

Master's thesis

# Physics-Consistent Surrogate Models with Uncertainty Quantification for Model Predictive Control

Jan-David Ridder

Matr. 217599

First examiner: Prof. Dr.-Ing. Sergio Lucia  
Second examiner: Prof. Dr.-Ing. Hannsjörg Freund  
Advisor: Niklas Kemmerling, M.Sc.  
Year: 2025  
Ident.: BCI-PAS-2025-M01

**Technische Universität Dortmund**  
Faculty of Biochemical and Chemical Engineering  
Laboratory for Process Automation Systems



# Eidesstattliche Versicherung

## (Affidavit)

Name, Vorname  
(surname, first name)

Matrikelnummer  
(student ID number)

☐ Bachelorarbeit  
(Bachelor's thesis)

☐ Masterarbeit  
(Master's thesis)

Titel  
(Title)

Ich versichere hiermit an Eides statt, dass ich die vorliegende Abschlussarbeit mit dem oben genannten Titel selbstständig und ohne unzulässige fremde Hilfe erbracht habe. Ich habe keine anderen als die angegebenen Quellen und Hilfsmittel benutzt sowie wörtliche und sinngemäße Zitate kenntlich gemacht. Die Arbeit hat in gleicher oder ähnlicher Form noch keiner Prüfungsbehörde vorgelegen.

I declare in lieu of oath that I have completed the present thesis with the above-mentioned title independently and without any unauthorized assistance. I have not used any other sources or aids than the ones listed and have documented quotations and paraphrases as such. The thesis in its current or similar version has not been submitted to an auditing institution before.

Ort, Datum  
(place, date)

Unterschrift  
(signature)

### Belehrung:

Wer vorsätzlich gegen eine die Täuschung über Prüfungsleistungen betreffende Regelung einer Hochschulprüfungsordnung verstößt, handelt ordnungswidrig. Die Ordnungswidrigkeit kann mit einer Geldbuße von bis zu 50.000,00 € geahndet werden. Zuständige Verwaltungsbehörde für die Verfolgung und Ahndung von Ordnungswidrigkeiten ist der Kanzler/die Kanzlerin der Technischen Universität Dortmund. Im Falle eines mehrfachen oder sonstigen schwerwiegenden Täuschungsversuches kann der Prüfling zudem exmatrikuliert werden. (§ 63 Abs. 5 Hochschulgesetz - HG - ).

Die Abgabe einer falschen Versicherung an Eides statt wird mit Freiheitsstrafe bis zu 3 Jahren oder mit Geldstrafe bestraft.

Die Technische Universität Dortmund wird ggf. elektronische Vergleichswerkzeuge (wie z.B. die Software „turnitin“) zur Überprüfung von Ordnungswidrigkeiten in Prüfungsverfahren nutzen.

Die oben stehende Belehrung habe ich zur Kenntnis genommen:

### Official notification:

Any person who intentionally breaches any regulation of university examination regulations relating to deception in examination performance is acting improperly. This offense can be punished with a fine of up to EUR 50,000.00. The competent administrative authority for the pursuit and prosecution of offenses of this type is the Chancellor of TU Dortmund University. In the case of multiple or other serious attempts at deception, the examinee can also be unenrolled, Section 63 (5) North Rhine-Westphalia Higher Education Act (*Hochschulgesetz, HG*).

The submission of a false affidavit will be punished with a prison sentence of up to three years or a fine.

As may be necessary, TU Dortmund University will make use of electronic plagiarism-prevention tools (e.g. the "turnitin" service) in order to monitor violations during the examination procedures.

I have taken note of the above official notification:\*

Ort, Datum  
(place, date)

Unterschrift  
(signature)

**\*Please be aware that solely the German version of the affidavit ("Eidesstattliche Versicherung") for the Bachelor's/ Master's thesis is the official and legally binding version.**



# Kurzzusammenfassung

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.



# Abstract

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.





# Contents



# Notation

## Numbers and Arrays

$a$	A scalar (integer or real)
$\mathbf{a}$	A vector
$\mathbf{a}_{[1:N]}$	A sequence of vectors from 1 to $N$ , such that $\mathbf{a}_{[1:N]} = (\mathbf{a}_1, \dots, \mathbf{a}_N)$
$\mathbf{A}$	A matrix
$\mathbf{A}$	A tensor
$\mathbf{I}_n$	Identity matrix with $n$ rows and $n$ columns
$\mathbf{I}$	Identity matrix with dimensionality implied by context
$\mathbf{e}^{(i)}$	Standard basis vector $[0, \dots, 0, 1, 0, \dots, 0]$ with a 1 at position $i$
$\text{diag}(\mathbf{a})$	A square, diagonal matrix with diagonal entries given by $\mathbf{a}$
$a$	A scalar random variable
$\mathbf{a}$	A vector-valued random variable
$\mathbf{A}$	A matrix-valued random variable

## Sets and Graphs

$\mathbb{A}$	A set
$\mathbb{R}$	The set of real numbers
$\{0, 1\}$	The set containing 0 and 1
$\{0, 1, \dots, n\}$	The set of all integers between 0 and $n$
$[a, b]$	The real interval including $a$ and $b$
$(a, b]$	The real interval excluding $a$ but including $b$

$\mathbb{A} \setminus \mathbb{B}$	Set subtraction, i.e., the set containing the elements of $\mathbb{A}$ that are not in $\mathbb{B}$
$\mathcal{G}$	A graph
$Pa_{\mathcal{G}}(\mathbf{x}_i)$	The parents of $\mathbf{x}_i$ in $\mathcal{G}$

## Indexing

$a_i$	Element $i$ of vector $\mathbf{a}$ , with indexing starting at 1.
$a_{-i}$	All elements of vector $\mathbf{a}$ except for element $i$
$A_{i,j}$	Element $i, j$ of matrix $\mathbf{A}$
$\mathbf{A}_{i,:}$	Row $i$ of matrix $\mathbf{A}$
$\mathbf{A}_{:,i}$	Column $i$ of matrix $\mathbf{A}$
$A_{i,j,k}$	Element $(i, j, k)$ of a 3-D tensor $\mathbf{A}$
$\mathbf{A}_{::,i}$	2-D slice of a 3-D tensor
$\mathbf{a}_i$	Element $i$ of the random vector $\mathbf{a}$

## Linear Algebra Operations

$\mathbf{A}^\top$	Transpose of matrix $\mathbf{A}$
$\mathbf{A}^+$	Moore-Penrose pseudoinverse of $\mathbf{A}$
$\mathbf{A} \odot \mathbf{B}$	Element-wise (Hadamard) product of $\mathbf{A}$ and $\mathbf{B}$
$\det(\mathbf{A})$	Determinant of $\mathbf{A}$

## Calculus

$\frac{dy}{dx}$	Derivative of $y$ with respect to $x$
$\frac{\partial y}{\partial x}$	Partial derivative of $y$ with respect to $x$
$\nabla_{\mathbf{x}} y$	Gradient of $y$ with respect to $\mathbf{x}$
$\nabla_{\mathbf{X}} y$	Matrix derivatives of $y$ with respect to $\mathbf{X}$
$\nabla_{\mathbf{x}} y$	Tensor containing derivatives of $y$ with respect to $\mathbf{X}$
$\frac{\partial f}{\partial \mathbf{x}}$	Jacobian matrix $\mathbf{J} \in \mathbb{R}^{m \times n}$ of $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$
$\nabla_{\mathbf{x}}^2 f(\mathbf{x})$ or $\mathbf{H}(f)(\mathbf{x})$	The Hessian matrix of $f$ at input point $\mathbf{x}$

$\int f(\mathbf{x})d\mathbf{x}$	Definite integral over the entire domain of $\mathbf{x}$
$\int_{\mathbb{S}} f(\mathbf{x})d\mathbf{x}$	Definite integral with respect to $\mathbf{x}$ over the set $\mathbb{S}$

## Probability and Information Theory

$a \perp b$	The random variables $a$ and $b$ are independent
$a \perp b \mid c$	They are conditionally independent given $c$
$P(a)$	A probability distribution over a discrete variable
$p(a)$	A probability distribution over a continuous variable, or over a variable whose type has not been specified
$a \sim P$	Random variable $a$ has distribution $P$
$\mathbb{E}_{\mathbf{x} \sim P}[f(x)]$	Expectation of $f(x)$ with respect to $P(\mathbf{x})$
$\text{Var}(f(x))$	Variance of $f(x)$ under $P(\mathbf{x})$
$\text{Cov}(f(x), g(x))$	Covariance of $f(x)$ and $g(x)$ under $P(\mathbf{x})$
$H(\mathbf{x})$	Shannon entropy of the random variable $\mathbf{x}$
$D_{\text{KL}}(P \parallel Q)$	Kullback-Leibler divergence of $P$ and $Q$
$\mathcal{N}(\mathbf{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma})$	Gaussian distribution over $\mathbf{x}$ with mean $\boldsymbol{\mu}$ and covariance $\boldsymbol{\Sigma}$

## Functions

$f : \mathbb{A} \rightarrow \mathbb{B}$	The function $f$ with domain $\mathbb{A}$ and range $\mathbb{B}$
$f \circ g$	Composition of the functions $f$ and $g$
$f(\mathbf{x}; \boldsymbol{\theta})$	A function of $\mathbf{x}$ parametrized by $\boldsymbol{\theta}$ . (Sometimes we write $f(\mathbf{x})$ and omit the argument $\boldsymbol{\theta}$ to lighten notation)
$\log x$	Natural logarithm of $x$
$\sigma(x)$	Logistic sigmoid, $\frac{1}{1 + \exp(-x)}$
$\zeta(x)$	Softplus, $\log(1 + \exp(x))$
$\ \mathbf{x}\ _p$	$L^p$ norm of $\mathbf{x}$
$\ \mathbf{x}\ $	$L^2$ norm of $\mathbf{x}$
$x^+$	Positive part of $x$ , i.e., $\max(0, x)$

## Datasets and Distributions

$p_{\text{data}}$	The data generating distribution
$\hat{p}_{\text{data}}$	The empirical distribution defined by the training set
$\mathbb{X}$	A set of training examples
$\boldsymbol{x}^{(i)}$	The $i$ -th example (input) from a dataset
$y^{(i)}$ or $\boldsymbol{y}^{(i)}$	The target associated with $\boldsymbol{x}^{(i)}$ for supervised learning
$\boldsymbol{X}$	The $m \times n$ matrix with input example $\boldsymbol{x}^{(i)}$ in row $\boldsymbol{X}_{i,:}$

# List of Figures





# List of Tables



# Chapter 1

## Introduction

In the chemical process industry, system models are a fundamental tool for design and operation. Due to its techno-scientific nature, chemical engineering is about modelling physical and chemical phenomena in an accurate enough manner to realize a save and profitable conversion process. These models vary in complexity and prediction potency from single equations such as the Bernoulli equation to detailed partial differential equation systems such as multi-phase Navier-Stokes. Physico-chemical models are at the core of optimal design or optimal control of most unit operations.

