# Set-based Multi-scale Modeling and Analysis of Signal Transduction Pathways

**Dissertation**
zur Erlangung des akademischen Grades
**Doktoringenieur (Dr.-Ing.)**

von
**Dipl.-Ing. Nadine Rudolph, geb. Strobel**
geboren am 20. April 1987 in Schleiz

genehmigt durch die Fakultät für Elektrotechnik und Informationstechnik der
Otto-von-Guericke-Universität Magdeburg

Gutachter:
Prof. Dr.-Ing. Rolf Findeisen
Prof. Dr. Fred Schaper
Dr. Heinrich Huber

eingereicht am 17. September 2019
Promotionskolloquium am 26. November 2019

# Contents

# Contents

# Abstract

Mathematical modeling of biological systems together with quantitative biochemical and cell biological analyses allow deep insights into the often complex molecular interconnections within a biological system. Furthermore, mathematical models allow to design and test new patient-specific intervention strategies for medical treatments.

The development of mathematical models is not trivial. Usually, not all interactions within a biological network are known or understood resulting in competing model candidates. These candidates have then to be tested whether they can represent available measurement data. In case a model cannot represent the data, it is deemed invalid. However, as measurement data are mostly uncertain this model discrimination task is challenging. In addition, model parameters are typically unknown and have to be calibrated for reliable model predictions. Finally, biological processes can span different time scales ranging from seconds to days. The combination and the integration of data from different time scales is challenging, however, needs to be considered for model predictions that allow insights into biology and eventually pathophysiological processes. For model analyses and parameter estimation, we apply and extend a set-based model analysis approach. Set-based estimation methods allow to tackle some of the mentioned challenges by considering bounded uncertainties in parameters, states and measurement data.

This thesis presents extensions to set-based estimation. First, we develop a method to estimate inner approximations of parameter sets. The inner approximation approach is based on an inversion of the problem for deriving outer approximations using mixed-integer linear programming. Second, we present a bilevel set-based experimental design approach which allows for guaranteed model invalidation when the available measurement data is not sufficient to discriminate between the different model candidates. Third, we tackle the problem of set-based combination of different time scales within one model using a phenomenological and a classification approach, respectively.

We apply the developed extensions to Interleukin-6-induced signal transduction. Interleukin-6 is a pleiotropic cytokine that is involved in a number of cellular functions. In particular, Interleukin-6 induces cellular growth, apoptosis and differentiation. Interleukin-6 mediates inflammatory effects, such as fever by stimulating the acute phase protein synthesis. Under certain conditions inflammatory effects can become uncontrolled and get out of control and chronic. We study Interleukin-6-induced signaling aiming to obtain deeper insights into its (patho-)physiology and finally, present a large application example showing that differences in Interleukin-6-induced signaling are only mediated by differences in the number of cell surface

receptors. Our results open a door for the design of new drug intervention strategies for the treatment of deregulated Interleukin-6-induced signaling and related chronic inflammatory responses.

# Deutsche Kurzfassung

Die mathematische Modellierung and Untersuchung biologischer Systeme zusammen mit quantitativen biochemischen und zellbiologischen Analysen erlauben tiefe Einblicke in komplexe molekularer Zusammenhänge innerhalb eines biologischen Systems. Mathematische Modelle werden vermehrt für die Entwicklung neuer, patientenspezifischer Interventionsstrategien genutzt.

Die Erstellung mathematischer Modelle zur Untersuchung komplexer molekularer Mechanismen birgt zahlreiche Herausforderungen. Oftmals sind nur limitierende Informationen über die molekularen Mechanismen vorhanden. Dies führt dazu, dass eine Vielzahl an Hypothesen existieren, welche anhand experimenteller Daten validiert werden müssen. Dies ist oftmals nicht einfach, da experimentelle Daten häufig große Fehlerabweichungen aufweisen. Des Weiteren sind die zugrundeliegenden kinetischen Parameter, welche die Reaktionsgeschwindigkeiten der betrachteten Mechanismen beschreiben, meistens unbekannt, so dass oftmals nur Unsicherheitsintervalle angegeben werden können. Die Prädiktionsmöglichkeiten eines mathematischen Modells hängen stark von diesen unsicheren und unbekannten Parametern ab. Schließlich spannen die zu untersuchenden biologischen Prozesse oftmals mehrere Zeitskalen, von Sekunden über Minuten zu Tagen. Diese Mehrskaligkeit erfordert, dass verschiedene Datensätze innerhalb eines Modells miteinander verbunden und interpretiert werden müssen. Wir verwenden in dieser Arbeit einen mengenbasierten Ansatz. Mengenbasierte Schätzmethoden erlauben es, einige der genannten Herausforderungen zu bewältigen, da sie endliche, unsichere Mengen der Parameter, Zustände und Messdaten betrachten.

Wir präsentieren zunächst Erweiterungen des verwendeten mengenbasierten Ansatzes. Wir entwickeln eine Methode für die Schätzung innerer Parametermengen, welche auf der Invertierung des Problems für die Schätzung äußerer Parametermengen besteht unter Verwendung einer gemischt-ganzzahligen linearen Programmierung. Danach stellen wir einen auf Bilevel-Optimierung gestützten, mengenbasierter Ansatz vor. Dieser erlaubt eine garantierte Diskriminierung konkurrierender Modelle. Schließlich widmen wir uns dem Problem der Mehrskaligkeit biologischer Prozesse und stellen einen mengenbasierten, phenomenologischen Ansatz vor. Dieser Ansatz erlaubt, schnelle Prozesse auf der Kurzzeitskala mit langsameren Prozessen auf der Langzeitskala zu verknüpfen. Des Weiteren beschäftigen wir uns damit, inwiefern mengenbasierte Ansätze mit Klassifizierungsmethoden verknüpft werden können. Die Idee ist hierbei, Informationen der Langzeitskala dazu zu verwenden, um Patientenkohorte in verschiedene Klassen (z. B.: Grad der Erkrankung) einzugruppieren.

Die entwickelten Ansätze werden für die Modellierung und Analyse der Interleukin-

6-induzierten Signaltransduktion exemplarisch verwendet. Interleukin-6 ist ein pleiotropes Zytokin, welches in einer Vielzahl von zellulären Prozessen involviert ist. So vermittelt Interleukin-6 zum Beispiel inflammatorische Effekte wie Fieber durch die Stimulation der Synthese von Proteinen der akuten Phase. Unter bestimmten Bedingungen können inflammatorische Effekte außer Kontrolle geraten und chronisch werden. Die Analyse der Interleukin-6-induzierten Signaltransduktion erlaubt einen tieferen Einblick in die (Patho-)Physiologie von derselben. Unsere Ergebnisse zeigen, dass Unterschiede in der Interleukin-6-induzierten Signaltransduktion einzig von der Expression membrangebundenener Rezeptoren abhängt. Diese Ergebnisse liefern einen wichtigen Beitrag für das Design neuer personalisierter Interventionstrategien bezüglich der Behandlung inflammatorischer Erkrankungen.

# Glossary

## Biological terms

| | |
|---|---|
| classic-signaling | activation of signal transduction by binding of IL-6 to membrane-bound IL-6R$\alpha$; mediates anti-inflammatory activities |
| cytokines | tissue hormones that bind to receptors and initiate intracellular signal transduction |
| dephosphorylation | the removal of a phosphate group from post-translationally modified proteins by a protein phosphatase |
| dissociation constant | $K_D$; describes the dissociation of two components in a complex, e. g. an antibody and its antigen; $K_D = \frac{k_{off}}{k_{on}}$ |
| ERK | extracellular signal-regulated kinase; MAPK that is involved in, e. g. the activation/phosphorylation of transcription factors regulating cellular growth |
| FACS | fluorescence-activated cell sorting by Flow Cytometry |
| Flow cytometry | a laser-based technology employed for the analysis of intracellular and membrane-bound proteins in single cells |
| gp130 | glycoprotein 130; a transmembrane signal-transducing receptor subunit that binds to IL-6:IL-6R$\alpha$ and Hyper-IL-6 |
| Grb2 | growth factor receptor-bound protein 2; adaptor protein that is constitutively associated with SOS; mediates binding of SOS to, e. g. SHP2; part of the MAPK signaling pathway |
| Hyper-IL-6 | designer fusion protein in which IL-6 is connected to sIL-6R$\alpha$ by a flexible peptide linker; induces trans-signaling |

| | |
|---|---|
| IL-6 | interleukin-6; a cytokine with pro- and anti-inflammatory activities; involved in immune and inflammatory responses, but also in the regulation of metabolic, regenerative, and neural processes |
| IL-6R$\alpha$ | interleukin-6 receptor $\alpha$; transmembrane receptor expressed on the surface of specific cell types; induces classic-signaling |
| IL-6:IL-6R$\alpha$ | dimeric protein complex consisting of IL-6 and IL-6R$\alpha$ |
| Jak | janus kinase; a family of intracellular, non-receptor tyrosine kinases; constitutively bound to gp130; transduces cytokine-mediated signals |
| Jak/STAT pathway | signal transduction pathway; regulates cell functions (cell growth) and the acute phase response; initiated by a wide variety of growth factors and cytokines, such as IL-6 |
| kinase | an enzyme that catalyzes protein phosphorylation |
| MAPK | mitogen-activated protein kinase, e. g. ERK; serine-threonine kinase |
| MAPK pathway | signal transduction pathway; regulates cell functions, such as growth, differentiation, mitosis and cell survival/apoptosis |
| Mek | threonine-tyrosine kinase; phophorylates ERK; part of the MAPK signaling pathway |
| phosphatase | an enzyme that catalyzes protein dephosphorylation |
| phosphorylation | the addition of a covalently bound phosphate to proteins by a protein kinase |
| mRNA | messenger ribonucleic acid; subtype of RNA conveying genetic information from DNA to the ribosome |
| (p)R$_{complex}$ | phosphorylated active receptor complex |

(p)STAT — phosphorylated STAT; phosphorylated by Jaks, dimerizes and acts as transcription factor to induce the expression of, e. g. SOCS3 protein and proteins regulating cellular responses, such as cell growth

patient stratification — in the framework of this theses: the classification of a patient cohort into subcategories of risk levels for developing inflammatory diseases

qRT-PCR — quantitative real-time polymerase chain reaction; technique to monitor amplification of complementary DNA in real-time

Raf — rapidely acclerated fibrosacroma; a serine-threonine kinase that activates Mek; part of the MAPK signaling pathway

Ras — a small G-protein that activates Raf; part of the MAPK signaling pathway

$R_{complex}$ — active receptor complex; consists of $2 \times$IL-6:IL-6R$\alpha$ and $2 \times$gp130 in classic-signaling and of $2 \times$IL-6:sIL-6R$\alpha$ and $2 \times$gp130 in trans-signaling

RT-PCR — reverse transcription polymerase chain reaction; a nucleic acid amplification technique in which RNA molecules are converted into their complementary DNA to measure their expression

SDS — sodium dodecyl sulfat; used to denature proteins within a cellular lysate

SDS-PAGE — SDS-polyacrylamide gel electrophoresis; used to separate proteins according their size by electrophoresis

SHP2 — SH2-containing protein tyrosine phosphatase 2; a signaling enzyme that contains two tandem Src homology-2 (SH2) domains that function as phospho-tyrosine binding domains; interacts with a variety of signaling intermediates such as Grb2

| | |
|---|---|
| signal transduction | a process where a biochemical signal is passed through a cascade of biological entities (proteins, metabolites, lipids, neucleotides) to achieve a certain biological response (protein production, cell growth, cell death) |
| sIL-6R$\alpha$ | soluble interleukin-6 receptor $\alpha$; recognizes and binds to IL-6 inducing the trans-signaling pathway |
| SOCS3 mRNA | suppressor of cytokine signaling 3 mRNA |
| SOCS3 | expressed during Jak/STAT signaling; negative feedback inhibitor that inhibits the activity of Jaks |
| SOS | son of sevenless; guanine nucleotide exchange factor; an enzyme that activates the small G-protein Ras |
| STAT | signal transducer and activator of transcription; transcription factor that resides in the cytoplasm until its activation via Jak |
| transcription | biochemical process by which a particular segment of DNA is transcribed into messenger RNA (mRNA) |
| transcription factor | binds to specific regulatory DNA sequences and controls the rate of transcription from DNA to mRNA |
| translation | biochemical process by which mRNA is decoded to synthesize proteins |
| trans-signaling | activation of signal transduction by binding of IL-6 to sIL-6R$\alpha$; mediates pro-inflammatory activities and is related to diseases, such as multiple sclerosis, rheumatoid arthritis and cancer |
| Western Blotting | transfer of proteins from a gel to a membrane; part of the protein detection and quantification process |

# Technical terms

| | |
|---|---|
| bisectioning | algorithm that is used to explore the feasible parameter space by dividing the space into partitions |
| bilevel optimization | optimization problem consisting of two embedded problems, i. e. the upper-level optimization task, and the lower-level optimization task |
| convex optimization | optimization of convex functions over convex sets |
| experiment design | in the context of this thesis: the design of a biological stimulus allowing for the discrimination of competing model hypotheses |
| false positive solution | a solution of the (MI)SDP or (MI)LP for which the original (MI)FP is not capable to represent the data |
| inner approximation | estimation approach for model parameters; contains only feasible solutions, but possibly not all |
| integrality constraint | constraint in a mixed-integer problem which can only take whole numbers, such as 0 or 1 |
| Lagrangian multipliers | a strategy for finding the local maxima and minima of a function subject to equality constraints |
| Lagrange duality | a method to solve an optimization problem under given constraints; obtained by forming the Lagrangian using nonnegative Lagrangian multipliers to add the constraints to the objective function; it is used for solving of some primal variable values (i. e. variable values in the (MI)FP) that minimize the Lagrangian; the solution to the dual problem provides a lower bound to the solution of the primal problem (i. e. the (MI)FP) |
| mixed-integer problem | a problem in which some of the variables are constrained to be integer (i. e. whole-numbered, such as 0 and 1) at the optimal solution |

| | |
|---|---|
| (MI)FP | (mixed-integer) feasibility problem; a mathematical problem that aims to determine feasible solutions that fulfill the given (integrality) constraints; a nonconvex problem |
| (MI)SDP | (mixed-integer) semidefinite program; a relaxed and convexified form of the (MI)FP; a problem of optimizing a linear objective function over the intersection of the cone of positive semidefinite matrices |
| (MI)LP | (mixed-integer) linear program; a relaxed form of the (MI)SDP; a problem of optimizing a linear objective function, subject to linear equality and linear inequality constraints |
| model calibration | the process of parametrizing a mathematical model |
| model discrimination | the process of separating models according to their outputs |
| model invalidation | the falsification of a model, i. e. proving that a model is not capable to represent measurement data |
| multi-scale model | a model that describes the system behaviour over several magnitudes of physical (space, time), biological (cells, tissues, organs) or logical entities (levels of details); in the context of this thesis: a model that describes processes from different time scales (i. e. from seconds to days) |
| objective function | an equation to be optimized (i. e. minimized or maximized) under given constraints |
| ODE | ordinary differential equation; a method to calculate the propagation of dynamical variables (here: over time) |
| optimization problem | a problem of finding the best solution from all feasible solutions |
| outer approximation | estimation approach for model parameters; encloses all feasible and possibly also false positive solutions |
| outer-bounding | algorithm to derive an upper and an lower bound of the feasible set of an (MI)FP, such as the model parameters |

QP

quadratic program; a problem of optimizing a quadratic function of several variables subject to linear constraints on these variables

relaxation

a technique for transforming a problem with hard constraints, that is difficult to solve into one with weaker constraints, that is easier to solve

# 1 Introduction

Model-based approaches using systems theoretical methods are the foundation for the development of quantitative and predictive models that allow in-depth insights into biological systems and signaling.

The devising of mathematical models faces several challenges. Firstly, limited information on the mechanisms of biological processes are usually available. This limitation leads often to a variety of **competing model hypotheses** requiring to discriminate between those that are capable to represent measurement data and those that cannot. As measurement data are usually **uncertain** the discrimination of different models is not trivial. Furthermore, even if a model describing the molecular mechanisms under study was established, the underlying dynamics of the biochemical reactions are often **not precisely known**. This leaves the **kinetic parameters** largely undetermined and restricts the predictive power of the generated models. Finally, biological processes can **span several time scales** from seconds over minutes to days, making it necessary to connect and interpret uncertain data on different time scales. This problem of temporal integration is often aggravated by the fact that quantitative data with high temporal resolution can only be obtained for short-term processes (e. g. biochemical kinetics). In contrast, only sparse data (e. g. on cell growth of tumor growth) or categorical data (e. g. on survival/death, stage of inflammation) and qualitative data are available for long-term processes.

To address these challenges, this thesis tackles the following issues (cf. Fig 1.1):

a) **estimation** of unknown model parameters for reliable and robust predictions

b) **design** of optimal experiments for model discrimination and selection

c) **combination** of processes acting on different time scales and the **integration** of uncertain data obtained at these time scales.

Throughout this work we focus on the analysis of signal transduction pathways as one important part in the field of systems biology. In particular, we use IL-6-induced signaling as a running example demonstrating applicability of the developed methods and aiming to obtain a deeper understanding of IL-6-induced (patho-)physiological processes (cf. Chapter 2). IL-6-induced signal transduction plays a crucial role during inflammatory processes by regulating the expression of a variety of target genes. These target genes can, for instance, initiate the acute phase response [73], which is important for initiating tissue repair. While under physiological conditions, IL-6-induced inflammatory processes are tightly regulated, a pathological misbalance in
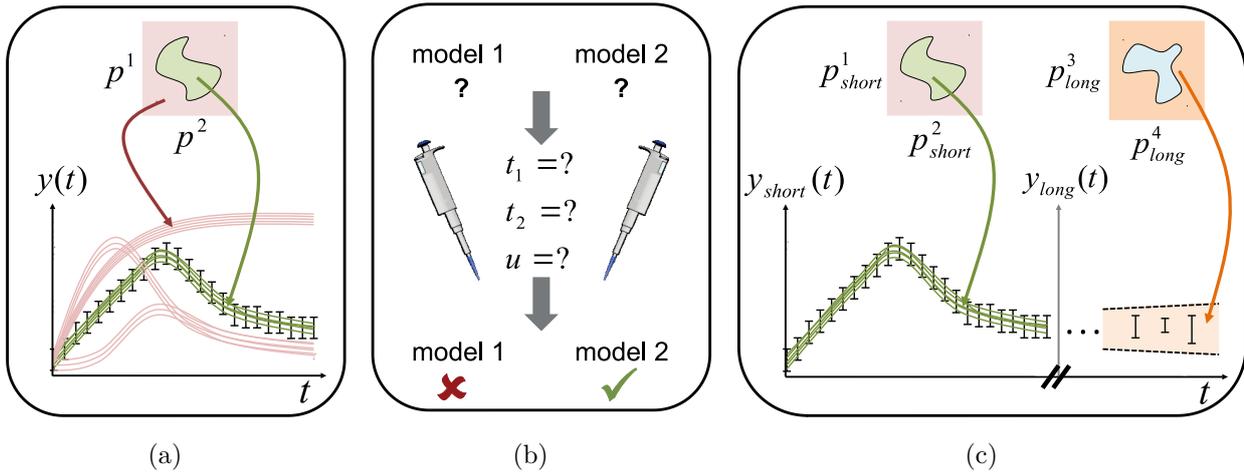
Figure 1.1: **Illustration of the topics addressed in this thesis.** (a) For the analysis and understanding of biochemical processes reliable parameter estimates and model predictions (green trajectories) are needed and false estimates leading to invalid predictions (red trajectories) have to be identified . (b) For established models the experimental design, e. g. of the sampling time $t$ or the model input $u$ allows us to further discriminate between valid (green check mark) and invalid models (red cross). (c) Valid models allow to describe and predict appropriate process trajectories (orange corridor) on the long-term time scale (right panel) based on processes on the short-term time scale (left panel).

inflammation-related effector proteins can cause the development of chronic and autoimmune diseases, such as rheumatoid arthritis, multiple sclerosis and Crohn's Disease [39, 54, 77, 109]. Furthermore, deregulated inflammation plays a crucial role in cancer progression [42]. To this end, we present a large application example for

d) model **estimation** and **analysis** of **IL-6-induced signaling pathways** that allow future design of new patient-specific intervention strategies.

Specifically, we apply and extend a set-based modeling framework. The set-based approach allows to consider uncertainties in model parameters and measurement data for guaranteed model invalidation, parameter estimation, and experimental design.

While for the example of IL-6-induced inflammatory signaling several computational studies have been performed [49, 106, 148, 150], none of these approaches considered set-based uncertainties and therefore, cannot give any guarantees for model invalidity and parameter estimation.

## 1.1 Overview on research topics

A large number of modeling frameworks are available for the estimation and analysis of biological processes including ordinary differential equations, Boolean or Petri networks

or stochastic approaches. Independent of the approach one has to keep in mind that measurement data are usually subject to (large) noise and model parameters as well as some of the model state variables are typically unknown and have to be inferred from the measurement data. Consequently, frameworks for modeling, estimation, and analysis of biological processes need to account for uncertainties in measurement data, model parameters, and states.

In the following, we review systems theoretic approaches that are frequently used for the analysis and estimation of uncertain biological processes and that are considered in this thesis.

**Model discrimination and selection**

The lack of precise biological information (e. g. limited knowledge about the underlying processes and/or noisy measurement data) leads often to competing model hypotheses which need to be compared against each other according to their capability to represent measurement data. The comparison of models based on their capability to represent measurement data is also denoted as model discrimination task. This task describes the discrimination between models that cannot represent the data (inconsistent, deemed as invalid), and those that allow to represent the data (consistent, deemed as possibly valid). To select models according to their capability to represent measurement data different measures exist allowing an assessment and subsequent ranking of competing model hypotheses. The so-called *minimum chi-square estimation* method, for example, is based on the *chi-squared test* ($\chi^2$ test) whose statistic describes the goodness-of-fit of the data to the model [124]. In particular, *minimum chi-square estimation* sums up the squared distance of the data and model prediction weighted with the standard measurement error. The $\chi^2$ test thereby aims to make the $\chi^2$ test statistic as small as possible [18]. Another approach is the *Akaike information criterion* assesses models by comparing the relative quality of statistical models against each other for a given set of data [3]. The Akaike information criterion value is computed by the sum of the natural logarithm of the likelihood (i. e. $log-likelihood$ - a measure for the model fit) and the number of parameters, i. e. $-2log(L)+Kp$. In the previous equation, $L$ is the likelihood function, $p$ is the number of parameters in the model, and $K$ is a regularization factor. Thereby, the regularization factor $K$ penalizes the amount of parameters $p$ and thus model complexity. In principal, models with a low Akaike information criterion value are more capable to represent the measurement data than models with a high Akaike information criterion value. In addition to that, the *Bayesian information criterion* is a function estimate of the posterior probability of a model being capable to represent measurement data under a certain Bayesian setup [21]. In comparison to Akaike information criterion, Bayesian information criterion penalizes model complexity more heavily than Akaike information criterion. The Bayesian information criterion is computed similar to Akaike information criterion, whereby the regularization factor $K$ is set to $log(N)$ with $N$ denoting the sample size

[126].

Although for all proposed criteria the model hypothesis with the lowest value of the respective criterion is suggested to be the most likely one, the described criteria are rather subjective [114]. This is due to the subjective choice of the threshold at which a model hypothesis is deemed as invalid (see e. g. [128]). In addition, the underlying statistic assumptions for each criterion (e. g. large sample size, prior distributions) are usually not met. Furthermore, the above scores only provide a probabilty of the validity of a model hypothesis. To definitely demonstrate validity of a model hypothesis, an infinite number of experiments need to be considered [5]. Therefore, model estimation, validation and selection methods restrain their efforts to demonstrate model invalidity in lieu of validity. Thereby, model invalidity refers to the fact that under all constraints and assumptions made, the model is not capable to represent the measurement data. As such methods for model invalidation, stochastic approaches to minimize the risk of false hypothesis acceptance, such as the *Neyman-Pearson Lemma*, have been developed and were discussed in [111]. Alternatively, the work presented in [127] provides a method for the (in)validation of continuous-time nonlinear models. The approach is based on so-called barrier-certificates that are functions of state, parameter and time. Such certificates allow guaranteed invalidation of models if the model trajectories do not intersect with the ranges of allowed measurement data. However, deriving such certificates is not trivial and their existence is not guaranteed.

**Parameter estimation**

Determination of a model that represents the data best, typically depends on the choice of the usually unknown model parameters. To this end, we give in the following an overview about parameter estimation and discuss systems theoretical approaches that are frequently used to infer these unknown model parameters.

Computational models for signal transduction processes are typically modeled using ordinary differential equations (ODEs) that describe the change of model states over time. Typically, these model states represent concentrations of chemical entities, such as proteins. For formulating equations that describe the considered system, widely accepted kinetic laws are applied, such as the law of mass action. As a consequence, the derived equations depend on kinetic parameters that describe the velocity of a reaction, e. g. the production or decay of proteins. Some of these parameters might be determined experimentally. However, most of the parameters are unknown and therefore, are usually determined by model calibration which is the process to parametrize a mathematical model.

The choice of an optimal model parametrization such that a model fits the measurement data best is a central challenge in mathematical modeling. The problem of parameter estimation can be seen as *inverse problem* which, in this context, means that model parameters have to be determined from experiments [161]. Inverse problems can be *well-* and *ill-posed*. A *well-posed* problem is characterized by three properties: a

solution exists, the solution is unique, and the behavior of the solution depends continuously on data and parameters. The problem of parametrizing mathematical models in biology is, that these criteria are rarely fulfilled (i. e. *ill-posed* problems) due to nonlinearities in the model equations, sparse data sets and large data uncertainties.

Typical approaches for estimating parameters of *ill-posed* problems are based on recasting the initial problem into an optimization problem of the following form:

$$
\begin{aligned}
\text{minimize} \quad & f_0(x) \\
\text{subject to} \quad & g_i(x) \le 0, \ i = 1, ..., n_{in} \\
& h_j(x) = 0, \ j = 1, ..., n_{eq}.
\end{aligned} \tag{1.1}
$$

In problem (1.1), the vector $x = (x_1, ..., x_n)$ describes the optimization variable, $f_0(x)$: $\mathbb{R}^n \to \mathbb{R}$ is the objective function to be minimized (or dependent on the problem formulation also maximized), $g_i(x)$ and $h_j(x)$ are inequality and equality constraints, respectively. In the field of systems biology (including parameter estimation for signal transduction pathways), $f_0(x)$ describes a measure for the model misfit between the measurement data and the model predictions. Furthermore, (in-)equality constraints represent the dynamics of the considered systems as well as prior knowledge about the system, such as kinetic parameters requiring to be non-negative.

For solving optimization problems as depicted in (1.1), different approaches and algorithms exist. The effectiveness of these approaches depends on many factors, such as the number of variables and constraints, or the structure of the problem (sparsity, linearity). Notably, problem (1.1) is in general difficult to solve [29].

When modeling biological processes, the resulting system becomes typically nonlinear as biochemical complexes are often of higher order (e. g. due to the assembly of multimeric protein complexes) and, therefore, products of the order of two or even higher appear. Furthermore, often complex kinetics, such as the Michaelis-Menten law are used. Due to these nonlinearities, the resulting optimization problem (1.1) becomes nonconvex. Consequently, the determination of parametrizations for which a model can represent measurement data is not trivial and crucially depends on several assumptions. These assumptions include initial concentrations of model variables or initial guesses for the parameters [174]. Nonconvexity of the optimization problem usually causes the convergence to local solutions (local minima). A local solution of an optimization problem is a solution that is only optimal within a neighborhood of candidate solutions. When it comes to model-based predictions, for instance, the prediction of unmeasurable model states, local solutions can lead to wrong conclusions as the determined parameters estimates may not be optimal.

To avoid the limitations of local optimization methods, global optimization approaches have been developed, for an overview see [103]. They allow the exploration of the whole parameter space, although a conversion of the solution to the global minimum of the optimization function in (1.1) is not guaranteed [63]. In principle,

methods for global optimization can be classified into deterministic and stochastic strategies. Determinstic approaches provide theoretical guarantees for reaching the global minimum within some defined tolerance [52, 68, 123]. Stochastic approaches [4, 163] are used more frequently and employ clustering methods, simulated-annealing or evolutionary algorithms, [9, 83, 85]. In contrast to determinstic methods, stochastic approaches have only weak theoretical guarantees of convergence to the global solution. Notably, also hybrid global optimization strategies exist. Hybrid methods are based on a combination of approaches for local and global optimization methods. In more detail, hybrid approaches apply global search strategies to approach the global solution and in the proximity of the global solution, the optimizer is switched from the global stochastic to a local deterministic search method. An example of such an approach can be comprehended in [11].

To increase the probability of finding *the best* parametrization, methods for local and global optimization are frequently combined with sampling-based approaches, such as Monte Carlo or bootstrapping. To do so, different combinations of initial conditions and parametrizations are tested for their capability of fitting the measurement data. For an overview refer to [99]. However, it may happen that important valid parametrizations are missed due to the stochastic character of the methods.

**Optimal experimental design**
Very often, available measurement data is not sufficient to clearly select a model or to infer the model parameters. Then, new experiments have to be planned (including e. g. the measurement time and stimulus administration) providing additional information about the system under study. The process of predicting the right measurement time and stimulus administration in an experiment is denotend as 'experimental design'. In the following, we review and discuss such model-based approaches for the purpose of model discrimination and parameter estimation.

The design of optimal experiments, or in other words the design of experiments that provide most valuable information helps to discriminate between competing model hypotheses and to identify the unknown kinetic parameters. In general, the more measurement data are available, the more invalid model hypotheses can be identified and the better model parameters can be estimated. However, experiments are typically time-consuming and expensive. Thus, their execution has to be planned carefully.

When planning experiments three relevant aspects should be taken into account: (i) which variables should be measured or manipulated, (ii) when should the measurement be performed, and (iii) which stimulus should be chosen in order to provide maximal information for further model invalidation and parameter estimation. To decipher which species should be measured and when this measurement should be taken (bullets (i)+(ii)) several approaches exist [89]. As one of the most prominent, the *Fisher Information Matrix* (FIM), determines the amount of information that a variable has about an unknown parameter [57]. Using the Fisher Information Matrix,

the measurement content can be evaluated by different measures. The D-optimality criterion, for instance, is calculated by maximizing the determinant of the covariance matrix (i. e. $\text{FIM}^{-1}$) and is used for the minimization of parameter variances. The E-optimality criterion is used for the minimization of the largest parameter variance by minimizing the largest eigenvalue of the covariance matrix $\text{FIM}^{-1}$. For alternative optimality criteria refer to [7, 8]. Depending on the problem, the computation and application of the measures has to be appropriately chosen. A prominent example for employing the Fisher Information Matrix for experimental design is given in [91]. There, the sampling time is optimized by minimizing the variance of the parameter estimation error. To do so, the authors define the Fisher Information Matrix and compute the covariance matrix of the parameter vector. To measure the accuracy of the parameter estimates, they compute the determinant of the Fisher Information Matrix which gives a scalar relating to the volume of the multidimensional simplex. The authors in [53] calculated and compared D-, E- and modified E-optimality criteria for optimal experimental design of the Mitogen-activated protein kinase (MAPK) signaling pathway. Application and computation of the E- and D-optimal experimental designs yielded the best results. Other applications using optimality criteria can be found in [48, 53, 61].

In biological systems, an important and independent variable is the stimulus to initiate a biological process, such as applying cytokines or administrating drugs. This external pertubation is considered as model input and can be easily modified in experiments (bullet (iii)). By applying different pertubation patterns, such as constant or pulse-like inputs in experiments, it can be studied whether or not the model is capable to represent such patterns in the corresponding data, or can be ruled out in the negative case. The design of optimal inputs has been studied, for example, in [40, 149] for model discrimination and in [10, 12, 129] for parameter estimation.

**Multi-scale modeling and data integration**
In the previous sections, we reviewed basic concepts that are used for model discrimination, parameter estimation and experimental design in systems biology. An aspect which was not discussed so far is the modeling and analysis of biological processes that act on different time scales, i. e. short- and long-term time scales. To this end, we provide in the following an overview about multi-scale modeling approaches and applications within the field of systems biology.

Complex biological processes typically span several time scales from second over minutes to days. Therefore, experimental analyses and measurements take into account rapid changes in phosphorylation levels (changing within seconds or minutes) as well as slower changes such as protein production or cell growth (changing from hours to days). Besides this, also the aspect of different data quality on different time scales needs to be considered and integrated into models. Typically, high frequency data for processes on the short-term time scale can be obtained. However, only sparse data

for processes on the long-term time scale are usually provided since it would be too expensive to measure the species of interest with the same frequency as on the short-term time scale. Furthermore, multi-scale modeling approaches have to consider a trade-off between model accuracy and manageable complexity considering the aspects of uncertainties in data, model parameters and state variables.

To devise a predictive mathematical model that represents data for multiple biological phenomena both, short- and long-term time scales need to be taken into account. However, modeling processes with different time scales and performing modeling and parameter estimation for all possible variables and reactions, can lead to an explosion of the problem size and, therefore, often becomes intractable and inapplicable. The choice of suitable methods for tackling such multi-scale modeling depends on the problem at hand. Several methods for multi-scale modeling have been developed over the last decades. Approaches range from the equation-free multiscale method over the multi-grid method to agent-based modeling and Cellular Automata [44]. Notably, also hybrid methods exist combining, for instance, agent-based modeling and Cellular Automata or partial differential equation methods [19].

In [119] and [169], 3-dimensional models for simulating tumor growth using agent-based modeling and cellular automata, respectively were developed. Notably, both approaches needed high computational efforts to solve the underlying models. Additionally, only unrealistic tumor sizes could be computed. Thus, the up-scaling of biological processes and the bridging of time scales is a limiting factor in multi-scale modeling.

The mentioned approaches consider highly complex processes, e. g. tumor growth and tumor vascularization. To understand the complexity how an extracellular signal is translated into the cell causing a certain decision (such as to grow, to shrink, to form vessels) on the long-term time scale, it is initially important to interpret how these signals are encoded during signal transduction on the short-term time scale. To this end, we present a *phenomenological modeling approach* that is tailored to signal transduction and related cellular responses.

**Multi-scale modeling tailored to signal transduction**
In general, *phenomenological modeling* is an approach for creating models, which describe the correlation of certain phenomena and which are difficult to model by first principles. To be more specific, we refer to *phenomenological model* as a mathematical model, which describes the relationship between signal transduction on the short-term time scale and cell fate decisions on the long-term time scale without having a detailed knowledge of the processes bridging the time scales.

For example, the authors in [143] used *phenomenological modeling* to study whether signal properties (i. e. the dynamics of protein activation/phosphorylation) of the MAPK isoforms ERK1/2 that can be used to predict cell fate decisions. They hypothesized that the strength of cell growth is encoded in the integrated response of

ERK phosphorylation, which was experimentally measured. They showed that ERK signaling contributes to cell growth in a dose-dependent and isoform-specific manner. Notably, since the processes between the time scales, such as gene transcription and translation were not modeled in detail, the approach in [143] is computationally low-demanding and thus, applicable to a large number of signal transduction networks, including also IL-6-induced signaling.

In [144], further approaches were proposed for linking short-term signal transduction and long-term cellular responses. As example, also the maximum peak height of a signal or the duration of a signal can be used to link both time scales, short- and long-term events.

**Short summary**
In the previous section, we reviewed approaches, that are frequently used for model invalidation, parameter estimation, and experimental design in systems biology. However, their application is limited due to several reasons. First, mathematical models describing biological processes are typically nonlinear resulting in noncovex optimization problems. Solving these nonconvex optimization problems is not trivial as usually only local parametrizations are obtained. Such local parameter estimates may not be optimal and potentially lead to wrong model predictions and conclusions. Furthermore, model parameters are usually unknown spanning several orders of magnitude and measurement data for inferring these parameters can be sparse, noisy and of different types, e. g. qualitative or categorical. For addressing the point of uncertainties in model parameters and data, often stochastic approaches are applied. These approaches require normally distributed noise as underlying assumption. However, this assumption is typically often not met for real data. As a remedy to these drawbacks, we apply set-based estimation methods in this thesis. The set-based framework allows us to incorporate uncertainties in parameters and measurement data via so-called *unknown-but-bounded* variables. In addition, the method allows to directly include qualitative data/information via a mixed-integer approach. By applying convex relaxations, we can estimate ranges for the unknown model parameters and proof model invalidity globally and with guarantees. To this end, we introduce in the following set-based methods and point out the main differences between classical systems theoretical approaches for model invalidation and parameter estimation and the set-based framework. Furthermore, we give an overview about existing set-based approaches.

## 1.2 Set-based model estimation and analysis

Two of the most important aspects during modeling and analysis of biological processes are the determination of valid paramterizations and the discrimination between model

hypotheses with respect to whether or not they are capable to represent measurement data. This is not trivial due to the mentioned challenges.

Set-based methods can overcome some of these challenges. Such approaches have been in the focus for the analysis, estimation and control of (non-)linear systems over the last years in several technical fields, [35]. Notably, the application of set-based methods in the context of multi-scale modeling has not been in the focus, yet.

### 1.2.1 Comparison of classical and set-based estimation methods

As reviewed in the previous section, many of the existing parameter estimation methods are based on fitting a model to measurement data, and hence, can be reformulated as mathematical optimization problems. The goal of optimization problems is to find the *optimal* solution out of all *valid* solutions. Depending on the obtained optimal parametrizations, different scores can be computed to rank competing model hypotheses accordingly to these scores to determine the most likely hypothesis. However, since model equations describing biological problems are typically nonlinear and the solution sets are usually nonconvex, optimal parametrizations are very hard to determine. It may happen that a local optimum is found leading in a first sight to a good model fit, while further analyses could lead to wrong conclusions regarding the validity of the biological system.

In contrast to classical parameter estimation and model discrimination methods, the goal of set-based approaches is not to find *the optimal* solution. Instead, their purpose is to determine *any* solution, such that the model is capable to represent the uncertain measurement data [135, 136]. Set-based approaches have been applied in different contexts. Notably examples are parameter estimation, hypothesis invalidation and state estimation [26, 34], fault diagnosis and isolation [118, 140], and the analysis of data outliers [25, 153].

For set-based methods, measurement uncertainties, uncertainties in state variables and parameters are described as *unknown-but-bounded* variables which means they belong to sets (characterized by lower and upper bounds) [95]. To approximate the set of all feasible solutions, relaxation techniques are applied resulting in convex problems that can then be solved using, for instance, Cplex, Gurobi or Sedumi [24] – commercial optimization solvers for integer, quadratic and linear programming. The exploration of the complete parameter space, therefore, allows conclusive statements about the invalidity of model hypotheses while guaranteeing that no valid solution is lost. In particular, if the sets of a model are found to be *empty*, it can be guaranteed that no solutions exist that are consistent with the measurement data. In this case, the model can be deemed as guaranteed invalid.

### 1.2.2 Set-based methods for estimation

Set-based approaches for linear dynamic systems have been, for instance, considered in [65] where ellipsoids for determining an enclosure of the model states were derived based on the work presented in [145] and [172]. Thereby, the estimated sets were obtained by applying linear programming algorithms [101]. For estimating nonlinear dynamic systems using set-based methods, approaches such as interval-analysis are often applied [80]. Interval analysis methods were initially developed to analyze and control numerical errors in mathematical computations [105]. Later, they were extended to parameter and state estimation. To do so, model variables and parameters are described via sets for which interval arithmetics can be applied. By checking whether subintervals can be related to the measurement data, enclosures of the variables and/or parameters can be derived guaranteeing that no solutions are lost. For applications see e. g. [80, 104].

The authors in [127] introduced a barrier-certificate set-based method for nonlinear model validation using convex optimization. The method is based on functions of state-parameter-time that are termed as barrier certificates. These barrier certificates are evocative of Lyapunov functions and allow for model (in)validation in the presence of parameter uncertainties in case the predicted model trajectories do not intersect with ranges of measurement data. Further set-based approaches for nonlinear models have been established for the purpose of system identification (i. e. the extraction of a mathematical model from measurement data) in [102], [17], and [55].

An alternative approach for set-based estimation was considered in [90]. Here, the parameter space was divided into valid and invalid regions based on semidefinite programming (see e. g. [116]). The approach was applied to a model with nonlinear mass action kinetics at steady state and consistent steady-state concentrations could be identified. The method has been extended to global steady-state sensitivity analyses for biochemical reaction networks [166]. As in some applications, however, it is not sufficient to only investigate the steady-state behavior of a system, Rumschinski et al. [136] extended the work presented in [90]. The authors demonstrated that their approach can be used for guaranteed model invalidation and parameter estimation of nonlinear dynamical models using convex relaxations.

### 1.2.3 Convex optimization and relaxation

In general, biological processes are highly nonlinear resulting in nonconvex problems for which the solution set is difficult to obtain. To circumvent this, convex optimization can be applied. In general, convex optimization problems are of the form [29]:

$$\begin{aligned}
\text{minimize} \quad & f_0(x) \\
\text{subject to} \quad & g_i(x) \leq 0, \ i = 1, ..., n_{in} \\
& h_j(x) = \alpha_j^T x - b_j, \ j = 1, ..., n_{eq}.
\end{aligned} \tag{1.2}$$

Compared with the general problem (1.1), a convex problem (1.2) must fulfill the following:

- the objective function $f_0(x)$ must be convex,

- the inequality constraint functions $g_i(x)$ must be convex, and

- the equality constraint functions $h_j(x)$ must be affine, where a function $h : \mathbb{R}^n \to R^m$ is affine, if it is a sum of a linear funtion plus a constant, i. e. $h(x) = Ax + b$, where $A \in \mathbb{R}^{m \times m}$ and $b \in \mathbb{R}^m$.

An important property of a convex optimization is, that the resulting feasible set (i. e. the solution space which contains all possible values) is also convex [29]. Thus, in a convex optimization problem, a convex objective function is minimized over a convex set. As a result it follows, that a determined local solution is also a global one.

The philosophy behind the applied set-based methods is to derive the set of all valid solutions by constructing a so-called *feasibility problem* (FP). A FP is a special case of (1.1) with the objective funtion $f_0(x) = 0$ [29] and can be denoted as

$$
\begin{aligned}
&\text{find} &&x \\
&\text{subject to} &&g_i(x) \leq 0, \ i = 1, ..., n_{in} \\
& &&h_j(x) = 0, \ j = 1, ..., n_{eq},
\end{aligned}
\tag{1.3}
$$

where $x \in \mathbb{R}^n$. Due to its nonconvexity properties, a solution for problem (1.3) is hard to derive. One can circumvent this by relaxing (1.3) into a semidefinite program (SDP). Semidefinite programming is a subfield of convex optimization, where a linear objective function is optimized over the intersection of the cone of positive semidefinite matrices with an affine space, denoted as a spectrahedron [116]. For applying SDP, a quadratic (or polynomial) representation of the original set of equalities and inequalities in (1.3) is required (for details see Chapter 4), which then allows a reformulation/relaxation in terms of symmetric matrices [165]:

$$
\begin{aligned}
&\text{minimize} &&tr(CX) \\
&\text{subject to} &&tr(Q_iX) \leq 0, \ i = 1, ..., n_{in,SDP}, \\
& &&tr(R_jX) = 0, \ j = 1, ..., n_{eq,SDP}, \\
& &&X \succeq 0.
\end{aligned}
\tag{1.4}
$$

In (1.4), $X \in \mathcal{S}^n$ and $C, Q_i, R_j \in \mathcal{S}^m$, where $\mathcal{S}^m$ denotes the set of real symmetric $m \times m$ matrices. The operator $\succeq$ denotes a generalized inequality, i. e. among the matrices $Q_1 \succeq Q_2$ implies $Q_1 - Q_2 \in \mathcal{S}^m_+$ with $\mathcal{S}^m_+$ representing the set of real symmetric positive semidefinite matrices [29, 117].

The obtained SDP problems are still convex, however restricted to the set of positive semidefinite matrices [116]. Due to the availability of efficient solvers, such as SeDuMi,

Gurobi and Cplex, problem (1.4) can be solved for a wide class of problems [1, 69, 158]. Yet, the size of the SDP is restricted by the numerical costs of these solvers.

The approach for relaxing problem (1.3) into a SDP and deriving a solution for the SDP can be computationally expensive. Thus, a further relaxation step is necessary. This step relaxes the SDP into a linear program (LP) capable to include much more variables and constraints [135, 152]. Linear programming is a method aiming to optimize the outcome in a mathematical model requiring linear dependencies in the model. To this end, LP optimizes a linear objective function subject to linear equality and inequality constraints. To relax problem (1.4) into a LP, one has to substitute $X \succeq 0$ with the weaker constraint $X \geq 0$ [93].

To sum up, the basic idea of a relaxation-based approach is to substitute nonlinearities with simpler expressions deriving the feasible solution set of an initial nonconvex problem. The substitution of constraints leads to an increase of additional (false positive) solutions. However, it can be guaranteed that any feasible point in (1.3) is also feasible for (1.4) and for the LP, but not vice versa (i.e. no false negative parametrizations). For a detailed description of the relaxation steps for the therein applied set-based methods, we refer to Chapter 4.

### 1.2.4 Outer and inner approximation of feasible sets

Besides model invalidation (i. e. to determine models that are not capable to represent data) using a relaxation approach, set-based methods also aim to determine an approximation of unknown feasible solution sets (including parameters and state variables) for possibly valid models. To this end, we introduce in the following the concept for outer and inner approximations of these feasible solution sets.

The set-based approach can be used to approximate the unknown parameters and/or state variables. In particular, so-called *outer* and *inner* approximations of feasible solution sets can be derived. In both approaches, the set over which a function is to be optimized is approximated in this work by polyhedra. In an outer approximation the polyhedra enclose the feasible set, while in an inner approximation the polyhedra are fully contained in the inner set. As a consequence, outer approximations describe an enclosure of all feasible solutions, but possibly also false positive samples. This might be advantageous for checking model invalidity. However, as not all samples are feasible ones, the approach for deriving outer approximations has to be combined with stochastic routines, such as Monte Carlo sampling. Thereby, the goal is to obtain valid samples, such that the model is capable to represent measurement data. Much more of an advantage are inner approximations for which it is guaranteed that only feasible solutions are contained, but most likely not all.

Inner and outer set approximations have been considered, for instance, in [79] in the context of bounded-error estimation using set inversion and interval analysis. The approach allows to characterize the feasible set for parameters by enclosing it between

internal and external unions of boxes. The estimation of inner approximations has been considered in [88] using occupation measures. These measures allow replacing the dynamics with linear (in-)equalities in an optimization problem over an infinite function space, which is then solved by a hierarchy of semidefinite programs.

## 1.2.5 Bilevel optimization

In this thesis, we also apply the concept of bilevel optimization, which is introduced next.

Bilevel optimization denotes a special kind of optimization problems, where one problem is embedded within another [38]. Bilevel optimization was initially considered for problems dealing with applications in the military field [30] as well as in production and marketing decision making [31]. Within the field of systems biology, bilevel optimization was, for instance, applied for identifying gene knockout strategies [32], optimization of metabolic pathways under stability considerations [37], and optimal profiles of genetic alterations in metabolic engineering [60]. One reason for using bilevel optimization in systems biology is, that the approach is especially suited when more than one decision has to be made in a hierarchical manner. As an example, in [32] bilevel optimization was used to suggest optimal gene deletion strategies (first problem/decision), such that an overproduction of succinate and fructose (second problem/decision) in *Escherichia coli* could be achieved.

We consider a bilevel problem which can be formulated as

$$
\begin{aligned}
&\underset{x,y}{\text{minimize}} && F_0(x,y) \\
&\text{subject to} && G(x,y) \leq 0 \\
& && \underset{y}{\text{minimize}} && f_0(x,y) \\
& && \text{subject to} && g(x,y) \leq 0,
\end{aligned}
\tag{1.5}
$$

where $x \in \mathbb{R}^{n_1}$ and $y \in \mathbb{R}^{n_2}$. The variables of problem (1.5) can be divided into two classes, the upper-level variables $x$ and the lower-level variables $y$. Akin to that, the functions $F_0 : \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \to \mathbb{R}$ and $f_0 : \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \to \mathbb{R}$ are the upper- and lower-level objective functions, respectively. Furthermore, the vector-valued functions $G : \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \to \mathbb{R}^{m_1}$, and $g : \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \to \mathbb{R}^{m_2}$ are called the upper- and lower-level constraints, respectively [14]. Notably, upper-level constraints involve variables from both levels.

