70, p. 658]. This result is formalised in Lemma 2.5.

Lemma 2.5 (Similarity transformation)
Consider a LTI system Σ that admits the state-space representation $(A_\Sigma,\ B_\Sigma,\ C_\Sigma,\ D_\Sigma)$. Let $P \in \mathbb{R}^{n_x \times n_x}$ be non-singular and consider a so-called similarity transformation $\tilde{x}_k = Px_k$. Then Σ admits another state-space form $(\hat{A}_\Sigma,\ \hat{B}_\Sigma,\ \hat{C}_\Sigma,\ \hat{D}_\Sigma)$, where

$$\begin{aligned} \hat{A}_\Sigma &= PA_\Sigma P^{-1}, & \hat{B}_\Sigma &= PB_\Sigma, \\ \hat{C}_\Sigma &= C_\Sigma P^{-1}, & \hat{D}_\Sigma &= D_\Sigma. \end{aligned} \tag{2.25}$$

This representation is said to be (algebraically) equivalent to the original state-space description of Σ and therefore results in the same input-output behaviour.

Proof. See Appendix A.3 □

2.4 Summary

In this chapter, the class of discrete-time Lur'e-type systems that is considered in this project for the purpose of system identification is exploited. The linear block Σ is assumed to admit a minimal state-space representation having a square nonlinearity input- output channel. Furthermore, it is explained how and why steady-state responses of Σ can be computed in the frequency domain and a method to compute the induced $\mathcal{L}_2$-gain of Σ. Moreover, it is assumed that the nonlinear block φ is static, decentralized and satisfying certain cone bounded sector constraints. The parametrisation of Lur'e-type systems is non-unique given the possibility of loop-scaling, loop-transformation and similarity transformation, which also allows to cast any Lur'e-type model into a normalised, symmetric form. For system identification purposes, Lur'e-type models are only accepted when they satisfy certain stability guarantees. This model class restriction is central to the upcoming chapter.

Chapter 3

Convergent Lur'e-type models

In the previous chapter, Lur'e-type models are introduced as the model class together with the most important assumptions and properties. A major challenge in nonlinear system identification, which has not been addressed so far, is to enforce a form of stability on the identified model. In Section 3.1 this challenge is explained and illustrated by a simple example. One can aim to constrain the model class further by guaranteeing a certain stability property for our identified Lur'e-type model. In this work, the global uniform exponential convergence property is enforced, which is discussed for generic nonlinear systems in Section 3.2 first. After that, numerically tractable conditions are provided to assess the global uniform exponential convergence property for models in our model class in Section 3.3. Finally, it is shown that satisfaction of the conditions for GEUC guarantees an upper bound to the gain of the linear block, which is used to develop efficient numerical methods for identification in Chapter 4.

3.1 Stability challenges for generic nonlinear systems

Let us consider a generic discrete-time nonlinear system

$$x(k+1) = f(x(k), w(k), k), \qquad \forall k \in \mathbb{Z}, \tag{3.1}$$

where $x(k) \in \mathbb{R}^{n_x}$ and $w(k) \in \mathbb{R}^{n_w}$ are the system's state and input, respectively. Assume that the nonlinear function $f : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{Z}_{\geq k_0} \to \mathbb{R}^n$ is continuous with respect to the two first arguments for any third one. One denotes the initial condition $x(k_0)$ as x_0 for $k_0 \in \mathbb{Z}$. Moreover, for ease of notation, let us introduce the solution of dynamics (3.1) as $\phi^w : \mathbb{Z}_{>k_0} \times \mathbb{Z} \times \mathbb{R}^{n_x} \to \mathbb{R}^{n_x}$ verifying

$$\phi^w(k+1, k_0, x_0) = f(\phi^w(k, k_0, x_0), w(k), k) \qquad \forall\ (k,\ k_0,\ x_0) \in \mathbb{Z}_{\geq k_0} \times \mathbb{Z} \times \mathbb{R}^n. \tag{3.2}$$

It is well-known that nonlinear systems can exhibit multiple stable solutions being attractive for different sets of initial conditions [27]. Moreover, depending on the applied excitation signal, the response of nonlinear models can become unstable. Consequently, even though the model explains a certain identification data set perfectly well, the model response to a new bounded excitation might be close to the true system response, but might as well be far off or even become unbounded. To interpret this behaviour even better, Example 3.1 illustrates the stability challenges for nonlinear systems.

Example 3.1 (Stability challenges)

Let us consider a scalar, nonlinear dynamical system (3.1) *subject to a constant input* w*, where*

$$f(x(k), w) = \frac{1}{10} x(k) w \left(x(k) + 1\right) \Big(\left(x(k) w\right)^2 - x(k) w - 1 \Big) + x(k). \tag{3.3}$$

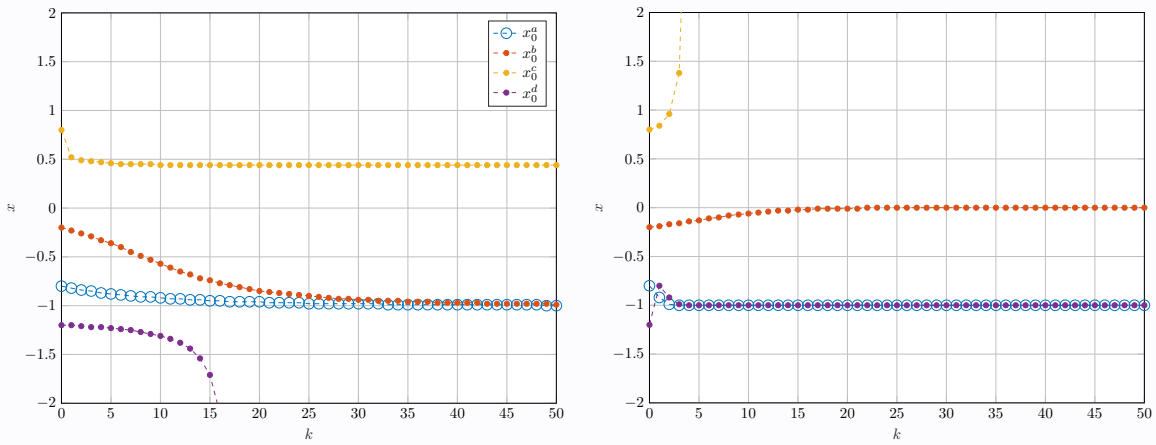
Our intention is to simulate these dynamics in multiple scenarios. To that end, one can first select a set of initial conditions $\left\{x_0^a,\ x_0^b,\ x_0^c,\ x_0^d\right\}$*, where*

$$x_0^a = -0.8, \qquad x_0^b = -0.2, \qquad x_0^c = 0.8, \qquad x_0^d = -1.2.$$

Also, a set of constant inputs $\left\{w^a, w^b\right\}$ *is selected, where*

$$w^a = -1.4, \qquad w^b = 2.1.$$

A simulation of the dynamics (3.3) *for the various initial conditions and inputs is shown in Figure 3.1.*



(a) Input: w^a. (b) Input: w^b.

Figure 3.1: Simulation of dynamics (3.3) for various initial conditions and inputs.

Taking a closer look at these simulation results, one can clearly identify the stability challenges for generic nonlinear systems that were introduced in this section. To start off, observe that the stability of solutions can indeed highly depend on the selected initial state. Looking at Figure 3.1a, it can be seen that a bounded response for $x_0^a,\ x_0^b,\ x_0^c$*, whereas the solution starting from* x_0^d *is unbounded. Undoubtedly, a bounded steady-state solution is not necessarily unique, since the solution for* x_0^c *does not converge in steady-state to the ones starting from* $x_0^a,\ x_0^b$*. Furthermore, taking also Figure 3.1b into account, one notices that stability of solutions is not necessarily preserved under input variations. The* x_0^d *solution stabilizes switching from* w^a *to* w^b*, whereas the* x_0^c *solution shows the exact opposite behaviour. Despite the* x_0^b *solution being bounded, it approaches a different steady-state than* $x = -1$*, namely* $x = 0$*, making the solutions less predictable. This in contrast to the* x_0^a *solution, which is not sensitive to this input variation and still converges to* $x = -1$ *in steady state. One could argue that such a predictable steady-state behaviour can be a desired model property and adds to the robustness of the model to initial conditions and to variations of the input, in particular, if such property is to be expected (or observed) for the system to be modelled.*

3.2 Stability notions for nonlinear systems

This section elaborates upon stability notions for nonlinear systems to deal with the challenges that were explained in Section 3.1. A nonlinear system that exhibits global asymptotic stability (GAS) solutions in absence of perturbations, can respond according to the four scenarios that are shown in Figure 3.2 in presence of bounded perturbations.

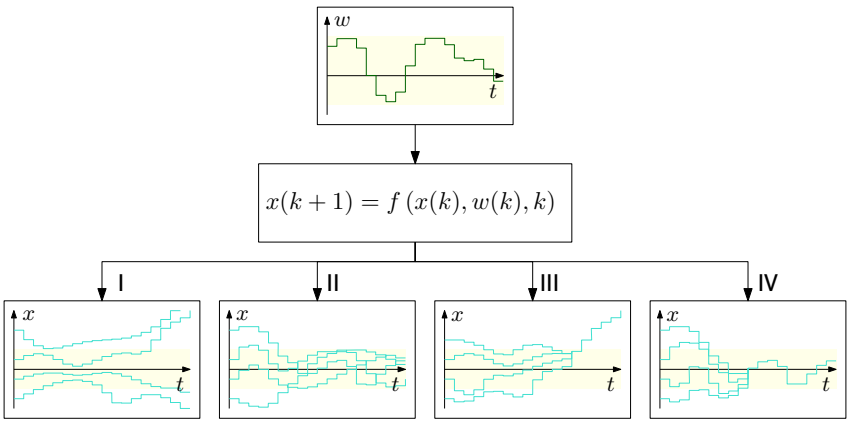


Figure 3.2: Stability notions on nonlinear systems in presence of perturbations.

In scenario I, solutions grow unbounded, despite the stability property of the unperturbed system. This kind of behaviour was observed in Example 3.1 before. In scenario II, the system exhibits the input-to-state stability (ISS) property, such that responses remain in some region around an equilibrium, but do not converge to each other and are, therefore, not predictable [71]. Obviously, in scenario II, one would be interested in quantifying how fast the response converges to the bounded region, together with the size of this convergence region. In scenario III, all responses converge to a unique, but possibly unbounded, steady-state response independent from the initial condition. This behaviour is referred to as *incremental stability*. Finally, in scenario IV, systems with the *convergence* property show bounded responses that converge to a unique steady-state solution independent from the initial condition. Due to the predictability and boundedness of the solution, scenario IV is enforced on the identified model by the identification algorithm in Chapter 4. The uniform convergence property is formally defined as follows.

Definition 3.1 (Uniform convergence [72, Def. 1])
A discrete-time nonlinear system (3.1) *is said to be* uniformly convergent *in a set* $\mathcal{X} \subset \mathbb{R}^{n_x}$ *for an input class* $\mathcal{W}$*, if, for every input* $w \in \mathcal{W}$*:*

- *There exists a unique* steady-state solution $\bar{x}^w(k) \in \mathcal{X}$*, that is defined and bounded for all* $k \in \mathbb{Z}$*.*
- $\bar{x}^w$ *is uniformly asymptotically stable in* $\mathcal{X}$*, i.e., there exists a* $\mathcal{KL}$*-function* β *such that*

$$\|\phi^w(k,\, k_0,\, x_0) - \bar{x}^w(k)\| \leq \beta(\|x_0 - \bar{x}^w(0)\|,\, k - k_0), \qquad \forall x_0 \in \mathcal{X}. \tag{3.4}$$

In the case where β *is exponential in the second argument of* (3.1)*, the convergence is called* exponential*. The set* $\mathcal{X}$ *is referred to as the convergence region. If* $\mathcal{X} = \mathbb{R}^n$*, then the convergence is* global*.*

Notice that $\bar{x}^w$ exists, and is bounded, on the entire time axis, so both in forward and backward time. An instrumental property of uniformly convergent systems addresses steady-state solution periodicity.

Property 3.1 ([73, Pr. 3.1]) *Suppose system* (3.1) *with a given input* w *is uniformly convergent. If the input* w *is constant, the corresponding steady-state solution* $\bar{x}^w$ *is also constant; if* w *is periodic with period* T*, then the corresponding steady-state solution* $\bar{x}^w$ *is also periodic with the same period* T*.*

This property turns out useful for system identification, where the steady-state dynamics of the nonlinear candidate models are calculated explicitly. The definition of uniform convergence does not provide us with a check to verify whether a particular system (3.1) is uniformly convergent, as it is not obvious to find the convergence region and β-function that prove this property according to Definition 3.1. It is desired to characterize the uniform convergence property such that it can be checked more easily. To that end, the notions of (compact) positively invariant sets and uniform asymptotic incremental stability are to be introduced first. In words, once a trajectory of the system (3.1) enters a positively invariant set at a certain time instant, then forward in time it never leaves it again. Such a set is called compact if it is closed and bounded. This notion is introduced formally in Definition 3.2 below.

Definition 3.2 (Positively invariant set [74, p. 4])
A set $\mathcal{S}_c \subset \mathbb{R}^n$ is said to be positively invariant *with respect to (i) the system dynamics* (3.1) *and (ii) inputs from a certain class $\mathcal{W}$, if for all $w \in \mathcal{W}$, $x_0 \in \mathcal{S}_c$, and $k,\ k_0 \in \mathbb{Z}$ such that $k \geq k_0$, it holds $\phi^w(k,\ k_0,\ x_0) \in \mathcal{S}_c$.*

What is left is to formulate a method for checking the existence of such a compact, positively invariant set. This can be done via a Lyapunov characterization, as explained in Lemma 3.1.

Lemma 3.1 (Compact, positively invariant set [75, Ass. 1])
Let us consider the $\mathcal{K}$-function $\gamma(\cdot)$, which characterizes the input class

$$\mathcal{W}_\gamma = \left\{ w \in \mathbb{R}^m,\ \|w\| \leq \gamma^{-1}(1) \right\}. \tag{3.5}$$

Suppose there exists a continuous function $V : \mathbb{R}^n \to \mathbb{R}_{>0}$, together with $\mathcal{K}_\infty$-functions $\alpha_1(\cdot)$, $\alpha_2(\cdot)$ and a nonnegative scalar $c \in \mathbb{R}_{\geq 0}$, such that the following conditions are satisfied for all $x \in \mathbb{R}^n$, $w \in \mathbb{R}^m$ and $k \in \mathbb{Z}$:

$$\alpha_1(\|x\|) \leq V(x) \leq \alpha_2(\|x\|), \tag{3.6a}$$
$$V(f(x, w, k)) \leq c, \text{ if } V(x) \leq c \text{ and } \gamma(\|w\|) \leq 1. \tag{3.6b}$$

Then, there exists a compact, positively invariant set

$$\mathcal{S}_c = \{ x \in \mathbb{R}^n,\ V(x) \leq c \}, \tag{3.7}$$

according to Definition 3.2, with respect to (i) the system dynamics (3.1) *and (ii) inputs from the class $\mathcal{W}_\gamma$.*

This characterisation is a modified version of the autonomous case in [74, pr. Th. 13]. A crucial property of these compact, positively invariant sets is explained by a discrete-time version of Yakubovich's lemma [76, Lem. 2].

Lemma 3.2 ([77, Lem. 2])
A compact positively invariant set $\mathcal{S}_c$ with respect to (i) the system dynamics (3.1) *and (ii) inputs from a certain class $\mathcal{W}$, contains at least one solution $\bar{x}^w$ that is defined on $\mathbb{Z}$ and satisfies $\bar{x}^w \in \mathcal{S}_c$ for all $w \in \mathcal{W}$ and $k \in \mathbb{Z}$.*

The next step is to formalise Definition 3.3, in which the uniform asymptotic incremental stability property is addressed. In words, this property guarantees that any two solutions of a nonlinear system having their initial conditions inside a certain set, asymptotically converge to each other regardless of their exact initial conditions.

Definition 3.3 (Uniform asymptotic incremental stability [75, Def. 2])
A discrete-time nonlinear system (3.1) *is said to be* uniformly asymptotically incrementally stable *in a positively invariant set $\mathcal{S} \subseteq \mathbb{R}^n$ for a class of inputs $\mathcal{W}$, if for any input $w \in \mathcal{W}$ there exists a $\mathcal{KL}$-function $\beta(\cdot,\cdot)$, such that*

$$\left\|\phi^w(k, k_0, x_0^a) - \phi^w(k, k_0, x_0^b)\right\| \leq \beta(\left\|x_0^a - x_0^b\right\|, k - k_0). \tag{3.8}$$

In case β is exponential in the second argument, the incremental stability is called exponential. *Finally, when $\mathcal{S} = \mathbb{R}^n$, then the incremental stability is global.*

Note that this definition is a slight modification of the definition for autonomous systems in [78, Def. 1]. Given the generic nonlinear dynamics (3.1) and inputs from a certain class $w \in \mathcal{W}$, a method for checking the uniform asymptotic incremental stability property is proposed via a Lyapunov characterization in Lemma 3.3.

Lemma 3.3 (Uniform asymptotic incremental stability [75, Ass. 2])
Consider a positively invariant set $\mathcal{S} \subseteq \mathbb{R}^n$ with respect to (i) the dynamics (3.1) *and (ii) inputs from a certain class $\mathcal{W}$, according to Definition 3.2. Suppose there exists a continuous function $V : \mathbb{Z} \times \mathcal{S} \times \mathcal{S} \to \mathbb{R}_{>0}$, referred to as an* incremental Lyapunov function, *together with $\mathcal{K}_\infty$-functions $\alpha_3(\cdot)$, $\alpha_4(\cdot)$ and $\mathcal{K}$-function $\alpha_5(\cdot)$ such that for all $x^a,\ x^b \in \mathcal{S}$, $w \in \mathcal{W}$ and $k \in \mathbb{Z}$, the following conditions are satisfied:*

$$\alpha_3(\left\|x^a - x^b\right\|) \leq V(k, x^a, x^b) \leq \alpha_4(\left\|x^a - x^b\right\|), \tag{3.9a}$$

$$sV\left(k+1,\ f(x^a, w, k),\ f(x^b, w, k)\right) - V(k,\ x^a,\ x^b) \leq -\alpha_5(\left\|x^a - x^b\right\|), \tag{3.9b}$$

Then the system is uniformly asymptotically incrementally stable in $\mathcal{S} \subseteq \mathbb{R}^n$ according to Definition 3.3. The incremental stability is global when $\mathcal{S} = \mathbb{R}^n$. Moreover, when $\alpha_3,\ \alpha_4$ and α_5 are quadratic functions, then system (3.1) *is exponentially incrementally stable.*

Again, this lemma is an extension of the autonomous dynamics case in [74, Th. 9]. Uniform convergence can be guaranteed by the existence of a compact positively invariant set together with the uniform asymptotic incremental stability property. Figure 3.3 helps to illustrate this behaviour via a two-dimensional state-space example.

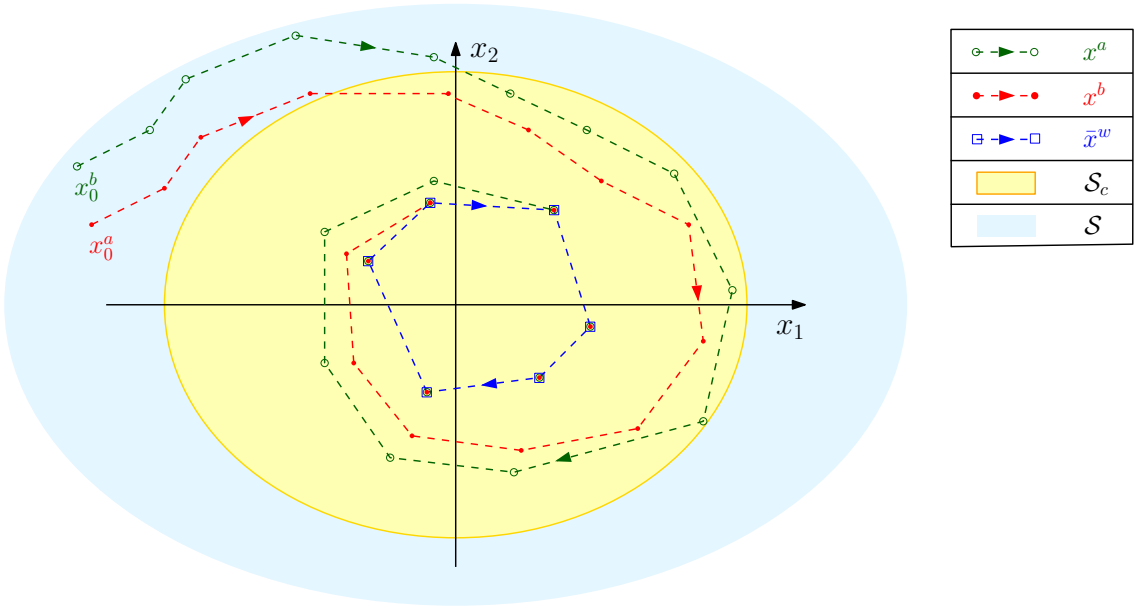


Figure 3.3: Uniform convergence via incremental stability and positively invariant sets.

Suppose there exists a compact, positively invariant set $\mathcal{S}_c$. By Lemma 3.2, it can be concluded that there

exists at least one bounded steady-state solution $\bar{x}^w \in \mathcal{S}_c$. Moreover, suppose there exists another (not necessarily compact) positively invariant set $\mathcal{S} \supseteq \mathcal{S}_c$, which is called the convergence region in which solutions are asymptotically incrementally stable according to Definition 3.3. In that case, any two solutions of dynamics (3.1) starting inside $\mathcal{S}$ must converge to each other. Given the existence of one solution that resides inside $\mathcal{S}_c$ for all time instances, namely the steady-state solution $\bar{x}^w$, all solutions having their initial conditions inside $\mathcal{S}$, converge for sure to this bounded solution. As an example, this behaviour is shown for solutions x^a and x^b in Figure 3.3, that converge to the steady-state solution $\bar{x}^w$. One observes that the two properties combined indeed imply the asymptotic stability and uniqueness of steady-state solution $\bar{x}^w$. Thereby, it guarantees uniform convergence in $\mathcal{S}$. Note that a compact, positively invariant set is required to exclude systems exhibiting incrementally stable solutions that grow unbounded from the class of uniformly convergent systems.

These insights are formalized in Theorem 3.4 below, which formulates a characterisation of uniform convergence via two adequate Lyapunov-like functions. On the one hand, a compact, positively invariant set is introduced as the level set of Lyapunov function V_1. On the other hand, incremental stability is characterized by the time-independent incremental Lyapunov function V_2.

Theorem 3.4 (Lyapunov characterisation of uniform convergence [75, Th. 5])
Consider a generic nonlinear discrete-time system (3.1)*. Suppose that there exists:*

- *A function $V_1 : \mathbb{R}^n \to \mathbb{R}_{>0}$, verifying the conditions in Lemma 3.1. Then, there exists a compact, positively invariant set*
$$\mathcal{S}_c = \{x \in \mathbb{R}^n,\, V_1(x) \leq c\}\,, \tag{3.10}$$
with respect to the system dynamics (3.1) *and inputs from a certain class $w \in \mathcal{W}_\gamma$, which satisfies the properties in Definition 3.2.*
- *A not-necessarily compact set $\mathcal{S} \subseteq \mathbb{R}^n$, which is also positively invariant with respect to the system dynamics* (3.1) *and inputs from the class $\mathcal{W}_\gamma$. Also, this set is a superset of $\mathcal{S}_c$ in the large sense, such that $\mathcal{S} \supseteq \mathcal{S}_c$.*
- *An incremental Lyapunov function $V_2 : \mathbb{Z} \times \mathcal{S} \times \mathcal{S} \to \mathbb{R}_{>0}$, verifying the conditions in Lemma 3.3. Then, the system* (3.1) *is (exponentially) uniformly asymptotically incrementally stable on the set $\mathcal{S}$.*

Then, the system (3.1) *is (exponentially) uniformly convergent on the set $\mathcal{S}$ for the class of inputs $\mathcal{W}_\gamma$ according to Definition 3.1. As a result, the system has a unique steady-state solution $\bar{x}^w$ for every $w \in \mathcal{W}_\gamma$ that verifies*
$$V_1(\bar{x}^w) \leq c. \tag{3.11}$$
Note that the system is said to be globally convergent when $S = \mathbb{R}^n$.

Proof. See [75, Th. 5]. □

This characterisation allows to search for two Lyapunov functions to conclude upon uniform convergence of system (3.1) and is, therefore, instrumental in the derivation of verifiable conditions to assess the uniform convergence property of nonlinear dynamical systems. In many existing results, the functions V_1 and V_2 are taken the same (and often of quadratic form). Theorem 3.4 shows that it is not necessary to choose them equally, which potentially reduces the conservatism of stability conditions. Nevertheless, it should be noted that finding the Lyapunov functions $V_1,\, V_2$ for generic nonlinear systems is still not an easy task. However, this project considers the class of Lur'e-type systems for which numerically tractable conditions can be derived for finding the Lyapunov functions that are required in Theorem 3.4. This topic is discussed in the next upcoming section.

3.3 Sufficient conditions for Lur'e-type systems

For the class of Lur'e-type systems, it has been shown that the assumptions of Theorem 3.4 can be verified by quadratic Lyapunov functions that can be found by solving a set of bilinear matrix inequality (BMI) conditions [75]. Note that all BMI conditions considered in this project reduce to LMI conditions once their respective scalar hyper-parameters are fixed. These LMI conditions can be cast into convex optimisation problems, for which there exist computationally efficient solution methods [79]. This methodology allows for structured approaches in analysing the stability of nonlinear systems, rather than searching for adequate Lyapunov functions. In this project, we restrict ourselves to the identification of discrete-time Lur'e-type systems (2.1), and, therefore, these BMI type of conditions are of interest to check numerically for the GEUC property.

Let us consider the class of Lyapunov functions that are quadratic in the system state:

$$\tilde{V}(x(k)) = \|x(k)\|^2_{\tilde{P}_a}, \tag{3.12}$$

where $0_n \prec \tilde{P}_a \in \mathbb{S}^n$. In [75], sufficient conditions for the GEUC property of discrete-time Lur'e-type systems (2.1) (with the additional assumption $D = 0_{n_y}$) are proposed via Lyapunov functions that are quadratic in the *system state*. This result is formalized in Theorem 3.5.

Theorem 3.5 (Global, exponential uniform convergence: I [75, Th. 7])
Consider a Lur'e-type system $(\Sigma,\ \varphi)$ according to Equation (2.1) and assume $D = 0_{n_y}$. Let Σ and φ verify Assumptions 2.1 and 2.2 respectively. Furthermore, consider symmetric matrices $P_1,\ P_2 \in \mathbb{S}^{n_y}$ and diagonal positive definite matrices $S_1,\ S_2 \in \mathbb{S}^{n_y}$. In addition, consider positive scalars $\tau_1,\ \tau_2,\ \sigma$ and c such that the following inequalities hold true:

$$\begin{bmatrix} A^\top \\ B^\top \\ L^\top \end{bmatrix} P_1 \begin{bmatrix} A^\top \\ B^\top \\ L^\top \end{bmatrix}^\top - \begin{bmatrix} \tau_1 P_1 & -C^\top \Omega S_1 & 0_{n_x \times n_w} \\ \star & 2S_1 & 0_{n_u \times n_w} \\ \star & \star & \tau_2 I_{n_w} \end{bmatrix} \prec 0_{n_x+n_u+n_w}, \tag{3.13a}$$

$$-c(1-\tau_1) + \tau_2 \sigma \leq 0, \tag{3.13b}$$

$$\begin{bmatrix} A^\top \\ B^\top \end{bmatrix} P_2 \begin{bmatrix} A^\top \\ B^\top \end{bmatrix}^\top - \begin{bmatrix} P_2 & -C^\top \bar{\Omega} S_2 \\ \star & 2S_2 \end{bmatrix} \prec 0_{n_x+n_u}. \tag{3.13c}$$

Then, the assumptions of Theorem 3.4 are verified by considering the (incremental) Lyapunov functions $V_1(x) = \|x\|^2_{P_1}$ and $V_2(x^a,\ x^b) = \|x^a - x^b\|^2_{P_2}$. According to Theorem 3.4, the Lur'e-type system is GEUC for the class of inputs $\mathcal{W}_\gamma = \{w(k) \in \mathbb{R}^m,\ \|w(k)\| \leq \sqrt{\sigma}\}$. Finally, the steady-state solution $\bar{x}^w$ belongs to the sellipsoid $\mathcal{E}(P_1,\ c)$.

Proof. An extended version of [75, pr. Th. 7] can be found in Appendix A.4. □

These conditions were taken as a starting point in formulating stability guarantees for Lur'e-type systems (2.1) in the scope of system identification. Obviously, due to a possible non-zero D matrix, our model class includes, but is not limited by, the dynamical systems considered in Theorem 3.5. Therefore, first and most importantly, it is shown that the assumption $D = 0_{n_y}$ can be dropped without any additional conservatism such that the feasibility of the conditions in Theorem 3.5 implies the convergence property for our model class (2.1), also for a non-zero feedthrough matrix D. Note that an additional system output (Equation (2.1c)) does not alter stability properties and is, therefore, not considered for stability analysis.

Lemma 3.4 *The feasibility of the conditions in Theorem 3.5 implies the global, exponential uniform convergence property as per Definition 3.1, for the class of arbitrary bounded inputs, of the dynamics $(\Sigma,\ \varphi)$ according to (2.1), where Σ and φ verify Assumptions 2.1 and 2.2, respectively.*

Proof. See Appendix A.5 □

As an alternative approach, one could alter the conditions in Theorem 3.5 to make the D matrix appear explicitly in the BMI constraints. One could then follow a proof similar to that of Theorem 3.5 to show that these updated constraints imply the convergence property for models inside our model class (2.1). This approach is presented as Theorem B.1 in Appendix B.1.

However, it can be shown that the conditions in Theorem 3.5 can be simplified by eliminating hyper parameters, which reduces the computational cost of stability checks in a system identification context. Therefore, this solution route is pursued and we first present the following argument:

Lemma 3.5 ([75]) *The feasibility of the conditions in Theorem 3.5 is irrespective of τ_2.*

Proof. See Appendix A.6. □

We can benefit from this argument by proposing an improved set of stability conditions in Theorem 3.6 below, from which the hyper-parameter τ_2 has been eliminated. This makes these new conditions preferred over the ones in Theorem 3.5 as we have reduced the number of optimisation variables.

Theorem 3.6 (Global, exponential uniform convergence: II [75])
Consider a Lur'e-type system $(\Sigma,\ \varphi)$ according to Equation (2.1). Let Σ and φ verify Assumptions 2.1 and 2.2, respectively. Furthermore, consider symmetric matrices $P_1,\ P_2 \in \mathbb{S}^{n_x}$ and diagonal positive definite matrices $S_1,\ S_2 \in \mathbb{S}^{n_y}$. In addition, consider the positive scalar $0 < \tau < 1$, such that

$$\begin{bmatrix} A^\top \\ B^\top \end{bmatrix} P_1 \begin{bmatrix} A^\top \\ B^\top \end{bmatrix}^\top - \begin{bmatrix} \tau P_1 & -C^\top S_1 \Omega \\ \star & 2S_1 \end{bmatrix} \prec 0, \tag{3.14a}$$

$$\begin{bmatrix} A^\top \\ B^\top \end{bmatrix} P_2 \begin{bmatrix} A^\top \\ B^\top \end{bmatrix}^\top - \begin{bmatrix} P_2 & -C^\top S_2 \bar{\Omega} \\ \star & 2S_2 \end{bmatrix} \prec 0. \tag{3.14b}$$

Then, via Lemmas 3.4 and 3.5, the assumptions in Theorem 3.5 are satisfied, such that the Lur'e system (2.1) is GEUC as per Definition 3.1 for the class of arbitrary bounded inputs by considering the (incremental) Lyapunov functions $V_1(x) = \|x\|_{P_1}$ and $V_2(x^a,\ x^b) = \left\|x^a - x^b\right\|_{P_2}$.

Proof. See Appendix A.7 □

Minimising τ (while ensuring feasible stability conditions) imposes a tighter bound on steady-state solution $\bar{x}^w$. A *bisection method* can be applied to find such a 'small' τ value as explained in Algorithm 1. Depending on the feasibility of the LMI conditions in a certain iteration m, one either increases the bisection lower-limit, or decreases the bisection upper-limit until the bisection interval becomes sufficiently tight.

Algorithm 1: Bisection method.

Initialise bisection interval $\{\tau_{\min}^{[1]},\ \tau_{\max}^{[1]} = \{0, 1\}$, stopping tolerance $\eta^* > 0$ and index $m = 1$.
Verify that the lower/upper limit indeed yield infeasible/feasible conditions in Theorem 3.6
while $(\tau_{\max}^{[m]} - \tau_{\min}^{[m]})/\tau_{\max}^{[m]} \geq \eta^*$ **do**
  Choose $\tau^{[m+1]} = \text{mean}\left(\tau_{\max}^{[m]},\ \tau_{\min}^{[m]}\right)$
  if *the conditions in Theorem 3.6 are feasible for* $\tau^{[m+1]}$ **then**
    Decrease upper-limit: $\{\tau_{\min}^{[m+1]},\ \tau_{\max}^{[m+1]}\} = \{\tau_{\min}^{[m]},\ \tau^{[m+1]}\}$
  else
    Increase lower-limit: $\{\tau_{\min}^{[m+1]},\ \tau_{\max}^{[m+1]}\} = \{\tau^{[m+1]},\ \tau_{\max}^{[m]}\}$
  end
  $m = m + 1$
end

For illustration purposes of the presented stability conditions together with the bisection method solution strategy, a simple instant of model class (2.1) is worked out, to which the conditions in Theorem 3.5 are applied to verify the GEUC property.

Example 3.2 (A dynamical system that satisfies the conditions in Theorem 3.6)
Let us consider a model $(\Sigma,\ \varphi)$ *according to Equation* (2.1)*, where* Σ *and* φ *verify Assumptions 2.1 and 2.2 respectively. The system's linear block* Σ *does not include a separate output equation and is parametrised by*

$$A = \begin{bmatrix} 0.5 & 0.1 \\ 0.3 & -0.4 \end{bmatrix}, \quad B = \begin{bmatrix} -0.3 & 0.6 \\ 0.6 & -0.3 \end{bmatrix}, \quad L = \begin{bmatrix} 0.2 \\ 0.4 \end{bmatrix}, \quad C = \begin{bmatrix} 0.5 & -0.5 \\ 0.6 & 0.2 \end{bmatrix}, \quad D = \begin{bmatrix} 0.2 \\ 0.1 \end{bmatrix}.$$

Besides that, the nonlinearity is defined as $\varphi = [\ \varphi_1(y_1)\ \ \varphi_2(y_2)\ ]^\top$*, where*

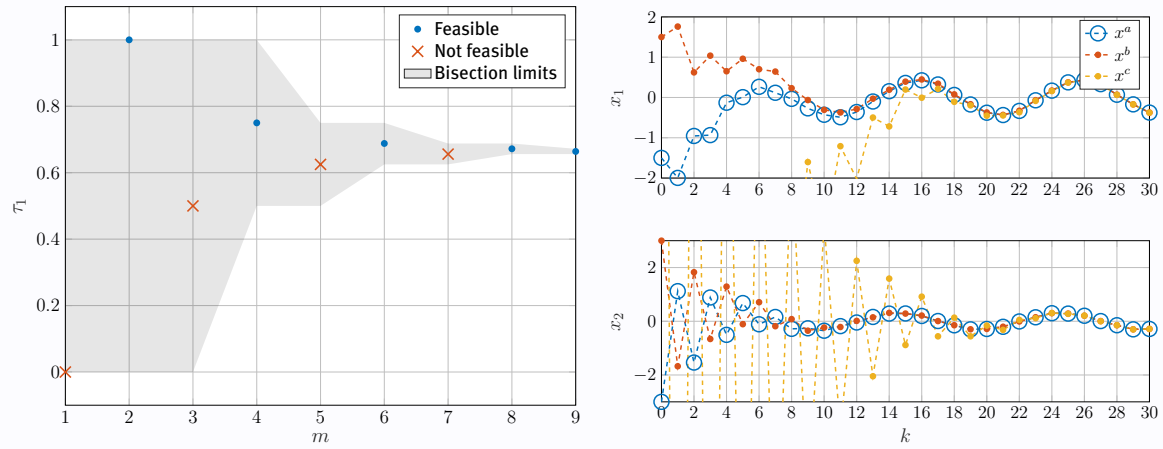
$$\varphi_i(y_i) = \bar{\Omega}_i \operatorname{sign}(y_i) \max(0, |y_i| - \delta). \tag{3.15}$$

The parameters $\bar{\Omega}_i = 0.8$ *and* $\delta_i = 0.1$ *for* $i \in \{1, \ldots, 2\}$ *are selected. One refers to Equation* (3.15) *as a deadzone-type nonlinearity, in which* δ_i *denotes the deadzone length and* $\bar{\Omega}_i$ *defines the nonlinearity slope outside of the deadzone region. The cone bounded sector constraints on* φ *are then characterized by* $\Omega = \bar{\Omega} = \operatorname{diag}\{\bar{\Omega}_1, \bar{\Omega}_1\}$*. Finally, the system is subject to the external excitation* $w(k) = \sin(\pi(k-1)/5)$*.*

Our intention is to verify the GEUC property for the example system via the conditions in Theorem 3.6. A bisection search on τ *has been performed, starting on the interval* $[0,\ 1]$ *for the stopping criterium* $\eta^* = 0.01$*. The iteration history of this search can be seen in Figure 3.4a and resulted in* $\tau = 0.66$ *being the smallest scalar for which a feasible solution of Theorem 3.6 exists. The solution of the corresponding LMI then reads*

$$P_1 = \begin{bmatrix} 1.41 & 0.14 \\ 0.14 & 0.67 \end{bmatrix}, \quad P_2 = \begin{bmatrix} 1.17 & 0.08 \\ 0.08 & 0.89 \end{bmatrix}, \quad S_1 = \begin{bmatrix} 0.74 & 0 \\ 0 & 1.24 \end{bmatrix}, \quad S_2 = \begin{bmatrix} 0.84 & 0 \\ 0 & 0.92 \end{bmatrix}.$$

Let us iterate the dynamics $(\Sigma,\ \varphi)$ *starting from* $x_0^a = [-1.5\ \ -3]^\top$ *or* $x_0^b = [1.5\ \ 3]^\top$*. We observe in Figure 3.4b convergence of solutions to a unique and bounded steady-state* $\bar{x}^w$*, which is periodic with the same period as our external excitation* w*. Even starting from an extreme initial condition* $x_0^c = [-50\ \ 50]^\top$ *yields similar conclusions, whereby the guaranteed stability property indeed makes a global impression.*

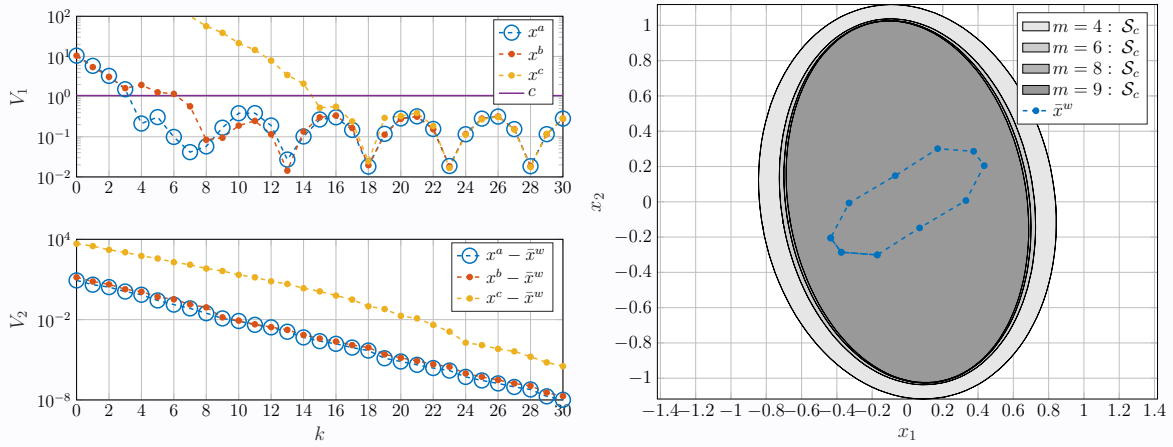


(a) Bisection search on hyper-parameter τ.