As these *first principle models* rely on intrinsic parameters such as activation energies, friction coefficients, etc. that are usually fitted by experiments, they exhibit some kind of prediction uncertainty. This uncertainty is heavily tied to the measurement uncertainty of the underlying experiments and the non-linearity in the propagation by the model. For example parametric uncertainties of  $\pm 50\%$  can occur in multi-component reaction systems [1]. As most phenomena in chemical engineering exhibit strong non-linearities, such as reaction kinetics or activity coefficients, uncertainty quantification represents a great challenge.

*Model Predictive Control* (MPC) is a potent optimal control algorithm that allows the optimal operation of various systems. The algorithm optimizes a custom loss or reward function under a set of constraints and the dynamics of the system model. The optimal inputs of the system are determined at every time instant as the solution of the optimization. MPC is able to deal with uncertain parameter scenarios in the form of *multi-stage* MPC, where all possible scenarios are optimized over simultaneously. This leads to tremendous computational effort especially for detailed system models. The computational effort grows multiplicatively with the amount of parameters and exponentially with the robust horizon [1]. These optimization tasks are impossible to solve in real time with current hardware.

*Neural Networks* (NNs) are able to approximate any continuous function that maps from a finite dimensional space to another up to an arbitrary small error [1]. Thus, NNs are a key component of machine learning advancements of the 2010s enabling data approximation of almost any kind such as autoregressive language generation [2], protein folding [1] or reduction of partial differential equations (PDEs). Their universal approximation ability, differentiability and fast evaluation makes NNs greatly suitable to reduce first principle models in the context of optimal control.

For NNs to be applied successfully and safely in process engineering, another measure is of great importance – The alignment of predictions with physical conservation laws such

as mass or energy. By definition, NN predictions seek to minimize a loss function to the ground truth data. The alignment of the predictions with real balance equations is only achieved until test accuracy. Deviations accumulate especially in propagation scenarios where predictions are propagated through different networks or through a time horizon in autoregressive manner. This leads to bad predictions or even unstable behavior in late propagation stages and makes MPC implementations challenging.

This work investigates the usage of physics-consistent (PC) NNs in the context of robust model predictive control under parametric uncertainty. The NNs in focus shall significantly reduce the computational effort due to different parameter scenarios as well as guarantee physical correctness of the predictions over the whole prediction horizon. To account for uncertainties, quantile regression is utilized which captures a prediction confidence interval. PC is enforced by implementing a recent type of activation function that elegantly maps a prediction into the allowed output space without altering the inference time [3].

The developed surrogate models are later evaluated regarding their simulation accuracy using a fictional plug flow tubular reactor (PFTR) model of the ethene oxide synthesis.

## Chapter 2

# Model Predictive Control under Uncertainty

Before diving into the application of a special model predictive (MPC) algorithm. Uncertain optimization is treated more generally. From economics, over self-driving cars to chemical plants, optimization problems are formulated in terms of uncertain variables in countless variations. These uncertainties arise in three different types – fluctuating disturbances such as wind or weather conditions, parametric offsets such as wrongly determined friction coefficients or model mismatches like unseen side reactions. As the straight maximization or minimization of an objective often puts a problem to its boundaries, unforeseen disturbances may lead to *constraint violation* or even *instability* [4]. Therefore, *robust* optimization formulations are needed, that are proof against mentioned challenges as well as retain the *performance* with respect to the objective function.

### 2.1 Robust Optimization

Considering the following optimization task (??) under an uncertainty vector  $\mathbf{w}$ .

**Definition 2.1.** Semi-infinite program with the uncertainty  $\mathbf{w}$ .

$$\begin{aligned} \min_{\mathbf{u}} \max_{\mathbf{w}} \quad & F_0(\mathbf{u}, \mathbf{w}) \\ \text{s.t.} \quad & F_i(\mathbf{u}, \mathbf{w}) \leq 0 \quad \forall \mathbf{w} \in \mathbb{W} \end{aligned}$$

Here, the optimization of the objective is formulated in a bilevel form. The scalar objective function  $F_0$  is to be minimized w.r.t the decision variables  $\mathbf{u}$  for the worst case scenario w.r.t the disturbance vector  $\mathbf{w}$ . It can be interpreted as a player-vs-player setup [4]. One player (the optimizer) seeks a minimum solution for the input  $\mathbf{u}$ , while the opposing player (nature) chooses a maximum solution for the disturbance variables  $\mathbf{w}$ . The inequality constraints  $F_i$  in addition, should be satisfied for all possible disturbances, that lie in a defined set  $\mathbb{W}$ . If this set is not finite, the number of possible constraints to be satisfied is infinitely big. Thus, this formulation is named as *semi-infinite-program* [5]. To reduce the number of constraints, the worst-case estimate can be applied here (??).

**Definition 2.2.** Bi-level formulation of the semi-infinite program

$$\begin{aligned} & \min_{\mathbf{u}} \max_{\mathbf{w}} F_0(\mathbf{u}, \mathbf{w}) \\ \text{s.t. } & \varphi_i(\mathbf{u}) := \max_{\mathbf{w}} F_i(\mathbf{u}, \mathbf{w}) \leq 0 \quad \mathbf{w} \in \mathbb{W} \end{aligned}$$

Now, the inequality constraints are also formulated in a bi-level manner. The maximum of  $F_i$  is constant in  $\mathbf{w}$  and only a function of the decision variables  $\mathbf{u}$ . The function  $\varphi_i$  exactly describes the worst-case for  $F_i$  parametrized by  $\mathbf{u}$ . In fact, if certain assumptions hold regarding the uncertainty set  $\mathbb{W}$  and the concavity of the maximization,  $\varphi_i$  can be exactly computed [6]. This would simplify the problem dramatically. In many real cases however, an analytical solution for  $\varphi_i$  is impossible to compute. But with the universal approximation capability of neural networks (NNs) in mind [7], a data-driven approach is promising, that delivers an approximated function  $\tilde{\varphi}_i(\mathbf{u})$ .

## 2.2 Model Predictive Control

Model predictive control (MPC) is a type of optimization, in which an optimal control problem is solved in a dynamic manner. The main goal of this control strategy is to find the best inputs to a system that maximize or minimize a defined objective. In a chemical process, for instance, the energy consumption for heating should be minimized. It should be minimized in such a way, that other variables remain within their bounds. For example, the product specification and, most importantly, the safety of the process needs to be ensured under all circumstances. In contrast to the *optimal design* of a process, that is done during the engineering phase, an MPC controller operates *online*. It acts on the process while it is running. Thus, it is coined *dynamic optimization*. To do so, the optimization problem needs to be solved recursively at given time instants. Chemical plants usually have large *time constants*, meaning, they react slowly to their inputs. Other physical systems with low time constants, like quadcopters behave much faster. So, the computation time of an MPC algorithm is to be kept low [8] to be able to react to the system. To predict the future behavior of the system, MPC requires a system model.

In total, an MPC controller is based on three pillars – an objective function, a set of constraints and, crucially, the model of the system.

**Definition 2.3.** Continuous-time formulation of an MPC problem.

$$\begin{aligned} & \min_{\mathbf{u}} J(t_0, t_{\text{end}}, \mathbf{x}, \mathbf{u}) \\ \text{s.t. } & \dot{\mathbf{x}} = f(t, \mathbf{x}, \mathbf{u}) \\ & F_i(t, \mathbf{x}, \mathbf{u}) \leq 0 \quad \forall t \in [t_0, t_{\text{end}}] \\ & \mathbf{x}_0 = \mathbf{x}_{\text{init}} \end{aligned}$$

Optimal system inputs  $\mathbf{u}$  are to be determined, that minimize the objective function  $J$ . This objective varies with the inputs and the states of the system  $\mathbf{x}$ . It is to be minimized over a time horizon from  $t_0$  to  $t_{\text{end}}$ , called the *prediction horizon*. The states of the system evolve along the dynamics of the system, which are described by the system model  $f(t, \mathbf{x}, \mathbf{u})$ . Thus, the quality of the controller heavily depends on the prediction accuracy of this model [9]. The system states however should remain in a

set of constraints at all times whereas the inputs are bounded, all summarized by the constraint functions  $F_i$ . In a physical sense, constraint violation can lead to instability and thus damage of the real system. Here is a vital example. For a chemical reactor a *reaction runaway* must be avoided at all costs. This happens, if the exothermic heat release exceeds the maximum possible cooling rate. Higher temperatures increase the reaction rate exponentially leading to a greater enthalpy release, which again causes a self-accelerating temperature rise [10]. That is why, stability and guaranteed mathematical constraint satisfaction are a great field of research in the discipline of optimal control.

## 2.3 Multi-Stage Model Predictive Control

Coming from this general perspective on robust optimization and basic MPC, *robust MPC* specializes *robustness* to a dynamic closed-loop control environment.

**Definition 2.4.** Continuous-time formulation of an MPC task under uncertainty.

$$\begin{aligned} & \min_{\mathbf{u}} \max_{\mathbf{w}} J(t_0, t_{\text{end}}, \mathbf{x}, \mathbf{u}, \mathbf{w}) \\ \text{s.t. } & \dot{\mathbf{x}} = f(t, \mathbf{x}, \mathbf{u}, \mathbf{w}) \\ & F_i(t, \mathbf{x}, \mathbf{u}, \mathbf{w}) \leq 0 \quad \forall t \in [t_0, t_{\text{end}}], \forall \mathbf{w} \in \mathbb{W} \\ & \mathbf{x}_0 = \mathbf{x}_{\text{init}} \end{aligned}$$

The only difference to the basic MPC formulation (??) is that the model, the objective and the constraints are functions of the uncertainty, denoted with  $\mathbf{w}$ . This MPC formulation (def. ??) is now infinite in time and in the uncertain parameters. To make computation feasible, the problem can be discretized in the time dimension either using single shooting, multiple shooting or orthogonal collocation on finite elements [11]. One possibility to roughly discretize the uncertain parameter set  $\mathbb{W}$  is done by *multi-stage-MPC* [12].

It generates a scenario tree (fig. ??) for all possible combinations of disturbances. In the algorithm design phase, scenario cases  $w_k^{(\cdot)}$  must be defined a priori for each independent disturbance that may act on the system. As different disturbances occur in different combinations, each possible scenario is represented by a branch. At every time step in the *robust horizon*, branches for every possible uncertainty scenario are generated. This procedure is coined *scenario generation*. In practice, the worst-cases for the disturbances are assumed and sometimes an additional expected or nominal value. Mostly, this yields good results with respect to controller performance and robustness [9].