Similar to problem (1.1), the embedded optimization problems in (1.5) can be nonconvex. As a consequence, for both problems in (1.5) only local solutions are most likely achievable. As remedy, bilevel optimization can be combined with convex relaxations. Due to the convexification of, e. g. the lower-level of the bilevel optimization problem, a certain behavior of the model can be globally guaranteed. As such it may be possible to make conclusive statements about the feasibility of a model to satisfy

given constraints, i. e. $g(x, y) \leq 0$. For methods and solution approaches of convex bilevel optimization problems, refer to [47, 108, 138].

## 1.3 Research contributions

This thesis uses and expands the results of Rumschinski et al. [135, 136] and Borchers et al. [24, 27] in which a set-based analysis framework was presented. It provides the following major contributions:

i) **Chapter 5 - Set-based estimation of inner parameter sets**
   We develop a method for deriving inner approximations of feasible parameter sets applying mixed-integer programming. We present two algorithms to determine inner approximations and apply these algorithms to two biological examples. The approach allows robust and guaranteed predictions under uncertainties. The results of this chapter were published in [156].

ii) **Chapter 6 - Set-based experimental design for model discrimination and selection**
   We extend the set-based framework by an experimental design approach based on bilevel optimization. The method allows for guaranteed model discrimination of valid and invalid models under uncertainties by designing an optimal input. The results appeared in [134].

iii) **Chapter 7 - Set-based multi-scale modeling and data integration**
   We present approaches for combining biological processes that act on different time scales as well as the integration of uncertain data obtained at these time scales using the set-based framework. We aim to predict long-term processes based on data and processes from the short-term time scale. We further combine the set-based framework with classification methods, which allows to stratify a patient cohort into risk categroies for developing inflammatory diseases based on data from the short-term time scale. The results of this chapter appeared in [133] and [132].

iv) **Chapter 8 - Data-driven set-based parameter estimation of IL-6-induced classic- and trans-signaling**
   We study the responsiveness of IL-6-induced classic- and trans-signaling. To this end, we implement set-based models describing classic- and trans-signaling. These models are tested with respect to their capability of reproducing available measurement data. Additionally, we derive guaranteed outer estimates for the unknown model parameters and study the impact of the cell surface receptors gp130 and IL-6R$\alpha$ on the responsiveness of classic- and trans-signaling. Our results lay the basis for potential approaches targeting IL-6-induced deregulated

signaling and related inflammatory diseases. The results of this study were published in [130].

.

## 1.4 Thesis outline

This thesis is structured as follows:

**Chapter 3** introduces the tasks of modeling and parameter estimation for biological systems under uncertainties.

In **Chapter 4**, the set-based analysis framework that is essential for the following chapters is presented. We review how the tasks for guaranteed model invalidation and parameter estimation can be addressed by relaxation of a feasibility problem.

**Chapter 5** proposes a method for estimation of inner parameter sets within the set-based framework. To do so, we show how to reformulate the feasibility problem from Chapter 4 using binary variables and combinations thereof. The performance of the developed algorithms for estimating inner parameter sets are studied using two examples.

**Chapter 6** presents a set-based experimental design approach for guaranteed discrimination of valid and invalid models. We develop a bilevel optimization approach allowing for the selection of competing model hypotheses under uncertainties.

**Chapter 7** introduces the problem of multi-scale modeling and data integration under uncertainties. We apply a phenomenological approach to combine processes from the short- and long-term time scale using the set-based framework. Furthermore, we introduce an unified framework that allows for the stratification of patients into risk categories for developing inflammatory diseases by combining set-based methods with classification approaches.

In **Chapter 8**, we present a large pathopyhsiological example describing IL-6-induced classic- and trans-signaling. The example demonstrates applicability of the set-based framework. Moreover, new biological insights in IL-6-induced signaling and a possible intervention strategy for targeting deregulated IL-6-induced classic- and trans-signaling are provided.

Finally, **Chapter 9** summarizes and concludes this thesis. The chapter further includes an outlook on possible research topics related to this thesis.

# 2 Running Example: IL-6-induced Signal Transduction

Throughout this thesis, we apply the developed set-based methods to IL-6-induced signaling. We focus on the Janus kinase/Signal Transducers and Activators of Transcription (Jak/STAT) and the Mitogen-Activated Protein Kinase (MAPK) pathways, which are introduced and explained next. For a comprehensive review of IL-6-induced signaling see [72].

**IL-6-induced Jak/STAT signaling**

Interleukin-6 (IL-6) is a pleiotropic cytokine which is involved in many cellular functions such as cell growth, apoptosis and differentiation [82]. Additionally, IL-6 is responsible for the stimulation of the acute phase protein (APP) synthesis [175], which is part of the innate immune response. APPs play an important role in mediating systemic inflammatory effects, such as fever and leukocytosis, which is an increase in the number of white blood cells. Increased expression of IL-6 and dysregulation of IL-6-induced signaling can lead to numerous pathological states including inflammatory diseases, such as rheumatoid arthritis and multiple sclerosis [39, 54].

Initiation of IL-6 signaling occurs through two different pathways, [131] (cf. Fig. 2.1). During *classic-signaling*, IL-6 binds in a first step to the membrane-bound receptor subunit gp80 (IL-6R$\alpha$) followed by the recruitment of the signal-transducing receptor subunit glycoprotein 130 (gp130). Thereby, a hexameric receptor complex (R$_{complex}$) containing two molecules of each IL-6, IL-6R$\alpha$ and gp130 is formed [28]. Also a soluble form of IL-6R$\alpha$, denoted as sIL-6R$\alpha$, exists [107]. SIL-6R$\alpha$ can also form active hexameric receptor complexes by binding to IL-6 and gp130, thereby initiating *trans-signaling*. Both pathways, classic- and trans-signaling are considered to converge in the activation of the Jak/STAT pathway. The formation of one of the active receptor complexes induces activation of receptor-associated Jak and subsequent phosphorylation of tyrosine residues (i. e. the addition of a phosphate group) within the cytoplasmatic part of gp130. The phosphorylated tyrosine residues (represented as (p)R$_{complex}$, light orange stars Fig. 2.1) serve as docking sides for molecules with SH2-domains, such as STAT. Upon recruitment to the receptor, monomeric STAT proteins become phosphorylated by Jaks. Subsequently, phosphorylated STAT proteins dimerize (not shown in Fig. 2.1) and translocate into the nucleus. There, they induce the transcription of pre-messenger ribonucleic acids (pre-mRNAs), which convey the genetic information from a particular section of the deoxyribonucleic acid
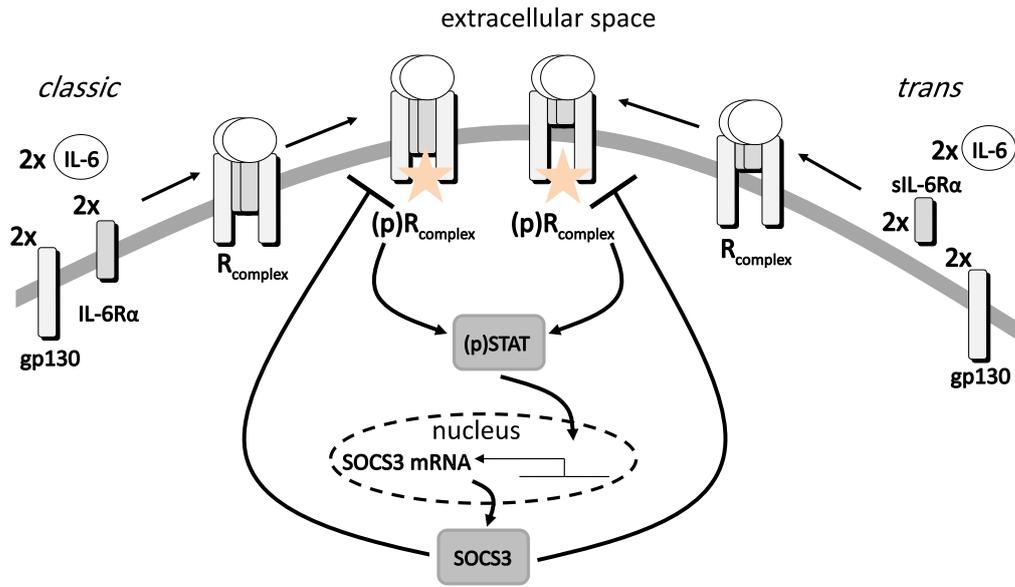
Figure 2.1: **Schematic representation of the IL-6-induced Jak/STAT signaling pathway.** During classic- and trans-signaling, IL-6 binds to the corresponding receptors causing receptor complex formation and activation (light orange stars). Receptor activation induces activation of Jak/STAT signaling and subsequent transcription of target genes.

(DNA) to the ribosom for protein synthesis. As such target protein suppressor of cytokine signaling 3 (SOCS3) is produced. SOCS3 is known as feedback inhibitor of Jak proteins [51]. This negative feedback leads, together with dephosphorylation of the active receptor complex via phosphatases, to a switch-off of Jak/STAT activation, resulting in a transient activation of Jak/STAT signaling. Notably, above we have only described processes of IL-6-induced pathway activation and translation of proteins (e. g. SOCS3) acting on the short-term time scale, i. e. within approximately 1.5 hours. However, IL-6 stimulation induces also the transcription of genes, that code for long-term processes, such as growth, differentiation or angiogenesis ranging from hours to days. Different long-term responses of both pathways, classic- and trans-signaling, have been described. It is well studied that while classic-signaling acts mainly anti-inflammatory [142, 160], trans-signaling has pro-inflammatory activities and is associated with inflammatory diseases [15, 96].

**IL-6-induced MAPK signaling**
Another major pathway of IL-6-induced signaling is the MAPK signaling cascade. MAPKs are protein kinases that specifically phosphorylate side chains of the amino acids serine and threonine. Several MAPK pathways exist. One of the most prominent is the Ras-Raf-Mek-ERK cascade [146]. Ras-Raf-Mek-ERK signaling is induced by various growth factors and cytokines including IL-6. Canonical activation of the Ras-Raf-Mek-ERK cascade can be described as follows (cf. Fig. 2.2):

Due to the activation of Jaks and the subsequent phosphorylation of gp130 akin to above Jak/STAT signaling, the SH2-containing protein tyrosine phosphatase 2 (SHP2) is recruited to phosphorylated tyrosine residues within the cytoplasmatic part of gp130. Subsequently, SHP2 is phosphorylated ((p)SHP2) acting as an adaptor protein for several proteins, including Growth factor receptor-bound protein 2 (Grb2). Grb2 is constitutive associated with SOS (Son Of Sevenless), which is a guanine nucleotide exchange factor activating the small G-protein Ras. Small G-proteins can be seen as signaling switches with 'active' and 'inactive'



Figure 2.2: Schematic representation of the IL-6-induced Ras-Raf-Mek-ERK cascade via the adaptor protein SHP2 and the Grb2/SOS complex.

states. In the 'inactive' state Ras is bound to the nucleotide guanosine diphosphate (GDP), while in the 'active' state, Ras is bound to guanosine triphosphate (GTP). SOS forces Ras to release GDP. Subsequently, Ras binds to GTP resulting in Ras activation (Ras$^*$). Ras$^*$ interacts with, and stimulates downstream signaling effectors, including Raf (Rapidly acclerated fibrosarcoma). Raf is activated to Raf$^*$ and phosphorylates Mek (Mitogen-activated protein kinase kinase), which is subsequently activated to Mek$^*$. Mek$^*$ in turn phosphorylates ERK (Extracellular signal-regulated kinase) to ERK$^*$. ERK$^*$ activates a number of transcription factors on the short-term time scale. Additionally, activated and phosphorylated ERK is known to be involved in the regulation of IL-6 synthesis [20] and its deregulation was associated with long-term joint damage [43].

## 2.1 Challenges and research contributions

IL-6 is a key regulator of inflammatory processes and has been identified as a therapeutic target, e. g. [113]. However, current strategies to block misbalanced IL-6-induced signaling, e. g. by monoclonal antibodies, such as sarilumab and siltuximab, that are directed against IL-6 itself or tocilizumab, that is specific to block IL-6R$\alpha$, can have many negative side effects for patients [62, 92]. This is because these antibodies block both, classic- and trans-signaling, and thus, also important physiological processes [33, 112]. Due to the negative side effects of existing therapeutic approaches, more suitable and targeted intervention strategies are needed. However, the development of such strategies is not trivial due to the complex underlying molecular signaling

processes whose misregulations in diseases are not yet fully understood. This work contributes to a deeper understanding of IL-6-induced signaling and deregulations on a systems level. Using set-based modeling and estimation together with biochemical analyses, this thesis aims to:

a) **derive quantitative and predictive models** of IL-6-induced pathway activation, receptor assembly and downstream signaling in classic- and trans-signaling.

b) **estimate unknown model parameters** that describe the dynamics of processes responsible for deregulated signaling.

c) **analyze differences in the responsiveness** of IL-6-induced classic- and trans-signaling.

The obtained results lay the basis for the development of novel and model-guided intervention strategies that are tailored to IL-6-induced misbalanced signaling and associated inflammatory diseases.

# 3 Modeling, Analysis and Parameter Estimation under Uncertainties

In this work we aim to derive a quantitative understanding of biological processes. Biochemical reaction networks in general and signal transduction pathways in particular are commonly modeled using mass action kinetics.

In the following, we provide the background for modeling such reaction networks. Moreover, we introduce the uncertainty description to account for errorneous data as well as the notation of competing model hypotheses and uncertain parameters.

## 3.1 Modeling biological systems

Typically, a reaction network consists of different state variables (denoted as 'species') $X_1, ..., X_{n_x}$, for instance proteins, metabolites or nucleic acids. The interactions among these species are denoted as *reactions*. Such reactions may result in the phosphorylation of proteins, the formation of protein or nucleic acid complexes, the degradation of species and many more [86]. Notably, model species are characterized in terms of concentrations, i. e. $[x_i]$ for $i \in \{1, ..., n_x\}$. For simplicity, we write

$$x = ([x_1], ...., [x_{n_x}])^T \in \mathbb{R}^{n_x} \tag{3.1}$$

comprising the concentration of the species $[x_i]$ that vary in time.

In the following, we assume the concentrations of the model species to be sufficiently large such that stochastic effects can be neglected. Additionally, we consider the species $X_i$ to be equally distributed assuming the absence of a spatial concentration gradient. Let us consider $n_x$ species interacting with each other. The species are converted by chemical reactions of the form

$$\alpha_{1j}X_1 + \alpha_{2j}X_2 + ... + \alpha_{n_xj}X_{n_x} \underset{r_j^-}{\overset{r_j^+}{\rightleftharpoons}} \beta_{1j}X_1 + \beta_{2j}X_2 + ... + \beta_{n_xj}X_{n_x}. \tag{3.2}$$

Here, $\alpha_{ij}$ and $\beta_{ij}$ for $i \in \{1, ..., n_x\}$ are the non-negative stoichiometric coefficients and $j \in \{1, ..., n_r\}$ is the reaction index. Furthermore, $r_j^+$ denotes the forward and $r_j^-$ the backward reaction rate for reaction $j$. For modeling reaction rates, several underlying kinetic laws are assumed (e. g. Michelis-Menten, Hill kinetics, for an overview see [41]). One simple and common approach to model reaction rates is to apply the law of mass action, where the reaction rate is assumed to be proportional to the substrate

concentrations, i. e.

$$r_j(t) = p_j^+ \prod_{i=1}^{n_x} x_i^{\alpha_{ij}}(t) - p_j^- \prod_{i=1}^{n_x} x_i^{\beta_{ij}}(t). \tag{3.3}$$

In Eq. (3.3), $p_j^+$ and $p_j^-$ denote the forward and backward reaction constants.

With this notation, the biochemical reaction network can be written in a compact form as

$$\dot{x}(t) = Nr(t), \tag{3.4}$$

where $\dot{x}(t)$ denotes the change of concentration of a species with time. Furthermore, all reaction rates are collected into the vector $r(t) = (r_1(t), ..., r_{n_r}(t))$. The matrix $N \in \mathbb{R}^{n_x \times n_r}$ is the so called stoichimetric matrix given by

$$N_{ij} = \beta_{ij} - \alpha_{ij}, \quad i \in \{1, ..., n_x\}, j \in \{1, ..., n_r\}. \tag{3.5}$$

Besides the time-dependent model states $x(t)$, we can furthermore include time-dependent inputs $u(t)$. In a biological context, they may represent external stimuli, such as application of cytokines and/or drugs, potentially with time-varying concentrations, such as applying a transient stimulus.

With the variables and definitions given, the overall dynamics are given by a system of ordinary differential equations, i. e.

$$\dot{x}(t) = Nr(t) = f(x(t), u(t), p). \tag{3.6}$$

In Eq. (3.6), $x(t) \in \mathbb{R}^{n_x}$, $u(t) \in \mathbb{R}^{n_u}$, and $p \in \mathbb{R}^{n_p}$ denotes the time-independent parameter vector collecting all rate constants $p_j^+$ and $p_j^-$. Throughout this work, we assume that $x(t)$ belongs to the set $\mathcal{X} \subseteq \mathbb{R}^{n_x}$, $u(t)$ to $\mathcal{U} \subseteq \mathbb{R}^{n_u}$, and $p$ to $\mathcal{P} \subseteq \mathbb{R}^{n_p}$, which means that $x(t)$, $u(t)$ and $p$ are bounded. Furthermore, in Eq. (3.6), $f : \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \times \mathbb{R}^{n_p} \to \mathbb{R}$ is a nonlinear vector function.

It is noteworthy, that often not all model species can be measured experimentally. Therefore, we distinguish between model states $x(t)$ that can be not measured and model outputs $y(t) \in \mathbb{R}^{n_y}$ that can be measured. For simplification, all outputs (measurements) are aggregated into the nonlinear vector function $g$, i. e.

$$y(t) = g(x(t), u(t), p) \tag{3.7}$$

where $g : \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \times \mathbb{R}^{n_p} \to \mathbb{R}$.

Summarizing Eq. (3.6) and Eq. (3.7), one obtains the following overall system

$$h : \begin{cases} \dot{x}(t) &= f(x(t), u(t), p) \\ y(t) &= g(x(t), u(t), p), \end{cases} \tag{3.8}$$

where $h$ stands for *hypothesis*.

For many biological processes the law of mass conservation applies, which means that the mass of a closed system does not change over time if no external supplies and losses are present. Mass conservation can be expressed by

$$x_j^T = \sum_{i=1}^{n_x} \gamma_{ij} x_i, \quad j = 1, ..., n_c, \tag{3.9}$$

where $x_j^T$ denotes the total, conserved amount of a model species $x_i$ and $\gamma_{ij}$ are non-negative coefficients. Conserved moieties can be used for expressing the model state variables $x_i$ by linear algebraic equations. These algebraic equations can then be inserted into the ODE system lowering its dimension by reducing the system's order from initial $n_x$ to $n_x$-$n_c$ state variables.

In the following, we exemplify the above described concepts for modeling biological processes by means of the MAPK pathway.

### 3.1.1 Example: IL-6-induced MAPK pathway

We consider in the following the MAPK pathway as described in Chapter 2. Thereby, a signal is transduced from the activated IL-6 receptor into the cytoplasm of a cell and eventually into its nucleus. This process is executed by enzymes, denoted as kinases and phosphatases leading either to the target proteins' activation or deactivation.

A simplified version of the Ras-Raf-Mek-ERK reaction network (i. e. without a detailed description of receptor assembly) can be described as

$$
\begin{aligned}
u + Ras &\underset{r_2}{\overset{r_1}{\rightleftharpoons}} Ras^* \\
Ras^* + Raf &\underset{r_4}{\overset{r_3}{\rightleftharpoons}} Raf^* \\
Raf^* + Mek &\underset{r_6}{\overset{r_5}{\rightleftharpoons}} Mek^* \\
Mek^* + ERK &\underset{r_8}{\overset{r_7}{\rightleftharpoons}} ERK^*,
\end{aligned}
\tag{3.10}
$$
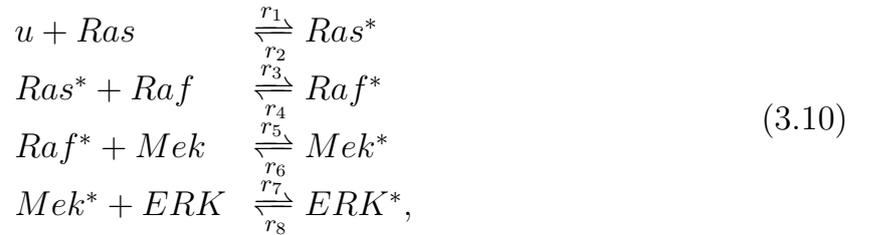
where $u$ describes the input to the system, i. e. IL-6.

The reaction network dynamics above can then be modeled as

$$
\begin{aligned}
[\dot{Ras}] &= r_2 - r_1 \\
[\dot{Ras^*}] &= r_1 - r_2 - r_3 \\
[\dot{Raf}] &= r_4 - r_3 \\
[\dot{Raf^*}] &= r_3 - r_4 - r_5 \\
[\dot{Mek}] &= r_6 - r_5 \\
[\dot{Mek^*}] &= r_5 - r_6 - r_7 \\
[\dot{ERK}] &= r_8 - r_7 \\
[\dot{ERK^*}] &= r_7 - r_8.
\end{aligned}
\tag{3.11}
$$

The states of the system are given by

$$x = \left([Ras], [Ras^*], [Raf], [Raf^*], [Mek], [Mek^*], [ERK], [ERK^*]\right)^T \in \mathbb{R}^8$$

and the reaction rates are $r_j$, $j = \{1, ..., 8\}$. Assuming an equal distribution of the proteins and by neglecting concentration gradients, one can apply the law of mass action, leading to the reaction rates:

$$r_1 = p_1 u Ras, \ r_2 = p_2 Ras^*, \ r_3 = p_3 Ras^* Raf, \ r_4 = p_4 Raf^*,$$
$$r_5 = p_5 Raf^* Mek, \ r_6 = p_6 Mek^*, \ r_7 = p_7 Mek^* ERK, \ r_8 = p_8 ERK^*,$$

where $p_j$, $j = \{1, ..., 8\}$ are the reaction constants. Under the assumption that no species of the system (3.10) is added or removed, the following relations for the conserved moieties hold:

$$\begin{aligned} Ras^T = Ras + Ras^*, \ Raf^T = Raf + Raf^*, \\ Mek^T = Mek + Mek^*, \ ERK^T = ERK + ERK^*. \end{aligned} \tag{3.12}$$

Using (3.12) the state vector $x$ can be reduced to

$$x^r = \left([Ras^*], [Raf^*], [Mek^*], [ERK^*]\right)^T \in \mathbb{R}^4.$$

We assume that the activity of all kinases and the small G-protein Ras* can be measured experimentally, the system outputs are given by

$$y = \left([Ras^*], [Raf^*], [Mek^*], [ERK^*]\right).$$

Having modeled the biological signal transduction process it is of interest if this model allows to describe the reality. This is based on measurement data. Thus, we describe in the following the data that are used to show (in-)validity of a mathematical model and to infer the usually unknown model parameters.

## 3.2 Uncertain data, models and parameters

When modeling and analyzing biological processes two major questions need to be addressed. Specifically, we ask for the possibility to distinguish between valid and invalid models and to determine parameter estimates for reliable and robust predictions. Answering these two questions is challenging due to the fact that the structure of the model itself can be uncertain (e. g. a species is mistakenly neglected or taken into account) or data that are used to infer biological models and parameters can be errorneous [94].

### 3.2.1 Measurement data and error description

In general, variation within measurement data is generated by technical and biological variations. Technical variations are due to uncertainties in measurement processes. These uncertainties can arise, for instance, from the imprecision of the measurement techniques and the fact that experiments cannot be repeated sufficiently often since they are time-consuming and expensive [136].

Biological variation occurs due to all factors in an experiment determined by the biochemistry in a living system under study. Due to, e. g. the stochasticity of translation/transcription which causes cell to cell variability, the results for replicated experiments can have large variations. These variations are also denoted as inherent noise. Notably, measurement errors can become very large due to the superposition of both, technical and biological variations.

Provided knowledge about the used measurement devices and there characteristics, technical errors can be modeled as follows.

*Absolute errors:* An absolute error is a measure for the distance between a measurement $\bar{y}_i(t_m)$ and the real value $y_i(t_m)$. It is resembled by adding a disturbance $\kappa_a \in \mathbb{R}_+$ to the measurement $\tilde{y}_i(t_m)$ leading to an uncertainty interval of

$$y_i(t_m) := [\tilde{y}_i(t_m) - \kappa_a, \tilde{y}_i(t_m) + \kappa_a], \tag{3.13}$$

where $t_m = \{0, 1, ..., n_t\}$ describes the vector of time-points where measurements are taken. Note, the choice of $\kappa_a$ is problem specific.

*Relative errors:* A relative error is an error where the uncertainty grows with the measured value itself. It can be modeled by a disturbance $\kappa_r \in \mathbb{R}_+$ satisfying $0 \leq \kappa_r \leq 1$. The respective uncertainty interval is then given by

$$y_i(t_m) := [(1 - \kappa_r)\tilde{y}_i(t_m), (1 + \kappa_r)\tilde{y}_i(t_m)]. \tag{3.14}$$

For demonstrating the usefulness of our methodological results, we use simulated measurement data with artificial relative errors within the framework of this thesis. In addition to that, we also use real data obtained by Western Blotting, flow cytometry and growth assays to compare the results. Usually, for all experimental readings, we give the mean value $\pm$ standard deviation at all time points $t_m$ to derive the 1-sigma confidence interval [176]. Hence, the measurement errors can be represented by lower bounds $\underline{y}_i$ and upper bounds $\overline{y}_i$ and we use the notation

$$y_i(t_m) := [\underline{y}_i(t_m), \overline{y}_i(t_m)]. \tag{3.15}$$

Eq. (3.15) gives a so-called worst-case uncertainty description which includes technical errors and biological noise.

### 3.2.2 Structural uncertainties and hypotheses

In a biological system it is often uncertain if any species have been missed that play an essential role during signaling. Furthermore, it is likely that the network structure does not resemble the structure of the biological system correctly. Therefore, usually several model hypotheses exist [136], which are given by:

$$h^{[i]} : \begin{cases} \dot{x}^{[i]}(t) & = & f^{[i]}(x^{[i]}(t), u(t), p^{[i]}) \\ y^{[i]}(t) & = & g^{[i]}(x^{[i]}(t), u(t), p^{[i]}), \end{cases} \tag{3.16}$$

where the superscript $[i]$ with $i \in \mathcal{I} := \{1, \ldots, n_h\}$ is used as an index of the different model hypotheses $\mathcal{H} := \{h^{[1]}, h^{[2]}, ..., h^{[n_h]}\}$. The different model hypotheses $h^{[i]}$ have then to be tested whether they are able to represent the measurement data or not.

### 3.2.3 Parameter uncertainties

To test whether a model hypothesis $h^{[i]}$ is (in-)valid, it is necessary to estimate the unknown parameters. Parameter uncertainties can be very large, since often limited or no knowledge about the real values is available [87]. We will use interval/set-based bounds for describing parameter uncertainties:

$$p_j := [\underline{p}_j, \overline{p}_j], \tag{3.17}$$

where $j \in \mathcal{J} := \{1, ..., n_p\}$, and $\underline{p}_j$ and $\overline{p}_j$ describe the lower and upper bounds for the $j-th$ parameter.

Usually in biology, some *a priori* knowledge about the processes and the corresponding reaction kinetics is available. As an example, phosphorylation of proteins are generally fast reactions as they occur within seconds to minutes [49, 76]. Hence, the *a priori* uncertainty interval from (3.17) can be reduced to smaller intervals.

## 3.3 Summary

We here described the background for modeling biological processes and exemplified the concepts by means of the MAPK pathway. Furthermore, uncertainty descriptions to account for errorneous measurement data and unknown parameters as well as the notation of competing model hypotheses were introduced.

In the following chapters, we present methods to tackle the questions of parameter estimation and model discrimination using set-based estimation methods. To do so, the employed set-based approach for model analysis and estimation is introduced in the next chapter.

# 4 Set-based Model Analysis and Estimation

In this chapter, we give an overview about the applied set-based estimation approach. The basic idea consists of a reformulation of the considered problem into a feasibility problem. The feasibility problem aims to determine feasible solutions that fulfill the given constraints including constraints for model parameters and measurement data. Usually, the feasibility problem is nonlinear and nonconvex, and thus hard to solve. To overcome this problem, the feasibility problem is relaxed into a semidefinite or linear program, which can be solved efficiently using state-of-the art solvers. The presented approach lies the foundation for our further methodological analyses and extensions which will be applied to IL-6-induced signaling.

The chapter is structured as follows. First, we describe the formulation of the model invalidation and parameter estimation problems in terms of a feasibility problem. Second, we review the relaxation steps to transfer the nonlinear and nonconvex feasibility problem into a convex semidefinite (or linear) program. After that, the problem for model invalidity and outer approximation of parameters is discussed. Before concluding the chapter, we give an example for application of set-based approaches in modeling of stepwise receptor assembly and activation – a common motif for initializing cell signaling pathways in biological systems.

## 4.1 Introduction

One of the most challenging issues during the modeling and estimation process of biological systems is the discrimination of mathematical models that are consistent from the ones that are inconsistent with the given uncertain measurement data. In addition, for models that are consistent with measurement data, the model parameters are often not experimentally determinable and have to be estimated from the uncertain data. Model invalidation and parameter estimation are crucial tasks for reliable model-based predictions. For example, robust solutions of valid models leading to reliable predictions are indispensable for generating new hypotheses. New hypotheses allow the discovery of molecular target and thus, allow the design of potentially new intervention strategies.

The goal in a set-based framework is not to find the *best* solution (e. g. parameters), but rather to determine *any* solution and, if solutions exist, to determine *all of them.* In this context, variables (e. g. model parameters and states) are described as *unknwon-but-bounded.*

Rumschinski et al. [136] introduced a set-based approach for dynamical systems, which we expand and use. By reformulating the model invalidation and parameter estimation task in terms of a feasibilty problem, the approach allows for conclusive statements about the invalidity of models. Furthermore, if a model is not proven to be invalid, the approach allows an efficient outer approximation of the feasible parameter space using a recursive bisectioning and a sequentially outer-bounding algorithm.

In the following, the set-based approach for model invalidation and parameter/state estimation from [136] is introduced and explained in more detail.

## 4.2 Feasibility problem formulation

Checking if a mathematical model is consistent with measurement data requires to test if there exists any solution leading to a consistent system behavior. Here consistent system behavior means that the model is capable to represent the measurement data, which in the context of this thesis means that it cannot be invalidated. To do so, in a first step we derive a discrete-time approximation of the considered model as in (3.8). This can be achieved by standard integration methods such as Euler or Runge-Kutta discretizations. Nevertheless, selecting an appropriate discretization scheme is very challenging as it introduces discretization errors and may result in numerical instability. As in depth consideration of this issue is out of the scope of this thesis, we refer to [157] and [136] and assume in the following that a model is given in its implicit formulation:

$$
h : \begin{cases} f(x(k{+}1), x(k), u(k), p) = 0, \\ g(y(k), x(k), u(k), p) = 0. \end{cases} \tag{4.1}
$$

Thereby, $x(k)$, $u(k)$, $y(k)$ and $p$ denote the model states, inputs, outputs and parameters. These variables are equivalent to the continuous-time system, but bounded by semialgebraic sets instead of continuous equality and inequality constraints. Furthermore $k \in \mathbb{N}$ denotes the discrete-time index, and $f$ and $g$ are polynomials.

To formulate the model invalidation and estimation task, all the information is gathered into a FP (cf. problem (1.3)):

$$
\text{FP} : \begin{cases} \text{find} & \xi \\ \text{subject to} & f(x(k{+}1), x(k), u(k), p) = 0, \\ & g(y(k), x(k), u(k), p) = 0, \\ & x(k) \in \mathcal{X}, \ k \in \mathcal{T}, \\ & u(k) \in \mathcal{U}, \ k \in \mathcal{T}, \\ & y(k) \in \mathcal{M}(t_k), \ k \in \mathcal{T}, \\ & p \in \mathcal{P}, \end{cases} \tag{4.2}
$$

where $\xi \in \mathbb{R}^{n_\xi}$ is a vector containing all time-dependent and time-independent variables

in the problem, i. e. $\xi = f(x, u, p)$ and where $\mathcal{M}(t_k)$ denotes the measurement data at time point $k$.

The general idea is now to check whether the FP admits a solution or not, which reflects the capability of model (4.1) to satisfy the given constraints and thus the measurements. In the following, we refer to the theorem for checking inconsistenty of FP as given in [24]:

**Theorem 1** *(FP inconsistency certificate).*
*If the FP admits no solution, then there exists no $p \in \mathcal{P}$ and $x(k) \in \mathcal{X}$ such that $y(k) \in \mathcal{M}(t_k)$ and $u(k) \in \mathcal{U}$, $k \in \mathcal{T}$.*

Until now, we focused on a quantitative description of measurement data for stating the FP and checking inconsistency of a model. Nevertheless, often also qualitative information exist that need to be integrated into the FP. To do so, special reformulations of the measurement set $\mathcal{M}(t_k)$ are necessary. These reformulations were presented in [135] and are shortly reviewed next.

### 4.2.1 Formulation of qualitative information

Often, only qualitative information on data are available. Such data might be given by *if-then* observations or statements, such as "if the input is applied, then the concentration of protein A increases". In classical parameter estimation and model invalidation approaches, the inclusion of such qualitative information is challenging. Nevertheless, theses types of data might be beneficial to obtain a valid model of the system under study and therefore, need to be taken into account.

The inclusion of qualitative information into the FP can be done by additional binary variables $\phi \in \{0, 1\}$ [135]. To then check if data represented by *unknown-but-bounded* variables $y(k)$ at time points $t_k$ of measurements are fulfilled, we have to formulate a constraint such that $\phi$ equals 1 if and only if $y(k) \in \mathcal{M}(t_k)$. This relation can be formulated as

$$(\phi = 1) \Leftrightarrow y(k) \in \mathcal{M}(t_k). \tag{4.3}$$

To be able to add (4.3) to the FP from the previous section, further reformulations of the above constraint are necessary. These reformulations depend on the definition of the set $\mathcal{M}(t_k)$. In case $\mathcal{M}(t_k)$ is a halfspace (with any $c \in \mathbb{R}^{n_y}$), one obtains $\mathcal{M}(t_k) = \{\alpha, y(k) \in \mathbb{R}^{n_y} : \alpha^T y(k) \geq c\}$. Consequently, (4.3) can be formulated as:

$$\begin{aligned} \phi &\geq \frac{\alpha^T y(k) - c}{M}, \\ \phi &\leq \frac{\alpha^T y(k) - c}{M} + 1, \end{aligned} \tag{4.4}$$

with $M > |\max_{y(k) \in \mathcal{M}(t_k)} \alpha^T y(k) - c|$ describing a suffiently large number. Notably, on the boundary $\phi$ can be either 0 or 1. This so-called ambiguity can be avoided by introducing further scaling factors [84, 135].

The equality constraint from (4.3) and the inequality constraints in (4.4) can be added to the FP enforcing that $\alpha^T y(k) \geq c$ holds. The above introduced constraints $\phi$ are denoted as integrality constraints taking either 0 or 1 as values. Due to these integrality constraints that allow to include qualitative information by binary variables, the FP is denoted as a mixed-integer FP (MIFP).

Also much more complicated sets, such as sets restricted by polytopes which are defined as the interesection of halfspaces, can be considered. We refer to [135] for more details.

Problem (4.2) (and thus, also the MIFP) is typically nonconvex. This is due to the nonlinearity of the equality constraint functions $f$ and $g$ and the possible nonconvex sets $\mathcal{X}, \mathcal{U}, \mathcal{M}$ and $\mathcal{P}$. Hence, it is difficult to derive a solution for the FP (and also the MIFP). However, (4.2) can be convexified by relaxing the FP into a SDP and LP, as explained next. Note, in the following we focus on the FP, yet, the same steps apply for relaxation of the MIFP.

## 4.3 Problem relaxation and infeasibility certificates

In general, the convexification and relaxation of optimization problems, such as in (4.2) requires several steps, which we detail in the following.

The approach for relaxing problem (4.2) into a SDP is derived from a relaxation method proposed in [90]. To this end, the first step consists of deriving a quadratic representation of (4.2) using *quadrification*, i. e. a quadratic reformulation of the constraints in (4.2) [147].

**Quadratic reformulation**
The basic idea for quadrification is to express every equation of the functions $f$ and $g$ in (4.2) in a quadratic form. The conditions for $f$ and $g$ are then written as $\xi^T Q_i \xi = 0$ and $\xi^T R_j \xi = 0$, respectively, with $i \in \{1, ..., n_f\}$ and $j \in \{1, ..., n_g\}$ by means of appropriate symmetric matrices $Q_i, R_j \in \mathcal{S}^{n_\xi}$. Moreover, bounds on the variables $x(k) \in \mathcal{X}$, $u(k) \in \mathcal{U}$, $y(k) \in \mathcal{M}(t_k)$, and $p \in \mathcal{P}$ can be easily expressed using a set of linear (non-negative) constraints $B\xi \geq 0$, for a suitable matrix $B$. As a consequence, in a quadratic representation, monomials are defined as products of lower degree monomials, i. e. $\xi^T D_j \xi = 0$, $j = \{1, ..., n_d\}$ with $n_d$ being the number of such dependencies.

Now, the FP can be rewritten in a quadratic manner of the form:

$$\text{QP}: \begin{cases} \text{find} & \xi \\ \text{subject to} & \xi^T Q_i \xi = 0, \\ & \xi^T R_j \xi = 0, \\ & \xi^T D_j \xi = 0, \\ & \xi_1 = 1, \\ & B\xi \geq 0. \end{cases} \tag{4.5}$$

Together with the equality constraint $\xi_1 = 1$ (4.5) leads to a quadratic, yet still nonconvex problem.

**Semidefinite relaxation**

As the QP in (4.5) is still nonconvex, the goal is to find suitable reformulations, such that QP can be convexified. By reformulating the constraints in (4.2) in terms of a symmetric matrix $X = \xi\xi^T$ with $X \in \mathcal{S}^{n_\xi}$, QP can be relaxed into a SDP. Thereby, the matrix $\mathcal{S}^{n_\xi}$ is composed of monomials needed to represent the inequalities and with the rank condition one. Further, the relation $\xi^T Q_i \xi = tr(Q_i \xi\xi^T)$ is used and the conditions $rank(X) = 1$ and $tr(X) \geq 1$ are replaced by the weaker constraint $X \succeq 0$. Consequently, problem (4.5) can be reformulated into a SDP:

$$\text{SDP}: \begin{cases} \text{find} & X \\ \text{subject to} & tr(Q_i X) = 0, \\ & tr(R_j X) = 0,, \\ & tr(D_j X) = 0, \\ & tr(e_1 e_1 X) = 1, \\ & BXe_1 \geq 0, \\ & BXB^T \geq 0, \\ & X \succeq 0, \end{cases} \tag{4.6}$$

where $e_1 = (1,0,...,0)^T \in \mathbb{R}^{n_\xi}$ are the unity base vectors. Problem (4.6) is a relaxation of (4.5) due to the weaker, but now convex constraints introduced [116].

It is noteworthy, that the relaxation from problem (4.2) into (4.6) leads to an increase of the space of feasible solutions. This means that in the case of parameter estimation, additional (false positive) parametrizations are introduced, potentially leading to wrong model-based predictions. However, it is always guaranteed that – if there exist feasible parametrizations – none of them are lost (no false negative parameter sets). One way to tighten the increased space of feasible solutions, which means to reduce the number of false positive solutions, is to introduce redundancy constraints $BXB^T \geq 0$ (see (4.6)) [90].

**Lagrange duality and infeasibility certificates**
A semidefinite program can be solved very efficiently using appropriate solvers. However, the problem sizes that might be considered can lead to computational challenges. A more tractable approach can be obtained by solving the Lagrangian dual of the SDP by standard-dual interior-point methods [110].

The key idea is to define a Lagrange function that includes the constraints of problem (4.2) into the cost function. Subsequently, the Lagrangian dual formulation for the SDP in (4.6) can be expressed by:

$$
\mathrm{L_D}: \begin{cases}
\text{maximize} & \omega \\
\text{subject to} & \sum\limits_{i=1}^{n_f} \nu_i Q_i + \sum\limits_{j=1}^{n_g} \nu_j R_j + \\
& + \sum\limits_{j=1}^{n_d} \phi_j D_j + \omega e_1 e_1^T + e_1 \lambda_1^T A + \\
& + A^T \lambda_1 e_1^T + A^T \lambda_2 A + \lambda_3 = 0, \\
& \lambda_1 \geq 0, \lambda_2 \geq 0, \lambda_3 \succeq 0.
\end{cases} \tag{4.7}
$$

Thereby, $\lambda_i$, $\nu_j$, $\omega$ and $\phi_j$ are the Lagrangian multipliers corresponding to the equality constraints in the SDP, and $\lambda_1, \lambda_2 \in \mathbb{R}^{2n_\xi}$, $\lambda_3 \in \mathbb{S}^{n_\xi}$ those corresponding to the remaining (i. e. inequality) constraints. The Lagrangian weak-duality theorem guarantees that, if the Lagrangian dual $L_D$ is unbounded (i. e. $\omega > 0$ ), the SDP admits no solution. Then, as consequence of the relaxation process, also the FP admits no solutions. Thus, the Lagrangian weak-duality theorem provides an efficient infeasibility certificate for model (4.1) (taken from [24, 136]):

**Theorem 2** *(Unboundedness of the Lagrangian dual).*
*If the Lagrangian dual $L_\mathrm{D}$ is unbounded, then there exists no $p \in \mathcal{P}$ and $x(k) \in \mathcal{X}$ such that $y(k) \in \mathcal{M}(t_k)$ and $u(k) \in \mathcal{U}$, $k \in \mathcal{T}$.*

The unboundedness of the Lagrangian dual provides a certificate for model invalidation, which means that a model is guaranteed to be inconsistent with the measurement data. If the Lagrangian dual is not unbounded then the unknown solution space can be further explored to estimate the unknown constraint set (including sets for model parameters and state variables). One way is to estimate so-called outer approximations of the model variable sets. To do so, the constraint set is divided into partitions and a (possibly nonconvex) outer approximation of the solution space of the FP is derived by excluding those partitions for which the $L_D$ is unbounded. Detailed explanations of algorithms to do so are outlined in the following.

**Comment to linear relaxation:** To handle very large problems up to several thousands variables, linear programs are very useful. To obtain such a linear program, the SDP in (4.6) can be relaxed into a LP, in which the constraint $X \succeq 0$ is replaced by the weaker constraint $X \geq 0$. The introduced constraint for the linear relaxation still

leads to a valid constraint as the polynomial inequalities describing the compact sets, i. e. $BXe_1 \geq 0$ already guarantee that the variables are positive [93]. Notably, the substitution of the constraints lead to a further increase of additional (false positive) solutions.

## 4.4 Outer approximations of feasible solution sets

The above relaxation approach can be employed to tackle the set-based estimation problem allowing us to check infeasibility of a model and to approximate the feasible solution sets (including parameters and state variables). In the following, two different algorithms are presented to derive such approximations (cf. [24, 136, 152]).

**Outer-bounding algorithm**
The first algorithm performs an outer approximation of the solution sets by sequentially and iteratively tightening the lower and upper bounds on single parameters or state variables [136] (cf. Fig. 4.1 exemplary for parameters). In more detail, such outer approximations can be obtained, when the feasibility problem (4.2) is replaced by an optimization problem in which the single parameters or state variables are minimized or maximized. By formulating

$$
\begin{aligned}
&\text{infimum} && \xi_i \\
&\text{subject to} && f(x(k{+}1), x(k), u(k), p) = 0, \\
&&& g(y(k), x(k), u(k), p) = 0, \\
&&& x(k) \in \mathcal{X}, \ k \in \mathcal{T}, \\
&&& u(k) \in \mathcal{U}, \ k \in \mathcal{T}, \\
&&& y(k) \in \mathcal{M}, \ k \in \mathcal{T}, \\
&&& p \in \mathcal{P},
\end{aligned}
\tag{4.8}
$$

a tighter lower bound on variable $\xi_i$ can be determined. Thus, after applying relaxations as described above to problem (4.8), a boxed-shaped outer approximation, i. e. lower and upper bounds on parameters and state variables can be determined. These bounds, however, include no information about the shape of the feasible solution sets. Knowing the shape of feasible solution sets would, however, be very beneficial as these shapes can provide knowledge about the correlation of individual parameters or state variables. The following bisectioning algorithm provides means to do so.
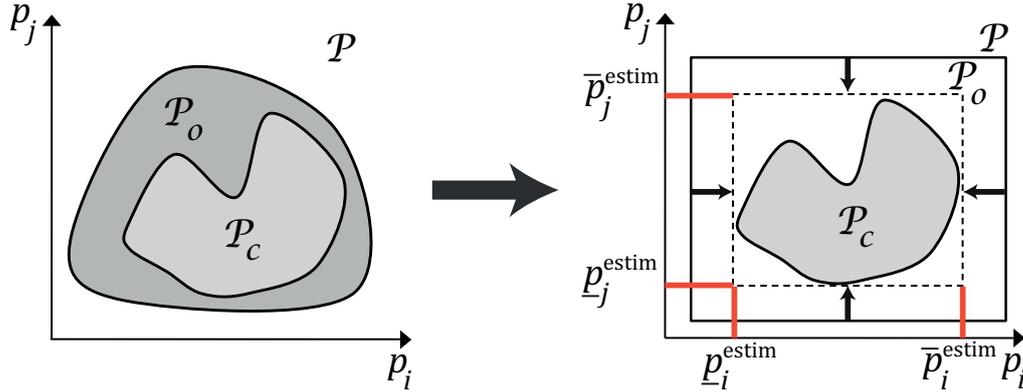
Figure 4.1: **Outer approximation of parameter sets using outer-bounding.**
Left: Relationship between the initial parameter space $\mathcal{P}$, the consistent parameter set $\mathcal{P}_c$ and an outer approximation $\mathcal{P}_o$. Right: Outer-bounding of parameters $p_j$ and $p_i$. The algorithm allows to estimate tight lower ($\underline{p}_j^{\text{estim}}$, $\underline{p}_i^{\text{estim}}$) and upper ($\overline{p}_j^{\text{estim}}$, $\overline{p}_i^{\text{estim}}$) bounds for the consistent parameter set $\mathcal{P}_c$ (dashed rectangle and red lines) by sequentially and iteratively tightening the initial parameter bounds (black arrows and black rectangle).

### Bisectioning algorithm

The bisectioning approach is a $2^n$-sectioning procedure, where $n$ denotes the total number of the variables of interest to be investigated. The initial solution space $\mathcal{P}$ is divided into partitions. In each iteration, the partition is checked for invalidity. If the partition is found to be invalid (i. e. $L_D$ is unbounded) it is not further considered. In contrast, if the partition is found to be not invalid (i. e. $L_D$ is not unbounded), it is divided into halves (cf. Fig. 4.2). The bisectioning algorithm can be implemented as a recursive procedure and ensures an outer approximation for the feasible regions with a desired precision.

Notably, the computational cost of the bisectioning algorithm depends exponentially on the desired level of precision, and on the number of variables to be investigated [136].
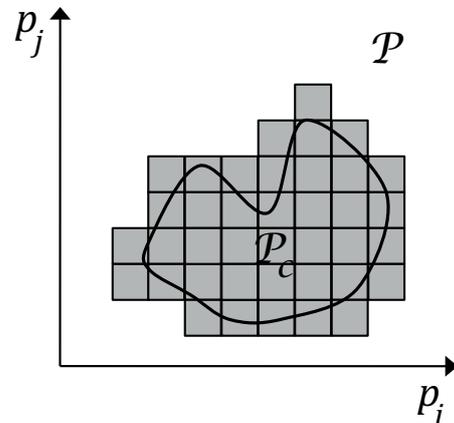


Figure 4.2: **Outer approximation of parameter sets using bisectioning.** Dark gray boxes describe a box-shaped outer approximation of the consistent parameter set $\mathcal{P}_c$.