(b) State trajectories.

Figure 3.4: Bisection search and state trajectories.

Given our state trajectories, we can also evaluate the trajectories of Lyapunov functions V_1 and V_2 and show them in Figure 3.5a. We can see the magnitudes decreasing in forward time, but the presence of excitation w prevents the Lyapunov functions from reaching zero. We can define a level set $\mathcal{S}_c := \{x \in \mathbb{R}^{n_x} \mid V_1(x) \leq c\}$ for $c = 1.06$. This level was found by choosing $\lambda = 0.1$ and using Equations (A.26) *and* (A.27) *with a ten percent margin with respect to the critical values of c and σ as was shown in the proof of Theorem 3.6 in Appendix A.7. We can clearly see in Figure 3.5a that once V_1 becomes smaller than c, it never exceeds this value again, which makes $\mathcal{S}_c$ positively invariant with respect to the dynamics (Σ, φ) and arbitrary bounded inputs.*



(a) Lyapunov functions trajectories.

(b) Positively invariant set.

Figure 3.5: Lyapunov functions and positively invariant sets.

The Lyapunov function level set $\mathcal{S}_c$ is shown in Figure 3.5b as the sellipsoid $\mathcal{E}(P_1,\ c)$ in state-space (x_1, x_2) for all bisection method iterations that yield feasible stability conditions. It can be seen that $\mathcal{S}_c$ includes $\bar{x}^w$. Also, a smaller value of τ provides a tighter bound on $\bar{x}^w$.

3.4 Efficient implementation of stability conditions for identification

The stability conditions in Theorem 3.5 allow for insights in a model's behaviour, e.g., an explicit characterization of $\mathcal{S}_c$, as was shown in Example 3.2. However, in this project we are only interested in verifying the GEUC property for a model inside our model class. We do not aim to find a bound on the steady-state solution while assessing stability properties, since we calculate the steady-state solution explicitly for system identification purposes; see Chapter 4. We can therefore propose one additional simplification of the stability conditions for the purpose of computational efficiency.

Theorem 3.7 (Global, exponential uniform convergence: III)
Consider a Lur'e-type system $(\Sigma,\ \varphi)$ according to Equation (2.1)*. Let Σ and φ verify Assumptions 2.1 and 2.2 respectively. Furthermore, consider a symmetric matrix $P \in \mathbb{S}^{n_x}$ and diagonal positive definite matrix $S \in \mathbb{S}^{n_y}$. In addition, the scalar $\tau^* = 0.99$ is selected, which obviously satisfies $0 < \tau^* < 1$, such that the following inequalities hold true:*

$$\begin{bmatrix} A^\top \\ B^\top \end{bmatrix} P \begin{bmatrix} A^\top \\ B^\top \end{bmatrix}^\top - \begin{bmatrix} \tau^* P & -C^\top S\bar{\Omega} \\ \star & 2S \end{bmatrix} \prec 0, \tag{3.16}$$

Then, the assumptions of Theorem 3.4 are verified by considering the (incremental) Lyapunov functions $V_1(x) = \|x\|_P$ and $V_2(x^a,\ x^b) = \left\|x^a - x^b\right\|_P$. According toTheorem 3.4, the Lur'e-type system (2.1) *is GEUC for the class of arbitrary bounded inputs as per Definition 3.1.*

Proof. See Appendix A.8. □

This theorem allows us to just search for a solution to Condition (3.16), which is an LMI, to characterise the Lyapunov functions V_1 and V_2. Because of this, it is no longer required to execute the bisection search and one can just solve a convex optimisation problem to verify the stability properties of a candidate model during system identification. This step imposes huge computational benefits and is, therefore, implemented for system identification. Notice that $\tau^* = 0.99$ is selected to show that our steady-state solution $\bar{x}^w$ lives in a relatively large, but finite positively invariant set. Via other stability conditions, e.g. Theorem 3.6, positively invariant sets that enclose the steady-state solution within a tighter margin can likely be selected. However, for the purpose of system identification, one is not necessarily interested in such a tight bound. This concludes our discussion on sufficient conditions for the GEUC property of Lur'e-type systems.

3.5 Bounded LTI block gain for convergent Lur'e-type systems

The main objective of this section is to show that the feasibility of the conditions in Theorem 3.6 implies the feasibility of the discrete-time Bounded-Real-Lemma condition applied to the LTI dynamics from $\tilde{u}$ to $\tilde{y}$ of the symmetric, normalised Lur'e-type model form $(\tilde{\Sigma},\ \tilde{\varphi})$ for $\gamma = 2$ in Lemma 2.1. This, in turn, guarantees an upper-bound on the $\mathcal{H}_\infty$ norm of these LTI dynamics, as is formalised in Lemma 3.6.

Lemma 3.6 ($\mathcal{H}_\infty$ norm of $\Sigma_{\tilde{u}y}$ implied by the convergence property)
Consider a Lur'e-type system $(\Sigma,\ \varphi)$ according to Equation (2.1). Let Σ and φ verify Assumptions 2.1 and 2.2 respectively. Suppose that $(\Sigma,\ \varphi)$ satisfies the conditions in Theorem 3.6 and is, therefore, GEUC as per Definition 3.1. Furthermore, suppose that the system has been transformed into its symmetric, normalised form $(\tilde{\Sigma},\ \tilde{\varphi})$ as per Lemma 2.4. Then, by applying the discrete-time Bounded-Real-Lemma (Lemma 2.1) to $\tilde{\Sigma}_{\tilde{y}\tilde{u}}(\varkappa)$, one can conclude that the following inequality holds true:

$$\left\| \mathcal{G}_{\tilde{\Sigma}_{\tilde{y}\tilde{u}}}(\varkappa) \right\|_{\mathcal{H}_\infty} < 2. \tag{3.17}$$

Proof. See Appendix A.9 □

This result is used extensively in a contraction analysis performed in Section 4.4.

3.6 Summary

This chapter started by explaining the challenges of stability analysis for nonlinear dynamical systems. By means of an example, it is motivated that, depending on system perturbations, these dynamics can exhibit multiple equilibria that are attractive for different sets of initial conditions. Steady-state solutions of nonlinear systems can either be bounded or unbounded and unique or non-unique. The stability notion of interest to this project is referred to as GEUC and guarantees the existence of a both bounded and unique steady-state solution that is globally exponentially stable. A Lyapunov characterisation of this stability notion is presented, which defines a system as convergent, if there exists a compact, positively invariant set in state-space, which is a subset of the domain on which the system exhibits the uniform incremental stability property. Sufficient conditions for GEUC of discrete-time Lur'e-type models (2.1) in terms of LMI conditions are presented and extended towards an efficient implementation of the stability checks as part of a system identification algorithm, which is to be discussed in the next chapter. Finally, it is shown that the feasibility of the sufficient conditions for GEUC imposes an upper-bound to the $\mathcal{H}_\infty$-norm on the linear block of the Lur'e-type system in its symmetric, normalised form. This property turns out instrumental for the implementation of a system identification algorithm as well, whose working principles and properties are discussed in the upcoming chapter.

Chapter 4

Identification of Lur'e-type systems

Via system identification techniques, we aim to find a descriptive model to a sequence of collected inputs and outputs from a certain dynamical system under study. This is usually done by solving a certain optimisation problem. Therefore, let us consider the standard form of a constrained optimisation problem, which reads

$$\underset{\theta \in \Theta}{\text{minimize}} \quad J(\theta). \tag{4.1}$$

In this description, one recognizes the decision variables θ that are constrained in a certain set $\Theta \subset \mathbb{R}^{n_\theta}$. It is desired to find the decision variables that minimise the objective function $J : \Theta \to \mathbb{R}$. In system identification problems, the decision variables parametrise a candidate model. Furthermore, Θ characterises a certain model class, containing all possible candidate models. More details on the specific optimisation problem at hand in the scope of system identification (explaining the definition of θ, Θ and J) can be found in Section 4.1. Once the system identification task has been formulated as an optimisation problem, we still do not know how to find the decision variables that solves the problem. Section 4.2 explains that we use a gradient-based method to solve the problem at hand starting from a well-chosen initial point. This solution method requires us to calculate the objective function gradient with respect to the decision variables, for which a parameter sensitivity model approach is presented in Section 4.3. An implementation of this solution method requires us to calculate many steady-state solutions for the candidate models and their respective (discrete-time variant of) parameter sensitivity models. We motivate the computationally efficient implementation of the system identification algorithm in Section 4.4, where the so-called MTF algorithm plays a central role.

4.1 Formulation of the system identification problem

Candidate models

In the scope of system identification, the n_θ decision variables in θ fully parametrise a globally exponentially uniformly convergent Lur'e-type candidate model $(\Sigma(\theta),\ \varphi(\theta))$. Such a candidate model is defined as

$$\Sigma(\theta) : \begin{cases} x(k+1,\theta) = A(\theta)x(k,\theta) + B(\theta)u(k,\theta) + L(\theta)w(k), \quad x(0,\theta) = x_0, & \text{(4.2a)} \\ y(k,\theta) = C(\theta)x(k,\theta) + D(\theta)w(k), & \text{(4.2b)} \\ z(k,\theta) = F(\theta)x(k,\theta) + G(\theta)u(k,\theta) + H(\theta)w(k), & \text{(4.2c)} \end{cases}$$

$$\varphi(\theta) : \{ \quad u(k,\theta) = \varphi(y(k,\theta),\theta). \tag{4.2d}$$

Note that the signals associated with the candidate model are written by calligraphic font. The relation between θ and the building blocks of Equation (4.2) depends on the specific problem at hand. One can adopt a black-box modelling approach, in which all entries of the linear block matrices and all nonlinear block parameters represent a decision variable. Contrarily, a white-box modelling routine can impose additional model structure as a result of first-principles modelling. The decision variable space $\Theta \subset \mathbb{R}^{n_\theta}$ is given by

$$\Theta = \Big\{ \theta \in \mathbb{R}^{n_\theta} \mid \exists (0 \prec P \in \mathbb{S}^{n_x}, 0 \prec S \in \mathbb{S}^{n_y}) : \text{Condition (3.16) in Theorem 3.7 holds true} \Big\}. \tag{4.3}$$

By Θ we thus denote the set of all candidate models that are globally exponentially uniformly convergent by verifying the conditions in Theorem 3.7. Because of Property 3.1, one can denote the steady-state solution of the candidate model under excitation by $w \in \ell_2^{n_w}(T)$ as $\big(\bar{x} \in \ell_2^{n_x}(T),\ \bar{u} \in \ell_2^{n_u}(T),\ \bar{y} \in \ell_2^{n_y}(T),\ \bar{z} \in \ell_2^{n_z}(T)\big)$.

True system under study

In this project, we assume the true system to be inside model class (4.2), as is formalised in Assumption 4.1 below.

Assumption 4.1 *There exists a decision variable set $\Theta_0 \subset \Theta$ describing the true system, i.e., the true system can be described by parametrising a model inside the class of globally exponentially uniformly convergent Lur'e-type models that verify Condition* (3.16) *in Theorem 3.7.*

Note that due to the non-uniqueness of the model parametrisation (see Section 2.3) Θ_0 is a subset of the decision variable space Θ rather than a single point in the n_θ-dimensional decision variable space. Because of this assumption, we can write the dynamics of the true system under study for some $\theta_0 \in \Theta_0$ as

$$\Sigma(\theta_0): \begin{cases} x_0(k+1,\theta_0) = A(\theta_0)x_0(k,\theta_0) + B(\theta_0)u_0(k,\theta_0) + L(\theta_0)w(k), \quad x_0(0,\theta_0) = x_0, & \text{(4.4a)} \\ y_0(k,\theta_0) = C(\theta_0)x_0(k,\theta_0) + D(\theta_0)w(k), & \text{(4.4b)} \\ z_0(k,\theta_0) = F(\theta_0)x_0(k,\theta_0) + G(\theta_0)u_0(k,\theta_0) + H(\theta_0)w(k), & \text{(4.4c)} \end{cases}$$

$$\varphi(\theta_0): \{ \quad u_0(k,\theta_0) = \varphi(y_0(k,\theta_0),\theta_0). \quad \text{(4.4d)}$$

Note that the signals associated with the true system under study are written using a zero subscript. Obviously, the decision variables θ_0 are unknown at this point. Because of Property 3.1, we can denote the steady-state solution of the true system as $\left(w \in \ell_2^{n_w}(T),\ \bar{x}_0 \in \ell_2^{n_x}(T),\ \bar{u}_0 \in \ell_2^{n_u}(T),\ \bar{y}_0 \in \ell_2^{n_y}(T),\ \bar{z}_0 \in \ell_2^{n_z}(T)\right)$ while being subject to excitation by a T-periodic input w. We assume that the true system is subject to additive output noise, such that we can define the measured true system output as follows from Assumption 4.2.

Assumption 4.2 *The measured steady-state output is denoted by $\bar{z}_0^e$ and assumed to be a summation of the output of $(\Sigma(\theta_0), \varphi(\theta_0))$ and and additive output noise, .i.e,*

$$\bar{z}_0^e(k,\theta_0) = \bar{z}_0(k,\theta_0) + e(k). \quad \text{(4.5)}$$

Here, $e(k)$ is (i) zero-mean Gaussian white noise with finite variance σ_e^2 and (ii) independent from $w(k)$.

Objective function

The next step is to define the objective function for the system identification problem. Here, one should carefully consider which model is a 'good' model given a set of observed system inputs and outputs. In literature, we can see two main branches of optimisation criteria, which are referred to as (i) one-step ahead prediction error and (ii) simulation error [12]. In the former methodology, we provide the candidate model with all previous system observations $\bar{z}_0^e(k-1,\theta_0)$, $\bar{z}_0^e(k-2,\theta_0)$, etc. and the current input $w(k)$ in order to make the candidate model predict the steady-state output $\bar{z}(k,\theta)$ at the current time instant, which we can then compare to the system observations $\bar{z}_0^e(k,\theta_0)$. This process of one-step-ahead prediction can be repeated along the entire measurement interval to allow for a quantification of the model performance. The latter methodology uses the measured input data only to predict system outputs along the entire measurement interval. This method allows the model to predict the system behaviour for new unseen input sequences at the cost of an increased algorithmic complexity. The nonlinear dynamics propagate along the measurement interval predictions, which might even cause unstable simulation results. This already indicates that having a stability guarantee on the candidate models turns out helpful. Furthermore, one should also take into account that computation of the objective function gradient with respect to the decision variables is more involved, since past outputs also depend (typically nonlinearly) on θ. In general, models learnt by the simulation error criteria perform better on long-term predictions than models that are learnt by the one-step ahead prediction error criteria [12]. Therefore, in this project, we select the simulation error to quantify the cost associated with a certain candidate model. The aforementioned challenges on model instability and objective function gradient computation are dealt with in Sections 4.3 and 4.4.

Let us define the steady-state model output errors with respect to:

1. *System observations* $\bar{z}_0^e$:

$$\bar{\epsilon}^e(k,\theta) = \bar{x}(k,\theta) - \bar{z}_0^e(k,\theta_0), \qquad \forall k \in \{1,\dots,N\}\,. \tag{4.6}$$

2. *Noiseless true system outputs* $\bar{z}_0$:

$$\bar{\epsilon}(k,\theta) = \bar{x}(k,\theta) - \bar{z}_0(k,\theta_0), \qquad \forall k \in \{1,\dots,N\}\,. \tag{4.7}$$

The steady-state output errors are illustrated for the j-th component of the outputs in Figure 4.1 and the error with respect to the *system observations* forms the basis to the objective function, which is commonly referred to as a least-squares criterium and is defined according to

$$J_N(\theta) = \frac{1}{N}\sum_{j=1}^{n_z}\sum_{k=1}^{N} \bar{\epsilon}_j^e(k,\theta)^2. \tag{4.8}$$

Note that the candidate model pays a penalty for any mismatch between its outputs and their respective system observations along the measurement interval. Let us define another mismatch measure, related to the *noiseless true system outputs*, as

$$V_N(\theta) = \frac{1}{N}\sum_{j=1}^{n_z}\sum_{k=1}^{N} \bar{\epsilon}_j(k,\theta)^2, \tag{4.9}$$

on which we impose the following assumption.

Assumption 4.3 *The unique global minimiser of* V_N *is* $\theta_0 \in \Theta_0$.

Note that the true parameter vector θ_0 is an element of a set Θ_0, which originates from the non-unique parametrisation of Lur'e-type models as discussed in Section 2.3. Then, we have discussed all building blocks to define the system identification estimator $\hat{\theta}_N$ as

$$\hat{\theta}_N = \underset{\theta \in \Theta}{\arg\min} \quad J_N(\theta). \tag{4.10}$$

The statistical properties of the estimator in Equation (4.10) together with Assumptions 4.1 to 4.3 have already been researched in the field of system identification [14]. It was concluded that the estimator is *consistent*, which means that $\hat{\theta}_N$ converges in probability to θ_0 for long measurement intervals. Also, the estimator is *asymptotically efficient*, implying that its variance with respect to θ_0 is minimised up to the Cramer-Rao lower-bound, which describes the best possible unbiased estimator [80].

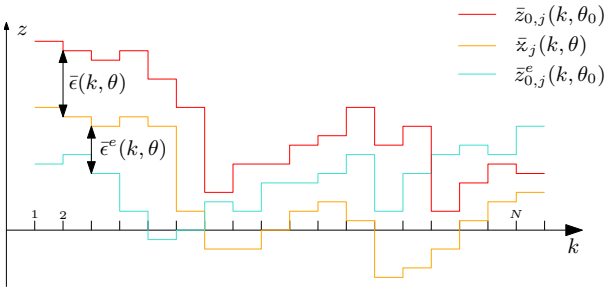


Figure 4.1: Steady-state output error definitions.

The system identification problem at hand is now fully defined as the constrained optimization problem (4.10). Nevertheless, up to this point, it is still unknown what kind of approaches can be taken to find $\hat{\theta}_N$ inside the model class. To that end, the solution strategy for the system identification problem is discussed next.

4.2 Solution to the system identification problem

The problem of nonlinear system identification has been translated into the standard-form constrained optimisation problem in Equation (4.10). There is an entire research area dedicated to solving these standardised problems, which we can benefit from in this project. In the system identification community, it is common practice to solve these problems via a two-step approach: (i) Decision variable initialisation and (ii) Gradient-based optimisation [13], [14], [48].

Step I: Decision variable initialisation

In the decision variable initialisation step, we aim to find a well-chosen initial parameter vector θ_{init} from which we can start the gradient-based search in the second step. The objective for θ_{init} is to be close to the set of true system parameters Θ_0, while being inside the feasible decision variable space Θ of globally exponentially uniformly convergent Lur'e-type models that verify the conditions in Theorem 3.7. In this project, we consider two methods for the purpose of decision variable initialisation, being the BLA and global optimisation, which are treated one after the other in the upcoming paragraphs.

Best Linear Approximation

This method relies on estimating the so-called BLA of the nonlinear system [81]. Here, we approximate the mapping from excitation w to system output z_0^e by the LTI model Σ_{BLA} that admits a state-space representation $(A_{BLA},\, B_{BLA},\, C_{BLA},\, D_{BLA})$. The system matrices in this state-space from can be found via the MATLAB routines N4SID or TFEST, and crucially, these routines can guarantee GAS of Σ_{BLA}, i.e., A_{BLA} is a Schur matrix such that $\lambda(A_{BLA}) \subset$.

Based on this LTI approximation of the nonlinear dynamics, one can initialise a Lur'e-type candidate model, which is denoted by $(\Sigma(\theta_{\text{init}}), \varphi(\theta_{\text{init}}))$. The linear block $\Sigma(\theta_{\text{init}})$ allows for the state-space representation $(A_{\Sigma(\theta_{\text{init}})},\, B_{\Sigma(\theta_{\text{init}})},\, C_{\Sigma(\theta_{\text{init}})},\, D_{\Sigma(\theta_{\text{init}})})$. Obviously, the convergence property is guaranteed for the initial Lur'e-type model if (i) $\mathcal{G}_{\Sigma(\theta_{\text{init}})_{zw}}(\infty) = \mathcal{G}_{\Sigma_{BLA}}(\infty)$ and (ii) nonlinearity output $u = 0$ or $\mathcal{G}_{\Sigma(\theta_{\text{init}})_{yu}}(\infty) = 0$. Recall that the transfer matrices for the linear block of a Lur'e-type model (2.1) are defined in Section 2.1.

Three possible initial candidate model parametrisations are identified that guarantee these two properties. Let us recall that matrices having their elements sampled from a uniform distribution on the $[-1, 1]$ interval are denoted by $\mathcal{U}_{i\times j} \in \mathbb{R}^{i\times j}$. Then, the three methods can be listed as follows:

1. **[Zero output matrices method]:** The nonlinearity is parametrised by random variables and the state-space representation of $\Sigma(\theta_{\text{init}})$ is defined as

$$\begin{aligned} A_{\Sigma(\theta_{\text{init}})} &= A_{BLA}, & B_{\Sigma(\theta_{\text{init}})} &= \begin{bmatrix} \mathcal{U}_{n_x\times n_u} & B_{BLA} \end{bmatrix}, \\ C_{\Sigma(\theta_{\text{init}})} &= \begin{bmatrix} 0_{n_y\times n_x} \\ C_{BLA} \end{bmatrix}, & D_{\Sigma(\theta_{\text{init}})} &= \begin{bmatrix} 0_{n_y} & 0_{n_y\times n_w} \\ \mathcal{U}_{n_z\times n_u} & D_{BLA} \end{bmatrix}, \end{aligned} \tag{4.11}$$

   such that there is zero nonlinearity input, which implies zero nonlinearity output due to the sector condition on the nonlinearity of the initial candidate model (see Definition 2.2).

2. **[Zero input matrices method]:** The nonlinearity is parametrised by random variables and the state-space representation of $\Sigma(\theta_{\text{init}})$ is defined as

$$\begin{aligned} A_{\Sigma(\theta_{\text{init}})} &= A_{BLA}, & B_{\Sigma(\theta_{\text{init}})} &= \begin{bmatrix} 0_{n_x\times n_u} & B_{BLA} \end{bmatrix}, \\ C_{\Sigma(\theta_{\text{init}})} &= \begin{bmatrix} \mathcal{U}_{n_y\times n_x} \\ C_{BLA} \end{bmatrix}, & D_{\Sigma(\theta_{\text{init}})} &= \begin{bmatrix} 0_{n_y} & \mathcal{U}_{n_y\times n_w} \\ 0_{n_z\times n_u} & D_{BLA} \end{bmatrix}, \end{aligned} \tag{4.12}$$

   such that a possibly non-zero nonlinearity output does not propagate through $\Sigma(\theta_{\text{init}})$ into the output.

3. **[Zero nonlinearity output method]:** A special nonlinearity parametrisation that realizes zero nonlinearity output is chosen and the state-space representation of $\Sigma(\theta_{\text{init}})$ is defined as

$$\begin{aligned} A_{\Sigma(\theta_{\text{init}})} &= A_{BLA}, & B_{\Sigma(\theta_{\text{init}})} &= \begin{bmatrix} \mathcal{U}_{n_x \times n_u} & B_{BLA} \end{bmatrix}, \\ C_{\Sigma(\theta_{\text{init}})} &= \begin{bmatrix} \mathcal{U}_{n_y \times n_x} \\ C_{BLA} \end{bmatrix}, & D_{\Sigma(\theta_{\text{init}})} &= \begin{bmatrix} 0_{n_y} & \mathcal{U}_{n_y \times n_w} \\ \mathcal{U}_{n_z \times n_u} & D_{BLA} \end{bmatrix}. \end{aligned} \tag{4.13}$$

   Note that a nonlinearity parametrisation might not be trivially tunable towards zero nonlinearity output. Consequently, this option is not applicable to every system identification problem.

Crucially, the nonlinear Lur'e-type model $(\Sigma(\theta_{\text{init}}), \varphi(\theta_{\text{init}}))$ is identical to the linear model Σ_{BLA} in terms of an equivalent input-output behaviour. The BLA framework is able to estimate initial model parameters fast and, additionally, it provides an analysis on nonlinear distortions and an estimate for the state dimension [81]. Also, this method was applied successfully as an initial point for nonlinear system identification in prior research, see [23] and references therein.

Global Optimisation

One can think of initialising decision variables via global optimisation routines in an attempt to identify an initial point that is sufficiently close to the global minimum of the objective function. The search can be stopped prematurely to save computational costs since a gradient-based method, starting from this initial point, can thereafter be used to effectively find the closest minimum of the cost function. Such global optimisation methods were researched extensively in literature and typically encompass (sophisticated) random walks through the decision variable space. These efforts resulted, among others, in Monte-Carlo, genetic-type and swarm intelligent algorithms [82]. Typically, these routines are considered computationally expensive as they tend to explore a significant share of the unfavourable regions in decision variable space [83]. Therefore, a computationally cheap method to evaluate the objective function value, such as the one proposed in Section 4.4, can be an enabler for the application of global optimisation routines. A controlled random search (CRS) method, as described in [84], [85], was successfully applied as an initialisation method for nonlinear system identification of continuous-time models [14], [86], [87]. The CRS version as implemented in this project is detailed in Algorithm 2.

Algorithm 2: Controlled random search algorithm [85, Alg. 1].

Set $k = 0$, $n_t = 0$ and $n_s = 0$.
Initialise a population R of n_p agents that are uniformly distributed on Θ.
Evaluate the cost associated with all agents in the population
while $(J^{\min} > J^*$ ***and*** $(J^{\max} - J^{\min})/J^{\max} > \eta^*$ ***and*** $k < k^*)$ ***or*** $k = 0$ **do**
    Denote the worst agent in R by $J^{\max} = J(r^{\max}) = \max_{r_i \in R} J(r_i)$
    Select a random subset of $n_m + 1$ agents in R such that $R_s = \{r_1, \ldots, r_{n_m+1}\}$
    Compute the centroid $\bar{G}$ of the first n_m points in R_s via $\bar{G} = \frac{1}{n_m} \sum_{i=1}^{n_m} r_i$
    Compute the primary point $\bar{P}_1$ according to $\bar{P}_1 = 2\bar{G} - r_{m+1}$
    Set $n_t = n_t + 1$
    if $\bar{P}_1 \in \Theta$ ***and*** $J(\bar{P}_1) < J^{\max}$ **then**
        Replace $r^{\max}$ by $\bar{P}_1$ in R `// Replace the worst agent by the primary point`
        Set $n_s = n_s + 1$
    else if $\frac{n_s}{n_t} < \alpha$ **then**
        Compute the secondary point: $\bar{P}_2 = \frac{1}{2}(\bar{G} + r_{n_m+1})$
        if $J(\bar{P}_2) < J^{\max}$ **then**
            Replace $r^{\max}$ by $\bar{P}_2$ in R `// Replace the worst agent by the secondary point`
            Set $n_s = n_s + 1$
        Set $n_t = n_t + 1$
    Best agent in the population: $J^{\min} = J(r^{\min}) = \min_{r_i \in R} J(r_i)$
    Set $k = k + 1$
end
return $\theta_{\text{init}} = r^{\min}$

The algorithm starts by population initialisation in line 2, in which n_p agents are selected that are uniformly distributed over the decision variable space Θ. Based on their objective function value, the agents can be ranked from worst to best in line 3. Consecutively, the population R is iteratively updated until stopping conditions are verified. A single iteration (lines 4-21) of the CRS algorithm can be explained via Figure 4.2 for a two-dimensional example, in which agents and trial points are denoted by dots and squares, respectively.

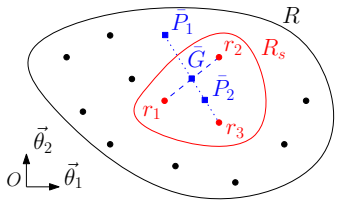


Figure 4.2: Primary and secondary point selection in a CRS iteration (for $n_m = 2$ and $n_p = 12$).

One firstly selects a random subset R_s of $n_m + 1$ agents from the population. Consecutively, the centroid $\bar{G}$ of the first n_m agents in the subset is computed. Thereafter, a so-called primary point $\bar{P}_1$ can be defined by mirroring r_{n_m+1} over the centroid $\bar{G}$. If the primary point lies inside Θ and yields an improvement, it replaces the worst agent in the population and the number of successful trials n_s is increased. Otherwise, if the success rate with respect to the number of trials n_t is insufficient, one also tries to update the population via a so-called secondary point $\bar{P}_2$ in the middle between r_{n_m+1} and $\bar{G}$. These steps iterate until (i) an agent is found with a sufficiently small corresponding objective function value, (ii) the objective function value of all agents has converged or (iii) a maximum number of iterations is performed. Finally, the CRS algorithm returns the best agent in the population as a solution to the decision variable initialisation problem.

Step II: Gradient-based optimisation

The intention is not to improve upon existing nonlinear optimisation techniques, but to use an off-the-shelf available solution method that fits the system identification problem well. Typically, the problems are high-dimensional, nonlinear in the parameters and non-convex. Therefore, finding the decision variables that minimise objective (4.8), even when starting close to the true system parameters, is not a trivial task. We limit ourselves to algorithms that are readily available in the MATLAB Optimisation toolbox to fit within our research context [88]. This toolbox provides us with two algorithms that can be used to solve the system identification problem, being (i) LSQNONLIN and (ii) FMINCON. The latter one has our preference since it allows for definition of additional (in)equality constraint functions next to the simple decision variable box constraints allowed by LSQNONLIN. Such (in)equality constraint functions can among other things be used in an attempt to eliminate non-unique model parametrisations as is explained in Chapter 5. The FMINCON algorithm applies an interior-point method approach, see [89], to solve the optimisation problem. Algorithm 3 describes the family of gradient-based optimization routines [90]. The differences between specific routines exist in, e.g., the computed search direction and step size in lines 5 and 6, respectively. These algorithms typically stop iterating upon either being close to a (local) minimum, indicated by small gradient magnitudes, or having converged to a fixed point of the iterative algorithm by taking small step sizes t.

Algorithm 3: Gradient-based optimisation with stopping criterium $\eta_{[i]}$ [90, p. 466].

1 Given a starting point $\theta_{\text{init}} \in \Theta$
2 Initialise iteration counter $i = 0$ and current iteration variables $\theta_{[0]} = \theta_{\text{init}}$
3 Initialise current stopping criterium $\eta_{[0]} = \infty$ and stopping criterium tolerance $\eta_* > 0$
4 **while** $\eta_{[i]} > \eta^*$ **do**
5 *Compute search direction:* $\Delta\theta$
6 *Compute step size:* $t_{[i]}$
7 *Update current iteration variables:* $\theta_{[i+1]} = \theta_{[i]} + t_{[i]}\Delta\theta$
8 *Update iteration counter:* $i = i + 1$.
9 **end**
10 **return** $\hat{\theta}_N = \theta_{[i]}$

Any gradient-based method requires the computation of the objective function gradient with respect to the decision variables to define its search direction in line 5. To that end, the gradient-based solver allows us to use finite difference approximations of the gradient. This method provides us with a non-exact representation of the gradient, whereby we take the steps in (slightly) incorrect directions, which could lead to an increased number of iterations upon completion of the optimisation algorithm, which is undesired from a computational point of view. Therefore, the approach is to calculate the exact objective function gradient, which can be derived as follows starting from Equation (4.8):

$$\frac{\partial}{\partial\theta}J_N(\theta) = \frac{2}{N}\sum_{j=1}^{n_z}\sum_{k=1}^{N}\underbrace{\bar{\epsilon}_j^e(k,\theta)}_{\text{Term I}}\underbrace{\frac{\partial}{\partial\theta}\bar{\epsilon}_j^e(k,\theta)}_{\text{Term II}}. \tag{4.14}$$

We can calculate term I by evaluating the steady-state output of the candidate model. Term II, however, is unknown up to this point. The gradient-based method iterates its decision variable until the stopping criterium is met. Finally, the final decision variable selection forms the output of the system identification solution method. Note that there exist no guarantees upon the global optimality for the outputs of these gradient-based optimisation approaches. These methods use only local information (functions and their gradients at a point) in their search process, so they converge only to a local objective function minimum.

In this section, we have presented the solution method to the system identification problem, which is split between an initialisation and a gradient-based optimisation step. The BLA and global optimisation methods were suggested for the first step, whereas an off-the-shelf gradient-based optimisation algorithm that uses an interior-point method is suggested for the second step. To calculate the exact objective function gradient for the gradient-based optimisation method, a parameter sensitivity model approach is presented next.

4.3 Parameter sensitivity model

The motivation for deriving so-called *parameter sensitivity* models originates from rewriting the unknown Term II in Equation (4.14) by using Equation (4.6) according to

$$\frac{\partial}{\partial\theta}\bar{\epsilon}_j^e(k,\theta) = \frac{\partial}{\partial\theta}\bar{z}_j(k,\theta) - \underbrace{\frac{\partial}{\partial\theta}\bar{z}_{0,j}^e(k,\theta_0)}_{=0} = \frac{\partial}{\partial\theta}\bar{z}_j^e(k,\theta). \tag{4.15}$$

Hence, the unknown term is equivalent to the sensitivity of the steady-state model output with respect to the decision variables. By definition of dynamics (4.2), these steady-state dynamics can be written as

$$\bar{z}(k) = F\bar{x}(k) + G\bar{u}(k) + Hw(k). \tag{4.16}$$

In order to evaluate the sensitivity of this equality with respect to an scalar element θ_i from the decision variables θ, one quickly admits that the sensitivity of the entire steady-state solution (w, $\bar{u}$, $\bar{x}$, $\bar{y}$, $\bar{z}$) with respect to this decision variable is required.

Throughout this section, the following partial derivatives notation is adopted to improve readability. Firstly, all explicit dependencies on θ are dropped in the notations. Secondly, where no ambiguity occurs, the θ_i and y subscripts denote a partial derivative with respect to the i-th decision variable and the nonlinearity input, respectively. Therefore, one can write $\varphi_{\theta_i}(y) := \frac{\partial}{\partial\theta_i}\varphi(y,\theta)$ and $\varphi_y(y) := \frac{\partial}{\partial y}\varphi(y,\theta)$ for the nonlinearity, as well as, e.g., $A_{\theta_i} := \frac{\partial}{\partial\theta_i}A(\theta)$ for a linear block system matrix. Finally, a θ_i superscript implies that the variable is associated with the i-th parameter sensitivity model. The following assumption holds true for partial derivatives of Lur'e-type candidate model components.

Assumption 4.4

- $\varphi_y(y)$ *exists and is continuous for all* $y \in \mathbb{R}^{n_y}$ *and* $\theta \in \Theta$.
- $\varphi_{\theta_i}(y)$, A_θ, B_θ, L_θ, C_θ, D_θ, F_θ, G_θ, H_θ *exist and are continuous for all* $y \in \mathbb{R}^{n_y}$ *and* $\theta \in \Theta$.

Using this assumption, together with the observations in Equations (4.15) and (4.16), the notion of a *parameter sensitivity model* together with its instrumental properties are formalised in Theorem 4.1. These results are inspired by the continuous-time results in [66].

Theorem 4.1 (Lur'e-type parameter sensitivity model)
Consider a Lur'e-type candidate model $(\Sigma,\ \varphi)$ *according to Equation* (4.2). *Suppose that* Σ *and* φ *verify Assumptions 2.1 and 2.2, respectively, and let the candidate model be globally, exponentially uniformly convergent by verifying the conditions in Theorem 3.7. Then, one can define a Lur'e-type, so-called* parameter sensitivity model $(\Sigma^{\theta_i},\ \varphi^{\theta_i})$ *that complies to the following dynamics*

$$\Sigma^{\theta_i}:\begin{cases} x_{\theta_i}(k+1) = \mathcal{A}x_{\theta_i}(k) + \mathcal{B}\breve{u}^{\theta_i}(k) + \mathcal{L}\breve{w}^{\theta_i}(k), \qquad x_{\theta_i}(0) = x_{\theta_i,0}, & \text{(4.17a)}\\ \mathit{y}_{\theta_i}(k) = \mathcal{C}x_{\theta_i}(k) + \mathcal{D}\breve{w}^{\theta_i}(k), & \text{(4.17b)}\\ \mathit{z}_{\theta_i}(k) = \mathcal{F}x_{\theta_i}(k) + \mathcal{G}\breve{u}^{\theta_i}(k) + \mathcal{H}\breve{w}^{\theta_i}(k), & \text{(4.17c)}\end{cases}$$

$$\varphi^{\theta_i}:\ \{\quad \breve{u}^{\theta_i}(k)\ = \varphi^{\theta_i}\big(\mathit{y}_{\theta_i}(k), \Psi(k)\big). \qquad \text{(4.17d)}$$

In these dynamics, one recognises the following components:

- *The external excitation* $\breve{w}$ *depends on* θ *and the steady-state solution of* $(\Sigma,\ \varphi)$ *via*

$$\breve{w}^{\theta_i}(k) = \begin{bmatrix} \breve{w}_1^{\theta_i}(k)^\top & \breve{w}_2^{\theta_i}(k)^\top & \breve{w}_3^{\theta_i}(k)^\top \end{bmatrix}^\top, \qquad \text{(4.18)}$$

  where

$$\breve{w}_1^{\theta_i}(k) = A_{\theta_i}\bar{x}(k) + B_{\theta_i}\bar{u}(k) + L_{\theta_i}w(k) + B\varphi_{\theta_i}(\bar{\mathit{y}}(k)), \qquad \text{(4.19a)}$$
$$\breve{w}_2^{\theta_i}(k) = C_{\theta_i}\bar{x}(k) + D_{\theta_i}w(k), \qquad \text{(4.19b)}$$
$$\breve{w}_3^{\theta_i}(k) = F_{\theta_i}\bar{x}(k) + G_{\theta_i}\bar{u}(k) + H_{\theta_i}w(k) + G\varphi_{\theta_i}(\bar{\mathit{y}}(k)). \qquad \text{(4.19c)}$$

- *The linear block* Σ^{θ_i} *can be written as function of the candidate model parametrisation via*

$$\begin{aligned} &\mathcal{A} = A, && \mathcal{B} = B, && \mathcal{L} = \begin{bmatrix} I_{n_x} & 0_{n_x\times n_u} & 0_{n_x\times n_w}\end{bmatrix},\\ &\mathcal{C} = C, && && \mathcal{D} = \begin{bmatrix} 0_{n_y\times n_x} & I_{n_y} & 0_{n_y\times n_w}\end{bmatrix},\\ &\mathcal{F} = F, && \mathcal{G} = G, && \mathcal{H} = \begin{bmatrix} 0_{n_z\times n_x} & 0_{n_z\times n_u} & I_{n_z}\end{bmatrix}. \end{aligned} \qquad \text{(4.20)}$$

- *The nonlinearity* φ^{θ_i} *is (incrementally) sector bounded within the same bounds as* φ *and respects*

$$\varphi^{\theta_i}\big(\mathit{y}_{\theta_i}(k), \Psi(k)\big) = \Psi(k)\mathit{y}_{\theta_i}(k), \qquad \text{(4.21)}$$

  in which one recognises the external, time-varying input term

$$\Psi(k) = \varphi_y\big(\bar{\mathit{y}}(k)\big). \qquad \text{(4.22)}$$

The parameter sensitivity model respects the following instrumental properties:

- *It is globally, exponentially uniformly convergent by verifying the conditions in Theorem 3.7.*
- *The unique* T*-periodic steady-state model output* $\bar{\mathit{z}}_{\theta_i}$ *is equivalent to* $\frac{\partial}{\partial\theta_i}\bar{\epsilon}^e$.

Proof. See Appendix A.10 □

The steady-state solution of a single parameter sensitivity model provides us with information on all output channels $j = \{1,\dots,n_z\}$ for a single parameter $i \in \{1,\dots,n_\theta\}$ in the definition of the objective function gradient (4.14). Therefore, n_θ steady-state solutions of parameter sensitivity models are to be computed.

4.4 Computationally efficient objective function (gradient) computation

In the scope of system identification, we are required to calculate the steady-state output of the candidate model and its respective parameter sensitivity models in order to evaluate the system identification problem's objective function and its gradient with respect to the decision variables. This task is required throughout each iteration of the gradient-based solver. Therefore, it is desired to assess this steady-state output both *accurately* and *computationally efficient*.