If the uncertainty is discrete and all cases are considered, the scenario tree formulation (fig. ??) is exact. In most applications however, when the uncertainties lie in a continuous space, the scenario tree is an extremely rough approximation. With a non-linear system, the scenarios are not monotonous regarding the disturbances. This means, even if the scenario branches cover the extreme values of the disturbances themselves, they can leave extreme scenarios caused by intermediate parameter combinations. Robust constraint satisfaction is not guaranteed and must be checked via a reachability

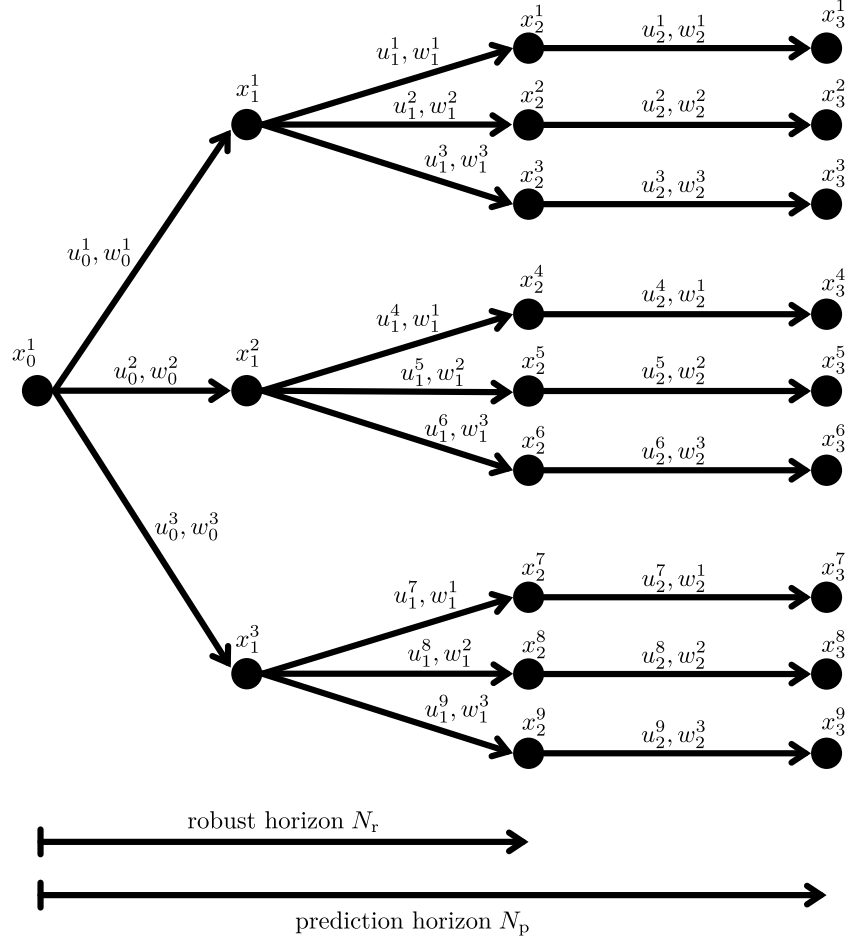


Figure 2.1: Multi-stage MPC discretizes the uncertainty set via user-defined possible cases  $w_k^{(\cdot)}$ . All possible case combinations are branched through the robust horizon  $N_r$ . To limit the complexity, the uncertainties are considered constant thereafter for the rest of the prediction horizon  $N_p$ .



analysis [12]. This can be an obstacle in safety-relevant applications. Another drawback of multi-stage MPC is, that it includes redundant scenarios not causing extreme cases due to the non-monotonicity. These scenarios still are part of the optimization leading to unnecessary complexity. On top of this, the number of optimization variables grows exponentially with the *robust horizon*  $N_r$ , the number of parameters and their respective values (eq. ??) [1].

$$n_s = \left( \prod_i^{n_w} \nu_i \right)^{N_r} \quad (2.1)$$

Here,  $n_w$  denotes the number of different disturbances and  $\nu_i$  their respective discrete values.  $n_s$  describes the number of scenarios in the scenario tree (fig. ??). High order uncertainties lead to optimization problems that are simply impossible to solve in real time.

## 2.4 Neural Networks

*Feed forward neural networks* (NNs), sometimes referred to as *multi-layer-perceptrons* (MLPs), provide a stunning capability of approximating any multivariate function from any finite dimensional space to another. According to the *Universal Approximation Theorem*, any finite dimensional function can be reproduced if the number of hidden layers and respective neurons is high enough, up to an arbitrarily small error [7]. Thus, NNs play a major role in recent developments in machine learning. The breakthrough of large language models such as ChatGPT heavily rely on neural networks [2]. NNs are not only successfully applied in natural language processing but in the chemical process industry. Applications range from thermodynamic property prediction over unit operation design to corporate level decision making [13]. Their ability to approximate any non-linear function while being differentiable, makes NNs especially suitable for dynamic optimization such as MPC. In contrast to detailed first principle models, NNs have the advantage of making a prediction in a fraction of the inference time of the first principle model. As MPC operates in real time parallel to the system to be controlled, the computation duration is a key measure for a successful MPC implementation. That is why, surrogate models in form of NNs are a crucial tool for dynamic optimization and heavily investigated in current research [14]. May a NN be given as an alternating series of function evaluations (def. ??).

**Definition 2.5.** Mathematical formulation of a neural network as a nested function evaluation.

$$\begin{aligned} \mathcal{NN}(\boldsymbol{\xi}, \boldsymbol{\theta}) &= g_{l+1} \circ h_{l+1} \circ \dots \circ g_1 \circ h_1(\boldsymbol{\xi}) \\ h(\mathbf{a}_{l-1}) &= \boldsymbol{\theta}_{wl} \mathbf{a}_{l-1} + \boldsymbol{\theta}_{bl} \\ \mathbf{a}_l &= g_l(h(\mathbf{a}_{l-1})) \end{aligned}$$

In fact, a NN is a series of function combinations denoted by the  $\circ$ -symbol alternating between a linear function  $h(\mathbf{a})$  and a *non-linear* activation function  $g(h(\mathbf{a}))$ . The linear transformation matrices  $\boldsymbol{\theta}_{wl}$  and offsets  $\boldsymbol{\theta}_{bl}$ , are called weights and biases respectively and are the trainable parameters of a NN. The activation functions  $g$  are important to introduce non-linearity into every linear projection. Otherwise, the complete function combination can be written as one singular linear transformation and the universal

approximation capability is lost. Given a set of training data of  $m$  data points, the optimal parameters for  $\boldsymbol{\theta}$  are found by solving an optimization task that involves the minimization of a prediction loss function (eq. ??).

$$\begin{aligned}\mathbb{X}_{\text{train}} &:= \{\boldsymbol{\xi}^{(0)}, \dots, \boldsymbol{\xi}^{(m-1)}\} \\ \mathbb{Y}_{\text{train}} &:= \{\mathbf{y}^{(0)}, \dots, \mathbf{y}^{(m-1)}\} \\ \mathcal{L}(\boldsymbol{\theta}, \mathbb{X}_{\text{train}}, \mathbb{Y}_{\text{train}}) &= \frac{1}{m} \sum_{i=0}^{m-1} \left\| \mathbf{y}^{(i)} - \tilde{\mathbf{y}}(\boldsymbol{\theta}, \boldsymbol{\xi}^{(i)}) \right\|_2^2\end{aligned}\tag{2.2}$$

In *supervised learning*, the training data is separated in *features*  $\mathbb{X}_{\text{train}}$ , the inputs of the NN, and *labels*  $\mathbb{Y}_{\text{train}}$ , the outputs of the NN. During training, the parameters  $\boldsymbol{\theta}$  are determined such that the NN approximates the data in best possible manner. In this example the loss function  $\mathcal{L}$  portrays the mean squared error (MSE), that is used in most regression tasks. The difference between the model prediction, denoted as  $\tilde{\mathbf{y}}(\boldsymbol{\xi}^{(i)})$ , and the true data point  $\mathbf{y}^{(i)}$  is penalized quadratically.

### 2.4.1 Quantile Regression

By utilizing a loss function in form of the MSE, a model is trained to achieve regression predictions in form of the quarred mean. To tweak the prediction characteristics of a NN for the task at hand, the type of training loss function is decisive. In a probabilistic environment, the prediction of confidence intervalls in the form of quantiles may be of mayor interest. To learn the quantiles of a distribution of data points, the *quantile loss* or *pinball loss* can be used (eq. ??).

$$\mathcal{L}_\tau = \begin{cases} \tau \cdot (y - \tilde{y}(\boldsymbol{\theta})), & y \geq \tilde{y}(\boldsymbol{\theta}) \\ (1 - \tau) \cdot (\tilde{y}(\boldsymbol{\theta}) - y), & y < \tilde{y}(\boldsymbol{\theta}) \end{cases}\tag{2.3}$$

The hyperparameter  $\tau$  denotes the type and therefore position of the quantile. A small value for  $\tau$ , e.g. 0.05 or 0.1, strongly penalizes data points that are below the threshold  $y < \tilde{y}(\boldsymbol{\theta})$  via the  $(1 - \tau)$ -term to keep the quantile small. A high value for  $\tau$ , e.g. 0.95, focuses on the penalization of data points above the threshold  $y \geq \tilde{y}(\boldsymbol{\theta})$  to cover a greater amount of points. For robust optimization, *quantile regression* can be applied to learn upper or lower bounds for the uncertainty-dependent constraints  $F_i(\mathbf{u}, \mathbf{w})$  (eq. ??) in terms of confidence quantiles. This would reduce the semi-infinite program to an uncertainty-finite-program.

## 2.5 NARX Models

Because of their universal approximation capability [7] and their differentiability, NNs are well suited as *surrogate models* in MPC algorithms to approximate the system behavior represented by the function  $f$  (def. ??). This can significantly reduce the solution time in contrast to discretized systems of ordinary differential equations (ODEs) [8].