In the following, we apply the presented approach to approximate feasible solution sets by bisectioning and demonstrate its usefulness for model invalidation and outer

approximation of feasible parameter sets. All model implementations and analyses within the framework of this thesis were performed using the Matlab-based toolbox ADMIT (*Analysis, Design, and Model Invalidation Toolbox*) [152]. In ADMIT, the formulation of feasibility problems, the involved relaxation steps, and the presented algorithms for deriving a solution of the problems by outer approximations are implemented. We note in addition that throughout this thesis all computations are conducted on a standard 2.4 GHz Intel desktop with 4 GB RAM.

## 4.5 Example: Stepwise receptor assembly

In this example, we consider IL-6-induced receptor assembly and activation during classic-signaling. Please refer to Chapter 2 for biological details on IL-6-induced receptor assembly and activation.

We develop and compare two model hypotheses, namely stepwise versus simplified receptor assembly and activation. The biological motivation as well as the models and results are published in [133]. In this study, we aimed to obtain a deeper understanding of IL-6-induced receptor assembly and activation as well as to identify possible model simplifications applying the set-based approach.

### 4.5.1 Model hypotheses and setup

To demonstrate the above set-based model invalidation and parameter estimation approach, we discuss in the following two competing model hypotheses describing IL-6-induced receptor assembly and activation. (cf. Fig. 4.3). We use model hypothesis 1 (i. e. $h^{[1]}$, described below) as a reference model. This reference model is compared with a more simplified version, i. e. hypothesis 2 ($h^{[2]}$) according their capability to represent the simulated measurement data within chosen parameter ranges.

**Model hypothesis 1:** The first model hypothesis depicts stepwise receptor assembly (cf. Fig. 4.3 upper panel). Thereby, IL-6 binds first to IL-6R$\alpha$ forming the complex IL-6:IL-6R$\alpha$. In the following, two molecules of gp130 are recruited to two complexes of IL-6:IL-6R$\alpha$ [28] (cf. Chapter 2). Receptor and ligand association leads to a formation of an active receptor complex $R_{complex}$. Formation of the active receptor complex induces activation of receptor-associated Jaks and subsequent phosphorylation of tyrosine residues in the cytoplasmatic domain of gp130 $\big($represented as (p)$R_{complex}$, light orange stars Fig. 4.3$\big)$.

Note, for model simplification we not explicitly considered Jaks, which were instead assumed to be represented as part of the gp130 species. This is an appropriate assumption as Jaks are constitutively associated with gp130 and the phosphorylation kinetics were shown to be identical [49].
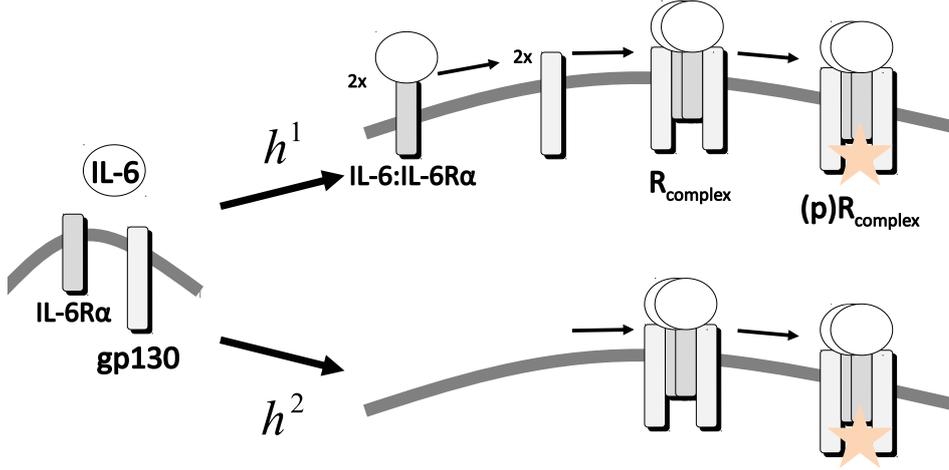
Figure 4.3: **Two competing model hypotheses for IL-6-induced receptor assembly in cells.** Model hypothesis 1 ($h^{[1]}$) describes a stepwise receptor assembly where in a first step IL-6 binds to receptor subunit IL-6R$\alpha$ and in a second step gp130 is recruited. Subsequently, formation of the active hexameric receptor complex ($R_{complex}$) leads to its phosphorylation (($p$)$R_{complex}$, light orange star). In model hypothesis 2 ($h^{[2]}$), the active hexameric receptor complex is immediately formed in the presence of IL-6, IL-6R$\alpha$ and gp130 resulting in receptor complex phosphorylation.

The reaction mechanisms for model hypothesis 1 can be described as follows:

$$h^{[1]} : \begin{cases} x_1(k+1) &= x_1(k) + \Delta t\Big(p_1 x_4(k)u - p_2 x_1(k) - 2p_3 x_2(k)^2 x_1(k)^2 + 2p_4 x_5(k)\Big) \\ x_2(k+1) &= x_2(k) + \Delta t\Big(2p_3 x_2(k)^2 x_1(k)^2 - 2p_4 x_5(k) - p_5 x_5(k) + p_6 x_3(k)\Big) \\ x_3(k+1) &= x_3(k) + \Delta t\Big(p_5 x_5(k) - p_6 x_3(k)\Big) \\ y(k) &= x_3(k), \end{cases}$$

$$(4.9)$$

where $\Delta t$ results from time discretization and $p_j \in \{1,2,3,4,5,6\}$ are the rate constants. Furthermore, the variables $x_1(k)$, $x_2(k)$, $x_3(k)$, and $u$ denote IL-6:IL-6R$\alpha$, $R_{complex}$, ($p$)$R_{complex}$ and IL-6, respectively. Moreover, $x_4(k)$ and $x_5(k)$ denote IL-6R$\alpha$ and gp130 which can be calculated from the following conservation laws:

$$\text{IL}-6\text{R}\alpha^{\text{Total}} = \text{IL}-6\text{R}\alpha + \text{IL}-6\text{:IL}-6\text{R}\alpha + 2\text{R}_{\text{complex}} + 2(\text{p})\text{R}_{\text{complex}} \qquad (4.10\text{a})$$

$$\text{gp130}^{\text{Total}} = \text{gp130} + 2\text{R}_{\text{complex}} + 2(\text{p})\text{R}_{\text{complex}}. \qquad (4.10\text{b})$$

To show applicability of the set-based approach, we simulated the model describing $h^{[1]}$ with a time step of $\Delta t$=1 min. All parameters were set to the nominal value 0.3. Further, a constant ligand concentration of IL-6=1 , total concentrations of IL-6R$\alpha$ and gp130, IL-6R$\alpha^{\text{Total}}$=1 and gp130$^{\text{Total}}$=1, and initial conditions of [IL-6:IL-6R$\alpha$(0),$R_{complex}$(0),($p$)$R_{complex}$(0)]=[0,0,0]. were assumed. We also assumed the phosphorylated receptor ($p$)$R_{complex}$ as observable model output, i. e. $y(k)$ and simulated measurement data with a relative error of 10% as given by the black dots

and bars in Fig. 4.4(b).

**Model hypothesis 2:** The second model hypothesis describes a simplification of IL-6-induced receptor assembly (cf. Fig. 4.3 lower panel). Typically, stepwise assembly of the active receptor complex as described above is not measurable experimentally in vitro. Hence, it is appropriate to assume that the hexameric receptor complex $R_{\text{complex}}$ is immediately formed in the presence of IL-6, IL-6R$\alpha$ and gp130. Thus, we aggregate binding of IL-6 to IL-6R$\alpha$ and binding of 2 molecules of IL-6:IL-6R$\alpha$ to two molecules gp130 into one step.

The reaction mechanisms for model hypothesis 2 can then be described as follows:

$$h^{[2]} : \begin{cases} x_1(k{+}1) & = x_1(k) + \Delta t\big(p_1 x_3(k)^2 x_4(k)^2 u^2 - (p_2 + p_5)x_1(k) + p_6 x_2(k)\big) \\ x_2(k{+}1) & = x_2(k) + \Delta t\big(p_5 x_1(k) - p_6 x_2(k)\big) \\ y(k) & = x_2(k), \end{cases}$$

(4.11)

where $x_1(k)$ and $x_2(k)$ denote $R_{\text{complex}}$ and $(\text{p})R_{\text{complex}}$, respectively. Additionally, the variables $x_3(k)$ and $x_4(k)$ describe IL-6R$\alpha$ and gp130 and can be extracted from the conservation laws:

$$\text{IL}-6R\alpha^{\text{Total}} = \text{IL}-6R\alpha + 2R_{\text{complex}} + 2(\text{p})R_{\text{complex}} \tag{4.12a}$$

$$\text{gp130}^{\text{Total}} = \text{gp130} + 2R_{\text{complex}} + 2(\text{p})R_{\text{complex}} \tag{4.12b}$$

For model simulations, the input concentration for IL-6, simulated measurement data for the phosphorylated receptor $(\text{p})R_{\text{complex}}$, initial conditions and time step $\Delta t$ were chosen according to hypothesis 1.

### 4.5.2 Set-based hypothesis invalidation

For the set-based analyses we assume all model parameters within the uncertainty interval $[0.03,3]$, which is equivalent to assumed half-life times of biomolecular reactions within seconds to minutes (according to the equation for the half-life time $t_{\frac{1}{2}}$ with $t_{\frac{1}{2},i} = \frac{ln\,2}{p_i}$, $i = \{1,2,3,4,5,6\}$ [41]).

For $h^{[2]}$ no parametrizations within the chosen uncertainty sets were determined for which the model was capable to explain the simulated measurement data. Therefore, $h^{[2]}$ is deemed as an invalid hypothesis. We then checked (in)validity of $h^{[1]}$. To this end, we performed in a first step an outer approximation for $p_4$, $p_5$ and $p_6$ applying the bisectioning algorithm (blue boxes in Fig. 4.4(a)). Notably, the parameter set was found to be non-empty and thus, $h^{[1]}$ is possibly a valid hypothesis.

As we only considered simulated measurement data for the dynamics of $(\text{p})R_{\text{complex}}$ (Fig. 4.4(b), black bars), the outer approximation of the parameters did not yield very tight parameter ranges. However, parameter regions could be determined that

are guaranteed invalid (light gray boxes in Fig. 4.4(a)). Furthermore, we determined parametrizations (red crosses in Fig. 4.4(a)) using Monte Carlo sampling for which the boundaries of measurement data were fulfilled (red bold trajectories Fig. 4.4(b)). Note, not all parametrizations derived from the outer approximation led to consistent solutions. Still invalid samples were obtained that led to inconsistent model-based predictions (cf. Fig. 4.4(a), blue crosses and Fig. 4.4(b), blue trajectories). This is due to the relaxation processes that lead to false positive parameters as described before.
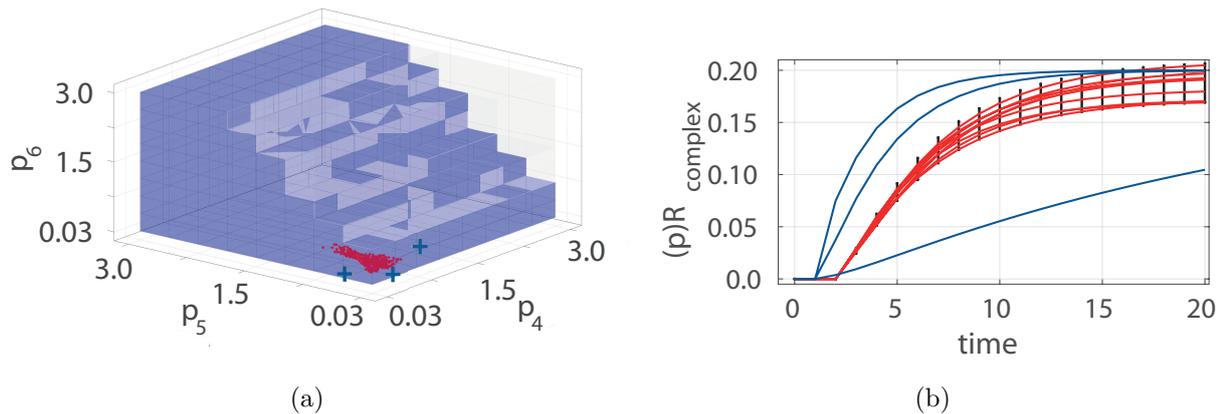


(a)                                              (b)

Figure 4.4: **Results for model invalidation and outer approximations of parameters for model hypothesis** $h^{[1]}$. (a) The parameters $p_4$, $p_5$ and $p_6$ were approximated using the bisectioning algorithm. While blue boxes denote the outer approximation, light gray boxes are guaranteed invalid regions. The cloud of red crosses describes valid parametrizations obtained by Monte Carlo sampling and blue crosses denote invalid samples leading to wrong predictions. (b) Valid (red trajectories) vs. invalid (blue trajectories) predictions based on the determined parametrizations in (a). Black bars describe the simulated measurement data assumed to have 10% relative error.

This example demonstrated the possibility to discriminate between different models for IL-6-induced receptor assembly and activation using the set-based approach. We wondered whether a simplified model for IL-6-induced receptor assembly and activation is also capable to explain simulated measurement data which were obtained by simulating the model for stepwise receptor assembly and activation with nominal values for parameters and initial conditions. As a result, model hypothesis 2 seems to be too simple for describing IL-6-induced receptor assembly and activation in an appropriate manner. Randomly determined parameter samples for model hypothesis 2 (not depicted) showed that the phosphorylation of the receptor complex is much quicker than receptor phosphorylation in model hypothesis 1.

Notably, one drawback of the presented set-based approach is that within the estimated outer approximation of the model parameters describing stepwise receptor assembly and activation also inconsistent parametrizations were obtained. Since it

cannot be guaranteed that every parametrization from the outer approximation is a valid solution, sampling routines, such as Monte Carlo sampling have to be implemented. One way to circumvent this ambiguity caused by false positive parameters resulting from relaxations are so-called inner approximations. Inner approximations obtain only valid parametrizations, thus, leading always to consistent model-based predictions. A set-based approach for deriving inner approximations of parameter sets are explained and exemplified in the following chapter.

## 4.6 Summary

In this chapter, we introduced the set-based estimation approach for discrete-time systems. The approach allows the inclusion of uncertainties in state variables, parameters and measurement data, and is based on formulating the task for checking model invalidity in terms of a nonlinear and nonconvex feasibility problem. As the feasibility problem cannot be solved directly, it is relaxed into a semidefinite or linear program. This relaxation of the feasibility problem into a semidefinite (and linear) program is conservative, nevertheless it can be used to draw guaranteed statements about model invalidity of the considered problem. For model hypotheses that could not be invalidated, outer approximations of model parameters were obtained including all consistent solutions.

We demonstrated applicability of the set-based methods by implementing model hypotheses for IL-6-induced receptor assembly. We were able to invalidate the hypothesis for a model decribing simplified receptor assembly. Furthermore, for the hypothesis describing a stepwise receptor assembly determine, outer bounds for the model parameters were estimated. We determined valid parametrizations by applying Monte Carlo sampling routines. For these parameter samples, the model was consistent with the simulated data. However, we also noted that false positive parametrizations were obtained that lead to inconsistent model predictions. This is due to the fact that during problem relaxations the size of the feasible solution space increases. Thus, also false positive parametrizations are obtained.

As a remedy, inner approximations of parameter sets can be derived. Inner approximations yield always consistent solutions and therefore, guarantee accurate and robust model predictions. For this reason, we extend in the following the introduced set-based approach for deriving inner approximations.

# 5 Approximation of Inner Parameter Sets for Reliable Model Predictions

In the previous chapter, we reviewed how infeasibility certificates can be used to derive outer approximations of parameter sets. As it cannot be guaranteed that each parametrization from these sets lead to reliable predictions, we derive in this chapter inner approximations of parameter sets that include only consistent parametrizations.

This chapter is structured as follows. First, we show how to reformulate quantitative measurement data using binary variables and logical operators. Second, we present the derivation of inner approximations of parameter sets by an inversion of the reformulated measurement constraints. Finally, two algorithms for estimating inner approximations are described. Before concluding the chapter, we present two examples.

The chapter is based on our work presented in [156].

## 5.1 Introduction

The presented set-based framework allows us to derive outer approximations of parameters considering uncertainties. Outer approximations can be very useful to determine sets of possible feasible parameters. However, as the relaxations (described in Chapter 4) lead to an increase in the problem size, they also contain inconsistent parametrizations. While the inconsistency might be less relevant when building models, it becomes relevant for model-based predictions and intervention strategies. As a remedy, we propose inner approximations of parameter sets which only contain such parameters that lead to solutions consistent with experiments (denoted as consistent solutions or consistent parameters in further). The goal for estimating inner approximation is to determine those parameter sets that always lead to consistent solutions and thus, reliable and robust predictions.

The estimation of inner approximations has been considered before. In [16], inner and outer approximations of convex polytopes in any finite dimension were computed. The authors determined two collections, i. e. an inner and an outer approximation for a given polytope $\mathcal{P}$ that can be described by a system of linear inequalities. The idea is to determine the collections of inner and outer approximations with non-overlapping interiors, such that all boxes in the inner approximation are contained in the polytope $\mathcal{P}$ (i. e. an inner approximation) and to approximate the union of all boxes in the outer approximation containing the polytope $\mathcal{P}$ (i. e. an outer approximation).

Thereby, the authors consider two objectives, i. e. (i) to minimize the box volume error, and (ii) to minmize the total number of generated boxes and aim to obtain the best trade-off between both. To this end, a collection of algorithms is introduced and the computational costs of the algorithms are compared against each other. Notably, the performance of the algorithms rapidly decreases as the dimension of the problem grows.

In [79], inner approximations in the context of bounded-error estimation were derived for nonlinear models. The problem of nonlinear bounded-error estimation is considered as one of *set inversion.* Set inversion describes the problem of characterizing sets such as $\mathcal{S}=\{p|f(p) \in \mathcal{Y}\}$, or alternatively $\mathcal{S}=f^{-1}(\mathcal{Y})$. To this end, in [79] the algorithm SIVIA (Set Inversion Via Interval Analysis) was presented, which makes it possible to approximate the feasible solution set $\mathcal{S}$ by enclosing it between inner and outer sets of boxes. The algorithm has been applied, for instance, for guaranteed parameter estimation [79], and robust stability analysis [168].

Henrion and co-workers developed methods for outer and inner approximations of consistent parameter sets and their regions of attraction of continuous-time systems [88, 155]. They thereby introduced so-called occupation measures which are used to cast the (nonlinear) dynamics of a equation system into a system with linear (in-)equalities in an optimization problem over an infinite function space. The problem is then solved by a hierarchy of semidefinite or linear programs. In particular, [155] exploited the approach of occupation measures to determine inner and outer approximations of the set containing all consistent initial conditions/parametrizations for nonlinear (polynomial) continuous-time systems. The authors incorporated unknown-but-bounded and pointwise-in-time state and output constraints into their problem formulation. Thereby, the problem was solved by a hierarchy of LMI (linear matrix inequality) relaxations that provide certificates in case no consistent initial condition/parametrization exists.

In this chapter, we present an extension of the set-based framework fom Rumschinski et al. [136] for deriving inner approximations of consistent parameter sets for nonlinear discrete-time systems. These set-based methods are particularly suited for proving infeasibility of models with guarantees. To demonstrate feasibility of a model, so far further approaches are needed, such as Monte Carlo sampling, as outer approximations might contain invalid (i. e. false positive) parametrizations due to the relaxations presented. Therefore, we propose an approach that allows us inverting the problem for deriving outer approximation, to determine solutions of the inner approximations. To this end, quantitative measurement data are reformulated using binary variables and combinations thereof. By an inversion of the measurement constraints and by proving infeasibility for the inverted problem we can guarantee that no constraints are violated which provides the inner approximation.

## 5.2 Mixed-integer reformulation of quantitative constraints

To derive inner approximations of parameter sets, the constraints representing uncertain quantitative measurement data need to be reformulated; cf. Chapter 4, Section 4.2, where we have assumed a collection of quantitative data at time instances $n_t$ to be given, i. e. $y(k) \in \mathcal{M}(t_k)$ for $k \in \mathcal{T} := \{0, 1, ..., n_t - 1\}$.

Typically, measurements of state variables are uncertain and are assumed to lie within uncertainty intervals. In biological experiments, such uncertainty intervals can be derived by computing, e. g. the mean value $\pm$ standard deviation, which gives the uncertainty sets $y_i(k) := [\underline{y}_i(k), \overline{y}_i(k)]$ for $i = 1, 2, ..., n_y$ (Chapter 3, Section 3.2.1).

Let in the following $m_y = [y_1(0), ..., y_1(n_t - 1), ..., y_{n_y}(0), ..., y_{n_y}(n_t - 1)] \in \mathbb{R}^{n_y n_t}$ be a vector containing the measurements $y$ at time instances $n_t$. With this notation, we can reformulate the initial measurement set $\mathcal{M}(t_k)$ in the following form

$$\mathcal{M} := \left\{ m_y : q_{y,i}(m_y) \geq 0, i = 1, 2, \ldots, c_m \right\} \subset \mathcal{Y}. \tag{5.1}$$

In Eq. (5.1), $q_{y_i}(m_y)$ are polynomial inequality constraints that correspond to the upper and lower bounds for $y_i(k)$, i. e. $\overline{y}_i(k)$ and $\underline{y}_i(k)$, respectively. Furthermore, $c_m$ denotes the number of inequalities.

The key idea is to express the inequality constraints (5.1) using logical relationships by additionally introduced binary variables $\phi_i \in \{0, 1\}, i = 1, 2, \ldots, n_y$ as reviewed in the previous chapter, Section 4.2.1. Briefly, the variable $\phi_i$ indicates whether the constraint $q_{y,i}(m_y)$ is satisfied ($\phi_i = 1$) or not ($\phi_i = 0$). To be able to add this to the FP, we introduce additional constraints such that $\phi_i$ takes the desired value. This can be achieved by replacing $\alpha^T y(k) - c$ in Eq. (4.4) with $q_{y,i}(m_y)$, yielding:

$$\begin{aligned}
\phi_i &\geq \frac{q_{y,i}(m_y)}{M_i}, \\
\phi_i &\leq \frac{q_{y,i}(m_y)}{M_i} + 1, \quad i = 1, 2, \ldots, c_m,
\end{aligned} \tag{5.2}$$

where $M_i > |\max_{\forall y \in \mathcal{M}} q_{y,i}(m_y)|$.

As we consider multiple sets $\mathcal{M}$ (i. e. at different time points $t_k$) described by $c_m$ inequalities, all binary variables $\phi_i$ need to be equal to 1, implying that all of the above equations for all variables $\phi_i$ are fulfilled. In other words, the model is capable to represent the measurement data if $\phi_i = 1, \forall i$. The above requirement that all constraints $\phi_i$ are fulfilled can then be expressed by

$$\sum_{i=1}^{c_m} \phi_i = c_m. \tag{5.3}$$

The derived constraint corresponds to a conjunction, which is a logical *and-*combination of the binary variables. As can be seen, if any $\phi_i$ is 0 (i. e. a measurement

constraint is not fulfilled), (5.3) is not satisfied.

To derive a solution for a problem including the mixed-integer reformulation from above, we first have to pose the corresponding feasibility problem.

## 5.3 Mixed-integer feasibility problem formulation and relaxation

The constraints (5.2) and (5.3) can be added to the FP (Eq. (4.2)), which is then denoted as a mixed-integer FP (MIFP) due to the inclusion of above integrality constraints.

Assuming that our goal is to derive inner approximations for the model parameter $p$, then the MIFP has the following form:

$$
\text{MIFP} : \begin{cases}
\text{find} & p \\
\text{s.t.} & f(x(k{+}1), x(k), u(k), p) = 0, \\
& g(y(k), x(k), u(k), p) = 0, \\
& x(k) \in \mathcal{X}, \forall k \in \mathcal{T}, \\
& u(k) \in \mathcal{U}, \forall k \in \mathcal{T}, \\
& p \in \mathcal{P}, \\
& \phi_i \geq \frac{q_{y,i}(m_y)}{M_i}, \\
& \phi_i \leq \frac{q_{y,i}(m_y)}{M_i} + 1, \\
& \sum_{i=1}^{c_m} \phi_i = c_m.
\end{cases}
\tag{5.4}
$$

With (5.4), one can state:

**Lemma 1** *(Equivalence of FP and MIFP).*
*The solution sets of FP and MIFP are equal, i. e. $\mathcal{P}_{FP} = \mathcal{P}_{MIFP}$.*

*Proof:* The proof is obvious and follows from the conversion of the FP into a MIFP.

In [24], a theorem was presented which gives inconsistency certifiates for the FP (see also Chapter 4). Using Lemma 1 and the theorem from [24] we can now state (without the need of a formal proof):

**Theorem 3** *(MIFP inconsistency certificate).*
*If the MIFP does not admit a solution, then there exists no $p \in \mathcal{P}$ and $x(k) \in \mathcal{X}$ such that $y(k) \in \mathcal{M}(t_k)$ and $u(k) \in \mathcal{U}, \ \forall k \in \mathcal{T}$.*

To solve the derived MIFP it can be, similarly to the FP, relaxed into a mixed-integer semidefinite problem (MISDP), which is a convexified form of the MIFP. To efficiently solve this MISDP it is further relaxed into a mixed-integer linear problem (MILP). A solution for the MILP can then be derived using appropriate solvers, such as Cplex.

As during the relaxation process the solution space of the MILP increases, it is likely to happen that, akin to outer approximations, also inconsistent (i. e. false positive) parametrizations are obtained. Consequently, constraint (5.3) cannot be proven with guarantees.

To avoid this, we present in the following a solution approach based on a constraint inversion of (5.3), which allows to derive guaranteed inner approximations of parameter sets.

## 5.4 Constraint inversion

Due to the convex relaxations of the MIFP into the MILP, it may occur that inconsistent solutions also for the inner approximations are obtained. Mathematically, there may exist parametrizations such that

$$\sum_{i=1}^{c_m} \phi_i \leq c_m - 1 \tag{5.5}$$

holds. In more detail, if (5.5) is fulfilled, then $y(k) \notin \mathcal{M}(t_k)$ holds for at least one $k$, and thus, the model is not capable to explain the measurement data.

Using the presented set-based approach, we can show infeasibility of a model with guarantees. However, due to the relaxation process, the same does not hold for proving feasibility. Therefore, the crucial idea for deriving inner approximations is to formulate the inversion of (5.3), which is expressed by (5.5). As a consequence it follows that if we can prove infeasibility of (5.5), then feasibility of (5.3) is confirmed. Feasibility of (5.3) refers to the model's capability to explain the measurement data at all time instances and, hence, an inner approximation $\mathcal{P}_i$ of the initial parameter set $\mathcal{P}$ is obtained.

In the following, the inverted constraint (5.5) is added to the MIFP instead of (5.3) which leads to:

$$\widehat{\mathrm{MIFP}} : \begin{cases} \text{find} & p \\ \text{s.t.} & f(x(k{+}1), x(k), u(k), p) = 0, \\ & g(y(k), x(k), u(k), p) = 0, \\ & x(k) \in \mathcal{X}, \forall k \in \mathcal{T}, \\ & u(k) \in \mathcal{U}, \forall k \in \mathcal{T}, \\ & p \in \mathcal{P}, \\ & \phi_i \geq \frac{q_{y,i}(m_y)}{M_i}, \\ & \phi_i \leq \frac{q_{y,i}(m_y)}{M_i} + 1, \\ & \sum_{i=1}^{c_m} \phi_i \leq c_m - 1. \end{cases} \tag{5.6}$$

Consequently, if the $\widehat{\text{MIFP}}$ in (5.6) admits no solution, we can guarantee that parametrizations $p \in \mathcal{P}$ exist, such that Eq. (5.3) holds and thus, $y(k) \in \mathcal{M}(t_k)$, $\forall k$.

To state the corresponding theorem, two assumptions have to be made:

**Assumption 1** *(Existence of solutions).*
*For every* $p \in \mathcal{P}$ *there exists a solution* $x(0), x(1), \ldots, x(n_t - 1)$ *and* $y(0), y(1), \ldots, y(n_t - 1)$.

Many systems fulfill this assumptions. In particular polynomial systems given in the explicit form $x(k+1) = \tilde{F}(x(k), p)$.

In addition, we have:

**Assumption 2** *(Bounds on solutions).*
*For all* $p \in \mathcal{P}$, *it holds that* $x_0 \in \mathcal{X}_0$ *and* $x(k) \in \mathcal{X}, \ \forall k \in \mathcal{T}$.

This assumption requires known guaranteed enclosures on the initial conditions and states on the time interval $\mathcal{T}$. This assumption is appropriate, as large uncertainty bounds can be chosen for $\mathcal{X}_0$ and $\mathcal{X}$. In addition, these bounds can be derived from system insight, such as from known relations on mass conservation as described in Chapter 3.

The main purpose of Assumption 2 is to reduce the number of constraints that have to be tested for validity. Then, one only needs to check and invert constraints for measurement data $q_{y,i}(m_y)$ and does not have to consider constraints for the initial conditions and model states. This implies the reduction of computational costs.

With Assumption 1 and 2 we can state the following theorem:

**Theorem 4** *(Consistency certificate)*
*If the Lagrangian-dual of the* $\widehat{\text{MILP}}$ *is unbounded, then it is guaranteed that there exists* $p \in \mathcal{P}$, *such that* $y(k) \in \mathcal{M}(t_k) \ \forall k \in \mathcal{T}$.

*Proof:* The weak-duality theorem and the relaxation process guarantee that if the dual $\widehat{\text{MILP}}$ is unbounded, then the $\widehat{\text{MIFP}}$ does not admit a solution. Due to Assumptions 1 and 2 it follows that (5.5) is not satisfied, hence (5.3) is satisfied for all $p \in \mathcal{P}$.

**Note on the increase of the problem size:** The reformulation of the original problem FP into a MIFP results in an increased problem size in terms of the number of variables and constraints. To be more specific, there are $c_m$ additional binary variables $\phi_i$, $2c_m$ additional nonlinear constraints (5.2), and one additional constraint representing the combination of all binary variables (5.5). However, an increase of the problem size does not restrict the numerical solvability of our analyses due to the existence of efficient and parallel implementations of branch-and-bound algorithms in mixed-integer solvers, such as Cplex.

By reformulating constraints for measurement data and their inversion, we are able to estimate inner approximations of parameters in the set-based context. In the following two algorithms for implementing Theorem 4 and for estimating inner approximations of parameter sets are presented.

## 5.5 Algorithms for determining inner parameter sets

This section outlines two algorithms to determine inner approximations of parameter sets $\mathcal{P}_i$ based on Theorem 4. Since the theorem builds on infeasibility, i. e. excluding infeasible regions, the method proceeds that entire parameter regions are either recursively (Algorithm 1) or incrementally (Algorithm 2) proven to be an inner approximation. Both algorithms terminate in finite time. A graphical illustration of the relationship between inner and outer approximations and the employed algorithms are given in Fig. 5.1 and 5.2.

### 5.5.1 Recursive inner approximation

The following algorithm assumes that the initial parameter set $\mathcal{P}$ is given as denoted in Eq. (3.17), which corresponds to a hyperrectangular description of the parameter space. An inner approximation $\mathcal{P}_i$ (cf. Fig. 5.1(a)) can then be determined by systematically exploring subregions of the initial parameter space $\mathcal{P}$. Algorithm 1 uses a bisectioning procedure and tests whether a defined hyperrectangle is an inner approximation or not. If the hyperrectangle is found to be no inner approximation, it is split into two parts along a heuristically or randomly chosen direction and the two obtained hyperrectangles are tested again. The recursion terminates either if a hyperrectangle is proved to be an inner approximation, or if a predefined recursion depth $N_{max}$ is reached. The algorithm is initially called with the parameter index set $\mathcal{I} := \{1, ..., n_p\}$, a set of $n_p$ counters $N_j$ which are all set to 1, and the initial uncertain parameter description as given in (3.17). The recursion depth $N_{max}$ and the dimension of $\mathcal{P}$ determine the number of hyperrectangles to be tested for invalidity and thereby define the accuracy of the inner approximation. Notably, the overall cost for Algorithm 1 grows exponentially with the number of hyperrectangles.

The algorithm is suited to explore the parameter space for inner approximations. However, it is computationally more demanding to prove a region to be an inner approximation than to check the same region whether it contains no solutions at all, by using cetrificates for outer approximations. It is therefore beneficial to combine both outer and inner approximations by, (i) using certficates for outer approximations up to a certain recursion depth and, (ii) checking the remaining boxes whether they are valid inner approximations or not (cf. Fig. 5.1(b)). Note, if an inner approximation $\mathcal{P}_i$ could be determined using Algorithm 1, a sampling-based routine (e. g. Monte Carlo sampling) can be used to determine guaranteed consistent parametrizations.
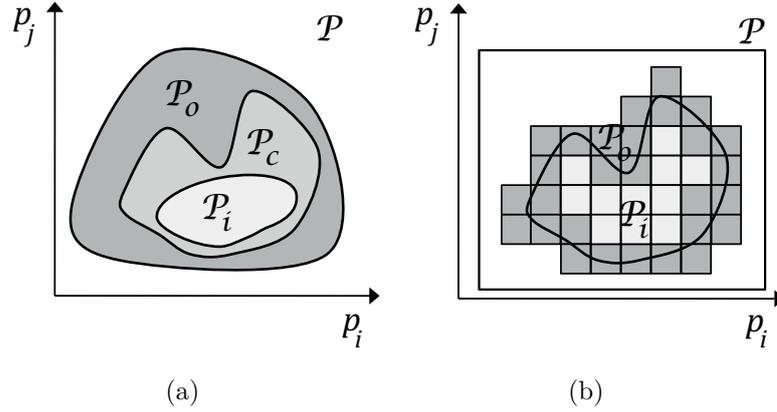
(a)                     (b)

Figure 5.1: **Inner approximation of parameter sets.** (a) Relationship between the initial parameter set $\mathcal{P}$, the consistent parameter set $\mathcal{P}_c$, an outer approximation $\mathcal{P}_o$ and an inner approximation $\mathcal{P}_i$. (b) Determination of an inner approximation for parameters $p_i$ and $p_j$ using the bisectioning Algorithm 1. While the dashed box describes a hyperrectangular outer approximation, the dark and light gray boxes show an outer and inner approximation of the initial parameter set $\mathcal{P}$, respectively.

---

**Algorithm 1**. (Recursive Inner Approximation).

---

```
Input:   index set I
         counters Nⱼ ∀ j ∈ I
         hyperrectangular bounds P
```

---

```
IF Lagrangian-dual of M̂ILP is unbounded
   ECHO 'P is an inner approximation'
   RETURN
END IF
WHILE I ≠ ∅
   PICK index j corresponding to
     largest interval width p̄ⱼ − pⱼ in P
   IF Nⱼ ≤ Nₘₐₓ:
     SET  Nⱼ ← Nⱼ + 1
     STORE and SET pⱼ ← pⱼ + ½ · (p̄ⱼ − pⱼ)
     CALL Algorithm 1 with Nⱼs, I and modified P
     RESET pⱼ
     STORE and SET p̄ⱼ ← p̄ⱼ − ½ · (p̄ⱼ − pⱼ)
     CALL Algorithm 1 with Nⱼs, I and modified P
     RESET p̄ⱼ
   ELSE
     SET I ← I \ j
     ECHO 'P is not considered an inner approximation'
   END IF
END WHILE
```

An alternative to Algorithm 1 is presented next.

## 5.5.2 Incremental polytopic expansion

Often, parameter samples leading to consistent model-based predictions are available in advance, e. g. by applying Monte Carlo sampling routines or by local nonlinear optimization (red cross in Fig. 5.2). The following algorithm uses such parameter samples to determine a polytopic inner approximation. The basic idea is to test whether a parameter set, which is defined by the polytoic convex hull of the samples, describes an inner approximation or not. If it does, the algorithm incrementally moves the facets of the polytope outwards until an inner approximation cannot be guaranteed anymore. In other words, Algorithm 2 increases the hypervolume of the polytope step by step (cf. Fig. 5.2 gray arrows).



Figure 5.2: **Algorithm for incremental polytopic sample expansion.** Determination of an inner approximation for parameters $p_i$ and $p_j$ using Algorithm 2.

We assume in the following that a polytopic convex hull of a series of parameter samples is given:

$$\mathcal{P} := \left\{ p : a_j^T p \geq b_j, \|a_j\|_2 = 1, \quad j = 1, 2, \ldots, c_f \right\}, \tag{5.7}$$

where $a_j$ are (outwards-facing) unit normal vectors of the $c_f$ facets of the polytope. Additionally, assume that $\mathcal{P}$ is bounded and represents a polytope of full dimension in $\mathbb{R}^{n_p}$.

Note that Algorithm 2 terminates after a finite number of steps if the consistent parameter set is bounded. Due to the random choice of the next facets and the step size $\nu$, the obtained polytopic description of the inner approximation is not unique and might not be the optimum in terms of maximal volume. However, the advantage of the algorithm is that, once an inner approximation has been found (based on a consistent sample), then no further sampling is required as it has been proved that the polytope describes an inner approximation using Algorithm 2.
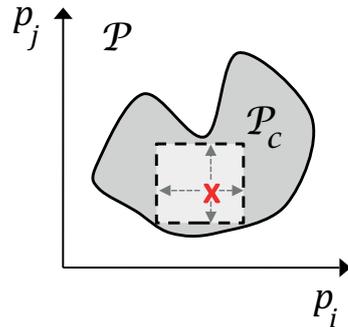
---

**Algorithm 2**. (Incremental Polytopic Expansion).

---

```
Input:   index set 𝓘 = 1, 2, …, c_f
         incremental step-size ν
         polytopic uncertainty description 𝓟
```

---

```
IF Lagrangian-dual of M̂ILP is NOT unbounded
   ECHO '𝓟 is not considered an inner approximation'
   RETURN
END IF
WHILE 𝓘 ≠ ∅
   PICK random j from 𝓘
   STORE AND REPLACE jth equation
     in 𝓟 by a_j^T p ≥ b_j + ν
   IF Lagrangian-dual of M̂ILP is NOT unbounded
     RESET jth equation
     SET 𝓘 = 𝓘 \ j
   END IF
END WHILE
```

---

## 5.6 Example: Michaelis-Menten kinetics

To show applicability of the presented framework and algorithms, we consider in the following two examples for determining inner approximations.

We consider an enzyme-catalyzed reaction network of the form

$$
\begin{aligned}
S_1 + E \underset{p_2}{\overset{p_1}{\rightleftharpoons}} C_1 \overset{p_3}{\rightarrow} P + E \\
S_2 + E \underset{p_5}{\overset{p_4}{\rightleftharpoons}} C_2.
\end{aligned}
\tag{5.8}
$$

This describes that an enzyme $E$ and a substrate $S_1$ reversibly form a complex $C_1$, in which the substrate is converted into the product $P$. The enzyme is bound by a second substrate $S_2$ forming the inhibitory complex $C_2$. The parameters $p_1, p_2, \ldots, p_5$ denote the unknown rate constants for which we aim to determine inner approximations.

**Model and setup:** The reaction mechanism (5.8) can be modeled as:

$$
\begin{aligned}
x_1(k{+}1) &= x_1(k) + \Delta t\big(p_1 x_4(k) x_5(k) - (p_2 + p_3) x_1(k)\big) \\
x_2(k{+}1) &= x_2(k) + \Delta t\big(p_4 x_5(k) x_6(k) - p_5 x_2(k)\big) \\
x_3(k{+}1) &= x_3(k) + \Delta t\big(p_3 x_1(k)\big)
\end{aligned}
\tag{5.9a}
$$

$$
y(k) = (x_1(k), x_2(k), x_3(k)),
\tag{5.9b}
$$

where $x_1(k)$, $x_2(k)$, $x_3(k)$, $x_4(k)$, $x_5(k)$ and $x_6(k)$ represent the concentrations $C_1$, $C_2$, $P$, $S_1$, $E$ and $S_2$, respectively, and $\Delta t$ is the discretization time, which is set to 0.1 h. We further assume that the following conservation relationships hold:

$$
\begin{aligned}
1 &= x_4(k) + x_1(k) + x_3(k) \\
1 &= x_5(k) + x_1(k) + x_2(k) \\
1 &= x_6(k) + x_2(k).
\end{aligned}
\tag{5.10a}
$$

To obtain outer and inner approximations of the model parameters, we used artificial measurement data for the model states, i. e. $x_1$, $x_2$ and $x_3$ (cf. Eq. (5.9b)). The artifical data were obtained simulating the system with the nominal initial condition $[x_1(0), x_2(0), x_3(0)]^{\mathrm{T}} = [0.05, 0.05, 0.05]^{\mathrm{T}}$. Parameter values $p_i$ were set to 2 for $i = 1, 2, 4, 5$ and to 5 for $i = 3$. To simulate data uncertainties, a random error of 5% was added to each simulated data point (black bars Fig. 5.3(d)). For set-based analyses, the initial parameter uncertainties were set to the interval $p_i{=}[0.1,10]$, i=1, ..., 5, which is equivalent to assume half-life times for the biomolecular reactions within seconds to minutes. For computational reasons, we express only the bounds for $x_3(k)$ using binary variables and inverted them as described before.

**Results:** Fig. 5.3 depicts the obtained results applying Algorithms 1 and 2. We first performed an outer approximation of the model parameters $p_1$-$p_5$ from the original problem (i. e. without mixed-integer formulations) using the proposed outer-bounding algorithm from [136] (cf. Fig. 5.3(a)). The obtained bounds were then used as inputs to the mixed-integer problem, reducing the parameter space that is tested for inner approximations. In Fig. 5.3(b) the results for Algorithm 1 are depicted. While blue boxes describe an outer approximation of the MIFP (cf. Eq. (5.4)), light red boxes describe the obtained inner approximation of parameters $p_1$ and $p_2$ for $\widehat{\mathrm{MIFP}}$ (cf. Eq. (5.6)) using Algorithm 1. To prove that the determined inner approximation is valid, we additionally performed a Monte Carlo sampling (green dots Fig. 5.3(b)). The results show a tight inner approximation for parameters $p_1$ and $p_2$ covering 82% of the valid Monte Carlo samples. Notably, to improve 82% coverage of the inner approximation, the recursion depth $N_{max}$ could be altered allowing for an enlargement of the inner approximation. The black rectangle shows the obtained polytopic inner approximation obtained with Algorithm 2 by expanding the boundaries of the consistent parametrization [2.06;2.26;4.9;1.9;1.8] (red cross, $\nu{=}0.02$) that was obtained by Monte Carlo sampling. The results for Algorithm 2 are additionally shown in Fig. 5.3(c) for the parameters $p_1$, $p_2$ and $p_3$. Finally, we depicted the obtained trajectories based on samples from the outer and the polytopic inner approximation. While in Fig. 5.3(d) not all samples from the outer approximation yield consistent model predictions (solid, light blue), samples from the obtained polytopic inner approximation (dashed, light red) are always consistent with the data (black bars).
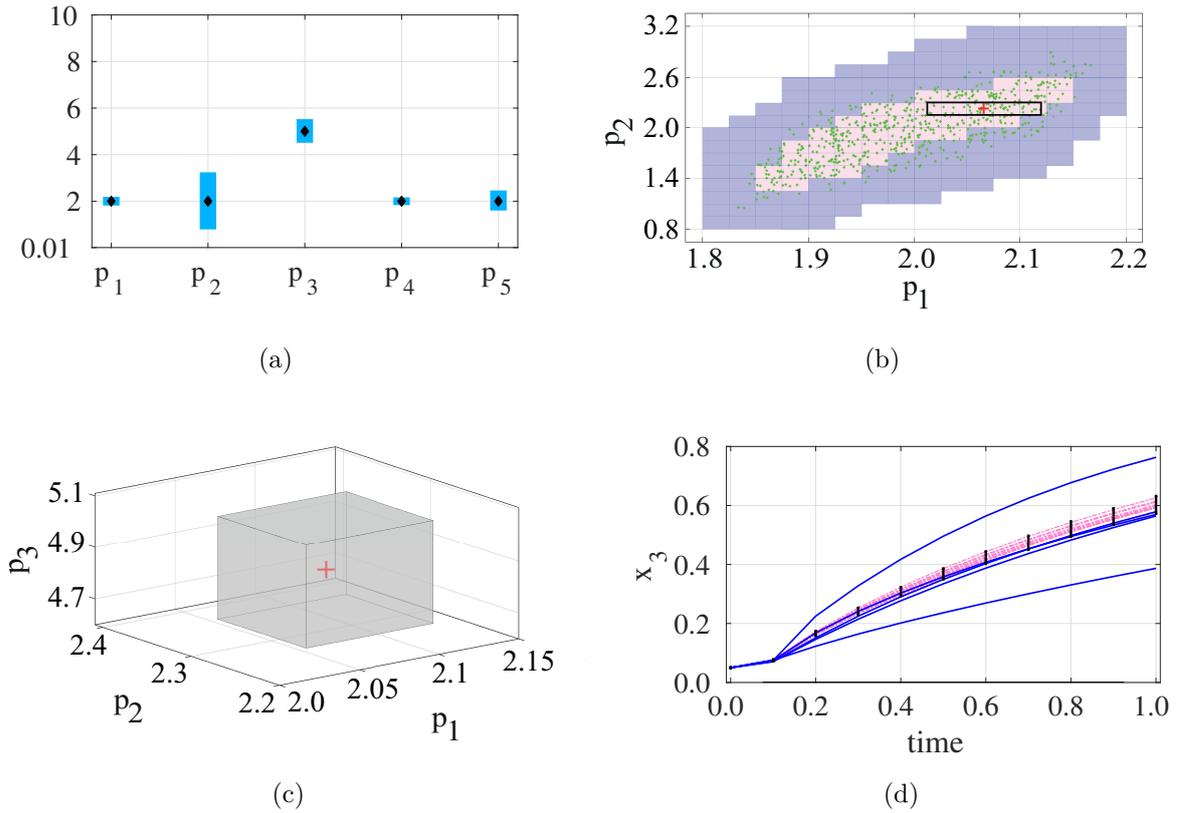
Figure 5.3: **Results for outer and inner approximations of parameter sets.**
(a) Estimated outer bounds (blue bars) for the model parameters using an
outer-bounding algorithm. The black dots show the nominal parameter sample
[2;2;5;2;2]. (b) Estimation of the consistent parameter set $\mathcal{P}_c$ using bisectioning
and Algorithm 1. An outer and an inner approximation are given by the blue
and light red boxes, respectively. Green dots correspond to valid Monte Carlo
samples. The black rectangle shows the obtained hyperrectangular inner approx-
imation using Algorithm 2 (see also (c)) projected onto ($p_1$-$p_2$)-space starting
with the valid sample [2.06;2.26;4.9;1.9;1.8] (red cross). (c) A polytopic expan-
sion using Algorithm 2 plotted in ($p_1$-$p_2$-$p_3$)-space starting with the valid sample
[2.06;2.26;4.9;1.9;1.8] (red cross) leads to the inner approximating hyperrectangle
(gray box) [2.02;2.10]×[2.2;2.3]×[4.7;5.1]×[1.7;2.2]×[1.5;2.3]. (d) Model simula-
tions using parameter samples from the outer (solid, blue) and inner approxima-
tions (dashed, light red). Black bars denote the simulated, uncertain measurement
data obtained with the nominal sample from (a).

## 5.7 Example: Interleukin-6-induced trans-signaling

In constrast to the simulated data in the previous example, real measurement data
are often sparse. Furthermore, usually not all model states are measurable due to,
for example, methodological restrictions. The following example demonstrates how

the proposed framework performs if not all model species are observable and if the measurement data are subject to large noise and different frequency.

