Mathematical Analysis and Modeling of Signaling Networks

Dissertation
zur
Erlangung des Doktorgrades (Dr. rer. nat.)
der
Mathematisch-Naturwissenschaftlichen Fakultät
der
Rheinischen Friedrich-Wilhelms-Universität Bonn

von
Benjamin Engelhardt
aus
Ansbach

Bonn 2017

Angefertigt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Rheinischen Friedrich-Wilhelms-Universität Bonn

1. Gutachter:	Dr. Holger Fröhlich
2. Gutachter:	Prof. Dr. Andreas Weber
Tag der Promotion:	10.10.2017
Erscheinungsjahr:	2017

"Dimidium facti, qui coepit, habet:
sapere aude, incipe."

Quintus Horatius Flaccus (65 v. Chr. – 8 v. Chr.)

Acknowledgments

I would like to acknowledge all the people who continued to believe in my work and who shared this journey with me. First and foremost, I want to express my deepest gratitude to my principal adviser Dr. Holger Fröhlich for his guidance, support, and the immense amount of knowledge he agreed to share. His continuous optimism for this work and his invaluable experience guided me through all ups and downs during the last three years and allowed me to grow both as an individual and as a research scientist. I could not imagine a better adviser.

I am greatly indebted to Prof. Dr. Andreas Weber for his strong support and valuable advice. Without his tireless effort and his believe in my work it would not have been possible to achieve the results presented in this thesis.

I sincerely thank Prof. em. Dr. med. Klaus Mohr for being part of my doctoral committee, for his insightful comments, encouragement and his continuous support and effort.

Furthermore, I would like to thank Prof. Dr. Thomas Schultz for agreeing to review this thesis and to join the doctoral committee.

For the very close cooperation and all the outstanding and fruitful discussions we had over the past years I owe special words of gratitude to Prof. Dr. Maik Kschischo. His brilliant comments and suggestions greatly helped to enhance this thesis. He proved himself to be a genius mentor, colleague and – last but not least – a good friend.

I gratefully acknowledge the funding I received from the Deutsche Forschungsgemeinschaft (DFG) towards my PhD as part of the Research Training Group 1873. I also owe special thanks to Prof. Dr. med. Alexander Pfeifer for giving me the opportunity to visit the laboratory of Prof. Dr. George S. Baillie. I would like to express my gratitude to Prof. Dr. George S. Baillie for establishing this successful cooperation. It was an honor for me to work with him, Jane E. Findlay and Dr. Christina Elliot in Glasgow. I will never forget this very intensive but highly rewarding phase of my work. I would also like to thank all my group members at the B-IT, at Prof. Mohr's Lab and the Research Training Group 1873 for the warm working atmosphere and the interesting discussions, which helped me in many ways.

This thesis is dedicated to my wife Anna and parents Hans and Monika who always supported me and took my absence from many family gatherings with a smile.

Abstract

Motivation

Mathematical models are in focus of modern systems biology and increasingly important to understand and manipulate complex biological systems. At the same time, new and improved techniques in metabolomics and proteomics enhance the ability to measure cellular states and molecular concentrations. In consequence, this leads to important biological insights and novel potential drug targets. Model development in systems biology can be described as an iterative process of model refinement to match the observed properties. The resulting research cycle is based on a well-defined initial model and requires careful model revision in each step.

Accomplishments and Results

As an initial step, a stoichiometry-based mathematical model of the muscarinic acetylcholine receptor subtype 2 (M_2 receptor)-induced signaling in Chinese hamster ovary (CHO) cells was derived. To validate the obtained initial model based on spatially accessible, not necessarily time-resolved data, the novel constrained flux sampling (CFS) is proposed in this work. The thus verified static model was then translated into a dynamical system based on ordinary differential equations (ODEs) by incorporating time-dependent experimental data.

To learn from the errors of systems biological models, the dynamic elastic-net (DEN), a novel approach based on optimal control theory, is proposed in this thesis. Next, the Bayesian dynamic elastic-net (BDEN), a systematic, fully algorithmic method based on the Markov chain Monte Carlo method was derived, which allows to detect hidden influences as well as missed reactions in ODE-based models. The BDEN allows for further validation of the developed M_2 receptor-induced signaling pathway and thus provides evidence for the completeness of the obtained dynamical system.

Conclusion

This thesis introduces the first comprehensive model of the M_2 receptor-induced signaling in CHO cells. Furthermore, this work presents several novel algorithms to validate and correct static and dynamic models of biological systems in a semi-automatic manner. These novel algorithms are expected to simplify the development of further mathematical models in systems biology.