**Definition 2.6.** Time-discrete form of a NN to predict the future state, based on past information.

$$\begin{aligned}\mathbf{x}_{k+1} &= f(\mathbf{x}_k, \mathbf{u}_k) \\ \tilde{\mathbf{x}}_{k+1} &= \mathcal{NN}(\mathbf{x}_k, \mathbf{x}_{k-1}, \dots, \mathbf{x}_{k-h}, \mathbf{u}_k, \mathbf{u}_{k-1}, \dots, \mathbf{u}_{k-h})\end{aligned}$$

In the discrete-time notation, the state at the next time instant  $\mathbf{x}_{k+1}$  is calculated by propagating the current state  $\mathbf{x}_k$  through a function  $f(\mathbf{x}_k, \mathbf{u}_k)$  that describes the system dynamic (def. ??). In the original MPC formulation (def. ??) this is some numerical integration scheme for a system of differential equations [11] with the abbreviation *right-hand-side* (rhs). The state propagation by the function  $f$  is a classical example of a time series extrapolation. NNs have shown significant capabilities in time series forecasting. Deep learning based architectures exist for time series forecasting in various complexities. Convolutional NNs (CNNs), Long-Short-Term-Memory (LSTMs) cells and transformers are the most prominent for time series prediction with varying complexity [2], [15]. These provide a great possibility to apply deep learning based forecasting within optimal control algorithms, where the next state  $\tilde{\mathbf{x}}_{k+1}$  is the output of a neural network  $\mathcal{NN}$ . In context of this work, shallow feedforward NNs or *single-layer perceptrons* (SLPs) are used for this purpose which technically do belong to *shallow learning* architectures.

In the realm of control, the non-linear system identification is summarized with the term *Non-linear AutoRegressive with exogenous inputs* (NARX). Here, the term *autoregressive* means, that the future prediction is a function of a series of past system states  $\{\mathbf{x}_k, \mathbf{x}_{k-1}, \dots, \mathbf{x}_{k-h}\}$ , while *exogenous* implies, that a series of outside inputs  $\{\mathbf{u}_k, \mathbf{u}_{k-1}, \dots, \mathbf{u}_{k-h}\}$  is passed to the function as well (def. ??). The hyperparameter  $h$  denotes the length of the time horizon of the past series. A high value for this parameter includes more past information, which possibly allows better predictions while increasing the network input. Between states and inputs,  $h$  can be chosen differently. In this work,  $h$  is set equal for states and inputs.

### 2.5.1 Input Compression

$$\phi_k := (\mathbf{x}_k, \mathbf{x}_{k-1}, \dots, \mathbf{x}_{k-h}, \mathbf{u}_k, \mathbf{u}_{k-1}, \dots, \mathbf{u}_{k-h}) \quad (2.4)$$

The concatenated input feature vector  $\phi_k$  contains all past states and inputs within the time horizon  $h$ , leading to a vector length of  $(n_x + n_u) \cdot (h + 1)$ , where  $n_x$  and  $n_u$  represent the number of states and inputs respectively. For large systems, the input length increases the network size unacceptably quick, while the input time series often exhibit strong redundancies. To reduce the input size, the feature vector  $\phi_k$  is passed to an *encoder* [16]. The encoder is based on feature compression via *principal component analysis* (PCA), that maps the feature space to a lower dimensional subspace. This subspace represents the space of the input data with the most information.

### 2.5.2 Singular Value Decomposition

*Singular value decomposition* (SVD) is a fundamental building block in linear algebra. Its applications range from image compression, over exploratory data analysis to model reduction in fluid mechanics. The power of SVD is, that it maps a high dimensional space into a low dimensional subspace. The mapping is achieved by a simplistic matrix transformation. It not only identifies a subspace, it also identifies underlying patterns in the data encoded via the respective *singular vectors*. The compression and component identification is guaranteed to exist for every matrix  $\mathbf{X} \in \mathbb{C}^{n \times m}$  in contrast to an eigendecomposition [17].

The SVD is extremely useful, when high dimensional data is presented, that embeds some kind of pattern. For example, simulations of partial differential equations (PDEs), images as collections of pixels and time serieses exhibit strong patterns within the data. This is especially true for data points in close temporal or spatial distance. Because the SVD is the basis for PCA to compress the feature vector  $\phi_k$  and *model order reduction* (MOR) of ODEs or PDEs it is explained in further detail. According to the SVD, any matrix  $\mathbf{X} \in \mathbb{C}^{n \times m}$  can be decomposed into three submatrices (th. ??).

**Theorem 2.1.** *Singular value decomposition of a data matrix  $\mathbf{X}$ .*

$$\mathbf{X} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^*$$

Here, the matrices  $\mathbf{U} \in \mathbb{C}^{n \times n}$  and  $\mathbf{V}^* \in \mathbb{C}^{m \times m}$  are quadratic, orthonormal matrices. The real, non-negative, diagonal matrix  $\mathbf{\Sigma} \in \mathbb{R}^{n \times m}$  contains the so called *singular values* of the matrix with the biggest as the top-left entry. The notation  $\mathbf{V}^*$  stands for the transposed complex-conjugate of  $\mathbf{V}$ . The computation in `python` can be done in `numpy` via the method `linalg.svd()` or directly in a training environment with `torch.linalg.svd()`. For large data matrices  $\mathbf{X}$ , that do not fit into memory, a batched or incremental SVD must be considered. A current `torch` implementation with GPU support exists under `torch_incremental_pca` [18]. The compression itself is done by truncating the yielded matrices by taking the first  $r$  columns of  $\mathbf{U}$ , the first  $r$  rows of  $\mathbf{V}^*$  and  $r$  diagonal elements of  $\mathbf{\Sigma}$  (eq. ??).

$$\begin{aligned} \mathbf{X} &\approx \mathbf{\Psi}\tilde{\mathbf{\Sigma}}\tilde{\mathbf{V}}^* \\ \text{with } \mathbf{\Psi} &:= \mathbf{U}_{:,r} \quad \tilde{\mathbf{\Sigma}} = \mathbf{\Sigma}_{:r,:r} \quad \tilde{\mathbf{V}}^* = \mathbf{V}_{:r,:}^* \end{aligned} \quad (2.5)$$

This has the advantage, that an approximation of the data  $\mathbf{X}$  can be stored via the truncated matrices  $\mathbf{\Psi}, \tilde{\mathbf{\Sigma}}, \tilde{\mathbf{V}}^*$  which take up less memory than the original matrix. It also allows to project the original matrix into a subspace of dimension  $r$  (eq. ??).

$$\tilde{\mathbf{X}} = \mathbf{\Psi}^\top \mathbf{X} \iff \mathbf{X} \approx \mathbf{\Psi}\tilde{\mathbf{X}} \quad (2.6)$$

The truncated matrix  $\mathbf{\Psi}$  acts as the transformation matrix between the low dimensional space and the original space (eq. ??). It is important to note, that the back transformation into the original space only yields an approximation of the original data. It can be proven, that this approximation is the best existing one with respect to the Frobenius Norm as it solves following minimization [17] (eq. ??).

$$\begin{aligned} \mathbf{\Psi} &= \underset{\mathbf{\Psi}}{\operatorname{argmin}} \quad \|\mathbf{X} - \mathbf{\Psi}\mathbf{\Psi}^\top \mathbf{X}\|_F \\ \text{s.t.} \quad &\operatorname{rank}(\mathbf{\Psi}) = r \end{aligned} \quad (2.7)$$

Reentering the concept of neural network NARX models, the large vector of input features  $\phi_k$  (eq. ??) can be compressed via SVD into a low dimensional feature vector  $\xi_k$  that retains the most dominant basis vectors of the original input (fig. ??). This is done by an *PCA encoder* placed before the first layer of a NARX model to reduce the network size.

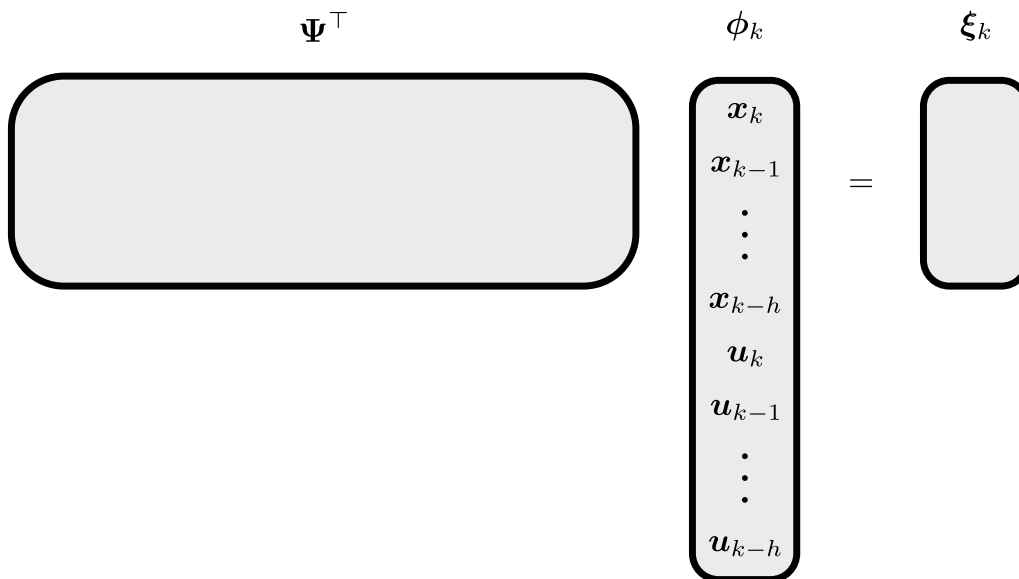


Figure 2.2: The PCA encoder uses SVD to derive a projection matrix  $\Psi$  that maps the high dimensional feature vector  $\phi_k$  to a lower dimensional feature vector  $\xi_k$  which is fed into a NARX model. This reduces the network size significantly, while retaining the most important information about the input.

## 2.6 Conservation of Atoms

The derived methodology will later be applied to control a *packed-bed tubular reactor* under parametric uncertainty which is a common realization for many chemical processes. For this sake, the *conservation of atoms* during chemical reactions is introduced. Similar to the conservation of mass or energy, the conservation of atoms demands, that all atoms that enter a chemical reaction must also leave this reaction. While they change the species they form, the atoms themselves are conserved. This can be expressed as follows.

**Theorem 2.2.** *Conservation of atoms during a chemical reaction.*

$$B\Delta n = 0$$

Here,  $B$  denotes the *element-species-matrix* and  $\Delta n$  the changes in the molar amounts of all participating specieses in a reaction network. The sum of all atoms in all species must be equal at all times regardless of how the underlying reactions take place. This builds the foundation of reaction network analysis and the determination of key and non-key-components [10].

$$B = \begin{pmatrix} \beta_{1,1} & \dots & \beta_{1,\alpha} \\ \vdots & \ddots & \vdots \\ \beta_{h,1} & \dots & \beta_{h,\alpha} \end{pmatrix} \quad (2.8)$$

The element-species-matrix contains the elements  $h$  as rows and the species  $\alpha$  as columns.  $\beta_{h,\alpha}$  describes the number of atoms of element  $h$  in the species  $\alpha$ . This implies, that when the initial molar amounts of all species are known, all possible combinations of molar amounts must lie in a pre-defined space.