Steady-state solution computation

One way of computing the steady-state solution for convergent nonlinear dynamical systems is simply to forward iterate the dynamics starting from an arbitrary initial condition for a sufficiently long time. The steady-state solution is globally exponentially stable, so the simulation always converges in forward time to the steady-state solution with any desired accuracy. The characterisation of the GEUC property did not provide us with an upper bound to the transient response time, so we have to find out while simulating whether we are sufficiently close to the steady state response already. Simulating a discrete time dynamical systems traces back to performing a sequence of algebraic operations, which is less computationally demanding than, e.g., solving the differential equations for its continuous-time counterpart. Nevertheless, we require to perform many simulations on typically large datasets and, hence, the steady-state model output calculation step is still expected to dominate the computation budget. Alternatively, the MTF algorithm is presented as a computationally efficient method to calculate steady-state model responses of globally exponentially uniformly convergent *continuous-time Lur'e-type systems with scalar signals* up to user-defined accuracy [57], [66]. This project extends this approach to *multivariable discrete-time* Lur'e-type systems $(\Sigma,\ \varphi)$ according to (2.1). The MTF algorithm can be used to accurately calculate steady-state solutions for normalised, symmetric Lur'e-type models and its main working principles are formalised in Theorem 4.2. Recall that steady-state operators for the linear block are defined in Section 2.1.

Theorem 4.2
Consider a Lur'e-type model (2.1) *that verifies the convergence conditions in Theorem 3.7 and is cast into its normalised, symmetric form* $(\tilde{\Sigma},\ \tilde{\varphi})$ *as per Lemma 2.4. Suppose that the system is subject to a periodic excitation* $w \in \ell_2^{n_w}(N)$*. Then, the sequence* $\left(\tilde{y}_{[i]}\right)_{i\in\mathbb{N}_0}$ *with* $\tilde{y}_{[0]}$ *being arbitrary inside* $\ell_2^{n_y}(N)$*, has a unique limit, denoted by* $\bar{\tilde{y}}$*, which results from iteratively applying the mappings*

$$\tilde{u}_{[i+1]} = \mathcal{F}_{\tilde{u}\tilde{y}} \circ \tilde{y}_{[i]}, \tag{4.23a}$$
$$\tilde{y}_{[i+1]} = \mathcal{F}_{\tilde{y}\tilde{u}} \circ \tilde{u}_{[i+1]} + \mathcal{F}_{\tilde{y}w} \circ w, \tag{4.23b}$$

because the composed operator $\mathcal{F}_{\tilde{y}\tilde{u}} \circ \mathcal{F}_{\tilde{u}\tilde{y}}$ *is a* contraction mapping*. Furthermore, this unique limit* $\bar{\tilde{y}}$ *coincides with the steady-state model output* $\bar{\tilde{y}}$ *of system* $(\tilde{\Sigma},\ \tilde{\varphi})$*.*

Proof. See Appendix A.11. □

Note that we have only shown the existence of a contraction mapping for Lur'e-type systems that were cast into their symmetric, normalised form. This theorem can be used to accurately compute the steady-state solution of a convergent Lur'e-type model candidate $(\Sigma(\theta), \varphi(\theta))$ and its parameter sensitivity systems $\left(\Sigma^{\theta_i}(\theta), \varphi^{\theta_i}(\theta)\right)$, which we require for the cost-function (gradient) evaluation. To improve upon the computational efficiency of the implementation of the MTF algorithm, we opt to evaluate the LTI dynamics via the *frequency domain steady-state operators* that are defined in Equation (2.10), while transforming the intermediate signals forth and back between the time- and frequency-domain using the DFT and IDFT (see Definition 1.4). The approach to calculate a steady-state solution for an arbitrary Lur'e-type model that verifies the convergence conditions in Theorem 3.7 is summarised in Algorithm 4.

Algorithm 4: Efficient steady-state solution computation for Lur'e-type systems (Σ, φ).

1 Apply the transformation of Lemma 2.4 on $(\Sigma,\ \varphi)$ to obtain $(\tilde{\Sigma},\ \tilde{\varphi})$.
2 Set iteration counter $i = 0$, stopping criterium $\eta_{[0]} = \infty$ and stopping tolerance $\eta^* > 0$.
3 Compute W from w using DFT
4 Evaluate LTI dynamics in frequency domain (assume $\bar{\tilde{u}}_{[0]} = 0_{N \times n_u}$): $\bar{\tilde{Y}}_{[0]} = \hat{\mathcal{F}}_{\tilde{Y}W} \circ W$.
5 Compute $\bar{\tilde{y}}_{[0]}$ from $\bar{\tilde{Y}}_{[0]}$ using IDFT.
6 **while** $\eta_{[i]} > \eta^*$ **do**
7 Evaluate the nonlinearity in time domain: $\bar{\tilde{u}}_{[i+1]} = \hat{\mathcal{F}}_{\tilde{u}\tilde{y}} \circ \bar{\tilde{y}}_{[i]}$.
8 Compute $\bar{\tilde{U}}_{[i+1]}$ from $\bar{\tilde{u}}_{[i+1]}$ using DFT.
9 Evaluate LTI dynamics in frequency domain: $\bar{\tilde{Y}}_{[i+1]} = \hat{\mathcal{F}}_{\tilde{Y}\tilde{U}} \circ \bar{\tilde{U}}_{[i+1]} + \bar{\tilde{Y}}_{[0]}$.
10 Compute $\bar{\tilde{y}}_{[i+1]}$ from $\bar{\tilde{Y}}_{[i+1]}$ using IDFT.
11 set $i = i + 1$.
12 **end**
13 Define $\bar{\tilde{Y}}$ and $\bar{\tilde{U}}$ as $\bar{\tilde{Y}}_{[i]}$ and $\bar{\tilde{U}}_{[i]}$.
14 Apply the inverse transformation of Lemma 2.4 on $\bar{\tilde{Y}}$ and $\bar{\tilde{U}}$ to define $\bar{Y}$ and $\bar{U}$.
15 Evaluate additional elements of Σ's steady-state solution in the frequency domain:
$\bar{Z} = \hat{\mathcal{F}}_{ZU} \circ \bar{U} + \hat{\mathcal{F}}_{ZW} \circ W, \qquad \bar{X} = \hat{\mathcal{F}}_{XU} \circ \bar{U} + \hat{\mathcal{F}}_{XW} \circ W$
16 Compute $\bar{u}$, $\bar{y}$, $\bar{z}$, $\bar{x}$ from $\bar{U}$, $\bar{Y}$, $\bar{Z}$, $\bar{X}$ using IDFT
17 **return** steady-state solution of $(\Sigma,\ \varphi)$: $(\bar{u}, w, \bar{x}, \bar{y}, \bar{z})$.

In line 1 of Algorithm 4, we transform (Σ, φ) into its normalised symmetric form $(\tilde{\Sigma}, \tilde{\varphi})$. Thereafter, in lines 2 to 13, we recognise a computationally efficient implementation of the discrete-time multivariable MTF algorithm, which can be better understood by taking a closer look at Figure 4.3. In this figure, we can see that first the contribution of excitation w into steady-state output $\bar{\tilde{Y}}$ is computed in the frequency domain. Consequently, we evaluate the i-th iteration's approximation of steady-state nonlinearity output $\bar{\tilde{U}}$ in the time-domain before calculating its contribution to the i'th iteration's approximation of $\bar{\tilde{Y}}$ in the frequency domain again. We then check whether the (normalised) 'distance' between the current and previous approximation of steady-state output $\bar{\tilde{Y}}$ meets the stopping threshold via the stopping criterium

$$\eta_{[i]} = \frac{\left\| \mathrm{vect}\left(\bar{\tilde{Y}}_{[i]} - \bar{\tilde{Y}}_{[i-1]}\right) \right\|_2}{\left\| \mathrm{vect}\left(\bar{\tilde{Y}}_{[i]}\right) \right\|_2}, \tag{4.24}$$

in which $\mathrm{vect}\,(\cdot)$ is an operator that reshapes all elements of a matrix into a single column vector with a column-major layout. In case the stopping tolerance is met by the criterium, we output the i'th iteration's approximations as $\bar{\tilde{Y}}$ and $\bar{\tilde{U}}$. Otherwise, we apply the mappings once more by a new loop iteration.

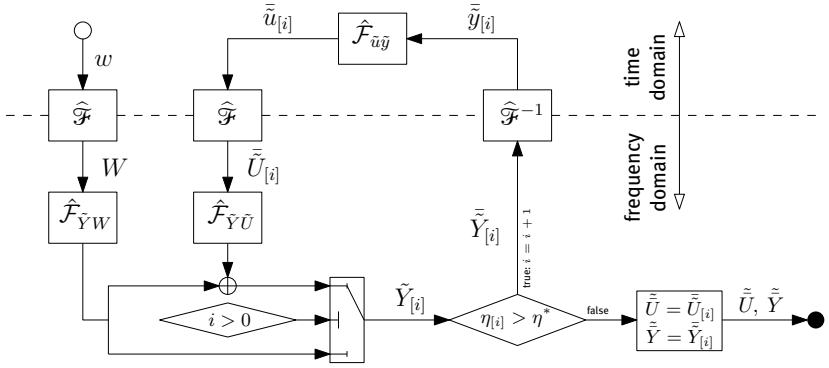


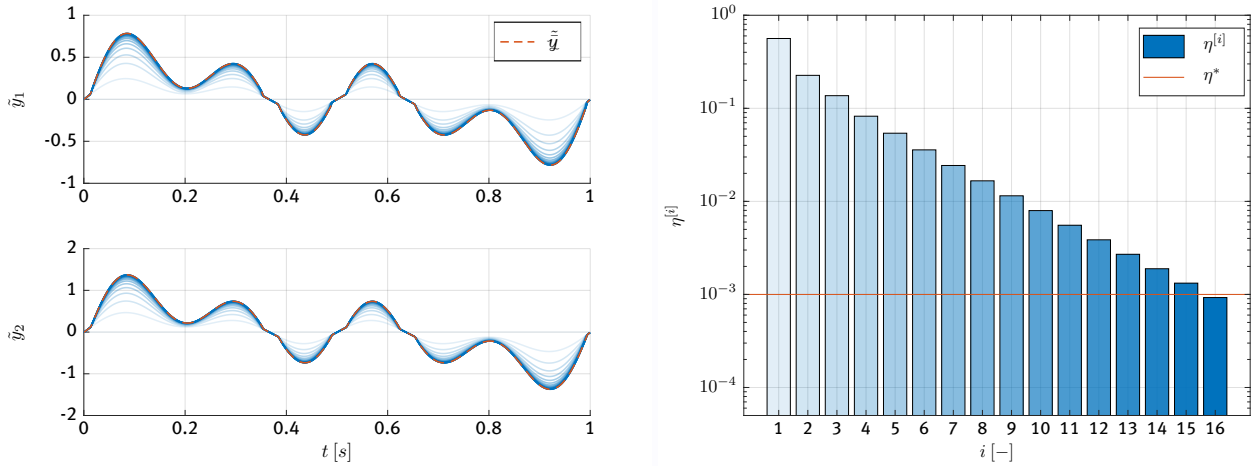
Figure 4.3: Discrete-time multivariable MTF algorithm, where $\hat{\mathcal{F}}$ is the N-point DFT operator.

We now shift the focus back to Algorithm 4. Upon completion of the MTF algorithm, in line 14, we inversely transform the nonlinear steady-state in- and outputs of $(\tilde{\Sigma}, \tilde{\varphi})$ into their (Σ, φ) equivalents. We then use frequency domain steady-state operators to calculate the missing elements of the steady-state solution in line 15 and transform everything back to the time-domain in line 16, which concludes the algorithm.

To illustrate the behaviour of the MTF algorithm and its computational benefits, we apply this method to calculate steady-state solutions for the Lur'e-type system in Example 4.1.

Example 4.1 (MTF algorithm example)

Let us once more consider the Lur'e-type model $(\Sigma,\ \varphi)$ of Example 3.2 and, additionally, select a sampling frequency $f_s = 1/2048$ Hz. The system is excited by the harmonic w, such that $w(k) = \frac{1}{10}\sum_{i=1}^{3}\sin(2\pi i k)$ on a measurement interval $k \in \{1, \ldots, N\}$ of length $N = 2048$ samples. Therefore, the measurement regards exactly one period of the harmonic excitation w. Previously, in Example 3.2, it was shown that this system is globally exponentially uniformly convergent by verifying the conditions in Theorem 3.6. We can transform $(\Sigma,\ \varphi)$ into its normalised, symmetric form $(\tilde{\Sigma},\ \tilde{\varphi})$ as per Lemma 2.4 in order to apply the MTF algorithm to calculate the steady-state solution of $(\tilde{\Sigma},\ \tilde{\varphi})$. For this example system, the MTF algorithm results are shown in Figure 4.4, where the color darkness of lines increases along the iteration history. In this figure, we also find the actual steady-state output $\bar{\tilde{y}}$ of $(\tilde{\Sigma},\ \tilde{\varphi})$, which was found by forward iterating the dynamics for an excessively long time.



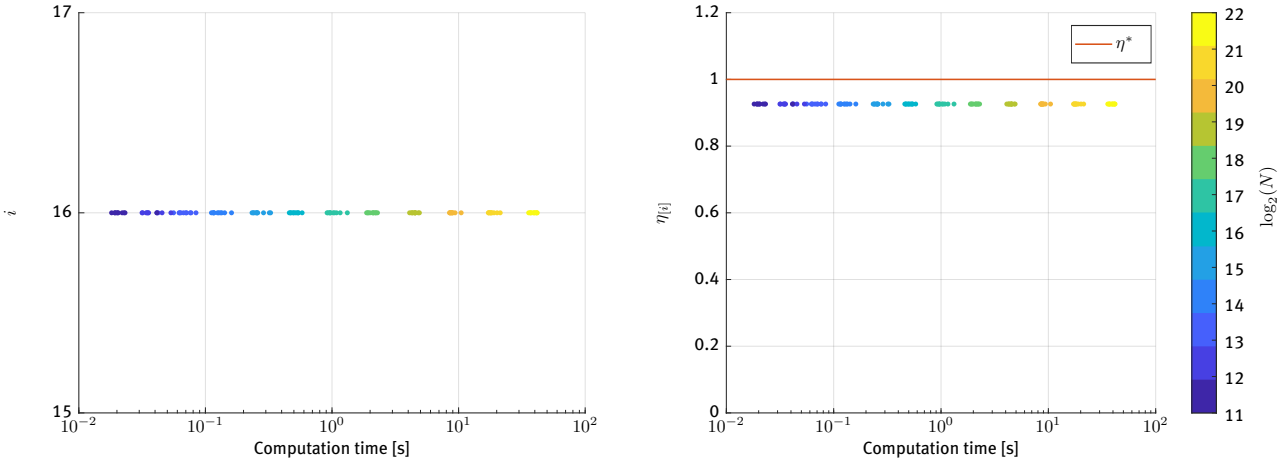
(a) $\bar{\tilde{y}}_{[i]}$ along the iteration history.

(b) $\eta_{[i]}$ along the iteration history.

Figure 4.4: MTF algorithm accuracy along its iteration history.

In Figure 4.4a, we can see that, within the deadzone region, $\bar{\tilde{y}}_{[i]}$ already coincides with $\bar{\tilde{y}}$ for $i = 1$. Outside of the deadzone region, the approximate steady-state output already takes a reasonable shape, but we still observe a clear difference with respect to the true steady-state solution $\bar{\tilde{y}}$. Throughout the iteration history, $\bar{\tilde{y}}_{[i]}$ resembles $\bar{\tilde{y}}$ better and better. Furthermore, the stopping criterium decreases in Figure 4.4b until it passes the threshold η^*, which means that the approximate steady-state outputs have converged sufficiently to the unique fixed point of the contraction mapping after nine MTF algorithm iterations.

To test the computational efficiency of the proposed method, we execute the MTF algorithm for various measurement interval lengths $N \in \{2^{11}, 2^{12}, \ldots, 2^{22}\}$ such that the measurement interval consists of an integer multiple of excitation periods. Each test case was executed ten times to assess the computational consistency of the algorithm and reject outliers. We define the return state of the MTF algorithm as the tuple $(i, \eta_{[i]})$ upon algorithm completion. This return state is shown for all instances of all testcases in Figure 4.5.



(a) Total number of MTF iterations.

(b) Stopping criterium value on final iteration.

Figure 4.5: MTF algorithm return state for many executions as function of measurement interval length N.

The most important conclusion from Figure 4.5a is that the number of MTF algorithm iterations appears independent from the measurement interval length N. Also, it can be seen that the computation time is roughly proportional to the measurement interval length. In all cases, the MTF algorithm has contracted sufficiently to the fixed point of the iterative mapping as can be seen in Figure 4.5b. Lastly, no large variations in computation time are observed along repetition of the MTF algorithm for constant measurement interval length.

The final contribution of this example is to quantify the computational incentive of using the MTF algorithm to compute steady-state solutions rather than forward iterating the nonlinear dynamics. The results of this analysis are shown in Figure 4.6.

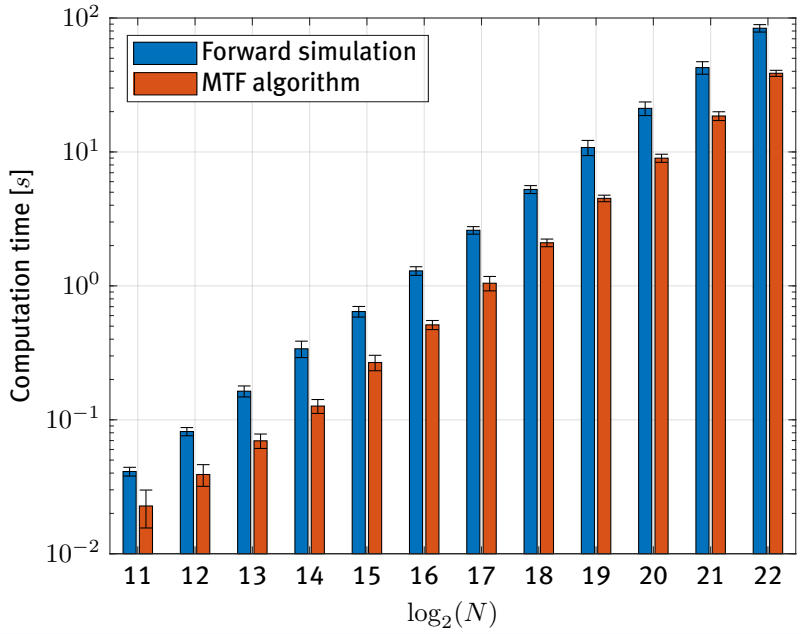


Figure 4.6: Computation time comparison.

The forward iteration method is implemented as follows: (i) Forward iterate the dynamics for two excitation periods starting from arbitrary initial state, (ii) Verify whether the last period is in steady-state, by comparing its contents to the previous period using a stopping criterium similar to Equation (4.24) and (iii) if necessary, forward iterate the dynamics for another excxitation period and repeat step (ii) until the steady-state is calculated sufficiently accurate. This solution method is used to calculate the steady-state solution of $(\tilde{\Sigma}, \tilde{\varphi})$ on the previously mentioned grid of measurement interval lengths. The average computation time for each test case of both methods is shown in Figure 4.6 together with its standard deviation along the ten repetitions. From this figure, one can conclude that the MTF algorithm manages to compute steady-state model responses roughly three times faster than the forward iteration method. Given the large computational budget for steady-state model response computation, this induces the expectation of a significant reduction for the total system identification algorithm runtime by roughly 66 %.

Efficient objective function (gradient) computation

In the previous paragraph, we have explained how to compute the unique steady-state solution of convergent Lur'e-type models computationally efficiently via Algorithm 4. However, up to this point, it is still an open question how this method connects to the solution for the system identification problem. We have implemented a four-stage strategy to evaluate the objective function and its gradient with respect to the decision variables. A schematic representation of this algorithm is shown in Figure 4.7.

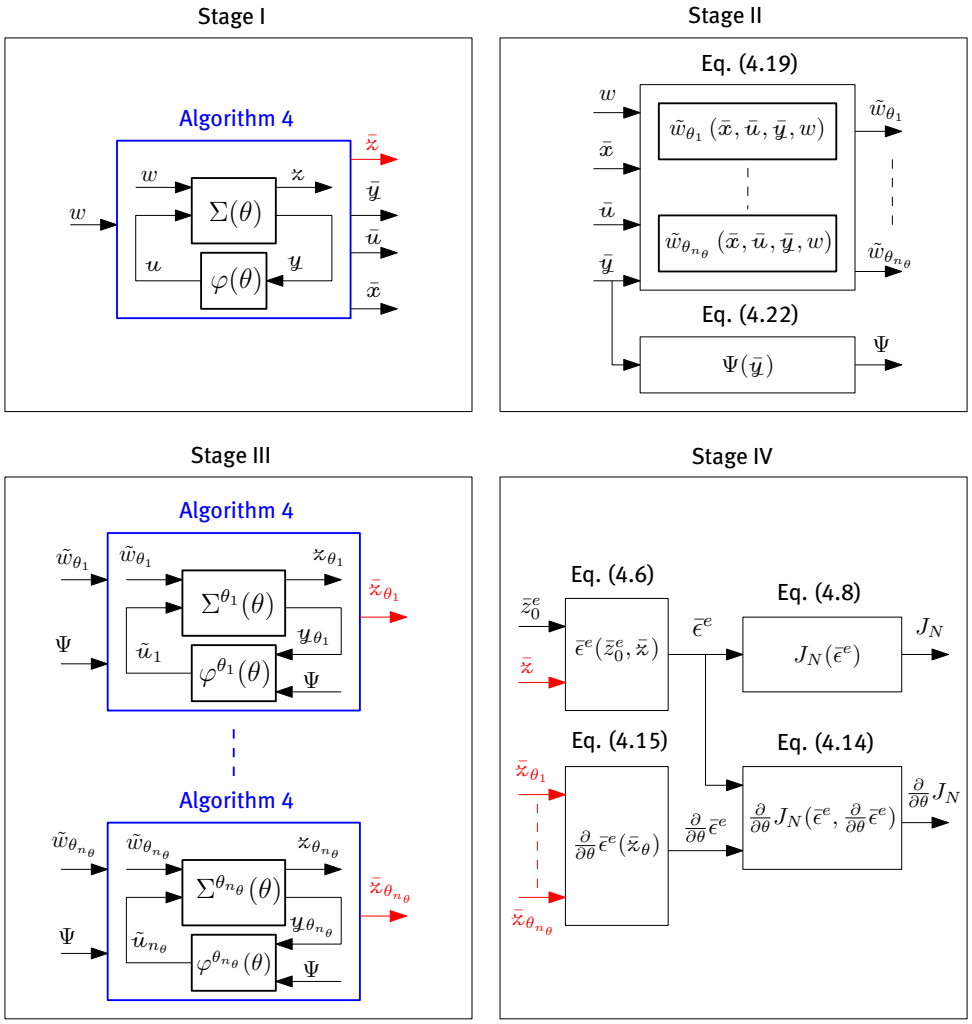


Figure 4.7: Efficient objective function (gradient) computation.

The computationally efficient objective function (gradient) computation strategy kicks off in **Stage I**, which is referred to as the *steady-state solution* stage, where we first interpret the decision variable θ as a Lur'e-type candidate model $(\Sigma(\theta),\ \varphi(\theta))$. Next, we check whether there exist stability guarantees for $(\Sigma(\theta),\ \varphi(\theta))$ in terms of the GEUC property by verifying the feasibility of the condition in Theorem 3.7. When there are no stability guarantees available, we instantly reject the proposed set of decision variables as a solution to the problem. Otherwise, we apply Algorithm 4 to this Lur'e-type model in order to calculate its steady-state solution $(w, \bar{u}, \bar{x}, \bar{y}, \bar{z})$. In **Stage II**, which is referred to as the *parameter sensitivity inputs* stage, we take the $(w, \bar{x}, \bar{u}, \bar{y})$ elements of the steady-state solution of $(\Sigma(\theta),\ \varphi(\theta))$ and translate them via the static mappings in Equations (4.19) and (4.22) into $(\tilde{w}_{\theta_1}, \ldots, \tilde{w}_{\theta_{n_\theta}})$ and Ψ as a preprocessing step before the third stage of the algorithm commences. In **Stage III**, the *sensitivity steady-state output stage*, we start by defining the set of Lur'e-type parameter sensitivity models associated with $(\Sigma(\theta),\ \varphi(\theta))$ as per Theorem 4.1. Note that in this stage, we require to define n_θ parameter sensitivity models to evaluate the sensitivity with respect to each element in θ. The next step is to calculate the steady-state output of all these globally, exponentially uniformly convergent parameter sensitivity models, for which we can again apply Algorithm 4. In this stage, we can improve upon the computational performance of the algorithm even further by realising that the computations on all sensitivity systems are independent. Therefore, we can benefit from *parallel computing* possibilities in MATLAB [91]. We can assign a task, i.e., computing the steady-state output of a parameter sensitivity model by executing Algorithm 4, to each core of the computer's multicore processor in order to run simulations in parallel and reduce the overall runtime of the system identification. **Stage IV**, the *objective function (gradient)* stage starts once the steady-state outputs of all parameter sensitivity models are calculated. In this stage, we first calculate the steady-state model output errors by comparing observations to the simulated steady-state model outputs in Equation (4.6), which provides all information to evaluate the objective function via Equation (4.8). The gradient of the steady-state output error with respect to the decision variables can be found by combining the results of Stage III with Equation (4.15). Then finally, we can calculate the objective function gradient itself by evaluating Equation (4.14), which concludes the computations within one iteration of the gradient-based nonlinear optimisation routine.

4.5 Summary

This chapter covers the definition of and solution to the nonlinear system identification problem considered in this project. In particular, the system identification problem is translated into a standard-form constrained optimisation problem, whose decision variables θ parametrise a candidate model (4.2). The decision variables can be selected from the set of globally, exponentially convergent discrete-time Lur'e-type models. The simulation error is desired to be minimised, which defines the objective function to the optimisation problem as the least-squares criterium in (4.8) and closes the system identification problem formulation. To solve such a problem, we adopt a two-step approach: 'Decision variable initialisation' and 'gradient-based optimisation'. The first step can be performed via a BLA of the nonlinear dynamics, or a global optimisation approach. The second step can be completed via an interior-point method that is readily available in the MATLAB optimisation toolbox. This method requires to evaluate the objective function, as well as its gradient with respect to the decision variables in each iteration. The prior term can be calculated via the steady-state output of the candidate model, whereas the latter term is found exactly by computing the steady-state output of a set of Lur'e-type parameter sensitivity models. Obviously, the computational performance of the system identification algorithm would benefit greatly from fast and accurate computation of all these steady-state solutions. To that end, the MTF algorithm is implemented, which is shown to contract iteratively to the unique steady-state solution of convergent Lur'e-type models. The computational performance of having such an MTF algorithm is illustrated by an example and this chapter ends by an explanation of the efficient objective function (gradient) computation algorithm that is validated on experimental and simulation case studies in the next chapter.

Chapter 5

Identification results

In this chapter, the proposed system identification algorithm is validated on benchmark experimental data in Section 5.1. The selected dataset is well-known in the nonlinear system identification field and can, therefore, be used to compare the estimated model performance quantitatively to existing solutions. The discussion of system identification results continues by a simulation case study for a first-order dynamical system in Section 5.2. Non-unique parametrisation issues and suggested solutions are illustrated. New insights in the behaviour of our methods can be gained in simulation studies, as the 'true' system is known.

5.1 Experimental study: Silverbox benchmark

The *Silverbox system* has been reported as an electronic implementation of the Duffing oscillator and its input-output behaviour is proposed as a benchmark dataset for the research field of nonlinear system identification [58], [59], [92]. Its behaviour can be modelled as the mechanical system that is shown schematically in Figure 5.1a. In this figure, we recognise a mass m_1 moving purely in the horizontal $\vec{x}_1$ direction over a frictionless surface. Its position is measured as a function of time and denoted by $x_1(t)$, relative to its initial position $x_1(0) = x_{1,0}$. The model dynamics are induced by a scalar harmonic force perturbation $w(t)$, a nonlinear spring characterised by coefficient k_3 and a linear spring-damper combination parametrised by k_1 and d_1. Newton's second law can be applied to this nonlinear system to derive the following second-order differential equation as the equation of motion for this mechanical system:

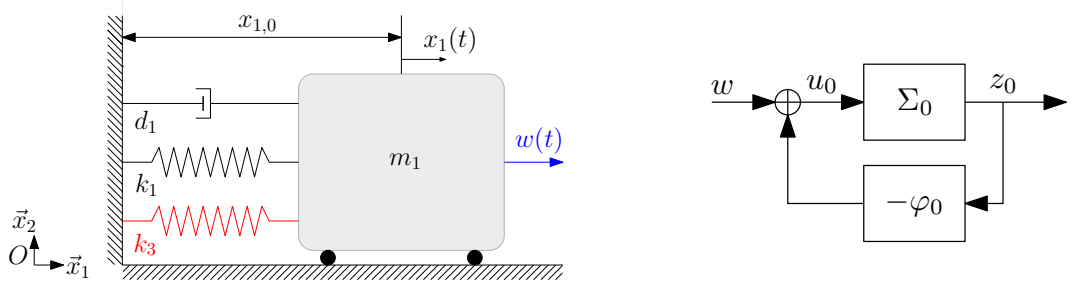
$$m_1\ddot{x}_1 + d_1\dot{x}_1(t) + k_1x_1(t) + k_3x_1^3(t) = w(t). \tag{5.1}$$

These nonlinear dynamics can be written into a continuous-time Lur'e-type form:

$$\Sigma_0 : \begin{cases} \dot{x}_0(t) = \begin{bmatrix} 0 & 1 \\ -\frac{k_1}{m_1} & -\frac{d_1}{m_1} \end{bmatrix} x_0(t) + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u_0(t), & \text{(5.2a)} \\ z_0(t) = \begin{bmatrix} 1 & 0 \end{bmatrix} x_0(t), & \text{(5.2b)} \end{cases}$$

$$u_0(t) = w(t) - \varphi_0(z_0(t)), \tag{5.2c}$$

as shown schematically in the block-diagram in Figure 5.1b.



(a) Mass-spring-damper model.

(b) Lur'e-type model.

Figure 5.1: Interpreting the Silverbox as a mass-spring-damper- and a Lur'e-type model.

In this Lur'e-type model, we have selected the state $x_0(t) = \begin{bmatrix} x_1(t) & \dot{x}_1(t) \end{bmatrix}^\top$, together with the scalar measured output $z_0(t) = x_1(t)$ and a cubic nonlinear function

$$\varphi_0(z_0(t)) = \frac{k_3}{m_1} z_0(t)^3. \tag{5.3}$$

The input force $u_0(t)$ consists of contributions by the external force perturbation and the nonlinear spring. Let us assume that $x_1(t)$ is measured at uniformly spaced time instants with sampling interval T_s. Also, a zero-order hold assumption is imposed to the external force excitation $w(t)$, such that

$$w(t) = w(kT_s), \qquad \forall\, kT_s \leq t < (k+1)T_s,\ k \in \mathbb{N}_0. \tag{5.4}$$

Because of these assumptions, well-known discretization techniques for LTI models can be applied to Equations (5.2a) and (5.2b), which together with the static relation in Equation (5.2c) results in a discrete-time Lur'e-type model describing the dynamics of the silverbox system [63, p. 119]. There has been extensive research on Duffing oscillators and, depending on the choice of constants m_1, k_1, d_1 and k_3, it is known that solutions to (5.1) can either be GES or chaotic with a high sensitivity to initial conditions [93].

In the remainder of this section, the system identification algorithm proposed in Chapter 4 is applied to data retrieved from the silverbox system, since prior knowledge confirms that these dynamics fit the model class (4.2). Additionally, one can guarantee that the identified model describes a duffing oscillator that operates in a convergent regime by enforcing feasibility of the condition in Theorem 3.7, which can be exploited, e.g., for controller design. The identification experiment is explained, followed by numerous system identification results. Thereafter, model performances are compared quantitatively to state-of-the-art benchmark results.

5.1.1 Identification experiment

The identification experiment that covers the Silverbox benchmark dataset is presented in Figure 5.2. In this figure, the scalar excitation w and output z_0^e are shown as function of time measured at a sampling frequency of $610.35\,\mathrm{Hz}$, both having a distinctive arrow-shape. The dataset can be divided into parts I and II, by which the arrow's 'head' and 'tail' are denoted, respectively. Part I, governing samples 1-40585, is a *band-limited Gaussian white noise* sequence between $0\,\mathrm{Hz}$ and $200\,\mathrm{Hz}$ with a linear increase in amplitude forward in time. Part II consists of ten realisations of a *random-phase odd multisine* sequence that are mutually interconnected by zero inputs. After part II, a sine-sweep sequence starts, but this part of the data is excluded from our analysis, which is in line with other approaches reported in literature [59].

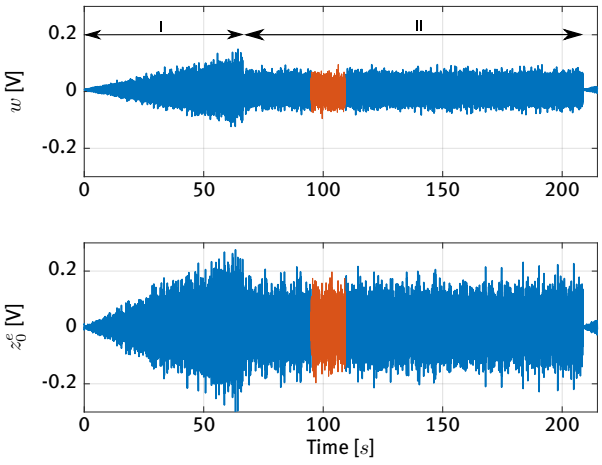


Figure 5.2: The arrow-shaped Silverbox dataset (——), consisting of a so-called 'head' and 'tail', respectively denoted by I and II. The 'tail' contains ten odd multisine realisations, one of which is denoted by (——).

Let us isolate the multisine realisation that was highlighted in Figure 5.2, and analyse its properties in time- and frequency-domain in Figure 5.3. The top plot of Figure 5.3 illustrates the division of the arrow's tail in ten sections that are separated by a sequence of zeros. One of these sections is shown in this figure and the previously mentioned multisine realisation is a subset of this section of length $N = 8192$, as shown in red. A frequency domain representation of the multisine realisation is included in the bottom plot of Figure 5.3 to verify whether the correct subsets of the arrow tail are selected as estimation data. It can be verified that all 1342 odd harmonics on a bandwidth between $0\,\mathrm{Hz}$ and $200\,\mathrm{Hz}$ are evenly excited, which indeed complies to the experiment design motivation in [58]. There is no information in the data of harmonics having a frequency above the passband until the Nyquist frequency f_N.

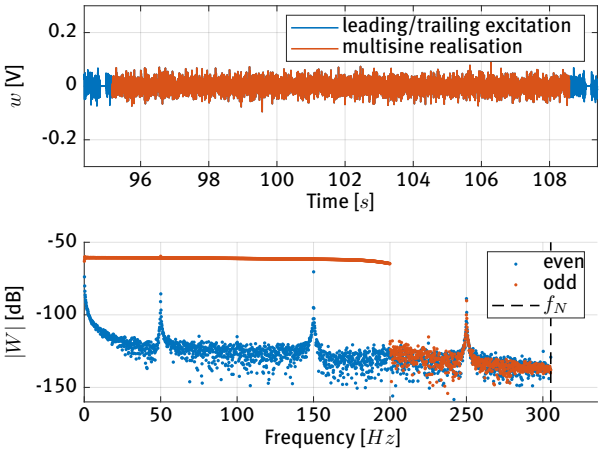


Figure 5.3: Random-phase odd-multisine excitation realisation in time- and frequency-domain.

As a final preprocessing step, the mean and first-order polynomial trends were removed from part I and the first nine multisines in part II. Then, the observations can be presented as the model estimation and validation datasets in Figure 5.4. The last multisine is omitted to train the model on comparable data with respect to existing solutions [92]. Also, note that the excitation amplitude of the validation dataset exceeds the amplitudes seen during training. This allows for testing extrapolation properties of identified models.

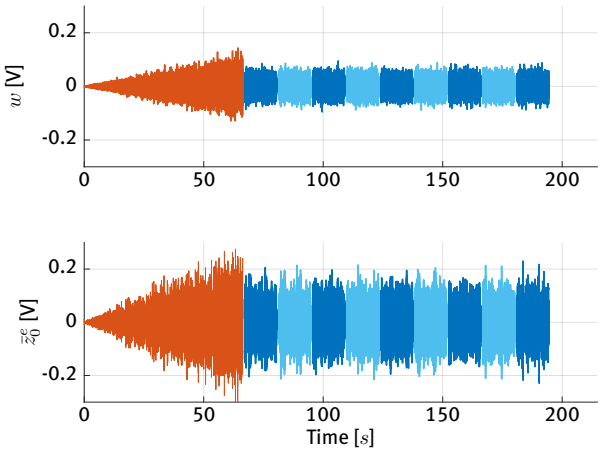


Figure 5.4: Silverbox identification experiment, split between the nine estimation (——), (——) and a single validation dataset (——).

5.1.2 Model structure selection

The main purpose of the model structure selection step is to define the relation between the decision variables θ and a Lur'e-type candidate model $(\Sigma(\theta), \varphi(\theta))$ for the specific system identification problem at hand. By selection of structure, we mean, e.g., definition of (i) the model order of $\Sigma(\theta)$, (ii) the nonlinearity $\varphi(\theta)$ itself and (iii) whether or not convergence is enforced. Not enforcing convergence is selected for the sake of benchmarking with reported results in the literature. Numerous model structure selections are proposed in Table 5.1, each of which is denoted by $(\Sigma(\theta_i), \varphi_i(\theta_i))$ for $i \in \{\mathsf{I}, \ldots, \mathsf{VIII}\}$.

Table 5.1: Model structure selections for Silverbox system identification.

ID	$\Sigma(\theta_i)$				$\varphi_i(\theta_i)$			Stability guarantees?	n_θ
	n_x	n_y	$z = y$	$w - \varphi_i(y) = u$	Name	Definition	Details		
I	2	1	✓	✓	zero	(5.8)		✓	5
II	2	1	✓	✓	polynomial	(5.9a)		✗	8
III	2	1	✓	✓	spline	(5.10a)		✗	7
IV	2	1	✓	✓	spline	(5.10a)		✓	7
V	2	1	✓	✓	NN	(5.11a)	$n_n = 3$	✗	14
VI	2	1	✓	✓	NN	(5.11a)	$n_n = 10$	✗	35
VII	2	1	✓	✓	NN	(5.11a)	$n_n = 3$	✓	14
VIII	2	1	✓	✓	NN	(5.11a)	$n_n = 10$	✓	35

The variety of model structures induces different candidate model parametrisations. The decision variables complying to a candidate model of a certain structure are characterised by

$$\theta_i = \begin{bmatrix} \theta_\Sigma \\ \theta_{\varphi_i} \end{bmatrix}, \qquad \forall\, i \in \{\mathsf{I}, \ldots, \mathsf{VIII}\}, \tag{5.5}$$

in which θ_Σ and θ_{φ_i} denote the parameters of the linear and nonlinear block for Lur'e-type candidate model $(\Sigma(\theta_i), \varphi_i(\theta_i))$ of class (4.2), respectively. The characterisation of these blocks is discussed next.