Contents

CHAPTER 1

Introduction

Even now, about 350 years after R. Hooke discovered empty contained spaces in cork and termed them cells, many biological phenomena on a micro- and molecular level remain unclear (Hooke, 1665). In contrast to the macroscopic properties of living systems, the view on molecular and even microscopic events is still restricted (Noble, 2012; Azeloglu and Iyengar, 2015). Today, there is no doubt that macroscopic biological properties, and many diseases, are ultimately caused by molecular interactions (Noble, 2012; Bruggeman and Westerhoff, 2007; Zhao and Iyengar, 2012). Although the continuous improvement of laboratory techniques has led to important insights into biological processes, there are still many unresolved questions. The improvement of laboratory techniques, however, has also sparked the emergence of various fields under the umbrella of life sciences (Blankenburg et al., 2009; Robyt and White, 1987; Cravatt et al., 2007). Starting with the form and structural features of organisms, the driving forces of molecular interactions are now in the focus of research (Bruggeman and Westerhoff, 2007; Noble, 2012; Palsson, 2006). Thus, the fields of metabolics, proteomics and, finally, genomics have become more and more important, with the overall aim to understand the molecular and thus non-observable mechanisms of life (Joyce and Palsson, 2006; Bruggeman and Westerhoff, 2007; Noble, 2012; Palsson, 2006; Sauer et al., 2007).

All these branches of modern life sciences focus on biological properties caused by molecular interactions, each at their own level of detail (Sauer et al., 2007; Noble, 2012; Aloy and Russell, 2006). One may think, for instance, of the interaction between the heartbeat and the respiratory rate or the influence of drugs on the cellular response caused by drug-receptor interactions (Yasuma and Hayano, 2004; Zhao and Iyengar, 2012). In the final analysis, there are interactions between several regions of deoxyribonucleic acid (DNA) as the carrier of the individuals' genetic information (Alberts et al., 2014). But this is only one part of the story. There are multifarious interactions across all levels of an organism (Noble, 2012). When considering the enormous number of cells in the human body, more than 10^{13} cells with about 30,000 genes, it becomes very clear how complex these interactions must be (Alberts et al., 2014). Furthermore, there are still a lot of yet unknown phenomena, which can often only be treated as random events (Saarinen et al., 2008). All this serves to emphasize that our view on biological systems, especially the human body and its cells, is restricted and only a certain part of these systems can be observed at the same time (Bruggeman and Westerhoff, 2007; Azeloglu and Iyengar, 2015; Aloy and Russell, 2006).

Pharmacology investigates the impact of drugs on the cell and the human body as the basis of pharmacy (Dale and Rang, 2011). Once the drug, a cocktail of specific molecules, enters the human body, e.g., by oral application, at some point it may be distributed throughout the whole body via the blood stream (Dale and Rang, 2011). The drug ultimately binds to structures, i.e., receptors, on the cell surface (Dale and Rang, 2011). This in turn initiates a flow of information involving single cells or cellular compartments (Dale and Rang, 2011; Alberts et al., 2014). Such signaling pathways consist of various proteins which interact in different ways (Aloy and Russell, 2006; Meier-Schellersheim et al., 2009; Alberts et al., 2014). The involved interactions and general protein structures are captured by the field of proteomics (Joyce and Palsson, 2006). In contrast, the field of metabolics focuses on chemical processes leading to intermediates and products of the metabolism (Joyce and Palsson, 2006). In turn, the distribution of proteins and thus the cellular structure is mainly determined by the genome (Sauer et al., 2007; Joyce and Palsson, 2006). Finally, genomics is the field of research dealing with the composition of the genome and, in consequence, gene expression (Sauer et al., 2007; Joyce and Palsson, 2006; Alberts et al., 2014).