## Chapter 3

# Surrogate Modelling

### 3.1 Physics-Constrained Neural Networks

By definition, NNs are designed to reproduce their training data in a best possible manner, defined by the characteristics of the loss function used during the training process. NN predictions minimize the loss towards the training data seen beforehand. Circumstances exist, in which the output of a NN should strictly satisfy a certain set of equations. In case of physical modeling, the prediction of a surrogate should follow the same conservation laws as a first principle model. Especially, when predictions are propagated through several NNs or simulated in an autoregressive manner, errors accumulate. Final predictions are untrustworthy or simulations exhibit instability [3]. As a MPC algorithm needs a reasonable approximation of the system behavior over the complete prediction horizon, the accumulation of conservation violations become an issue.

To realize physics consistent NN outputs, several strategies exist. Physics-informed neural networks (*PINNs*) for instance, embed physical balance equations into the loss function similar to a regularization term [19]. This serves as a soft constraint and comes with two downsides. The NN output does not strictly follow the constraint equation, because the physical consistency is fulfilled up to the trade off between data fit and constraint satisfaction. Every new balance equation to be included introduces a new hyperparameter worsening the ability to include network parameter regularization. To guarantee strict constraint satisfaction and thus perfect physics consistency up to machine precision, a new type of activation function is presented equivalently to the work by Chen et al. [3]. This activation function serves as a linear projection from the NN output space into the physical allowed subspace without adding neither learnable parameters nor hyperparameters. Let the output of a NN be defined as well as a set of linear equality constraints that summarize some physical conservation quantity (def. ??).

**Definition 3.1.** Crude NN output with a given set of linear equality constraints.

$$\begin{aligned}\tilde{\mathbf{y}}(\boldsymbol{\theta}, \boldsymbol{\xi}) &:= \mathcal{NN}(\boldsymbol{\theta}, \boldsymbol{\xi}) \\ \mathbf{A}(\mathbf{y} - \mathbf{z}_0) &= \mathbf{b} \\ \text{with } \tilde{\mathbf{y}}(\boldsymbol{\theta}, \boldsymbol{\xi}), \mathbf{z}_0 &\in \mathbb{R}^{n_{\text{out}}} \quad \mathbf{A} \in \mathbb{R}^{n_c \times n_{\text{out}}} \quad \mathbf{b} \in \mathbb{R}^{n_c}\end{aligned}$$

$\tilde{\mathbf{y}}$  denotes the NN output,  $\boldsymbol{\xi}$  the input and the matrices  $\mathbf{A}$  and  $\mathbf{b}$  portray the feasible space of points. The variable  $\mathbf{z}_0$  illustrates an external input into the constraint. This

may be a NN input, a constant, or another batched reference point. In a physical sense,  $\mathbf{z}_0$  is often introduced as a boundary or initial condition. A feasible point  $\mathbf{y}$  now needs to be determined that fullfills the set of linear equality constraints. For a straight forward implementation, the formulation appears in such a way that it supports batching of a certain batch size  $n_B$ .  $n_{\text{out}}$  and  $n_c$  describe the number of output features and numbers of linear equality constraints respectively. The main idea is, to find a point in the feasible subspace that is closest to the NN prediction. This is forced by following quadratic programm (QP) (def. ??).

**Definition 3.2.** Minimization of the euclidian distance towards the NN output as a linear equality constrained optimization that supports batching.

$$\begin{aligned} \mathbf{y}^* &= \underset{\mathbf{y}}{\operatorname{argmin}} \|\mathbf{y} - \tilde{\mathbf{y}}\|_2^2 \\ \text{s.t. } & (\mathbf{y} - \mathbf{z}_0)\mathbf{A}^\top - \mathbf{b} = \mathbf{0} \end{aligned}$$

This means, the output of the network  $\tilde{\mathbf{y}}$ , the optimal solution  $\mathbf{y}^*$  and  $\mathbf{z}_0$  are matrices of dimension  $n_B \times n_{\text{out}}$ . Computing frameworks that support broadcasting such as `numpy` or `pytorch` allow the implementation of  $\mathbf{z}_0$  in the shape of  $1 \times n_{\text{out}}$  to reduce memory usage as long as it remains constant [1]. The transposed constraint matrix  $\mathbf{A}^\top$  must be of shape  $n_c \times n_B$ . The vector  $\mathbf{b}$  represents the offset of the linear equality constraints. This optimization task can be solved analytically by setting up the *Karush-Kuhn-Tucker* (KKT) conditions (th. ??).

**Theorem 3.1.** *Langrangian and KKT conditions of the QP*

$$\begin{aligned} \mathcal{L}(\mathbf{y}, \boldsymbol{\nu}) &= \|\mathbf{y} - \tilde{\mathbf{y}}\|_2^2 + \boldsymbol{\nu}^\top (\mathbf{A}(\mathbf{y} - \mathbf{z}_0)^\top - \mathbf{b}) \\ \nabla_{\mathbf{y}} \mathcal{L}(\mathbf{y}, \boldsymbol{\nu}) &= 2(\mathbf{y}^* - \tilde{\mathbf{y}}) + \boldsymbol{\nu}^\top \mathbf{A} \stackrel{!}{=} \mathbf{0} \\ \nabla_{\boldsymbol{\nu}} \mathcal{L}(\mathbf{y}, \boldsymbol{\nu}) &= (\mathbf{y}^* - \mathbf{z}_0)\mathbf{A}^\top - \mathbf{b} \stackrel{!}{=} \mathbf{0} \end{aligned}$$

The Lagrangian  $\mathcal{L}$  arises as the sum of the objective function and all linear equality constraints. The vector  $\boldsymbol{\nu}$  contains the dual variables of the linear equality constraints. The optimality conditions lead to a linear system of equations (cor. ??).

**Corollary 3.1.** *The optimal solution in the feasible space as projection of an activation function*

$$\begin{aligned} \begin{pmatrix} \mathbf{y}^* & (\boldsymbol{\nu}^*)^\top \end{pmatrix} \begin{pmatrix} 2\mathbf{I} & \mathbf{A}^\top \\ \mathbf{A} & \mathbf{0} \end{pmatrix} &= \begin{pmatrix} 2\tilde{\mathbf{y}} & \mathbf{z}_0\mathbf{A}^\top + \mathbf{b} \end{pmatrix} \\ g_{pc}(\tilde{\mathbf{y}}, \mathbf{z}_0) &:= \begin{pmatrix} 2\tilde{\mathbf{y}} & \mathbf{z}_0\mathbf{A}^\top + \mathbf{b} \end{pmatrix} \begin{pmatrix} 2\mathbf{I} & \mathbf{A}^\top \\ \mathbf{A} & \mathbf{0} \end{pmatrix}^{-1} = \begin{pmatrix} \mathbf{y}^* & (\boldsymbol{\nu}^*)^\top \end{pmatrix} \end{aligned}$$

The optimal solution denoted with "\*" is directly yielded by projecting the NN output  $\tilde{\mathbf{y}}$  using the additional input  $\mathbf{z}_0$  and a projection matrix. This projection matrix only contains the constraint matrix  $\mathbf{A}$  can be computed before runtime which drastically speeds up forward and backward passes in contrast to implicit methods [20]. The solution elegantly serves as an activation function  $g_{pc}(\tilde{\mathbf{y}}, \mathbf{z}_0)$ , that maps the crude NN



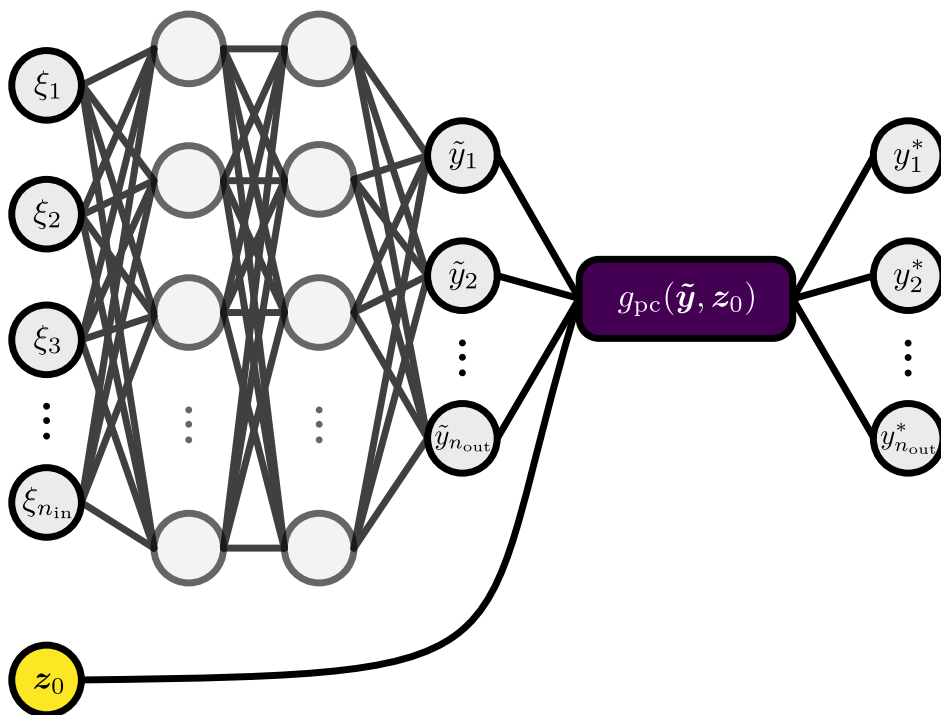


Figure 3.1: The physics constrained NN turns the input vector  $\xi$  into the crude output  $\tilde{y}$  via a feedforward NN, which is then projected by the physics constrained activation function  $g_{pc}$  into the feasible output  $y^*$ . It allows to include data points such as boundary conditions as needed in the form of  $z_0$

output into a subspace defined by  $\mathbf{A}$  and  $\mathbf{b}$ . The final output for  $y^*$  satisfies the linear equality constraints by definition, while it is located as close to the crude output  $\tilde{y}$  as possible. This is not only applicable in the field of chemical engineering but a general formulation that allows to integrate linear equality constraints into NNs. Additionally, the type of network architecture is independent of the mapping. It can be appended to any network type such as recurrent NNs, convolutional NNs or transformer-based architectures.

Due to its general applicability, the activation function could be used within a network to correct intermediate values. In most use cases however, it is applied as a final correction function at the output (fig. ??). The input  $\xi$  is propagated through a feedforward NN with the physics constrained activation function behind the last layer (purple). It takes the reference value  $z_0$  (yellow) without being forwarded through the network. In recent literature, NNs that utilize such a type of activation function can be found as *KKT-hPINN* [3], which is short for *Karush-Kuhn-Tucker-hard-physics-informed-neural-network*. For later reference, this activation function  $g_{pc}$  will be referred to as *physics constrained activation* (cor. ??) to have a short name in the context of chemical engineering. Furthermore, this procedure could be extended to support arbitrary non-linear

equality constraints by splitting the network output in frozen and unfrozen variables [21]. It enables the integration of enthalpy balances that are of non-linear nature. As this work combines the implementation of physics consistent NNs with parametric uncertainty for MPC, the focus remains on linear equality constraints.