Linear block structure

All of the selected model structures use the same parametrisation for the linear block $\Sigma(\theta_i)$. Given the prior knowledge on the Silverbox dynamics (5.2), it was decided to adopt a discrete-time second-order LTI state-space model that is interconnected via feedback by a SISO nonlinearity, such that $n_x = 2$ and $n_y = 1$. Furthermore, it is assumed that the measured output channel z also serves as nonlinearity input channel y. Finally, the nonlinearity output u and the external excitation w are assumed to be summable towards a scalar input of $\Sigma(\theta_i)$ as illustrated in Figure 5.1b. All constraints on the linear block have been explained, so the next step is to introduce a candidate model's linear block parametrisation as

$$\theta_\Sigma = \begin{bmatrix} a_1 & a_2 & b_0 & b_1 & b_2 \end{bmatrix}^\top, \tag{5.6}$$

whose entries characterise the unique SISO transfer function of $\Sigma(\theta_i)$ as

$$\mathcal{G}_{\Sigma(\theta_i)}(z) = \frac{b_0 z^2 + b_1 z + b_2}{z^2 + a_1 z + a_2}. \tag{5.7}$$

A transfer function coefficients parametrisation reduces the number of model parameters compared to a full parametrisation of all matrix elements in a state-space model. Nevertheless, the system identification algorithm requires a characterization of the linear block in terms of a state-space form in which we distinguish between external and nonlinear inputs. Therefore, a state-space representation in *controllability canonical form* (see [63, p. 252]) can be derived from $\mathcal{G}_{\Sigma(\theta_i)}(z)$ in (5.7) and characterises $\Sigma(\theta_i)$ into a state-space form that fits the model class (4.2) with

$$A(\theta) = \begin{bmatrix} 0 & 1 \\ -a_2 & -a_1 \end{bmatrix}, \quad B(\theta) = -\begin{bmatrix} 0 \\ 1 \end{bmatrix}, \qquad L(\theta) = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \qquad C(\theta) = \begin{bmatrix} b_2 - a_2 b_0 \\ b_1 - a_1 b_0 \end{bmatrix}^\top, \qquad D(\theta) = b_0.$$

Note that the state space matrices associated with the output equation are, respectively constrained by $F(\theta) = C(\theta)$, $G(\theta) = 0_{n_y}$ and $H(\theta) = D(\theta)$, due to the $\varkappa = y$ assumption. Also, note that $B = -L$ since the input of Σ as defined in Equation (5.2c) is split between an external and a nonlinear input in the model class. Because the linear block is constrained to this canonical form, the possibility of non-unique parametrisation due to similarity transformations is eliminated from the system identification problem.

Nonlinear block structures

The different candidate model parametrisations are induced by the various nonlinearity structures $\varphi_i(\theta_i)$ and are now treated one after the other. Instances of all nonlinearity structures are shown in Figure 5.9, which helps to understand the different characterisations. The first considered nonlinearity is actually not a 'real' nonlinear function, but simply ensures a zero output via

$$\varphi_{\text{I}}(y(k)) = 0. \tag{5.8}$$

Obviously, this nonlinearity adds no additional model parameters. In an attempt to improve representation capabilities beyond those of a linear model, a *cubic polynomial* is selected as the nonlinearity of structure II. This nonlinear function fits the prior knowledge on the Silverbox system well and its definition can be written together with its parametrisation according to

$$\varphi_{\text{II}}(y(k)) = \sum_{i=1}^{3} k_i y(k)^i, \tag{5.9a}$$

$$\theta_{\varphi_{\text{II}}} = \begin{bmatrix} k_1 & k_2 & k_3 \end{bmatrix}^\top \in \mathbb{R}^3. \tag{5.9b}$$

A downside of the polynomial nonlinearity is that structure II does not allow for stability guarantees, since its nonlinear function can obviously not be (incrementally) sector bounded within finite bounds.

To address this problem, the third and fourth model structures implement a so-called '*spline*' nonlinearity, which behaves as a pure cubic polynomial on a certain domain including the origin and is linear outside. One can define this nonlinearity and its parametrisation as

$$\varphi_{\text{III,IV}}(y(k)) = \begin{cases} k_1 y(k)^3 & \text{for} \quad |y(k)| \leq y_* \\ \alpha y(k) + \beta & \text{for} \quad y(k) > y_* \\ \alpha y(k) - \beta & \text{for} \quad y(k) < -y_*, \end{cases} \tag{5.10a}$$

$$\theta_{\varphi_{\text{III,IV}}} = \begin{bmatrix} k_1 & y_* \end{bmatrix}^\top \in \mathbb{R}^2_{\geq 0}, \tag{5.10b}$$

where $\alpha = 3k_1 y_*^2$ and $\beta = -2k_1 y_*^3$. This nonlinear function verifies Assumption 2.2 and its cone-bounded sector constraints are characterised by $\Omega = \bar{\Omega} = 3k_1 y_*^2$ for nonnegative constants k_1 and y_*.

Both previously mentioned nonlinearity definitions are heavily dependent on prior knowledge of the Silverbox system dynamics. Consequently, the previous model structures are not expected to be applicable to data obtained from systems other than the Silverbox. This motivates the introduction of a *feed-forward neural network* as the nonlinearity of structures V-VIII, as the universal approximation theorem states that such a network can approximate any continuous function for inputs in a specific range (see [37], [94]).

A schematic view on these feed-forward neural networks with a single hidden layer can be found in Figure 5.5. In this figure, it can be seen that the neural network input $y(k)$ splits inside the hidden layer into n_n branches, each of which is scaled linearly with its associated input weight. Consecutively, in each branch the input bias is added, whereafter the nonlinear activation function is applied. In the output layer, each branch is scaled by an output weight, whereafter all branches are summed together with the output bias to define the neural network output $u(k)$. In summary, the neural network nonlinearity and its parametrisation can be defined as

$$\varphi_{\text{V-VIII}}(y(k)) = {W^{[2]}}^\top \sigma\left(W^{[1]} y(k) + b^{[1]}\right) + b^{[2]}, \tag{5.11a}$$

$$\theta_{\varphi_{\text{V-VIII}}} = \begin{bmatrix} W^{[1]} & b^{[1]} & W^{[2]} \end{bmatrix}^\top \in \Theta_{\varphi_{\text{V-VIII}}} \subset \mathbb{R}^{3n_n}, \tag{5.11b}$$

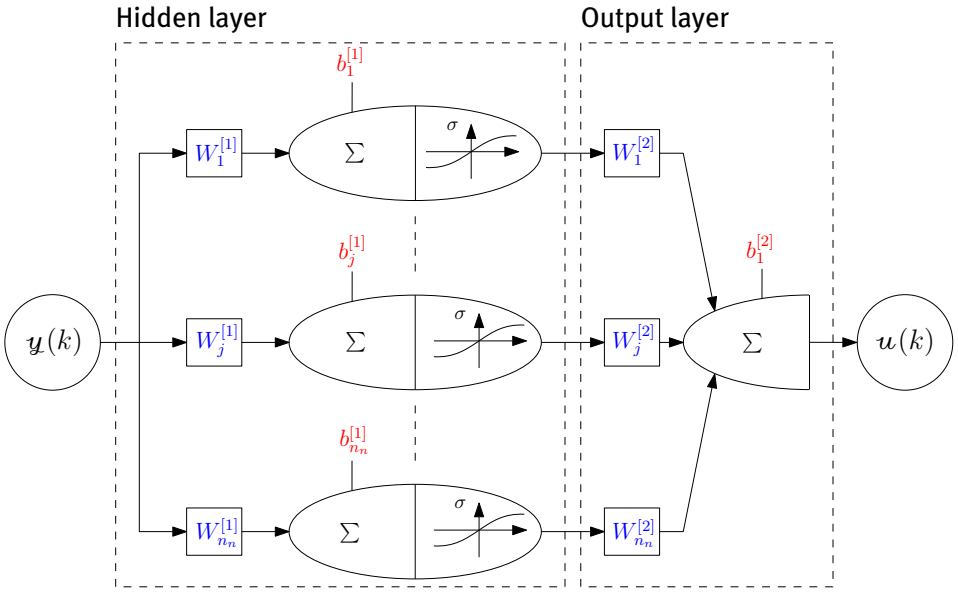


Figure 5.5: Schematic representation of a feed-forward neural network with a single hidden layer. Subscripts denote vector element indices.

in which $\sigma(\cdot)$ is the nonlinear activation function for which a hyperbolic tangent was selected in this project. The hidden layer consists of n_n neurons and, hence, we recognise the input- and output-weight matrices $W^{[1]}, W^{[2]} \in \mathbb{R}^{n_n}_{\geq 0}$ together with input-bias matrix $b^{[1]} \in \mathbb{R}^{n_n}$ in the neural network parametrisation. Thereby, the decision variable space associated with the neural network parametrisation is defined as

$$\Theta_{\varphi\text{-VIII}} = \left\{ \theta \in \mathbb{R}^{3n_n} \mid W^{[1]}(\theta),\ W^{[2]}(\theta) \in \mathbb{R}^{n_n}_{\geq 0} \right\}. \tag{5.12}$$

Let us recall that the sector condition in Definition 2.2 requires $\varphi(0) = 0$, which can be satisfied by constraining the output bias to

$$b^{[2]} = -{W^{[2]}}^{\top} \tanh\left(b^{[1]}\right). \tag{5.13}$$

As a final step, the existence of cone bounded sector constraints on $\varphi_{\text{V-VIII}}$ is investigated. Let us define the partial derivative with respect to its input y as

$$\frac{\partial}{\partial y} \varphi_{\text{V-VIII}}(y(k)) = {W^{[2]}}^{\top} M W^{[1]}, \tag{5.14}$$

in which

$$M = \operatorname{diag}\left(\left(\mathbb{1}_{n_y} - \tanh^2\left(W^{[1]} y(k) + b^{[1]} \right) \right)^{\top} \right). \tag{5.15}$$

The diagonal elements of M are bounded on $[0, 1]$, which allows us to upper-bound the nonlinearity slope by $\gamma = {W^{[2]}}^{\top} W^{[1]}$. This bound is tight for zero input bias and conservative otherwise. In conclusion, $\varphi_{\text{V-VIII}}$ verifies Assumption 2.2 and its cone-bounded sector constraints are characterised by $\Omega = \bar{\Omega} = \gamma$.

All model structure selections are now explained together with their respective parametrisations. The first model structure is linear and hence has limited representation capabilities. Model structure II improves upon this by allowing for a cubic nonlinearity, but GEUC cannot be guaranteed. The structures III and IV with a spline nonlinearity allow for a trade-off, being in a finite (incremental) sector, as well as cubic on a domain near the origin to reflect reality. The expected representation capabilities are even further improved by considering neural network nonlinearities in structures V-VIII. Note that an improved representation capability usually goes hand in hand with an increased number of model parameters, which complicates the system identification procedure. This trade-off between model complexity and representation capability is illustrated in the following, where model estimation results are presented for all model structure selections.

5.1.3 Model estimation

The problem of system identification of discrete-time Lur'e-type models for the Silverbox benchmark is solved for the model structures that were introduced before. Model structure I is linear and can thus be identified via linear system identification techniques. The BLA method that was introduced in Section 4.2 is implemented, but provides us with a non-structured SISO LTI state-space model $(A_{BLA},\ B_{BLA},\ C_{BLA}, D_{BLA})$. Previously, it was explained that the linear block's state-space representation is expected in *controllability canonical form*. Under the condition that the pair $(A_{BLA},\ B_{BLA})$ is controllable, one can always apply a special similarity transformation as per Lemma 2.5 to the BLA, which renders the LTI model into the desired form [63, p. 252]. The required transformation matrix is given by

$$P = \begin{bmatrix} q^\top & (qA_{BLA})^\top & \dots & \left(qA_{BLA}^{n-1}\right)^\top \end{bmatrix}^\top, \tag{5.16}$$

where q denotes the bottom row of C_{BLA}. Consequently, we retrieve an equivalent state-space representation $(\tilde{A}_{BLA},\ \tilde{B}_{BLA},\ \tilde{C}_{BLA},\ \tilde{D}_{BLA})$, which is in controllability canonical form and allows to define the initial values for θ_Σ via the *zero nonlinearity output method* explained in Section 4.2. Via this technique, one initialises a GEUC Lur'e-type model, whose input-output behaviour is equivalent to the BLA. The selected initial nonlinearity parametrisations that guarantee zero nonlinearity output are summarised in Table 5.2.

Table 5.2: Initialisation of nonlinearity parametrisation for different model structures.

ID	I	II	III,IV	V-VIII
initialisation of θ_{φ_i}	$\begin{bmatrix}\end{bmatrix}$	$0_{3\times1}$	$\begin{bmatrix} 0 & \lvert\mathcal{U}_{1\times1}\rvert \end{bmatrix}^\top$	$\begin{bmatrix} \lvert\mathcal{U}_{n_n\times1}\rvert^\top & \mathcal{U}_{n_n\times1}^\top & 0_{n_n\times1}^\top \end{bmatrix}^\top$

Note that the spline width y_* in structures III, IV, as well as the elements of the input weights and bias matrices in structures V-VIII are initialised as (nonnegative) random variables from a uniform distribution, which is denoted by $\mathcal{U}_{i\times j}$. This approach ensures a non-zero initial objective function gradient with respect to the nonlinearity parameters. What remains in the model estimation stage is to solve the system identification problem via the gradient-based method from Section 4.2. To that end, Algorithm 4 is implemented to calculate steady-state model responses of GEUC Lur'e-type candidate models and their respective parameter sensitivity models. In case no stability guarantees are enforced, i.e., for model structures II, III, V and VI, we rely on forward iterating the nonlinear dynamics as demonstrated in Example 4.1, until the solution has converged to the steady-state solution. For unstable models, the cost associated with the their 'steady-state' solution is assumed to be excessively large, whereby such decision variable selections are rejected by the gradient-based solution method. Figure 5.6 shows a typical progression of the gradient-based method.

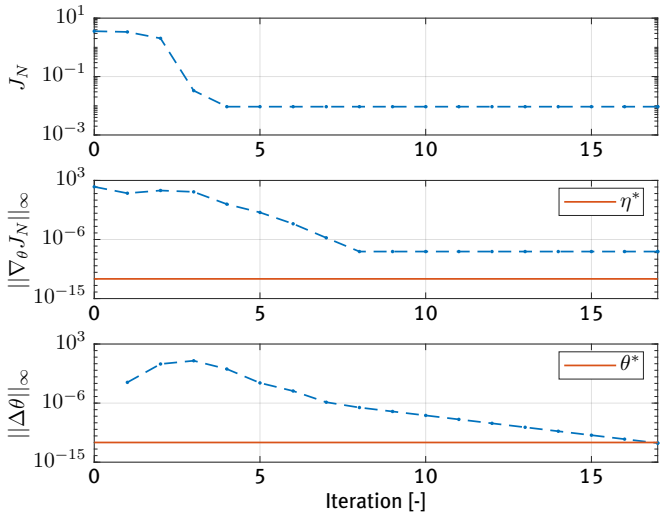
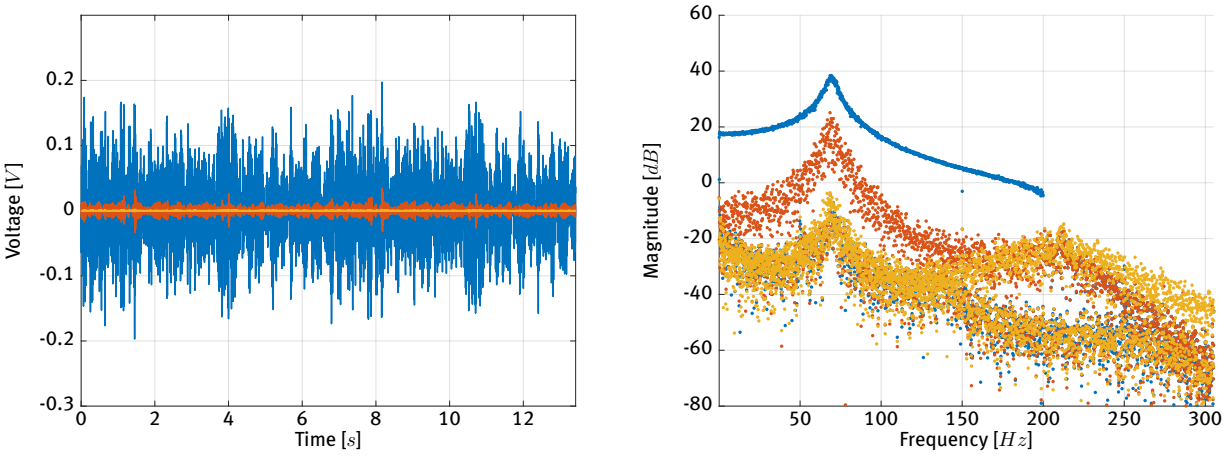


Figure 5.6: Iteration history of the estimation method for model structure II.

In this figure, it can be validated that the objective function value of the initial parametrisation complies to the cost associated with the BLA. Starting from this initial point, the gradient-based optimisation method quickly reduces the objective function value by several orders of magnitude. Also, the first-order-optimality measure, $\|\nabla J_N\|_\infty$, decreases drastically, indicating that the estimation method converges towards a (local) minimum. Simultaneously, the maximum step size $\|\Delta\theta\|_\infty$ shrinks until it passes the minimum step size threshold θ^* in the 17^{th} iteration as can be seen in the bottom plot of Figure 5.6, concluding the model estimation step for structure II. To interpret the results of this model estimation step, the steady-state model output error $\bar{\epsilon}^e$ of model structures I and II are shown in Figure 5.7 for the third multisine realisation.



(a) Time domain.

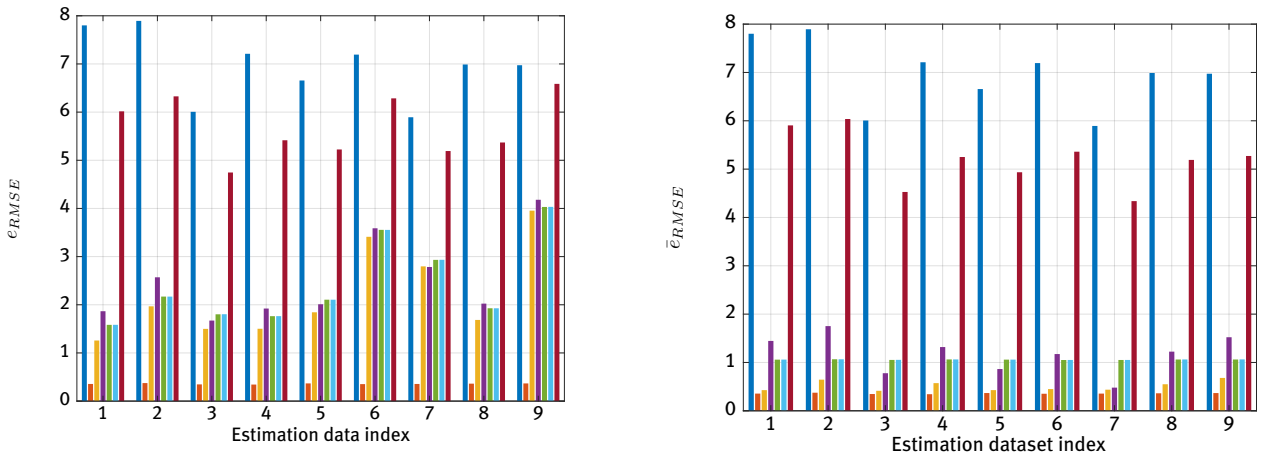
(b) Frequency domain.

Figure 5.7: Estimated model performance on estimation data 3.
The plots include $\bar{z}_0^e$ (——), $\bar{\epsilon}^e$ for (BLA) structure I (——) and $\bar{\epsilon}^e$ for (cubic) structure II (——).

Note that this figure displays steady-state information instead of a simple simulation of the model subject to the multisine excitation sequence. It can be seen that structure I captures the system dynamics relatively well, but is clearly outperformed by structure II. In order to quantitatively compare the performance between different estimated models, a performance measure is defined in terms of Root Mean Squared Error (RMSE), which (depending on a steady-state assumption for the considered measurements) reads as

$$e_{\text{RMSE}} = \sqrt{\frac{1}{N}\sum_{k=1}^{N}\left(z_0^e(k) - \varkappa(k)\right)^2}, \qquad (5.17) \qquad \bar{e}_{\text{RMSE}} = \sqrt{\frac{1}{N}\sum_{k=1}^{N}\left(\bar{z}_0^e(k) - \bar{\varkappa}(k)\right)^2}. \qquad (5.18)$$

Numerous system identification approaches for this benchmark dataset adopt this performance measure (see, e.g., Table 5.3). In case Equation (5.17) is used, the model response is computed via forward iteration of the model dynamics throughout one excitation signal period. Contrarily, Equation (5.18) requires the steady-state model response to be computed via, e.g., the MTF algorithm. Both performance measures are presented for model structures I-VII in Figure 5.8 and estimation datasets 1-9. In this figure, it can clearly be seen that the performance for the different model structures can be ranked from best to worst in the order II-III-V-VI-IV-VII-I. Note that the model estimation step for structure VIII was unsuccessful and, therefore, no results are presented. The cubic structure II is the clear winner, which is in line with expectations due to our prior knowledge on the Silverbox system. It can be seen that performance of models with stability guarantees (IV and VII) deteriorates significantly with respect to their unconstrained equivalents (III and V). Also, the performance measure shows much more consistency throughout different estimation datasets for $\bar{e}_{\text{RMSE}}$ in Figure 5.8b, compared to e_{RMSE} in Figure 5.8a, which suffers from an unmodelled initial condition to the internal state. The model structure color coding is preserved throughout the remainder of this section.



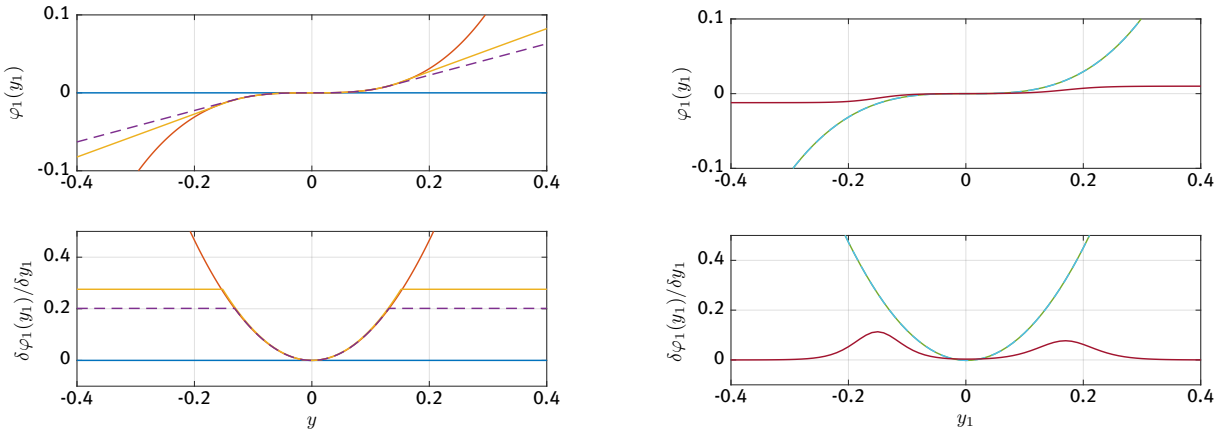
(a) e_{RMSE}.

(b) $\bar{e}_{\text{RMSE}}$.

Figure 5.8: Estimation data performance of I (——), II (——), III (——), IV (——), V (——), VI (——), VII (——).

To interpret the characteristics of the estimated models even further, the nonlinearity graph and the Bode magnitude diagram of the involved LTI block's transfer function are compared in Figures 5.9 and 5.10, respectively. Note that the estimates, even though being structurally different, show many similarities.

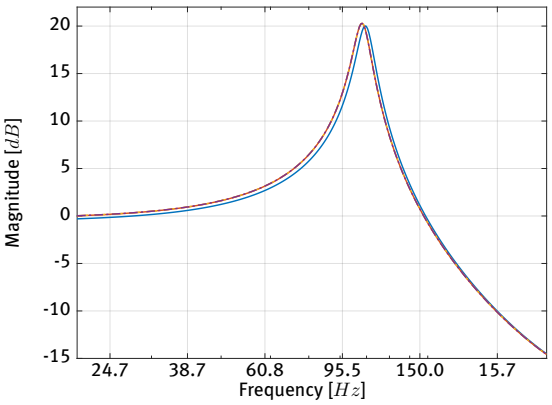
Let us start by discussing the nonlinearities in Figure 5.9a. Structure I has a trivial zero nonlinearity and needs no further clarification. For structure II, one clearly recognises the cubic polynomial shape that is expected to perform well on this dataset by the prior knowledge on the Silverbox system. Furthermore, it is well received that spline nonlinearities III, IV behave exactly like the cubic polynomial of II inside their spline widths. Outside of this domain, the spline structures in III and IV behave linear to allow for a finite bound on their (incremental) sector. The spline width of III is slightly larger than that of IV, which can be explained by the additional stability constraints on IV. Continuing to Figure 5.9b, it can be seen that also the neural networks without stability guarantees (V and VI) have fitted the shape of the cubic polynomial within the input region of interest. Because there is no prior knowledge on the nonlinearity design of the Silverbox system included in these model structures, this is an outstanding performance by the identification algorithm.



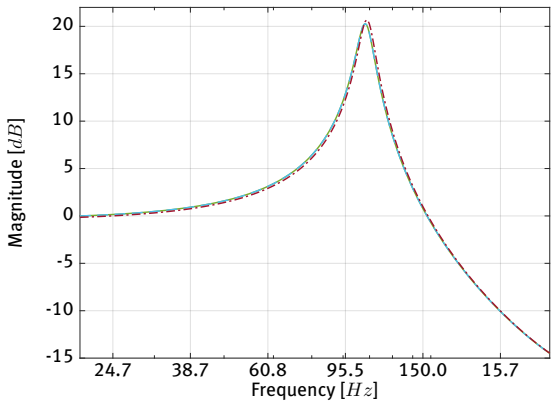
(a) I: (——), II: (——), III: (——), IV: (- - -)

(b) V: (——), VI: (- - -), VII: (——)

Figure 5.9: Estimated nonlinearity graph for various model structures.



(a) I: (——), II: (——), III: (- - -), IV: (-·-·-)



(b) V: (——), VI: (- - -), VII: (-·-·-)

Figure 5.10: Estimated linear block transfer function for various model structures.

It can be concluded from Figure 5.10 that the characteristics of the estimated linear blocks are similar to those of the BLA initialisation. The figure shows Bode magnitude diagrams of the estimated models in close proximity to the one of structure I. This marks the end of the discussion on the estimated models, which are subjected to a validation test next.

5.1.4 Model validation

Does the estimated model actually solve our system identification problem and/or is it in conflict with either the validation data or prior knowledge? These questions are answered in the model validation step [12]. The validation test is defined as forward iterating the estimated model dynamics throughout one period of the preprocessed arrow-'head' excitation as was introduced in Figure 5.4 and comparing the simulation results to observed data. The estimated model performances can then be compared to state-of-the-art results on this benchmark dataset, for which an overview is given in Table 5.3. In this table, (i) indicates the simulation RMSE, defined as e_{RMSE} in (5.17), on the validation set (1-40585). Furthermore, (ii) and (iii) denote the validation test on set (1-40000) and a special validation set (iii) that is mentioned in the table, respectively. The results presented in this work are written in blue and empty cells indicate unavailability of the information.

Table 5.3: Model performance in terms of validation RMSE, compared to the state of the art [59].

Author	Method	n_θ	(i)	(ii)	(iii)	Validation (iii)	Estimation set
Ljung et al. (2004)	Neural network + cubic regr.	> 700			0.3	1 - 40495	40586 - 127410
Hjalmarsson (2004)	Physical block-oriented	5		0.96			40001 - Body
Verdult (2004)	Weighted local linear SS	16			1.3	1 - 40495	40585 - 49192
Sragner et al. (2004)	Special MLP	100		7.8			40001 - 125000
Espinoza et al. (2004)	FS-PLS	500		0.318			40001 - 85000
Espinoza Tapia (2006)	PL-LSSVM			0.271			40001 - 85000
Paduart et al. (2010)	PNLSS	37			0.26	1 - 40700	40701 - Body
Marconato et al. (2012)	Nonlinear state-space	23		0.34			40001 - 80000
Mulders et al. (2013)	Poly-LFR	12			0.35	Head	Body
Mattos et al. (2017)	RVFL-NLMM			10			40001 - End
Santos et al. (2018)	FS-LSSVR			0.76			40001 - 85000
Maroli (2019)	TCN (fair)	4753	2.18				40586 - 85000
Maroli (2019)	TCN	4753		1.74			40001 - 127500
This work (2021)	I: BLA	5	13.7	13.5			40586 - 118314
This work (2021)	II: Cubic	8	0.55	0.54			40586 - 118314
This work (2021)	III: Spline	7	2.78	2.79			40586 - 118314
This work (2021)	IV: Stable spline	7	4.87	4.85			40586 - 118314
This work (2021)	V: NN (3 neurons, no stab.)	14	1.14	1.12			40586 - 118314
This work (2021)	VI: NN (10 neurons, no stab.)	35	1.15	1.13			40586 - 118314
This work (2021)	VII: NN (3 neurons, stab.)	14	11.5	11.3			40586 - 118314

It is only possible to compare our estimated models to results in literature quantitatively when there is consensus on the performance measure, as well as the considered validation dataset. Our intention in this work is to follow the guidelines for fair benchmarking on the Silverbox dataset in [59] and validate on the data in set (i), which allows to quantify the model extrapolation performance. Unfortunately, [59] is the only work that is known to the author that validates on the exact same set. The model performances in [59] are easily outperformed by structures II,V and VI, despite them being parametrised by an excessive number of 4753 parameters. In order to compare our results to more works in literature, the validation test results on set (ii) are added. Clearly, the BLA in structure I can no longer catch up with the state-of-the-art nonlinear models. Also, the model structures IV and VII cannot compete with the the best results in literature in terms of extrapolation performance, but these models come with stability guarantees on the identified model. Arguably, these benefits render them (especially IV) in the sweet spot of being relatively accurate and applicable to reliably predict on input sequences that were not seen during model training. Our high performance structures II, V and VI are in close proximity to the best performing models that are reported in literature, non of which come with some guaranteed form of stability. Note that our models typically require significantly less parameters compared to their competition.

Next to these quantitative results for the model validation step of our system identification outcomes, a qualitative understanding of the model behaviours in the validation test is desired. To that end, the simulation results of structures I-IV are compared in Figures 5.11 and 5.12 in time and frequency domain, respectively. The time-domain results shown some initial error due to an unmodelled initial condition for the internal state. Furthermore, it can be seen that the cubic structure II is a good representation of the observations. Even in the extrapolation regime, which starts roughly after 50 seconds, only a slight increase in model error is observed. The behaviour of the spline models III, IV in the model validation test is perfectly in line with expectations. For small excitation magnitudes, their nonlinearities behave in the cubic regime and the models match the performance of structure I. For larger excitation magnitudes, the nonlinearity acts in its linear regime, whereby it differs from the prior knowledge on the Silverbox system, inducing model performance loss. This effect is greater for structure IV, since a smaller spline width is selected for this model to verify its stability constraint. Finally, it can be seen in Figure 5.12 that the nonlinear model errors are concentrated around the resonance frequency and close to DC.

The qualitative analysis on the model validation step continues in Figures 5.13 and 5.14, in which validation data simulation results of the neural network structures V-VII are shown in time and frequency domain, respectively. The BLA results for structure I are repeated to calibrate our observations with respect to Figures 5.11 and 5.12. The results for structures V and VI show much similarity despite their difference in hidden layer dimension. Note that these simulation results do not suffer significantly from the extrapolation effects that were seen for the cubic and spline nonlinearity structures. Finally, the stable neural network structure VII appears barely as an improvement with respect to a linear model. It can be seen that the stability constraint deteriorated the model performance significantly, which can partly be explained by the conservative upper-bound on the nonlinearity slope that is used to characterise the cone-bounded sector constraints. A tighter upper-bound enlarges the feasible region in decision variable space, whereby a better (local) minimum may become available.

Hereby, the extensive Silverbox case study comes to an end. Prior knowledge on the true system dynamics and the identification experiment is explained. Consecutively, multiple model structures are proposed in an attempt to model the Silverbox system in the class of discrete-time GEUC Lur'e-type models. It was decided to investigate a cubic nonlinearity structure, which fits the prior knowledge well. In order to allow for stability guarantees, a spline nonlinearity structure is proposed, followed by a variety of single hidden layer feedforward neural networks. The identification methodology of Chapter 4 is applied and estimation, as well as extrapolation results are motivated both quantitatively and qualitatively. Also, the estimated models are compared to state-of-the-art models in literature. In conclusion, neural networks and cubic nonlinearities without stability guarantees accurately model the Silverbox system dynamics having their performance measures close to the best results in literature while using significantly less model parameters. Finally, the GEUC spline nonlinearity structure is shown to be a significant improvement with respect to the (linear) models in literature with stability guarantees.

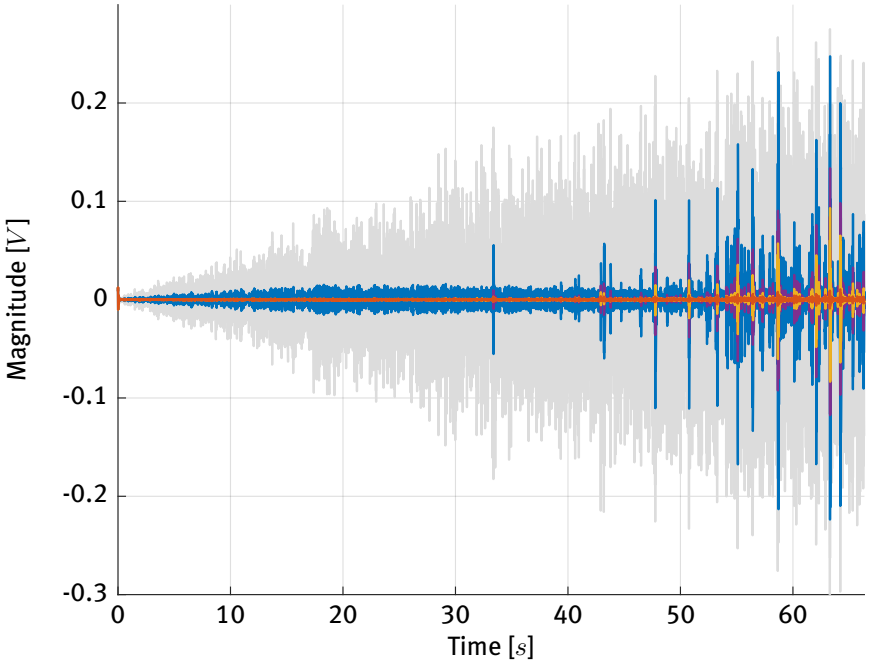


Figure 5.11: Validation test results ϵ^e in time domain for structures I: (—), II: (—), III: (—) IV: (—).

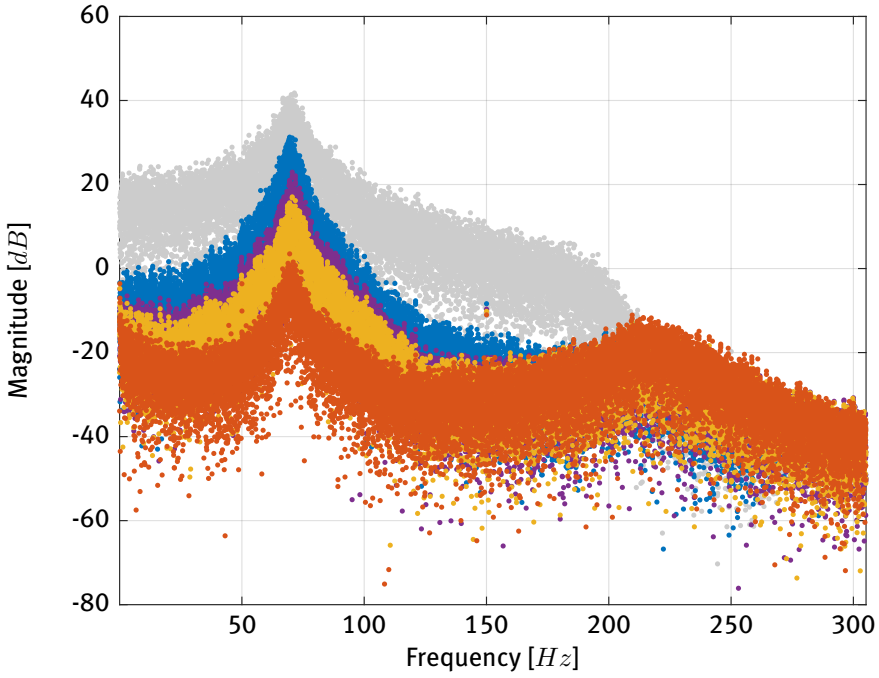


Figure 5.12: Validation test results ϵ^e in frequency domain for structures I: (—), II: (—), III: (—) IV: (—).

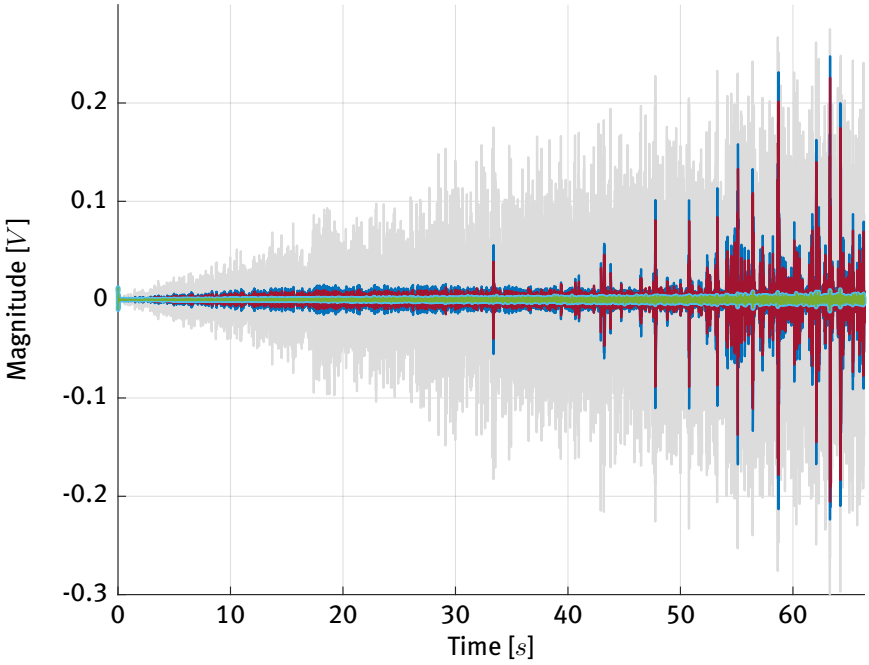


Figure 5.13: Validation test results ϵ^e in time domain for structures I: (—), V: (—), VI: (—) VII: (—).

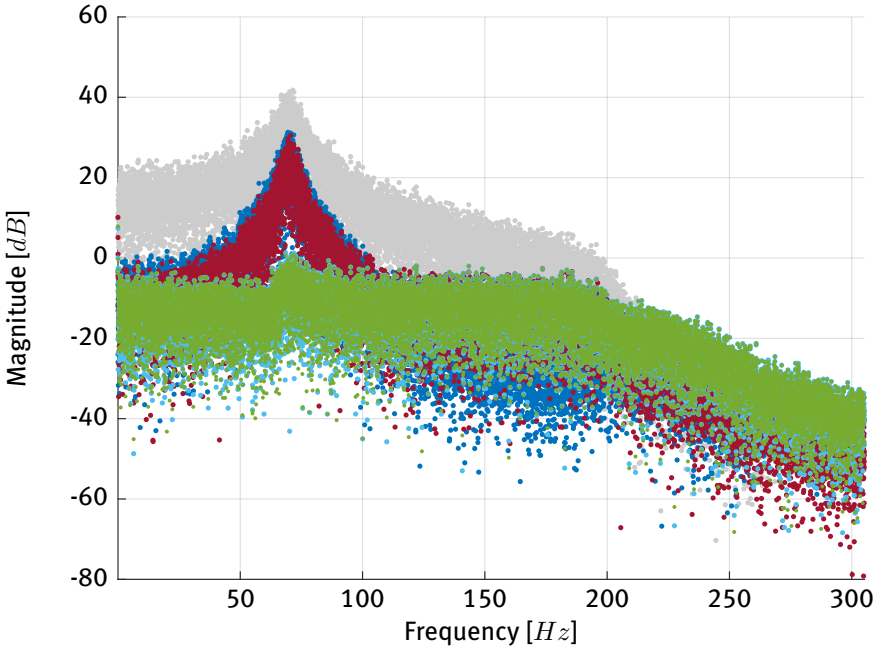


Figure 5.14: Validation test results ϵ^e in frequency domain for structures I: (—), V: (—), VI: (—) VII: (—).

5.2 Simulation study: First order dynamics

In this simulation case study, a system identification problem is solved for the purpose of estimating a first-order Lur'e-type model, which experiences feedback from a smooth deadzone-type nonlinearity. The true system is known to the user, such that steady-state measurements of excitation w and the true model response can be computed. This simulated 'observed data' serves as an input to the proposed system identification approach with the intent to identify the parameters describing the true system.