This thesis deals with mathematical models of signaling pathways. A general signaling cascade based on protein-protein interactions is illustrated in Figure 1.1. Protein-protein interactions are a direct consequence of several mechanisms driven by chemical relations (Sauer et al., 2007; Joyce and Palsson, 2006; Berg et al., 2013). With respect to the degree of detail, the interactions between proteins and, in general, between proteins and other molecules are considered from different angles (Sauer et al., 2007; Joyce and Palsson, 2006; Aloy and Russell, 2006). Following this, protein-protein interactions affect the three-dimensional structure of the involved proteins, e.g., the binding of several proteins often leads to larger protein complexes (Aloy and Russell, 2006; Berg et al., 2013). Other protein-protein interactions may result in a change of the protein's chemical composition (Aloy and Russell, 2006; Berg et al., 2013). For instance, protein phosphorylation causes the association of an additional phosphate-residue (Aloy and Russell, 2006; Berg et al., 2013). Such slight modifications may significantly alter the behavior of the proteins (Aloy and Russell, 2006; Berg et al., 2013). In the context of signaling pathways, phosphorylation is the major kind of interaction (Aloy and Russell, 2006; Berg et al., 2013; Heinrich et al., 2002). It generally results in an activation of the target protein (Aloy and Russell, 2006; Berg et al., 2013; Heinrich et al., 2002). As mentioned above, there are about 30,000 genes in a single cell (Alberts et al., 2014). Each gene contains the genetic blueprint for at least one protein, leading to very complex protein networks (Alberts et al., 2014). Hence it becomes clear that only certain parts of protein-protein interaction networks can be considered for a detailed mathematical modeling of such systems (Bruggeman and Westerhoff, 2007; Noble, 2012; Azeloglu and Iyengar, 2015).

In the past, a lot of effort has been devoted to the characterization and description of proteins and their functions (Sauer et al., 2007; Aloy and Russell, 2006; Kanehisa et al., 2016). But even