---

**Algorithm 1** Forward pass of the physics constrained activation function.

---

**Require:**  $\mathbf{K} = \begin{pmatrix} 2\mathbf{I} & \mathbf{A}^\top \\ \mathbf{A} & \mathbf{0} \end{pmatrix}^{-1}$ ,  $\tilde{\mathbf{y}}, z_0, \mathbf{A}, \mathbf{b}$

$\mathbf{h} \leftarrow \begin{pmatrix} 2\tilde{\mathbf{y}} & z_0\mathbf{A}^\top + \mathbf{b} \end{pmatrix}$

$\mathbf{y}^* \leftarrow \mathbf{h}\mathbf{K}$

$\mathbf{y}^* \leftarrow \mathbf{y}^*_{:,n_{\text{out}}}$

**return**  $\mathbf{y}^*$

---

The forward pass through the physics constrained activation involves the setup of a temporary vector  $\mathbf{h}$  (alg. ??) and the truncation of the dual variables in the solution. The matrix inversion to yield  $\mathbf{K}$  should be done offline and then saved as untrainable model parameter. In `pytorch` this is can be done using the `register_buffer()` method. An implementation example of the physics consistent activation can be found in the appendix (XX).

# Chapter 4

## Case Study

The goal of this work is to derive a physics-consistent surrogate modelling technique to be applied within a robust MPC algorithm. Physics consistency, prediction accuracy and uncertainty quantification will be assessed using a case study of a catalytic *packed-bed tubular reactor*. This reactor type is common in the chemical industry in many processes such as ethylene oxidation, carbon dioxide methanization or steam reforming. Especially with regard to *load flexible* processes using renewable energy sources, dynamic control such as MPC can enable optimal operation. For example, the energy supply can be closely tied to weather conditions or the feed composition may fluctuate due to biological resources.

For exothermal reactions, *transient behavior* is a mayor safety concern of packed-bed tubular reactors [10]. In a dynamic operating environment, the thermal inertia of the catalytic packed bed can abruptly heat up unreacted feed mixture. This can cause a self-accelerating temperature runaway leading to thermal damage or even explosion. Furthermore, high temperatures can cause several catalyst deactivation effects such as sintering. Therefore, the temperature is considered a safety-crucial state that needs to obey to a maximum constraint in the sense of  $F_i \leq 0$  in the original optimization (def. ??) for all possible disturbances  $\mathbf{w}$ . The maximum of  $F_i$  should be approximated by  $\tilde{\varphi}_i(\mathbf{u})$ . This will be achieved by training a NARX model that predicts a tuned upper and lower bound for  $\tilde{\varphi}_i(\mathbf{u})$  independent from the uncertainty  $\mathbf{w}$  using quantile regression. Apart from safety, the maximization of the selectivity towards the value product is a reasonable control objective, as the selectivity is most correlated to process profitability in many cases. To achieve a reasonable turnover from educts to products, a minimum conversion is set to be required via a constraint. In total, following general optimization structure is proposed.

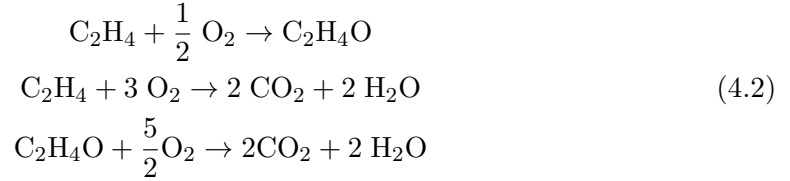
$$\begin{aligned} & \max S \\ \text{s.t. } & T \leq T_{\max} \\ & X \geq X_{\min} \end{aligned} \tag{4.1}$$

### 4.1 Reaction Kinetics

To focus on the surrogate modelling methodology, the example system is kept rather simplistic. The *ethylene oxidation* process consists of only two parallel reactions. Parametric

data is readily available due to extensive literature about this reaction system. Further, the reaction and reactor configuration is taken from an optimal reactor design study by Pietschak et al. [22] and adjusted for the needs at hand. Thus, some proposed operating conditions may be unrealistic due to safety concerns or over-simplifications.

Ethene is partially oxidized to ethylene oxide, while the total oxidation (combustion) of ethene takes place as a parallel reaction to carbon dioxide and water (eq. ??). According to the kinetics of Al-Saleh et al. [23], the consecutive reaction from ethylene oxide to water and carbon dioxide can be neglected.



The reaction rates can be described using following rate expression (eq. ??).

$$r_j = \frac{k_j p_{\text{E}}^{n_j^{(\text{E})}} p_{\text{O}_2}^{n_j^{(\text{O}_2)}}}{1 + K_j p_{\text{CO}_2}} \quad (4.3)$$

Here,  $r_j$  denotes the reaction rate of reaction  $j$ ,  $k_j$  the collision factor,  $p_\alpha$  the partial pressures of component  $\alpha$  and  $K_j$  the adsorption coefficient for carbon dioxide. The collision factor and the adsorption coefficient follow simple Arrhenius correlations (eq. ??, ??).

$$k_j = k_{0,j} \exp\left(\frac{-E_{\text{A},j}}{RT}\right) \quad (4.4)$$

$$K_j = K_{0,j} \exp\left(\frac{T_{\text{ads},j}}{T}\right) \quad (4.5)$$

These introduce a strong non-linearity in the temperature  $T$  and the kinetic parameters, the upright letter  $R$  represents the universal gas constant,  $E_{\text{A}}$  the activation energy,  $k_{0,j}$  and  $K_{0,j}$  the respective pre-exponential factors and  $T_{\text{ads},j}$  the adsorption temperature. To simulate these kinetics, the partial pressures of the rate determining species ethene, oxygen and carbon dioxide must be known.

## 4.2 Packed-bed tubular Reactor Model

The reactor as a packed-bed tubular reactor (TR) is assumed to be already existend with a fixed design to clearly separate the model derivation from an optimal design task (fig. ??). To ensure thermal stability, the additional temperatureprofile derivation is mandatory.

### 4.2.1 PDE System Derivation

The technical reactor itself is equipped with four wall cooling sections ( $T_{\text{w},1}$  to  $T_{\text{w},4}$ ) to manipulate the temperature profile. The input flow velocity  $u$  can fluctuate due to flexible feed availability such as in carbon dioxide methanization, where the throughput

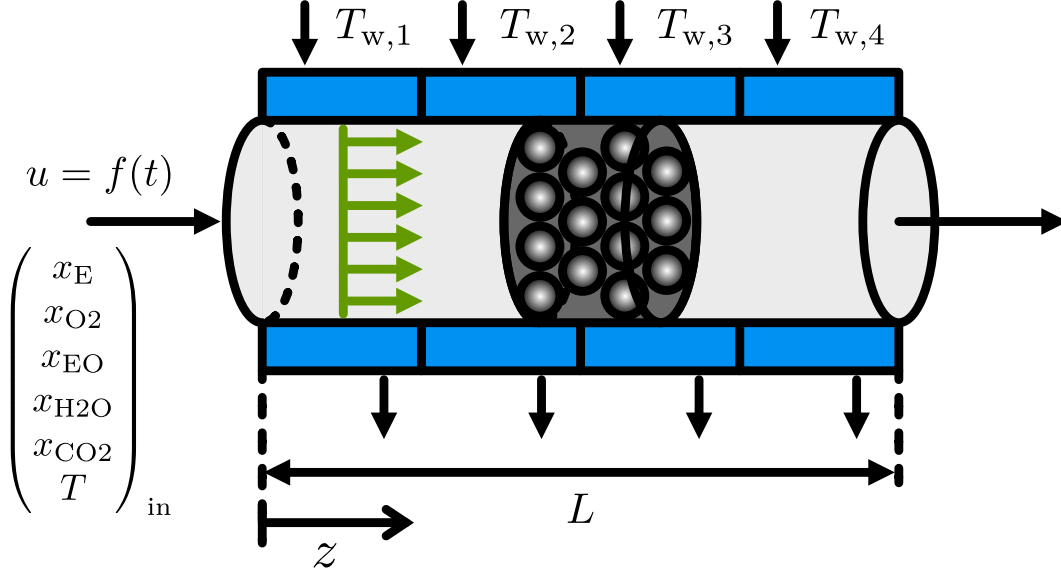


Figure 4.1: The packed-bed reactor is assumed to exhibit plug flow behavior (green). Catalyst and gaseous phase are lumped into one pseudo-homogenous reaction phase. The inlet conditions remain constant for the temperature  $T$  and the molar fractions of each component  $x_{(\cdot)}$ , while the inlet flow speed  $u$  can fluctuate due to flexible feed availability. Four cooling sections (blue) are integrated.

heavily relies on the hydrogen supply [\[empty citation\]](#) Because isobaric conditions and a non-compressible phase are assumed, the simulated states narrow down to the molar fractions of the rate-determining species  $x_\alpha$  and the temperature  $T$ . These are considered to be constant at the inlet, which is important for the later application of linear equality constraints in a NN. Using the component balance in local form (eq. ??), a PDE can be derived that describes the dynamic concentration gradients of the components  $\alpha$ . Furthermore, the enthalpy balance shall be covered to calculate the temperature field regarding the exothermic reactions.

$$\frac{\partial \rho_\alpha}{\partial t} = -\frac{\partial}{\partial z_k}(\rho_\alpha u_k + j_{\alpha,k}) + \sigma_\alpha \quad (4.6)$$

The general component density  $\rho_\alpha$  is expressed in  $\text{mol}/\text{m}^3$ , while  $\partial z_k$  indicates the spatial derivative in direction  $k$ .  $u_k$  and  $j_{\alpha,k}$  denote the convective velocity and the diffusive mass transport in direction  $k$ . The source or sink term  $\sigma_\alpha$  allows to consider chemical reactions. Ideal gas behavior is considered (eq. ??).

$$p_\alpha = \rho_\alpha R T \quad (4.7)$$

Moreover, *plug flow* and constant pressure are assumed. Under the consideration of a 1D-pseudo-homogenous model with gradients only in axial direction and no diffusive mass

transfer, this equation reduces to an ideal *plug flow tubular reactor* (PFTR) (eq. ??). A detailed derivation of the model equations can be found in the appendix (XX).

$$\begin{aligned} \frac{\partial \chi_\alpha}{\partial t} &= -u \frac{\partial \chi_\alpha}{\partial z} + (1 - \epsilon) \frac{RT_*}{px_*} \sum_j \nu_{\alpha,j} r_j(\chi_\alpha, T') \\ \text{with } \chi_\alpha &= \frac{x'_\alpha}{T'} \quad x'_\alpha = \frac{x_\alpha}{x_*} \quad T' = \frac{T}{T_*} \end{aligned} \quad (4.8)$$

To improve numerical stability, the mole fractions  $x_\alpha$  and the temperature occur in a dimensionless form indicated by a prime (') symbol. This is done by diving by a reference value, denoted with a star (\*) (eq. ??). It not only improves numerical simulation stability, for later training of surrogate models, it also eliminates the need for an *output scaler* which simplifies implementation and training. These reference values are defined as follows (eq. ??).

$$x_* = x_{E,\text{in}} \quad T_* = T_{\text{in}} \quad (4.9)$$

Due to the ideal gas behavior (eq. ??), the molar density  $\rho_\alpha$  varies not only with the molar fraction but also with the temperature. Therefore, an artificial state  $\chi_\alpha$  is introduced to enable an easy and explicit formulation of the PDE via substitution (eq. ??).