First, the selected first-order model structure is presented. Thereafter, the true system design is motivated together with the simulated identification experiment. Then, the system identification problem is solved in the model estimation step. The identified model leads to new insights in the parametrisation of the selected model structure, resulting into model structure improvements.

First-order model structure selection

In this simulation case study, candidate models $(\Sigma^\theta,\ \varphi^\theta)$ according to (4.2) with dimensions $n_x = n_y = n_u = n_z = n_w = 1$ are considered. A black-box modeling approach is adopted, such that each matrix element of the LTI state-space form represents a model parameter. Also, a smooth-deadzone type nonlinearity was selected, which can describe real-world phenomena such as worn steering wheel linkages in cars, or coulomb frictions in DC motors. This nonlinearity defines φ^θ according to

$$\varphi^\theta(y(k)) = y(k) - \delta(\theta)\tanh\big(\frac{1}{\delta}y(k)\big), \tag{5.19}$$

which is parametrised by the scalar deadzone length $\delta(\theta)$. All parameters in the selected model structure are collected in

$$\theta = \begin{bmatrix} A(\theta) & B(\theta) & L(\theta) & C(\theta) & D(\theta) & F(\theta) & G(\theta) & H(\theta) & \delta(\theta) \end{bmatrix}^\top \in \Theta_F \subset \mathbb{R}^9, \tag{5.20}$$

in which decision variable selections are constrained to the set

$$\Theta_F = \Big\{\theta \in \mathbb{R}^9 \mid (A(\theta),\ [B(\theta)\ L(\theta)]) \text{ controllable},\ (A(\theta),\ [C(\theta)^\top\ F(\theta)^\top]^\top) \text{ observable},\ \delta(\theta) \geq 0\Big\}. \tag{5.21}$$

Controllability and observability of the pairs $(A(\theta),\ [B(\theta)\ L(\theta)])$ and $(A(\theta),\ [C(\theta)^\top\ F(\theta)^\top]^\top)$, respectively, implies the minimality of Σ^θ's state-space representation and, hence, Assumption 2.1 is verified. Furthermore, nonlinearities with nonnegative deadzone lengths are considered, such that φ^θ is static, decentralised and respects the cone bounded (incremental) sector constraints that are characterised by $\Omega = \bar{\Omega} = 1$. These properties together imply the satisfaction of Assumption 2.2. Given that the assumptions on the linear and nonlinear block are verified, the decision variable space Θ can now be defined as

$$\Theta = \Big\{\theta \in \Theta_F \mid \exists\,(P \in \mathbb{R}_{>0},\ S \in \mathbb{R}_{>0}) : \text{Condition (3.16) in Theorem 3.7 holds true}\Big\}, \tag{5.22}$$

guaranteeing the GEUC property for models inside this model class, i.e., $\theta \in \Theta$. This concludes the discussion on the selected model structure in this simulation case study. The identification experiment design is explained next.

Identification experiment

In simulation case studies, the true system dynamics are known to the user. In this example, it was decided to design the true system dynamics inside of the proposed first-order model structure selection. More specifically, a parametrisation $\theta_0 \in \Theta$ was chosen, defining the true system dynamics $(\Sigma_0,\ \varphi_0)$ as (4.4) with

$$\begin{aligned} A(\theta_0) &= 0.5, & B(\theta_0) &= 0.78, & L(\theta_0) &= 0.3, \\ C(\theta_0) &= 0.35, & & & D(\theta_0) &= 0.93, \\ F(\theta_0) &= 1, & G(\theta_0) &= 0.62, & H(\theta_0) &= 2, \\ & & \delta(\theta_0) &= 1. & & \end{aligned}$$

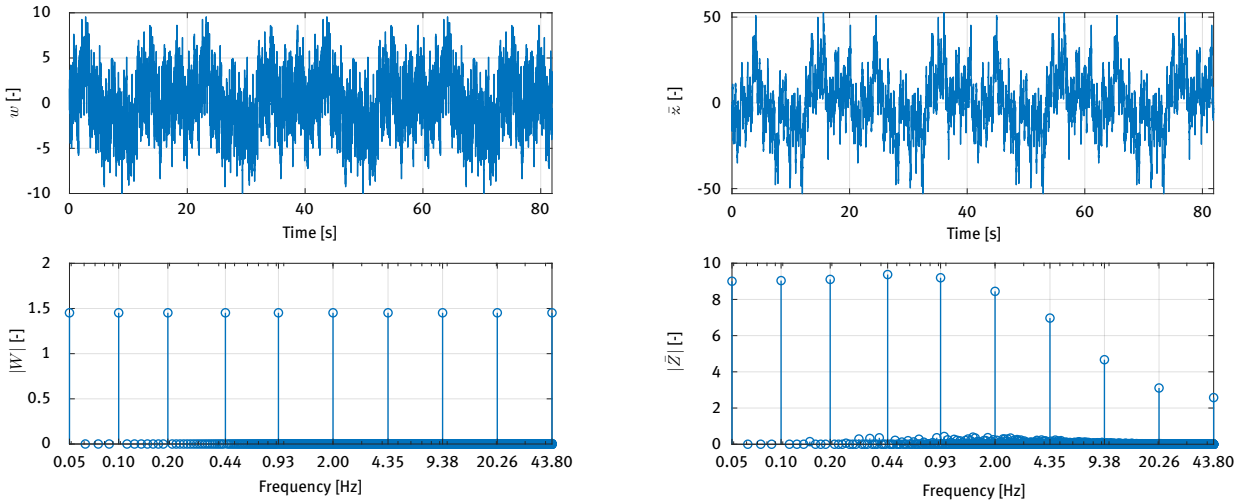
The true-system parametrisation θ_0 indeed represents a GEUC model, since the positive scalars $P = 9.42$ and $S = 9.11$ yield feasibility of the condition (3.16) in Theorem 3.7.

The next step is to design an excitation signal that covers the system operation domain of interest and brings out all essential system features of interest [12]. Important characteristics of the true system are displayed in Figure 5.16. It is common practice to apply multisine excitation signals for the purpose of system identification. Particularly, a random-phase quasi-logarithmic multisine excitation design with constant amplitude distribution is generated via the Frequency Domain System Identification Toolbox (FDIDENT) toolbox (see, e.g., [95]) with the following settings.

Table 5.4: Quasi-logarithmic multisine excitation settings.

	T_s	N	N_{per}	Range	Trials	N_f	$f_{\min}$	$f_{\max}$
w	0.01 [s]	2048	4	$[-10,\ 10]$	50	10	0.004 [Hz]	50 [Hz]

The excitation signal design retrieved from the toolbox is analysed in time- and frequency domain in Figure 5.15a. The observed characteristics comply to the settings in Table 5.4, since ten harmonic contributions can be recognised that appear equidistant on a quasi-logarithmic grid between 0.004 Hz and 50 Hz with an even amplitude distribution. Also, the excitation amplitude varies in time-domain between -10 and 10 as was intended. In conclusion, the excitation signal design is deemed successful and Algorithm 4 can be applied to compute the steady-state model response for the dynamics $(\Sigma_0,\ \varphi_0)$. The results are shown in time- and frequency-domain in Figure 5.15b. This completes the identification experiment discussion and the next step is to estimate a model based on the identification experiment and the proposed model class.



(a) Excitation.

(b) Steady-state output.

Figure 5.15: Identification experiment design for the first-order simulation case study.

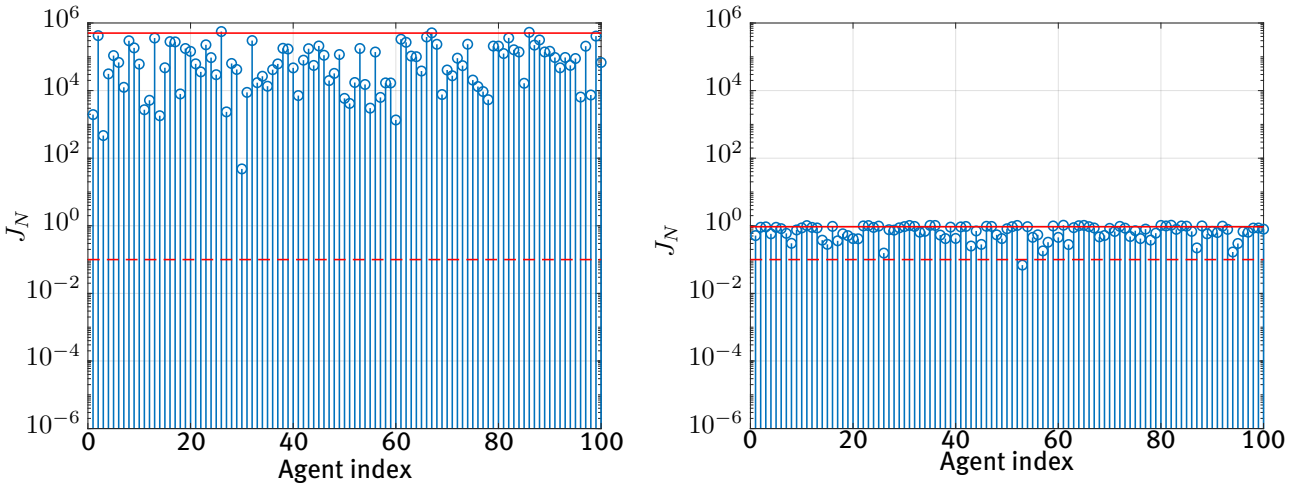
Model estimation

The model estimation step is performed for the simulation case study via a CRS global optimisation approach in the decision variable initialisation step, followed by a gradient-based optimisation routine. The main working principles of the CRS are treated in Section 4.2 and the algorithm was executed with the following settings that are based on [85].

Table 5.5: CRS settings.

n_p	n_m	α	J^*	η^*	k^*
100	9	0.5	0.1	0.1	1000

The progresssion of the CRS algorithm is explained in Figure 5.16. The plot in Figure 5.16a shows the initial population of hundred agents. The agents were randomly selected from Θ_F and, only under feasibility of the LMI stability condition, it is passed as an agent to the population, so typically a large objective function value is associated with these initial agents. The CRS algorithm was performed for $k = 855$ iterations until the final population is found as shown in Figure 5.16b. It can be seen that the difference in objective function value between the agents was greatly reduced and the best agent in the population even crossed the J^* tolerance, which made the algorithm come to a stop and concluded the decision variable initialisation.



(a) Initial population.

(b) Final population after $k = 855$ iterations.

Figure 5.16: The initial and final population of the CRS decision variable initialisation step. The objective function value is shown for each agent in the population. Also, the stopping conditions η^* (——) and J^* (- - -) are shown, which are activated if all agents are above, or when one agent is below, respectively.

The next step is to use the parameters corresponding to the best CRS agent as the initial point for the gradient-based optimisation routine that is explained in Section 4.2. The iteration history of the gradient-based method is shown in Figure 5.17 below.

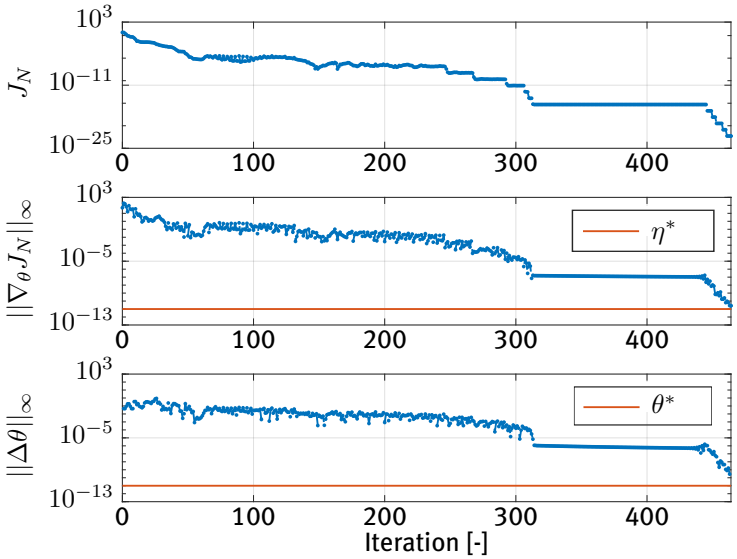


Figure 5.17: Gradient-based optimisation.

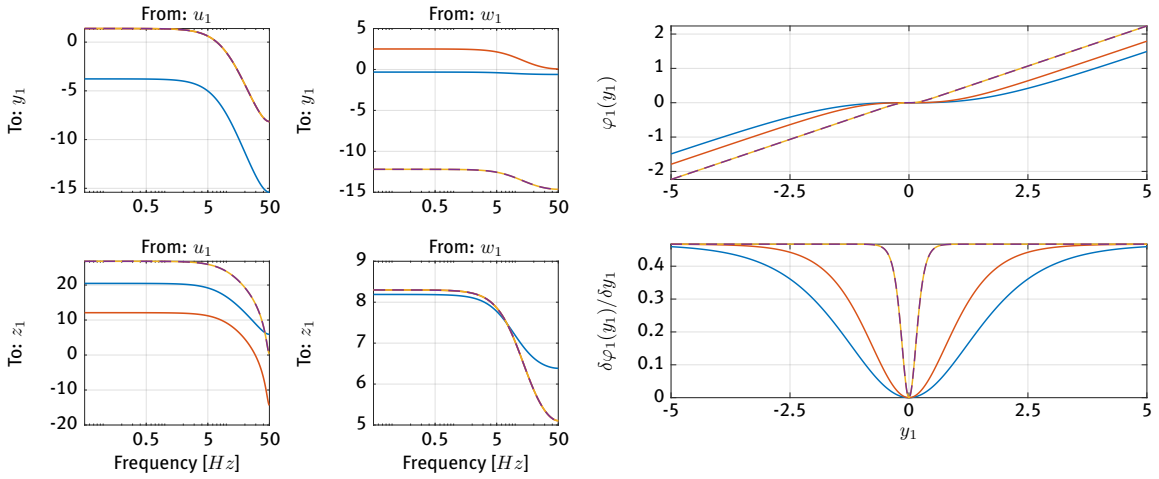
A first observation is that the optimisation can be deemed successful, since the gradient-based method managed to decrease the objective function value up to numerical precision. Also, it can be seen that the largest element of the objective function gradient becomes small in magnitude, indicating that the iterative algorithm converged to a (local) minimum in decision variable space. It should be remarked that the estimation required more than 400 iterations of the gradient-based solution method, which is unfavourable computationally. Also, notice that in most of these iterations (see, e.g, iteration 300-400) barely any progress was made. Hence, there is still room for improvements here.

To analyse the properties of the identified model $\hat{\theta}_N$, it is compared to θ_{init} and θ_0 in terms of Bode magnitude diagrams of the linear block and the nonlinearity graph in Figure 5.18. The results in this figure are surprising, since the estimated model $\hat{\theta}_N$ does not coincide with the characteristics of the true dynamics θ_0 despite explaining the training data perfectly well. Apparently, the model parametrisation contains non-uniqueness other than those induced by similarity transformations (that do not appear in Bode magnitude diagrams). The estimated deadzone length is clearly off and despite correctly identifying the Bode magnitude diagrams for $\mathcal{G}_{\Sigma_{zw}}(\varkappa)$ and $\mathcal{G}_{\Sigma_{yu}}(\varkappa)$, the transfer functions $\mathcal{G}_{\Sigma_{yw}}(\varkappa)$ and $\mathcal{G}_{\Sigma_{zu}}(\varkappa)$ are off by a constant gain. These observation are, possibly, due to parameter non-uniqueness via loop scaling (see Section 2.3).

Indeed, it can be shown that the parametrisation is non-unique due to gain-exchanges between the linear and the nonlinear block. Let us apply a loop-scaling transformation as per Lemma 2.2 to the dynamics $(\Sigma,\ \varphi)$ for $\Psi_2 = \Psi_1 = \gamma$. The resulting Lur'e-type model $(\tilde{\Sigma},\ \tilde{\varphi})$ can then be written as

$$\begin{cases} x_{k+1} = Ax(k) + B\gamma^{-1}\tilde{u}(k) + Lw(k), & \text{(5.23a)} \\ \tilde{y}_k = \gamma^{-1}Cx(k) + \gamma^{-1}Dw(k), & \text{(5.23b)} \\ z_k = Fx(k) + G\gamma^{-1}\tilde{u}_k + Hw(k), & \text{(5.23c)} \\ \tilde{u}_k = \tilde{y}(k) - \gamma^{-1}\delta \tanh\left(\delta^{-1}\gamma\tilde{y}(k)\right). & \text{(5.23d)} \end{cases}$$

Remarkably, the blue terms in Equation (5.23) represent a new deadzone length $\hat{\delta} = \gamma^{-1}\delta$ and, hence, one can select any invertible loop scaling by γ while still generating the same input-output behaviour. It is quite likely that the estimation method exploited many of these equivalent points in decision variable space, which reduced its convergence rate drastically. The result of loop-scaling $\hat{\theta}^N$ towards the deadzone length of θ_0 is shown as $\tilde{\hat{\theta}}_N$ in Figure 5.18 and indeed also ensures a bode magnitude diagram match with θ_0. As a solution, the redundant degree of freedom can be eliminated by fixing the deadzone length in our model class, without inducing any additional conservatism on the model's representation capabilities. This proposal marks the end of the first order system simulation case study.



(a) Bode magnitude diagrams.

(b) Nonlinearity graph.

Figure 5.18: Lur'e-type model characteristics for θ_{init} (—), $\hat{\theta}_N$ (—), θ_0 (—) and $\tilde{\hat{\theta}}_N$ (- - -).

5.3 Summary

The proposed system identification methodology validation in this chapter is twofold. First, an experimental validation was performed on the Silverbox benchmark dataset that is well-known in the research domain of nonlinear system identification. Model structures inside our model class including cubic, spline and feedforward neural networks with or without stability guarantees are proposed. After that, the system identification methodology is deployed and results are explained both quantitatively and qualitatively on estimation and validation datasets. A comparison between our model performances and state-of-the-art Silverbox models in literature is presented. In conclusion, our methodology enables to identify models having cubic or neural network nonlinearities without stability guarantees that approximately match the performance of the best models in literature, albeit being parametrised by fewer parameters. Also, a significant improvement on modeling with stability guarantees is presented in terms of the GEUC spline structure, which easily outperforms stable BLA models. The second system identification method validation involves a simulation case study of first-order dynamical systems with SISO deadzone nonlinearities. A quasi-logarithmic excitation was designed in order to simulate an identification experiment. The system identification is initialised via a CRS global optimisation routine. The proposed system identification methodology manages to successfully explain all details of the identification experiment data. However, it is shown that this model structure contains unexpected non-unique parametrisations in terms of gain exchanges between the deadzone length and the linear block matrices. This non-uniqueness can be eliminated by fixing the deadzone length before starting the model estimation. In conclusion, the system identification methodology is successfully implemented on both experimental data and a simulation case study. It is a logical next research objective to validate the methodologies on identification of MIMO Lur'e-type models with stability guarantees on the identified model.

Chapter 6

Conclusions and recommendations

In this thesis, a method for system identification of discrete-time state-space nonlinear models with stability guarantees is developed. Existing methods in this research field suffer typically from the computational complexity of solving high-dimensional nonlinear optimisation problems. Also, these approaches generally do not enforce some form of stability on the identified model. This work contributes to the open challenges in the field of nonlinear system identification by aiming to achieve the following main research goal:

"*Develop an experimentally validated, computationally efficient system identification methodology for flexible discrete-time nonlinear state-space models with stability guarantees.*"

These main research objectives have mostly been achieved and the outcomes of the work presented in this thesis are split into conclusions and recommendations for future research.

6.1 Conclusions

This work considers the class of discrete-time multiple-input multiple-output (MIMO) Lur'e-type models. Such models consist of an linear time-invariant (LTI) block having a minimal state-space representation that is placed in feedback with a static, decentralised nonlinearity that verifies cone bounded sector constraints. The model class has extensive representation capabilities, as it includes, but is not limited by many well-known block-oriented model classes such as (parallel) Wiener- and Hammerstein-structures.

Contributions to developments in stability analysis of discrete-time MIMO Lur'e-type models are presented in this thesis. It first extends existing sufficient conditions for the global exponential uniform convergence (GEUC) property to a sufficient condition that is computationally cheap to check. Thereafter, it shows that satisfaction of the sufficient conditions for GEUC also guarantees an upper-bound to the $\mathcal{H}_\infty$-norm of the LTI block's transfer function related to the input-output channels of the MIMO nonlinearity.

Another main contribution of this work involves the development of a mixed-time-frequency (MTF) algorithm for discrete-time MIMO Lur'e-type models in order to compute steady-state responses in a *computationally efficient* manner. This method is inspired by existing works on continuous-time single-input single-output (SISO) Lur'e-type models. It is shown that under the sufficient condition for GEUC, the discrete-time MIMO MTF algorithm globally contracts towards the steady-state model response. Furthermore, it is shown in a numerical case study that a 66 % computation time reduction can be achieved by implementing the MTF algorithm rather than forward iterating the nonlinear dynamics to calculate steady-state model responses.

A *system identification methodology* is developed to estimate models inside the model class. This task is tackled as a constrained optimisation problem with the objective to minimise the simulation error between observations and steady-state model outputs along a measurement interval. Candidate models are constrained to the class of discrete-time MIMO Lur'e-type models that are GEUC by verifying the proposed condition. System identification is performed via a two-stage algorithm. First, a good initial point in decision variable space is selected by existing methods such as best linear approximation (BLA) and controlled random search (CRS) global optimisation. Starting from this initial point, a gradient-based optimisation method is used to solve the system identification problem. This work aims to improve the computational efficiency of such methods. It does so by computing steady-state candidate model responses via the MTF algorithm in order to evaluate the objective function. Also, it is shown that the exact objective function gradient with respect to the decision variables can be evaluated by calculation of steady-state responses via the MTF algorithm for a set of GEUC parameter sensitivity models associated with the candidate model.

The final contributions of this work concern *validation* of the proposed system identification methodology. First, an experimental validation was performed on the Silverbox benchmark dataset that is well-known in the research domain of nonlinear system identification. Estimated models with cubic or neural network nonlinearities and without stability guarantees are presented that approximately match the performance of the best Silverbox models in literature. Also, significant improvements on stable Also, a significant improvement on modeling with stability guarantees is presented in terms of the GEUC spline structure, which easily outperforms stable BLA models. The second validation involves a simulation case study of first-order dynamical systems with SISO deadzone nonlinearities. The identification data can be explained perfectly by our estimated models and unexpected parameter non-uniqueness is shown for the model class in terms of gain exchanges that can easily be eliminated by fixing the deadzone length of the nonlinearity.

6.2 Recommendations

Some good progress was made in this work on nonlinear system identification. However, throughout this project, several possible further improvements were identified as recommendations for future work that are listed below without a specific order.

- **System identification methodology validation:**
  As a short-term extension of the current system identification methodology validation, a tighter upper-bound on the slope of a single hidden layer feedforward neural network is desired. It is expected that such an advancement leads to improved Silverbox models with stability guarantees as the current incremental sector bounds appear to limit the model structure's representation capabilities. An obvious second extension of this thesis involves the validation of our system identification methodology beyond models having SISO nonlinearities. One can think of the challenging F-16 aircraft ground vibration benchmark for nonlinear system identification as a starting point, since prior knowledge on the true system dynamics indicates the existence of multiple saturation-type nonlinearities [96]. Such a system fits our model class well and the benchmark allows for comparison with state-of-the-art modelling results in literature. Also, the simulation study that is presented in this work could be extended to the MIMO nonlinearity case, for example by simulation of mass-spring-damper systems with nonlinear springs and dampers. Such a step would strengthen the confidence of our system identification methodology being able to express complex real-world systems into accurate mathematical models.

- **Model class generalisation:**
  In this thesis, the model class is limited to discrete-time Lur'e-type models consisting of a linear block with a minimal state-space representation that is interconnected via feedback by a static, decentralised MIMO nonlinearity with cone-bounded sector constraints. In case of multiple nonlinearities, the similarity transformation towards a controllability canonical form as used in the Silverbox case study no longer holds. Therefore, a different approach on eliminating similarity transformations of the linear block is to be developed. Also, the assumption of having decentralised nonlinearities limits the model class by excluding, e.g., nonlinearities that involve cross-products of system states. An extension of the model class to nonlinearities with coupled inputs allows us to implement (deep) neural networks in the feedback loop. The Lur'e-type model structure can still be of use in stability analysis, where innovations in the field of artificial intelligence, such as improvements on tight estimation of Lipschitz constants to deep neural networks, can be combined with, e.g., small-gain arguments originating from the field of (robust) control theory. This bridges the gap between two rapidly growing research areas allowing for many more new research steps.

- **Stability condition conservatism reduction:**
  The sufficient conditions for GEUC that are used in this work are conservative in the sense that there exist GEUC systems for which our conditions cannot guarantee the property. Our stability conditions prove the convergence property via two (incremental) Lyapunov functions that are quadratic in the internal system state. As a future research step, it could be investigated whether the conservatism can be reduced by considering an extended Lyapunov function class. Also, it could be investigated whether local stability analysis reduces conservatism on the stability conditions. Many systems only act in specific operating regimes. Therefore, it is sufficient to prove stability for these operation modes and omit the possibility of instabilities at unachievable system states.

Appendices

Appendix A

Proofs

A.1 Proof of Lemma 2.2

Proof.
Let us substitute the equalities $u_k = \Psi_1 \tilde{u}_k$ and $y_k = \Psi_2^{-1} \tilde{y}_k$ into our model class (2.1) such that the updated nonlinearity output- and input, respectively $\tilde{u}$ and $\tilde{y}$, appear explicitly in the model equations

$$\begin{cases} x_{k+1} = Ax(k) + B\Psi_1 \tilde{u}(k) + Lw(k), & \text{(A.1a)} \\ \Psi_2^{-1} \tilde{y}_k = Cx(k) + Dw(k), & \text{(A.1b)} \\ z_k = Fx(k) + G\Psi_1 \tilde{u}_k + Hw(k), & \text{(A.1c)} \\ \Psi_1 \tilde{u}_k = \varphi(\Psi_2^{-1} \tilde{y}(k), k). & \text{(A.1d)} \end{cases}$$

We can rewrite Equations (A.1b) and (A.1d), such that we find the system

$$\tilde{\Sigma}: \begin{cases} x(k+1) = Ax(k) + B\Psi_1 \tilde{u}(k) + Lw(k), \quad x(0) = x_0, & \text{(A.2a)} \\ \tilde{y}(k) = \Psi_2 Cx(k) + \Psi_2 Dw(k), & \text{(A.2b)} \\ z(k) = Fx(k) + G\Psi_1 \tilde{u}(k) + Hw(k), & \text{(A.2c)} \end{cases}$$

$$\tilde{\varphi}: \{ \quad \tilde{u}(k) = \Psi_1^{-1} \varphi(\Psi_2^{-1} \tilde{y}(k), k). \quad \text{(A.2d)}$$

Indeed, this system can be interpreted as a new Lur'e-type structure $(\tilde{\Sigma}, \tilde{\varphi})$ with $\tilde{\Sigma}$ having a state-space representation $(A_{\tilde{\Sigma}}, B_{\tilde{\Sigma}}, C_{\tilde{\Sigma}}, D_{\tilde{\Sigma}})$ according to Equation (2.19) and a nonlinearity $\tilde{\varphi}$ according to Equation (2.20) satisfying the (incremental) sector bounds mentioned in Lemma 2.2. □

A.2 Proof of Lemma 2.3

Proof. (See [97, p. 51]).
Let us substitute the equality $u_k = \tilde{u}_k + \Psi_3 y_k$ into our model class (2.1) such that the updated nonlinearity output $\tilde{u}$ appears explicitly in the model equations

$$\begin{cases} x_{k+1} = Ax(k) + B(\tilde{u}(k) + \Psi_3 y(k)) + Lw(k), & \text{(A.3a)} \\ y_k = Cx(k) + Dw(k), & \text{(A.3b)} \\ z_k = Fx(k) + G(\tilde{u}_k + \Psi_3 y(k)) + Hw(k), & \text{(A.3c)} \\ (\tilde{u}(k) + \Psi_3 y(k)) = \varphi(y(k), k). & \text{(A.3d)} \end{cases}$$

We can substitute Equation (A.3b) into Equations (A.3a) and (A.3c), which reads after rearranging terms as

$$\tilde{\Sigma}: \begin{cases} x(k+1) = (A + B\Psi_3 C)x(k) + B\tilde{u}(k) + (L + B\Psi_3 D)w(k), \quad x(0) = x_0, & \text{(A.4a)} \\ y(k) = Cx(k) + Dw(k), & \text{(A.4b)} \\ z(k) = (F + G\Psi_3 C)x(k) + G\tilde{u}(k) + (H + G\Psi_3 D)w(k), & \text{(A.4c)} \end{cases}$$

$$\tilde{\varphi}: \{ \quad \tilde{u}(k) = \varphi(y(k), k) - \Psi_3 y(k). \quad \text{(A.4d)}$$

Note that the assumption of having no direct feedthrough from u into y simplified the equalities in Equation (A.4). Indeed, this system can be interpreted as a new Lur'e-type structure $(\tilde{\Sigma}, \tilde{\varphi})$ inside our model class with $\tilde{\Sigma}$ having a state-space representation $(A_{\tilde{\Sigma}},\ B_{\tilde{\Sigma}},\ C_{\tilde{\Sigma}},\ D_{\tilde{\Sigma}})$ according to Equation (2.21) and a nonlinearity $\tilde{\varphi}$ according to Equation (2.22) satisfying the (incremental) sector bounds mentioned in Lemma 2.3. □

A.3 Proof of Lemma 2.5

Proof. (See [70, p. 658]).
The transfer matrix of a discrete-time LTI system Σ is denoted by $\mathcal{G}_\Sigma(\varkappa)$ and can as function of its state-space formulation $(A_\Sigma, B_\Sigma, C_\Sigma, D_\Sigma)$ via

$$\mathcal{G}_\Sigma(\varkappa) = C_\Sigma(\varkappa I_{n_x} - A_\Sigma)^{-1} B_\Sigma + D_\Sigma. \quad \text{(A.5)}$$

Moreover, the transfer function associated with the updated state-space representation $(\hat{A}_\Sigma, \hat{B}_\Sigma, \hat{C}_\Sigma, \hat{D}_\Sigma)$, denoted by $\hat{\mathcal{G}}_\Sigma(\varkappa)$, can be expressed as

$$\begin{aligned} \hat{\mathcal{G}}_\Sigma(\varkappa) &= \hat{D}_\Sigma + \hat{C}_\Sigma(\varkappa I_{n_x} - \hat{A}_\Sigma)^{-1}\hat{B}_\Sigma \\ &= D_\Sigma + (C_\Sigma P^{-1})(\varkappa I_{n_x} - PA_\Sigma P^{-1})^{-1}(PB_\Sigma) \\ &= D_\Sigma + C_\Sigma \left[P^{-1}(\varkappa I_{n_x} - PA_\Sigma P^{-1})P\right]^{-1} B_\Sigma \\ &= D_\Sigma + C_\Sigma(\varkappa I_{n_x} - A_\Sigma)^{-1}B_\Sigma \\ &= \mathcal{G}_\Sigma(\varkappa), \end{aligned} \quad \text{(A.6)}$$

Replace $(\hat{A}_\Sigma, \hat{B}_\Sigma, \hat{C}_\Sigma, \hat{D}_\Sigma)$; $(AB)^{-1} = B^{-1}A^{-1}$; *Use* $P^{-1}P = I_{n_x}$

which indeed shows that both state-space representations describe the same transfer function and hence have an identical input-output behaviour. □

A.4 Proof of Theorem 3.5

Proof.
This proof verifies the validity of the assumptions of the sufficient part of Theorem 3.4 and is split in two parts. Part 1 addresses the existence of a compact, positively invariant set for the class of inputs $\mathcal{W}_\gamma$. Consecutively, the global incremental stability property is discussed in part 2.

Part 1: Existence of a compact positively invariant set

The existence of a compact, positively invariant set is guaranteed if the assumptions in Lemma 3.1 are verified. Therefore, this part of the proof aims at showing that solutions to Theorem 3.5 imply satisfaction of the conditions (3.6) for adequate choices of functions $\gamma(\cdot)$, $\alpha_1(\cdot)$, $\alpha_2(\cdot)$ and scalar c.

- Firstly, we consider condition (3.6a) in Lemma 3.1. Recall that V_1 is positive definite by definition. Then, we can apply Rayleigh's inequality as in Lemma B.4, such that

$$\lambda_{\min}(P_1)\|x\|^2 \leq V_1(x) \leq \lambda_{\max}(P_1)\|x\|^2, \tag{A.7}$$

  where $\lambda_{\min}(\cdot)$ and $\lambda_{\max}(\cdot)$ refer to a matrix's smallest and largest eigenvalue respectively. Obviously, this expression indeed satisfies condition (3.6a) for the quadratic $\mathcal{K}_\infty$-functions $\alpha_1(s) = \tilde{\alpha}_1 s^2$ and $\alpha_2(s) = \tilde{\alpha}_2 s^2$, with positive coefficients $\tilde{\alpha}_1 = \lambda_{\min}(P_1)$, $\tilde{\alpha}_2 = \lambda_{\max}(P_1)$ and argument $s = \|x\| \in \mathbb{R}_{>0}$.

- Consecutively, we consider condition (3.6b) in Lemma 3.1. To that end, one can define

$$\eta := \begin{bmatrix} x \\ u \\ w \end{bmatrix}. \tag{A.8}$$

  Pre- and post-multiplication of condition (3.13a) by $\eta^\top$ and η, respectively, yields for all $\eta \neq 0$:

$$\begin{aligned}
&\begin{bmatrix} A^\top \\ B^\top \\ L^\top \end{bmatrix} P_1 \begin{bmatrix} A^\top \\ B^\top \\ L^\top \end{bmatrix}^\top - \begin{bmatrix} \tau_1 P_1 & -C^\top \Omega S_1 & 0 \\ \star & 2S_1 & 0 \\ \star & \star & \tau_2 I_m \end{bmatrix} \prec 0, && \text{pre- and post-multiply with } \eta^\top \text{ and } \eta \\
\implies \quad &\eta^\top \begin{bmatrix} A^\top \\ B^\top \\ L^\top \end{bmatrix} P_1 \begin{bmatrix} A^\top \\ B^\top \\ L^\top \end{bmatrix}^\top \eta - \eta^\top \begin{bmatrix} \tau_1 P_1 & -C^\top \Omega S_1 & 0 \\ \star & 2S_1 & 0 \\ \star & \star & \tau_2 I_m \end{bmatrix} \eta < 0, && \text{Use } V_1 \text{ and (2.16)} \\
\implies \quad &V_1\left(f(x,w)\right) - \tau_1 V_1(x) - \tau_2 \|w\|^2 - 2 s_\Omega(S_1, Cx) < 0 && \text{Add (3.13b)} \\
\implies \quad &V_1\left(f(x,w)\right) - c - \tau_1(V_1(x) - c) - \tau_2(\|w\|^2 - \sigma) - 2 s_\Omega(S_1, Cx) < 0 && s_\Omega \leq 0, \|w\|^2 \leq \sigma \\
\implies \quad &V_1\left(f(x,w)\right) - c - \tau_1(V_1(x) - c) < 0 \\
\implies \quad &V_1\left(f(x,w)\right) < c \text{ if } V_1(x) < c \text{ and } \|w\|^2 \leq \sigma.
\end{aligned} \tag{A.9}$$

  This expression indeed verifies condition (3.6b) for the quadratic $\mathcal{K}$-function $\gamma(s) = \frac{1}{\sigma} s^2$ given the positive coefficient σ and argument $s = \|w\| \in \mathbb{R}_{>0}$.

From this analysis, we can conclude that condition (3.6)in Lemma 3.1 holds true and hence, there exists a compact set $\mathcal{S}_c = \{x \in \mathbb{R}^{n_x} \mid V_1(x) \leq c\}$, which is positively invariant according to Definition 3.2 with respect to (i) the system dynamics (??) and (ii) inputs from class $\mathcal{W} = \{w \in \mathbb{R}^{n_w} \mid \|w\|^2 \leq \sigma\}$.

Part 2: Global exponential uniform asymptotic incremental stability

The global exponential uniform asymptotic incremental stability property is guaranteed if the assumptions in Lemma 3.3 are verified for $\mathcal{S} = \mathbb{R}^{m_x}$. This part of the proof shows that the conditions in Theorem 3.5 imply the satisfaction of the conditions (3.9) for the time-independent incremental Lyapunov function $V_2(x^a,\, x^b) = \left\|x^a - x^b\right\|_{P_2}^2$ and adequate choices of functions $\alpha_3(\cdot)$, $\alpha_4(\cdot)$ and $\alpha_5(\cdot)$.

- Firstly, we consider (3.9a) in Lemma 3.3. Recall that V_2 is positive definite by definition. Then, we can apply Rayleigh's inequality as in Lemma B.4, such that

$$\lambda_{\min}(P_2)\left\|x^a - x^b\right\|^2 \leq V_2(x^a,\, x^b) \leq \lambda_{\max}(P_2)\left\|x^a - x^b\right\|^2. \tag{A.10}$$

  Obviously, this expression indeed satisfies condition (3.9a) for the quadratic $\mathcal{K}_\infty$-functions $\alpha_3(s) = \tilde{\alpha}_3 s^2$ and $\alpha_4(s) = \tilde{\alpha}_4 s^2$, with strictly positive coefficients $\tilde{\alpha}_3 = \lambda_{\min}(P_2)$, $\tilde{\alpha}_4 = \lambda_{\max}(P_2)$ and argument $s = \left\|x^a - x^b\right\| \in \mathbb{R}_{>0}$.

- Consecutively, we consider condition (3.9b) in Lemma 3.3. One can define

$$\mu := \begin{bmatrix} x^a - x^b \\ u^a - u^b \end{bmatrix}. \tag{A.11}$$

  Note that μ has a similar role as η in (A.69) with the difference that η depends on the input w, whereas μ is independent of w. One can pre- and post-multiply linear matrix inequality (LMI) (3.13c) by $\mu^\top$ and μ respectively, such that for all $\mu \neq 0$:

$$\begin{aligned}
&\begin{bmatrix} A^\top \\ B^\top \end{bmatrix} P_2 \begin{bmatrix} A^\top \\ B^\top \end{bmatrix}^\top - \begin{bmatrix} P_2 & -C^\top \bar{\Omega} S_2 \\ \star & 2S_2 \end{bmatrix} \prec 0_{n+p+m} && \text{pre-and post-multiply with } \mu^\top \text{ and } \mu \\
\implies \quad &\mu^\top \begin{bmatrix} A^\top \\ B^\top \end{bmatrix} P_2 \begin{bmatrix} A^\top \\ B^\top \end{bmatrix}^\top \mu - \mu^\top \begin{bmatrix} P_2 & -C^\top \bar{\Omega} S_2 \\ \star & 2S_2 \end{bmatrix} \mu < 0 && \text{use } V_2 \text{ and (2.16)} \\
\implies \quad &V_2(f(x^a,\, w),\, f(x^b,\, w)) - V_2(x^a,\, x^b) - 2s_{\bar{\Omega}}(S_1, y^a, y^b) < 0 && s_{\bar{\Omega}} \leq 0 \\
\implies \quad &V_2\left(f(x^a,\, w),\, f(x^b,\, w)\right) - V_2(x^a,\, x^b) < 0 && \text{strict inequality} \\
\implies \quad &V_2\left(f(x^a,\, w),\, f(x^b,\, w)\right) - V_2(x^a,\, x^b) < \tilde{\alpha}_5 \left\|x^a - x^b\right\|.
\end{aligned}$$

  This inequality indeed verifies condition (3.9b) for the quadratic $\mathcal{K}_\infty$-function $\alpha_5(s) = \tilde{\alpha}_5 s^2$, with some positive coefficient $\tilde{\alpha}_5$ and argument $s = \left\|x^a - x^b\right\| \in \mathbb{R}_{>0}$.

Now all assumptions of Lemma 3.3 have been checked for solutions of Theorem 3.5 and hence the system is globally exponentially uniformly asymptotically incrementally stable with respect to dynamics (??).

In conclusion, this proof has shown the existence of a compact, positively invariant set, as well as the global exponential uniform asymptotic incremental stability property. As a result, Theorem 3.4 can be used to show that the system is globally exponentially uniformly convergent with respect to dynamics (??) for the class of inputs $\mathcal{W}_\gamma$. □

A.5 Proof of Lemma 3.4

Proof.