**Model for early IL-6-induced trans-signaling:** In the following we focus on a model for IL-6-induced trans-signaling and activation of Jak/STAT3 signal transduction. To derive a mathematical model for the biological mechanisms described in Chapter 2, we make the following assumptions:

- Jak kinases are represented by gp130 species.

- STATs are phosphorylated by the active and phosphorylated receptor complex $(\text{p})\text{R}_{\text{complex}}$ (consisting of Jak and gp130).

- To mimic trans-signaling in experiments, Hyper-IL-6 was used as stimulus. As a consequence, binding of IL-6 to its soluble receptor sIL-6R$\alpha$ is not considered and, hence, the corresponding kinetic parameters $p_1$ and $p_2$ are set to 0.

Note, the third assumption allows us to reduce the model for IL-6-induced trans-signaling by one dynamic equation.

For complexity reduction, we consider only the first 15 minutes of IL-6-induced Jak/STAT3 signaling. Consequently, the transcription of SOCS3 mRNA and the translation of SOCS3 protein due to the activation of STAT3 is disregarded. A suitable model for the dynamic processes of IL-6-induced trans-signaling and activation of Jak/STAT3 signal transduction can then be described by:

$$
\begin{aligned}
x_1(k+1) &= x_1(k) + \Delta t\big(2p_4 x_4(k) - 2p_3 x_1(k)^2 u^2\big) \\
x_2(k+1) &= x_2(k) + \Delta t\big(p_5 x_4(k) - p_6 x_2(k)\big) \\
x_3(k+1) &= x_3(k) + \Delta t\big(p_7 x_5(k) x_2(k) - p_8 x_3(k)\big)
\end{aligned}
\tag{5.11a}
$$

$$
y(k) = x_3(k). \tag{5.11b}
$$

In Model (5.11), the variables $x_1(k)$, $x_2(k)$, $x_3(k)$, $x_4(k)$, and $x_5(k)$ describe the species gp130, $(\text{p})\text{R}_{\text{complex}}$, phosphorylated STAT3 (i. e. (p)STAT3), $\text{R}_{\text{complex}}$, and STAT3, respectively. Furthermore, $u$ is the constant model stimulus, i. e. Hyper-IL-6 and $x_3(k)$ is the model output (cf. (5.11b)).

Note that the variables $x_4(k)$ ($\text{R}_{\text{complex}}$) and $x_5(k)$ (STAT3) can be derived from the following conserved moeties:

$$
\text{gp130}^{\text{Total}} = \text{gp130} + 2\text{R}_{\text{complex}} + 2(\text{p})\text{R}_{\text{complex}} \tag{5.12a}
$$

$$
\text{STAT3}^{\text{Total}} = \text{STAT3} + (\text{p})\text{STAT3}. \tag{5.12b}
$$

**Measurement data:** To derive inner approximations for the unknown model parameters $p_i$, $i = \{3, ..., 8\}$, quantitative experiments in HepG2 cells were performed (cf. [130] and Appendix A). The absolute amount of STAT3 phosphorylation for a constant Hyper-IL-6 stimulation of 0.17 nM was measured over a time horizon of 15 minutes using methods presented in Appendix A. According to Fig. 5.4(a) the phosphorylation of STAT3 steadily increases up to 15 minutes. In addition to that, the total concentrations of gp130 and STAT3 amounted to 16.8±3.1 nM (16198±2965 molecules per cell, n=4 assuming a cell volume of $V_{cell}$=1.6×10$^{-12}$ L) and 958±445 nM (921000±428000 molecules per cell, n=7), respectively (cf. Fig. 5.4(b)).
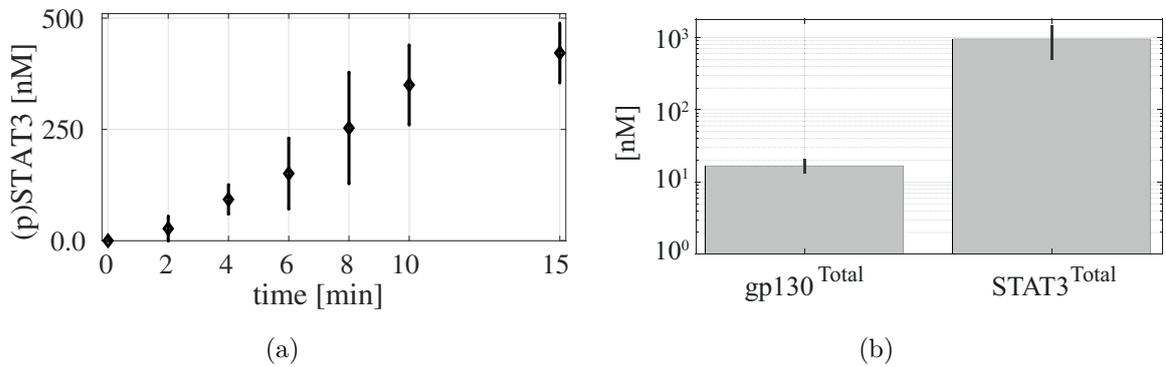


Figure 5.4: **Measurement data for determining inner approximations of parameter sets.** (a) Absolute quantification of STAT3 phosphorylation over a time horizon of 15 minutes. Diamonds correspond to the mean values and bars denote the determined standard deviations for n=6 independent replicates. (b) Absolute quantification of total amounts for gp130 and STAT3. Bars denote the determined standard deviations for n=4 and n=7 independent replicates, respectively.

**Model setup:** For set-based model simulations, the boundaries for gp130$^{\text{Total}}$ and STAT3$^{\text{Total}}$ are set to the uncertainty intervals of [13.7,19.9] and [513,1403], respectively. Furthermore, ranges for the model parameters $p_i$, $i = \{3, ..., 8\}$ were set to [10$^{-3}$,10$^1$] equivalent to assumed half-times for biomolecular reactions within seconds to minutes ($t_{\frac{1}{2},i} = \frac{ln\ 2}{p_i}$). Additionally, the IL-6 input concentration $u$ was fixed to 0.17 and the time discretization constant $\Delta t$ was set to 1 min.

To obtain inner approximation of parameters $p_i$ the measurement data $y(k)$=(p)STAT3 were expressed using binary variables and inverted as described in Sections 5.2 and 5.4.

**Results:** We aimed to estimate an inner approximation for the parameters $p_7$ and $p_8$ as they describe the velocity of STAT3 phosphorylation and dephosphporylation, respectively. Using Algorithm 1, which allows for bisectioning of the parameter space, no inner approximation was determined. Then, we applied Algorithm 2 starting from

a valid parameter sample that was determined by a local optimization routine using the Matlab inbuild function *fmincon*. Briefly, we generated different initial parameter samples using the Matlab inbuild function *rand* and performed least square fitting. Finally, we stored the obtained parametrization which gave the minimum distance between the model simulation and the data. The procedure was repeated three times. The obtained parametrizations were used as starting point for Algorithm 2. To this end, parameter $\nu$ which defines the step-sizes for moving the facets outwards, was set to 0.001. Fig. 5.5 depicts the results for outer approximations and the polytopic sample expansion algorithm for three valid parametrizations. Note that the results for parameter estimation are depicted on a logarithmic scale. As can be seen the polytopic sample expansion algorithm partly yielded very small inner approximations (see e. g. Fig. 5.5(a) and (b), yellow and green samples as well as Fig. 5.5(c), red and green samples). These results indicate that, depending on the valid parameter sample, the system is very sensitive with respect to small changes in the parameter values. Additionally and as already mentioned, no inner approximations could be determined using the bisectioning algorithm (Algorithm 1). The reason might be, other than for the example of an enzyme-catalyzed reaction that there exist no connected inner approximation. Indeed, the recursion depth $N_{max}$ in Algorithm 1 could be defined, such that the determined hyperrectangular inner approximations depicted in Fig. 5.5(a), (b), and (c) are met. However, this is not in relation to the computational effort needed to determine this inner approximation.

Fig. 5.5(d) depicts 40 model trajectories (red bold lines) obtained by sampling within the determined hyperrectangular inner approximation $[0.54;1.18] \times [0.32;0.82] \times [0.22;0.37] \times [0.30;0.57] \times [0.024;0.032] \times [0.012;0.019]$ (black rectangles Fig. 5.5(a)-(c)). This hyperrectangular inner approximation was obtained by expanding the valid sample $s_1$ (red crosses Fig. 5.5(a)-(c)) using Algorithm 2. All trajectories satisfy the boundaries for the model output (p)STAT3. Thus, Algorithm 2 allows to determine an inner approximation at least in the surrounding area of valid parametrizations. Indeed, these polytopic inner approximations may not be unique and of maximum volume. However, they allow for consistent and reliable model predictions. Notably, further studies about the identifiability of the determined polytopic inner approximations would be beneficial.

It is sometimes intricate to estimate inner approximations for real applications such as IL-6-induced signaling. One important reason is that the algorithms used to determine inner (and outer) approximations are numerically very high-demanding. As an example, when using bisectioning, the overall cost grows exponentially with the number of hyperrectangles. Therefore, the number of parameters for which an inner approximations can be derived, need to be considered carefully. Nevertheless, often it is sufficient to determine inner approximations for a subset of the model parameters, such as the most sensitive ones. Then, the computational effort can be reduced to this subset.

(a)

(b)

(c)

(d)
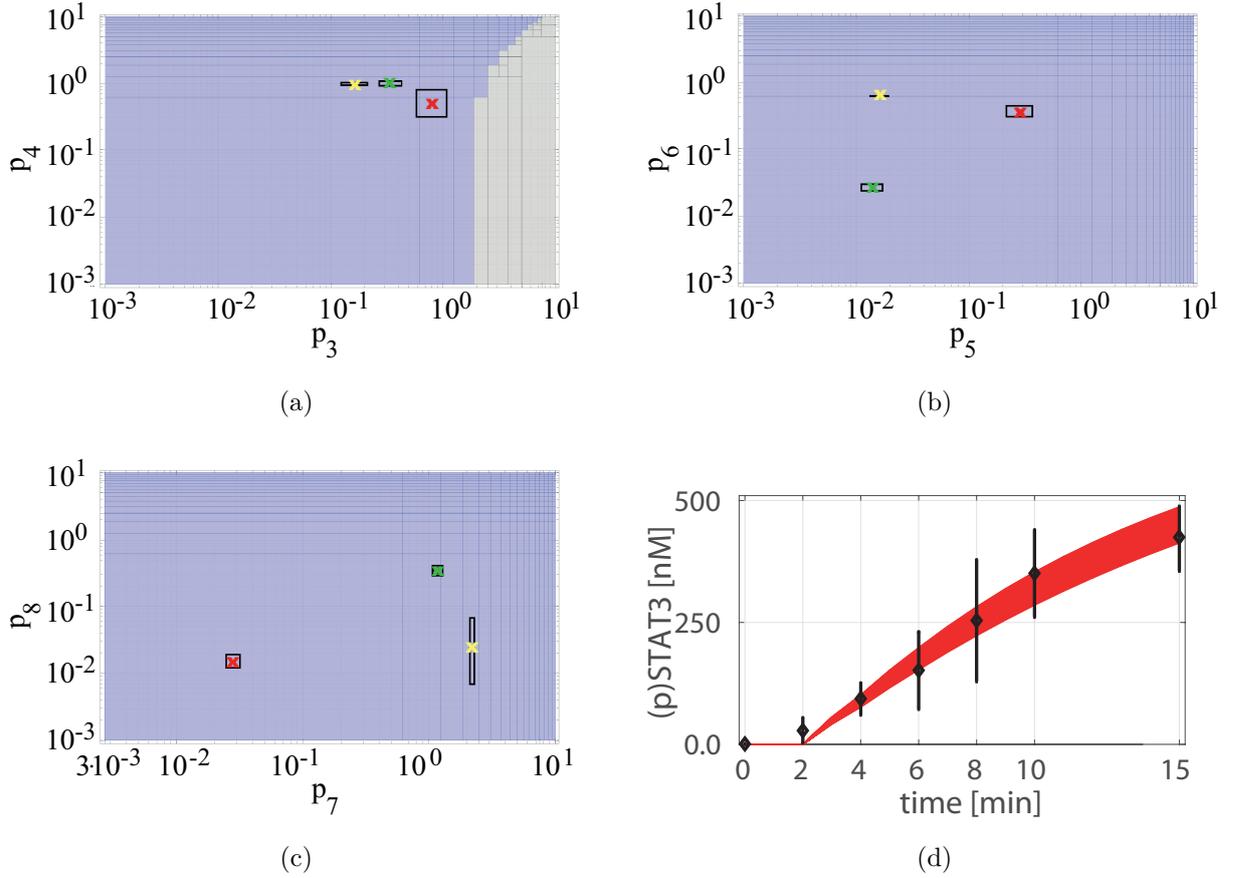
Figure 5.5: **Results for parameter estimation.** Outer approximations (blue boxes) for parameters (a) $p_3$ and $p_4$, (b) $p_5$ and $p_6$, as well as (c) $p_7$ and $p_8$. Gray boxes denote guaranteed invalid parameter ranges, black rectangles show the obtained hyperrectangular inner approximations using Algorithm 2 starting with the valid samples $s_1 = [0.78; 0.50; 0.28; 0.37; 0.028; 0.015]$ (red crosses), $s_2 = [0.16; 0.98; 0.01; 0.62; 2.18; 0.02]$ (yellow crosses), and $s_3 = [0.33; 0.98; 0.01; 0.02; 1.15; 0.35]$ (green crosses). (d) 40 model simulations (red trajectories) using parameter samples from the hyperrectangular inner approximations for $p_3$-$p_8$ obtained with the valid sample $s_1$ (red crosses in (a)-(c)).

The presented approach leads to a problem that scales linearly with the number of measurement constraints with respect to which an inner approximation is searched for. As a consequence, it might lead to an intractable problem if the number of constraints is large. As a remedy, the inner approximation can be refined and possibly enlarged by considering new or additional measurement constraints in a recursive manner.

## 5.8 Summary

In this chapter, we proposed an extension of the set-based estimation approach by determining inner approximations of parameter sets. To this end, quantitative measurement data were reformulated using binary variables. As these binary variables indicate whether the measurement constraints are simultanously fulfilled or violated, they were combined via a logical *and*-combination. Since the relaxation processes entail an increase of the solution space (i. e. obtaining false positive solutions), our crucial idea consisted in inverting the measurement constraints. If, in such case, the inverted problem admitted no solution, we could guarantee that an inner approximation was determined. Furthermore, we proposed two algorithms that implement the presented approach.

We demonstrated applicability of the method by means of two examples, i. e. an enzyme-catalyzed reaction and early IL-6-induced trans-signaling and activation of Jak/STAT3 signal transduction. For the first example, we were able to determine tiny inner approximations of the model parameters, possibly due the assumption that measurement data for all species are available. However, the second example showed that an inner approximation may not be a largly connected area, rather only a small range around different valid parametrizations that are obtained by local optimization.

So far, we discussed set-based methods for parameter estimation and model invalidation. However, it is likely to happen that several hypotheses for the considered system under study exist, that cannot be discriminated by means of the available experimental data. Then, the modeler needs to propose meaningful experiments that allow to distinguish invalid hypotheses from the probably valid ones. To this end, we propose in the following an approach to tackle the problem of experimental design under uncertainties.

# 6 Optimal Experimental Design for Model Discrimination and Selection

Usually, when considering biological systems not all interactions of the underlying molecular mechanisms are fully understood and known. As a consequence, several competing model hypotheses are developed and need to be evaluated for (in-)validity. We provided in the previous chapters means how set-based methods and infeasibility certificates can be used to check whether models are capable to explain measurement data or not. In case a model cannot represent given measurement data, it is deemed to be invalid and can be ruled out with guarantees. The model discrimination, however, might be inconclusive as model parameters are uncertain or even unknown and measurement data are possibly very noisy. Thus, it may not be possible to discriminate between valid and invalid models. Consequently, more informative experiments are needed to select a potentially valid model.

In this chapter, we propose a set-based experimental design approach for guaranteed discrimination of competing models. It is based on bilevel optimization. In the outer program an input that minimizes a given norm and satisfies input constraints is determined. In turn, the inner program certifies that the reachable output sets of the models are non-overlapping for at least one time point for the determined input. This allows model discrimination by separation of the variable sets facilitates the task of model selection

The chapter is structured as follows. First, the problem of model discrimination and selection is introduced. Second, the proposed bilevel optimization approach is explained. An illustrative example is presented considering IL-6-induced receptor formation as shown in Chapter 4.

The results presented in this chapter are based on our work in [134].

## 6.1 Motivation - Design of Experiments

The design of experiments builds an important bridge between mathematical modeling and experimental analysis allowing a deeper understanding of the system under study. In particular, when designing new experiments one of the major question is, how to perturb a system, such that the most valuable information, e. g. for model discrimination can be obtained.

Several approaches for experimental design in the field of control theory have been developed and applied to biological problems. A noteworthy approach for model dis-

crimination and selection was presented in [6]. The method allows for the design of a stimulus (i. e. the model input) to facilitate model discrimination for signaling. Specifically, the authors presented two formulations of a model-based controller that is used to design a dynamic stimulus. In this context, a controller is a model that monitors a certain process (e. g. a dynamical model) to drive a certain output behavior. In [6], for example, the designed input signals were applied to the model candidates driving the model output to a target trajectory. Such a target trajectory can be the time-dependent change of a measured state variable, such as the dynamic phosphorylation of STAT3. The quality of a model candidate is assessed by the ability of the corresponding controller to drive the system to follow the specific target trajectory.

The model discrimination task is generally difficult to solve due to the uncertainties in parameters and/or state variables. In the past, experimental design methods for model discrimination under uncertainties were proposed. For example, in [154] uncertainties of the parameters and state variables are taken into account using a polynomial chaos approach. Polynomial chaos is a non-sampling-based method to determine the propagation of uncertainty in a dynamical system. The method is used in [154] to approximate the probability density functions of the model outputs. To discriminate between the uncertain nonlinear models, a measure based on the Bhattacharyya coefficient is introduced. The Bhattacharyya coefficient describes a measure of the amount of overlap between two statistical samples or populations [22]. The approach is applied to an example considering the Michaelis-Menten and Henri mechanisms as two model hypotheses for an enzyme-catalyzed reaction. The approach presented in [58] furthermore allows model discrimination under experimental uncertainties. The authors account for such uncertainities by introducing an optimal control-based approach and applying the sigma-point method [81] to reduce parameter uncertainties. As a measure for model discrimination, the overlapping areas of the expected response probability distribution functions (PDFs) are computed.

Another approach for model discrimination and experimental design was presented in [129]. By applying a profile likelihood approach, the model information is maximized for robust parameter identification. Also Bayesian approaches for the design of experiments are frequently applied [36]. Bayesian experimental design is based on Bayesian inference, in which Bayes' law (i. e. the probability of an event based on prior knowledge) is used to update the probability for a hypothesis as more data becomes available. Notably, stochastic approaches require data which follow a normal distribution. However, this is rarely fulfilled.

In this chapter, we devise a novel approach for optimal experimental design for the purpose of guaranteed model discrimination using set-based methods and bilevel optimization. In particular, our approach determines a model input to discriminate between competing model hypotheses. The approach consists of two programs, i. e. an *inner* and an *outer program.* Thereby, we aim to determine a locally optimal input

(outer program) such that two competing model hypotheses can be globally discriminated (inner program), whereby 'globally' means with guarantees and considering uncertainities in parameters and model state variables.

Notably, the determination of optimal time points at which the output sets can be separated and at which measurements shall be taken, is not directly in the focus of this chapter. However, in general the approaches for both, the design of an optimal stimulus and optimal measurement time points cannot be clearly separated and can be easily integrated in the approach.

## 6.2 Bilevel optimization for model discrimination

In the following, we present the proposed solution approach, and the inner and outer programs allowing for model discrimination under uncertainties.

We consider $n_h$ competing discrete-time models as described in (4.1) and similar to (3.16):

$$h^{[i]} : \begin{cases} f^{[i]}(x^{[i]}(k+1), x^{[i]}(k), u(k), p^{[i]}) = 0, \\ g^{[i]}(y^{[i]}(k), x^{[i]}(k), u(k), p^{[i]}) = 0. \end{cases} \tag{6.1}$$

The model hypotheses $h^{[i]}$ differ in their functions $f^{[i]}$ and/or $g^{[i]}$. Furthermore, each model has its own set of variables $x^{[i]}(k)$, $y^{[i]}(k)$ and $p^{[i]}$. Notably, however, all models have the same input $u(k)$, which should be designed. In the following, we do not consider measurement data $\mathcal{M}(t_k)$, but rather the reachable sets for the model outputs $y(k)$, i. e. $y(k) \in \mathcal{Y}(k) \subseteq \mathbb{R}^{n_y}$.

The general problem of set-based experimental design and the derived solution approach are illustrated in Fig. 6.1: Let in the following the set of competing models be $\mathcal{H} := \{h^{[1]}, h^{[2]}\}$. Then, $\mathcal{Y}^{[1]}(k)$ and $\mathcal{Y}^{[2]}(k)$ for $k \in \mathcal{T}$ denote the corresponding output sets (cf. Fig. 6.1(a), dark and light gray bars, respectively). Uncertainties in the output sets result basically from i) uncertainties in model parameters and/or model states that have to be propagated through the system, and ii) the applied input sequence $\boldsymbol{u} := \{u(0), \dots, u(n_t - 1)\}$ belonging to the uncertainty set $\mathcal{U}(k)$. As can be seen, both output sets $\mathcal{Y}^{[1]}(4)$ and $\mathcal{Y}^{[2]}(4)$ overlap for the time point of interest, i. e. $k=4$ (cf. Fig. 6.1(a), rectangle with rounded corners, left and right panel). Consequently, data at non-overlapping area would be needed to discriminate between both hypotheses and to allow guaranteed invalidation of at least one model. Therefore, one needs to design a new input sequence $\hat{\boldsymbol{u}} := \{\hat{u}(0), ..., \hat{u}(n_t - 1)\}$, such that the output sets do not overlap at least at one time point (cf. Fig. 6.1(b), rectangle with rounded corners, left and right panel). If then a single measurement would be taken at the considered time point, it is guaranteed that at least one of the models can be invalidated if the taken measurement does not fall into one of the output sets.

Notably, since it would be too restrictive or even infeasible to demand output set
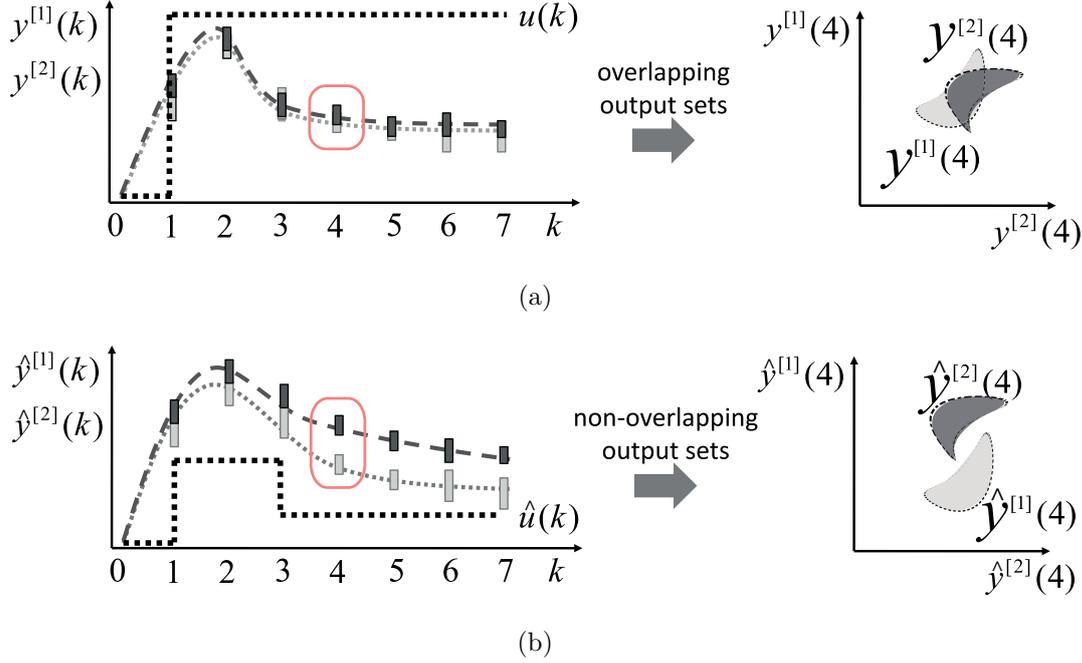
Figure 6.1: **Solution approach for set-based optimal experimental design for model hypotheses 1 and 2.** Left panel: temporal trajectories for both models and the uncertain model output sets $\mathcal{Y}^{[1]}$ and $\mathcal{Y}^{[2]}$. Right panel: Reachable output sets in state space obtained for each model using the dedicated parameter set. (a) Overlapping output sets $\mathcal{Y}^{[1]}$ and $\mathcal{Y}^{[2]}$ at time point $k{=}4$ (red rectangle with rounded corners, left panel) for an initial applied input $\mathbf{u}$. (b) Non-overlapping output sets $\hat{\mathcal{Y}}^{[1]}$ and $\hat{\mathcal{Y}}^{[2]}$ at $k{=}4$ (red rectangle with rounded corners, left panel) for a newly designed input $\hat{\mathbf{u}}$.

separation for all time points $k$, it is sufficient to consider only a subset or even a single time point.

To realize the described approach, we apply the set-based methods introduced in Chapter 4 combined with bilevel optimization. Therefore, we introduce in the following an inner and an outer program. While the goal of the inner program is to check output set separation for defined time points, the outer program is used to design an input (sequence) for model discrimination. The designed input is steadily supplied to the inner program until output set separation can be guaranteed for the considered time point. Please note that we consider in the following two competing model hypotheses for simplicity of notation and presentation of the optimal experimental design approach. The approach, however, is valid for more than two model hypotheses to be considered.

**Output set separation – Inner program:** The inner program aims to check output set separation for a considered time point (or several time points). To this end, we introduce the variable $\tau \in \mathcal{T}$ describing those time points for which output

set separation shall be reached.

Let $\hat{\mathcal{Y}}^{[1]}(k)$ and $\hat{\mathcal{Y}}^{[2]}(k)$, $k \in \mathcal{T}$ denote the output sets for hypothesis 1 and 2 resulting from the applied input sequence $\hat{\boldsymbol{u}}$ and uncertainties. Then, the distance $\hat{\delta}$ of the model output sets can be defined into an optimization problem of the form:

$$
\begin{aligned}
\hat{\delta} := \min \quad & \left(\hat{y}^{[1]}(\tau) - \hat{y}^{[2]}(\tau)\right)^2 \\
\text{s.t.} \quad & \hat{y}^{[1]}(\tau) \in \hat{\mathcal{Y}}^{[1]}(\tau) \\
& \hat{y}^{[2]}(\tau) \in \hat{\mathcal{Y}}^{[2]}(\tau).
\end{aligned} \tag{6.2}
$$

For guaranteed model discrimination, $\hat{\delta} > 0$ is required, which is the condition for output set separation at time point(s) $\tau$. The output sets $\hat{\mathcal{Y}}^{[1]}(k)$ and $\hat{\mathcal{Y}}^{[2]}(k)$ are not known explicitly. Therefore, we reformulate in the following the optimization problem (6.2) using the model equations (6.1) and the uncertainty descriptions of both models $h^{[1]}$ and $h^{[2]}$ according to Chapter 3, Section 3.1:

$$
\begin{aligned}
\hat{\delta} := \min \quad & \left(\hat{y}^{[1]}(\tau) - \hat{y}^{[2]}(\tau)\right)^2 \\
\text{s.t.} \quad & \text{model equations and uncertainties for } h^{[1]}: \\
& f^{[1]}(\hat{x}^{[1]}(k{+}1), \hat{x}^{[1]}(k), p^{[1]}, \hat{u}(k)) = 0 \\
& g^{[1]}(\hat{y}^{[1]}(k), \hat{x}^{[1]}(k), p^{[1]}, \hat{u}(k)) = 0 \\
& \hat{y}^{[1]}(k) \in \hat{\mathcal{Y}}^{[1]}(k), \forall k \in \mathcal{T}, \\
& \hat{x}^{[1]}(k) \in \hat{\mathcal{X}}^{[1]}(k), \forall k \in \mathcal{T}, \\
& p^{[1]} \in \mathcal{P}^{[1]} \\
& \hat{u}(k) \in \mathcal{U}(k), \forall k \in \mathcal{T} \\
& \text{model equations and uncertainties for } h^{[2]}: \\
& f^{[2]}(\hat{x}^{[2]}(k{+}1), \hat{x}^{[2]}(k), p^{[2]}, \hat{u}(k)) = 0 \\
& g^{[2]}(\hat{y}^{[2]}(k), \hat{x}^{[2]}(k), p^{[2]}, \hat{u}(k)) = 0 \\
& \hat{y}^{[2]}(k) \in \hat{\mathcal{Y}}^{[2]}(k), \forall k \in \mathcal{T}, \\
& \hat{x}^{[2]}(k) \in \hat{\mathcal{X}}^{[2]}(k), \forall k \in \mathcal{T}, \\
& p^{[2]} \in \mathcal{P}^{[2]}, \\
& \hat{u}(k) \in \mathcal{U}(k), \forall k \in \mathcal{T}.
\end{aligned} \tag{6.3}
$$

Again, we have defined, the dynamical model equations $f^{[1]}$, $g^{[1]}$ and $f^{[2]}$, $g^{[2]}$, the state variables $\hat{x}^{[1]}(k)$ and $\hat{x}^{[2]}(k)$, the model outputs $\hat{y}^{[1]}(k)$ and $\hat{y}^{[2]}(k)$, the model parameters $p^{[1]}$ and $p^{[2]}$, for both models 1 and 2, and the common input $\hat{u}(k)$, respectively.

Problem (6.3) is in general a nonlinear and nonconvex optimization problem. Since it is difficult to derive a solution, (6.3) is relaxed into a convex semidefinite optimization problem as described in Chapter 4. The solution that is obtained by relaxing and solving (6.3) is denoted as $\hat{\delta}_{\text{relax}}$ which satisfies the relation $\hat{\delta} \geq \hat{\delta}_{\text{relax}}$. Thus, if the output sets $\hat{\mathcal{Y}}^{[1]}(\tau)$ and $\hat{\mathcal{Y}}^{[2]}(\tau)$ do not overlap, i. e. $\hat{\delta} > 0$, then the convex relaxation approach guarantees output set separation due to the relations $\hat{\delta} \geq \hat{\delta}_{\text{relax}} > 0$.

**Determining a separating input – Outer program:** The outer program is used to determine an input sequence $\hat{\boldsymbol{u}}$ of minimum norm that satisfies input constraints and for which output set separation can be certified by the relaxed inner program, i. e. $\hat{\delta}_{\text{relax}} > 0$. The resulting bilevel optimization problem is given by

$$
\begin{aligned}
\min \quad & \sum_{k=0}^{n_t-1} \|\hat{u}(k)\| \\
\text{s.t.} \quad & \hat{u}(k) \in \mathcal{U}(k), \forall k \in \mathcal{T} \\
& \hat{\delta}_{\text{relax}} > 0.
\end{aligned}
\tag{6.4}
$$

The basic idea is now to solve the (nonconvex) outer program using a deterministic local nonlinear solver. The obtained input sequence $\hat{\boldsymbol{u}}$ is subsequently supplied to the relaxed inner program. The procedure is repeated until a (locally) optimal input sequence is obtained.

## 6.3 Example: Stepwise receptor assembly

In the following, we illustrate the proposed approach considering IL-6-induced receptor assembly and activation as shown in Chapter 4, Section 4.5. We considered two hypotheses, which describe simplified versus stepwise receptor assembly and activation. Measurement data were generated by simulating the model describing stepwise receptor assembly with nominal values. Thereby, we were able to demonstrate invalidity of the model describing a simplified receptor assembly.

We here assume that no (simulated) measurement data for (p)R$_{\text{complex}}$ are available. Using the presented approach, we aim to design an optimal experiment that allows to discriminate between both hypotheses. In particular, one aims to design an optimal IL-6 stimulus concentration to separate the output sets for (p)R$_{\text{complex}}$ for both hypotheses with guarantees.

**Model setup:** The model and the setup for optimal experimental design is akin to the setup described in Chapter 4, Section 4.5. Notably, it is assumed that all model parameters lie within the uncertainty intervals $p_j=[0.1,1]$, $j = \{1,...,6\}$.

By applying the set-based experimental design approach an input $\hat{\mathbf{u}}=\widehat{\text{IL}-6}$ is designed, such that both model hypotheses can be clearly discriminated. Notably, for output set separation, a constant IL-6 input, i. e. IL-6=$u(0) = \ldots = u(n_t - 1)$ within the uncertainty range IL-6=[0,1] is used. Notably, the inner program is implemented and solved using the set-based approach and ADMIT. As a consequence, model discrimination can be globally guaranteed. In contrast to that, the outer program is computed locally using the Matlab inbuild optimization routine *fmincon*, which is based on the Nelder-Mead algorithm.

**Results:** The results for experimental design are depicted in Fig. 6.2. For an initial input concentration of IL-6=1, the output sets for $(p)R_{complex}^{[1]}$ (cf. Fig. 6.2(a), blue bars) and $(p)R_{complex}^{[2]}$ (cf. Fig. 6.2(a), red bars) overlapped for all time instances $k$. Overlapping for all time instances indicates that no model discrimination is possible for these conditions. After application of the proposed framework, a new IL-6 input concentration was determined that allows a clear separation of the output sets at $\tau_1=2$ min and $\tau_2=3$ min (cf. Fig. 6.2(b), inlet). The designed IL-6 input was determined to one fifth of the initial IL-6 concentration, i. e. $\widehat{IL-6}=0.2$.
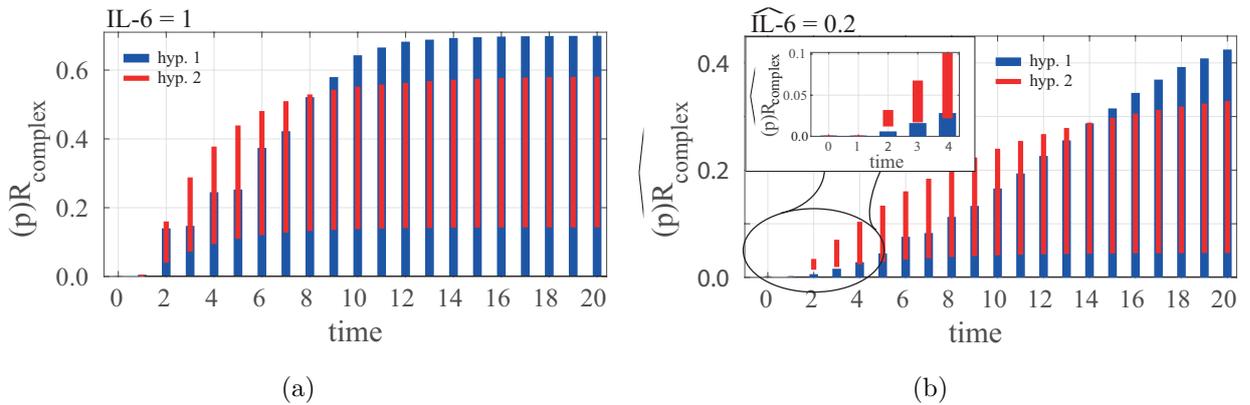


Figure 6.2: **Results for set-based experimental design.** (a) Outer approximation of overlapping model output sets $(p)R_{complex}$ for hypothesis 1 and 2 (blue and red bars, respectively) for an initial input IL-6 = 1. (b) Outer approximation of non-overlapping model output sets $(p)\widehat{R_{complex}}$ for hypothesis 1 and 2 (blue and red bars, respectively) for a designed input $\widehat{IL-6} = 0.2$. Output set separation was achieved for time points $\tau_1 = 2$ min and $\tau_2 = 3$ min (inlet, see non-overlapping area).

The results for set separation strongly depend on the intital chosen parameter uncertainties. Separation of the output sets $(p)R_{complex}^{[1]}$ and $(p)R_{complex}^{[2]}$ would not have been possible for larger uncertainty intervals as assumed. Consequently, for the given example, the approach requires good initial knowledge of the parameter ranges. This, however, is often non-restrictive as usually initial experiments allow an outer or even an inner approximation of the underlying parameters. Hence, the initial uncertainty ranges can be reduced to a reasonable interval and then, the experimental design approach can be applied.

In the example, the time points at which both output sets $(p)\widehat{R_{complex}^{[1]}}$ (cf. Fig. 6.2(b), blue bars) and $(p)\widehat{R_{complex}^{[2]}}$ (cf. Fig. 6.2(b), red bars) could be clearly separated were determined to be $\tau_1=2$ min and $\tau_2=3$ min. Receptor assembly and activation due to IL-6 stimulation describes very fast processes and hence, occurs within the first minutes of signaling. Intuitively, if certain time points are important to decipher

valuable information about receptor assembly, they must be temporarily close after start of IL-6 stimulation.

To underpin that especially early time points after IL-6 stimulation contain the most valuable information to discriminate between two model hypotheses describing IL-6-induced receptor assembly and activation, we compute in the following the FIM (Fisher Information Matrix), cf. Chapter 1: The FIM determines the amount of information that a measurement contains about an unknown parameter. In general, if the FIM has large values, then the information content is high. In contrast, if only small values are obtained for the FIM, then the information content is low.

To this end, we first determined the sensitivities $S$ for the (theoretical) measurable model species $(p)R_{complex}$ at each time point $t_j$, $j = \{0, ..., 20\}$ with respect to the model parameters $p_i$, $i = \{1, ..., 6\}$. The sensitivity matrix for the considered system was obtained by computing:

$$S = \begin{bmatrix} \frac{\partial(p)R_{complex}}{\partial p_1}(t_0) & \cdots & \frac{\partial(p)R_{complex}}{\partial p_6}(t_0) \\ \vdots & \cdots & \vdots \\ \frac{\partial(p)R_{complex}}{\partial p_1}(t_{20}) & \cdots & \frac{\partial(p)R_{complex}}{\partial p_6}(t_{20}) \end{bmatrix}.$$

Then, the FIM was calculated by

$$\text{FIM} = \sum_{j=0}^{20} \text{SS}^T.$$

Fig. 6.3 depicts the obtained results.
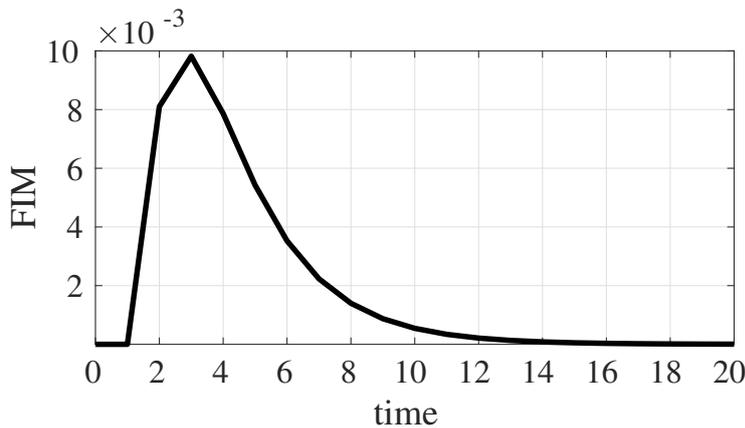


Figure 6.3: **Computation of the FIM (black bold line) for optimal experiment design.** The FIM reaches its maximum at t = 3 min indicating that early time points contain the most valuable information to discriminate between two model hypotheses describing IL-6-induced receptor assembly and activation.

The FIM increases at t=1 min and reaches its maximum at t=3 min. After that, the FIM decreases and reaches a minimum after 20 minutes. Consequently, early

time points, i. e. at 2, 3 and 4 min have the most valuable impact on the design of experiments that allow for model discrimination between simplified versus stepwise receptor assembly and activation. These findings are in line with our results for set-based experimental design.

The presented set-based approach does not directly consider the determination of optimal time points for which the most meaningful information can be obtained. We note that the time points $\tau$ for which output set separation shall be reached, have to be chosen carefully dependent on the biological system under study. Otherwise, the method could fail. As an example, model discrimination for receptor assembly and activation would not have been possible considering measurement time points $\tau$ greater than 10 minutes. Therefore, computing first the FIM to determine the information content of a measurement over time, followed by the proposed bilevel approach is beneficial to decipher an optimal stimulus concentration allowing for successful model discrimination.

Currently, the model for IL-6-induced receptor assembly and activation contains $(p)R_{complex}$ as model output. The experimental quantification of $(p)R_{complex}$ is challenging. As potential remedy and to validate our results for simplified versus stepwise receptor assembly and activation, one could analyse phosphorylated Jak (i. e. (p)Jak, which can be detected by Western Blotting) as a surrogate for $(p)R_{complex}$. One important assumption would then be that (p)Jak serves as an indicator for receptor activation and phosphorylation. We refer the interested reader to [49] for the corresponding experimental setup.

## 6.4 Summary

In this chapter, we presented a set-based experimental design approach which allowed for guaranteed model discrimination and selection under uncertainties. The method is based on bilevel optimization and consists of an inner and an outer program. The inner program fosters the output set separation by combining competing model hypotheses into one optimization problem. As it is difficult to directly derive a solution (e. g. due to model nonlinearities), the optimization problem was reformulated and subsequently relaxed into a semidefinite convex optimization problem. For the convexified optimization problem output set separation was checked and could be guaranteed globally using appropriate convex solvers. In the outer program, we formulated a nonconvex optimization problem which aimed to determine new model inputs. These inputs were steadily supplied to the convex inner program checking whether or not the model output sets for the determined inputs overlap.

This method was applied to two model hypotheses for IL-6-induced receptor assem-

bly and activation. The model hypotheses described simplified receptor assembly and activation and stepwise receptor assembly and activation, respectively. The output sets for both hypotheses were considered to be the phosphorylated and activated receptor complex. Notably, for an IL-6 input concentration of 1 both output sets overlapped initially. After having designed an optimal IL-6 input concentration of 0.2 the output sets for both hypotheses were clearly separated. In particular, the study indicated that especially early time points of IL-6-induced receptor assembly and activation are important for experimental design and model separation as they contain the most information. Thus, our results allow in further to plan experiments accordingly gaining new insights into IL-6-induced receptor assembly and activation.

The outlined approaches allow an in-depth study of signal transduction pathways on the short-term time scale, i. e. biological processes ranging from minutes to approximately one hour. For a deeper understanding of IL-6-induced (patho-)physiological processes also processes acting on the long-term time scale (i. e. over days) have to be taken into account. Set-based methods are restricted with respect to the problem size, which scales directly with the sampling time and the time horizon. Further approaches are needed to consider both, the short- and the long-term time scale at the same time. Therefore, we present in the following chapter methods allowing for the combination of short-term signaling with long-term processes using the set-based framework.

# 7 Set-based Multi-scale Modeling and Data Integration

The modeling and estimation of biological processes acting on different time scales as well as the integration of data along these temporal scales is a key challenge in systems biology and systems medicine. The combination of both, short- and long-term time scale, is indispensable to understand the development of diseases based on molecular deregulations on the short-term time scale.

We present and apply a framework that allows to combine processes and data on the short- and long-term time scale within one model preserving biological functionality under uncertainties. In addition, we expand the approach with an unified framework that allows for the stratification of a patient cohort into subcatgeories of high, medium and low risk. In more detail, the unified patient stratification framework is based on a fusion of the set-based multi-scale modeling framework with classification approaches. The fusion of set-based methods with classification approaches inaugurates a new field for possible applications of set-based methods and the usage of the information obtained by set-based estimation.

The chapter is structured as follows. First, we review a phenomenological approach to combine short-term signaling events with long-term cellular responses. We apply this approach considering the influence of IL-6-induced Jak/STAT3 signaling on cell growth. Second, we derive a framework for patient stratification. The approach is applied to models that describe IL-6-induced Jak/STAT3 and MAPK trans-signaling and their influence on inflammation.

The chapter is based on our results presented in [133] and [132].

## 7.1 Introduction

Biological processes often possess subprocesses of different temporal dynamics [46, 167]. The understanding of all processes from different time scales, such as short-term signaling as well as long-term cellular responses, is important when it comes to the model-guided development and design of new intervention strategies. However, the integration of short-term signaling events and long-term cellular responses into mathematical models is challenging due to the multiple biological factors that influence this integration. These factors are often not yet well characterized and understood and thereby cannot be modeled in detail. Furthermore, even if all biological factors are known, the modeling of biological processes from both time scales leads to computa-

tionally demanding problems with a large number of model states and parameters to be estimated. The introduced set-based approach is restricted in the problem size that can be considered [136]. Consequently, for the set-based framework and for the therein presented methodological extensions often only *medium-size* models can be considered. The term *medium-size* cannot be clearly defined and is problem-dependent. For modeling, estimating and analyzing IL-6-induced signaling an alternative approach is needed to understand the resulting (patho-)physiological long-term effects.

In this chapter, we first outline a so-called phenomenological approach towards multi-scale modeling. This approach allows to combine short-term signaling events with long-term responses using the set-based framework. We use the set-based phenomenological approach for analyzing and fusing it with classification approaches. In more detail, we demonstrate how set-based information, such as lower and upper bounds on short-term model parameters can be used to stratify a patient cohort into risk subcategories for developing inflammatory diseases on the long-term time scale. The specific stratification problem considered in this work is based on individual short-term signaling profiles of phophorylated STAT3 acting as an inflammatory upstream marker and its integrated response which serves as the long-term outcome.

## 7.2 Combining short-term signaling events with long-term cellular responses

The following sections built upon the work presented in [143]. The approach proposed to combine early short-term signaling events with long-term cellular responses using a linear correlation of shape properties of activated signaling molecules, e. g. the maximum peak height or the integral of a signal, with the strength of long term responses (cf. Fig. 7.1). In the following, we apply these ideas using the set-based framework.

### 7.2.1 Decoding approaches

We extend the model for IL-6-induced short-term signaling using the approaches described in [143] and as illustrated in Fig 7.1. One approach for combining short-term signaling events with long-term cellular responses is to correlate the integral of the trajectory of a time-dependent signal S(t) linearily with the strength of the long-term response $L_{int}$:

$$L_{int} = \alpha \cdot \int_{t_0}^{t_{end}} S(t)\,dt.$$ 

(7.1)

Here, $\alpha$ is an uncertain parameter describing the linear dependency of the integral and the strength of the long-term response, and $t_0$ as well as $t_{end}$ are the initial and final time points for computing the integral.
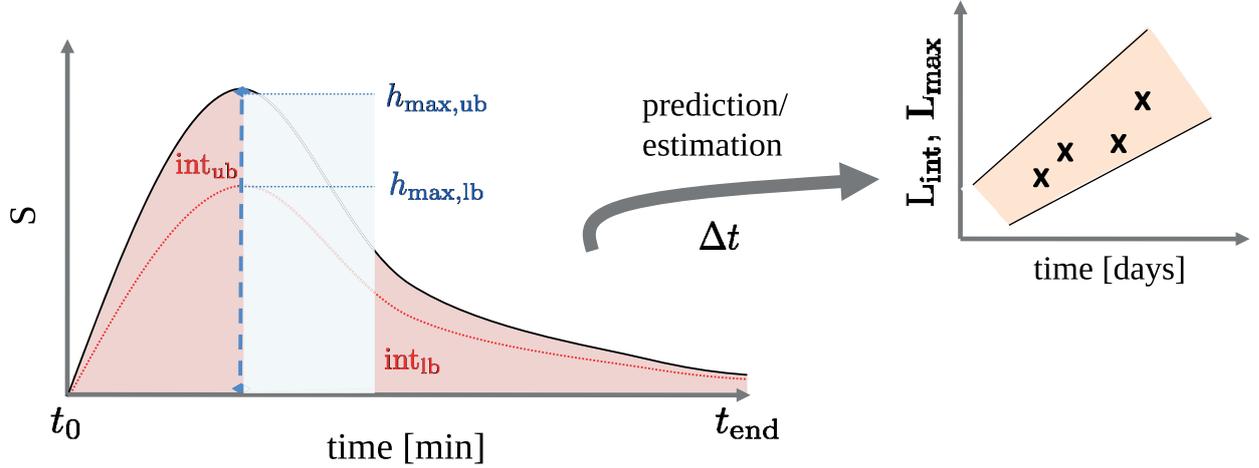
Figure 7.1: **Schematic representation of the approaches for decoding a short-term signal S into a long-term response L.** Left panel: The black bold and the dotted red line represent two time-dependent kinetics for a time-dependent short-term signal S. Further depicted are the integrals for the corresponding upper ($\text{int}_{\text{ub}}$) and lower ($\text{int}_{\text{lb}}$) trajectory as well as the maximum peak heights ($\text{h}_{\text{max}}$) for the corresponding upper ($\text{h}_{\text{max,ub}}$) and lower ($\text{h}_{\text{max,lb}}$) trajectory (dashed light blue lines and light blue box). Right panel: By computing the corresponding parameters from the left panel our goal is to combine the short- and long-term time scale to estimate, predict and understand uncertain long-term (patho-)physiological responses $\text{L}_{\text{int}}$ and $\text{L}_{\text{max}}$ (black crosses within orange corridor).