The enthalpy balance in local form can be significantly reduced under isobaric conditions with neither volume forces, nor friction, nor dispersive, nor conductive enthalpy transfer.

$$\frac{\partial(\rho h)}{\partial t} - \frac{\partial p}{\partial t} = -\frac{\partial}{\partial z_k} \left( \rho h u_k + q'_k \right) + u_k \frac{\partial p}{\partial z_k} + \sum_i j_{k,i} f_{k,i} - \Pi_{j,k} \frac{\partial u_j}{\partial z_k} + \sigma_h \quad (4.10)$$

The final PDE to describe the temperature field in the reactor is yielded (eq. ??).

$$\frac{\partial T'}{\partial t} = -u \frac{\partial T'}{\partial z} + \frac{1}{\rho c_p} \left( \frac{1 - \epsilon}{T_*} \sum_j (-\Delta h_{R,j}) r_j + \frac{4}{d_t} \alpha \left( \frac{T_w}{T_*} - T' \right) \right) \quad (4.11)$$

Important parameters are the bed void fraction  $\epsilon$ , the tube diameter  $d_t$ , the enthalpy of reaction  $\Delta h_{R,j}$ , the total density  $\rho$  and the total heat capacity  $c_p$ . The wall heat transfer is assumed to only be governed by the fluid-wall heat transfer coefficient  $\alpha$ , which is a function of the fluid velocity and the radial heat conductivity of the catalytic bed (eq. ??). In fact, this parameter abbreviated with  $\lambda_{\text{bed}}$  is crucial for the temperature behavior of the TR. Underlying dependencies like Reynolds and Prandl expressions are part of the appendix to keep the derivation tidy.

$$\alpha = f(\text{Re}, \text{Pr}, \lambda_{\text{bed}}) \quad (4.12)$$

The system of non-linear PDEs is discretized in  $z$ -direction via *backward finite differences*.

**Definition 4.1.** Discretized system of PDEs using backward finite differences to yield an ODE system

$$\begin{aligned}\frac{d\chi_{\alpha,i}}{dt} &= -u \frac{\chi_{\alpha,i} - \chi_{\alpha,i-1}}{\Delta z} + (1 - \epsilon) \frac{RT_*}{px_*} \sum_j \nu_{\alpha,j} r_j(\chi_{\alpha,i}, T'_i) \\ \frac{dT'_i}{dt} &= -u \frac{T'_i - T'_{i-1}}{\Delta z} + \frac{1}{\rho c_p} \left( \frac{1 - \epsilon}{T_*} \sum_j (-\Delta h_{R,j}) r_j(\chi_{\alpha,i}, T'_i) + \frac{4}{d_t} \alpha \left( \frac{T_{w,i}}{T_*} - T'_i \right) \right)\end{aligned}$$

### 4.3 Parametric Uncertainty

Because model parameters such as the activation energy or the pre-exponential collision factor are mostly fitted by laboratory experiments that exhibit some kind of measurement uncertainty, the parameters themselves cannot be exactly determined. They lie in a confidence space. Furthermore, fitting results with counteracting mechanisms produce cross-correlations between parameters. For example, when a higher activation energy is determined, simultaneously the according pre-exponential factor must be also higher to compensate for the reduced reaction speed. In this case study, five first principle model parameters should be treated as uncertain. These are the radial bed heat conductivity  $\lambda_{\text{bed}}$ , the activation energies  $E_{A,1}, E_{A,2}$  and their respective pre-exponential collision factors  $k_{0,1}, k_{0,2}$  for the main and the side reaction.

#### 4.3.1 Radial bed heat conductivity

The radial bed heat conductivity is set to follow a normal distribution with a mean of  $0.404 \text{ W m}^{-1} \text{ K}^{-1}$  calculated by a correlation by Zehner et al. [24] and a rather high standard deviation of  $0.08 \text{ W m}^{-1} \text{ K}^{-1}$  (fig. ??).

**Definition 4.2.** Normal distribution of the radial heat conductivity of the catalytic bed

$$\lambda_{\text{bed}} \sim \mathcal{N}(\mu_\lambda, \sigma_\lambda)$$

#### 4.3.2 Activation Energies and Pre-exponential factors

The activation energies and pre-exponential factors are assumed to arise from a multi-variate normal distribution, which can be determined for a cross-correlated parameter fitting (def. ??).

**Definition 4.3.** Multi-variate normal distribution of cross-correlated kinetic parameters

$$\begin{pmatrix} \ln k_{0,1} \\ E_{A,1} \\ \ln k_{0,2} \\ E_{A,2} \end{pmatrix} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$$

To calculate the mean vector  $\boldsymbol{\mu}$  and the covariance matrix  $\boldsymbol{\Sigma}$ , an actual parameter fitting with emulated measurement noise has been executed using a *berty* reactor model. All probabilistic terms are summarized (tab. ??).

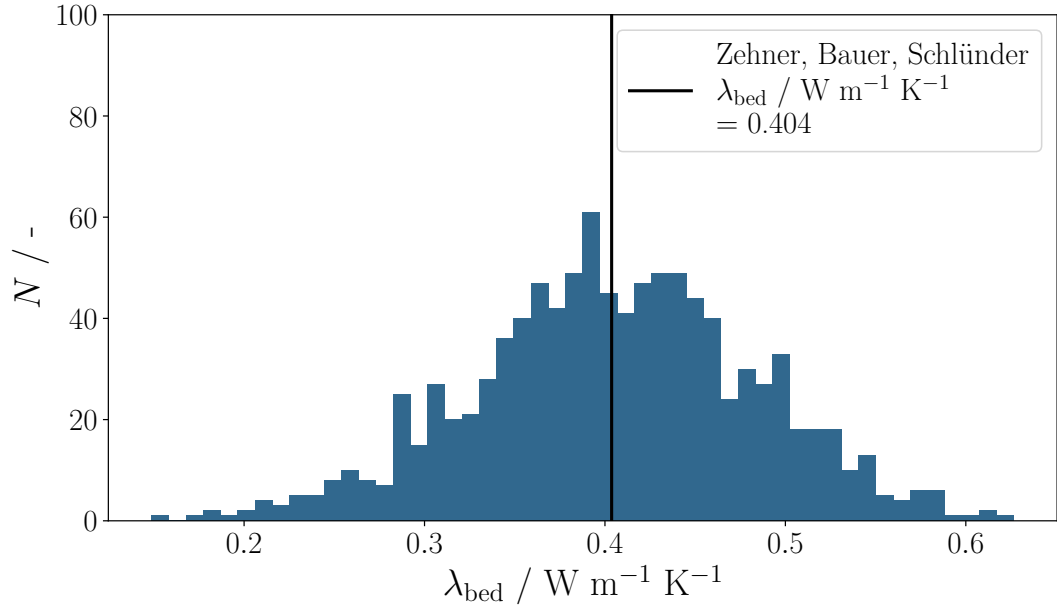


Figure 4.2: The radial heat conductivity of the packed bed follows a normal distribution with its mean calculated by the correlation of Zehner et al. [24] and a standard deviation of  $\sigma = 0.08 \text{ Wm}^{-1}\text{K}^{-1}$ , shown for 1000 sampling points.

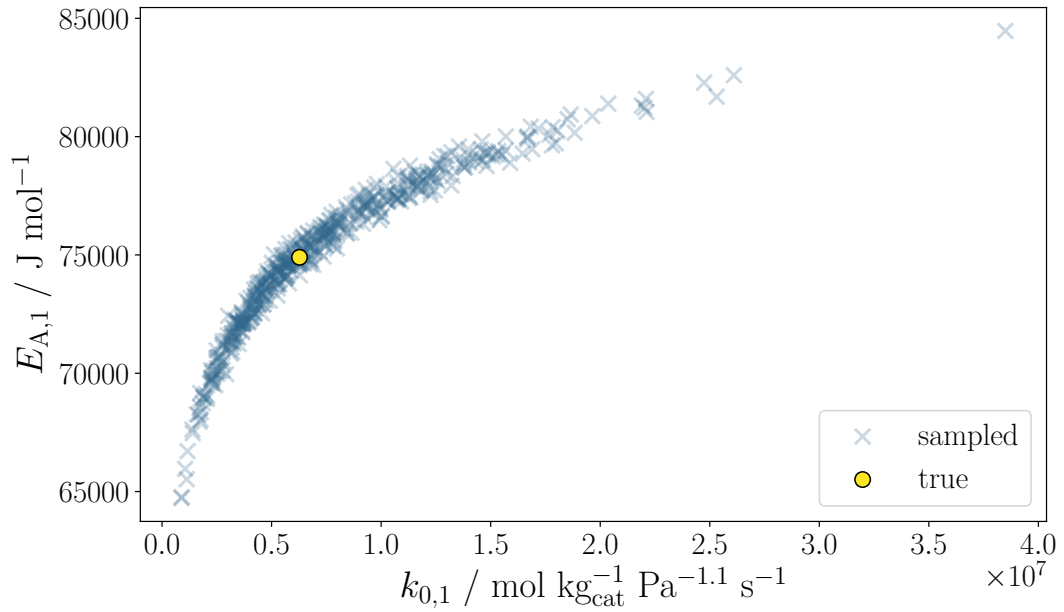


Figure 4.3: Due to the multi-variate normal distribution caused by an experimental parameter fitting and the non-linearity of the Arrhenius-Ansatz (eq. ??), the activation energy and the pre-exponential factor show severe non-linear uncertainty. The nominal value is marked in yellow.