This proof shows that the GEUC property of the dynamics (2.1) for the class of arbitrary bounded inputs is implied by the feasibility of the conditions in Theorem 3.5. It does so by admitting that the incremental Lyapunov function is invariant under absence or presence of feedthrough matrix D and the following steps are taken.

- Consider a Lur'e-type model

$$\begin{cases} x(k+1) = Ax(k) + Bu(k) + Lw(k) & \text{(A.12a)} \\ y(k) = Cx(k), & \text{(A.12b)} \\ u(k) = \varphi(y(k)), & \text{(A.12c)} \end{cases}$$

  and let this model be convergent under the conditions of Theorem 3.5 for the class of bounded inputs with a given bound.

- Under this convergence condition, there exists an incremental Lyapunov function $V = \|x_1 - x_2\|_P^2$ for some $0 \prec P \in \mathbb{S}^{n_x}$, such that

$$\Delta V \leq -\epsilon V, \qquad \forall x_1, x_2 \in \mathbb{R}^{n_x}, \tag{A.13}$$

  in which $\epsilon \in \mathbb{R}_{>0}$ is a strictly positive scalar.

- For obvious reasons, the incremental stability property is invariant under transformations from dynamics (A.12) into dynamics of type (A.12). Therefore, V can also be used as an incremental Lyapunov function for the dynamics (A.12).

- Let us introduce the Lyapunov function candidate $V_2 = \|x\|_P^2$, which can be interpret as a special case of V for $x_1 = x$ and $x_2 = 0$. Then, the incremental Lyapunov function loses its incremental appearance and reduces to $V = \|x\|_P$. The Lyapunov function increment can be written as

$$\begin{aligned} \Delta V &= \|f(x,\ u)\|_P^2 - \|x\|_P^2 \\ &= \|f(x,\ u) - f(0,\ u) + f(0, u)\|_P^2 - \|x - 0\|_P^2 \\ &\leq \|f(x,\ u) - f(0)\|_P^2 + \|f(0,\ u)\|_P^2 - \|x - 0\|_P^2 \\ &\leq -\epsilon V + \|f(0,\ u)\|_P^2. \end{aligned} \tag{A.14}$$

- The red term in the previous inequality is bounded for arbitrary bounded inputs, because of the incremental sector condition on the nonlinearity. This can be seen as follows:

$$\begin{aligned} f(x(k),\ u(k)) &= x(k+1) \\ &= Ax(k) + Bu(k) + Lw(k), \end{aligned} \tag{A.15}$$

  and hence

$$\begin{aligned} f(0,\ u(k)) &= Bu(k) + Lw(k) \\ &= B\varphi(y(k) + Dw(k)) + Lw(k) \\ &= B\varphi(Dw(k)) + Lw(k). \end{aligned} \tag{A.16}$$

  This equality is indeed bounded for bounded excitations $w(k)$ and an incrementally sector bounded nonlinearity φ.

- Then, V_2 is a Lyapunov function that guarantees the existence of a positively invariant set. Together with the global exponential incremental stability property, this in turn implies the GEUC property for the dynamics (2.1) and arbitrary bounded inputs, which closes the proof.

□

A.6 Proof of Lemma 3.5

Proof.

This proof starts by assuming that there exists a set of variables $\{P_1,\ P_2,\ S_1, S_2, \tau_1, \tau_2\}$ such that the conditions in Theorem 3.5 hold true. To show that the feasibility of these conditions is irrespective from hyper-parameter τ_2, we first split the conditions into two independent sets:

- Set I is affected by $\{P_1,\ S_1,\ \tau_1,\ \tau_2\}$) and consists of Inequalities (3.13a) and (3.13b) (.
- Set II is affected by $\{P_2,\ S_2\}$ and consists of Inequality (3.13c).

Consecutively, we analyse the redundancy of τ_2 for feasibility purposes in these two sets of conditions one after the other.

- Next, we show that the feasibility of set I is independent from τ_2.
  - Let us multiply (3.13a) by a positive scalar $\epsilon > 0$, such that

$$\begin{bmatrix} A^\top \\ B^\top \\ F^\top \end{bmatrix} (\epsilon P_1) \begin{bmatrix} A^\top \\ B^\top \\ F^\top \end{bmatrix}^\top - \begin{bmatrix} \tau_1 (\epsilon P_1) & -C^\top \Omega (\epsilon S_1) & 0_{n_x \times n_w} \\ \star & 2(\epsilon S_1) & 0_{n_u \times n_w} \\ \star & \star & (\epsilon \tau_2) I_{n_w} \end{bmatrix} \prec 0_{n_x+n_u+n_w}. \tag{A.17}$$

    From this inequality, we can conclude that (3.13a) is feasible for any variable set: $\{\epsilon P_1,\ \epsilon S_1,\ \tau_1, \epsilon\tau_2\}$. This makes the feasibility of (3.13a) irrespective of hyper-parameter τ_2.
  - Let us analyse Inequality (3.13b) for arbitrary $0 < \tau_1 < 1$ and $\tau_2 > 0$. Admit that one can always select positive scalar variables c and σ, such that Inequality (3.13b) is feasible. Indeed, this observation makes the feasibility of Theorem B.1 independent from hyper-parameter τ_2
- The feasibility of set II is trivially irrespective from τ_2, because these conditions are not parametrised by this hyper-parameter.

We have shown that the feasibility of sets I and II does not depend on τ_2. Thereby, the feasbility of Theorem 3.5 is irrespective from τ_2. □

A.7 Proof of Theorem 3.6

Proof.
This proof verifies the validity of the assumptions of the sufficient part of Theorem 3.4 and is split in two parts. Part 1 addresses the existence of a compact, positively invariant set for the class of bounded inputs. Consecutively, the global exponential incremental stability property is discussed in part 2. Note that there exists a strong parallel between this proof and the proof of Theorem 3.5. For readability purposes, the exact parallels are not repeated.

Part 1: Existence of a compact positively invariant set

The existence of a compact, positively invariant set is guaranteed if the assumptions in Lemma 3.1 are verified. Therefore, this part of the proof aims at showing that solutions to Theorem 3.6 imply satisfaction of the conditions (3.6) for adequate choices of functions $\gamma(\cdot)$, $\alpha_1(\cdot)$, $\alpha_2(\cdot)$ and scalar τ.

- Firstly, we consider condition (3.6a) in Lemma 3.1. This condition can be verified as presented in the proof of Theorem 3.5. There exist functions $\alpha_1(s) = \tilde{\alpha}_1 s^2$ and $\alpha_2(s) = \tilde{\alpha}_2 s^2$ with positive scalars $\tilde{\alpha}_1 := \lambda_{\min}(P_1)$, $\tilde{\alpha}_2 := \lambda_{\max}(P_1)$ and argument $s = \|x\|$. Trivially, these quadratic $\mathcal{K}_\infty$-functions verify condition (3.6a).

- Consecutively, we consider condition (3.6b) in Lemma 3.1. To that end, we introduce a nonzero scalar λ, such that we can apply Lemma B.1. This lemma guarantees $M - \kappa^* N \prec 0$, which for $\kappa^* = \lambda^{-1}$ and choosing the left-hand-side of (3.14a) as M, together with a possibly indefinite choice for matrix N reads

$$\underbrace{\begin{bmatrix} A^\top \\ B^\top \end{bmatrix} P_1 \begin{bmatrix} A^\top \\ B^\top \end{bmatrix}^\top - \begin{bmatrix} \tau P_1 & -C^\top S_1 \Omega \\ \star & 2S_1 \end{bmatrix}}_{M} - \underbrace{\left(-\frac{1}{\lambda}\right)}_{\kappa^*} \underbrace{\begin{bmatrix} A^\top P_1 L \\ B^\top P_1 L \end{bmatrix} \begin{bmatrix} A^\top P_1 L \\ B^\top P_1 L \end{bmatrix}^\top}_{N} \prec 0. \tag{A.18}$$

  We can rewrite this expression into the form

$$\underbrace{\begin{bmatrix} A^\top \\ B^\top \end{bmatrix} P_1 \begin{bmatrix} A^\top \\ B^\top \end{bmatrix}^\top - \begin{bmatrix} \tau P_1 & -C^\top S_1 \Omega \\ \star & 2S_1 \end{bmatrix}}_{\bar{A}} - \underbrace{\begin{bmatrix} A^\top P_1 L \\ B^\top P_1 L \end{bmatrix}}_{\bar{B}} \underbrace{\left(-\frac{1}{\lambda} I_m\right)}_{\bar{D}^{-1}} \underbrace{\begin{bmatrix} A^\top P_1 L \\ B^\top P_1 L \end{bmatrix}^\top}_{C} \prec 0. \tag{A.19}$$

  From this expression, we observe that $\bar{D} \prec 0$ by definition. Therefore, we can apply Lemma B.3 as a necessary and sufficient condition for

$$\begin{bmatrix} \bar{A} & \bar{B} \\ \bar{C} & \bar{D} \end{bmatrix} \prec 0. \tag{A.20}$$

  After rearranging terms and substituting the original variables, this inequality is equivalent to

$$\begin{bmatrix} A^\top \\ B^\top \\ L^\top \end{bmatrix} P_1 \begin{bmatrix} A^\top \\ B^\top \\ L^\top \end{bmatrix}^\top - \begin{bmatrix} \tau P_1 & -C^\top S_1 \Omega & 0_{n\times m} \\ \star & 2S_1 & 0_{p\times m} \\ \star & \star & \lambda I_m + L^\top P_1 L \end{bmatrix} \prec 0. \tag{A.21}$$

  Subsequently, one can define

$$\eta := \begin{bmatrix} x^\top & u^\top & w^\top \end{bmatrix}^\top, \tag{A.22}$$

such that pre- and post-multiplication of Inequality (A.21) by $\eta^\top$ and η, respectively, yields for all $\eta \neq 0$:

$$\begin{aligned}
&\begin{bmatrix} A^\top \\ B^\top \\ L^\top \end{bmatrix} P_1 \begin{bmatrix} A^\top \\ B^\top \\ L^\top \end{bmatrix}^\top - \begin{bmatrix} \tau P_1 & -C^\top S_1 \Omega & 0_{n\times m} \\ \star & 2S_1 & 0_{p\times m} \\ \star & \star & \lambda I_m + L^\top P_1 L \end{bmatrix} \prec 0 && \text{pre-and post-multiply} \\
\implies \quad &\eta^\top \begin{bmatrix} A^\top \\ B^\top \\ L^\top \end{bmatrix} P_1 \begin{bmatrix} A^\top \\ B^\top \\ L^\top \end{bmatrix}^\top \eta - \eta^\top \begin{bmatrix} \tau P_1 & -C^\top S_1 \Omega & 0_{n\times m} \\ \star & 2S_1 & 0_{p\times m} \\ \star & \star & \lambda I_m + L^\top P_1 L \end{bmatrix} \eta < 0 && \text{Use } V_1 \text{ and (2.16)} \\
\implies \quad &V_1\left(f(x,w)\right) - \tau V_1(x) - 2 s_\Omega(S_1, Cx) - \lambda \|w\|_2^2 - \|Lw\|_{P_1}^2 < 0 && s_\Omega \leq 0 \\
\implies \quad &V_1\left(f(x,w)\right) - \tau V_1(x) - \lambda \|w\|_2^2 - \|Lw\|_{P_1}^2 < 0. && \text{(A.23)}
\end{aligned}$$

The next step is to assume that there always exist positive scalars $\sigma,\ c \in \mathbb{R}_{>0}$, such that

$$-c(1-\tau) + \sigma < 0, \tag{A.24}$$

A summation of Inequalities (A.23) and (A.24) can then be written as

$$V_1\left(f(x,w)\right) - c - \tau\left(V_1(x) - c\right) + \underbrace{\left(\sigma - \|Lw\|_{P_1}^2 - \lambda \|w\|_2^2\right)}_{X} < 0. \tag{A.25}$$

One can always guarantee $X \geq 0$ by choosing

$$\sigma \geq \sigma^* := \lambda \|w\|_2^2 + \|Lw\|_{P_1}^2. \tag{A.26}$$

Note that this decision induces the following requirement, which originates from (A.24):

$$c > c^* := \frac{\sigma^*}{1-\tau}. \tag{A.27}$$

Given that $X \geq 0$, this term can be excluded from Inequality (A.25), such that the remaining terms imply

$$V_1\left(f(x,w)\right) < c \mid \left\{V_1(x) \leq c\right\}, \tag{A.28}$$

for the input class

$$\mathcal{W}_\gamma = \left\{ w \in \mathbb{R}^m,\ \|w\|_2^2 \leq \frac{\sigma}{\rho(L^\top P_1 L) + \lambda} \right\}. \tag{A.29}$$

These expressions indeed verify condition (3.6b) for the quadratic $\mathcal{K}$-function $\gamma(s) = \tilde{\gamma} s^2$ with argument $s = \|w\|_2$ and coefficient $\tilde{\gamma} = \frac{\rho(L^\top P_1 L) + \lambda}{\sigma}$. Given the positive scalars $\sigma,\ \lambda$, it holds $\tilde{\gamma} \in \mathbb{R}_{>0}$.

From this analysis, we can conclude that condition (3.6) in Lemma 3.1 holds true and hence, there exists a compact, positively invariant set $\mathcal{S}_c = \left\{ x \in \mathbb{R}^n \mid V_1(x) \leq c,\ \|w\|_2 \leq \gamma^{-1}(1) \right\}$.

Part 2: Global exponential uniform asymptotic incremental stability

The global exponential uniform asymptotic incremental stability property is guaranteed if the assumptions in Lemma 3.3 are verified for $\mathcal{S} = \mathbb{R}^n$. This part of the proof shows that solutions to Theorem 3.6 imply conditions (3.9) for the time-independent incremental Lyapunov function $V_2(x^a,\ x^b) = \left\|x^a - x^b\right\|_{P_2}^2$ and adequate choices of functions $\alpha_3(\cdot)$, $\alpha_4(\cdot)$ and $\alpha_5(\cdot)$. Notice that the LMI (3.14b) is equivalent to (3.13c) in Theorem 3.5. Therefore, part 2 of that theorem's proof can be followed to guarantee that there exists a quadratic $\mathcal{K}_\infty$-function $\alpha_5 = \tilde{\alpha}_5 s^2$ with some positive scalar $\tilde{\alpha}_5$ satisfying condition (3.9) in Lemma 3.3. This makes the dynamics (2.1) globally exponentially uniformly asymptotically incrementally stable.

In conclusion, this proof has shown the existence of a compact, positively invariant set, as well as the global exponential uniform asymptotic incremental stability property. As a result, Theorem 3.4 defines the system as globally exponentially uniformly convergent with respect to dynamics (2.1) for the class of inputs $\mathcal{W}_\gamma$. □

A.8 Proof of Theorem 3.7

Proof.
In this proof, we show that the feasibility of condition (3.16) in Theorem 3.7 implies the inequality (3.14). We write all three conditions on a scalar level and then show that Condition (3.16) is more strict than the conditions (3.14).

- Assume that there exist variables $P \succ 0,\ S \succ 0$ such that inequality (3.16) holds true. Let us now introduce the variable

$$\mu_1 = \begin{bmatrix} x^\top & u^\top \end{bmatrix}^\top. \tag{A.30}$$

Pre- and post-multiplication of Condition (3.16) by $\mu_1^\top$ and μ_1, respectively, yields for all $\mu_1 \neq \underline{0}_{n_x+n_y}$:

$$\begin{aligned}
& \begin{bmatrix} A^\top \\ B^\top \end{bmatrix} P \begin{bmatrix} A^\top \\ B^\top \end{bmatrix}^\top - \begin{bmatrix} \tau^* P & -C^\top \bar{\Omega} S \\ \star & 2S \end{bmatrix} \prec 0_{n_x+n_y} && \text{pre- and post-multiply} \\
\implies & \mu_1^\top \left(\begin{bmatrix} A^\top \\ B^\top \end{bmatrix} P \begin{bmatrix} A^\top \\ B^\top \end{bmatrix}^\top - \begin{bmatrix} \tau^* P & -C^\top \bar{\Omega} S \\ \star & 2S \end{bmatrix} \right) \mu_1 < 0 && \\
\widehat{=}\ & \underbrace{x^\top A^\top P A x + 2x^\top A^\top P B u + u^\top \left(B^\top P B - 2S\right) u}_{:=\alpha} - \tau^* x^\top P x + 2x^\top C^\top S \bar{\Omega} u < 0 && \\
\widehat{=}\ & \alpha - \tau^* x^\top P x + 2x^\top C^\top S \bar{\Omega} u < 0 && \text{(A.31)}
\end{aligned}$$

- Let us propose the same variables as a solution to Condition (3.14b) such that we can substitute $P_2 = P$ and $S_2 = S$. Furthermore, let us also pre-and post-multiply this inequality by $\mu_1^\top$ and μ_1, respectively. Then we find for all $\mu_1 \neq \underline{0}_{n_x+n_y}$:

$$\begin{aligned}
& \begin{bmatrix} A^\top \\ B^\top \end{bmatrix} P_2 \begin{bmatrix} A^\top \\ B^\top \end{bmatrix}^\top - \begin{bmatrix} P_2 & -C^\top \bar{\Omega} S_2 \\ \star & 2S_2 \end{bmatrix} \overset{?}{\prec} 0_{n_x+n_y} && \text{Substitute solution} \\
& \begin{bmatrix} A^\top \\ B^\top \end{bmatrix} P \begin{bmatrix} A^\top \\ B^\top \end{bmatrix}^\top - \begin{bmatrix} P & -C^\top \bar{\Omega} S \\ \star & 2S \end{bmatrix} \overset{?}{\prec} 0_{n_x+n_y} && \text{pre- and post-multiply} \\
\implies & \mu_1^\top \left(\begin{bmatrix} A^\top \\ B^\top \end{bmatrix} P \begin{bmatrix} A^\top \\ B^\top \end{bmatrix}^\top - \begin{bmatrix} P & -C^\top \bar{\Omega} S \\ \star & 2S \end{bmatrix} \right) \mu_1 \overset{?}{<} 0 && \\
\widehat{=}\ & \underbrace{x^\top A^\top P A x + 2x^\top A^\top P B u + u^\top \left(B^\top P B - 2S\right) u}_{\alpha} - x^\top P x + 2x^\top C^\top S \bar{\Omega} u \overset{?}{<} 0 && \\
\widehat{=}\ & \alpha - x^\top P x + 2x^\top C^\top S \bar{\Omega} u \overset{?}{<} 0 && \text{(A.32)}
\end{aligned}$$

- The same reasoning can be applied to Condition (3.14b) such that we find for all $\mu_1 \neq \underline{0}_{n_x+n_y}$:

$$\begin{aligned}
& \begin{bmatrix} A^\top \\ B^\top \end{bmatrix} P_1 \begin{bmatrix} A^\top \\ B^\top \end{bmatrix}^\top - \begin{bmatrix} \tau P_1 & -C^\top \Omega S_1 \\ \star & 2S_1 \end{bmatrix} \overset{?}{\prec} 0_{n_x+n_y} && \text{Substitute solution} \\
& \begin{bmatrix} A^\top \\ B^\top \end{bmatrix} P \begin{bmatrix} A^\top \\ B^\top \end{bmatrix}^\top - \begin{bmatrix} \tau P & -C^\top \Omega S \\ \star & 2S \end{bmatrix} \overset{?}{\prec} 0_{n_x+n_y} && \text{Pre- and post-multiply} \\
\implies & \mu_1^\top \left(\begin{bmatrix} A^\top \\ B^\top \end{bmatrix} P \begin{bmatrix} A^\top \\ B^\top \end{bmatrix}^\top - \begin{bmatrix} \tau P & -C^\top \Omega S \\ \star & 2S \end{bmatrix} \right) \mu_1 \overset{?}{<} 0 && \\
\widehat{=}\ & \underbrace{x^\top A^\top P A x + 2x^\top A^\top P B u + u^\top \left(B^\top P B - 2S\right) u}_{\alpha} - \tau x^\top P x + 2x^\top C^\top S \bar{\Omega} u \overset{?}{<} 0 && \\
\widehat{=}\ & \alpha - \tau x^\top P x + 2x^\top C^\top S \Omega u \overset{?}{<} 0 && \text{(A.33)}
\end{aligned}$$

- let us denote the left-hand-side of an inequality by $\mathrm{LHS}(\cdot)$. To arrive at the desired result, we verify

$$\begin{aligned}
&\mathrm{LHS}(\text{(A.32)}) \overset{?}{\leq} \text{(A.31)} && \\
\widehat{=}\quad & \alpha - x^\top P x + 2x^\top C^\top S\bar{\Omega}u \overset{?}{\leq} \alpha - \tau^* x^\top P x + 2x^\top C^\top S\bar{\Omega}u \ < 0 && \text{Substitute variables} \\
\widehat{=}\quad & (\tau^* - 1)\left\|x\right\|_P^2 \overset{?}{\leq} 0 && \text{Rearrange terms}
\end{aligned} \tag{A.34}$$

  which indeed holds true because (i) $\tau^* < 1$ and (ii) $P \succ 0$. Furthermore, let us verify

$$\begin{aligned}
&\mathrm{LHS}(\text{(A.33)}) \overset{?}{\leq} \text{(A.31)} && \\
\widehat{=}\quad & \alpha - \tau x^\top P x + 2x^\top C^\top S\Omega u \overset{?}{\leq} \alpha - \tau^* x^\top P x + 2x^\top C^\top S\bar{\Omega}u \ < 0 && \text{Substitute variables} \\
\widehat{=}\quad & (\tau^* - \tau)\left\|x\right\|_P^2 + 2u^\top S\left(\Omega - \bar{\Omega}\right)Cx \overset{?}{\leq} 0 && \text{Rearrange terms} \\
\Longrightarrow\quad & (\tau^* - \tau)\left\|x\right\|_P^2 + 2u^\top \underbrace{S\left(\Omega - \bar{\Omega}\right)}_{\mathcal{P}} y \overset{?}{\leq} 0 && y = Cx
\end{aligned} \tag{A.35}$$

  which indeed holds true, because (i) $P \succ 0$, (ii) there always exists a τ such that $\tau^* < \tau < 1$ and, hence, $\tau^* - \tau < 0$, (iii) $\mathcal{P} \preceq 0$ since $\bar{\Omega} \succeq \Omega$ and (iv) u_i and y_i always have equal signs. The third and fourth item in this list follow from the sector condition on φ as explained in Section 2.2.

- As a result, we have shown that the feasibility of condition (3.16) in Theorem 3.7 implies the feasibility of the inequalities (3.14) in Theorem 3.6 for the Lur'e-type dynamics (2.1) with the additional $D = 0_{n_y}$ assumption made in Equation (A.35). As a final step, by Lemma 3.4 we conclude upon global exponential uniform convergence for any class of bounded inputs with a possibly non-zero feedthrough matrix D.

□

A.9 Proof of Lemma 3.6

Proof.
This proof verifies the validity of the assumptions of the sufficient part of Lemma 3.6 and is split in three parts. In part 1, we manipulate the convergence conditions and the Bounded-Real-Lemma condition towards a scalar form. Thereafter, in part 2 we show that the feasibility of the Bounded-Real-Lemma condition for $\tilde{\Sigma}_{\tilde{y}\tilde{u}}$ is implied for $\gamma \geq 2$ to conclude upon an upper-bound to the $\mathcal{H}_\infty$ norm on the dynamics of interest.

Part 1: Convergence conditions manipulation

Assume that there exist variables $P \succ 0,\ S \succ 0$ such that Inequality (3.16) in Theorem 3.7 holds true. In the proof of Theorem 3.7 (see Appendix A.8) we have seen that this assumption implies the feasibility of Condition (3.14b) in Theorem 3.6 for $P_2 = P$ and $S_2 = S$. Let us first define the variable

$$\mu_1 := \begin{bmatrix} x^\top & u^\top \end{bmatrix}^\top. \tag{A.36}$$

Pre- and post-multiplication of Condition (3.14b) by $\mu_1^\top$ and μ_1, respectively, yields for all $\mu_1 \neq \underline{0}_{n+p}$:

$$\begin{aligned}
& \begin{bmatrix} A^\top \\ B^\top \end{bmatrix} P \begin{bmatrix} A^\top \\ B^\top \end{bmatrix}^\top - \begin{bmatrix} P & -C^\top \bar{\Omega} S \\ \star & 2S \end{bmatrix} \prec 0_{n+p} \qquad \left.\begin{array}{l} \\ \\ \end{array}\right\} \text{pre- and post-multiply} \\
\implies \quad & \mu_1^\top \left(\begin{bmatrix} A^\top \\ B^\top \end{bmatrix} P \begin{bmatrix} A^\top \\ B^\top \end{bmatrix}^\top - \begin{bmatrix} P & -C^\top \bar{\Omega} S \\ \star & 2S \end{bmatrix} \right) \mu_1 < 0 \\
\widehat{=} \quad & \underbrace{x^\top \left(A^\top P A - P \right) x + 2x^\top A^\top P B u + u^\top B^\top P B u}_{=:\alpha} - 2u^\top S u + 2x^\top C^\top S \bar{\Omega} u < 0 \\
\widehat{=} \quad & 2u^\top S u - 2x^\top C^\top S \bar{\Omega} u > \alpha
\end{aligned} \tag{A.37}$$

Part 2: Bounded real Lemma manipulation

Recall that our Lur'e-type system is cast into its normalised, symmetric form $(\tilde{\Sigma},\ \tilde{\varphi})$ as per Lemma 2.4. Our focus lies on the LTI dynamics from nonlinearity output $\tilde{u}$ to nonlinearity input $\tilde{y}$, denoted by $\tilde{\Sigma}_{\tilde{u}y}$.

- We have seen before that this system respects the transfer function representation

$$\tilde{\Sigma}_{\tilde{u}\tilde{y}}(\varkappa) = \left[\begin{array}{c|c} A + \frac{1}{2} B \bar{\Omega} C & B \\ \hline \bar{\Omega} C & 0_{n_y} \end{array}\right], \tag{A.38}$$

  such that a state-space formulation of these dynamics reads

$$\begin{cases} x_{k+1} = \underbrace{(A + \frac{1}{2} B \bar{\Omega} C)}_{\bar{A}} x_k + B \tilde{u}_k, & \text{(A.39a)} \\ \tilde{y}_k = \underbrace{\bar{\Omega} C}_{\bar{C}} x_k, & \text{(A.39b)} \end{cases}$$

  Also, recall that the nonlinearity output $\tilde{u}$ of $\tilde{\Sigma}$ is related to the signals in Σ via

$$\tilde{u} = u - \frac{1}{2} \bar{\Omega} y \tag{A.40}$$

- Let us first introduce the variable

$$\mu_2 := \begin{bmatrix} x^\top & \tilde{u}^\top & \tilde{y}^\top \end{bmatrix}^\top, \tag{A.41}$$

which allows us via pre- and post-multiplication by $\mu_2^\top$ and μ_2 to write the Bounded real Lemma LMI (2.12) for *minimal state-space realisations* on a scalar level according to

$$\begin{bmatrix} \bar{A}^\top P \bar{A} - P & \bar{A}^\top P B & \bar{C}^\top \\ * & B^\top P B - \gamma I_{n_y} & 0_{n_y} \\ * & * & -\gamma I_{n_y} \end{bmatrix} \overset{?}{\preceq} 0 \qquad \text{pre- and post-multiply}$$

$$\implies \quad \mu_2^\top \begin{bmatrix} \bar{A}^\top P \bar{A} - P & \bar{A}^\top P B & \bar{C}^\top \\ * & B^\top P B - \gamma I_{n_y} & 0_{n_y} \\ * & * & -\gamma I_{n_y} \end{bmatrix} \mu_2 \overset{?}{\leq} 0$$

$$\widehat{=} \quad \underbrace{x^\top \left(\bar{A}^\top P \bar{A} - P\right) x}_{\text{Term 1}} + \underbrace{2 x^\top \bar{A}^\top P B \tilde{u}}_{\text{Term 2}} + \underbrace{2 \tilde{y}^\top \bar{C} x - \gamma \tilde{y}^\top \tilde{y}}_{\text{Term 3}} + \underbrace{\tilde{u}^\top (B^\top P B - \gamma I_{n_y}) \tilde{u}}_{\text{Term 4}} \overset{?}{\leq} 0. \tag{A.42}$$

For readability of the sequel, this inequality was split into four terms, which will all be manipulated (if necessary) according to the following two steps: Step I: Substitution of the normalised, symmetric transformation to express the inequality into the parametrisation of $(\Sigma,\ \varphi)$. Step II: Substitution of the nonlinearity input equation Equation (A.39b).

Term 1:

$$\begin{aligned} & x^\top \left(\bar{A}^\top P \bar{A} - P\right) x && \text{I} \\ \widehat{=} \quad & x^\top \left(\left(A + \frac{1}{2} B \bar{\Omega} C\right)^\top P \left(A + \frac{1}{2} B \bar{\Omega} C\right) - P \right) x && \text{I} \\ \widehat{=} \quad & x^\top \left(A^\top P A - P\right) x + x^\top A^\top P B \bar{\Omega} C x + \frac{1}{4} x^\top C^\top \bar{\Omega} B^\top P B \bar{\Omega} C x && \text{II} \\ \widehat{=} \quad & x^\top \left(A^\top P A - P\right) x + \underline{x^\top A^\top P B \bar{\Omega} y} + \underline{\frac{1}{4} y^\top \bar{\Omega} B^\top P B \bar{\Omega} y} \end{aligned} \tag{A.43}$$

Term 2:

$$\begin{aligned} & 2 x^\top \bar{A}^\top P \bar{B} \tilde{u} && \text{I} \\ \widehat{=} \quad & 2 x^\top \left(A + \frac{1}{2} B \bar{\Omega} C\right)^\top P B \left(u - \frac{1}{2} \bar{\Omega} y\right) && \text{I} \\ \widehat{=} \quad & 2 x^\top A^\top P B u - x^\top A^\top P B \bar{\Omega} y + x^\top C^\top \bar{\Omega} B^\top P B u - \frac{1}{2} x^\top C^\top \bar{\Omega} B^\top P B \bar{\Omega} y && \text{II} \\ \widehat{=} \quad & 2 x^\top A^\top P B u - \underline{x^\top A^\top P B \bar{\Omega} y} + \underline{y^\top \bar{\Omega} B^\top P B u} - \underline{\frac{1}{2} y^\top \bar{\Omega} B^\top P B \bar{\Omega} y} \end{aligned} \tag{A.44}$$

Term 3:

$$\begin{aligned} & 2 \tilde{y}^\top \bar{C} x - \gamma \tilde{y}^\top \tilde{y} && \text{I} \\ \widehat{=} \quad & 2 y^\top \bar{\Omega} \bar{\Omega} C x - \gamma y^\top \bar{\Omega} \bar{\Omega} y && \text{II} \\ \widehat{=} \quad & (2 - \gamma)\, y^\top \bar{\Omega} \bar{\Omega} y \end{aligned} \tag{A.45}$$

Term 4:

$$\begin{aligned} & \tilde{u}^\top (\bar{B}^\top P \bar{B} - \gamma I_{n_y}) \tilde{u} && \text{I} \\ \widehat{=} \quad & \left(u - \frac{1}{2} \bar{\Omega} y\right)^\top \left(B^\top P B - \gamma I_{n_y}\right) \left(u - \frac{1}{2} \bar{\Omega} y\right) && \text{I} \\ \widehat{=} \quad & u^\top B^\top P B u - \frac{1}{4} \gamma y^\top \bar{\Omega} \bar{\Omega} y - \underline{y^\top \bar{\Omega} B^\top P B u} + \gamma u^\top \bar{\Omega} y + \underline{\frac{1}{4} y^\top \bar{\Omega} B^\top P B \bar{\Omega} y} - \gamma u^\top u \end{aligned} \tag{A.46}$$

Notice that all underlined terms in Equations (A.43) to (A.46) cancel with respect to the other terms with a matching color. Incorporating all these changes, we can write Inequality (A.42) as

$$\underbrace{x^\top \left(A^\top P A - P\right) x + 2x^\top A^\top P B u + u^\top B^\top P B u}_{=\alpha} + \left(2 - \frac{5}{4}\gamma\right) y^\top \bar{\Omega}\bar{\Omega} y + \gamma u^\top \bar{\Omega} y - \gamma u^\top u \overset{?}{\leq} 0 \tag{A.47}$$

- In this expression, we recognize the α-term that we saw before in the convergence conditions manipulation step. We found an upper-bound to α in Inequality (A.37), which we can apply here and hence

$$2u^\top S u - 2x^\top C^\top S \bar{\Omega} u + \left(2 - \frac{5}{4}\gamma\right) y^\top \bar{\Omega}\bar{\Omega} y + \gamma u^\top \bar{\Omega} y - \gamma u^\top u \overset{?}{\leq} 0. \tag{A.48}$$

We can once more substitute Equation (A.39b), such that

$$2u^\top S u - 2y^\top \bar{\Omega} S u + \left(2 - \frac{5}{4}\gamma\right) y^\top \bar{\Omega}\bar{\Omega} y + \gamma u^\top \bar{\Omega} y - \gamma u^\top u \overset{?}{\leq} 0. \tag{A.49}$$

Furthermore, by the cone bounded sector constraints on φ, we know that $u^\top u \leq y^\top \bar{\Omega} u \leq y^\top \bar{\Omega}\bar{\Omega} y$. We can upper-bound the blue term in Inequality (A.49) by this relation to arrive at

$$\underline{2y^\top \bar{\Omega} S u} - \underline{2y^\top \bar{\Omega} S u} + \left(2 - \frac{5}{4}\gamma\right) y^\top \bar{\Omega}\bar{\Omega} y + \gamma u^\top \bar{\Omega} y - \gamma u^\top u \overset{?}{\leq} 0. \tag{A.50}$$

The first two terms cancel out and what remains is

$$\left(2 - \frac{5}{4}\gamma\right) y^\top \bar{\Omega}\bar{\Omega} y + \gamma u^\top \bar{\Omega} y - \gamma u^\top u \overset{?}{\leq} 0. \tag{A.51}$$

Observe that due to the diagonality of $\bar{\Omega}$ we can interpret this inequality as

$$y^\top \begin{bmatrix} (2-\frac{5}{4}\gamma)\bar{\Omega}_1^2 & & \\ & \ddots & \\ & & (2-\frac{5}{4}\gamma)\bar{\Omega}_{n_y}^2 \end{bmatrix} y + u^\top \begin{bmatrix} \gamma\bar{\Omega}_1 & & \\ & \ddots & \\ & & \gamma\bar{\Omega}_{n_y} \end{bmatrix} y - u^\top \begin{bmatrix} \gamma & & \\ & \ddots & \\ & & \gamma \end{bmatrix} u \overset{?}{\leq} 0, \tag{A.52}$$

which can be rewritten into

$$\sum_{i=1}^{n_y} (2 - \frac{5}{4}\gamma)\bar{\Omega}_i^2 y_i^2 + \gamma \bar{\Omega}_i u_i y_i - \gamma u_i^2 \overset{?}{\leq} 0. \tag{A.53}$$

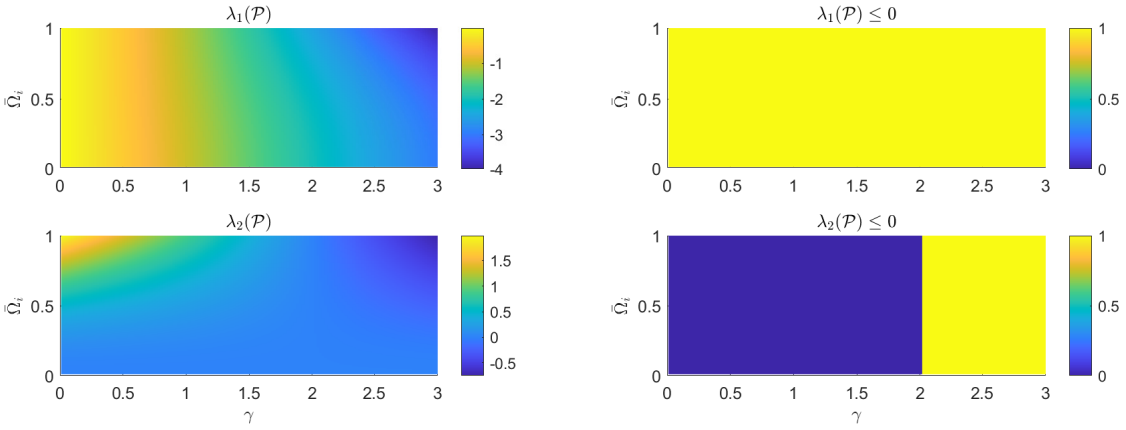
A sufficient condition for Inequality (A.53) to hold, reads

$$(2 - \frac{5}{4}\gamma)\bar{\Omega}_i^2 y_i^2 + \gamma \bar{\Omega}_i u_i y_i - \gamma u_i^2 \leq 0, \qquad \forall i \in \{1, \dots, n_y\}, \tag{A.54}$$

which admits the quadratic form

$$\begin{bmatrix} u_i \\ y_i \end{bmatrix}^\top \underbrace{\begin{bmatrix} -\gamma & \frac{1}{2}\bar{\Omega}_i \\ \star & (2-\frac{5}{4}\gamma)\bar{\Omega}_i^2 \end{bmatrix}}_{\mathcal{P}(\bar{\Omega}_i, \gamma)} \begin{bmatrix} u_i \\ y_i \end{bmatrix} \overset{?}{\leq} 0, \qquad \forall i \in \{1, \dots, n_y\}. \tag{A.55}$$

This inequality holds true whenever $\mathcal{P}$ is negative semi-definite. To assess this property, we analyse the eigenvalues of $\mathcal{P}$ as function of $\bar{\Omega}_i$ and γ in Figure A.1.



(a) Eigenvalues. (b) Nonpositive?

Figure A.1: Eigenvalues of $\mathcal{P}$ on the $(\gamma,\ \bar{\Omega})$ grid.

It can clearly be seen that $\mathcal{P}(\bar{\Omega}_i, \gamma)$ is negative semi-definite for all $\Omega_i, \gamma \in \mathbb{R}_{>0}$ as long as $\gamma \geq 2$. On the boundary, $\mathcal{P}$ is still negative semi-definite, as can be concluded from its non-positive eigenvalues

$$\lambda_1(\mathcal{P}(\bar{\Omega}_i, 2)) = 0, \qquad \lambda_2(\mathcal{P}(\bar{\Omega}_i, 2)) = -2 - \frac{1}{2}\bar{\Omega}_i^2.$$

In conclusion, the feasibility of the non-strict discrete-time bounded-real lemma conditions on $\mathcal{G}_{\Sigma_{\tilde{y}\tilde{u}}}(\varkappa)$ for $\gamma \geq 2$ is implied by the feasibility of the convergence conditions. Therefore, it holds

$$\left\|\mathcal{G}_{\Sigma_{\tilde{y}\tilde{u}}}(\varkappa)\right\|_{\mathcal{H}_\infty} < 2, \tag{A.56}$$

which closes this proof.

□

A.10 Proof of Theorem 4.1

Proof.