Another approach is to correlate the maximum peak height of a signal S(t) linearly with the strength of the long-term response $\text{L}_{\text{max}}$:

$$\text{L}_{\text{max}} = \beta \cdot \text{h}_{\text{max}}. \tag{7.2}$$

In Eq. (7.2), $\beta$ is an uncertain parameter describing the linear dependency of the maximum signal peak height and the strength of the long-term response.

In the following, we integrate Eq. (7.1) and (7.2) into the set-based framework using IL-6-induced STAT3 phosphorylation as the short-term outcome S(t) and IL-6-induced cell growth as the long-term responses $\text{L}_{\text{int}}$ and $\text{L}_{\text{max}}$, respectively.

## 7.2.2 Example

We apply the approach to unravel the influence of IL-6-induced short-term Jak/STAT3 classic-signaling on IL-6-induced long-term cell growth. First, we describe the model as well as measurement data and the set-based problem setup.

**Model for IL-6-induced Jak/STAT3 classic-signaling and cell growth:** We adapt the model as depicted in (4.9). In addition, we consider SOCS3 mRNA transcription and the translation to SOCS3 protein. Furthermore, we include the ideas

given by Eq. (7.1) and (7.2) leading to:

$$
\begin{aligned}
x_1(k{+}1) &= x_1(k) + \Delta t\Big(p_1 u x_7(k) - p_2 x_1(k) \\
&\quad - 2p_3 x_2(k)^2 x_1(k)^2 + 2p_4 x_8(k)\Big) \\
x_2(k{+}1) &= x_2(k) + \Delta t\Big(2p_4 x_8(k) - 2p_3 x_2(k)^2 x_1(k)^2\Big) \\
x_3(k{+}1) &= x_3(k) + \Delta t\Big(\frac{p_5 x_8(k)}{1 + p_{13} x_6(k)} - p_6 x_3(k)\Big) \\
x_4(k{+}1) &= x_4(k) + \Delta t\Big(p_7 x_9(k) x_3(k) - p_8 x_4(k)\Big) \\
x_5(k{+}1) &= x_5(k) + \Delta t\Big(p_9 x_4(k) - p_{10} x_5(k)\Big) \\
x_6(k{+}1) &= x_6(k) + \Delta t\Big(p_{11} x_5(k) - p_{12} x_6(k)\Big)
\end{aligned}
\tag{7.3a}
$$

$$
y^s(k) = x_4(k) \tag{7.3b}
$$

$$
\begin{aligned}
G_{\text{int}} &= \frac{\alpha}{2} \cdot \sum_{k=0}^{k_{end}} \Big(x_4(k{+}1) + x_4(k)\Big) \cdot \Delta t \\
G_{\text{max}} &= \beta \cdot \max\Big(x_4(k)\Big)
\end{aligned}
\tag{7.3c}
$$

$$
y^l = \big(G_{\text{int}}, G_{\text{max}}\big). \tag{7.3d}
$$

Here the variables $x_1(k)$, $x_2(k)$, $x_3(k)$, $x_4(k)$, $x_5(k)$, $x_6(k)$, and $u$ denote the species IL-6:IL-6R$\alpha$, gp130, (p)R$_{\text{complex}}$, (p)STAT3, SOCS3 mRNA, SOCS3 protein as well as the constant model stimulus, i. e. IL-6. Note, that the variables $x_7(k)$ (IL-6R$\alpha$), $x_8(k)$ (R$_{\text{complex}}$) and $x_9(k)$ (STAT3) can be derived from the following conserved moeties:

$$
\text{IL}{-}6\text{R}\alpha^{\text{Total}} = \text{IL}{-}6\text{R}\alpha + \text{IL}{-}6\text{:IL}{-}6\text{R}\alpha + 2\text{R}_{\text{complex}} + 2(\text{p})\text{R}_{\text{complex}} \tag{7.4a}
$$

$$
\text{gp130}^{\text{Total}} = \text{gp130} + 2\text{R}_{\text{complex}} + 2(\text{p})\text{R}_{\text{complex}} \tag{7.4b}
$$

$$
\text{STAT3}^{\text{Total}} = \text{STAT3} + (\text{p})\text{STAT3}. \tag{7.4c}
$$

Negative feedback inhibition through SOCS3 is modeled by a rational term (cf. (7.3a)).

In (7.3c), the variable $G_{\text{int}}$ denotes cell growth obtained by calculating the integral of (p)STAT3 over the considered time horizon by applying the trapezoidal method. $G_{\text{max}}$ describes cell growth obtained by determining the maximum peak height of (p)STAT3.

On the short-term time scale (cf. (7.3a)) the parameters are denoted by $p_i$, $i = \{1, ..., 13\}$. Additionally, we assume $x_4(k)$ as model output (i. e. $y^s(k)$, cf. (7.3b)). On the long-term time scale, the parameters are denoted by $\alpha$ and $\beta$ (cf. (7.3c)), and $G_{\text{int}}$ as well as $G_{\text{max}}$ are the corresponding long-term model outputs (i. e. $y^l(k)$, cf. (7.3d)).

**Measurement data**: Stimulation of HepG2 cells with IL-6 induces no cell growth. Therefore, Ba/F3 (murine pro B) cells were used instead to obtain measurement data on the short- and long-term time scale. Ba/F3 cells were cultivated and stimulated with

four different IL-6 concentrations, i. e. IL-6={0.004, 0.04, 0.2, 0.4} nM over a time horizon of 45 minutes. The dynamics for STAT3 phosphorylation were quantified for the time points $t_{\text{meas}} = \{0, 15, 30, 45\}$ min (cf. Fig. 7.2(a), daten taken from [59]). According to Fig. 7.2(a), STAT3 becomes phosphorylated in an IL-6 concentration-dependent manner. After 30 minutes, the amount of STAT3 phosphorylation decreases due to the negative SOCS3 feedback inhibition and dephosphorylation of phosphorylated STAT3 by phosphatases. We note, that the depicted data on STAT3 phosphorylation was normalized to the maximum phosphorylation, i. e. at $t$=15 min and IL-6=0.4 nM.



(a)

(b)

(c)
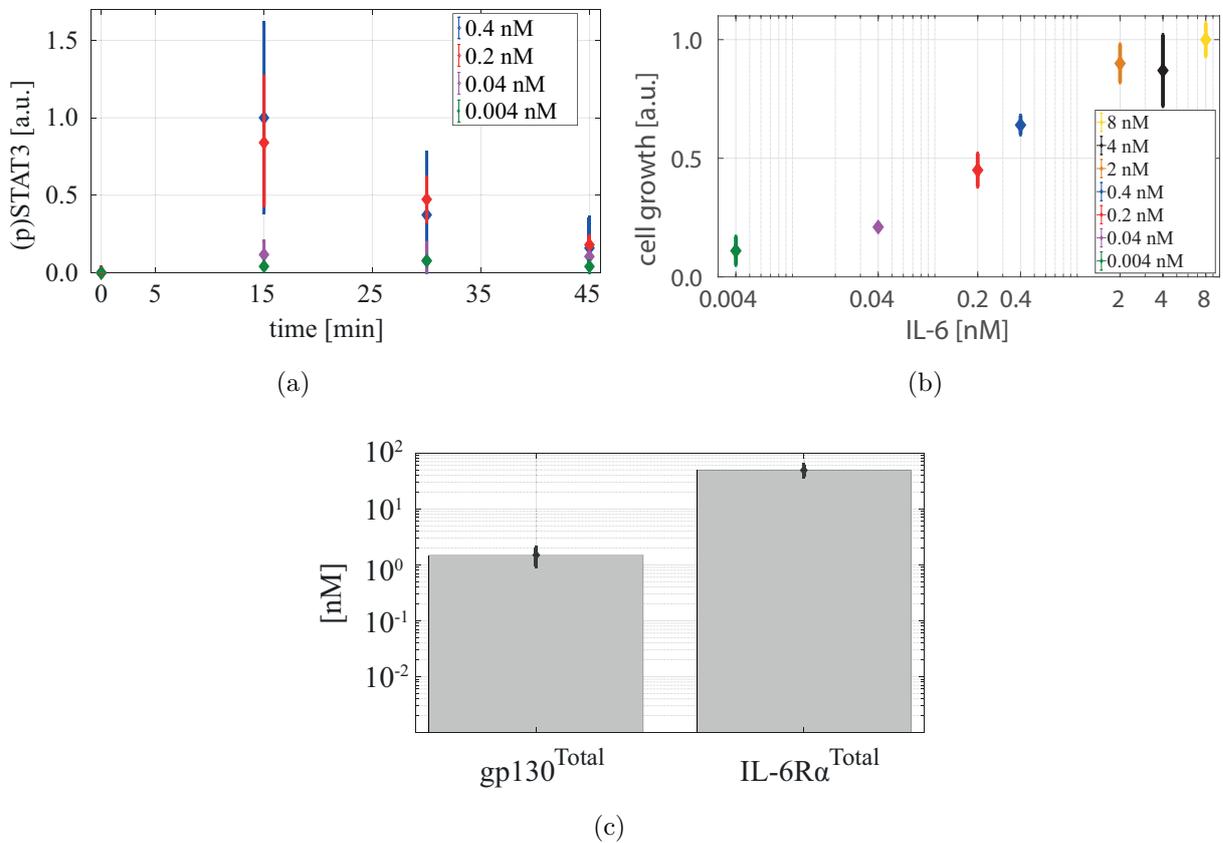
Figure 7.2: **Measurement data for IL-6-induced signaling in Ba/F3-gp130-IL-6R$\alpha$ cells.** (a) The cells were stimulated with four different concentration of IL-6 and STAT3 phosphorylation was measured for the time points 0, 15, 30 and 45 minutes (data taken from [59]). Diamonds correspond to the mean values and bars denote the determined standard deviations for n=2 (0.004 nM), n=3 (0.04 nM), n=3 (0.4 nM) and n=6 (0.4 nM) independent replicates. (b) The cells were stimulated with seven different IL-6 concentrations and cell growth was measured after 48 hours (data taken from [100]). Diamonds correspond to the mean values and bars denote the determined standard deviations for n=3 independent replicates, respectively. (c) Absolute quantification for the expression of gp130 and IL-6R$\alpha$. Bars denote the determined standard deviations for n=4 independent replicates, respectively.

To obtain data for cell growth on the long-term time scale, the cells were stimulated with seven different IL-6 concentrations, i. e. IL-6={0.004, 0.04, 0.2, 0.4, 2, 4, 8} nM and cultivated for 48 hours. Cell growth was quantified using the CellTiter-Blue Cell Viability Assay reagent as described in Appendix A (cf. Fig. 7.2(b), data taken from [100]). According to Fig. 7.2(b), the relative amount of cell growth (normalized to maximal cell growth, i. e. at IL-6=8 nM) increases with increasing IL-6 concentrations.

In addition to the above data, the total concentrations for the receptors gp130 and IL-6R$\alpha$ were quantified (cf. Appendix A for experimental methods) and amounted to 1.2$\pm$0.7 nM (378$\pm$224 molecules per cell, assuming a cell volume of $V_{cell}$=0.5$\times$10$^{-12}$ L) and 49.2$\pm$12.04 nM (15568$\pm$3809 molecules per cell), respectively (Fig. 7.2(c)).

**Problem setup:** For the following set-based analyses, we used a sampling time of $\Delta t$=1 min, $p_i$=[0.01,1] for $i = \{1, ..., 13\}$ (equivalent to assumed half-times for biomolecular reactions within seconds to minutes), and $\alpha$, $\beta$=[10$^{-4}$,10$^{4}$]. Furthermore, the intitial conditions for the variables form the short-term time scale are fixed to $x_0$=(0,gp130$^{\text{Total}}$,0,0,0,0)$^{\text{T}}$.

Note, to decode (p)STAT3 into cell growth using the approach in Eq. (7.1), we use the Matlab-based function *trapz*. The function *trapz* computes the approximate integral of the corresponding upper and lower bounds on $x_4(k)$ ((p)STAT3) for $k_1$=0 and $k_{\text{end}}$=45 via the trapezoidal method.

**Results and discussion:** The short- and long-term experimental results for concentrations of IL-6={0.004, 0.04, 0.2, 0.4} nM (green, purple, red and blue bars in Fig. 7.2(a) and (b)) were used to calibrate the model. In particular, we aimed to estimate tighter ranges for the model parameters $p_i$ as well as $\alpha$ and $\beta$, such that the the model is capable to represent the available measurement data on both time scales. The remaining long-term data for the concentrations of IL-6={2, 4, 8} nM (brown, black and yellow bars in Fig. 7.2(a) and (b)) were subsequently used for validation of the obtained results.

The results for the prediction of short-term STAT3 phosphorylation are depicted in Fig. 7.3(a). We first estimated the short-term parameters $p_i$ using the outer-bounding algorithm as presented in Chapter 3. Then, we performed Monte Carlo sampling to obtain valid parametrizations within the outer boundaries, such that the model is capable to represent the available short-term data for STAT3 phosphorylation. We determined 200 valid parametrizations leading to the depicted colored corridors in Fig. 7.3(a). As can be seen, the setup model together with the assumed parameter ranges is capable to represent IL-6-induced short-term STAT3 phosphorylation for the different IL-6 concentrations.

We next computed the integrals as well as the maximum peak heights for those trajectories that describe the outer margin of the colored corridors in Fig. 7.3(a) aiming to derive an outer approximation of the long-term parameters $\alpha$ and $\beta$. As a result that
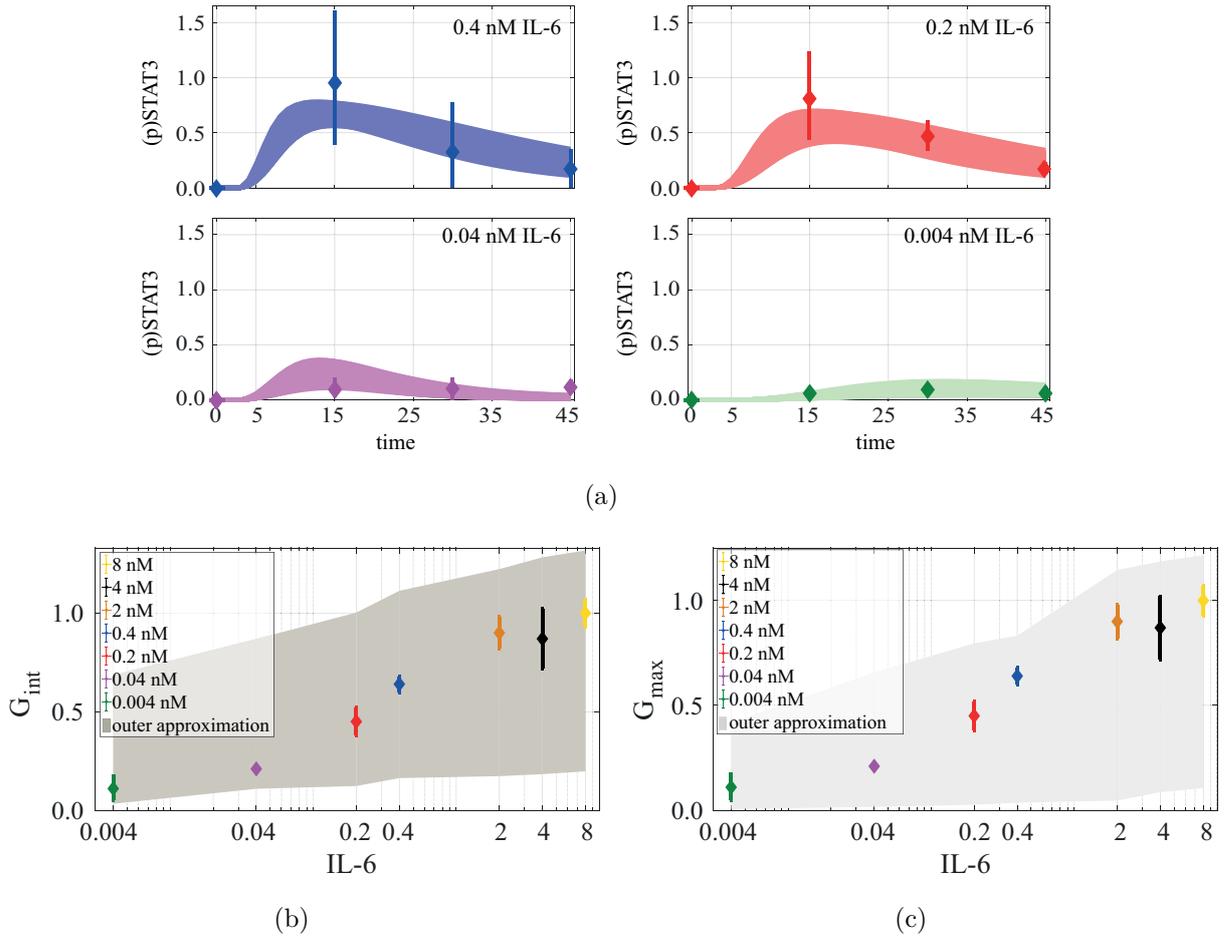
(a)



(b)



(c)

Figure 7.3: **Results for decoding short-term STAT3 phosphorylation into long-term cell growth.** (a) Dynamic short-term STAT3 phosphorylation for four different IL-6 concentrations. Colored lines depict the obtained model trajectories for the corresponding IL-6 concentration by simulating the model with 200 determined parameter samples. Colored bars and diamonds denote the corresponding uncertain short-term measurement data. (b) IL-6-induced cell growth. The dark gray area depicts the obtained outer approximations for cell growth using the integral for (p)STAT3 as decoding parameter. Colored bars and diamonds denote uncertain long-term data for cell growth. (c) The light gray area depicts the obtained outer approximations for cell growth using the maximum peak height of (p)STAT3 as decoding parameter.

the long-term parameters were estimated to $\alpha=[0.01,0.31]$ and $\beta=[0.37,1.18]$ leading to an outer approximation of long-term cellular growth, depicted as dark and light gray corridors in Fig. 7.3(b) and (c), respectively. We were capable to represent the available long-term data for the corresponding concentrations of IL-6=$\{0.004, 0.04, 0.2, 0.4\}$ nM. While Fig. 7.3(b) depicts the obtained results for correlating the integral of the (p)STAT3 signal with cell growth $G_{int}$, Fig. 7.3(c) shows the results for correlating the maximum peak height of the (p)STAT3 signal with cell growth $G_{max}$.

At this point, we assumed to have determined a calibrated model. Next, we used the determined 200 parametrizations to predict STAT3 phosphorylation for the remaining

IL-6 concentrations, i. e. IL-6={2, 4, 8} nM (trajectories not shown). Based on these predictions, we calculated the corresponding integrals and maximum peak heights as before and aimed to predict experimentally determined cell growth. As can be seen in Fig. 7.3(b) and (c), the long-term data could be predicted for both approaches very well, i. e. the data lie within the determined outer approximations (dark and light gray corridors Fig. 7.3(b) and (c)).

### 7.2.3 Summary

In the previous sections, we presented how biological processes from the fast, short-term time scale can be combined with cellular responses on the slow, long-term time scale. Typically, biological processes from signaling over gene expression to cell growth are very complex and cannot be modeled in detail. The combination of short- and long-term processes reduces these complexities by describing the main phenomena with only few equations and parameters. In particular, we here employed approaches for correlating specific shape properties of the kinetics of IL-6-induced STAT3 phosphorylation with cell growth considering processes such as transcription, translation, and growth as a black box. We showed that both approaches, i. e. a linear dependency of the maximum peak height of STAT3 phosphorylation as well as a linear dependency of the integral of the STAT3 phosphorylation signal with cell growth can be used to combine short- and long-term time scales within the set-based estimation approach.

We note, that the adapted approaches for combining short-term STAT3 phosphorylation with long-term cell growth are very specific approaches which need additional experimental validation. Nevertheless, we showed that it is possible to combine different time scales within the set-based framework avoiding a disproportional increase in the model size and the number of unknown parameters to be estimated. Prospectively, the combination of short- and long-term time scales deepens our understanding of how cells integrate signals, such as IL-6-induced STAT3 phosphorylation into distinct (patho-)physiological cell decisions, such as IL-6-induced cell growth.

An improved knowledge about the development of pathophysiological processes lay the foundation for the establishment of new medical treatments concerning, e. g. IL-6-induced misbalanced signaling and related inflammatory diseases. The selection of appropriate medical treatments concerning misbalanced IL-6-induced signaling is not trivial and its success depends on individual, patient-specific characteristics (such as genetic background). Therefore, so-called stratification approaches are often applied by classifying patients into different subcategories aiming to design the optimal patient's treatment.

In the following, we present a new and unified framework that works towards the stratification of patients for individual medical treatments.

# 7.3 Model-supported patient stratification using set-based information

Stratification describes the identification of groups with common characteristics, such as age, gender, social, ethnical or medical background [71]. In medicine, stratification of patients aims to assess potential risk for which certain diseases propagate or develop. To this end, risk factors are collected and associated, for instance, with clinical disease-specific symptoms [45, 78, 115]. Based on this stratification, physicians can decide upon personalized intervention strategies in an optimal manner [171].

In the following, we approach the task of patient stratification by combining set-based estimation methods for short-term molecular pathways with classification approaches of long-term disease development. In particular, we propose an unified framework that allows for the stratification of patients into subcategories of high, medium and low risk levels for developing inflammatory diseases. The framework is demonstrated by means of IL-6-induced Jak/STAT3 and MAPK trans-signaling and corresponding long-term responses.

## 7.3.1 Main idea and algorithms

Our approach combines the set-based estimation framework presented in [136] for data on the fast, short-term time scale with classification methods for patient stratification using data on the slow, long-term time scale as detailed in Figure 7.4. In a first step, we assume the availability of short-term data obtained from biochemical surrogates of patients (such as cells extracted from biopsies or blood serum). These surrogates are probed by biochemical stimulation (e. g. excitation of a certain pathway) and point us to changes in proteins or genes indicative of disease status within seconds to minutes (upper, left-hand box). We apply the set-based estimation framework to obtain parameter sets that can describe the biochemical data (lower, left-hand box). These parameter sets are subsequently transformed (piped) through a classification algorithm (e. g. an artificial neural network, support vector machines, etc.), resulting in a set of transformed parameters (in our example: weights and thresholds in a neural network; lower blue box) to match the long-term response. Long-term responses can be seen as the physiolgcial patient outcome (e. g. cell growth) and may describe the disease status over days (upper, right-hand box). In the second step of the framework, we aim to predict the long-term outcome based on model parameters determined in the first step. The physiological patient outcome can be either experimentally determined or may be obtained by integrating the short-term patient outcome as illustrated in the previous sections. To predict the long-term patient outcome, we arrange the sequence of long-term data into ordered tupels for each patient. The thus obtained tupels are associated to classes of long-term patient responses and serve as output categories for the classification approach. If the long-term data only consists
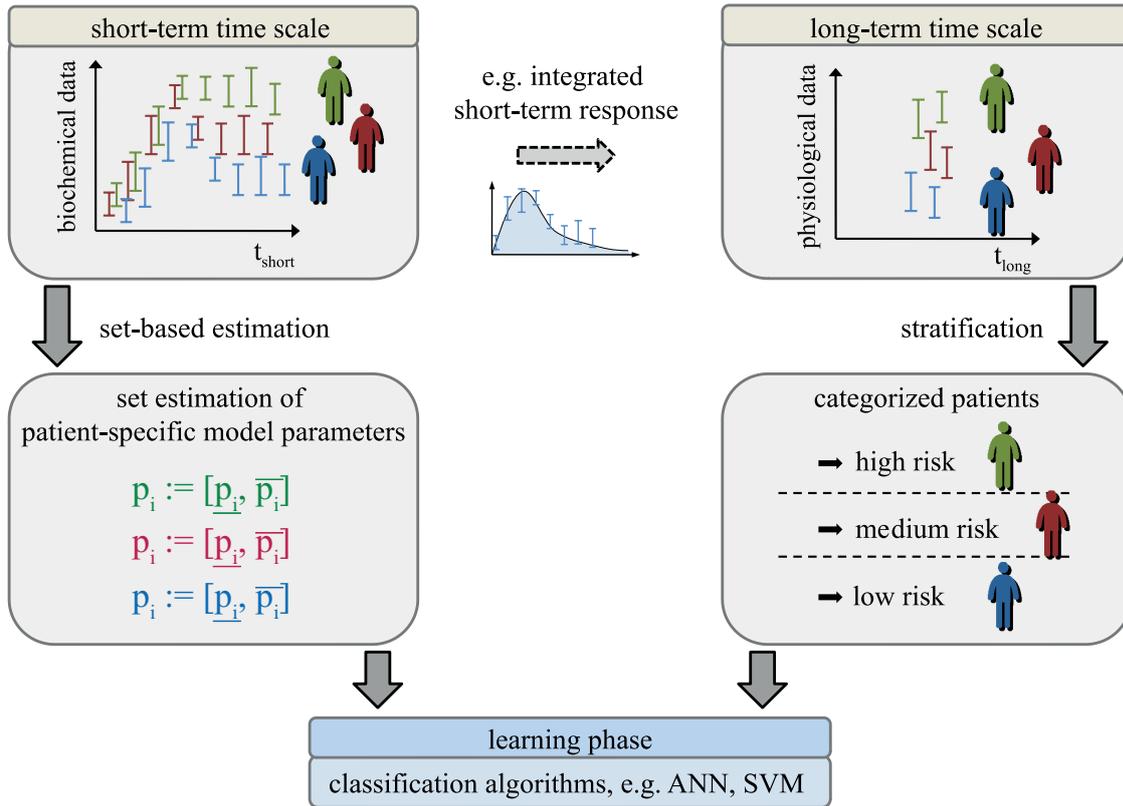
Figure 7.4: **Proposed workflow for combining uncertain biochemical processes on the short-term time scale with uncertain physiological responses on the long-term time scale and patient stratification.** The short-term time scale data often describe the patient diagnosis over longer time and may be seen as a means for patient strata for disease development (e. g. low, medium and high risk). For the stratification of patient cohorts into different subcategories, algorithms such as Artificial Neural Networks (ANN) or Support Vector Machines (SVM) can be used.

of a single endpoint (as in our example), its value are directly associated with a group such as providing a stratification for low, medium or high risk based on predefined thresholds for the respective data (lower, right-hand box). For stratification, the patients are split into groups for training, validation and test which is in line with standard procedures of classification approaches [50]. The analysis is performed for each patient of each group separately.

The presented workflow can be systematically formalized in Algorithms 1 and 2. The first algorithm is used to obtain a trained classifier based on the short-term patient data and the corresponding strata thresholds for high, medium, and low risk. Once Algorithm 1 has produced a classifier that performs satisfyingly, the obtained result is applied in a second step to stratify new patients according their long-term profiles using Algorithm 2.

---

**Algorithm 1** Set-based classifier training

**Input:**
    Short- and long-term time scale data for each patient;
    A dynamical model representing the data on the short-term time scale;
    Threshold values defining each risk group;
    An untrained classifier;

**Output:**
    A trained classifier;

1: Perform set-based parameter estimation using the short-term time scale data and the dynamical model
2: Process the data for each time instance on the long-term time scale into the risk subcategories
3: Stratify the patients into the corresponding risk subcategories
4: Arrange the corresponding parameter bounds of each patient to the corresponding risk subcategory for each long-term time instance
5: Split patients into groups for training, validation and test
6: Train the chosen classifier
7: Verify the quality of the classifier and if needed adjust and re-train the classifier

---

**Algorithm 2** Patient stratification

**Input:**
    Short-term time scale data for each patient;
    The same dynamical model representing the short-term time scale data as in Algorithm 1;
    The trained classifier from Algorithm 1;

**Output:**
    A prediction of the patient-specific risk category;

1: Perform set-based parameter estimation using the short-term time scale data and the dynamical model
2: Input the set-based parameter estimation results into the trained classifier
3: Obtain the stratification results

---

### 7.3.2 Example

To demonstrate applicability of the presented framework and algorithms, we use simulated measurements for the short- and long-term time scale for a group of 50 patients. Simulated short-term data consists of a biochemical data set, assumed to be provided by biochemically testing the response of extracted patient tissue to a certain stimulation. In particular, we assume to avail of protein changes of patient

tissue associated with two important signaling pathways, IL-6-induced Jak/STAT3 and MAPK trans-signaling. As long-term response, we assume the availability of a single endpoint at a later time point for each patient, which we associate to low, medium and high risk for developing an IL-6-induced inflammatory disease.

**Model for Jak/STAT3 trans-signaling:** The reaction mechanisms for the Jak/STAT3 pathway are described as:

$$
\begin{aligned}
x_1(k+1) &= x_1(k) + \Delta t\big(p_1 x_7(k)u - p_2 x_1(k) \\
&\quad - 2p_3 x_2(k)^2 x_1(k)^2 + 2p_4 x_8(k)\big) \\
x_2(k+1) &= x_2(k) + \Delta t\big(2p_4 x_8(k) - 2p_3 x_2(k)^2 x_1(k)^2\big) \\
x_3(k+1) &= x_2(k) + \Delta t\Big(\frac{p_5 x_8(k)}{1 + p_{13} x_6(k)} - p_6 x_3(k)\Big) \\
x_4(k+1) &= x_4(k) + \Delta t\big(p_7 x_3(k)x_9(k) - p_8 x_4(k)\big) \\
x_5(k+1) &= x_5(k) + \Delta t\big(p_9 x_4(k) - p_{10} x_5(k)\big) \\
x_6(k+1) &= x_6(k) + \Delta t\big(p_{11} x_5(k) - p_{12} x_6(k)\big).
\end{aligned}
\tag{7.5}
$$

Thereby, the variables $x_1(k)$, $x_2(k)$, $x_3(k)$, $x_4(k)$, $x_5(k)$, $x_6(k)$ and $u$ denote IL-6:sIL-6R$\alpha$, gp130, (p)R$_{\text{complex}}$, (p)STAT3, SOCS3 mRNA, SOCS3, and IL-6. Furthermore $x_7(k)$, $x_8(k)$, $x_9(k)$ describe the entities sIL-6R$\alpha$, R$_{\text{complex}}$ and STAT3, respectively which can be extracted from the following conservation laws:

$$
\begin{aligned}
\text{sIL}-6\text{R}\alpha^{\text{Total}} &= \text{sIL}-6\text{R}\alpha + \text{IL}-6\text{:sIL}-6\text{R}\alpha + 2\text{R}_{\text{complex}} + 2\text{pR}_{\text{complex}} \\
\text{gp130}^{\text{Total}} &= \text{gp130} + 2\text{R}_{\text{complex}} + 2\text{pR}_{\text{complex}} \\
\text{STAT3}^{\text{Total}} &= \text{STAT3} + \text{pSTAT3}.
\end{aligned}
\tag{7.6}
$$

Notably, for implementing Model (7.5) we assume, apart from the mentioned assumptions in Chapter 5 that STAT3 activation represents both, phosphorylation and dimerization of STAT3 proteins.

**Model for MAPK trans-signaling:** The reaction mechanisms for the MAPK pathway are described as:

$$x_1(k+1) = x_1(k) + \Delta t \big( p_1 x_7(k) u - p_2 x_1(k)$$
$$- 2 p_3 x_2(k)^2 x_1(k)^2 + 2 p_4 x_8(k) \big)$$
$$x_2(k+1) = x_2(k) + \Delta t \big( 2 p_4 x_8(k) - 2 p_3 x_2(k)^2 x_1(k)^2 \big)$$
$$x_3(k+1) = x_2(k) + \Delta t \big( p_5 x_8(k) - p_6 x_3(k) \big)$$
$$x_4(k+1) = x_4(k) + \Delta t \big( p_7 x_3(k) x_8(k) - p_8 x_4(k) \big) \quad (7.7)$$
$$x_5(k+1) = x_5(k) + \Delta t \big( p_9 x_4(k) x_9(k) - p_{10} x_5(k) \big),$$
$$x_6(k+1) = x_6(k) + \Delta t \big( p_{11} x_5(k) x_{10}(k) - p_{12} x_6(k) \big)$$
$$x_7(k+1) = x_7(k) + \Delta t \big( p_{13} x_6(k) x_{11}(k) - p_{14} x_7(k) \big).$$

In (7.7), the variables $x_1(k)$, $x_2(k)$ and $x_3(k)$ are similar to (7.5), whereby the same conserved moieties hold for sIL-6R$\alpha$ and R$_{\text{complex}}$. Moreover, the variables $x_4(k)$, $x_5(k)$, $x_6(k)$, and $x_7(k)$ denote Ras$^*$, Raf$^*$, Mek$^*$ and ERK$^*$, respectively. The inactive forms Ras, Raf, Mek and ERK denoted as $x_8(k)$, $x_9(k)$, $x_{10}(k)$, and $x_{11}(k)$ can be extracted from the conservation laws:

$$\text{Ras}^{\text{Total}} = \text{Ras} + \text{Ras}^*$$
$$\text{Raf}^{\text{Total}} = \text{Raf} + \text{Raf}^*$$
$$\text{Mek}^{\text{Total}} = \text{Mek} + \text{Mek}^* \quad (7.8)$$
$$\text{ERK}^{\text{Total}} = \text{ERK} + \text{ERK}^*.$$

We note that Grb2 was not explicitly modelled but considered as an integral part of the phosphorylated receptor.

**Simulated short- and long-term patient data:** We assume a cohort of 50 patients and simulate measurement data for the proteins (p)STAT3 and ERK$^*$, acting as upstream surrogates for inflammatory diseases. For generating data on (p)STAT3, the input IL-6 was fixed to 0.2 and STAT3$^{\text{Total}}$ was set to 10. Furthermore, we fixed the kinetic parameters $p_i^{\text{Jak/STAT3}}$ with $i := \{1, \dots, 13\}$ to the nominal values

$$p_i^{\text{Jak/STAT3}} = (0.075, 0.056, 0.01, 0.00015, 0.25, 0.09, 1.5, 0.01, 0.1, 0.1, 1, 0.1, 5)^{\text{T}}$$

and the initial conditions were set to

$$x^{\text{Jak/STAT3}}(0) = (0, \text{gp130}^{\text{Total}}, 0, 0, 0, 0, 0).$$

Values for the total concentrations of gp130$^{\text{Total}}$ and IL-6R$\alpha^{\text{Total}}$ were randomly generated within bounds of gp130$^{\text{Total}}$=[1,5] and sIL-6R$\alpha^{\text{Total}}$=[0.5,2] obtaining 50 patient-specific profiles on (p)STAT3.

For data generation on ERK$^*$, we fixed IL-6 to 0.2, gp130$^{\text{Total}}$ to 5 and IL-6R$\alpha^{\text{Total}}$

to 2. The kinetic parameters $p_i^{\text{MAPK}}$ with $i := \{1, \ldots, 14\}$ were set to the nominal values

$$p_i^{\text{MAPK}} = (0.075, 0.056, 0.01, 0.00015, 0.25, 0.09, 0.1, 0.1, 0.1, 0.1, 0.1, 0.1, 0.1, 0.1)^{\text{T}}$$

and the initial conditions were set to

$$x^{\text{MAPK}}(0) = (0, \text{gp130}^{\text{Total}}, 0, 0, 0, 0, 0).$$

Values for the total concentrations of $\text{Ras}^{\text{Total}}$, $\text{Raf}^{\text{Total}}$, $\text{Mek}^{\text{Total}}$, and $\text{ERK}^{\text{Total}}$ were randomly generated within the bounds [1,10] obtaining 50 patient-specific profiles on $\text{ERK}^*$.

As long-term patient outcome, we assumed the integrated response of the activation of both pathways as a disease surrogate. We therefore first calculated the integrals of the (p)STAT3 and $\text{ERK}^*$ trajectories for each patient for a time horizon of 60 minutes. Then, both integrals were multiplied to assess a collective effect for inflammation on the long-term time scale. These results were then grouped over the patient cohort to classify them into patients with values lower than 25%-quantile "low risk patients", such between 25%- and 75%-quantiles "medium risk patients" and such with higher than the 75%-quantile "high risk patients".

Note, that the simulated short- ((p)STAT3 and $\text{ERK}^*$) and long-term time scale data (corresponding integrals) together with the patient stratification that is used for classification are shown exemplary for 10 patients in Fig. 7.5.

**Set-based problem setup:** The set-based approach was used to remodel the short-term time scale data under the given initial conditions for each pathway in order to render a set of outer-bounded parameters for each patient and pathway. For solving the feasibility problem of the dynamical system, the time discretization $\Delta t$ was set to 2 minutes for a time horizon of 60 minutes. To account for experimental uncertainties, errors of $\pm 10\%$ were added to the parameter vectors $p_i^{\text{Jak/STAT3}}$ and $p_i^{\text{MAPK}}$ as well as to the simulated short-term data for (p)STAT3 and $\text{ERK}^*$. Based on this generated short-term data, the patient-specific parameters $\text{gp130}^{\text{Total}}$ and $\text{sIL-6R}\alpha^{\text{Total}}$ in the Jak/STAT3 pathway and $\text{Ras}^{\text{Total}}$, $\text{Raf}^{\text{Total}}$, $\text{Mek}^{\text{Total}}$, and $\text{ERK}^{\text{Total}}$ in the MAPK pathway were estimated for each patient using the outer-bounding approach as described in Chapter 4.

**Structure of the network and setup of the classifier:** For the classification part of our method, several approaches can be used (cf. [50, 97]). To demonstrate the proposed framework, we opted to use Artificial Neural Networks (ANN), see e. g. [50, 121] and Fig. 7.6 for a schematic representation. Yet, also Support Vector Machines [2] or boosting methods [141] can be used. An ANN is based on a collection of nodes (colored circles Fig. 7.6). The different neurons are connected to one another
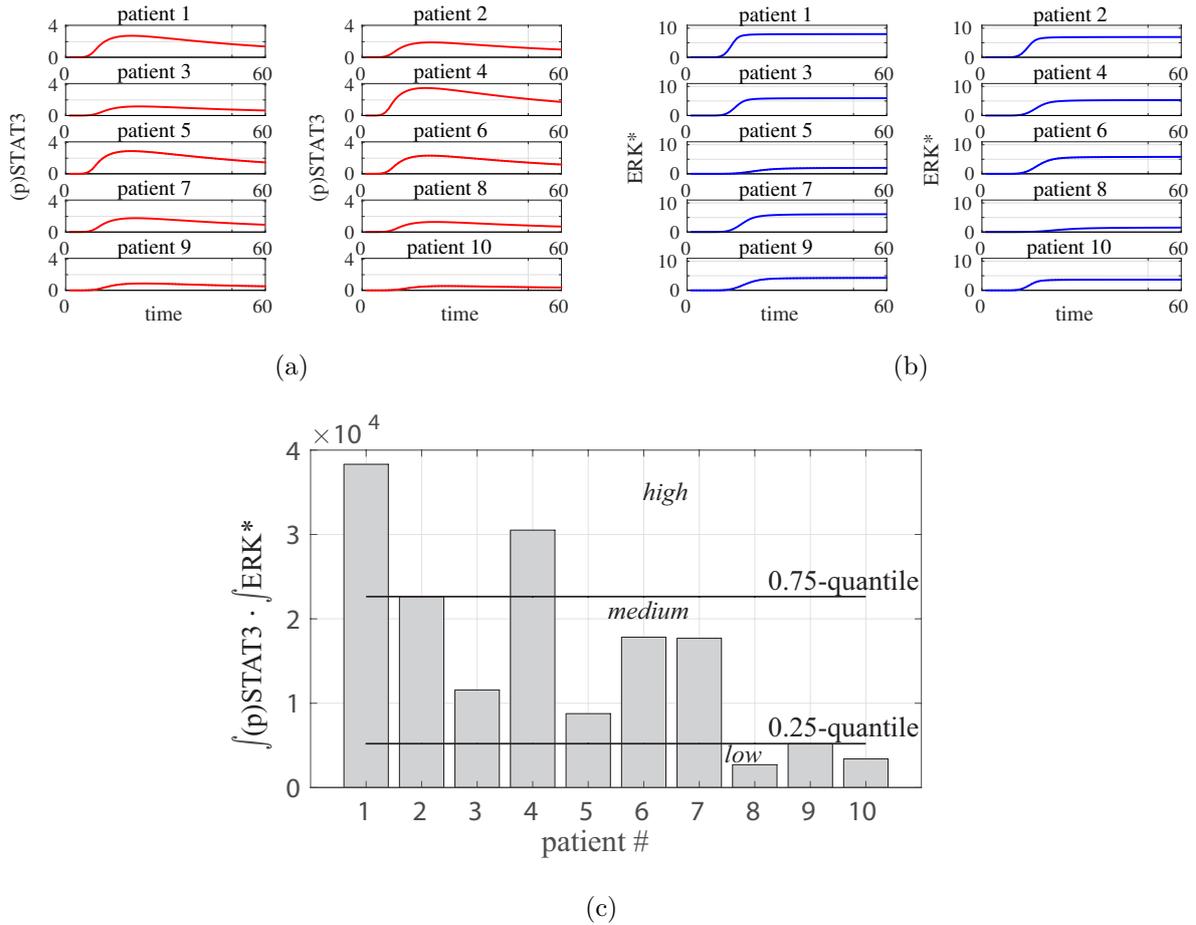
Figure 7.5: **Simulated patient data exemplarily depicted for 10 patients.** (a) Dynamic (p)STAT3 profiles, red trajectories. (b) Dynamic ERK* profiles, blue trajectories. (c) Integrated response of (p)STAT3 and ERK* as a disease surrogate. Horizontal black lines denote the thresholds of the patient cohort grouped into the corresponding risk categories.

(simulating a simplified version of a synapse, black arrows Fig. 7.6) allowing the neurons to transmit a signal from one to another. Notbaly, during the transmission, the signal is processed (strengthened, weakened). In ANN implementations, the signal at a connection between neurons is a real number, and the output of each neuron is calculated by a nonlinear function of the sum of its inputs. Neurons and connections typically have a weight that adjusts as learning proceeds. The weight increases or decreases the strength of the signal at a connection. Neurons may have a threshold such that the signal is sent only if the aggregate signal crosses that particular threshold. Typically, neurons are organized in layers that may perform different kinds of transformations on their inputs. Signals travel from the input (first) layer via one or several hidden layers to the output (last) layer, possibly after traversing the layers multiple times.

In our example, the chosen ANN consisted of one hidden layer with 10 artificial neurons, to which the patient-specific set-based input information are fed (Fig. 7.6
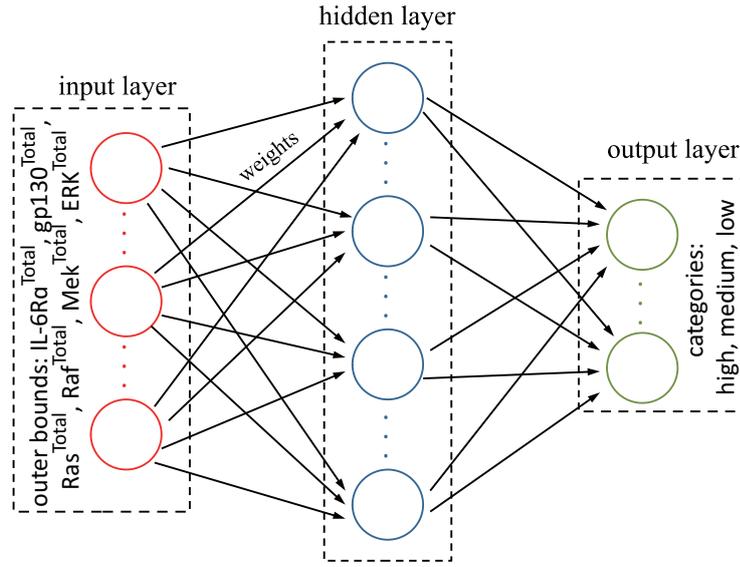
Figure 7.6: **Schematic representation of an Artificial Neural Network as used in this work.** Artificial Neural Networks consist of an input, hidden and output layer, in which neurons (colorored circles) are connected to one another (arrows). Patient-specific parameters (e. g. IL-6R$\alpha^{\text{Total}}$) are used as inputs to the network, while the output layer describes three different categories.

input layer, red circles). Furthermore, the ANN consisted of one output layer, which provides the stratification results (Fig. 7.6 output layer, green circles). All neurons in the hidden layer have sigmoid activation functions, in contrast to the output layer having softmax activation functions [159]. The use of softmax neurons normalizes the outcome such that all outcomes add to 1, and hence the patient category as outcome can be interpreted as probability function [67].

### 7.3.3 Results and discussion

First, we estimated the corresponding parameters (total receptor and kinase amounts) for each patient and pathway using the set-based approach. In Table 7.1 the outer-bounding results for 4 simulated patients (cf. Fig. 7.5, patients 1-4) are presented. As can be seen, the parameters could be approximated to tight ranges with respect to the initial chosen uncertainty ranges.

The estimated outer bounds for the considered parameters serve subsequently as inputs to the neural network. The inputs are vectors of 4 inputs per patient for the Jak/STAT3 pathway comprising the upper and lower boundary value (reflecting the uncertainty of the data) of the parameters sIL-6R$\alpha^{\text{Total}}$ and gp130$^{\text{Total}}$. In addition, for the MAPK pathway we have vectors of 8 inputs per patient comprising again the upper and lower boundary value of the parameters Ras$^{\text{Total}}$, Raf$^{\text{Total}}$, Mek$^{\text{Total}}$ and ERK$^{\text{Total}}$, respectively. Thus, in total we have a vector of 12 inputs per patient to the

network. We note, that all guarantees obtained by set estimation of the patient-specific parameters are lost when piping through the ANN.

Table 7.1: Exemplary outer-bounding results for 4 patients for the Jak/STAT3 (cf. Fig. 7.5(a)) and the MAPK signaling pathway (cf. Fig. 7.5(b)).

| patient # | Jak/STAT3 pathway | | MAPK pathway | | | |
| --- | --- | --- | --- | --- | --- | --- |
| | sIL-6R$\alpha^{\text{Total}}$ | gp130$^{\text{Total}}$ | Ras$^{\text{Total}}$ | Raf$^{\text{Total}}$ | Mek$^{\text{Total}}$ | ERK$^{\text{Total}}$ |
| 1 | [1.2,1.8] | [3.7,4.9] | [7.6,9.3] | [6.7,9.2] | [7.2,9.8] | [8.3,9.6] |
| 2 | [0.9,1.6] | [3.2,4.3] | [6.8,9.2] | [6.0,9.0] | [6.4,9.6] | [7.4,8.5] |
| 3 | [0.7,1.1] | [2.7,3.6] | [5.9,8.0] | [5.2,8.7] | [5.6,8.4] | [6.5,7.4] |
| 4 | [1.4,2.0] | [4.1,5.0] | [3.9,5.3] | [3.2,5.4] | [5.7,8.5] | [5.9,6.8] |

For stratification, the simulated patients were splitted into a training, validation, and test group according a split of 60%, 20% and 20%, respectively. The resulting confusion matrices are presented in Fig. 7.7. The results demonstrate a 83.3% correct classification of the training set (Fig. 7.7(a)), 90% of the cross-validation set (Fig. 7.7(b)), 80% of the test set (Fig. 7.7(c)) and an overall correctness of 84% (Fig. 7.7(d)).