$\Sigma$		$\mu$			
x	1	$\ln k_{0,1}$	$E_{A,1}$	$\ln k_{0,2}$	$E_{A,2}$
$\ln k_{0,1}$	15.621453	0.351544	1768.109910	0.380453	1913.304578
$E_{A,1}$	74856.339000	1768.109910	9031744.980000	1913.445056	9773020.700000
$\ln k_{0,2}$	16.272557	0.380453	1913.445056	0.411795	2070.851800
$E_{A,2}$	89853.839400	1913.304578	9773020.700000	2070.851800	10576522.400000



## Chapter 5

# Results on Surrogate Model Performance

Please choose a more expressive title for this chapter. If it makes sense for your work you can also have multiple chapters presenting your results.

### 5.1 Training Effectivity

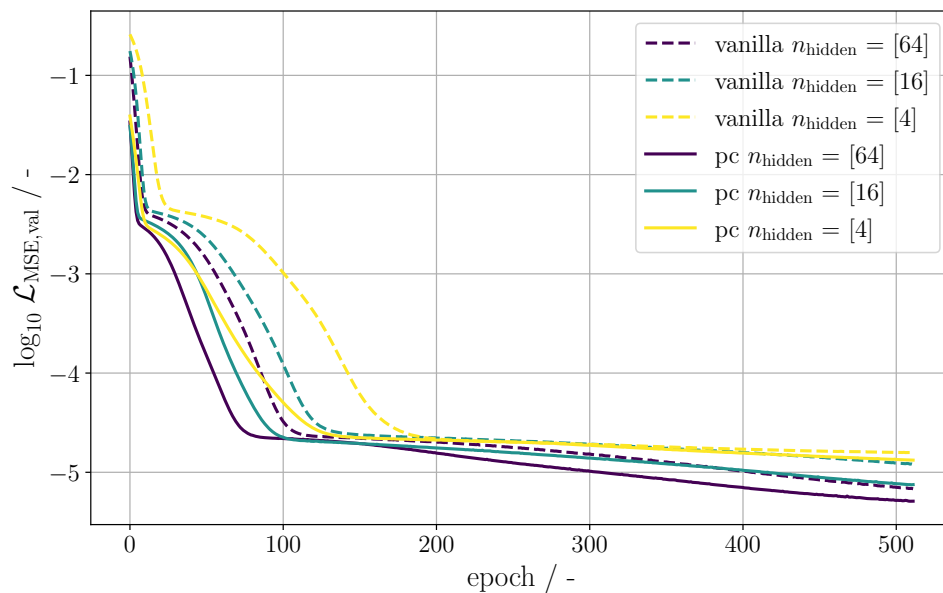


Figure 5.1: The validation loss calculated as MSE  $\mathcal{L}_{\text{MSE, val}}$  during the training of the vanilla NARX (dashed line) and the physics-constrained NARX (solid line). The three data rows each indicate *one hidden layer* with 64, 16 and 4 hidden neurons respectively. The physics-constrained NARX converges visibly faster than the vanilla NARX.

## Chapter 6

# Conclusions



# Appendix A

## Datasheets

Did your work require long tables of parameters, e.g. to configure a system model. This table would distract the main part of the thesis but including it is an important aspect of scientific work. Please attach it here.

### A.1 Parameter for ...

...





## Appendix B

### Proofs

Did you introduce lemmas or theorems with lengthy proofs? Including them in the main part of the thesis might be distracting. You can also include them in the appendix.

#### B.1 Proof of Theorem X

...



# Bibliography

- [1] JD, *Placeholder citation*, 0.
- [2] A. Vaswani, N. Shazeer, N. Parmar, *et al.*, *Attention is all you need*, 2023. arXiv: [1706.03762](https://arxiv.org/abs/1706.03762) [cs.CL]. [Online]. Available: <https://arxiv.org/abs/1706.03762>.
- [3] H. Chen, G. E. C. Flores, and C. Li, “Physics-informed neural networks with hard linear equality constraints,” *Computers & Chemical Engineering*, vol. 189, p. 108764, 2024. DOI: <https://doi.org/10.1016/j.compchemeng.2024.108764>.
- [4] B. Houska, “Robust optimization of dynamic systems,” Ph.D. dissertation, KU Leuven, 2011, ISBN: 978-94-6018-394-2.
- [5] O. Stein, “Bi-level methods for semi-infinite programming,” in *Bi-Level Strategies in Semi-Infinite Programming*. Boston, MA: Springer US, 2003, pp. 145–169, ISBN: 978-1-4419-9164-5. DOI: [10.1007/978-1-4419-9164-5\\_5](https://doi.org/10.1007/978-1-4419-9164-5_5). [Online]. Available: [https://doi.org/10.1007/978-1-4419-9164-5\\_5](https://doi.org/10.1007/978-1-4419-9164-5_5).
- [6] M. Diehl, “Robust optimization,” Lecture slides, Summer School on Robust Model Predictive Control with CasADi, University of Freiburg, September 15–19, 2025. Slides jointly developed with Titus Quah, Katrin Baumgärtner, and Florian Messerer, based on joint work with B. Houska and some illustrations from his PhD thesis., University of Freiburg, Germany, 2025.
- [7] K. Hornik, “Approximation capabilities of multilayer feedforward networks,” *Neural Networks*, vol. 4, no. 2, pp. 251–257, 1991. DOI: [https://doi.org/10.1016/0893-6080\(91\)90009-T](https://doi.org/10.1016/0893-6080(91)90009-T).
- [8] T. Salzmann, E. Kaufmann, J. Arrizabalaga, M. Pavone, D. Scaramuzza, and M. Ryll, “Real-time neural mpc: Deep learning model predictive control for quadrotors and agile robotic platforms,” *IEEE Robotics and Automation Letters*, vol. 8, no. 4, pp. 2397–2404, Apr. 2023. DOI: [10.1109/lra.2023.3246839](https://doi.org/10.1109/lra.2023.3246839).
- [9] S. Lucia and S. Engell, “Robust nonlinear model predictive control of a batch bioreactor using multi-stage stochastic programming,” in *2013 European Control Conference (ECC)*, 2013, pp. 4124–4129. DOI: [10.23919/ECC.2013.6669521](https://doi.org/10.23919/ECC.2013.6669521).
- [10] R. Güttel and T. Turek, *Chemische Reaktionstechnik*, 1st ed. 2021. Berlin, Heidelberg: Springer Berlin Heidelberg and Springer Spektrum, 2021, ISBN: 978-3-662-63150-8.
- [11] “Dynamic optimization methods with embedded dae solvers,” in *Nonlinear Programming*, ch. 9, pp. 251–286. DOI: [10.1137/1.9780898719383.ch9](https://doi.org/10.1137/1.9780898719383.ch9). eprint: <https://epubs.siam.org/doi/pdf/10.1137/1.9780898719383.ch9>. [Online]. Available: <https://epubs.siam.org/doi/abs/10.1137/1.9780898719383.ch9>.

- [12] S. Lucia and S. Engell, “Potential and limitations of multi-stage nonlinear model predictive control,” *IFAC-PapersOnLine*, vol. 48, no. 8, pp. 1015–1020, 2015. DOI: <https://doi.org/10.1016/j.ifacol.2015.09.101>.
- [13] K. McBride and K. Sundmacher, “Overview of surrogate modeling in chemical process engineering,” *Chemie Ingenieur Technik*, vol. 91, no. 3, pp. 228–239, 2019. DOI: <https://doi.org/10.1002/cite.201800091>. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/cite.201800091>.
- [14] T. Salzmann, J. Arrizabalaga, J. Andersson, M. Pavone, and M. Ryll, “Learning for casadi: Data-driven models in numerical optimization,” in *Learning for Dynamics and Control Conference (L4DC)*, 2024.
- [15] K. Benidis, S. S. Rangapuram, V. Flunkert, *et al.*, “Deep learning for time series forecasting: Tutorial and literature survey,” *ACM Computing Surveys*, vol. 55, no. 6, pp. 1–36, Dec. 2022. DOI: [10.1145/3533382](https://doi.org/10.1145/3533382).
- [16] N. Kemmerling and S. Lucia, “Uncertainty-aware reduced-order modeling for robust control of zeolite crystallization,” Manuscript in preparation, TU Dortmund University, Dortmund, Germany, 2025.
- [17] S. L. Brunton and J. N. Kutz, “Reduced order models (roms),” in *Data-Driven Science and Engineering: Machine Learning, Dynamical Systems, and Control*, POD for model reduction, Cambridge University Press, 2019, pp. 375–402.
- [18] Anonymous, *Torch-incremental-pca: A pytorch implementation of incremental pca*, <https://pypi.org/project/torch-incremental-pca/>, PyTorch implementation inspired by scikit-learn’s IncrementalPCA and PCAonGPU. Provides GPU-accelerated, memory-efficient incremental PCA for large datasets., 2025.
- [19] M. Raissi, P. Perdikaris, and G. Karniadakis, “Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations,” *Journal of Computational Physics*, vol. 378, pp. 686–707, 2019. DOI: <https://doi.org/10.1016/j.jcp.2018.10.045>.
- [20] A. Agrawal, B. Amos, S. Barratt, S. Boyd, S. Diamond, and Z. Kolter, “Differentiable convex optimization layers,” in *Advances in Neural Information Processing Systems*, 2019.
- [21] G. Lastrucci, T. Karia, Z. Gromotka, and A. M. Schweidtmann, *Picard-kkt-hpinn: Enforcing nonlinear enthalpy balances for physically consistent neural networks*, 2025. arXiv: [2501.17782](https://arxiv.org/abs/2501.17782) [cs.LG]. [Online]. Available: <https://arxiv.org/abs/2501.17782>.
- [22] A. Pietschak, M. Kaiser, and H. Freund, “Tailored catalyst pellet specification for improved fixed-bed transport characteristics: A shortcut method for the model-based reactor design,” *Chemical Engineering Research and Design*, vol. 137, pp. 60–74, 2018. DOI: [10.1016/j.cherd.2018.06.043](https://doi.org/10.1016/j.cherd.2018.06.043).
- [23] M.A. Al-Saleh M.S. Al-Ahmadi M.A. Shalabi, *Kinetic study of ethylene oxidation in a berty reactor*, 1988. DOI: [10.1016/0300-9467\(88\)80004-2](https://doi.org/10.1016/0300-9467(88)80004-2). [Online]. Available: <https://www.sciencedirect.com/science/article/pii/0300946788800042>.

- [24] P. Zehner and E. U. Schlünder, “Wärmeleitfähigkeit von schüttungen bei mäßigen temperaturen,” *Chemie Ingenieur Technik*, vol. 42, no. 14, pp. 933–941, 1970. DOI: <https://doi.org/10.1002/cite.330421408>. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/cite.330421408>.