Part 1: Derivation of the parameter sensitivity model

The product- and chain rule for differentiation with indirect dependencies are applied to the entire steady-state dynamics of (4.2), such that

$$\begin{cases} \bar{x}_{\theta_i}(k+1) = A_{\theta_i}\bar{x}(k) + A\bar{x}_{\theta_i}(k) + B_{\theta_i}\bar{u}(k) + B\bar{u}_{\theta_i}(k) + L_{\theta_i}w(k) + Lw_{\theta_i}(k), & \text{(A.57a)} \\ \bar{y}_{\theta_i}(k) = C_{\theta_i}\bar{x}(k) + C\bar{x}_{\theta_i}(k) + D_{\theta_i}w(k) + Dw_{\theta_i}(k), & \text{(A.57b)} \\ \bar{z}_{\theta_i}(k) = F_{\theta_i}\bar{x}(k) + F\bar{x}_{\theta_i}(k) + G_{\theta_i}\bar{u}(k) + G\bar{u}_{\theta_i}(k) + H_{\theta_i}w(k) + Hw_{\theta_i}(k), & \text{(A.57c)} \\ \bar{u}_{\theta_i}(k) = \varphi_{\theta_i}(\bar{y}(k)) + \varphi_y\big(\bar{y}(k)\big)\bar{y}_{\theta_i}(k) & \text{(A.57d)} \end{cases}$$

Obviously, $w_\theta = 0$, thanks to the external excitation being independent from the decision variables. Consequently, the blue terms in (A.57) can be omitted. Let us now substitute Equation (A.57d) into Equations (A.57a) and (A.57c), such that after rearranging terms one finds

$$\begin{cases} \bar{x}_{\theta_i}(k+1) = A\bar{x}_{\theta_i}(k) + B\varphi_y\big(\bar{y}(k)\big)\bar{y}_{\theta_i}(k) + A_{\theta_i}\bar{x}(k) + L_{\theta_i}w(k) + B_{\theta_i}\bar{u}(k) + B\varphi_{\theta_i}\big(\bar{y}(k)\big), & \text{(A.58a)} \\ \bar{y}_{\theta_i}(k) = C\bar{x}_{\theta_i}(k) + C_{\theta_i}\bar{x}(k) + D_{\theta_i}w(k), & \text{(A.58b)} \\ \bar{z}_{\theta_i}(k) = F\bar{x}_{\theta_i}(k) + G\varphi_y\big(\bar{y}(k)\big)\bar{y}_{\theta_i}(k) + F_{\theta_i}\bar{x}(k) + H_{\theta_i}w(k) + G_{\theta_i}\bar{u}(k) + G\varphi_{\theta_i}\big(\bar{y}(k)\big), & \text{(A.58c)} \end{cases}$$

Note that the red terms in this expression can be interpret as a nonlinear steady-state input $\breve{\bar{u}}(k) = \varphi^{\theta_i}(\bar{y}_{\theta_i}, \Psi(k))$ associated with the nonlinearity $\varphi^{\theta_i}(\bar{y}_{\theta_i}(k), \Psi(k)) = \Psi(k)\bar{y}_{\theta_i}(k)$ that is subject to a *time-varying* external input $\Psi(k) = \varphi_y(\bar{y}(k))$. By the incremental sector condition on φ, it is known that $\Psi(k) \in [0, \bar{\Omega}]$. Therefore, the nonlinearity φ^{θ_i} is also incrementally sector bounded within $[0, \bar{\Omega}]$. The blue terms depend on the steady-state solution of (Σ, φ) and can, therefore, be interpret as exogenous inputs $\breve{w}_1^{\theta_i}$, $\breve{w}_2^{\theta_i}$, $\breve{w}_3^{\theta_i}$ (see Equation (4.19)) to these steady-state dynamics. Consequently, Equation (A.58) can be rewritten into

$$\begin{cases} \bar{x}_{\theta_i}(k+1) = A\bar{x}_{\theta_i}(k) + B\bar{\tilde{u}} + \breve{w}_1^{\theta_i}(k,\theta), & \text{(A.59a)} \\ \bar{y}_{\theta_i}(k) = C\bar{x}_{\theta_i}(k) + \breve{w}_2^{\theta_i}(k,\theta), & \text{(A.59b)} \\ \bar{z}_{\theta_i}(k) = F\bar{x}_{\theta_i}(k) + G\bar{\tilde{u}} + \breve{w}_3^{\theta_i}(k,\theta), & \text{(A.59c)} \\ \bar{\tilde{u}}_i(k) = \varphi^{\theta_i}(\bar{y}_{\theta_i}(k), \Psi(k)), & \text{(A.59d)} \end{cases}$$

Note that the Equations (A.57) to (A.59) only hold true in steady-state. Moreover, notice that Equation (A.59) describes the steady-state solution of the model

$$\Sigma^{\theta_i}: \begin{cases} x_{\theta_i}(k+1) = Ax_{\theta_i}(k) + B\breve{u}^{\theta_i}(k) + \breve{w}_1^{\theta_i}(k), & \text{(A.60a)} \\ y_{\theta_i}(k) = Cx_{\theta_i}(k) + \breve{w}_2^{\theta_i}(k), & \text{(A.60b)} \\ z_{\theta_i}(k) = Fx_{\theta_i}(k) + G\breve{u}^{\theta_i}(k) + \breve{w}_3^{\theta_i}(k), & \text{(A.60c)} \end{cases}$$

$$\varphi^{\theta_i}: \{ \quad \breve{u}^{\theta_i}(k) = \varphi^{\theta_i}(\bar{y}_{\theta_i}(k), \Psi(k)), \quad \text{(A.60d)}$$

which is equivalent to the parameter sensitivity system Equation (4.17) up to a renaming of variables via Equations (4.18) and (4.20).

Part 2: Argumentation for the properties of parameter sensitivity models

The sensitivity of steady-state model outputs $\bar{z}$ with respect to a decision variable θ_i can thus be described by the steady-state solution of the parameter sensitivity model (A.60). Via Equations (4.14) and (4.15), this sensitivity is directly related to the objective function gradient with respect to the decision variables. Finally, recall that (Σ, φ) is convergent by verifying the condition in Theorem 3.7. This condition is dependent on the variables $\{A, B, C, \bar{\Omega}\}$, and because all these variables are shared between (Σ, φ) and $(\Sigma^{\theta_i}, \varphi^{\theta_i})$, the parameter sensitivity model is convergent as well. □

A.11 Proof of Theorem 4.2

Proof.
Let us consider two arbitrary steady-state nonlinearity input signals $\bar{\tilde{y}}^a_{[i]},\ \bar{\tilde{y}}^b_{[i]} \in \ell_2^{n_y}(N)$ to Lur'e-type model $(\tilde{\Sigma}, \tilde{\varphi})$, defined at arbitrary iteration i of the MTF algorithm. Our approach is to execute one iteration of the MTF algorithm and then compare the 'distance' between $\bar{\tilde{y}}^a$ and $\bar{\tilde{y}}^b$ measured in $\ell_2^{n_y}$-norm before and after this iteration to show the contraction mapping property for $(\hat{\mathcal{F}}_{\tilde{y}\tilde{u}} \circ \hat{\mathcal{F}}_{\tilde{u}\tilde{y}})$.

- To that end, let us apply Equation (4.23) on $\bar{\tilde{y}}^a_{[i]}$ and $\bar{\tilde{y}}^b_{[i]}$ such that for $p \in \{a,\ b\}$, we write

$$\bar{\tilde{y}}^p_{[i+1]} = \left(\hat{\mathcal{F}}_{\tilde{y}\tilde{u}} \circ \hat{\mathcal{F}}_{\tilde{u}\tilde{y}}\right) \circ \bar{\tilde{y}}^p_{[i]} + \hat{\mathcal{F}}_{\tilde{y}w} \circ w, \tag{A.61}$$

  in which we recognise the linear steady-state operators $\hat{\mathcal{F}}_{\tilde{y}\tilde{u}}$ and $\hat{\mathcal{F}}_{\bar{\tilde{y}}\tilde{w}}$ as defined for LTI dynamics in Equation (2.5). Furthermore, we recognise a nonlinear steady-state operator $\hat{\mathcal{F}}_{\tilde{u}\tilde{y}}$, defined as

$$\bar{u} = \hat{\mathcal{F}}_{\tilde{u}\tilde{y}} \circ \bar{\tilde{y}}, \qquad \text{where} \qquad \left(\hat{\mathcal{F}}_{\tilde{u}\tilde{y}} \circ \bar{\tilde{y}}\right)(m) = \tilde{\varphi}(\bar{\tilde{y}}(m)). \tag{A.62}$$

  Consequently, the 'distance' between $\bar{\tilde{y}}^a_{[i+1]}$ and $\bar{\tilde{y}}^b_{[i+1]}$ can be expressed as

$$\left\|\bar{\tilde{y}}^a_{[i+1]} - \bar{\tilde{y}}^b_{[i+1]}\right\|_{\ell_2^{n_y}} = \left\|\left(\hat{\mathcal{F}}_{\tilde{y}\tilde{u}} \circ \hat{\mathcal{F}}_{\tilde{u}\tilde{y}}\right) \circ \bar{\tilde{y}}^a_{[i]} - \left(\hat{\mathcal{F}}_{\tilde{y}\tilde{u}} \circ \hat{\mathcal{F}}_{\tilde{u}\tilde{y}}\right) \circ \bar{\tilde{y}}^b_{[i]}\right\|_{\ell_2^{n_y}} \tag{A.63}$$

- Let us now present the following two instrumental inequalities:
  1. By Definition 2.3 we know that the following inequality holds true for any $\bar{\tilde{u}}_1, \bar{\tilde{u}}_2 \in \ell_2^{n_u}(N)$:

$$\left\|\bar{\tilde{u}}_1 - \bar{\tilde{u}}_2\right\|_{\ell_2^{n_u}} = \left\|\hat{\mathcal{F}}_{\tilde{u}\tilde{y}} \circ \bar{\tilde{y}}_1 - \hat{\mathcal{F}}_{\tilde{u}\tilde{y}} \circ \bar{\tilde{y}}_2\right\|_{\ell_2^{n_u}} \leq K_{\tilde{\varphi}} \left\|\bar{\tilde{y}}_1 - \bar{\tilde{y}}_2\right\|_{\ell_2^{n_y}}. \tag{A.64}$$

     Since $\tilde{\varphi}$ is incrementally sector bounded within $\left[-\frac{1}{2}I_{n_y},\ \frac{1}{2}I_{n_y}\right]$, it is known that $K_{\tilde{\varphi}} = \frac{1}{2}$.
  2. Also, we can use Property 2.2 to derive the following inequality for any $\bar{\tilde{u}}_1, \bar{\tilde{u}}_2 \in \ell_2^{n_u}(N)$:

$$\left\|\bar{\tilde{y}}_1 - \bar{\tilde{y}}_2\right\|_{\ell_2^{n_y}} = \left\|\hat{\mathcal{F}}_{\tilde{y}\tilde{u}} \circ \bar{\tilde{u}}_1 - \hat{\mathcal{F}}_{\tilde{y}\tilde{u}} \circ \bar{\tilde{u}}_2\right\|_{\ell_2^{n_u}} \leq K_{\tilde{\Sigma}} \left\|\bar{\tilde{u}}_1 - \bar{\tilde{u}}_2\right\|_{\ell_2^{n_u}}. \tag{A.65}$$

     By the result of Lemma 3.6, we know that $K_{\tilde{\Sigma}} = \left\|\tilde{\Sigma}_{\bar{\tilde{y}}\bar{\tilde{u}}}(\varkappa)\right\|_{\mathcal{H}_\infty} < 2$.

- Let us combine Inequalities Equations (A.64) and (A.65) such that for any $\bar{\tilde{y}}_1, \bar{\tilde{y}}_2 \in \ell_2^{n_y}(N)$:

$$\left\|\left(\hat{\mathcal{F}}_{\tilde{y}\tilde{u}} \circ \hat{\mathcal{F}}_{\tilde{u}\tilde{y}}\right) \circ \bar{\tilde{y}}_1 - \left(\hat{\mathcal{F}}_{\tilde{y}\tilde{u}} \circ \hat{\mathcal{F}}_{\tilde{u}\tilde{y}}\right) \circ \bar{\tilde{y}}_2\right\|_{\ell_2^{n_y}} \leq K_{\tilde{\varphi}} K_{\tilde{\Sigma}} \left\|\bar{\tilde{y}}_1 - \bar{\tilde{y}}_2\right\|_{\ell_2^{n_y}}. \tag{A.66}$$

  in which we can substitute $\bar{\tilde{y}}_1 = \bar{\tilde{y}}^a_{[i]}$ and $\bar{\tilde{y}}_2 = \bar{\tilde{y}}^b_{[i]}$ to identify the RHS of Equation (A.63) such that we conclude

$$\left\|\bar{\tilde{y}}^a_{[i+1]} - \bar{\tilde{y}}^b_{[i+1]}\right\|_{\ell_2^{n_y}} \leq \underbrace{K_{\tilde{\varphi}} K_{\tilde{\Sigma}}}_{:=\alpha\,<\,1} \left\|\bar{\tilde{y}}^a_{[i]} - \bar{\tilde{y}}^b_{[i]}\right\|_{\ell_2^{n_y}}. \tag{A.67}$$

  This inequality verifies the contraction mapping theorem condition (B.17) in Theorem B.3 for $\alpha < 1$. Indeed, $\hat{\mathcal{F}}_{\tilde{y}\tilde{u}} \circ \hat{\mathcal{F}}_{\tilde{u}\tilde{y}}$ is a contraction mapping having a unique fixed point $\bar{\tilde{y}}$. Finally, verifying that the true steady-state solution $\bar{\tilde{y}}$ of $(\tilde{\Sigma}, \tilde{\varphi})$ is a fixed point of the iterative procedure in (4.23), i.e.,

$$\bar{\tilde{y}} = \left(\hat{\mathcal{F}}_{\tilde{y}\tilde{u}} \circ \hat{\mathcal{F}}_{\tilde{u}\tilde{y}}\right) \circ \bar{\tilde{y}} + \hat{\mathcal{F}}_{\tilde{y}w} \circ w, \tag{A.68}$$

  we conclude that $\bar{\tilde{y}} = \bar{\tilde{y}}$. Hence, the limit $\bar{\tilde{y}}$ of the iterative procedure (4.23) equals the true steady-state solution $\bar{\tilde{y}}$ of $(\tilde{\Sigma}, \tilde{\varphi})$.

□

A.12 Proof of Theorem B.1

Proof.
This proof verifies the validity of the assumptions of the sufficient part of Theorem 3.4 and is split in two parts. Part 1 addresses the existence of a compact, positively invariant set for the class of inputs $\mathcal{W}_\gamma$. Consecutively, the global exponential incremental stability property is discussed in part 2. Note that there exists a strong parallel between this proof and the proof of Theorem 3.5. For readability purposes, the exact parallels are not repeated here.

Part 1: Existence of a compact positively invariant set

The existence of a compact, positively invariant set is guaranteed if the assumptions in Lemma 3.1 are verified. Therefore, this part of the proof aims at showing that the conditions in Theorem B.1 imply satisfaction of the conditions (3.6) for adequate choices of functions $\gamma(\cdot)$, $\alpha_1(\cdot)$, $\alpha_2(\cdot)$ and scalars c, τ_1, τ_2 and σ.

- Firstly, we consider condition (3.6a) in Lemma 3.1. This condition can be verified as presented in the proof of Theorem 3.5. There exist functions $\alpha_1(s) = \tilde{\alpha}_1 s^2$ and $\alpha_2(s) = \tilde{\alpha}_2 s^2$ with positive scalars $\tilde{\alpha}_1 := \lambda_{\min}(P_1)$, $\tilde{\alpha}_2 := \lambda_{\max}(P_1)$ and argument $s = \|x\|_2$. In conclusion, these quadratic $\mathcal{K}_\infty$-functions verify condition (3.6a).
- Consecutively, we consider condition (3.6b) in Lemma 3.1. To that end, one can define

$$\eta := \begin{bmatrix} x^\top & u^\top & w^\top \end{bmatrix}^\top. \tag{A.69}$$

Pre- and post-multiplication of LMI (B.1a) by $\eta^\top$ and η, respectively, yields for all $\eta \neq 0$:

$$\begin{aligned} & \begin{bmatrix} A^\top \\ B^\top \\ F^\top \end{bmatrix} P_1 \begin{bmatrix} A^\top \\ B^\top \\ F^\top \end{bmatrix}^\top - \begin{bmatrix} \tau_1 P_1 & -C^\top \Omega S_1 & 0_{n\times m} \\ \star & 2S_1 & -S_1\Omega D \\ \star & \star & \tau_2 I_m \end{bmatrix} \prec 0_{n+p+m} && \text{pre-and post-multiply} \\ \implies \quad & \eta^\top \begin{bmatrix} A^\top \\ B^\top \\ F^\top \end{bmatrix} P_1 \begin{bmatrix} A^\top \\ B^\top \\ F^\top \end{bmatrix}^\top \eta - \eta^\top \begin{bmatrix} \tau_1 P_1 & -C^\top \Omega S_1 & 0_{n\times m} \\ \star & 2S_1 & -S_1\Omega D \\ \star & \star & \tau_2 I_m \end{bmatrix} \eta < 0 && \text{Use } V_1 \text{ and (2.16)} \\ \implies \quad & V_1\left(f(x,w)\right) - \tau_1 V_1(x) - \tau_2 \|w\|_2^2 - 2 s_\Omega(S_1, Cx + Dw) < 0. && \end{aligned} \tag{A.70}$$

This expression is of the same type as Equation (A.9) in the proof of Theorem 3.5. Therefore, that proof can be followed from this point onwards to verify condition (3.6b) via the quadratic $\mathcal{K}$-function $\gamma(s) = \tilde{\gamma} s^2$ with argument $s = \|w\|_2$ and coefficient $\tilde{\gamma} = \frac{1}{\sigma}$. Given the positive scalar σ, it holds $\tilde{\gamma} \in \mathbb{R}_{>0}$.

From this analysis, we can conclude that condition (3.6) in Lemma 3.1 holds true and hence, there exists a compact, positively invariant set $\mathcal{S}_c = \{x \in \mathbb{R}^{n_x} \mid V_1(x) \leq c\}$ given input from the class $\mathcal{W}_\gamma = \left\{w \in \mathbb{R}^{n_w} \mid \|w\|_2^2 \leq \sigma\right\}$.

Part 2: Global exponential uniform asymptotic incremental stability

The global exponential uniform asymptotic incremental stability property is guaranteed if the assumptions in Lemma 3.3 are verified for $\mathcal{S} = \mathbb{R}^{n_x}$. This part of the proof shows that solutions to Theorem 3.5 imply conditions (3.9) for the time-independent incremental Lyapunov function V_2 and adequate choices of functions $\alpha_3(\cdot)$, $\alpha_4(\cdot)$ and $\alpha_5(\cdot)$. Notice that the LMI (B.1c) is equivalent to (3.13c) in Theorem 3.5. Therefore, part 2 of that theorem's proof can be followed to guarantee that here exists a quadratic $\mathcal{K}_\infty$-function $\alpha_5 = \tilde{\alpha}_5 s^2$ with some positive scalar $\tilde{\alpha}_5$ satisfying condition (3.9) in Lemma 3.3. This makes the dynamics (2.1) globally exponentially uniformly asymptotically incrementally stable.

In conclusion, this proof has shown the existence of a compact, positively invariant set, as well as the global exponential uniform asymptotic incremental stability property. As a result, Theorem 3.4 shows that the system is globally exponentially uniformly convergent with respect to dynamics (2.1) for the class of inputs $\mathcal{W}_\gamma$. □

A.13 Proof of Theorem B.2

Proof.
This proof verifies the validity of the assumptions of the sufficient part of Theorem 3.4 and is split in two parts. Part 1 addresses the existence of a compact, positively invariant set for the class of inputs $\mathcal{W}_\gamma$. Consecutively, the global exponential incremental stability property is discussed in part 2. For readibility purposes we denote the evaluation of a function at time instant k by a subscript k.

Part 1: Existence of a compact positively invariant set

The existence of a compact positively invariant set is guaranteed if the assumptions in Lemma 3.1 are verified. Therefore, this part of the proof aims at showing that solutions to Theorem B.2 imply satisfaction of the conditions (3.6) for adequate choices of functions $\gamma(\cdot)$, $\alpha_1(\cdot)$, $\alpha_2(\cdot)$ and scalar c. To that end, compliance to conditions (3.6a) and (3.6b) is guaranteed one after the other.

- Firstly, we consider condition (3.6a) in Lemma 3.1. A suitable lower bound to V_1 can be thought of as the quadratic $\mathcal{K}_\infty$-function $\alpha_1(r) = \tilde{\alpha}_1 r^2$ for some $\tilde{\alpha}_1 > 0$. This lower-bound exists as can be seen by pre- and post-multiplying inequality (B.7a) by $\xi(x)^\top$ and $\xi(x)$, respectively, which yields for all $\|x\|_2 \neq 0$:

$$
\begin{aligned}
P_1 + G_\Omega(S_0) &\succ 0_{n_x+n_y} && \big\downarrow \textit{Pre- and post-multiply by } \xi(x)^\top \textit{ and } \xi(x)\\
\|\xi(x)\|_{P_1}^2 + \|\xi(x)\|_{G_\Omega(S_0)}^2 &> 0 && \big\downarrow \textit{Substitute (B.6a)}\\
V_1(x) + 2s_\Omega(S_0, Cx) &> 0 && \big\downarrow \textit{Strict inequality}\\
V_1(x) + 2s_\Omega(S_0, Cx) &> \tilde{\alpha}_1 \|x\|_2^2 && \big\downarrow s_\Omega(S,y) \leq 0\\
V_1(x) &\geq \alpha_1(r)
\end{aligned}
$$

  Indeed, this expression satisfies the lower bound in (3.6a) for some positive scalar $\tilde{\alpha}_1$ and argument $r = \|x\|_2$.

  As an intermediate step, observe that the cone bounded sector constraint (2.16) guarantees $\varphi(y) \leq \Omega y$ and can be used as follows:

$$
\begin{aligned}
\|\varphi(y)\|_2^2 &= \varphi(y)^\top \varphi(y) && \big\downarrow \textit{Use } \varphi(y) \leq \Omega y\\
&\leq y^\top \Omega\Omega y && \big\downarrow \textit{Use } y = Cx\\
&= x^\top C^\top \Omega\Omega C x && \big\downarrow \textit{Rayleigh's inequality}\\
&\leq \rho\left(C^\top \Omega\Omega C\right) \|x\|_2^2,
\end{aligned}
\tag{A.71}
$$

  A suitable upper bound to V_1 can be thought of as the quadratic $\mathcal{K}_\infty$-function $\alpha_2(r) = \tilde{\alpha}_2 r^2$ for some $\tilde{\alpha}_2 > 0$. This upper-bound exists as can be seen by applying Rayleigh's inequality, see Lemma B.4, to the quadratic form $V_1(x)$. We have

$$
\begin{aligned}
V_1(x) &= \xi(x)^\top P_1 \xi(x) && \big\downarrow \textit{Rayleigh's inequality}\\
&\leq \rho(P_1)\|\xi(x)\|_2^2 && \big\downarrow \|\xi(x)\|_2^2 = \|x\|_2^2 + \|\varphi(y)\|_2^2\\
&= \rho(P_1)\left(\|x\|_2^2 + \|\varphi(y)\|_2^2\right) && \big\downarrow \textit{Use (A.71)}\\
&\leq \rho(P_1)\left(\|x\|_2^2 + \rho(C^\top \Omega\Omega C)\|x\|_2^2\right)\\
&= \tilde{\alpha}_2 \|x\|_2^2\\
&= \alpha_2(r)
\end{aligned}
$$

  which indeed satisfies the upper bound in (3.6a) for the positive scalar $\tilde{\alpha}_2 := \rho(P_1)(1 + \rho(C^\top \Omega\Omega C)$ and argument $r = \|x\|_2$.

- Consecutively, we consider condition (3.6b) in Lemma 3.1. One can define

$$\begin{pmatrix} \xi_k \\ \xi_{k+1} \\ w_k \end{pmatrix} = M_1 \eta_k, \tag{A.72}$$

where

$$\eta_k := \begin{pmatrix} \xi_k \\ \varphi(CAx_k + B\varphi(Cx_k) + Fw_k) \\ w_k \end{pmatrix}, \tag{A.73}$$

If we now pre- and post-multiply inequality (B.7b) by $\eta_k^\top$ and η_k, respectively, then we find for all $x_k \neq 0$:

$$\begin{aligned} & \eta^\top M_1^\top \begin{bmatrix} -\tau_1 P_1 - G_\Omega(S_1) & 0_{n_x+n_y} & 0_{(n_x+n_y)\times n_w} \\ 0_{n_x+n_y} & P_1 - G_\Omega(S_2) & 0_{(n_x+n_y)\times n_w} \\ 0_{n_w\times(n_x+n_y)} & 0_{n_w\times(n_x+n_y)} & -\tau_2 I_{n_w} \end{bmatrix} M_1 \eta < 0 \\ \Longrightarrow \quad & \begin{pmatrix} \xi_k \\ \xi_{k+1} \\ w_k \end{pmatrix}^\top \begin{bmatrix} -\tau_1 P_1 - G_\Omega(S_1) & 0_{n_x+n_y} & 0_{(n_x+n_y)\times n_w} \\ 0_{n_x+n_y} & P_1 - G_\Omega(S_2) & 0_{(n_x+n_y)\times n_w} \\ 0_{n_w\times(n_x+n_y)} & 0_{n_w\times(n_x+n_y)} & -\tau_2 I_{n_w} \end{bmatrix} \begin{pmatrix} \xi_k \\ \xi_{k+1} \\ w_k \end{pmatrix} < 0 \quad \text{Use (A.72)} \\ \widehat{=} \quad & \xi_k^\top \left(-\tau_1 P_1 - G_\Omega(S_1)\right) \xi_k + \xi_{k+1}^\top \left(P_1 - G_\Omega(S_2)\right) \xi_{k+1} - \tau_2 w_k^\top w_k < 0 \\ \widehat{=} \quad & \xi_{k+1}^\top P_1 \xi_{k+1} - \tau_1 \xi_k^\top P_1 \xi_k - \tau_2 w_k^\top w_k - 2 s_\Omega(S_1, Cx_k) - 2 s_\Omega(S_2, Cx_{k+1}) < 0, \quad \text{Use (B.6a)} \end{aligned}$$

which after rearranging terms results in

$$V_1(f(x_k,\, w_k)) - \tau_1 V_1(x_k) - \tau_2 \|w_k\|^2 - 2 s_\Omega(S_1, Cx_k) - 2 s_\Omega(S_2, Cx_{k+1}) < 0. \tag{A.74}$$

One can now add Inequality (B.7c) such that we find the form

$$W(x_k,\, w_k) - S(x_k,\, w_k) + g_4(x_k,\, w_k) < 0, \tag{A.75}$$

where

$$W(x_k,\, w_k) = V_1(f(x_k, w_k)) - c \tag{A.76a}$$

$$S(x_k, w_k) = \sum_{i=1}^{2} \tau_i G_i(x_k, w_k), \quad G_1(x_k, w_k) = V_1(x_k) - c, \quad G_2(x_k, w_k) = w_k^\top w_k - \sigma \tag{A.76b}$$

$$g_4(x_k, w_k) = -2 s_\Omega(S_1,\, Cx_k) - 2 s_\Omega(S_2,\, Cf(x_k, w_k)) \tag{A.76c}$$

Thanks to sector condition (2.16), we conclude $g_4(x_{k+1}) \geq 0$. Thus, $W(f(x_k, w_k)) - S(x_k, w_k) < 0$ is necessary to satisfy (A.75). Therefore, we can apply the $\mathcal{S}$-procedure as Lemma B.5 to conclude that

$$\begin{aligned} W(x_k,\, w_k) < 0 \mid & \{G_1(x) \leq 0,\, G_2(x) \leq 0\} \\ V_1(f(x_k, w_k)) < c \mid & \left\{V_1(x_k) \leq c,\, \|w\|^2 \leq \sigma\right\}, \end{aligned} \quad \text{Use (A.76)}$$

This expression indeed verifies condition (3.6b) for the quadratic $\mathcal{K}$-function $\gamma(s) = \frac{1}{\sigma} s^2$ given the positive coefficient σ and argument $s = \|w_k\| \in \mathbb{R}_{>0}$.

From this analysis, we can conclude that condition (3.6)in Lemma 3.1 holds true and hence, there exists a compact, positively invariant set $\mathcal{S}_c = \left\{x \in \mathbb{R}^n \mid V_1(x) \leq c,\, \|w\|_2^2 \leq \gamma\right\}$.

Part 2: Global exponential uniform asymptotic incremental stability

The global exponential uniform asymptotic incremental stability property is guaranteed if the assumptions in Lemma 3.1 are verified for $\mathcal{S} = \mathbb{R}^n$. This part of the proof shows that solutions to Theorem 3.5 imply conditions (3.9) for the time-independent incremental Lyapunov function $V_2(\xi(x^a),\, \xi(x^b)) = \big\|\xi(x^a) - \xi(x^b)\big\|_{P_2}^2$ and adequate choices of functions $\alpha_3(\cdot)$, $\alpha_4(\cdot)$ and $\alpha_5(\cdot)$. To that end, conditions (3.9a) and (3.9b) are checked one after the other.

- Firstly, we consider condition (3.9a) in Lemma 3.3. A suitable lower bound to V_2 can be thought of as the quadratic $\mathcal{K}_\infty$-function $\alpha_3(r) = \tilde{\alpha}_3\, r^2$ for some $\tilde{\alpha}_3 > 0$. This lower-bound exists as can be seen by pre- and post-multiplying the LMI (B.7d) by $\big(\xi(x^a) - \xi(x^b)\big)^\top$ and $\big(\xi(x^a) - \xi(x^b)\big)$, which yields for all $\big\|x^a - x^b\big\| \neq 0$:

$$
\begin{aligned}
& & P_2 + G_{\bar{\Omega}}(S_3) &\succ 0_{n_x+n_y} && \big) \text{ Pre- and post-multiply}\\
\implies & \big(\xi(x^a) - \xi(x^b)\big)^\top \big(P_2 + G_{\bar{\Omega}}(S_3)\big)\big(\xi(x^a) - \xi(x^b)\big) &> 0 && \big) \text{ Substitute (B.6b)}\\
\implies & \big\|\xi(x^a) - \xi(x^b)\big\|_{P_2}^2 + 2s_{\bar{\Omega}}(S_3,\, Cx^a,\, Cx^b) &> 0 && \big) \text{ Strict inequality}\\
\implies & \big\|\xi(x^a) - \xi(x^b)\big\|_{P_2}^2 + 2s_{\bar{\Omega}}(S_3,\, Cx^a,\, Cx^b) &> \tilde{\alpha}_3\big\|x^a - x^b\big\|^2 && \big) \; s_{\bar{\Omega}} \leq 0\\
\implies & \big\|\xi(x^a) - \xi(x^b)\big\|_{P_2}^2 &\geq \tilde{\alpha}_3\big\|x^a - x^b\big\|^2 &&\\
\implies & V_2(x^a, x^b) &\geq \alpha_3(r), &&
\end{aligned}
$$

  which indeed satisfies the lower bound in (3.9a) for some positive scalar $\tilde{\alpha}_3$ and argument $r = \big\|x^a - x^b\big\|$.

  A suitable upper bound to V_2 can be thought of as the quadratic $\mathcal{K}_\infty$-function $\alpha_4(r) = \tilde{\alpha}_4\, r^2$ for some $\tilde{\alpha}_4 > 0$. This upper-bound exists as can be seen by applying the Rayleigh inequality, see Lemma B.4, to the quadratic form $V_2(\xi(x^a),\, \xi(x^b))$. We have

$$
\begin{aligned}
V_2(\xi(x^a), \xi(x^b)) &= \big(\xi(x^a) - \xi(x^b)\big)^\top P_2 \big(\xi(x^a) - \xi(x^b)\big) && \big) \text{ Rayleigh's inequality}\\
&\leq \rho(P_2)\big\|\xi(x^a) - \xi(x^b)\big\|_2^2 && \big) \; \|\xi\|_2^2 = \|x\|_2^2 + \|\varphi(y)\|_2^2\\
&= \rho(P_2)\left(\big\|x^a - x^b\big\|_2^2 + \big\|\varphi(y^a) - \varphi(y^b)\big\|_2^2\right) && \big) \text{ Substitute (A.71)}\\
&\leq \rho(P_2)\left(\big\|x^a - x^b\big\|_2^2 + \rho(C^\top \bar{\Omega}\bar{\Omega} C)\big\|x^a - x^b\big\|_2^2\right) &&\\
&= \tilde{\alpha}_4\big\|x^a - x^b\big\|_2^2 &&\\
&= \alpha_4(r) &&
\end{aligned}
$$

  which indeed satisfies the upper bound in (3.9a) for a positive scalar $\tilde{\alpha}_4 := \rho(P_2)(1 + \rho(C^\top \bar{\Omega}\bar{\Omega} C)$ and the argument $r = \big\|x^a - x^b\big\|_2$.

- Consecutively, we consider condition (3.9b) in Lemma 3.3. One can define

$$
\begin{pmatrix} \xi(x_k^a) - \xi(x_k^b) \\ \xi(x_{k+1}^a) - \xi(x_{k+1}^b) \end{pmatrix} = M_0 \mu_k, \tag{A.77}
$$

  where

$$
\mu_k := \begin{pmatrix} \xi(x_k^a) - \xi(x_k^b) \\ \varphi(Cx_{k+1}^a) - \varphi(Cx_{k+1}^b) \end{pmatrix}. \tag{A.78}
$$

  note that μ_k has a similar role as η_k in (A.73) with the difference that η_k depends on the input w, whereas μ_k is independent of w.

One can now pre- and post-multiply LMI (B.7e) by $\mu_k^\top$ and μ_k respectively, such that

$$\mu_k^\top M_0^\top \begin{bmatrix} -P_2 - G_{\bar{\Omega}}(S_4) & 0_{n_x+n_y} \\ 0_{n_x+n_y} & P_2 - G_{\bar{\Omega}}(S_5) \end{bmatrix} M_0 \mu_k \prec 0 \qquad \text{Use (A.77)}$$

$$\begin{pmatrix} \xi(x_k^a) - \xi(x_k^b) \\ \xi(x_{k+1}^a) - \xi(x_{k+1}^b) \end{pmatrix}^\top \begin{bmatrix} -P_2 - G_{\bar{\Omega}}(S_4) & 0_{n_x+n_y} \\ 0_{n_x+n_y} & P_2 - G_{\bar{\Omega}}(S_5) \end{bmatrix} \begin{pmatrix} \xi(x_k^a) - \xi(x_k^b) \\ \xi(x_{k+1}^a) - \xi(x_{k+1}^b) \end{pmatrix} < 0 \qquad \text{Use (B.6b)}$$

$$\left\|\xi(x_{k+1}^a) - \xi(x_{k+1}^b)\right\|_{P_2}^2 - \left\|\xi(x_k^a) - \xi(x_k^b)\right\|_{P_2}^2 - 2s_{\bar{\Omega}}(S_4, Cx^a, Cx^b) - 2s_{\bar{\Omega}}(S_5, Cx^a, Cx^b) < 0$$

Due to the strict inequality and cone bounded sector condition (2.17), we can write this inequality as

$$V_2(x_{k+1}^a,\, x_{k+1}^b) - V_2(x_k^a,\, x_k^b) < -\tilde{\alpha}_5 \left\|x_k^a - x_k^b\right\|^2 \tag{A.79}$$

This inequality indeed verifies condition (3.9b) for the quadratic $\mathcal{K}_\infty$-function $\alpha_5(s) = \tilde{\alpha}_5 s^2$, with some positive coefficient $\tilde{\alpha}_5$ and argument $s = \left\|x^a - x^b\right\| \in \mathbb{R}_{>0}$.

Now all assumptions of Lemma 3.3 have been checked for solutions of Theorem B.2 and hence the system is globally exponentially uniformly asymptotically incrementally stable with respect to dynamics (??).

In conclusion, this proof has shown the existence of a compact, positively invariant set, as well as the global exponential uniform asymptotic incremental stability property. As a result, Theorem 3.4 defines the system as GEUC with respect to the considered dynamics for the class of inputs $\mathcal{W}_\gamma$. □

A.14 Proof of Lemma B.1

Proof.
By definition of $M \succ 0$, the condition is trivially satisfied for $N = 0$. The case $\rho(N) > 0$ can be considered as follows:

- Because of the strict inequality $M \succ 0$, there exists a scalar $\kappa_1 > 0$, such that
$$M - \kappa_1 I \succ 0. \tag{A.80}$$
- Select
$$\kappa^* = \frac{\kappa_1}{\rho(N)}, \tag{A.81}$$
such that for all $k \in [-\kappa^*,\ \kappa^*]$, by definition of the spectral radius, it holds
$$-\rho(\kappa N)I \leq \kappa N \leq \rho(\kappa N)I. \tag{A.82}$$
By adding $\kappa_1 I$ to these inequalities, we find:
$$-\rho(\kappa N)I + \kappa_1 I \leq \kappa N + \kappa_1 I \leq \rho(\kappa N)I + \kappa_1 I, \tag{A.83}$$
in which
$$\rho(\kappa N) = |\kappa|\rho(N) \leq \kappa^* \rho(N) = \kappa_1. \tag{A.84}$$
From Equations (A.83) and (A.84), we conclude $-\rho(\kappa N) + \kappa_1 \geq 0$. This in turn implies
$$\kappa N + \kappa_1 I \succeq 0. \tag{A.85}$$
- Finally, we observe that
$$M + \kappa N = (M - \kappa_1 I) + (\kappa N + \kappa_1 I) \succ 0, \tag{A.86}$$
which concludes this proof.

□

A.15 Proof of Lemma B.2

Proof.

- For a positive definite matrix $X \succ 0$, it holds
$$z^T X z > 0, \qquad \forall z \neq 0. \tag{A.87}$$
- Define a new variable y, such that
$$z = Ty. \tag{A.88}$$
Furthermore, notice that since T is invertible, such that $z = 0 \iff y = 0$.
- Substitute Equation (A.88) in the positive definite form (A.87), such that
$$\begin{aligned} & z^\top X z \geq 0, && \forall z \neq 0 \\ \iff \quad & (Ty)^\top X (Ty) \geq 0, && \forall y \neq 0 \\ \iff \quad & y^\top \left(T^\top X T\right) y \geq 0, && \forall y \neq 0 \\ \iff \quad & T^\top X T \succ 0. \end{aligned}$$

□

A.16 Proof of Lemma B.3

Proof.

- Choose the following block-structured matrices:

$$T_1 = \begin{bmatrix} I_p & 0 \\ -D^{-1}C & I_q \end{bmatrix}, \qquad \text{(A.89)} \qquad T_2 = \begin{bmatrix} I_p & -A^{-1}B \\ 0 & I_q \end{bmatrix}. \qquad \text{(A.90)}$$

- Observe that T_1 and T_2 are block upper- and lower-triangular matrices with identities on their block diagonal. As a result, all eigenvalues of T_1 and T_2 equal one and therefore both matrices are invertible by definition.
- Notice that

$$T_1^\top M T_1 = \begin{bmatrix} M/D & 0 \\ 0 & D \end{bmatrix}, \qquad \text{(A.91)} \qquad T_2^\top M T_2 = \begin{bmatrix} A & 0 \\ 0 & M/A \end{bmatrix}. \qquad \text{(A.92)}$$

- From Lemma B.3 and the structures of Equations (A.91) and (A.92), we can now conclude that indeed Equation (B.13) holds true.

□

A.17 Proof of Lemma B.4

Proof. (See [98])
Since A is real and symmetric, it is orthogonally diagonalizable. Specifically, define the n orthonormal eigenvectors of A as $[v_1,\, v_2,\, \ldots,\, v_n] \in \mathbb{R}^n$, and let $[\lambda_1,\, \lambda_2,\, \ldots,\, \lambda_n] \in \mathbb{R}$ be their corresponding eigenvalues. Let P be the matrix whose i^{th} column is v_i and let D be a diagonal matrix whose i^{th} diagonal entry is λ_i. Then

$$A = PDP^\top, \qquad \text{and} \qquad P^\top P = I.$$

Choose $v := P^\top u$, such that

$$v^\top v = \left(P^\top u\right)^\top \left(P^\top u\right) = u^\top \left(PP^\top\right) u = u^\top u.$$

Furthermore, notice that

$$u^\top A u = u^\top \left(PDP^\top\right) u = \left(P^\top u\right) D \left(P^\top u\right) = v^\top D v = \sum_{i=1}^{n} \lambda_i v_i^2.$$

Without loss of generality, assume $\lambda_n \leq \cdots \leq \lambda_i \leq \cdots \leq \lambda_1$, such that

$$\begin{aligned} \sum_{i=1}^{n} \lambda_n v_i^2 \leq \sum_{i=1}^{n} \lambda_i v_i^2 \leq \sum_{i=1}^{n} \lambda_1 v_i^2 \\ \implies \lambda_n v^\top v \leq \sum_{i=1}^{n} \lambda_i v_i^2 \leq \lambda_1 v^\top v \\ \implies \lambda_n u^\top u \leq u^\top A u \leq \lambda_1 u^\top u \end{aligned}$$

which was the intended result. Finally, notice that the provided bounds are tight, because of

$$v_n^\top A v_n = \lambda_n v_n^\top v_n \qquad \wedge \qquad v_1^\top A v_1 = \lambda_1 v_1^\top v_1.$$

□

Appendix B

Additional stability conditions

In this appendix we discuss additional stability conditions that are researched as part of this project, but were not pursued into a main contribution to the system identification algorithm. First we show a set of conditions *with an explicit dependency on feedthrough matrix* D that guarantee global, exponential uniform convergence via Lyapunov functions that are quadratic in the system state. Thereafter, we discuss conditions that characterise stability via Lyapunov functions that are *quadratic in an extended state* in an attempt to reduce the conservatism of our stability analysis. The final section of this appendix addresses some auxiliary lemmas that were used to proof of at least one set of stability conditions.