To ensure reproducibility of the results different runs were carried out. The integrated random algorithm that was used for choosing which patient falls in which group (i. e. for training, validation and testing, repectively), demonstrated that the results could only improve. Furthermore, the different runs showed that the results are independent of the grouping choice. Also an increase in the size of the hidden layer did not lead to an improvement of the stratification results.

## 7.4 Summary

We presented a modeling framework that allows for combining processes and data on the short- and long-term time scale under the umbrella of the set-based approach. The combination of both time scales together with the application of classification methods allowed in particular the model-guided stratification of patients into risk categories for high, medium or low risk for developing inflammatory diseases. The stratification of patients is based on dynamic processes and parameters on the short-term time scale, however, does not demand a deeper knowledge about processes on the long-term time scale. Thus, an advantage of the therein proposed framework is that only one model for the short-term time scale is needed. The presented framework can be seen as an extension of methods that provide a feasibility set for measurement data with inherent uncertainties. The sets obtained for a fast, short-term and often pathway-based description of the disease process are transformed by piping it through a classification algorithm to provide a prediction of long-term time scale data. With
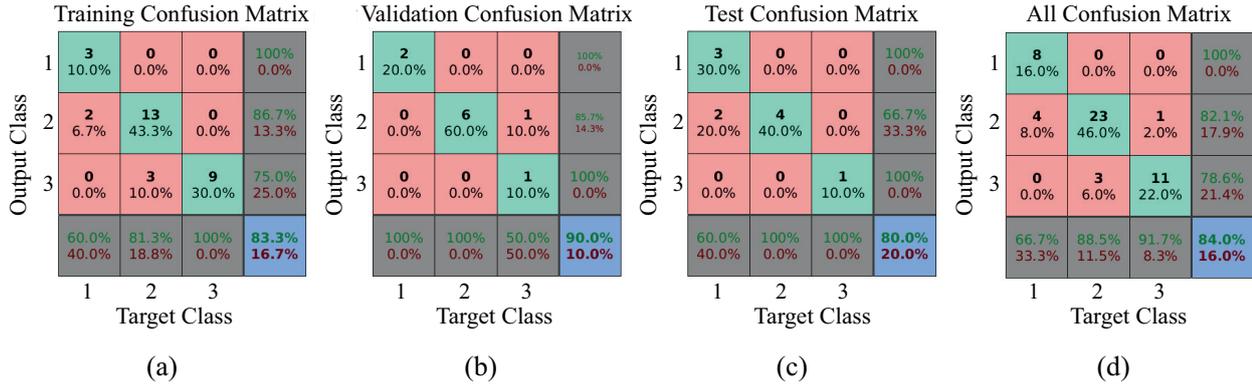
Figure 7.7: **Results for classification of a simulated patient cohort into sub-categories of risk levels for developing inflammatory diseases.** Results for (a) training, (b) validation, and (c) testing of the classifier over the three patient categories, i. e. 1: high, 2: medium, and 3: low risk. Red and green highlighted areas correspond to a wrong and correct classification, respectively. Dark gray highlighted areas show the classification results for the corresponding category 1, 2, and 3, while blue highlighted areas describe the overall result comprising all three categories. (d) Overall conclusion comprising the results from (a), (b) and (c).

this, we aimed to provide a shift in reasoning over feasibility sets to define them as a super-class for explaining data under uncertainty, while covering a process of detailed dynamical modeling and more abstract stratification approaches at the same time.

By stratifying patients into risk categories, physicians are able to select optimal treatments and intervention strategies tailored to the individual patient. Tailoring of medical treatment to the individual characteristics of each patient (i. e. precision medicine) has been a part of healthcare for many years. Model-guided stratification approaches can help to support physicians in their decisions reducing, for example, complications during medical treatments.

In the previous Chapters 4-7 we showed how the set-based estimation framework [136] can be used and extended for parameter estimation, experimental design and the combination of different time scales. We applied the approaches using IL-6-induced signaling for obtaining a deeper understanding about the underlying complex molecular processes and relating deregulations. In the next chapter, we consider both, IL-6-induced classic- and trans-signaling aiming to gain new insights into differences of the two pathways. To achieve our goals, we apply the set-based estimation method together with quantitative biochemcial analyses.

# 8 Fusing Experimental Insights and Dynamic Modeling: Response to IL-6 Trans- and Classic-signaling is Determined by the Ratio of the IL-6R$\alpha$ to gp130 Expression

In this chapter, we present a comprehensive application example for set-based model invalidation and parameter estimation considering IL-6-induced classic- and trans-signaling. Notably, for the first time the set-based approach is applied to a realistic-sized pathophysiological problem. The chapter is based on our work presented in [130]. We note, that measurement data used to study classic- and trans-signaling were generated specifically by M. Sc. Heike Reeh and M. Sc. Ulrike Billing (group of Prof. Fred Schaper and Dr. Anna Dittrich, Department of Systems Biology, Magdeburg). We focus on the modeling and set-based estimation, details on experimental methodologies can be found in Appendix A.

## 8.1 Introduction

IL-6 is, among many other processes, involved in the regulation of inflammatory responses. It is well studied that dysregulated IL-6-induced signaling is associated with the development of immunological and proliferative diseases, such as rheumatoid arthritis, inflammatory bowel disease and colon cancer [39, 42, 54, 77, 109]. Initiation of IL-6-signaling occurs through two different pathways, i. e. classic- and trans-signaling. While classic-signaling induces regenerative and protective responses, trans-signaling is related to pro-inflammatory responses [142] and is thus, associated with the development of inflammatory diseases.

In order to analyse trans-signaling independently of classic-signaling the fusion protein Hy-IL-6 was developed. In HepG2 cells, Hy-IL-6 induces maximal expression of acute-phase proteins at molar concentrations substantially lower than those needed for IL-6 [56]. In mouse models, injection of Hy-IL-6 results in a significantly stronger induction of acute-phase proteins than IL-6 [120]. These observations led to the conclusion that trans-signaling is a stronger activator of Jak/STAT3 signaling than classic-signaling. However, the molecular basis for this hypothesis has not been identified yet. In this chapter, we use set-based mathematical modeling as well as biochemical and cell biological analyses to study the differences between IL-6-induced classic- and trans-siganling. Notably, although several computer models of the JaK/STAT3

pathway have already been published ([23, 49, 128, 148, 162, 173]), they focus on classic-signaling. Differences between IL-6 classic- and trans-signaling have not yet been addressed and systematically modeled.

The basic question of interest is how close or different the signaling machinery of both pathways is. Our results show that the differences between IL-6-induced classic- and trans-signaling are only mediated by differences in the expression and molar ratios of receptor components whereas canonical intracellular signaling is indifferent in both pathways. Our computational study lays the basis for potential intervention strategies targeting IL-6-induced misbalanced signaling at the receptor level.

## 8.2 Measurement data

To study differences between IL-6-induced trans- and classic-signaling we make use of the specific trans-signaling inducer Hy-IL-6. We first verified the applicability of Hy-IL-6 to induce trans-signaling in lieu of IL-6 and sIL-6Rα. Using the law of mass action and considering the dissociation constant $K_D$=0.5 nM of the IL-6:IL-6Rα complex [177], we calculated how much IL-6:sIL-6Rα complex is formed for given amounts of IL-6 and sIL-6Rα. Next, we compared the strength of trans-signaling induced by either IL-6:sIL-6Rα complex or an equimolar amount of Hy-IL-6. Exemplary, HepG2 expressing both gp130 and membrane-bound IL-6Rα are stimulated with either 0.17 nM IL-6 + 100 nM sIL-6Rα, forming 0.17 nM IL-6:sIL-6Rα complex, or 0.17 nM Hy-IL-6. After 15 min of stimulation phosphorylation of STAT3 was analysed by intracellular flow cytometry in stimulated and unstimulated HepG2 cells as control. As result, no obvious difference in STAT3 activation in response to trans-signaling induced by Hy-IL-6 or by the IL-6:sIL-6Rα complex exists (Appendix B, Fig. B.1). Hence, Hy-IL-6 can be used as specific stimulus for trans-signaling.

To investigate whether the strength of IL-6-induced classic-signaling is different to trans-signaling we compared next the kinetics of STAT3 phosphorylation, as well as SOCS3 mRNA and SOCS3 protein expression in HepG2 cells stimulated with either 0.17 nM (Fig. 8.1(a)) or 0.08 nM (Fig. 8.1(b)) of IL-6 to induce classic-signaling or Hy-IL-6 to induce trans-signaling. As result, IL-6-induced classic- and trans-signaling result in transient phosphorylation of STAT3. However, trans-signaling-induced STAT3 activation (red) is more pronounced than classic-signaling-induced STAT3 phosphorylation (blue). SOCS3 mRNA and protein expression follow the peak of STAT3 phosphorylation. Both SOCS3 mRNA and protein induction are higher in response to trans-signaling than to classic-signaling. Data normalization was performed as described in Appendix B. In addition to the data presented, the number of total IL-6Rα and gp130 receptor proteins on the cell surface was determined by a bead-based FACS assay as 2.2 ± 0.3 nM (2099 ± 347 receptors/cell) and 16.8 ± 3.1 nM (16198 ± 2965 receptors/cell), respectively (Fig. 8.1(c)). The number of total STAT3 in HepG2 cells

was measured by quantitative Western blotting and amounted $958 \pm 445$ nM ($9.2 \cdot 10^5 \pm 4.2 \cdot 10^5$ molecules/cell).
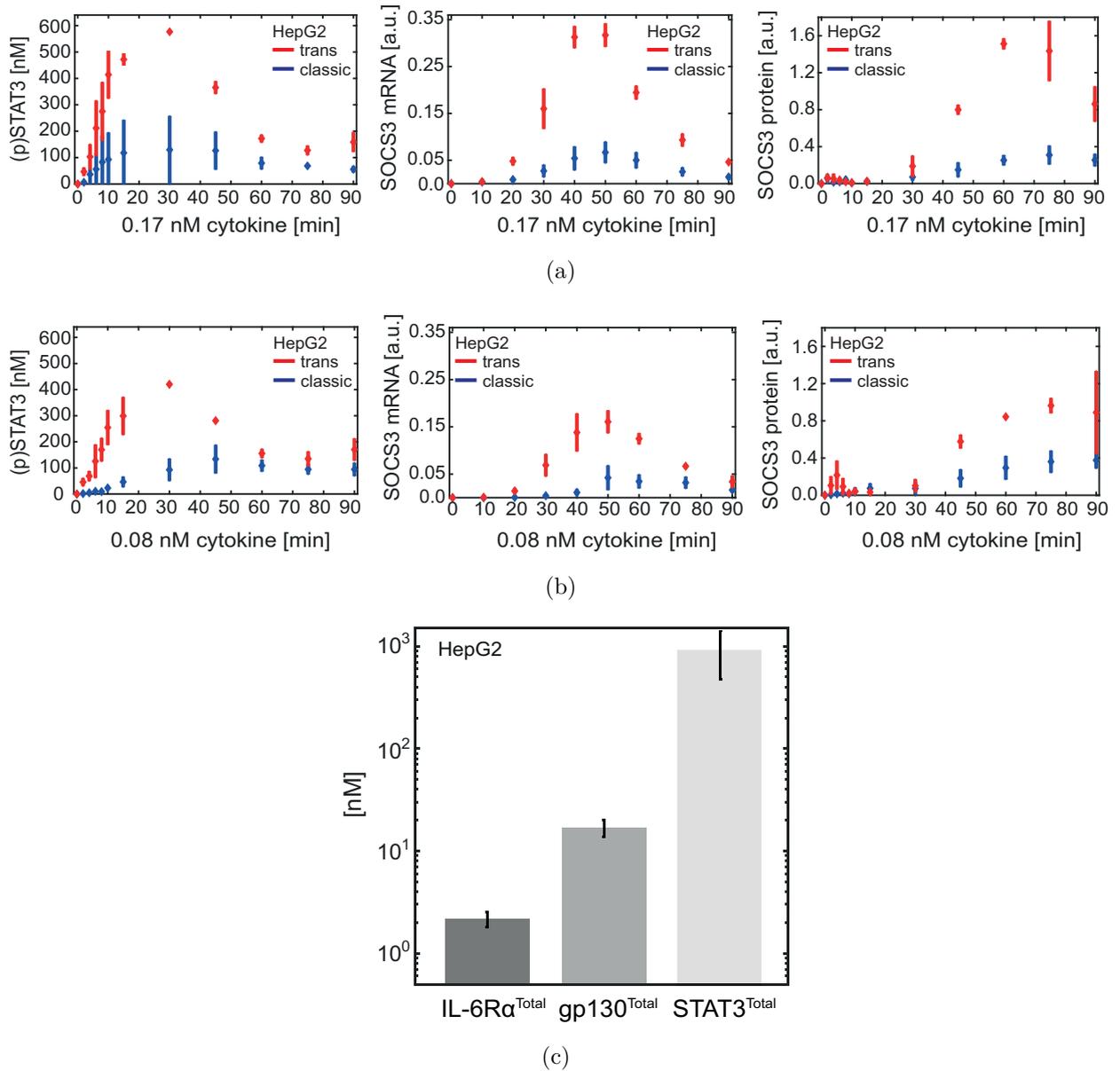


Figure 8.1: **Measurement data for IL-6-induced classic- and trans-signaling in HepG2 cells (blue and red, respectively).** (a) HepG2 cells were stimulated with 0.17 nM IL-6 (classic) and Hy-IL-6 (trans), respectively. STAT3 phosphorylation, SOCS3 mRNA and SOCS3 protein expression were analysed. Data are given as mean±STD from n=3-4 experiments. (b) HepG2 cells were stimulated with 0.08 nM IL-6 (classic) and Hy-IL-6 (trans), respectively. (c) The expression of gp130 and IL-6Rα of HepG2 cells was quantified by flow cytometry. Mean±STD from n=4 independent experiments is shown. Activation and expression of STAT3 in HepG2 cells were quantified using recombinant calibrator proteins. Mean±STD from n=7 independent experiments is shown.

## 8.3 Models for IL-6-induced classic- and trans-signaling

So far, it is unknown whether the observed differences between classic- and trans-signaling (Fig. 8.1(a) and (b)) are caused by different strength of receptor activation or by different signaling mechanisms as well as kinetics downstream to receptor activation. To test these alternative hypotheses we use set-based modeling. To this end, we consider three different models. While the first model describes a combination of classic- and trans-signaling (Fig. 8.2(a)), the second and third models describe classic- and trans-signaling by two separated sub-models (Fig. 8.2(b) and (c)). The detailed model assumptions can be found in Appendix C.

Note, we model classic-signaling by considering that IL-6 first binds to IL-6Rα followed by binding of the IL-6:IL-6Rα complex to gp130. In case of trans-signaling Hy-IL-6 associates directly with gp130.
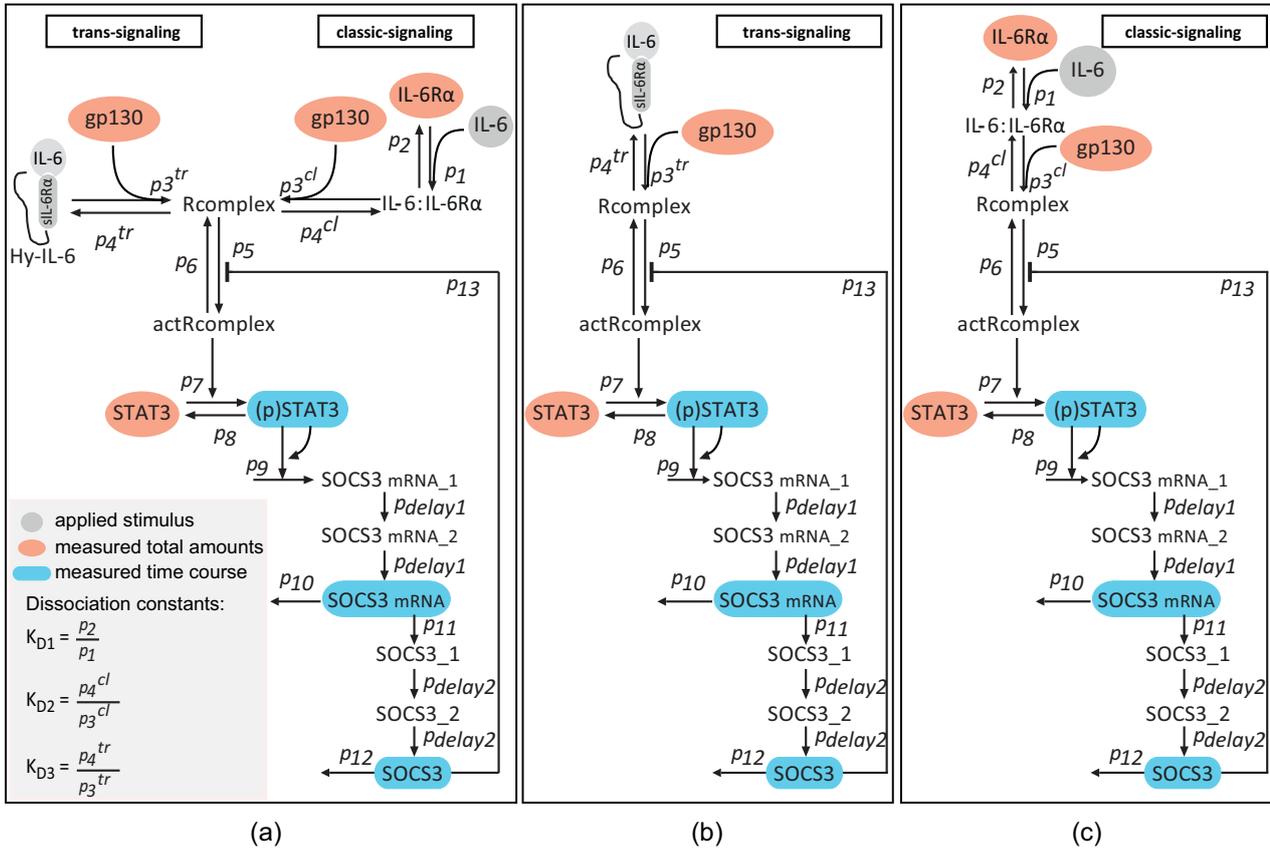


Figure 8.2: **Assumed network topologies.** Initial models describing (a) both, classic- and trans-signaling and (b) trans-signaling, only and (c) classic-signaling, only. Classic-signaling is induced by binding of IL-6 to membrane-bound IL-6Rα. The complex associates with gp130. Trans-signaling is induced by binding of Hy-IL-6 to gp130. The active receptor complex initiates Jak/STAT3 signaling and SOCS3 expression.

Based on the modeling assumptions (Appendix C) and the topology given in Fig. 8.2(b), the differential equations describing Hy-IL-6-induced trans-signaling are

given by:

$$\frac{\mathrm{d}[\mathrm{gp130}]}{\mathrm{dt}} = 2v_{4b} - 2v_{3b}$$

$$\frac{\mathrm{d}[\mathrm{actRcomplex}]}{\mathrm{dt}} = v_5 - v_6$$

$$\frac{\mathrm{d}[\mathrm{(p)STAT3}]}{\mathrm{dt}} = v_7 - v_8$$

$$\frac{\mathrm{d}[\mathrm{SOCS3\ mRNA\_1}]}{\mathrm{dt}} = v_9 - v_{10}$$

$$\frac{\mathrm{d}[\mathrm{SOCS3\ mRNA\_2}]}{\mathrm{dt}} = v_{10} - v_{11} \qquad (8.1)$$

$$\frac{\mathrm{d}[\mathrm{SOCS3\ mRNA}]}{\mathrm{dt}} = v_{11} - v_{12}$$

$$\frac{\mathrm{d}[\mathrm{SOCS3\_1}]}{\mathrm{dt}} = v_{13} - v_{14}$$

$$\frac{\mathrm{d}[\mathrm{SOCS3\_2}]}{\mathrm{dt}} = v_{14} - v_{15}$$

$$\frac{\mathrm{d}[\mathrm{SOCS3}]}{\mathrm{dt}} = v_{15} - v_{16}.$$

Additonally, the differential equations describing IL-6-induced classic-signaling are given as:

$$\frac{\mathrm{d}[\mathrm{IL{-}6{:}IL{-}6R}\alpha]}{\mathrm{dt}} = v_1 - v_2 - 2v_{3a} + 2v_{4a}$$

$$\frac{\mathrm{d}[\mathrm{gp130}]}{\mathrm{dt}} = 2v_{4a} - 2v_{3a}$$

$$\frac{\mathrm{d}[\mathrm{actRcomplex}]}{\mathrm{dt}} = v_5 - v_6$$

$$\frac{\mathrm{d}[\mathrm{(p)STAT3}]}{\mathrm{dt}} = v_7 - v_8$$

$$\frac{\mathrm{d}[\mathrm{SOCS3\ mRNA\_1}]}{\mathrm{dt}} = v_9 - v_{10}$$

$$\frac{\mathrm{d}[\mathrm{SOCS3\ mRNA\_2}]}{\mathrm{dt}} = v_{10} - v_{11} \qquad (8.2)$$

$$\frac{\mathrm{d}[\mathrm{SOCS3\ mRNA}]}{\mathrm{dt}} = v_{11} - v_{12}$$

$$\frac{\mathrm{d}[\mathrm{SOCS3\_1}]}{\mathrm{dt}} = v_{13} - v_{14}$$

$$\frac{\mathrm{d}[\mathrm{SOCS3\_2}]}{\mathrm{dt}} = v_{14} - v_{15}$$

$$\frac{\mathrm{d}[\mathrm{SOCS3}]}{\mathrm{dt}} = v_{15} - v_{16}.$$

Notably, to obtain the model describing both, classic- and trans-signaling (topology given in Fig. 8.2(a)), (8.1) and (8.2) are combined into one system.

Descriptions of the flux expressions $v_i$ are given in Appendix D, Table D.1. Concentrations for the quantities [IL-6:IL-6Rα], [Rcomplex] and [STAT3] can be derived from the following algebraic equations:

$$[\text{IL}-6\alpha^{\text{Total}}] = [\text{IL}-6\text{R}\alpha] + [\text{IL}-6\text{:IL}-6\text{R}\alpha] + 2[\text{Rcomplex}] + 2[\text{actRcomplex}]$$
$$[\text{gp130}^{\text{Total}}] = [\text{gp130}] + 2[\text{Rcomplex}] + 2[\text{actRcomplex}] \tag{8.3}$$
$$[\text{STAT3}^{\text{Total}}] = [\text{STAT3}] + [\text{(p)STAT3}],$$

In (8.3), the total amounts of the receptors and protein are set according the experimentally determined values (see Fig. 8.1(c)).

Note, a comprehensive description of the single model state variables, parameters and initial conditions are given in Appendix D, Tables D.2 and D.3.

## 8.4 Overall workflow for development of calibrated models and invalidity test for hypothesis

Our goal is to develop a model that is in line with the hypothesis for identical topologies and kinetics downstream of receptor activation in trans- and classic-signaling-induced Jak/STAT3 signaling. To achieve this goal we developed the following workflow (Fig. 8.3). First, parameter estimation is performed for the initial models (see topologies as given in Fig. 8.2). Subsequently, these results serve as inputs to reduced models lacking SOCS3 synthesis and negative feedback. The aim of this second step is to further confine the initial parameter ranges using again set-based parameter estimation. Next, the results of both set-based parameter estimation rounds are merged which results in calibrated models with reduced parameter ranges. To finally test whether our calibrated model cannot be invalidated and hence supports our initial hypothesis a yes/no workflow is applied. Notably, the results of the yes/no workflow applied in this study are depicted in bold black arrows in Fig. 8.3, while alternative workflows are given by dotted arrows. We first ask whether the obtained parameter ranges for the three calibrated models (Fig. 8.3 box 1) overlap. In case the ranges are disjoint, the initial hypothesis is deemed invalid. In case the ranges overlap, we next ask whether the model combining both, trans- and classic-signaling yields the smallest and the same ranges as at least one of the models describing trans-signaling only and classic-signaling only (box 2). If this question is neglected, a Monte Carlo sampling analysis is subsequently performed for all three models to check whether individual parametrizations can be found that overlap between all thre models (boxes 2a and b). In case the individual parameter sets are disjoint, we can state that our initial hypothesis is invalid. If in contrast, the obtained clouds of samples overlap, we deem
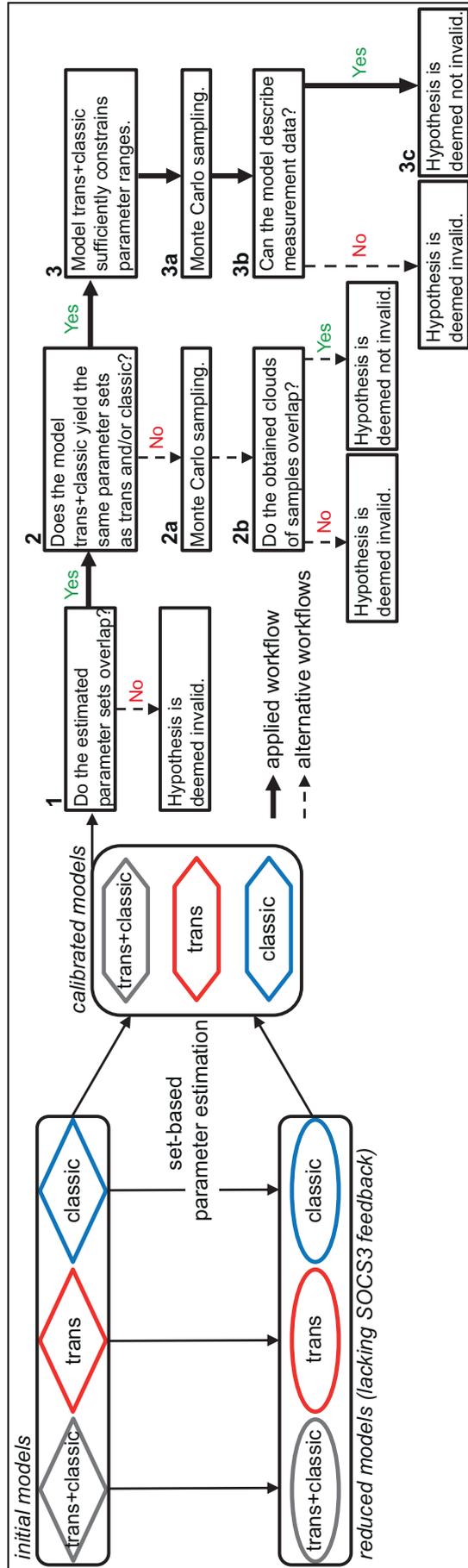
Figure 8.3: **Modeling workflow.** Model parameters of the initial models were estimated using the set-based approach and served subsequently as inputs to the reduced models, which lack SOCS3 protein synthesis and feedback loop. The obtained results by estimating the reduced models and the estimation results for parameters describing SOCS3 synthesis and negative feedback from the initial models are merged, which results in calibrated models. To finally test hypothesis invalidity, the given yes/no workflow is applied. Notably, the bold black arrow depicts the applied workflow, while dotted arrows depict alternative workflows.

this hypothesis as not invalid.

We above ask whether the model combining both, trans- and classic-signaling yields the smallest and same parameter ranges as at least one of the models describing trans-signaling only and classic-signaling only. If this applies, we can state that the model which combines both, trans- and classic-signaling constrains parameter ranges best (Fig. 8.3 box 3) and can be used for further Monte Carlo sampling analyses (box 3a). If finally parametrizations are determined, such that the model is capable to represent all measurement data (box 3b), we cannot invalidate the hypothesis that trans- and classic-signaling-induced Jak/STAT3 signaling employ the same pathway topology downstream of receptor activation (box 3c). Subsequently, the developed and not invalid model can be used for further analyses, while in negative case, the hypothesis is deemed invalid and is rejected.

## 8.5 Set-based modeling and parameter estimation

For set-based analyses of IL-6-induced classic-signaling and Hy-IL-6-induced trans-signaling, a discrete-time approximation of the models is derived applying a first order Euler discretisation scheme. We use a step size of 2 min for the first 30 minutes of stimulation and a step size of 2.5 min for the remaining time horizon, i. e. 30 to 90 minutes as these time points show less dynamic changes compared to the first 30 minutes.

As constraints, we employ the measurement data presented in Fig. 8.1 and initial boundaries of the 17 parameters (15 for classic-signaling and 13 for trans-signaling) are specified in a global range of $10^{-9}$-$10^3$ covering all biologically-justified parameter values. We further consider the range of 0.5-50 nM for the dissociation constant of the IL-6:IL-6Rα complex ($K_{D1}=\frac{p_2}{p_1}$) [13, 74, 75, 122, 170] and the range of 0.01-0.05 nM for both, the IL-6:IL-6Rα:gp130 complex ($K_{D2}=\frac{p_4^{cl}}{p_3^{cl}}$) and the Hy-IL6:gp130 complex ($K_{D3}=\frac{p_4^{tr}}{p_3^{tr}}$) [74, 122] (Fig. 8.2).

**Set-based estimation:** Based on these constraints, we applied the set-based approach and the outer-bounding algorithm to estimate the unknown parameter sets. Notably, empty parameter sets (i. e. no parametrizations exist that can represent measurement data) refer to an invalidity of our initial hypothesis that trans- and classic-signaling-induced Jak/STAT3 signaling employ the same pathway topology downstream of receptor activation. We started with the initial model that combines both, classic- and trans-signaling. The first round of set-based parameter estimation provided restrictions on the model parameters $p_3^{tr}$, $p_4^{tr}$, and $p_7$-$p_{12}$, while other parameters cannot ($p_5$, $p_6$ and $p_{13}$) or only marginally be restricted ($p_1$, $p_2$, $p_3^{cl}$ and $p_4^{cl}$) (Fig. 8.4 dark gray bars compared to initial parameter intervals in black and Table D.4, second column). Thus, parameter sets were found to be non-empty.
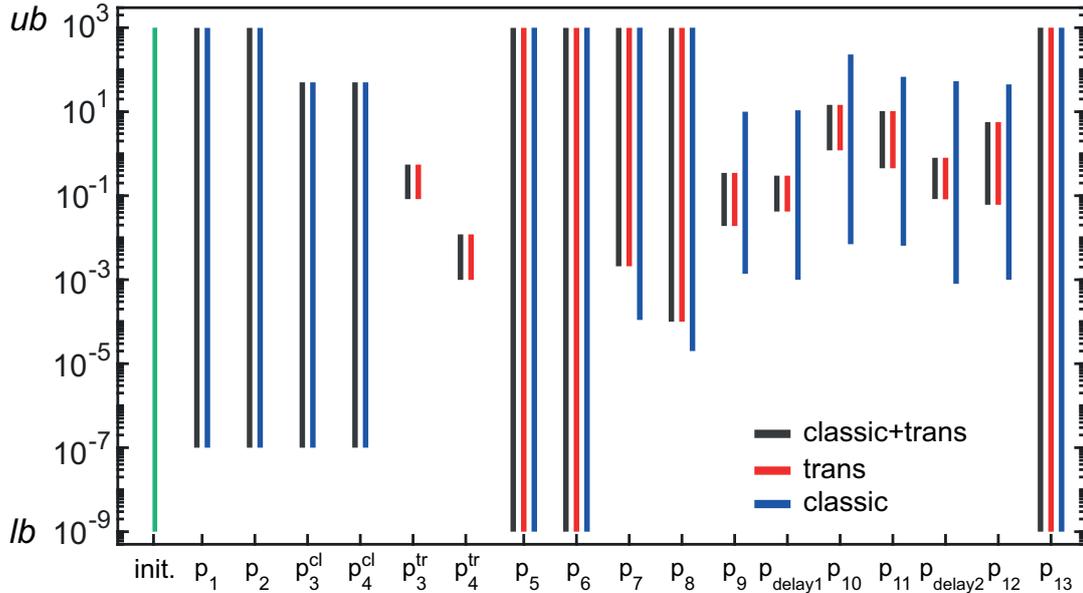
Figure 8.4: **Results for outer-bounding of model parameters for initial models.** Initial parameter bounds (green bar) range from $10^{-9}$ (lower bound, lb) to $10^3$ (upper bound, ub). Dark gray, red and blue bars depict ranges for the parameters after set-based analysis of the initial models.

Including both, classic- and trans-signaling simultaneously in a single model may constrain the parameter boundaries in comparison to specific models on classic- or trans-signaling. Thus, in a second step of our analyses, we tested whether separated implementation of models describing either classic- or trans-signaling reveals the same or different results for parameter estimation compared to those derived from a single model describing both, classic- and trans-signaling. We casted classic- and trans-signaling induced Jak/STAT3 signaling into two separated models. The results for set-based parameter estimation of these two models are depicted in Fig. 8.4 (red for trans-signaling only; blue for classic-signaling only) and Tables D.5 and D.6, second columns, respectively.

Our initial hypothesis that classic- and trans-signaling-induced Jak/STAT3 signaling employ the same pathway topology downstream of receptor activation therefore cannot be deemed invalid as also for separated implementation of models describing either trans- or classic-signaling parameter sets were found to be non-empty. We could restrict 14 ($p_1$, $p_2$, $p_3^{cl}$, $p_4^{cl}$, $p_3^{tr}$, $p_4^{tr}$, $p_7$, $p_8$, $p_9$, $p_{delay1}$, $p_{10}$, $p_{11}$, $p_{delay2}$, $p_{12}$) out of 17 model parameter ranges to at least one boundary. However, for the remaining three parameters ($p_5$, $p_6$ and $p_{13}$) no further restrictions could be made (Fig. 8.4). In summary the first round of set-based parameter estimation did not render our models invalid and enabled us to restrict most of the unknown parameters. This result counts for all three initial models.

## 8.6 Decoupling of fast and slow processes for improved parameter estimation

The so far unrestricted or only marginally restricted parameters are important to describe the initial and fast activation of the pathway. Thus, we analyzed these model parameters in the reduced models that decouple the early and fast receptor activation from the subsequent slow reactions including synthesis of SOCS3 protein and the SOCS3-dependent negative feedback. We used the estimated parameter bounds obtained by analyzing the initial models (Fig. 8.4) as inputs to the reduced models. Thereby we exploited the fact that biochemical parameters of interacting proteins are independent from the network topology. As before, one model describes classic- and trans-signaling (Fig. 8.5(a)) and two additional reduced models describe either trans- (Fig. 8.5(b)) or classic-signaling (Fig. 8.5(c)). By setting the parameters $p_{11}$, $p_{delay2}$, $p_{12}$ and $p_{13}$ to zero we assumed the production of SOCS3 protein - and hence the resulting negative feedback - to be blocked.

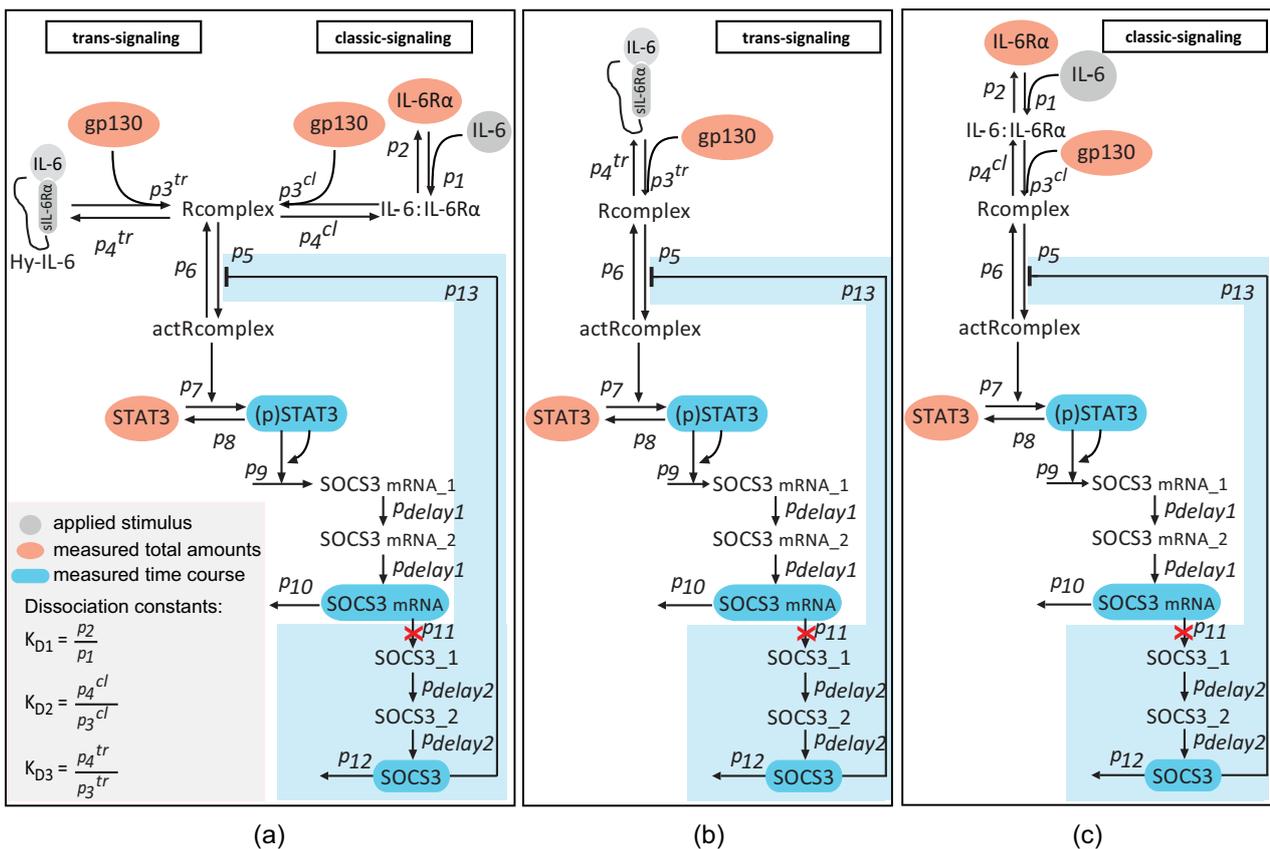

Figure 8.5: **Assumed network topology of the reduced models neglecting SOCS3 protein expression (red crosses) and negative feedback by SOCS3.** The blue boxes depict the network parts which are neglected in the reduced models by setting the corresponding parameter values to zero. (a) Model combining trans- and classic-signaling. (b) Model for trans-signaling. (c) Model for classic-signaling.

**Refined measurement data:** To match the above assumptions experimentally, we analysed the kinetics of Jak/STAT3 signaling in HepG2 cells stimulated with either 0.08 nM or 0.17 nM IL-6 and Hy-IL-6, respectively while blocking the synthesis of SOCS3 protein with cycloheximide (CHX) (Fig. B.2 in the appendix). In the presence of CHX, both IL-6 and Hy-IL-6-induced SOCS3 protein expression was blocked and consequently cytokine-induced STAT3 phosphorylation was strongly increased. Consequently, IL-6 and Hy-IL-6-induced phosphorylation of STAT3 was not transient in the presence of CHX but reached a plateau after 60 min of stimulation. Cytokine-induced expression of SOCS3 mRNA rose continuously until the end of the experiment. Notably, trans-signaling is stronger than classic-signaling also in the presence of CHX. These additional measurement data is used for parameter estimation based on the reduced models as described.

Compared to the analyses of the initial models, ranges of parameters $p_1$, $p_2$, $p_3^{cl}$, $p_4^{cl}$, $p_7$ and $p_8$ could be further reduced for all three models lacking the SOCS3 feedback loop (Fig. 8.6, compare light colours (w/o SOCS3 feedback) with the corresponding dark colours (including SOCS3 feedback); Table D.4, D.5 and D.6 fourth columns, respectively). Notably, ranges for parameters $p_5$, $p_6$ and $p_{13}$ could not be restricted, neither using the initial models and corresponding data, nor using the reduced models with the additional data.
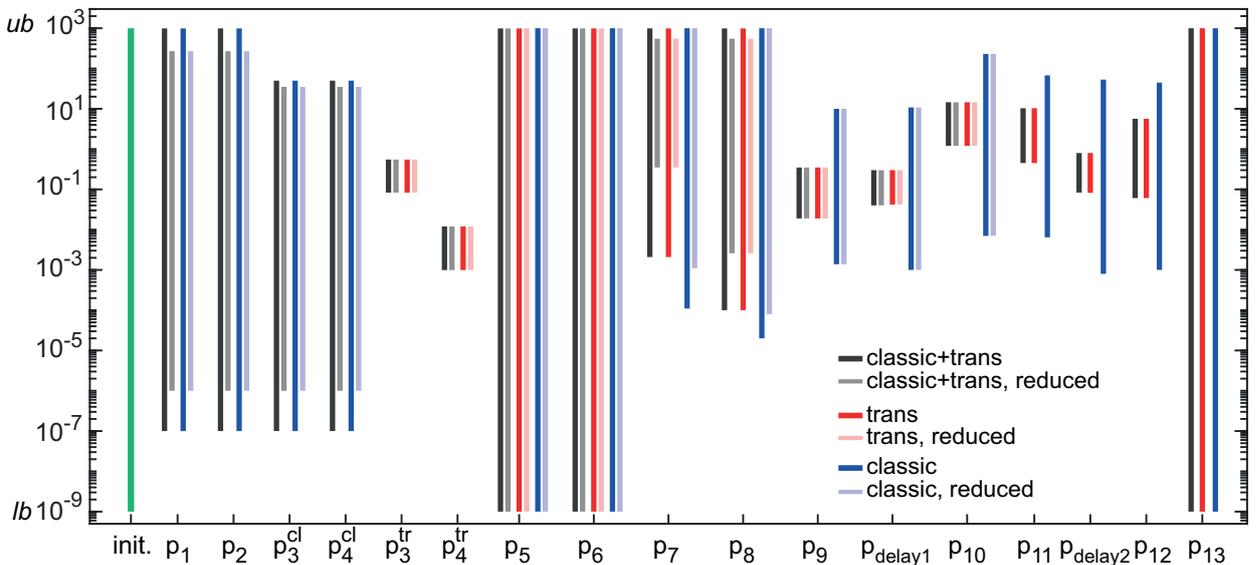


Figure 8.6: **Results for outer-bounding of model parameters for initial and reduced models.** Initial parameter bounds (green bar) range from $10^{-9}$ (lower bound, lb) to $10^3$ (upper bound, ub). Dark gray, blue and red bars depict parameter ranges for the individual parameters after set-based analysis of the initial models. Light gray, light blue and light red bars depict parameter ranges after parameter estimation of the reduced models neglecting SOCS3 mediated feedback.

We next merged the results from first and second set-based parameter estimation rounds by choosing the smallest obtained parameter ranges from both rounds (compare workflow Fig. 8.3) and thereby obtained calibrated models (Tables D.4, D.5 and

D.6, fourth columns, respectively). Subsequently, we followed the flow chart as given in Fig. 8.3 to test (non-)invalidity of our initial hypothesis that signaling mechanisms downstream of receptor activation do not differ between trans- and classic-signaling-induced Jak/STAT3 signaling. Our analyses showed that the obtained parameter sets for the models describing both, trans- and classic-signaling, trans-signaling only and classic-signaling only, overlap (Fig. 8.3 box 1). Furthermore, the parameter ranges estimated from the model describing specifically trans-signaling correspond to those ranges estimated from the model describing both, trans- and classic-signaling (compare dark and light gray and red bars in Fig. 8.6) (Fig. 8.3 box 2). Parameter ranges estimated from the model describing specifically classic-signaling were less restricted compared to the model describing trans-signaling. Obviously, the model which combines both, trans- and classic-signaling constrains the parameter ranges best (Fig. 8.3 box 3).

## 8.7 Monte Carlo sampling and set-based refinements of parameter ranges

Using the set-based approach parameter ranges were tightened, which led to fewer false positive solutions. However, to achieve our goal for developing a feasible model which supports our hypothesis that trans- and classic-signaling-induced Jak/STAT3 signaling employ the same pathway topology downstream of receptor activation, verification of the parameter sets within the given boundaries (Fig. 8.6) was required. We therefore applied Monte Carlo sampling using the calibrated model that describes both, trans- and classic-signaling (Fig. 8.3 box 3a). Out of 150,000 parametrizations, we derived the 150 parametrizations within the estimated parameter ranges (Fig. 8.6; light and dark gray bars), (Table D.4 fourth column) with lowest square deviation between our model predictions and the measurement data (exemplary magenta plus signs Fig. 8.7(a)). These 150 parametrizations allowed predictions, which are in line with the measurement data (Fig. 8.3 box 3b; Fig. 8.7(b)). Specifically, in Fig. 8.7(b) model predictions for the kinetics of trans- and classic-signaling-induced STAT3 phosphorylation, SOCS3 mRNA expression, and SOCS3 protein expression for up to 90 min are depicted in dark and light gray corridors, respectively. These corridors result from simulations of the model with the determined 150 parametrizations. Measurement data are given in red for trans-signaling and blue for classic-signaling. As the model was capable to represent all measurement data using the obtained parametrizations, we could not invalidate our initial hypothesis that signaling mechanisms downstream of receptor activation do not differ between trans- and classic-signaling as we wanted to show (Fig. 8.3 box 3c).

For most of the parameters the derived parametrizations did not cover the complete
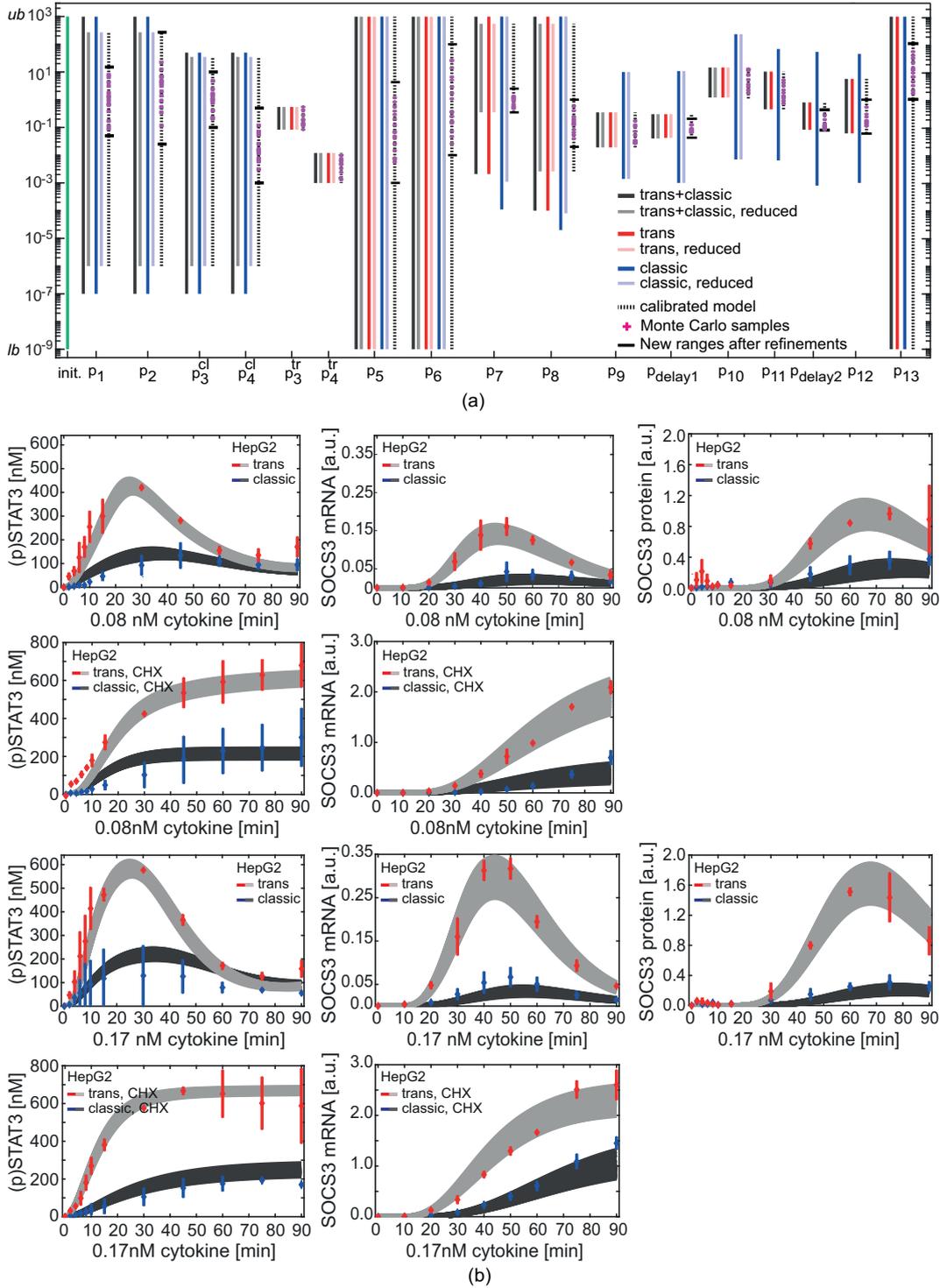
Figure 8.7: **Monte Carlo sampling and parameter refinements.** (a) Results for outer-bounding of model parameters for the intital (bold colors) and reduced models (light colors). Dotted bars depict final results for the calibrated model, magenta plus signs depict exemplary valid Monte Carlo samples and black horizontal lines show the newly obtained parameter ranges after refinements. (b) Resulting trajectories based on Monte Carlo samples from (a). Model outputs (light and dark gray corridors) were plotted against measurement data (red and blue lines).

estimated ranges (Fig. 8.7(a)). This, however, is no proof for the non-existence of valid solutions within these regions where no samples could be determined. Due to computational limits, it was not possible to exhaustively examine the parameter ranges for valid parametrizations using Monte Carlo sampling. In the following, we confirmed our results from Monte Carlo sampling. Specifically, we aimed to demonstrate that regions where no samples were determined, are invalid ones. We provided an iterative procedure that is built upon successive refinements of the lower and upper parameter bounds (Table D.4 fourth column) to tighten the overall parameter ranges. Starting with parameter $p_1$, we moved the previously estimated lower and upper bounds of $p_1$ inwards, while testing at each step if the model is deemed invalid. By this, we obtained refined and tightened parameter bounds for $p_1$. Remarkably, boundaries for $p_2$ were automatically restricted after refining $p_1$ as the ratio of $p_1$ to $p_2$ represent the dissociation constant of the IL-6:IL-6Rα complex. We proceeded with parameter $p_{cl}^3$ similar as to $p_1$. As result, also parameter $p_{cl}^4$ could be further restricted. The procedure was repeated for the remaining parameters resulting in a further refinement of the estimated parameter ranges depicted as black horizontal lines in Fig. 8.7(a) (Table D.4 fifth column). These refined ranges comprised all determined parametrizations. Thus, our results from Monte Carlo sampling could be confirmed.