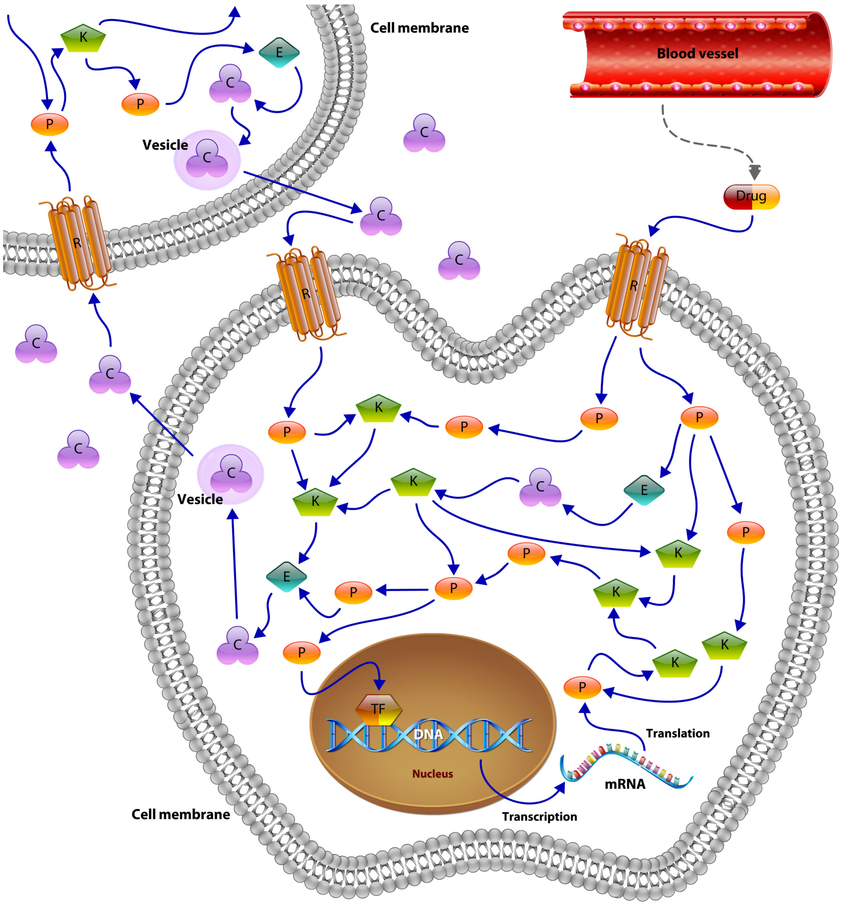


Figure 1.1: Illustration of a generic signaling pathway stimulated by a single drug and intercellular communication. Compounds (C), e.g., neurotransmitters or hormones, bind to specific receptors (R), which leads to the modulation of subsequent intracellular proteins (P). These compounds are released by other cells as a way of intercellular communication. As illustrated, the produced or modulated messenger compounds are uptaken by vesicles and then transported to the cell membrane. Afterwards, the compounds are released via exocytosis, i.e., the fusion of the vesicle with the cellular membrane. On the other hand, drugs distributed through, for instance, the blood stream, bind to specific receptors to induce signal cascades as one way to modulate intercellular communication. Once the signaling pathway is stimulated, there are multiple ways of signal transduction. Here, the general principle is explained using an arbitrary signaling pathway as an example. The blue arrows indicate the interactions or the dependencies between the involved actors. These interactions can be phosphorylation, protein-binding, catalytic or inhibitory events. Most proteins are further characterized, e.g., as kinases (K) or enzymes (E). Kinases are able to phosphorylate specific subsequent proteins which in turn can then be dephosphorylated by phosphatases. In contrast, enzymes can catalyze chemical reactions to produce or modulate other intracellular compounds, i.e., metabolites. In addition, proteins or other molecules can enter the nucleus to modulate DNA transcription. This is mediated via specific transcription factors (TF). In particular, the DNA gets transcripted into messenger ribonucleic acid (mRNA) which can then leave the nucleus and gets translated into the specific protein. The protein distribution is regulated individually in each cell via this mechanism. Please note that this illustration serves as a principal schematic illustration of a signaling pathway stimulated by a single drug and intercellular communication. There are many other mechanisms of inter- and intracellular signaling which are not described at this point.

if all proteins including their various functions were fully characterized, it would still not be possible to observe these structures and their interactions in sufficient detail in real-time (Bruggeman and Westerhoff, 2007; Noble, 2012; Azeloglu and Iyengar, 2015). However, today there exist a lot of techniques to quantify the number of proteins and their composition at fixed time points (Blankenburg et al., 2009; Robyt and White, 1987; Cravatt et al., 2007). More often than not the quantification is not possible in real-time for a huge number of proteins and thus the processes have to be stopped and conserved in order to proceed with the actual quantification (Robyt and White, 1987; Schröder et al., 2010; Cravatt et al., 2007). Aggravated by the fact that it is not possible to observe these interactions with the human eye, in contrast to other phenomena which occur on a microscopic level, this is where computational biology in general and systems biology in particular comes in (Meier-Schellersheim et al., 2009; Kitano, 2002).

Computational biology provides methods to capture otherwise unobservable compounds and mechanisms of such biological properties (Machado et al., 2011; Bruggeman and Westerhoff, 2007; Kitano, 2002). It provides tools to analyze experimental data even if the underlying cause is not observable or if the data is too complex for manual analysis (Bruggeman and Westerhoff, 2007; Meier-Schellersheim et al., 2009; Kitano, 2002). Determined by the nature of the accessible experimental data, the gained insights are often limited. When studying protein-protein interactions, often only the amount of active and inactive proteins is measurable (Bruggeman and Westerhoff, 2007; Meier-Schellersheim et al., 2009; Kitano, 2002). Therefore, direct assumptions regarding the underlying mechanisms are not possible. Nevertheless, it is possible to make relative statements about correlations, e.g., whether an increase in one protein concentration affects the concentration of another protein or not, even though the cause remains unclear (Machado et al., 2011; Bruggeman and Westerhoff, 2007; Kitano, 2002). Here, computational biology provides multifunctional methods to further interpret these relations (Bruggeman and Westerhoff, 2007; Meier-Schellersheim et al., 2009; Machado et al., 2011; Kitano, 2002).

This thesis aims for a detailed mechanistic modeling of protein-protein interaction networks to provide a physiological explanation for observed biological properties. In more detail a mathematical model for the muscarinic acetylcholine receptor subtype 2 (M_2 receptor)-induced signaling cascade is developed. Recent in vitro studies outlined an amplification of the drug-induced whole-cell response, as drafted in Figure 1.2 (Schrage et al., 2013). The variation of the optical density of the cell caused by dynamic mass redistribution (DMR) is a common measure for characterizing affinities and the efficiency of a given drug (Schrage et al., 2013; Schröder et al., 2011). So far, there is no explanation of the observed amplification, although it may lead to novel therapeutic approaches (Schrage et al., 2013; Schröder et al., 2011). The regulators of G protein signaling (RGS) identified in this work are well known for their therapeutic potential in the context of heart diseases (Kimple et al., 2011). Besides others, a lack of RGS4 results in a decrease of the heart rate in mice, which underlines the therapeutic potential of this class of RGS (Kimple et al., 2011).