B.1 Explicit feedthrough matrix dependency

Theorem B.1 (Sufficient conditions for global, exponential convergence)
Consider a Lur'e-type system $(\Sigma,\ \varphi)$ *according to Equation* (2.1)*. Let* Σ *and* φ *verify Assumptions 2.1 and 2.2 respectively. Furthermore, consider symmetric matrices* $P_1,\ P_2 \in \mathbb{S}^{n_x}$ *and positive scalars* $\tau_1,\ \tau_2,\ \sigma$ *and* c *such that the following inequalities hold true:*

$$\begin{bmatrix} A^\top \\ B^\top \\ F^\top \end{bmatrix} P_1 \begin{bmatrix} A^\top \\ B^\top \\ F^\top \end{bmatrix}^\top - \begin{bmatrix} \tau_1 P_1 & -C^\top \Omega S_1 & 0_{n_x \times n_w} \\ \star & 2S_1 & -S_1 \Omega D \\ \star & \star & \tau_2 I_{n_w} \end{bmatrix} \prec 0_{n_x+n_y+n_w}, \tag{B.1a}$$

$$-c(1-\tau_1) + \tau_2 \sigma \leq 0, \tag{B.1b}$$

$$\begin{bmatrix} A^\top \\ B^\top \end{bmatrix} P_2 \begin{bmatrix} A^\top \\ B^\top \end{bmatrix}^\top - \begin{bmatrix} P_2 & -C^\top \bar{\Omega} S_2 \\ \star & 2S_2 \end{bmatrix} \prec 0_{n_x+n_y}. \tag{B.1c}$$

Then, the assumptions of Theorem 3.4 are verified by considering $V_1(x) = \|x\|_{P_1}^2$ *and* $V_2(x^a,\ x^b) = \left\|x^a - x^b\right\|_{P_2}^2$*. According to Theorem 3.4, the Lur'e system is globally exponentially uniformly convergent with respect to the dynamics* (2.1) *for the class of inputs* $\mathcal{W}_\gamma = \{w_k \in \mathbb{R}^{n_w},\ \|w_k\| \leq \sqrt{\sigma}\}$*. Furthermore, the steady-state solution* $\bar{x}_k^w$ *belongs to the sellipsoid* $\mathcal{E}(P_1,\ c)$*.*

Proof. See Appendix A.12. □

As an alternative to the conditions, we can verify the conditions in Theorem 3.5 together with Lemma 3.4.

B.2 Quadratic Lyapunov functions in the extended state

The next conditions can be used as a starting point to reduce the conservatism in our stability conditions. Here, we once more assume no direct feedthrough from w into y, hence $D = 0_{n_y}$. We can extend our scope of Lyapunov function candidates by going beyond quadratic functions in the system state. To that end, let us define an extended state $\xi_k : \mathbb{R}^{n_x} \to \mathbb{R}^{n_x+n_y}$ as

$$\xi_k = \begin{pmatrix} x_k \\ u_k \end{pmatrix} = \begin{pmatrix} x_k \\ \varphi(Cx_k) \end{pmatrix}, \tag{B.2}$$

such that the class of generalized Lyapunov functions can be introduced as

$$\tilde{V}(x_k) = \|\xi_k\|^2_{\tilde{P}}, \qquad \tilde{P} = \begin{bmatrix} \tilde{P}_a & \tilde{P}_b \\ \star & \tilde{P}_c \end{bmatrix}, \tag{B.3}$$

where $0_n \prec \tilde{P}_a \in \mathbb{S}^{n_x}$. Furthermore, $\tilde{P}_b \in \mathbb{R}^{n_x \times n_y}$ and $\tilde{P}_c \in \mathbb{R}^{n_y \times n_y}$ are indefinite. The commonly used Lur'e type Lyapunov functions are inside this class for $\tilde{P}_b = C^\top R\Omega$ and $\tilde{P}_c = 0$, given some diagonal $R \succ 0$. For clarification and readability of the sequel, the following notation is adopted:

- Matrices $G_\Omega(S)$ and $G_{\bar{\Omega}}(S)$ are introduced as

$$G_\Omega(S) = \begin{bmatrix} 0_n & -C^\top S\Omega \\ \star & 2S \end{bmatrix}, \tag{B.4}$$

$$G_{\bar{\Omega}}(S) = \begin{bmatrix} 0_n & -C^\top S\bar{\Omega} \\ \star & 2S \end{bmatrix}. \tag{B.5}$$

  As a result, the cone bounded sector constraints (2.16) and (2.17) can be written into the following negative semidefinite quadratic forms for any diagonal matrix $0 \prec S \in \mathbb{S}^{n_y}$:

$$2s_\Omega(S,\, Cx) = \|\xi(x)\|^2_{G_\Omega(S)} \leq 0, \tag{B.6a}$$

$$2s_{\bar{\Omega}}(S,\, Cx^a,\, Cx^b) = \left\|\xi(x^a) - \xi(x^b)\right\|^2_{G_{\bar{\Omega}}(S)} \leq 0, \tag{B.6b}$$

- Matrices $M_0,\ M_1$ depend solely on a priori known system matrices of dynamics (??) and are defined as

$$M_0 := \begin{bmatrix} I_{n_x} & 0_{n_x \times n_y} & 0_{n_x \times n_y} \\ 0_{n_y \times n_x} & I_{n_y} & 0_{n_x \times n_y} \\ A & B & 0_{n_x \times n_y} \\ 0_{n_y \times n_x} & 0_{n_y} & I_{n_y} \end{bmatrix}, \qquad M_1 := \left[\begin{array}{c|c} M_0 & \begin{matrix} 0_{(n_x+n_y)\times n_w} \\ F \\ 0_{n_y \times n_w} \end{matrix} \\ \hline 0_{n_w \times (n_x+n_y)} & I_{n_w} \end{array}\right].$$

We can now formally introduce the conditions in Theorem B.2 that are sufficient for the global exponential uniform convergence property via Lyapunov-like functions that are quadratic in the *extended state*.

Theorem B.2 (Quadratic Lyapunov functions in the extended state [75, Th. 6])
Consider a Lur'e-type system $(\Sigma,\, \varphi)$ according to Equation (2.1) and assume $D = 0_{n_y}$. Let Σ and φ verify Assumptions 2.1 and 2.2 respectively. Furthermore, consider symmetric matrices $\tilde{P}_1,\ P_2 \in \mathbb{S}^{n_x+n_y}$ and diagonal positive definite matrices $S_0,\ S_1,\ S_2, S_3,\ S_4,\ S_5 \in \mathbb{S}^{n_y}$. In addition, consider positive scalars $\tau_1,\ \tau_2,\ \sigma$ and c such that the following inequalities hold true:

$$\tilde{P}_1 + G_\Omega(S_0) \succ 0_{n_x+n_y} \tag{B.7a}$$

$$M_1^\top \begin{bmatrix} -\tau_1\tilde{P}_1 - G_\Omega(S_1) & 0_{n_x+n_y} & 0_{(n_x+n_y)\times n_w} \\ 0_{n_x+n_y} & \tilde{P}_1 - G_\Omega(S_2) & 0_{(n+p)\times n_w} \\ 0_{m\times(n+p)} & 0_{m\times(n+p)} & -\tau_2 I_{n_w} \end{bmatrix} M_1 \prec 0_{n_x+2n_y+n_w} \tag{B.7b}$$

$$-c(1-\tau_1) + \tau_2\sigma \leq 0 \tag{B.7c}$$

$$P_2 + G_{\bar{\Omega}}(S_3) \succ 0_{n_x+n_y} \tag{B.7d}$$

$$M_0^\top \begin{bmatrix} -P_2 - G_{\bar{\Omega}}(S_4) & 0_{n_x+n_y} \\ 0_{n_x+n_y} & P_2 - G_{\bar{\Omega}}(S_5) \end{bmatrix} M_0 \preceq 0_{n_x+2n_y} \tag{B.7e}$$

Then, the assumptions of Theorem 3.4 are verified by considering the (incremental) Lyapunov functions $\tilde{V}_1(x) = \|x\|^2_{\tilde{P}_1}$ and $\tilde{V}_2(x^a,\, x^b) = \left\|x^a - x^b\right\|^2_{P_2}$. By Theorem 3.4, the considered Lur'e-type system is globally exponentially uniformly convergent for the class of inputs $\mathcal{W}_\gamma = \{w(k) \in \mathbb{R}^m,\ \|w(k)\| \leq \sqrt{\sigma}\}$.

Proof. See Appendix A.13 □

B.3 Auxiliary lemmas

This section features a selection of theorems and lemmas that are required to understand the reasoning in the main text of this report.

Schur complements are often of interest in the analysis of positive definite matrices. The results in Lemmas B.1 to B.3 were used in the proof of Theorem 3.6.

Lemma B.1
For a symmetric, positive definite matrix $M \succ 0$, and a (possibly indefinite) matrix N of the same dimension, there exists a sufficiently small scalar κ^*, such that for any $\kappa \in [-\kappa^*,\ \kappa^*]$, it holds:

$$M + \kappa N \succ 0. \tag{B.8}$$

Proof. See Appendix A.14 □

Lemma B.2 (Sylvester Law of Inertia [99, Th. 8.1.17])
Consider an invertible matrix T. Then:

$$X \succ 0 \iff T^\top X T \succ 0. \tag{B.9}$$

Proof. See Appendix A.15 □

Lemma B.3 (Schur complement and positive definite matrices [100, Prop. 16.1])
Suppose $p,\ q$ are nonnegative integers such that $A \in \mathbb{R}^{n_y \times n_y}$, $B \in \mathbb{R}^{q \times q}$ and M is a symmetric block-structured matrix

$$M = \begin{bmatrix} A & B \\ C & D \end{bmatrix}. \tag{B.10}$$

Note that $C = B^\top$ due to the symmetric property of M. Schur complements of M can then be defined as follows:

- *If D is invertible, then the Schur complement of the block D in M is defined by*

$$M/D := A - BD^{-1}C \tag{B.11}$$

- *If A is invertible, then the Schur complement of the block A in M is defined by*

$$M/A := D - CA^{-1}B \tag{B.12}$$

For these Schur complements, the following properties hold true:

$$M \succ 0 \iff A \succ 0 \ \wedge \ M/A \succ 0 \tag{B.13a}$$
$$M \succ 0 \iff D \succ 0 \ \wedge \ M/D \succ 0 \tag{B.13b}$$

Proof. See Appendix A.16 □

Rayleigh's inequality is applied multiple times in the proofs of Theorem B.2 that can be found in Appendix A.13.

Lemma B.4 (Rayleigh's inequality [101, p. 34])
Let A be a $n \times n$ real symmetric matrix. Furthermore, let λ_{min} and λ_{max} be the minimum and maximum eigenvalues of A, with v_{min} and v_{max} being their corresponding eigenvectors. Then

$$\lambda_{\min} u^\top u \leq u^\top A u \leq \lambda_{\max} u^\top u \qquad \forall u \in \mathbb{R}^n. \tag{B.14}$$

Proof. See Appendix A.17. □

The well-known $\mathcal{S}$-procedure is applied in the proof of Theorem B.2, which can be found in Appendix A.13.

Lemma B.5 ($\mathcal{S}$-procedure [102, p. 103])
Let us consider the statement

$$W(x) < 0 \mid \{G_1(x) \leq 0,\, G_2(x) \leq 0,\, \ldots\, , G_k(x) \leq 0,\, x \neq 0\}. \tag{B.15}$$

Suppose that we define a weighted sum of G_i's with weighting factors λ_i according to

$$S(x) = \sum_{i=1}^{k} \lambda_i G_i(x), \qquad \lambda_i \geq 0,\, \forall i \in \{1,\, \ldots\, , k\}.$$

Furthermore, suppose that there exist nonnegative λ_i's such that the following statement holds true

$$W(x) - S(x) < 0 \mid \{x \neq 0\}. \tag{B.16}$$

Then statement (B.16) *implies* (B.15).

Finally, the well-known contraction mapping theorem is applied in the proof of Theorem 4.2, which can be found in Appendix A.11.

Theorem B.3 (Contraction Mapping Theorem [103, Th. A.3.1])
Let X be a Banach space, T a mapping from X to X, $m \in \mathbb{N}$, and $\alpha < 1$. Suppose that T satisfies

$$\|T^m(x_1) - T^m(x_2)\| \leq \alpha \|x_1 - x_2\| \qquad \forall x_1, x_2 \in X, \tag{B.17}$$

Then:

1. *There exists a unique fixed point of T: $x^* \in X$, such that $T(x^*) = x^*$.*
2. *For any $x_0 \in X$, the sequence $\{x_n, n \geq 1\}$ defined by $x_n := T^n(x_0)$ converges to x^* as $n \to \infty$.*

Proof. See reference in [103, Th. A.3.1] □

Appendix C

TU/e Code of Scientific Conduct

Declaration concerning the TU/e Code of Scientific Conduct for the Master's thesis

I have read the TU/e Code of Scientific Conduct[i].

I hereby declare that my Master's thesis has been carried out in accordance with the rules of the TU/e Code of Scientific Conduct

Date

16/09/2021

Name

N. (Niels) Vervaet

ID-number

0941902

Signature

Submit the signed declaration to the student administration of your department.

[i] See: https://www.tue.nl/en/our-university/about-the-university/organization/integrity/scientific-integrity/
The Netherlands Code of Conduct for Scientific Integrity, endorsed by 6 umbrella organizations, including the VSNU, can be found here also. More information about scientific integrity is published on the websites of TU/e and VSNU

February 21, 2020

Bibliography

[1] J. Spruyt. "General Dynamics (SABCA) F-16AM Fighting Falcon (401) - Belgium - Air Force | Airliners.net." (2018), [Online]. Available: https://www.airliners.net/photo/Belgium-Air-Force/General-Dynamics-SABCA-F-16AM-Fighting-Falcon-401/5540981 (visited on 09/15/2021).

[2] W. D. Ross, *Aristotle's Metaphysics: Books ZN*. Clarendon Press, 1924.

[3] S. Gould, "Nonoverlapping magisteria," *Natural history*, vol. 106, no. 2, pp. 16–22, 1997.

[4] T. Ritchey, "Analysis and synthesis: On scientific method – based on a study by Bernhard Riemann," *Systems Research*, vol. 8, no. 4, pp. 21–41, 1991, ISSN: 10991735. DOI: 10.1002/sres.3850080402.

[5] L. Von Bertalanffy, "The History and Status of General Systems Theory," *Academy of Management Journal*, vol. 15, no. 4, pp. 407–426, 1972, ISSN: 0001-4273. DOI: 10.5465/255139.

[6] Y. Bar-Yam, "General Features of Complex Systems," *Knowledge Management, Organisational Intelligence and Learning and Complexity*, vol. I, no. 1, pp. 1–10, 1997.

[7] L. Ljung, "System identification—Theory for the user," *PTR Prentice Hall, Upper Saddle River, NJ*, vol. 28, 1999, ISSN: 00051098.

[8] B. Edmonds, C. Le Page, M. Bithell, *et al.*, "Different modelling purposes," *JASSS*, vol. 22, no. 3, Jun. 2019, ISSN: 14607425. DOI: 10.18564/jasss.3993.

[9] X.-Q. Zhao, *Dynamical Systems in Population Biology (CMS Books in Mathematics)*. Springer, 2011, p. 276, ISBN: 1441918159. DOI: 10.1007/978-0-387-21761-1.

[10] W.-B. Zhang, *Discrete dynamical systems, bifurcations and chaos in economics*. Elsevier, 2006, ISBN: 9780444521972.

[11] J. Gleick and R. C. Hilborn, "Chaos, Making a New Science," *American Journal of Physics*, vol. 56, no. 11, pp. 1053–1054, Nov. 1988, ISSN: 0002-9505. DOI: 10.1119/1.15345.

[12] J. Schoukens and L. Ljung, "Nonlinear System Identification: A User-Oriented Road Map," *IEEE Control Systems*, vol. 39, no. 6, pp. 28–99, 2019, ISSN: 1941000X. DOI: 10.1109/MCS.2019.2938121. arXiv: 1902.00683.

[13] M. F. Shakib, A. Y. Pogromsky, A. Pavlov, and N. van de Wouw, "Fast identification of continuous-time Lur'e-type systems with stability certification," in *IFAC-PapersOnLine*, vol. 52, Elsevier B.V., Sep. 2019, pp. 227–232. DOI: 10.1016/j.ifacol.2019.11.783.

[14] M. Shakib, A. Pogromsky, A. Pavlov, and N. van de Wouw, "Computationally Efficient Identification of Continuous-Time Lur'e-type Systems with Stability Guarantees [Accepted]," *Automatica*,

[15] R. K. Pearson and Ü. Kotta, "Nonlinear discrete-time models: State-space vs. I/O representations," *Journal of Process Control*, vol. 14, no. 5, pp. 533–538, Aug. 2004, ISSN: 09591524. DOI: 10.1016/j.jprocont.2003.09.007.

[16] G. F. Shoukry, "State-space realization for nonlinear systems," Ph.D. dissertation, Georgia Institute of Technology, 2008.

[17] J. Slotine and W. Li, *Applied Nonlinear Control*. Prentice hall Englewood Cliffs, NJ, 1991, ISBN: 0130408905.

[18] C. C. Pantelides and J. G. Renfro, "The online use of first-principles models in process operations: Review, current status and future needs," *Computers and Chemical Engineering*, vol. 51, pp. 136–148,

2013, ISSN: 00981354. DOI: 10.1016/j.compchemeng.2012.07.008.

[19] S. L. Brunton and J. N. Kutz, *Data-Driven Science and Engineering: Machine learning, dynamical systems, and control*. Cambridge University Press, 2019. DOI: 10.1017/9781108380690.

[20] G. F. Sirca and H. Adeli, *System identification in structural engineering*, Dec. 2012. DOI: 10.1016/j.scient.2012.09.002.

[21] J. Reinders, B. Hunnekens, F. Heck, T. Oomen, and N. van de Wouw, "Adaptive Control for Mechanical Ventilation for Improved Pressure Support," *IEEE Transactions on Control Systems Technology*, pp. 1–14, Feb. 2020, ISSN: 1063-6536. DOI: 10.1109/tcst.2020.2969381.

[22] F. Giri and E.-W. Bai, *Block-oriented Nonlinear System Identification*, ser. Lecture Notes in Control and Information Sciences. London: Springer London, 2010, vol. 404, ISBN: 978-1-84996-512-5. DOI: 10.1007/978-1-84996-513-2.

[23] M. Schoukens and K. Tiels, "Identification of block-oriented nonlinear systems starting from linear approximations: A survey," *Automatica*, vol. 85, pp. 272–292, 2017, ISSN: 00051098. DOI: 10.1016/j.automatica.2017.06.044. arXiv: 1607.01217.

[24] A. De Kraker, N. Van De Wouw, H. L. Van Den Bosch, and D. H. Van Campen, "Identification of nonlinear phenomena in a stochastically excited beam system with impact," in *Proceedings of the 23rd International Conference on Noise and Vibration Engineering, ISMA*, 1998, pp. 569–576, ISBN: 9073802679.

[25] S. Boyd and L. O. Chua, "Fading Memory and the Problem of Approximating Nonlinear Operators with Volterra Series," *IEEE transactions on circuits and systems*, vol. CAS-32, no. 11, pp. 1150–1161, 1985, ISSN: 00984094. DOI: 10.1109/tcs.1985.1085649.

[26] G. Palm, "On representation and approximation of nonlinear systems - Part II: Discrete time," *Biological Cybernetics*, vol. 34, no. 1, pp. 49–52, Sep. 1979, ISSN: 03401200. DOI: 10.1007/BF00336857.

[27] H. Khalil, *Nonlinear systems*, 2nd ed. Upper Saddle River NJ: Prentice Hall, 1996, ISBN: 9780132280242.

[28] A. Van Mulders, L. Vanbeylen, and K. Usevich, "Identification of a block-structured model with several sources of nonlinearity," in *2014 European Control Conference, ECC 2014*, Institute of Electrical and Electronics Engineers Inc., Jul. 2014, pp. 1717–1722, ISBN: 9783952426913. DOI: 10.1109/ECC.2014.6862455.

[29] P. Hu and F. Ding, "Multistage least squares based iterative estimation for feedback nonlinear systems with moving average noises using the hierarchical identification principle," *Nonlinear Dynamics*, vol. 73, no. 1-2, pp. 583–592, Jul. 2013, ISSN: 0924090X. DOI: 10.1007/s11071-013-0812-0.

[30] A. Van Mulders, J. Schoukens, and L. Vanbeylen, "Identification of systems with localised nonlinearity: From state-space to block-structured models," in *Automatica*, vol. 49, Pergamon, May 2013, pp. 1392–1396. DOI: 10.1016/j.automatica.2013.01.052.

[31] L. Vanbeylen, "Nonlinear LFR block-Oriented model: Potential benefits and improved, user-Friendly identification method," *IEEE Transactions on Instrumentation and Measurement*, vol. 62, no. 12, pp. 3374–3383, 2013, ISSN: 00189456. DOI: 10.1109/TIM.2013.2272868.

[32] J. Paduart, G. Horvatti, and J. Schoukens, "Fast identification of systems with nonlinear feedback," in *IFAC Proceedings Volumes (IFAC-PapersOnline)*, IFAC Secretariat, Sep. 2004, pp. 381–385. DOI: 10.1016/S1474-6670(17)31253-3.

[33] J. Schoukens, J. G. Nemeth, P. Crama, Y. Rolain, and R. Pintelon, "Fast approximate identification of nonlinear systems," *Automatica*, vol. 39, no. 7, pp. 1267–1274, Jul. 2003, ISSN: 00051098. DOI: 10.1016/S0005-1098(03)00083-9.

[34] T. H. Van Pelt and D. S. Bernstein, "Nonlinear system identification using Hammerstein and nonlinear feedback models with piecewise linear static maps - Part 1: theory," in *Proceedings of the American Control Conference*, vol. 1, IEEE, 2000, pp. 225–229. DOI: 10.1109/acc.2000.878840.

[35] K. S. Narendra Fellow and K. Parthasarathy, "Identification and Control of Dynamical Systems Using Neural Networks," *IEEE TRANSACTIONS ON NEURAL NETWORKS*, 1990.

[36] M. Schoukens and R. Toth, "On the Initialization of Nonlinear LFR Model Identification with the Best Linear Approximation," *IFAC-PapersOnLine*, vol. 53, no. 2, 2020. DOI: 10.1016/j.ifacol.2020.12.142.

[37] J. Park and I. W. Sandberg, "Universal Approximation Using Radial-Basis-Function Networks," *Neural Computation*, vol. 3, no. 2, pp. 246–257, 1991, ISSN: 0899-7667. DOI: 10.1162/neco.1991.3.2.246.

[38] J. Adlers and G. Pihl, "Prediction of training time for deep neural networks in TensorFlow," Ph.D. dissertation, KTH, 2018.

[39] L. Ljung, "Perspectives on system identification," *Annual Reviews in Control*, vol. 34, no. 1, pp. 1–12, Apr. 2010, ISSN: 13675788. DOI: 10.1016/j.arcontrol.2009.12.001.

[40] M. M. Tobenkin, I. R. Manchester, and A. Megretski, "Convex Parameterizations and Fidelity Bounds for Nonlinear Identification and Reduced-Order Modelling," *IEEE Transactions on Automatic Control*, vol. 62, no. 7, pp. 3679–3686, Jul. 2017, ISSN: 00189286. DOI: 10.1109/TAC.2017.2694820. arXiv: 1701.06652.

[41] R. G. Hakvoort and P. M. J. Van Den Hof, "Frequency domain curve fitting with maximum amplitude criterion and guaranteed stability," *International Journal of Control*, vol. 60, no. 5, pp. 809–825, 1994, ISSN: 13665820. DOI: 10.1080/00207179408921496.

[42] T. D'Haene, R. Pintelon, and P. Guillaume, "Stable approximations of unstable models," in *Conference Record - IEEE Instrumentation and Measurement Technology Conference*, Institute of Electrical and Electronics Engineers Inc., 2007, ISBN: 1424410800. DOI: 10.1109/imtc.2007.379369.

[43] L. Balogh and R. Pintelon, "Stable approximation of unstable transfer function models," *IEEE Transactions on Instrumentation and Measurement*, vol. 57, no. 12, pp. 2720–2726, 2008, ISSN: 00189456. DOI: 10.1109/TIM.2008.926050.

[44] T. Van Gestel, J. A. K. Suykens, P. Van Dooren, and B. De Moor, "Identification of stable models in subspace identification by using regularization," *IEEE Transactions on Automatic Control*, vol. 46, no. 9, pp. 1416–1420, Sep. 2001, ISSN: 00189286. DOI: 10.1109/9.948469.

[45] S. L. Lacy and D. S. Bernstein, "Subspace identification with guaranteed stability using constrained optimization," *IEEE Transactions on Automatic Control*, vol. 48, no. 7, pp. 1259–1263, Jul. 2003, ISSN: 00189286. DOI: 10.1109/TAC.2003.814273.

[46] J. Umenberger and I. R. Manchester, "Specialized Interior-Point Algorithm for Stable Nonlinear System Identification," *IEEE Transactions on Automatic Control*, vol. 64, no. 6, pp. 2442–2456, Mar. 2019, ISSN: 15582523. DOI: 10.1109/TAC.2018.2867358. arXiv: 1803.01066.

[47] J. Decuyper, T. De Troyer, M. C. Runacres, K. Tiels, and J. Schoukens, "Nonlinear state-space modelling of the kinematics of an oscillating circular cylinder in a fluid flow," *Mechanical Systems and Signal Processing*, vol. 98, pp. 209–230, Jan. 2018, ISSN: 10961216. DOI: 10.1016/j.ymssp.2017.04.048. arXiv: 1804.08383.

[48] J. Decuyper, M. C. Runacres, J. Schoukens, and K. Tiels, "Tuning nonlinear state-space models using unconstrained multiple shooting," *IFAC-PapersOnLine*, vol. 53, no. 2, pp. 334–340, 2020. DOI: 10.1016/j.ifacol.2020.12.182.

[49] H. Zhu, M. Akrout, B. Zheng, *et al.*, *Benchmarking and analyzing deep neural network training*, 2018. DOI: 10.1109/IISWC.2018.8573476.

[50] J. Dean, G. S. Corrado, R. Monga, *et al.*, "Large scale distributed deep networks," in *Advances in Neural Information Processing Systems*, vol. 2, 2012, pp. 1223–1231, ISBN: 9781627480031.

[51] L. G. J. van de Kamp, "Nonlinear System Identification in Mechanical Ventilation," Eindhoven University of Technology, Tech. Rep., 2019.

[52] A. Pavlov, A. Pogromsky, N. Van De Wouw, and H. Nijmeijer, "Convergent dynamics, a tribute to Boris Pavlovich Demidovich," *Systems & Control Letters*, vol. 52, pp. 257–261, 2004. DOI: 10.1016/j.sysconle.2004.02.003.

[53] R. Pintelon and J. Schoukens, *System Identification: A Frequency Domain Approach, Second Edition*. Wiley, 2012, p. 743, ISBN: 9780470640371. DOI: 10.1002/9781118287422.

[54] J. Sjoberg, Q. Zhang, L. Ljung, *et al.*, *Nonlinear Black-Box Modeling in System Identification: a Unified Overview*, 12. Linkoping University, 1995, vol. 31, pp. 1691–1724. DOI: 10.1016/0005-1098(95)00120-8.

[55] M. Revay and I. R. Manchester, "Contracting Implicit Recurrent Neural Networks: Stable Models with Improved Trainability," *Learning for Dynamics and Control*, vol. 120, pp. 1–11, 2020. arXiv: 1912.10402.

[56] M. Revay, R. Wang, and I. R. Manchester, "Convex Sets of Robust Recurrent Neural Networks," *arXiv preprint arXiv:2004.05290*, 2020. arXiv: 2004.05290.

[57] A. Pavlov and N. Van De Wouw, "Fast computation of frequency response functions for a class of nonlinear systems," *Proceedings of the IEEE Conference on Decision and Control*, pp. 1180–1186, 2008, ISSN: 01912216. DOI: 10.1109/CDC.2008.4739265.

[58] T. Wigren and J. Schoukens, "Three free data sets for development and benchmarking in nonlinear system identification," in *2013 European Control Conference, ECC 2013*, IEEE Computer Society, 2013, pp. 2933–2938, ISBN: 9783033039629. DOI: 10.23919/ecc.2013.6669201.

[59] J. M. Maroli, Ü. Özgüner, and K. Redmill, "Nonlinear System Identification Using Temporal Convolutional Networks: A Silverbox Study," in *IFAC-PapersOnLine*, vol. 52, Elsevier, Jan. 2019, pp. 186–191. DOI: 10.1016/j.ifacol.2019.12.642.

[60] T. Yoshikawa and J. Innes, *Functional analysis*. 2013, pp. 71–81, ISBN: 9780203101261. DOI: 10.2307/3610182.

[61] D. E. Edmunds and Y. Netrusov, "Schütt's theorem for vector-valued sequence spaces," *Journal of Approximation Theory*, vol. 178, pp. 13–21, Sep. 2014, ISSN: 00219045. DOI: 10.1016/j.jat.2013.11.001. arXiv: 1309.7885.

[62] R. Baraniuk, "Signals and systems," *Applied and Numerical Harmonic Analysis*, no. 9780817643904, pp. 1–15, 2007, ISSN: 22965017. DOI: 10.1007/978-0-8176-4540-3_1.

[63] P. J. Antsaklis and A. N. Michel, *A linear systems primer*. Springer Science & Business Media, 2007, pp. 1–517, ISBN: 9780817644604. DOI: 10.1007/978-0-8176-4661-5.

[64] S. W. Smith, *The scientist and engineer's guide to digital signal processing*. California Technical Pub, 1997, p. 626, ISBN: 0966017633 9780966017632.

[65] T. McKelvey, "On the finite length DFT of input-output signals of multivariable linear systems," in *Proceedings of the IEEE Conference on Decision and Control*, vol. 5, 2000, pp. 5190–5191. DOI: 10.1109/cdc.2001.914780.

[66] A. Pavlov, B. G. B. Hunnekens, N. V. D. Wouw, and H. Nijmeijer, "Steady-state performance optimization for nonlinear control systems of Lur'e type," *Automatica*, vol. 49, no. 7, pp. 2087–2097, 2013, ISSN: 00051098. DOI: 10.1016/j.automatica.2013.04.017.

[67] A. Megretski. "MIT EECS 6.241: Dynamic Systems and Control - Lecture 6: "L2 Gain"." (2006), [Online]. Available: https://web.mit.edu/course/6/6.241/ameg_www_fall2006/www/images/L06L2gain.pdf.

[68] P. Gahinet and P. Apkarian, "A linear matrix inequality approach to H infinity control," *International Journal of Robust and Nonlinear Control*, vol. 4, no. 4, pp. 421–448, Jan. 1994, ISSN: 10991239. DOI: 10.1002/rnc.4590040403.

[69] F. Zhang, H. L. Trentelman, and J. M. A. Scherpen, "Fully distributed robust synchronization of net-

worked Lur'e systems with incremental nonlinearities," *Automatica*, vol. 50, no. 10, pp. 2515–2526, Oct. 2014, ISSN: 00051098. DOI: 10.1016/j.automatica.2014.08.033.

[70] J. O. Smith, *Introduction to digital filters with audio applications*. W3K Publishing, 2007, ISBN: 978-0-9745607-1-7.

[71] Z. P. Jiang and Y. Wang, "Input-to-state stability for discrete-time nonlinear systems," *Automatica*, vol. 37, no. 6, pp. 857–869, Jun. 2001, ISSN: 00051098. DOI: 10.1016/S0005-1098(01)00028-0.

[72] A. Pavlov and N. Van De Wouw, "Steady-state analysis and regulation of discrete-time nonlinear systems," *IEEE Transactions on Automatic Control*, vol. 57, no. 7, pp. 1793–1798, 2012, ISSN: 00189286. DOI: 10.1109/TAC.2011.2178790.

[73] A. Pavlov and N. Van de Wouw, "Convergent systems: Nonlinear simplicity," in *Lecture Notes in Control and Information Sciences*, vol. 470, 2017, pp. 51–77, ISBN: 9783319303567. DOI: 10.1007/978-3-319-30357-4_3.

[74] D. N. Tran, B. S. Rüffer, and C. M. Kellett, "Convergence Properties for Discrete-Time Nonlinear Systems," *IEEE Transactions on Automatic Control*, vol. 64, no. 8, pp. 3415–3422, Aug. 2019, ISSN: 15582523. DOI: 10.1109/TAC.2018.2879951. arXiv: 1612.05327.

[75] M. Jungers, N. van de Wouw, and M. Shakib, "Discrete-time Convergent nonlinear systems [Unpublished]," 2021.

[76] V. A. Yakubovich, "Method of matrix unequalities in theory of nonlinear control systems stability. I. Forced oscillations absolute stability," *Avtomat. i Telemekh.*, vol. 25, no. 7, pp. 1017–1029, 1964.

[77] A. Pavlov and N. Van De Wouw, "Convergent discrete-time nonlinear systems: The case of PWA systems," in *Proceedings of the American Control Conference*, 2008, pp. 3452–3457, ISBN: 9781424420797. DOI: 10.1109/ACC.2008.4587027.

[78] D. N. Tran, B. S. Rüffer, and C. M. Kellett, "Incremental stability properties for discrete-time systems," in *2016 IEEE 55th Conference on Decision and Control, CDC 2016*, 2016, pp. 477–482, ISBN: 9781509018376. DOI: 10.1109/CDC.2016.7798314.

[79] J. F. Sturm, "Using SeDuMi 1.02, A Matlab toolbox for optimization over symmetric cones," *Optimization methods and software*, vol. 11, no. 1, pp. 625–653, 2008. DOI: 10.1080/10556789908805766.

[80] S. Theodoridis, "Learning in Parametric Modeling: Basic Concepts and Directions," in *Machine Learning: A Bayesian and Optimization Perspective*, Academic Press, Jan. 2020, pp. 67–120, ISBN: 978-0-12-818803-3. DOI: 10.1016/b978-0-12-818803-3.00012-x.

[81] J. Schoukens, R. Pintelon, T. Dobrowiecki, and Y. Rolain, "Identification of linear systems with nonlinear distortions," in *Automatica*, vol. 41, Pergamon, Mar. 2005, pp. 491–504. DOI: 10.1016/j.automatica.2004.10.004.

[82] M. Locatelli and F. Schoen, *Global optimization: theory, algorithms, and applications*. Philadelphia: Society for Industrial and Applied Mathematics, 2013, ISBN: 9781611972665. DOI: 10.1137/1.9781611972672.

[83] K. Gallagher, M. Sambridge, and G. Drijkoningen, "Genetic algorithms: An evolution from Monte Carlo Methods for strongly non-linear geophysical optimization problems," *Geophysical Research Letters*, vol. 18, no. 12, pp. 2177–2180, 1991, ISSN: 19448007. DOI: 10.1029/91GL02368.

[84] W. L. Price, "A controlled random search procedure for global optimisation," *The Computer Journal*, vol. 20, no. 4, pp. 367–370, Apr. 1977, ISSN: 0010-4620. DOI: 10.1093/comjnl/20.4.367.

[85] E. M. Hendrix, P. M. Ortigosa, and I. García, "On success rates for controlled random search," *Journal of Global Optimization*, vol. 21, no. 3, pp. 239–263, Nov. 2001, ISSN: 09255001. DOI: 10.1023/A:1012387510553.

[86] S. L. Speetjens, J. D. Stigter, and G. van Straten, "Physics-based model for a water-saving greenhouse,"

Biosystems Engineering, vol. 105, no. 2, pp. 149–159, Feb. 2010, ISSN: 15375110. DOI: 10.1016/j.biosystemseng.2009.06.026.

[87] G. C. Karras, C. P. Bechlioulis, M. Leonetti, *et al.*, "On-line identification of autonomous underwater vehicles through global derivative-free optimization," in *IEEE International Conference on Intelligent Robots and Systems*, 2013, pp. 3859–3864, ISBN: 9781467363587. DOI: 10.1109/IROS.2013.6696908.

[88] *Matlab optimization toolbox*, Version 9.0 (R2020b).

[89] R. H. Byrd, M. E. Hribar, and J. Nocedal, "An interior point algorithm for large-scale nonlinear programming," *SIAM Journal on Optimization*, vol. 9, no. 4, pp. 877–900, Jul. 1999, ISSN: 10526234. DOI: 10.1137/S1052623497325107.

[90] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge: Cambridge University Press, 2004, ISBN: 9780521833783. DOI: 10.1017/CBO9780511804441.

[91] G. Sharma and J. Martin, "MATLAB®: A language for parallel computing," *International Journal of Parallel Programming*, vol. 37, no. 1, pp. 3–36, Oct. 2009, ISSN: 08857458. DOI: 10.1007/s10766-008-0082-5.

[92] A. Marconato, J. Sjöberg, J. Suykens, and J. Schoukens, "Identification of the silverbox benchmark using nonlinear state-space models," in *IFAC Proceedings Volumes (IFAC-PapersOnline)*, vol. 16, IFAC Secretariat, 2012, pp. 632–637, ISBN: 9783902823069. DOI: 10.3182/20120711-3-BE-2027.00135.

[93] G. Chen and X. Dong, "On Feedback Control of Chaotic Continuous-Time Systems," *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, vol. 40, no. 9, pp. 591–601, 1993, ISSN: 10577122. DOI: 10.1109/81.244908.

[94] M. A. Nielsen, "A visual proof that neural nets can compute any function," in *Neural Networks and Deep Learning*, Determination Press, 2015, ch. 4.

[95] I. Kollár, J. Schoukens, R. Pintelon, G. Simon, and G. Román, "Extension for the Frequency Domain System Identification Toolbox: Graphical User Interface, Objects, Improved Numerical Stability," *IFAC Proceedings Volumes*, vol. 33, no. 15, pp. 699–702, Jun. 2000, ISSN: 14746670. DOI: 10.1016/s1474-6670(17)39833-6.

[96] J.-P. Noël and M. Schoukens, "F-16 aircraft benchmark based on ground vibration test data," *2017 Workshop on Nonlinear System Identification Benchmarks*, no. April, pp. 19–23, 2017.

[97] C. A. Desoer and M. Vidyasagar, *Feedback Systems: Input-Output Properties*. Society for Industrial and Applied Mathematics, Jan. 2009. DOI: 10.1137/1.9780898719055.

[98] TheoremDep. "Bounding matrix quadratic form using a quadratic eigenvalues form." (2019), [Online]. Available: https://sharmaeklavya2.github.io/theoremdep/nodes/linear-algebra/matrices/bounding-quadratic-form-using-eigenvalues.html (visited on 01/03/2021).

[99] N. Lord, G. H. Golub, and C. F. V. Loan, "Matrix Computations," *The Mathematical Gazette*, vol. 83, no. 498, p. 556, 1999, ISSN: 00255572. DOI: 10.2307/3621013.

[100] J. Gallier, "Schur Complements and Applications," in Springer, New York, NY, 2011, pp. 431–437. DOI: 10.1007/978-1-4419-9961-0_16.

[101] E. K. P. Chong and S. H. Zak, *An Introduction to Optimization: Second Edition*. John Wiley & Sons, 2004, pp. 1–495, ISBN: 9781118033340. DOI: 10.1002/9781118033340.

[102] I. R. Petersen, V. A. Ugrinovskii, and A. V. Savkin, "Robust Control Design Using H-Infinity Methods," Communications and Control Engineering, 2000. DOI: 10.1007/978-1-4471-0447-6.

[103] R. F. Curtain and H. Zwart, *An Introduction to Infinite-Dimensional Linear Systems Theory*, ser. Texts in Applied Mathematics. Springer New York, 1995, vol. 21. DOI: 10.1007/978-1-4612-4224-6.

PO Box 513
5600 MB Eindhoven
The Netherlands
tue.nl

DEPARTMENT OF MECHANICAL ENGINEERING

DYNAMICS & CONTROL

TU/e EINDHOVEN UNIVERSITY OF TECHNOLOGY