As summary, Monte Carlo sampling and a subsequent refinement of parameter ranges allowed us to develop a mathematical model with tight and valid parameter ranges.

**Predictive capacity of the model:** We finally challenged the predictive capacity of our model and therefore, calculated the dose-dependent phosphorylation of STAT3 expected after 30 min of stimulation with either IL-6 or Hy-IL-6. For experimental validation, we stimulated HepG2 cells with 13 different equimolar concentrations of Hy-IL-6 and IL-6 for 30 min and monitored STAT3 phosphorylation by intracellular flow cytometry. Both classic- and trans-signaling induced phosphorylation of STAT3 dose-dependently. The experimental conditions used in Fig. 8.7(b) (stimulation with 0.08 nM and 0.17 nM cytokine for 30 min) were included and again resulted in stronger STAT3 phosphorylation in response to trans-signaling than in response to classic-signaling. Notably, trans-signaling was stronger than classic-signaling for all cytokine concentrations tested (Fig. 8.8).
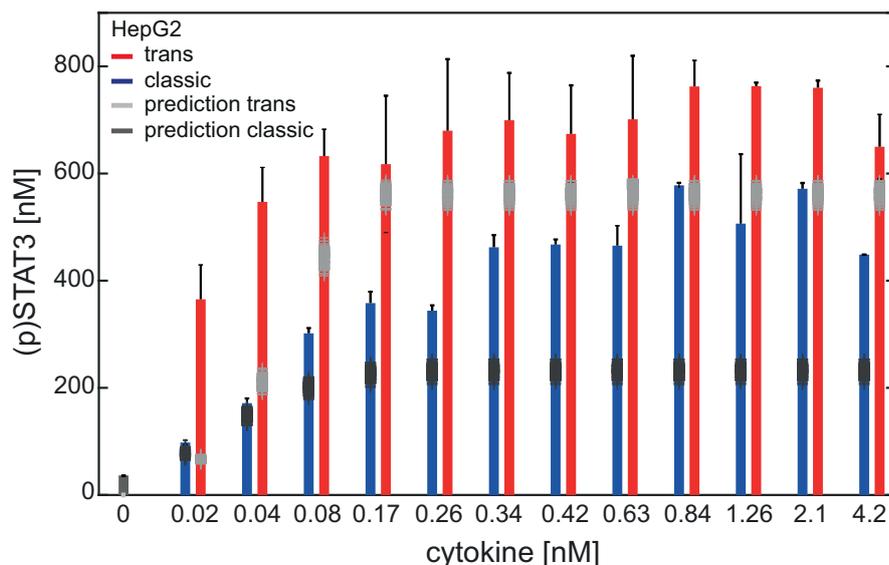
Figure 8.8: **Predictive capability of the model.** Model predictions of classic- (dark gray) and trans-signaling (light gray) for dose-dependent phosphorylation of STAT3 after 30 min stimulation with either IL-6 (blue bars) or Hy-IL-6 (red bars) using the 150 Monte Carlo samples from Fig. 8.7.

## 8.8 Model prediction reveals that differences between classic- and trans-signaling are caused by the ratio of gp130 to IL-6Rα on the cell surface

As we could not invalidate the hypothesis that topology and kinetics of Jak/STAT3 signaling downstream to the receptor activation are the same for classic- and trans-signaling, we next asked which components of the pathways are responsible for the observed differences in STAT3 activation in response to classic- and trans-signaling. To analyse whether the amount of receptors on the cell surface affects the ratio of classic- to trans-signaling, we varied the start values of gp130 and membrane-bound IL-6Rα. With these input variables, we performed model predictions using the obtained 150 Monte Carlo parameter samples. We predicted the ratio of trans- to classic-signaling-induced STAT3 phosphorylation after 30 minutes of cytokine stimulation. A ratio of 1 means that both signaling modes are equally strong activated (red bold line Fig. 8.9), whereas a ratio $> 1$ means that trans-signaling is stronger than classic-signaling.

First, we fixed IL-6Rα$^{\text{Total}}$ and STAT3$^{\text{Total}}$ to their determined mean concentration values 2.2 nM and 958 nM (Fig. 8.1(c)), respectively, and varied the mean value of gp130$^{\text{Total}} = 16.8$ nM $\pm$ one order of magnitude. Notably, for endogenous gp130 concentrations (white area) the model rendered well the high ratio of trans- to classic-signaling (Fig. 8.9(a)). For increasing amounts of gp130 the ratio of trans- to classic-signaling further increased, whereas, for lower amounts of gp130 the ratio of trans- to classic-signaling decreased. Next we varied the amount of IL-6Rα$^{\text{Total}} \pm$ one order of magnitude, i. e. ranging the IL-6Rα$^{\text{Total}}$ concentration from 0.22 nM to 22 nM
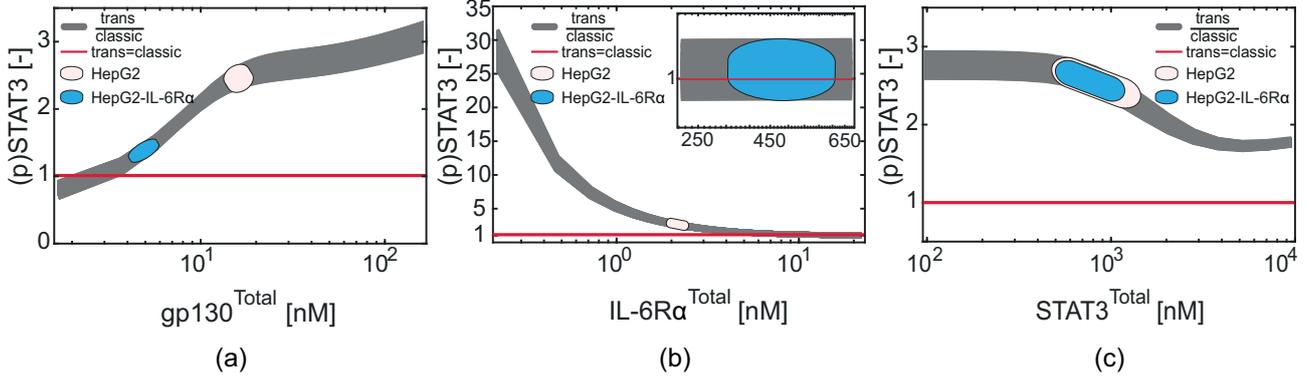
Figure 8.9: **Model predictions for the ratio of STAT3 phosphorylation in classic- and trans-signaling after 30 min cytokine stimulation with 0.17 nM using the obtained parametrizations from Fig. 8.7.** (a) The amount of $gp130^{Total}$ was changed from 1.68 nM to 168 nM while expression of IL-6Rα$^{Total}$ (2.2 nM) and STAT3$^{Total}$ (958 nM) were fixed. (b) The amount of IL-6Rα$^{Total}$ was changed from 0.22 nM to 22.2 nM while expression of $gp130^{Total}$ (16.8 nM) and STAT3$^{Total}$ (958 nM) were fixed. (c) The amount of STAT3$^{Total}$ was changed from 95.8 nM to 9580 nM while expression of $gp130^{Total}$ (16.8 nM) and IL-6Rα$^{Total}$ (2.2 nM) were fixed. Gray corridors correspond to model predictions. Red line depicts equal strength of classic- and trans-signaling. The white areas describe the receptor as well as protein amounts in HepG2 cells as presented in Fig. 8.1(c). The blue areas describe the receptor as well as protein amounts in HepG2-IL-6Rα cells as presented in Fig. 8.10

(Fig. 8.9(b)) while keeping the concentration of $gp130^{Total}$ and STAT3$^{Total}$ constant. The white area depicts the concentration $\pm$ STD of endogenous IL-6Rα at which trans-signaling is two to three times stronger than classic-signaling. Interestingly, for higher concentrations of IL-6Rα$^{Total}$ the difference between STAT3 phosphorylation during trans- and classic-signaling was completely ablated (Fig. 8.9(b)).

From these observations we conclude that the ratio of IL-6Rα to gp130 on the cell surface crucially determines the strength of classic- and trans-signaling. When IL-6Rα is expressed at lower numbers than gp130, trans-signaling allows formation of more active receptor complexes than classic-signaling and hence trans-signaling is stronger than classic-signaling. When, in contrast, gp130 is expressed at lower numbers than IL-6Rα, gp130 acts as a bottleneck and trans-signaling cannot surpass classic-signaling. In line with this hypothesis HepG2 cells express more gp130 than IL-6Rα (Fig. 8.1(c)).

The results from set-based modeling do not argue for classic- or trans-signaling-specific signal transduction downstream of the respective activated receptor complex. To substantiate this hypothesis we applied our model to predict the influence of the extent of STAT3 expression on the ratio of trans- to classic-signaling-induced STAT3 activation. We predicted STAT3 phosphorylation for changing amounts of STAT3$^{Total}$ ranging from 95.8 to 9580 nM (Fig. 8.9(c)) and fixed expression for IL-6Rα$^{Total}$ and $gp130^{Total}$. For low concentrations of STAT3$^{Total}$ trans-signaling-induced STAT3

phosphorylation was in average three time stronger than classic-signaling. With increasing amounts of STAT3$^{\text{Total}}$ the difference between trans- and classic-signaling decreased. However, STAT3 phosphorylation in trans-signaling was still more than two times higher than in classic-signaling at an concentration of 9580 nM STAT3$^{\text{Total}}$. This supports our hypothesis that intracellular signaling is not causative for the differences between classic- and trans-signaling. As result of our model predictions, we hypothesize that the ratio of gp130 to IL-6R$\alpha$ determines differences between classic- and trans-signaling.

**Experimental validation of model predictions:** To challenge the validity of the above predictions we generated HepG2 cells that stably overexpress IL-6R$\alpha$ (HepG2-IL-6R$\alpha$). Surface expression of IL-6R$\alpha$ and gp130 was quantified by FACS analysis (Fig. 8.10(a)). In contrast to HepG2 cells that express approximately 8 times more gp130 than IL-6R$\alpha$, HepG2-IL-6R$\alpha$ cells express approximately 10 times more IL-6R$\alpha$ than gp130. HepG2-IL-6R$\alpha$ cells therefore reflect a situation in which regarding to our mathematical simulation classic- and trans-signaling do not differ (Fig. 8.9 blue area). We next measured the dynamics of STAT3 phosphorylation, SOCS3 mRNA as well as SOCS3 protein expression for 0.17 nM IL-6- and Hy-IL-6 in HepG2-IL-6R$\alpha$ cells (Fig. 8.10(b)). Furthermore, the dose-dependent phosphorylation of STAT3 after 30 min of stimulation with 13 different concentrations of Hy-IL-6 and IL-6 was monitored (Fig. 8.10(c)). As shown by our model predictions, classic- and trans-signaling resulted in equal activation of Jak/STAT3 signaling in all analysed cases in HepG2-IL-6R$\alpha$ cells.

## 8.9 Summary and conclusion

We employed the set-based estimation framework for studying a large biochemical problem, i. e. signaling mechnanisms during IL-6-induced classic- and trans-signaling. To this end, we developed three different set-based models, i. e. one, which combines both, classic- and trans-signaling; one, which describes trans-signaling only; one, which describes classic-signaling only and analyzed these models according their capability of reproducing measurement data. As main assumption on all three models, we defined that intracellular processes and corresponding kinetics are the same in classic- and trans-signaling. As a result, none of the three models could be invalidated. We further derived outer-bounds for the unknown kinetic parameters. We noticed that the model which combines both, classic- and trans-signaling yields the same outer-bounds compared to the model, which describes trans-signaling only. The outer parameter bounds determined using the model which describes classic-signaling only were found to be larger compared to the remaining two models. Notably, the estimated parameter ranges were all overlapping and thus, our intitial hypothesis for common downstream

(a)

(c)



(b)
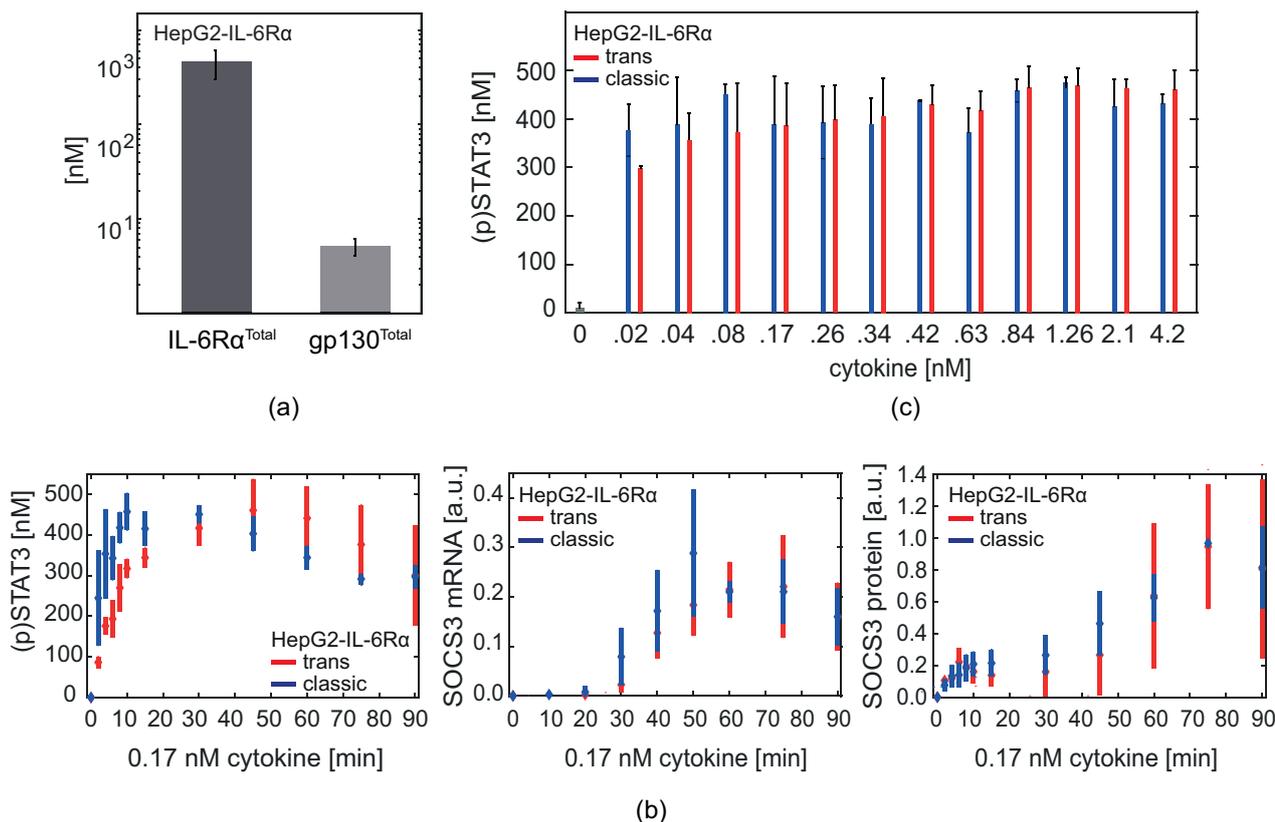
Figure 8.10: **High IL-6Rα/gp130 receptor ratio in HepG2-IL-6Rα cell lines
ablates difference between classic- and trans-signaling.** (a) Absolute quan-
tification of total amounts of IL-6Rα and gp130 expression in HepG2-IL-6Rα cells.
Bars denote the determined standard deviation for n=4 independent replicates,
respectively. (b) HepG2-IL-6Rα cells were stimulated with 0.17 nM IL-6 and Hy-
IL-6, respectively. STAT3 phosphorylation, SOCS3 mRNA and SOCS3 protein
expression were analysed. Data are given as mean±STD from n=4 experiments.
(c) HepG2-IL-6Rα cells were stimulated with indicated concentrations of IL-6 or
Hy-IL-6, respectively. STAT3 phosphorylation was evaluated. Data are given as
mean±STD from n=3 experiments.

signaling mechanisms during classic- and trans-signaling could not be invalidated.

Set-based analyses showed that ranges from 14 out of 17 parameters could be
restricted to at least one boundary. Ranges of three remaining parameters could be
not restricted at all. We proposed an iterative refinement of the obtained parameter
ranges based on Monte Carlo sampling. By this, we could further tighten the
parameter ranges improving predictibility of our model.

Our model-guided predictions together with biochemical experiments showed for
the first time that differences in the responsiveness of IL-6-induced classic- and trans-
signaling are only caused by the ratio of the IL-6-receptor subunits on the cell surface
of the responding cell but not by intracellular differences. When the amount of gp130
exceeds the amount of IL-6Rα trans-signaling is stronger than classic-signaling. In
contrast, when IL-6Rα exceeds gp130 both pathways are equally strong. Our results

call for intervention strategies that directly interfere with receptor activation instead of intracellular signaling to specifically block pro-inflammatory trans-signaling.

# 9 Conclusions and Outlook

To understand and unravel the complexity of biological systems with the objective to identify and design promising intervention strategies, model-guided analyses and predicitons are an important tool. The development and parametrization of mathematical models, however, is not trivial. Usually, kinetic parameters are unknown and have to inferred from the data. These data are typically subject to (large) noise. Noisy data makes it hard to decide whether a model reproduces the data or not. As result, often several competing model candidates exist due to several biological hypotheses, that need to be invalidated against each other. Moreover, biological processes often span several time scales, which need to be considered and their interrelation need to be unraveled for a deeper understanding of the underlying molecular interactions.

## 9.1 Summary

The present thesis expands the set-based estimation framework presented in [135, 136] and [24, 27] towards reliable parameter estimation, experimental design and multi-scale modeling. The used set-based framework enables to directly consider uncertainites in data, parameters and initial conditions and therefore, allows to provide guarantees.

Throughout the thesis, we applied the developed methods to understand IL-6-induced receptor assembly and activation as well as downstream signaling. IL-6 stimulates inflammatory and auto-immune processes in a number of diseases, such as rheumatoid arthritis, multiple sclerosis and Crohn's Diseases. Due to the involvement of IL-6 in the development of inflammatory diseases, there is a great interest in the development of anti-IL-6 therapies. We applied our methods to model and understand IL-6-induced signaling in order to obtain deeper insights into its (patho-)physiology and to present possible approaches targeting IL-6-induced misbalanced signaling.

After a review of the research topics, an introduction of the running example and an overview about the concepts for modeling biological systems (**Chapters 1-3**), we introduced the set-based estimation framework and two approaches for the estimation of unknown parameter sets in **Chapter 4**. The framework is based on a nonconvex feasibility problem that includes all information and equations of the underlying biological problem. To allow for conclusive statements about model invalidity, the nonconvex problem is relaxed into a convex semidefinte (or linear) program. These programs can be solved efficiently due to the availability of state-of-the art solvers. As example, we considered IL-6-induced receptor assembly and activation during classic-signaling. In

particular, we considered two competing model candidates, i. e. one model, which described stepwise receptor assembly and activation and another model, which described simplified receptor assembly and activation. Using simulated uncertain measurement data and initial parameter ranges, we demonstrated that the simplified model candidate is not capable to reproduce the data. Thus, this model candidate was deemed invalid. We were further able to derive an outer approximation of the valid parameter set for the model candidate that describes stepwise receptor assembly and activation. Our results underpin the current perception of IL-6-induced receptor activation [72], i. e. IL-6 forms in a first step a complex with the receptor subunit IL-6R$\alpha$ and in a second step, the receptor subunit gp130 binds to the complex of IL-6 and IL-6R$\alpha$.

Outer approximations still include invalid parametrizations due to the relaxation processes. As remedy, we derived so-called inner approximations of parameter sets in **Chapter 5**. The derivation of inner approximations is built upon a reformulation of available quantitative measurement data using binary variables and logical operators. An inversion of the reformulated measurement constraints allowed the estimation of inner approximations by proving invalidity of the inverted problem. We proposed two algorithms, i. e. a recursive algorithm and an incremental polytopic sample expansion algorithm. We demonstrated applicability considering two examples. For the first example, an enzyme-catalyzed reaction, tiny inner approximations were derived. For the second example describing early IL-6-induced trans-signaling, the recursive algorithm did not lead to an inner approximation. However, we derived valid parameter samples by local optimization and applied the polytopic sample expansion algorithm.

**Chapter 6** dealt with the design of experiments for the discrimination of valid and invalid models under uncertainties. We developed an approach based on bilevel optimization. In the inner program the different models were captured into one optimization problem. A constraint was formulated calculating the distance between the model output sets. As the inner program was nonconvex and thus, difficult to solve, the crucial idea consisted in the relaxation of the inner program using convex solvers. In the outer program - a nonconvex optimization problem - an input (sequence) was generated und subsequently supplied to the inner program checking whether or not the derived input leads to output sets, which do not overlap. While, the nonconvex inner program was convexified and solved using the set-based method, the outer program was solved using *fmincon*. We applied the proposed bilevel approach to two competing models describing IL-6-induced receptor assembly and activation. We designed an optimal IL-6 concentration for which the output sets of both model were separated. Our results call for additional biochemical experiments that allow for a deeper study of model candidates describing IL-6-induced receptor assembly and activation.

In **Chapter 7** approaches for the combination of biological processes that act on different time scales as well as the integration of uncertain data obtained at these time scales were presented. In the first part of the chapter, we introduced and applied a phenomenological approach for the combination of short-term signaling events (i. e.

dynamic phosphorylation of proteins) with long-term responses (i. e. cell growth). In particular, we correlated specific shape properties of short-term signals with cell growth using the set-based framework. We applied the approach to IL-6-induced classic-signaling and correlated the shape properties of the kinetics of IL-6-induced STAT3 phosphorylation with IL-6-induced cellular growth. We computed the maximum peakt height as well the the integrated response of the STAT3 phosphorylation signal and mapped the obtained results with cell growth. Our results showed that specific shape properties from the short-term time scale can be used to predict long-term cellular responses. Although this approach is very specific and needs further experimental validation, we showed that phenomenological approaches are very useful when computing time is a limiting factor and the size of the models needs to be kept small.

In the second part of the chapter, we proposed an approach which combines the set-based framework with classification methods. The main goal for combining both frameworks is the stratification of patients into different categories for the long-term risk of developing a certain disease based on uncertain data from the short-term time scale. To this end, we simulated a cohort of 50 patients and short-term profiles for IL-6-induced STAT3 and ERK phosphorylation. Both proteins, STAT3 and ERK, can be seen as upstream markers for developing inflammatory diseases. The corresponding patient- and pathways-specific parameters served as inputs to the neural network and as result, 84% of the patients were stratified correctly into the corresponding categories.

In Chapters 4-7 we applied our developed extensions for the set-based approach to rather small/medium-sized models describing IL-6-induced signaling. In **Chapter 8**, the set-based framework was applied to a large pathophysiological problem describing IL-6-induced receptor activation, Jak/STAT3 signaling and negative feedback inhibition by SOCS3 during classic- and trans-signaling. We studied differences in IL-6-induced classic- and trans-signaling using set-based parameter estimation together with model-guided predictions and biochemical analyses. We showed, that differences in IL-6-induced classic- and trans-signaling are only mediated by differences in the expression of the receptors of the responding cell.

## 9.2 Outlook

The topics addressed in this thesis lay the basis for a series of research directions. The most limiting factor using the set-based methods is the computational demand for composing and solving the semidefinite and linear programs. Especially, in the application example (Chapter 8) we reached computational limits solving the model which combines both, classic- and trans-signaling. A reason for that is, that the set-based framework exponentially depends on both, the problem size and the length of the considered time horizon. A possible solution approach for the reduction of computational costs is the exploitation of certain system properties, such as symmetry [137, 164]. This

allows the decomposition of large semidefinite programs into smaller subproblems. The application of the concept of symmetry is, however, problem-dependent and may not be applied directly due to the problem structure in biological examples. Additional concepts for complexity reduction within the set-based framework were proposed in [139] and showed promising results for a multiple tank system.

We presented set-based approaches for the estimation of inner parameter sets, the design of experiments and the combination of processes acting on different time scales. All approaches were applied to the example of IL-6-induced signaling. Further research should focus on validation of the obtained results. As an example, additional measurement data will allow an invalidation of one of the proposed model candidates (Chapter 6) and will deepen our knowledge about IL-6-induced pathway activation and receptor assembly. Additional future research should also focus on the combination of conventional methods, such as the set-based framework with classification approaches. The stratification of patient cohorts becomes more and more important as the demand of new intervention strategies increases. Furthermore, the heterogeneity in the data landscape (e. g. quantitative, sparse data versus qualitative, densed data) need to be considered. Model-guided analyses (including pathway modeling but also machine learning strategies) as well as the combination of the different approaches allow the extraction of helpful information for physicians to design personalized intervention strategies [66].

We worked towards the goal for the design of new personalized intervention strategies by developing and analyzing set-based models that describe IL-6-induced signaling pathways. Future research should focus on further validation of our results, that differences in the responsiveness of IL-6-induced classic- and trans-signaling are only mediated by the membrane-bound receptor subunits IL-6R$\alpha$ and gp130. Our model-guided analyses lay the basis for the development of new individualized intervention strategies by taking into account patient-specific ratios of IL-6R$\alpha$ and gp130. So far, approved antibodies, such as *tocilizumab*, have many negative side effects. A main reason for that is, that a therapy with *tocilizumab* also blocks IL-6 classic-signaling, which in turn leads to a reduced capability of the body to cope with bacterial infections. Based on our knowledge that differences in the responsiveness of IL-6 classic- and trans-siganling are only caused by differences in the receptor ratios, we will be able to design new and improved targeted treatments of inflammatory diseases caused by trans-signaling. At the same time partially or complete blockade of IL-6 classic-signaling can be prevented. Thus, improving the quality of life for patients.

# Appendix

## A Generation of measurement data

All data presented were generated by M. Sc. Heike Reeh and M. Sc. Ulrike Billing (group of Prof. Fred Schaper and Dr. Anna Dittrich, Department of Systems Biology, Magdeburg). To obtain measurement data, a human liver cancer cell line (HepG2) and a mouse pro B cell line (Ba/F3) were used. To induce classic- and trans-signaling, the cells were stimulated with equimolar concentrations of IL-6 and Hyper-IL-6, respectively. Hyper-IL-6 is an artificially generated fusion protein of IL-6 and sIL-6R$\alpha$ which is used to mimic trans-signaling [56] (cf. Fig.B.1 for experimental validation).

**Relative quantification of protein phosphorylation:** For quantifications of cellular proteins, such as SOCS3 and (phosphorylated) STAT3, Western Blotting was performed. Western Blotting is a technique where proteins within cellular lysates are denaturated in the presence of sodium dodecyl sulfate (SDS) and loaded on a gel. The gel is then used to separate the proteins according their size by electrophoresis (SDS-PAGE). Subsequently, the proteins are transferred on PVDF (polyvinylidene fluoride) membranes and stained with specific antibodies against the protein of interest.

**Absolute quantification of intracellular proteins:** Absolute amounts of STAT3 and phosphorylated STAT3 per cell were analysed by quantitative immunoprecipitation (IP). Briefly, STAT3 and phosphorylated STAT3 were isolated from cellular lysates by immunoprecipitation, a technique to pull out all factors binding to a protein by using a specific antibody to the respective protein. Isolated proteins together with known amounts of a recombinant STAT3 calibrator protein were analyzed by SDS-PAGE and Western Blotting followed by detection with an anti-STAT3 antibody specific for an epitope present in STAT3, phosphorylated STAT3 and recombinant STAT3.

**Absolute quantification of membrane proteins:** To determine total numbers of the receptors IL-6R$\alpha$ and gp130 on the cell membrane, a flow cytometry assay was used. Flow cytometry is a laser-based technology employed to detect cell surface antigens. The amount of gp130 and IL-6R$\alpha$ on the cell surface was analysed using the bead-based flow cytometry assay QIFIKIT [125] according to manufacturer's protocol.

**mRNA quantification:** For relative quantification of mRNA, quantitative RT-PCR

(real-time polymerase chain reaction, qRT-PCR) was performed. Briefly, RNA (or mRNA) molecules are converted into their complementary DNA (cDNA) sequences by reverse transcriptase. Reverse transcription is followed by a RT-PCR to quantify the amount of input RNA.

**Growth assays:** Cell growth was measured using the CellTiter-Blue Cell Viability Assay reagent (Promega, Karlsruhe, Germany). To this end, the cells are stimulated and cultivated for 48 hours. Then, cell growth is quantified by measuring the extinction of the cells which determines the number of viable cells in a sample.

For details on the experimental methods, we refer the reader to [130] and references therein.

# B Measurement data

## B.1 IL-6 versus Hy-IL-6 stimulation
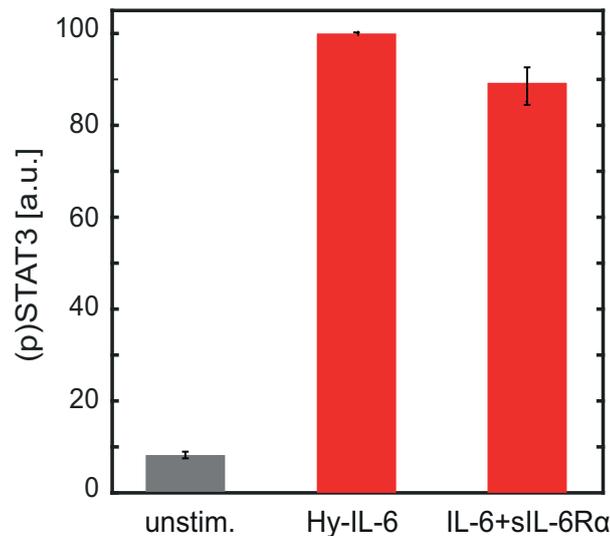


Figure B.1: HepG2 cells were stimulated with 0.17 nM Hy-IL-6 or a mixture of 0.17 nM IL-6 and 100 nM sIL-6R$\alpha$. STAT3 phosphorylation was evaluated by intracellular flow cytometry using specific fluorescent antibodies against STAT3 (p)Y705. For independent experiments mean fluorescence of 10000 cells per time point was calculated and maximal mean fluorescence was normalized to 100%. Data are given as mean±STD from n=3 independent experiments.

## B.2 Treatment of cells with cycloheximid



(a)



(b)
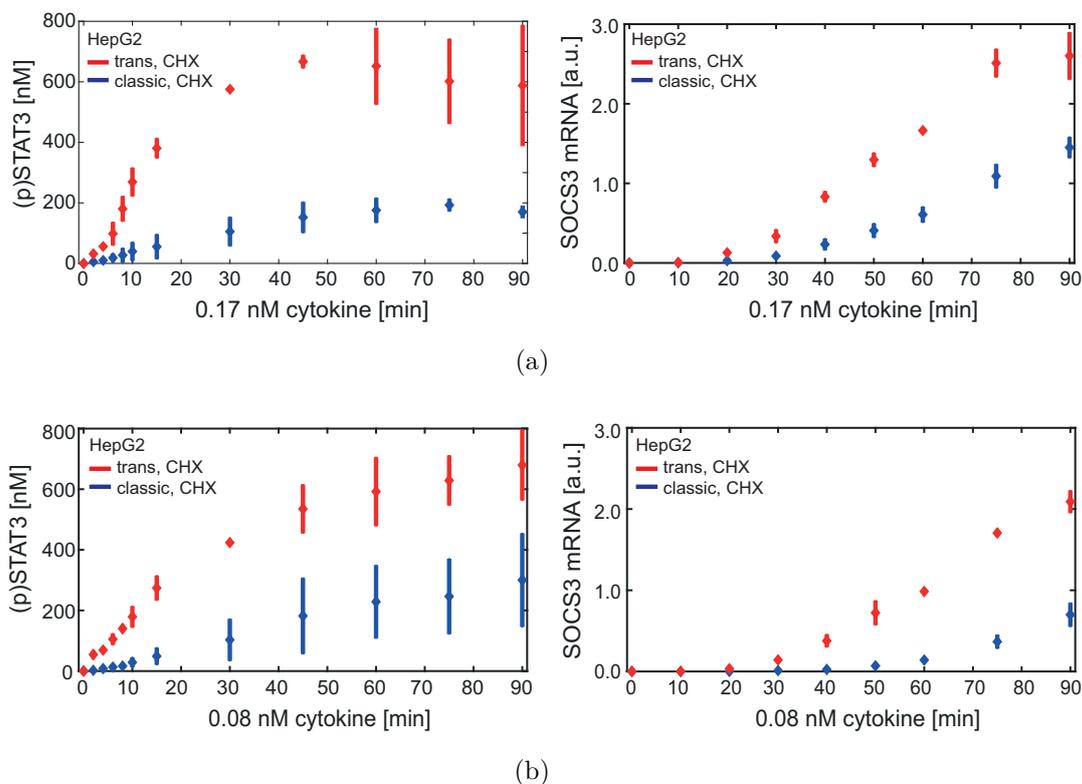
Figure B.2: HepG2 cells were stimulated with (a) 0.08 nM IL-6 (blue lines) and Hy-IL-6 (red lines), respectively and (b) 0.17 nM IL-6 (blue lines) and Hy-IL-6 (red lines), respectively. STAT3 phosphorylation and SOCS3 proetin expression were evaluated by Western Blotting. The expression of SOCS3 mRNA was analysed using qRT-PCR. Data are given as mean±STD from n= experiments.

## B.3 Data normalization

**Proteins:** All experiments were performed in a minimum of three independent biological replicates. For time series, numerical values obtained by quantification of Western Blots of the respective proteins were divided by those of the loading controls at each time point. Resulting values for each individual experiment were normalized to the value at time point t = 30 min (analysis of STAT3 phosphorylation) or t = 60 min (analysis of SOCS3 expression) Hy-IL-6 (HepG2) or IL-6 (HepG2-IL-6R$\alpha$) or to the value of either IL-6 or Hy-IL-6 (Ba/F3-gp130-IL-6R$\alpha$), respectively. Subsequently, the resulting value at time point t = 0 min was subtracted from the values at each other time point within an individual experiment. Mean values and standard deviation over all biological replicates were calculated.

For analysis of STAT3 activation resulting mean values were then normalized to absolute amounts of (p)STAT3. First, the concentration dependency of STAT3 phosphorylation was determined. To do so, the strength of STAT3 phosphorylation after

stimulation of HepG2 cells for 30 min with 0.08 nM, 0.17 nM, and 0.42 nM Hy-IL-6 (n = 3 each) was determined by Western Blotting. Second, the resulting mean value of STAT3 phosphorylation induced by 0.42 nM Hy-IL-6 was equated with the absolute amount of (p)STAT3 obtained by quantitative immunoprecipitation. Absolute amounts of phosphorylated STAT3 after stimulation with 0.08 nM and 0.17 nM Hy-IL-6 were adapted according to the determined dose-dependency.

To normalize relative STAT3 activation to absolute numbers in HepG2-IL-6R$\alpha$ cells and Ba/F3-gp130-IL-6R$\alpha$ cells these cells were stimulated with 0.08 nM, 0.17 nM, and 0.42 nM IL-6 and calculations were performed as described for HepG2 cells.

To analyse concentration dependency of SOCS3 protein expression HepG2 cells were stimulated for 60 min with 0.08 nM and 0.17 nM Hy-IL-6 and SOCS3 expression was analysed by Western Blotting (n = 3 each). The ratio of SOCS3 expression induced by these different amounts of Hy-IL-6, i. e. $r^{SOCS3} = \frac{SOCS3(0.17nM)}{SOCS3(0.08nM)}$ was subsequently calculated for the mean of three independent replicates. The resulting ratio was used to normalize kinetics of SOCS3 protein expression.

**mRNA:** qRT-PCR results were normalized to time point t = 60 min Hy-IL-6 + CHX treatment for HepG2 cells and to time point t = 60 min IL-6 + CHX treatment for HepG2-IL-6R$\alpha$ cells. Subsequently, the value at t = 0 min was subtracted from all other time points. Mean value and standard deviation of the resulting values over all biological replicates were calculated.

To obtain the concentration dependency of SOCS3 mRNA expression, the cells were pretreated with CHX and stimulated for 60 min with 0.08 nM and 0.17 nM Hy-IL-6 (HepG2) and IL-6 (HepG2-IL-6R$\alpha$), respectively (n = 3 each). The ratio of SOCS3 mRNA expression induced by these different amounts of Hy-IL-6 or IL-6, i. e. $r^{SOCS3mRNA} = \frac{SOCS3mRNA(0.17nM)}{SOCS3mRNA(0.08nM)}$ was subsequently calculated for the mean of three independent experiments. The resulting ratio was used to normalize kinetics of SOCS3 mRNA expression.

# C Model assumptions

To cast IL-6-induced classic- and trans-signaling into ODE-based models (Chapter 8, we had to make the following modeling assumptions:

i) The system is well mixed. Because of the large number of initial cells we neglect stochastic effects.

ii) Since no obvious difference in STAT3 activation in response to trans-signaling induced by Hy-IL-6 or by the IL-6:sIL-6R$\alpha$ complex exists (Fig. B.1 in the Appendix), we did not incorporate binding of IL-6 to soluble IL-6R$\alpha$ during trans-signaling in our model.

iii) IL-6 and Hy-IL-6 are assumed as constant model inputs as no change in the amount of cytokine in the supernatant was observed experimentally during the considered time horizon of 90 minutes (data not shown).

iv) The hexameric receptor complex (IL-6:IL-6R$\alpha$:gp130)$_2$ formed during classic-signaling and the hexameric receptor complex (Hy-IL-6:gp130)$_2$ formed during trans-signaling are considered as active receptor complexes (actRcomplex), [28].

v) Formation of the active receptor complex induces activation of receptor-associated Jaks [98, 151] and subsequent phosphorylation of tyrosine residues in the cytoplasmatic domain of gp130 [70]. However, as the interaction of Jaks and gp130 is very tight [64], we do not explicitly consider Jaks in our model. Rather, we assume Jaks to be represented as part of the gp130 state variable.

vi) STATs are phosphorylated by the active receptor complex. The activated receptor complexes represent activated Jaks [151].

vii) To describe nonlinear dynamcis of SOCS3 mRNA transcription, we add a positive feedback to (p)STAT3-induced SOCS3 mRNA transcription [49].

viii) Negative feedback inhibition via SOCS3 (58) is modeled by a rational term to allow for an inhibition of the receptor activity [150].

ix) The species SOCS3 mRNA_1, SOCS3 mRNA_2, SOCS3_1 and SOCS3_2 are modelled to simulate the delay caused by SOCS3 mRNA transcription, mRNA processing and translation, respectively. To this end, we apply a linear chain, where the delays are distributed in two steps each with the kinetic rates $p_{delay1}$ and $p_{delay2}$, respectively [150].

# D  Description of models and parameters

## D.1  Model fluxes

Considered model fluxes are given in the table:

Table D.1: Expression and description of considered model fluxes.

| Flux | Equation | Description |
|------|----------|-------------|
| $v_1$ | $p_1 \cdot [\text{IL-6}] \cdot [\text{IL-6R}\alpha]$ | Association of IL-6 to IL-6R$\alpha$ |
| $v_2$ | $p_2 \cdot [\text{IL-6:IL-6R}\alpha]$ | Dissociation of IL-6:IL-6R$\alpha$ |
| $v_3$ | | Association of hexameric receptor complex |
| $v_{3a}$ | $p_3^{cl} \cdot [\text{IL-6:IL-6R}\alpha]^2 \cdot [\text{gp130}]^2$ | classic (IL-6:IL-6R$\alpha$:gp130)$_2$ |
| $v_{3b}$ | $p_3^{tr} \cdot [\text{Hy-IL-6}]^2 \cdot [\text{gp130}]^2$ | trans (Hy-IL-6:gp130)$_2$ |
| $v_4$ | | Dissociation of hexameric receptor complex |
| $v_{4a}$ | $p_4^{cl} \cdot [\text{Rcomplex}]$ | classic |
| $v_{4b}$ | $p_4^{tr} \cdot [\text{Rcomplex}]$ | trans |
| $v_5$ | $\frac{p_5 \cdot [\text{Rcomplex}]}{1 + p_{13} \cdot [\text{SOCS3}]}$ | Receptor complex activation and negative feedback inhibition by SOCS3 |
| $v_6$ | $p_6 \cdot [\text{actRcomplex}]$ | Receptor complex deactivation |
| $v_7$ | $p_7 \cdot [\text{actRcomplex}] \cdot [\text{STAT3}]$ | STAT3 phosphorylation |
| $v_8$ | $p_8 \cdot [\text{(p)STAT3}]$ | STAT3 dephosphorylation |
| $v_9$ | $p_9 \cdot [\text{(p)STAT3}]^2 \cdot (1\text{nM}+[\text{(p)STAT3}])$ | Initiation of SOCS3 mRNA transcription |
| $v_{10}$ | $p_{delay1} \cdot [\text{SOCS3 mRNA\_1}]$ | Transcriptional delay |
| $v_{11}$ | $p_{delay1} \cdot [\text{SOCS3 mRNA\_2}]$ | Transcriptional delay |
| $v_{12}$ | $p_{10} \cdot [\text{SOCS3 mRNA}]$ | SOCS3 mRNA degradation |
| $v_{13}$ | $p_{11} \cdot [\text{SOCS3 mRNA}]$ | Initiation of SOCS3 protein synthesis |
| $v_{14}$ | $p_{delay2} \cdot [\text{SOCS3\_1}]$ | SOCS3 protein synthesis delay |
| $v_{15}$ | $p_{delay2} \cdot [\text{SOCS3\_2}]$ | SOCS3 protein synthesis delay |
| $v_{16}$ | $p_{12} \cdot [\text{SOCS3}]$ | SOCS3 protein degradation |

## D.2 Description of state variables and initial conditions

A description of all model state variables including initial conditions for set-based analyses can be taken from Table D.2.

Table D.2: Description of state variables and initial conditions.

| State variable | Description | Initial condition [unit] |
|---|---|---|
| IL–6, Hy–IL–6 | free IL–6, Hyper IL–6 | according to experimental setup [nM] |
| IL–6Rα | membrane-bound receptor subunit IL–6Rα | [1.8,2.54] [nM] |
| gp130 | membrane-bound receptor subunit gp130 | [13.7,19.9] [nM] |
| IL–6:IL–6α | dimerized receptor complex classic-signaling | 0 [nM] |
| Rcomplex | non-active hexameric receptor complex | 0 [nM] |
| actRcomplex | active hexameric receptor complex | 0 [nM] |
| STAT3 | signal transducer and activator of transcription 3 | [478,1428] [nM] |
| (p)STAT3 | phosphorylated STAT3 | 0 [nM] |
| SOCS3 mRNA | suppressor of cytokine signaling 3 mRNA | 0 [a.u.] |
| SOCS3 mRNA _ 1 | and intermediate states _ 1 | 0 [a.u.] |
| SOCS3 mRNA _ 2 | and _ 2 | 0 [a.u.] |
| SOCS3 | suppressor of cytokine signaling 3 | 0 [a.u.] |
| SOCS3 _ 1 | and intermediate states _ 1 | 0 [a.u.] |
| SOCS3 _ 2 | and _ 2 | 0 [a.u.] |

## D.3 Set-based parameter estimation results

To identify the unknown parameters within the set-based framework the ODE-systems (8.1), (8.2) and the combination of both were discretized using a first order Euler scheme and subsequently implemented in ADMIT. In total, the FP describing classic- and trans-signaling in one model consisted of 41 time steps, 17 unknown parameters, 10 state variables for classic-signaling, 9 state variables for trans-signaling and the size of the problem was about 1335 variables large. For a separated analysis of classic- and trans-signaling, each FP consisted of 41 time steps, 15 unknown parameters for classic-signaling, 13 unknown parameters for trans-signaling and the size of the problem was about 720 variables for classic-signaling and 635 variables for trans-signaling.

Table D.3, Table D.4 as well as Tables D.5 and Table D.6 depict a description of all model parameters and the obtained outer-bounding results for the considered models, respectively.

Table D.3: Description of model parameters.

| Parameter | Unit | Description |
| --- | --- | --- |
| $p_1$ | $nM^{-1}min^{-1}$ | Association of IL-6:IL-6R$\alpha$ complex |
| $p_2$ | $min^{-1}$ | Dissociation of IL-6:IL-6R$\alpha$ complex |
| $p_3^{cl}$ | $nM^{-3}min^{-1}$ | Association of (IL-6:IL-6R$\alpha$:gp130)$_2$ |
| $p_3^{tr}$ | $nM^{-3}min^{-1}$ | and (Hy-IL-6:gp130)$_2$ complexes |
| $p_4^{cl}$ | $min^{-1}$ | Dissociation of (IL-6:IL-6R$\alpha$:gp130)$_2$ |
| $p_4^{tr}$ | $min^{-1}$ | and (Hy-IL-6:gp130)$_2$ complexes |
| $p_5$ | $min^{-1}$ | Activation of the receptor complex |
| $p_6$ | $min^{-1}$ | Deactivation of the receptor complex |
| $p_7$ | $nM^{-1}min^{-1}$ | Phosphorylation of STAT3 |
| $p_8$ | $min^{-1}$ | Dephosphorylation of STAT3 |
| $p_9$ | $\mu M^{-3}min^{-1}$ | Transcription of SOCS3 mRNA |
| $p_{delay1}$ | $min^{-1}$ | Transcriptional delay |
| $p_{10}$ | $min^{-1}$ | Degradation of SOCS3 mRNA |
| $p_{11}$ | $min^{-1}$ | Translation of SOCS3 protein |
| $p_{delay2}$ | $min^{-1}$ | Translational delay |
| $p_{12}$ | $min^{-1}$ | Degradation of SOCS3 protein |
| $p_{13}$ | $a.u.^{-1}$ | Negative feedback inhibition |

Table D.4: **Description of model parameters, initial uncertainty intervals and set-based estimation results for the model describing both, classic- and trans-signaling.** $\underline{p_i}$ and $\overline{p_i}$ indicate the estimated lower and upper bounds for the individual parameters using an outer-bounding algorithm. The first column depicts the corresponding parameter names. The second, third, and fourth columns depict the set-based parameter estimation results for the initial model (Fig. 8.2(a)), the reduced model (Fig. 8.5(a)) and the calibrated model, respectively. Parameter ranges obtained after iterative refinements based on Monte Carlo sampling are depicted in the fifth column.

| Parameter | $[\underline{p_i},\overline{p_i}]$ initial model | $[\underline{p_i},\overline{p_i}]$ reduced model | $[\underline{p_i},\overline{p_i}]$ calibrated model | $[\underline{p_i},\overline{p_i}]$ refined bounds |
|---|---|---|---|---|
| $p_1$ | $[10^{-7},9.9\cdot10^2]$ | $[10^{-6},2.7\cdot10^2]$ | $[10^{-6},2.7\cdot10^2]$ | $[5\cdot10^{-2},1.5\cdot10^1]$ |
| $p_2$ | $[10^{-7},9.9\cdot10^2]$ | $[10^{-6},2.7\cdot10^2]$ | $[10^{-6},2.7\cdot10^2]$ | $[2.5\cdot10^{-3},2.7\cdot10^2]$ |
| $p_3^{cl}$ | $[10^{-7},5\cdot10^1]$ | $[10^{-6},3.5\cdot10^1]$ | $[10^{-6},3.5\cdot10^1]$ | $[10^{-1},10^1]$ |
| $p_3^{tr}$ | $[8.3\cdot10^{-2},5.5\cdot10^{-1}]$ | $[8.3\cdot10^{-2},5.5\cdot10^{-1}]$ | $[8.3\cdot10^{-2},5.5\cdot10^{-1}]$ | no refinement |
| $p_4^{cl}$ | $[10^{-7},5\cdot10^1]$ | $[10^{-6},3.5\cdot10^1]$ | $[10^{-6},3.5\cdot10^1]$ | $[10^{-3},5\cdot10^{-1}]$ |
| $p_4^{tr}$ | $[9.9\cdot10^{-4},10^{-2}]$ | $[9.9\cdot10^{-4},10^{-2}]$ | $[9.9\cdot10^{-4},10^{-2}]$ | no refinement |
| $p_5$ | $[10^{-9},10^3]$ | $[10^{-9},10^3]$ | $[10^{-9},10^3]$ | $[10^{-3},5]$ |
| $p_6$ | $[10^{-9},10^3]$ | $[10^{-9},10^3]$ | $[10^{-9},10^3]$ | $[10^{-2},10^2]$ |
| $p_7$ | $[2\cdot10^{-3},9.9\cdot10^2]$ | $[3.5\cdot10^{-1},5.5\cdot10^2]$ | $[3.5\cdot10^{-1},5.5\cdot10^2]$ | $[3.5\cdot10^{-1},5]$ |
| $p_8$ | $[4\cdot10^{-4},9.9\cdot10^2]$ | $[2.6\cdot10^{-3},5.5\cdot10^2]$ | $[2.6\cdot10^{-3},5.5\cdot10^2]$ | $[2\cdot10^{-2},1]$ |
| $p_9$ | $[1.9\cdot10^{-2},3.5\cdot10^{-1}]$ | $[1.9\cdot10^{-2},3.5\cdot10^{-1}]$ | $[1.9\cdot10^{-2},3.5\cdot10^{-1}]$ | no refinement |
| $p_{delay1}$ | $[4\cdot10^{-2},3\cdot10^{-1}]$ | no improvement | $[4\cdot10^{-2},3\cdot10^{-1}]$ | $[4\cdot10^{-2},2\cdot10^{-1}]$ |
| $p_{10}$ | $[0.12\cdot10^1,1.4\cdot10^1]$ | no improvement | $[0.12\cdot10^1,1.4\cdot10^1]$ | no refinement |
| $p_{11}$ | $[4.5\cdot 10^{-1},10^1]$ | — | $[4.5\cdot 10^{-1},10^1]$ | no refinement |
| $p_{delay2}$ | $[8\cdot10^{-2},8\cdot10^{-1}]$ | — | $[8\cdot10^{-2},8\cdot10^{-1}]$ | $[8\cdot10^{-2},3\cdot10^{-1}]$ |
| $p_{12}$ | $[6\cdot10^{-2},5.64]$ | — | $[6\cdot10^{-2},5.64]$ | $[6\cdot10^{-2},1]$ |
| $p_{13}$ | $[10^{-9},10^3]$ | — | $[10^{-9},10^3]$ | $[1,300]$ |

Table D.5: **Set-based estimation results for models describing trans-signaling.** $\underline{p_i}$ and $\overline{p_i}$ indicate the estimated lower and upper bounds for the individual parameters using an outer-bounding algorithm. The first column describes the corresponding parameter name. The second, third and fourth columns depict the set-based parameter estimation results for the initial (Fig. 8.2(b)), the reduced (Fig. 8.5(b)), and the calibrated model.

| Parameter | $[\underline{p_i},\overline{p_i}]$ initial model | $[\underline{p_i},\overline{p_i}]$ reduced model | $[\underline{p_i},\overline{p_i}]$ calibrated model |
|---|---|---|---|
| $p_3^{tr}$ | $[8.3{\cdot}10^{-2},5.5{\cdot}10^{-1}]$ | $[8.3{\cdot}10^{-2},5.5{\cdot}10^{-1}]$ | $[8.3{\cdot}10^{-2},5.5{\cdot}10^{-1}]$ |
| $p_4^{tr}$ | $[9.9{\cdot}10^{-4},10^{-2}]$ | $[9.9{\cdot}10^{-4},10^{-2}]$ | $[9.9{\cdot}10^{-4},10^{-2}]$ |
| $p_5$ | $[10^{-9},10^3]$ | $[10^{-9},10^3]$ | $[10^{-9},10^3]$ |
| $p_6$ | $[10^{-9},10^3]$ | $[10^{-9},10^3]$ | $[10^{-9},10^3]$ |
| $p_7$ | $[2{\cdot}10^{-3},9.9{\cdot}10^2]$ | $[3.5{\cdot}10^{-1},5.5{\cdot}10^2]$ | $[3.5{\cdot}10^{-1},5.5{\cdot}10^2]$ |
| $p_8$ | $[4{\cdot}10^{-4},9.9{\cdot}10^2]$ | $[2.6{\cdot}10^{-3},5.5{\cdot}10^2]$ | $[2.6{\cdot}10^{-3},5.5{\cdot}10^2]$ |
| $p_9$ | $[1.9{\cdot}10^{-2},3.5{\cdot}10^{-1}]$ | no improvement | $[1.9{\cdot}10^{-2},3.5{\cdot}10^{-1}]$ |
| $p_{delay1}$ | $[4{\cdot}10^{-2},3{\cdot}10^{-1}]$ | no improvement | $[4{\cdot}10^{-2},3{\cdot}10^{-1}]$ |
| $p_{10}$ | $[0.12{\cdot}10^1,1.4{\cdot}10^1]$ | no improvement | $[0.12{\cdot}\ 10^1,1.4{\cdot}10^1]$ |
| $p_{11}$ | $[4.5{\cdot}10^{-1},10^1]$ | - | $[4.5{\cdot}10^{-1},10^1]$ |
| $p_{delay2}$ | $[8{\cdot}10^{-2},8{\cdot}10^{-1}]$ | - | $[8{\cdot}10^{-2},8{\cdot}10^{-1}]$ |
| $p_{12}$ | $[6{\cdot}10^{-2},5.64]$ | - | $[6{\cdot}10^{-2},5.64]$ |
| $p_{13}$ | $[10^{-9},10^3]$ | - | $[10^{-9},10^3]$ |

Table D.6: **Set-based estimation results for models describing classic-signaling.** $\underline{p_i}$ and $\overline{p_i}$ indicate the estimated lower and upper bounds for the individual parameters using an outer-bounding algorithm. The first column describes the corresponding parameter name. The second, third and fourth columns depict the set-based parameter estimation results for the initial (Fig. 8.2(c)), the reduced (Fig. 8.5(c)), and the calibrated model.

| Parameter | $[\underline{p_i},p_i]$ initial model | $[\underline{p_i},p_i]$ reduced model | $[\underline{p_i},p_i]$ calibrated model |
|---|---|---|---|
| $p_1$ | $[10^{-7},9.9{\cdot}10^2]$ | $[10^{-6},2.7{\cdot}10^2]$ | $[10^{-6},2.7{\cdot}10^2]$ |
| $p_2$ | $[10^{-7},9.9{\cdot}10^2]$ | $[10^{-6},2.7{\cdot}\,10^2]$ | $[10^{-6},2.7{\cdot}10^2]$ |
| $p_3^{cl}$ | $[10^{-7},5{\cdot}10^1]$ | $[10^{-6},3.5{\cdot}10^1]$ | $[10^{-6},3.5{\cdot}10^1]$ |
| $p_4^{cl}$ | $[10^{-7},5{\cdot}10^1]$ | $[10^{-6},3.5{\cdot}10^1]$ | $[10^{-6},3.5{\cdot}10^1]$ |
| $p_5$ | $[10^{-9},10^3]$ | $[10^{-9},10^3]$ | $[10^{-9},10^3]$ |
| $p_6$ | $[10^{-9},10^3]$ | $[10^{-9},10^3]$ | $[10^{-9},10^3]$ |
| $p_7$ | $[1.1{\cdot}10^{-4},10^3]$ | $[2{\cdot}10^{-3},9.9{\cdot}10^2]$ | $[2{\cdot}10^{-3},9.9{\cdot}10^2]$ |
| $p_8$ | $[2{\cdot}10^{-5},10^3]$ | $[8{\cdot}10^{-5},9.9{\cdot}10^2]$ | $[8{\cdot}10^{-5},9.9{\cdot}10^2]$ |
| $p_9$ | $[1.3{\cdot}10^{-2},10^1]$ | no improvement | $[1.3{\cdot}10^{-2},10^1]$ |
| $p_{delay1}$ | $[10^{-3},1.1{\cdot}10^1]$ | no improvement | $[10^{-3},1.1{\cdot}10^1]$ |
| $p_{10}$ | $[7{\cdot}10^{-3},20.3{\cdot}10^1]$ | no improvement | $[7{\cdot}10^{-3},20.3{\cdot}10^1]$ |
| $p_{11}$ | $[6.4{\cdot}\,10^{-3},6.7{\cdot}10^1]$ | - | $[6.4{\cdot}\,10^{-3},6.7{\cdot}10^1]$ |
| $p_{delay2}$ | $[8{\cdot}10^{-4},5.3{\cdot}10^1]$ | - | $[8{\cdot}10^{-4},5.3{\cdot}10^1]$ |
| $p_{12}$ | $[10^{-3},4.4{\cdot}10^1]$ | - | $[10^{-3},4.4{\cdot}10^1]$ |
| $p_{13}$ | $[10^{-9},10^3]$ | - | $[10^{-9},10^3]$ |

## D.4 Determining valid parameters using Monte Carlo sampling

Due to relaxations, set-based outer approximations still include false positive parametrizations that lead to wrong model predictions. Therefore, we applied Monte Carlo sampling to obtain parameter samples that can reasonable represent the experimental data. To do so, we determined random parametrizations of 150 000 samples within the estimated outer bounds of all parameters using the MATLAB-based function random and a log2-uniform distribution. Furthermore, values for IL-6R$\alpha^{\text{Total}}$, gp130$^{\text{Total}}$ and STAT3$^{\text{Total}}$ were randomly determined within their experimentally measured uncertainty ranges, i. e. mean±STD. The obtained parametrizations were then tested whether they can represent the measurement data or not using the continuous-time models. 150 simulations that are in line with the data and that result in the lowest achievable quadratic distance between simulations and data were bundled into corridors for the quantities [(p)STAT3], [SOCS3 mRNA] and [SOCS3].

# Bibliography

[1] ILOG CPLEX 11.0 User's Manual. *Gentilly, France: ILOG SA*, 2007.

[2] S. Abe. *Support vector machines for pattern classification*, volume 53. Springer, 2005.

[3] H. Akaike. Information theory and an extension of the maximum likelihood principle. In *Breakthroughs in statistics*, pages 610–624. Springer, 1992.

[4] M. Ali, C. Storey, and A. Törn. Application of stochastic global optimization algorithms to practical problems. *Journal of Optimization Theory and Applications*, 95(3):545, 1997.

[5] J. Anderson and A. Papachristodoulou. On validation and invalidation of biological models. *BMC Bioinformatics*, 10(1):132, 2009.

[6] J. F. Apgar, J. E. Toettcher, D. Endy, F. M. White, and B. Tidor. Stimulus design for model selection and validation in cell signaling. *PLoS Computational Biology*, 4(2):e30, 2008.

[7] A. Atkinson. Developments in the design of experiments, correspondent paper. *International Statistical Review/Revue Internationale de Statistique*, pages 161–177, 1982.

[8] A. Atkinson, A. Donev, and R. Tobias. *Optimum experimental designs, with SAS*, volume 34. Oxford University Press, 2007.

[9] T. Bäck and H.-P. Schwefel. An overview of evolutionary algorithms for parameter optimization. *Evolutionary Computation*, 1(1):1–23, 1993.

[10] E. Balsa-Canto, A. A. Alonso, and J. R. Banga. Computational procedures for optimal experimental design in biological systems. *IET Systems Biology*, 2(4):163–172, 2008.

[11] E. Balsa-Canto, M. Peifer, J. R. Banga, J. Timmer, and C. Fleck. Hybrid optimization method with general switching strategy for parameter estimation. *BMC Systems Biology*, 2(1):26, 2008.

[12] S. Bandara, J. P. Schlöder, R. Eils, H. G. Bock, and T. Meyer. Optimal experimental design for parameter estimation of a cell signaling model. *PLoS Computational Biology*, 5(11):e1000558, 2009.

[13] P. Baran, S. Hansen, G. H. Waetzig, M. Akbarzadeh, L. Lamertz, H. J. Huber, M. R. Ahmadian, J. M. Moll, and J. Scheller. The balance of Interleukin (IL)-6, IL-6:soluble IL-6 receptor (IL-6R) and IL-6:sIL-6R:sgp130 complexes allows simultaneous classic and trans-signaling. *Journal of Biological Chemistry*, pages

6762–6775, 2018.

[14] J. F. Bard. *Practical bilevel optimization: algorithms and applications*, volume 30. Springer Science & Business Media, 2013.

[15] T. Barkhausen, T. Tschernig, P. Rosenstiel, M. van Griensven, R. Vonberg, M. Dorsch, A. Mueller-Heine, A. Chalaris, J. Scheller, and S. Rose-John. Selective blockade of interleukin-6 trans-signaling improves survival in a murine polymicrobial sepsis model. *Critical Care Medicine*, 39(6):1407, 2011.

[16] A. Bemporad, C. Filippi, and F. D. Torrisi. Inner and outer approximations of polytopes using boxes. *J. Comput. Geom.*, 27(2):151–178, 2004.

[17] A. Bemporad, A. Garulli, S. Paoletti, and A. Vicino. A bounded-error approach to piecewise affine system identification. *IEEE Transactions on Automatic Control*, 50(10):1567–1580, 2005.

[18] J. Berkson. Minimum chi-square, not maximum likelihood! *The Annals of Statistics*, pages 457–487, 1980.

[19] M. O. Besenhard, M. Jarzabek, A. C. O'Farrell, J. J. Callanan, J. H. Prehn, A. T. Byrne, and H. J. Huber. Modelling tumour cell proliferation from vascular structure using tissue decomposition into avascular elements. *Journal of Theoretical Biology*, 402:129–143, 2016.

[20] R. Beyaert, A. Cuenda, W. V. Berghe, S. Plaisance, J. Lee, G. Haegeman, P. Cohen, and W. Fiers. The p38/RK mitogen-activated protein kinase pathway regulates interleukin-6 synthesis response to tumor necrosis factor. *The EMBO Journal*, 15(8):1914, 1996.

[21] H. S. Bhat and N. Kumar. On the derivation of the Bayesian Information Criterion. *School of Natural Sciences, University of California*, 2010.

[22] B. K. Bhattacharya and G. T. Toussaint. An upper bound on the probability of misclassification in terms of matusita's measure of affinity. *Annals of the Institute of Statistical Mathematics*, 34(1):161–165, 1982.

[23] M. A. Blätke, A. Dittrich, C. Rohr, M. Heiner, F. Schaper, and W. Marwan. Jak/Stat signalling–an executable model assembled from molecule-centred modules demonstrating a module-oriented database concept for systems and synthetic biology. *Molecular BioSystems*, 9(6):1290–1307, 2013.

[24] S. Borchers, P. Rumschinski, S. Bosio, R. Weismantel, and R. Findeisen. A set-based framework for coherent model invalidation and parameter estimation of discrete time nonlinear systems. In *Proceeding of the IEEE Conference on Decision and Control*, pages 6786–6792, 2009.

[25] S. Borchers and R. Findeisen. Outlier detection for polynomial systems using semidefinite relaxations. In *Proceedings of the IFAC Symposium on Nonlinear Control Systems*, pages 761–766, 2013.

[26] S. Borchers, S. Freund, A. Rath, S. Streif, U. Reichl, and R. Findeisen. Identifi-

cation of growth phases and influencing factors in cultivations with AGE1. HN cells using set-based methods. *PLoS One*, 8(8):e68124, 2013.

[27] S. Borchers, P. Rumschinski, S. Bosio, R. Weismantel, and R. Findeisen. Model invalidation and system identification of biochemical reaction networks. In *Proceedings of the IFAC Vienna Conference on Mathematical Modelling*, pages 787–795, 2009.

[28] M. Boulanger, D. Chow, E. Brevnova, and K. Garcia. Hexameric structure and assembly of the interleukin-6/IL-6-receptor/gp130 complex. *Signal Transduction Knowledge Environment - Science*, 300(5628):2101, 2003.

[29] S. Boyd and L. Vandenberghe. *Convex optimization.* Cambridge University Press, 2004.

[30] J. Bracken and J. T. McGill. Defense applications of mathematical programs with optimization problems in the constraints. *Operations Research*, 22(5):1086–1096, 1974.

[31] J. Bracken and J. McGill. Production and marketing decisions with multiple objectives in a competitive environment. *Journal of Optimization Theory and Applications*, 24(3):449–458, 1978.

[32] A. P. Burgard, P. Pharkya, and C. D. Maranas. Optknock: a bilevel programming framework for identifying gene knockout strategies for microbial strain optimization. *Biotechnology and Bioengineering*, 84(6):647–657, 2003.

[33] L. Campbell, C. Chen, S. S. Bhagat, R. A. Parker, and A. J. Östör. Risk of adverse events including serious infections in rheumatoid arthritis patients treated with tocilizumab: a systematic literature review and meta-analysis of randomized controlled trials. *Rheumatology*, 50(3):552–562, 2010.

[34] L. Carius, P. Rumschinski, T. Faulwasser, D. Flockerzi, H. Grammel, and R. Findeisen. Model-based derivation, analysis and control of unstable microaerobic steady-states - Considering *Rhodospirillum rubrum* as an example. *Biotechnology and Bioengineering*, 111(4):734–747, 2014.

[35] B. Chachuat, B. Houska, R. Paulen, N. Peri'c, J. Rajyaguru, and M. E. Villanueva. Set-theoretic approaches in analysis, estimation and control of nonlinear systems. In *Proceedings of the IFAC Symposium on Advanced Control of Chemical Processes*, pages 981–995, 2015.

[36] K. Chaloner and I. Verdinelli. Bayesian experimental design: A review. *Statistical Science*, pages 273–304, 1995.

[37] Y. Chang and N. V. Sahinidis. Optimization of metabolic pathways under stability considerations. *Computers & Chemical Engineering*, 29(3):467–479, 2005.

[38] B. Colson, P. Marcotte, and G. Savard. An overview of bilevel optimization. *Annals of Operations Research*, 153(1):235–256, 2007.

[39] A. Compston and A. Coles. Multiple sclerosis. *Lancet*, 372(9648):1502–17, 2008.

[40] M. Cooney and K. McDonald. Optimal dynamic experiments for bioreactor model discrimination. *Applied Microbiology and Biotechnology*, 43(5):826–837, 1995.

[41] A. Cornish-Bowden. *Principles of enzyme kinetics*. Elsevier, 2014.

[42] L. M. Coussens and Z. Werb. Inflammation and cancer. *Nature*, 420(6917):860, 2002.

[43] C. D. Cruz, F. L. Neto, J. Castro-Lopes, S. B. McMahon, and F. Cruz. Inhibition of ERK phosphorylation decreases nociceptive behaviour in monoarthritic rats. *Pain*, 116(3):411–419, 2005.

[44] J. O. Dada and P. Mendes. Multi-scale modelling and simulation in systems biology. *Integrative Biology*, 3(2):86–96, 2011.

[45] T. J. DeGraba, J. M. Hallenbeck, K. D. Pettigrew, A. J. Dutka, and B. J. Kelly. Progression in acute stroke: value of the initial NIH stroke scale score on patient stratification in future trials. *Stroke*, 30(6):1208–1212, 1999.

[46] T. S. Deisboeck, Z. Wang, P. Macklin, and V. Cristini. Multiscale cancer modeling. *Annual Review of Biomedical Engineering*, 13:127–155, 2011.

[47] S. Dempe and S. Franke. On the solution of convex bilevel optimization problems. *Computational Optimization and Applications*, 63(3):685–703, 2016.

[48] H. Dette, V. B. Melas, and N. Strigul. Design of experiments for microbiological models. *Applied Optimal Designs*, page 137, 2005.

[49] A. Dittrich, T. Quaiser, C. Khouri, D. Görtz, M. Mönnigmann, and F. Schaper. Model-driven experimental analysis of the function of SHP-2 in IL-6-induced Jak/STAT signaling. *Molecular BioSystems*, 8(8):2119–2134, 2012.

[50] G. Dougherty. *Pattern Recognition and Classification: An Introduction*. Springer-Verlag New York, 2013.

[51] T. A. Endo, M. Masuhara, M. Yokouchi, R. Suzuki, H. Sakamoto, K. Mitsui, A. Matsumoto, S. Tanimura, M. Ohtsubo, H. Misawa, et al. A new protein containing an SH2 domain that inhibits JAK kinases. *Nature*, 387(6636):921–924, 1997.

[52] W. R. Esposito and C. A. Floudas. Global optimization for the parameter estimation of differential-algebraic systems. *Industrial & Engineering Chemistry Research*, 39(5):1291–1310, 2000.

[53] D. Faller, U. Klingmüller, and J. Timmer. Simulation methods for optimal experimental design in systems biology. *Simulation*, 79(12):717–725, 2003.

[54] M. Feldmann, F. Brennan, and R. Maini. Role of cytokines in rheumatoid arthritis. *Annual Review of Immunology*, 14(1):397–440, 1996.

[55] J. L. Figueroa, S. I. Biagiola, and O. E. Agamennoni. An approach for identification of uncertain wiener systems. *Mathematical and Computer Modelling*,

48(1):305–315, 2008.

[56] M. Fischer, J. Goldschmitt, C. Peschel, J. Brakenhoff, K. Kallen, A. Wollmer, J. Grötzinger, and S. Rose-John. A bioactive designer cytokine for human hematopoietic progenitor cell expansion. *Nature Biotechnology*, 15(2):142–145, 1997.

[57] R. A. Fisher. On an absolute criterion for fitting frequency curves. *Statistical Science*, 12(1):39–41, 1997.

[58] R. J. Flassig and K. Sundmacher. Optimal design of stimulus experiments for robust discrimination of biochemical reaction networks. *Bioinformatics*, 28(23):3089–3096, 2012.

[59] K. Franzen. Kinetics of IL-6-induced classic- and trans-signaling. Bachelor's Thesis, Otto-von-Guericke Universität, 2013.

[60] K. G. Gadkar, F. J. Doyle Iii, J. S. Edwards, and R. Mahadevan. Estimating optimal profiles of genetic alterations using constraint-based models. *Biotechnology and Bioengineering*, 89(2):243–251, 2005.

[61] K. G. Gadkar, R. Gunawan, and F. J. Doyle. Iterative approach to model identification of biological networks. *BMC Bioinformatics*, 6(1):155, 2005.

[62] C. Garbers, S. Heink, T. Korn, and S. Rose-John. Interleukin-6: designing specific therapeutics for a complex cytokine. *Nature Reviews Drug Discovery*, 2018.

[63] C. Gardiner. Stochastic methods. *Springer Series in Synergetics (Springer-Verlag, Berlin, 2009)*, 1985.

[64] B. Giese, C.-K. Au-Yeung, A. Herrmann, S. Diefenbach, C. Haan, A. Küster, S. Wortmann, C. Roderburg, P. C. Heinrich, I. Behrmann, et al. Long-term association of the cytokine receptor gp130 and the janus kinase Jak1 revealed by FRAP analysis. *Journal of Biological Chemistry*, 2003.

[65] J. Glover and F. Schweppe. Control of linear dynamic systems with set constrained disturbances. *IEEE Transactions on Automatic Control*, 16(5):411–423, 1971.

[66] A. M. Gonzalez-Angulo, B. T. Hennessy, and G. B. Mills. Future of personalized medicine in oncology: a systems biology approach. *Journal of Clinical Oncology*, 28(16):2777, 2010.

[67] I. Goodfellow, Y. Bengio, and A. Courville. *Deep learning*. MIT Press Cambridge, 2016.

[68] I. E. Grossmann. *Global Optimization in engineering design*, volume 9. Springer Science & Business Media, 2013.

[69] Gurobi. Gurobi optimizer reference manual. *URL: http://www. gurobi. com*, 2012.

[70] D. Guschin, N. Rogers, J. Briscoe, B. Witthuhn, D. Watling, F. Horn, S. Pellegrini, K. Yasukawa, P. Heinrich, and G. Stark. A major role for the protein tyrosine kinase JAK1 in the JAK/STAT signal transduction pathway in response to interleukin-6. *The EMBO Journal*, 14(7):1421–1429, 1995.

[71] R. M. Hauser and D. L. Featherman. *The process of stratification: Trends and analyses*. Elsevier, 2013.

[72] P. Heinrich, I. Behrmann, S. Haan, H. Hermanns, G. Müller-Newen, and F. Schaper. Principles of interleukin (IL)-6-type cytokine signaling and its regulation. *The Biochemical Journal*, 374(1):1–20, 2003.

[73] P. C. Heinrich, J. V. Castell, and T. Andus. Interleukin-6 and the acute phase response. *Biochemical Journal*, 265(3):621, 1990.

[74] M. Hibi, M. Murakami, M. Saito, T. Hirano, T. Taga, and T. Kishimoto. Molecular cloning and expression of an IL-6 signal transducer, gp130. *Cell*, 63(6):1149–1157, 1990.

[75] M. Honda, S. Yamamoto, M. Cheng, K. Yasukawa, H. Suzuki, T. Saito, Y. Osugi, T. Tokunaga, and T. Kishimoto. Human soluble IL-6 receptor: its detection and enhanced release by hiv infection. *The Journal of Immunology*, 148(7):2175–2180, 1992.

[76] J. C. Houtman, R. A. Houghtling, M. Barda-Saad, Y. Toda, and L. E. Samelson. Early phosphorylation kinetics of proteins involved in proximal TCR-mediated signaling pathways. *The Journal of Immunology*, 175(4):2449–2458, 2005.

[77] C. A. Hunter and S. A. Jones. Il-6 as a keystone cytokine in health and disease. *Nature Immunology*, 16(5):448, 2015.

[78] J. S. Hyams, G. D. Ferry, F. S. Mandel, J. D. Gryboski, P. M. Kibort, B. S. Kirschner, A. M. Griffiths, A. J. Katz, R. J. Grand, and J. T. Boyle. Development and validation of a pediatric Crohn's disease activity index. *Journal of Pediatric Gastroenterology and Nutrition*, 12(4):439–447, 1991.

[79] L. Jaulin and E. Walter. Set inversion via interval analysis for nonlinear bounded-error estimation. *Automatica*, 29(4):1053–1064, 1993.

[80] L. Jaulin. *Applied interval analysis: with examples in parameter and state estimation, robust control and robotics*, volume 1. Springer Science & Business Media, 2001.

[81] S. Julier, J. Uhlmann, and H. F. Durrant-Whyte. A new method for the nonlinear transformation of means and covariances in filters and estimators. *IEEE Transactions on Automatic Control*, 45(3):477–482, 2000.

[82] D. Kamimura, K. Ishihara, and T. Hirano. IL-6 signal transduction and its physiological roles: the signal orchestration model. In *Reviews of Physiology, Biochemistry and Pharmacology*, pages 1–38. Springer, 2003.

[83] A. R. Kan and G. T. Timmer. Stochastic global optimization methods part I:

Clustering methods. *Mathematical Programming*, 39(1):27–56, 1987.

[84] S. Karaman, R. G. Sanfelice, and E. Frazzoli. Optimal control of mixed logical dynamical systems with linear temporal logic specifications. In *Proceedings of the IEEE Conference on Decision and Control*, pages 2117–2122, 2008.

[85] S. Kirkpatrick, C. Gelatt J.R., and M. Vecchi. Optimization by simulated annealing. *Science*, 220(4598):671–680, 1983.

[86] H. Kitano. Computational systems biology. *Nature*, 420(6912):206–210, 2002.

[87] H. Kitano. *Foundations of systems biology*. MIT Press Cambridge, 2001.

[88] M. Korda, D. Henrion, and C. N. Jones. Inner approximations of the region of attraction for polynomial dynamical systems. In *Proceedings of the IFAC Symposium on Nonlinear Control Systems*, pages 534–539, 2013.

[89] C. Kreutz and J. Timmer. Systems biology: experimental design. *FEBS Journal*, 276(4):923–942, 2009.

[90] L. Kuepfer, U. Sauer, and P. Parrilo. Efficient classification of complete parameter regions based on semidefinite programming. *BMC Bioinformatics*, 8:12, 2007.

[91] Z. Kutalik, K.-H. Cho, and O. Wolkenhauer. Optimal sampling time selection for parameter estimation in dynamic pathway modeling. *Biosystems*, 75(1):43–55, 2004.

[92] V. R. Lang, M. Englbrecht, J. Rech, H. Nüsslein, K. Manger, F. Schuch, H.-P. Tony, M. Fleck, B. Manger, G. Schett, and J. Zwerina. Risk of infections in rheumatoid arthritis patients treated with tocilizumab. *Rheumatology*, 51(5):852–857, 2011.

[93] J. B. Lasserre. Semidefinite programming vs. LP relaxations for polynomial programming. *Mathematics of Operations Research*, 27(2):347–360, 2002.

[94] G. Lillacci and M. Khammash. Parameter estimation and model selection in computational biology. *PLoS Computational Biology*, 6(3):e1000696, 2010.

[95] L. Ljung. System identification. In *Signal analysis and prediction*, pages 163–173. Springer, 1998.

[96] C. Lo, M. Chen, M. Hsiao, S. Wang, C. Chen, S. Hsiao, J. Chang, T. Lai, S. Rose-John, and M. Kuo. IL-6 trans-signaling in formation and progression of malignant ascites in ovarian cancer. *Cancer Research*, 71(2):424–434, 2011.

[97] D. Lu and Q. Weng. A survey of image classification methods and techniques for improving classification performance. *International Journal of Remote Sensing*, 28(5):823–870, 2007.

[98] C. Lutticken, U. M. Wegenka, J. Yuan, J. Buschmann, C. Schindler, A. Ziemiecki, A. G. Harpur, A. F. Wilks, K. Yasukawa, T. Taga, et al. Association of transcription factor APRF and protein kinase Jak1 with the Interleukin-6

signal transducer gp130. *Science*, 263(5143):89–92, 1994.

[99] B. F. Manly. *Randomization, bootstrap and Monte Carlo methods in biology*, volume 70. CRC Press, 2006.

[100] J. Meiling. Quantitative Analyse der Unterschiede im IL-6-induzierten classic- und trans-signaling. Bachelor's Thesis, Otto-von-Guericke Universität, 2013.

[101] M. Milanese and G. Belforte. Estimation theory and uncertainty intervals evaluation in presence of unknown but bounded errors: Linear families of models and estimators. *IEEE Transactions on Automatic Control*, 27(2):408–414, 1982.

[102] M. Milanese and C. Novara. Set membership identification of nonlinear systems. *Automatica*, 40(6):957–975, 2004.

[103] C. Moles, P. Mendes, and J. Banga. Parameter estimation in biochemical pathways: a comparison of global optimization methods. *Genome Research*, 13(11):2467–2474, 2003.

[104] R. E. Moore. *Methods and applications of interval analysis*. SIAM, 1979.

[105] R. E. Moore and C. Yang. Interval analysis I. *Technical Document LMSD-285875, Lockheed Missiles and Space Division, Sunnyvale, CA, USA*, 1959.

[106] C. Moya, Z. Huang, P. Cheng, A. Jayaraman, and J. Hahn. Investigation of IL-6 and IL-10 signalling via mathematical modelling. *IET Systems Biology*, 5(1):15–26, 2011.

[107] J. Mülberg, H. Schooltink, T. Stoyan, M. Günther, L. Graeve, G. Buse, A. Mackiewicz, P. C. Heinrich, and S. Rose-John. The soluble interleukin-6 receptor is generated by shedding. *European Journal of Immunology*, 23(2):473–480, 1993.

[108] L. D. Muu and N. Van Quy. A global optimization method for solving convex quadratic bilevel programming problems. *Journal of Global Optimization*, 26(2):199–219, 2003.

[109] S. Nancey, N. Hamzaoui, D. Moussata, I. Graber, J. Bienvenu, and B. Flourie. Serum interleukin-6, soluble interleukin-6 receptor and crohns disease activity. *Digestive Diseases and Sciences*, 53(1):242–247, 2008.

[110] Y. Nesterov and A. Nemirovski. Interior-point polynomial algorithms in convex programming. *SIAM Studies in Applied Mathematics*, 13, 1994.

[111] J. Neyman and E. S. Pearson. On the problem of the most efficient tests of statistical hypotheses. In *Breakthroughs in statistics*, pages 73–108. Springer, 1992.

[112] J. R. O'dell. Therapeutic strategies for rheumatoid arthritis. *New England Journal of Medicine*, 350(25):2591–2602, 2004.

[113] A. C. of Rheumatology. Tocilizumab (Actemra). pages 1–3, 2012.

[114] N. Oreskes and K. Belitz. Philosophical issues in model assessment. *Model validation: Perspectives in hydrological science*, 23, 2001.

[115] J. S. Parker, M. Mullins, M. C. Cheang, S. Leung, D. Voduc, T. Vickery, S. Davies, C. Fauron, X. He, Z. Hu, et al. Supervised risk predictor of breast cancer based on intrinsic subtypes. *Journal of Clinical Oncology*, 27(8):1160, 2009.

[116] P. Parrilo. Semidefinite programming relaxations for semi-algebraic problems. *Mathematical Programming*, 96(2):293–320, 2003.

[117] P. A. Parrilo and S. Lall. Semidefinite programming relaxations and algebraic optimization in control. *European Journal of Control*, 9(2):307 – 321, 2003.

[118] J. A. Paulson, D. M. Raimondo, R. Findeisen, R. D. Braatz, and S. Streif. Guaranteed active fault diagnosis for uncertain nonlinear systems. In *Proceedings of the IEEE European Control Conference*, pages 926–931, 2014.

[119] H. Perfahl, H. M. Byrne, T. Chen, V. Estrella, T. Alarcón, A. Lapin, R. A. Gatenby, R. J. Gillies, M. C. Lloyd, P. K. Maini, M. Reuss, and M. R. Owen. Multiscale modelling of vascular tumour growth in 3D: the roles of domain size and boundary conditions. *PLoS One*, 6(4):e14790, 2011.

[120] M. Peters, G. Blinn, F. Solem, M. Fischer, K.-H. M. zum Büschenfelde, and S. Rose-John. In vivo and in vitro activities of the gp130-stimulating designer cytokine Hyper-IL-6. *The Journal of Immunology*, 161(7):3575–3581, 1998.

[121] P. Picton. *Introduction to Neural Networks*. Macmillan Education UK, London, 1994.

[122] D. Pietzko, D. Zohlnhöfer, L. Graeve, D. Fleischer, T. Stoyan, H. Schooltink, S. Rose-John, and P. Heinrich. The hepatic Interleukin-6 receptor. studies on its structure and regulation by phorbol 12-myristate 13-acetate-dexamethasone. *Journal of Biological Chemistry*, 268(6):4250–4258, 1993.

[123] J. Pintér. Global optimization in action. continuous and lipschitz optimization: Algorithms, implementation and applications, 1996.

[124] R. L. Plackett. Karl Pearson and the chi-squared test. *International Statistical Review/Revue Internationale de Statistique*, pages 59–72, 1983.

[125] P. Poncelet and P. Carayon. Cytofluorometric quantification of cell-surface antigens by indirect immunofluorescence using monoclonal antibodies. *Journal of Immunological Methods*, 85(1):65–74, 1985.

[126] D. Posada and T. R. Buckley. Model selection and model averaging in phylogenetics: advantages of Akaike Information Criterion and Bayesian approaches over likelihood ratio tests. *Systematic biology*, 53(5):793–808, 2004.

[127] S. Prajna. Barrier certificates for nonlinear model validation. *Automatica*, 42(1):117–126, 2006.

[128] T. Quaiser, A. Dittrich, F. Schaper, and M. Mönnigmann. A simple work flow for biologically inspired model reduction-application to early jak-stat signaling. *BMC Systems Biology*, 5.

[129] A. Raue, V. Becker, U. Klingmüller, and J. Timmer. Identifiability and observability analysis for experimental design in nonlinear dynamical models. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 20(4):045105, 2010.

[130] H. Reeh, N. Rudolph, U. Billing, H. Christen, S. Streif, E. Bullinger, M. Schliemann-Bullinger, R. Findeisen, F. Schaper, H. J. Huber, et al. Response to IL-6 trans-and IL-6 classic signalling is determined by the ratio of the IL-6 receptor $\alpha$ to gp130 expression: fusing experimental insights and dynamic modelling. *Cell Communication and Signaling*, 17(1):46, 2019.

[131] S. Rose-John and M. Neurath. IL-6 trans-signaling: the heat is on. *Immunity*, 20(1):2–4, 2004.

[132] N. Rudolph, P. Andonov, H. J. Huber, and R. Findeisen. Model-supported patient stratification using set-based estimation methods. In *Proceedings of the IFAC Symposium on Advanced Control of Chemical Processes*, pages 892–897, 2018.

[133] N. Rudolph, T. Meyer, K. Franzen, C. Garbers, F. Schaper, S. Streif, A. Dittrich, and R. Findeisen. A two-level approach for fusing early signaling events and long term cellular responses. In *Proceedings of the IFAC Symposium on Advanced Control of Chemical Processes*, pages 1228–1233, 2015.

[134] N. Rudolph, S. Streif, and R. Findeisen. Set-based experiment design for model discrimination using bilevel optimization. In *Proceedings of the Foundations of Systems Biology in Engineering Conference*, pages 295–299, 2016.

[135] P. Rumschinski, S. Streif, and R. Findeisen. Combining qualitative information and semi-quantitative data for guaranteed invalidation of biochemical network models. *International Journal of Robust and Nonlinear Control*, 22:1157–1173, 2012.

[136] P. Rumschinski, S. Borchers, S. Bosio, R. Weismantel, and R. Findeisen. Set-base dynamical parameter estimation and model invalidation for biochemical reaction networks. *BMC Systems Biology*, 4(1):69, 2010.

[137] P. Rumschinski, D. S. Laila, and R. Findeisen. Set-based parameter estimation for symmetric network motifs. *IFAC Proceedings Volumes*, 44(1):10454–10459, 2011.

[138] S. Sabach and S. Shtern. A first order method for solving convex bilevel optimization problems. *SIAM Journal on Optimization*, 27(2):640–660, 2017.

[139] A. Savchenko. *Efficient Set-based Process Monitoring and Fault Diagnosis*. PhD thesis, Shaker Verlag Aachen, 2017.

[140] A. Savchenko, P. Rumschinski, S. Streif, and R. Findeisen. Complete diagnosability of abrupt faults using set-based sensitivities. In *Proceeding of the IFAC Symposium on Fault Detection, Supervision and Safety of Technical Processes*, pages 860–865, 2012.

[141] R. E. Schapire. The boosting approach to machine learning: An overview. In *Nonlinear estimation and classification*, pages 149–171. Springer, 2003.

[142] J. Scheller, A. Chalaris, D. Schmidt-Arras, and S. Rose-John. The pro-and anti-inflammatory properties of the cytokine Interleukin-6. *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research*, 1813(5):878–888, 2011.

[143] M. Schilling, T. Maiwald, S. Hengl, D. Winter, C. Kreutz, W. Kolch, W. D. Lehmann, J. Timmer, and U. Klingmüller. Theoretical and experimental analysis links isoform-specific ERK signalling to cell fate decisions. *Molecular Systems Biology*, 5(1):334, 2009.

[144] A. Schneider, U. Klingmüller, and M. Schilling. Short-term information processing, long-term responses: Insights by mathematical modeling of signal transduction. *Bioessays*, 34(7):542–550, 2012.

[145] F. Schweppe. Recursive state estimation: Unknown but bounded errors and system inputs. *IEEE Transactions on Automatic Control*, 13(1):22–28, 1968.

[146] R. Seger and E. G. Krebs. The MAPK signaling cascade. *The FASEB Journal*, 9(9):726–735, 1995.

[147] H. D. Sherali and C. H. Tuncbilek. Comparison of two reformulation-linearization technique based linear programming relaxations for polynomial programming problems. *Journal of Global Optimization*, 10(4):381–390, 1997.

[148] A. Singh, A. Jayaraman, and J. Hahn. Modeling regulatory mechanisms in IL-6 signal transduction in hepatocytes. *Biotechnology and Bioengineering*, 95(5):850–862, 2006.

[149] D. Skanda and D. Lebiedz. An optimal experimental design approach to model discrimination in dynamic biochemical systems. *Bioinformatics*, 26(7):939–945, 2010.

[150] S. Sobotta, A. Raue, X. Huang, J. Vanlier, A. Jünger, S. Bohl, U. Albrecht, M. J. Hahnel, S. Wolf, N. S. Mueller, et al. Model based targeting of IL-6-induced inflammatory responses in cultured primary hepatocytes to improve application of the JAK inhibitor ruxolitinib. *Frontiers in Physiology*, 8:775, 2017.

[151] N. Stahl, T. G. Boulton, T. Farruggella, N. Y. Ip, S. Davis, B. A. Witthuhn, F. W. Quelle, O. Silvennoinen, G. Barbieri, S. Pellegrini, et al. Association and activation of Jak-Tyk kinases by CNTF-LIF-OSM-IL-6 beta receptor components. *Science*, 263(5143):92–95, 1994.

[152] S. Streif, A. Savchenko, P. Rumschinski, and R. Findeisen. ADMIT: a toolbox for guaranteed model invalidation, estimation and qualitative-quantitative modeling. *Bioinformatics*, 28(9):1290–1291, 2012.

[153] S. Streif, M. Karl, and R. Findeisen. Outlier analysis in set-based estimation for nonlinear systems using convex relaxations. In *Proceedings of the IEEE European Control Conference*, pages 2921–2926, 2013.

[154] S. Streif, F. Petzke, A. Mesbah, R. Findeisen, and R. D. Braatz. Optimal experimental design for probabilistic model discrimination using polynomial chaos. In *Proceedings of the IFAC World Congress*, pages 4103–4109, 2014.

[155] S. Streif, P. Rumschinski, D. Henrion, and R. Findeisen. Estimation of consistent parameter sets for continuous-time nonlinear systems using occupation measures and LMI relaxations. In *Proceedings of the IEEE Conference on Decision and Control*, pages 6379–6384, 2013.

[156] S. Streif, N. Strobel, and R. Findeisen. Inner approximations of consistent parameter sets by constraint inversion and mixed-integer programming. In *Proceedings of the IFAC Symposium of Computer Applications in Biotechnology*, pages 326–331, 2013.

[157] A. Stuart and A. R. Humphries. *Dynamical systems and numerical analysis*, volume 2. Cambridge University Press, 1998.

[158] J. F. Sturm. Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones. *Optimization Methods and Software*, 11(1-4):625–653, 1999.

[159] R. S. Sutton and A. G. Barto. *Reinforcement learning: An introduction.* MIT Press Cambridge, 1998.

[160] T. Tanaka, M. Narazaki, and T. Kishimoto. Therapeutic targeting of the Interleukin-6 receptor. *Annual review of pharmacology and toxicology*, 52:199–219, 2012.

[161] A. Tarantola. *Inverse problem theory and methods for model parameter estimation.* SIAM, 2005.

[162] F. Theis, S. Bohl, and U. Klingmüller. Theoretical analysis of time-to-peak responses in biological reaction networks. *Bulletin of Mathematical Biology*, 73(5):978–1003, 2011.

[163] A. Törn, M. M. Ali, and S. Viitanen. Stochastic global optimization: Problem classes and solution techniques. *Journal of Global Optimization*, 14(4):437–447, 1999.

[164] F. Vallentin. Symmetry in semidefinite programs. *Linear Algebra and its Applications*, 430(1):360–369, 2009.

[165] L. Vandenberghe and S. Boyd. Semidefinite programming. *SIAM Review*, 38(1):49–95, 1996.

[166] S. Waldherr, R. Findeisen, and F. Allgöwer. Global sensitivity analysis of biochemical reaction networks via semidefinite programming. In *Proceedings of the IFAC World Congress*, pages 9701–9706, 2008.

[167] J. Walpole, J. A. Papin, and S. M. Peirce. Multiscale computational models of complex biological systems. *Annual Review of Biomedical Engineering*, 15:137–154, 2013.

[168] E. Walter and L. Jaulin. Guaranteed characterization of stability domains via

set inversion. *IEEE Transactions on Automatic Control*, 39(4):886–889, 1994.

[169] R. Wcisło, W. Dzwinel, D. A. Yuen, and A. Z. Dudek. A 3-d model of tumor progression based on complex automata driven by particle dynamics. *Journal of Molecular Modeling*, 15(12):1517, 2009.

[170] O. Weiergräber, U. Hemmann, A. Küster, G. Müller-Newen, J. Schneider, S. Rose-John, P. Kurschat, J. P. Brakenhoff, M. H. Hart, S. Stabel, et al. Soluble human interleukin-6 receptor: expression in insect cells, purification and characterization. *European journal of biochemistry*, 234(2):661–669, 1995.

[171] A. D. Weston and L. Hood. Systems biology, proteomics, and the future of health care: toward predictive, preventative, and personalized medicine. *Journal of Proteome Research*, 3(2):179–196, 2004.

[172] H. S. Witsenhausen. *Minimax Controls of Uncertain Systems*. MIT press, 1966.

[173] S. Wormald, J.-G. Zhang, D. L. Krebs, L. A. Mielke, J. Silver, W. S. Alexander, T. P. Speed, N. A. Nicola, and D. J. Hilton. The comparative roles of suppressor of cytokine signaling-1 and-3 in the inhibition and desensitization of cytokine signaling. *Journal of Biological Chemistry*, 281(16):11135–11143, 2006.

[174] S. J. Wright and J. Nocedal. Numerical optimization. *Springer Science*, 35(67-68):7, 1999.

[175] Z. Xing, J. Gauldie, G. Cox, H. Baumann, M. Jordana, X.-F. Lei, and M. K. Achong. IL-6 is an antiinflammatory cytokine required for controlling local or systemic acute inflammatory responses. *Journal of Clinical Investigation*, 101(2):311, 1998.

[176] T. Yamane. Statistics: An introductory analysis. 1973.

[177] D. Zohlnhöfer, L. Graeve, S. Rose-John, H. Schooltink, E. Dittrich, and P. Heinrich. The hepatic interleukin-6 receptor down-regulation of the interleukin-6 binding subunit (gp80) by its ligand. *FEBS Letters*, 306(2):219–222, 1992.