In addition to the developed M_2 receptor-induced signaling pathway, several methodological innovations are presented that are of general interest for model development in systems

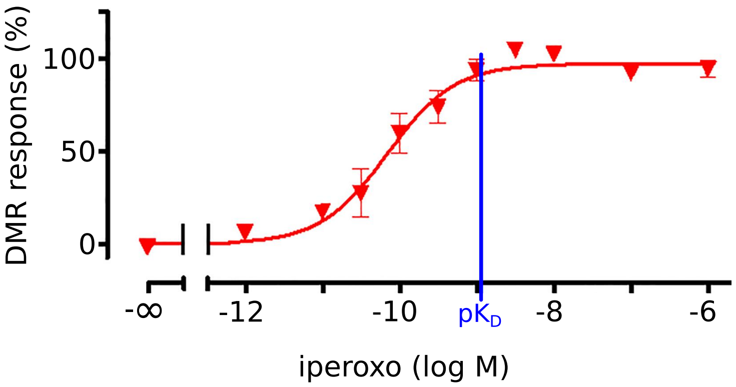


Figure 1.2: Ligand-induced whole-cell response in CHO cells. The cells were stimulated with different iperoxo concentrations. The blue line indicates the 50% receptor binding level (pK$_D$). The curve clearly suggests a concentration-based signal-amplification. In addition, a time-dependent amplification of the ligand-induced DMR response was observed (Engelhardt et al., 2017; Schrage et al., 2013).

biology (Engelhardt et al., 2017; Kitano, 2002; Azeloglu and Iyengar, 2015). Once the first sketch of the assumed protein-protein interaction network has been derived, an experimental validation of the model is required (Engelhardt et al., 2017; Kitano, 2002; Azeloglu and Iyengar, 2015). Such data should ideally reflect all relevant components of the model in a time- and space-resolved manner (Klipp and Liebermeister, 2006). In practice, this is often difficult to achieve due to technical, financial and time restrictions. Hence, it is desirable to first check the principal feasibility of the model to reproduce some expected key properties (Engelhardt et al., 2017; Kitano, 2002; Azeloglu and Iyengar, 2015). This way, the driving and thus important proteins of the network can be localized and preferably measured, especially in cases were resources are scarce and only a small number of the proteins are experimentally accessible (Engelhardt et al., 2017; Kitano, 2002; Azeloglu and Iyengar, 2015). In order to address these issues, I proposed a novel technique called constrained flux sampling (CFS) which allows for the generation of a first hypothesis and the ranking of the involved structures and players regarding their importance (Engelhardt et al., 2017). This allows for an improved experimental planning with respect to the derived main actors (Engelhardt et al., 2017; Kitano, 2002; Azeloglu and Iyengar, 2015).

The CFS analyzes the elementary flux modes (EFMs), the extreme pathways and the flux sampling to reveal the principal behavior of the system (Engelhardt et al., 2017). In this work, this allowed to identify biologically important subnetworks, which have been described in detail in literature (Engelhardt et al., 2017). Moreover, CFS allowed to confirm that our developed mathematical model can reproduce the experimentally observable increase in cyclic adenosine monophosphate (cAMP) production after receptor stimulation, which affects the cytoskeleton structure and, in consequence, changes the optical density of the cells (Engelhardt et al., 2017). This demonstrates that mathematical tools developed for metabolic network analysis can also be applied to mixed metabolic and signaling models (Engelhardt et al., 2017). This is very helpful for performing a priori model analyses with little effort and for the generation of hypotheses for further research (Engelhardt et al., 2017; Kitano, 2002; Azeloglu and Iyengar, 2015).

For further validation and to achieve a better understanding of the temporal aspect of the signaling process, a series of quantitative protein measurements were conducted. With the help of the gathered data, it was possible to extend and reformulate the original model. The refined mathematical model (formulated as a system of ordinary differential equations (ODEs)) reflected the observed temporal behavior of the M_2 receptor-dependent signaling cascade.

In the context of ODE-based mathematical model development, a frequent concern is that, on the one hand, all relevant molecules and processes are included and, on the other hand, there are neither wrongly modeled nor missing molecular interactions. In order to address this general problem within this thesis, a novel algorithm called dynamic elastic-net (DEN) is proposed. The DEN uses both an initial ODE system and experimental measurements as input and estimates which molecules in the ODE system are most likely influenced by factors that are not captured in the model (Engelhardt et al., 2016). This allows to estimate hidden influences at specific points in the system (Engelhardt et al., 2016). The methodology is based on optimal control theory (OCT) and can be applied to any mathematical system of ODEs (Engelhardt et al., 2016). Therefore, the proposed approach could be applied to other fields of natural science as well (Engelhardt et al., 2016). As it is common for deterministic approaches, the DEN does not include uncertainties of the estimates. To address this issue, a full Bayesian extension is developed in this thesis. The Bayesian dynamic elastic-net (BDEN) takes the uncertainties of the obtained estimates into account and leads to more reliable solutions (**?**). Moreover, the BDEN is principally able to detect wrong and missing reactions in ODE-based biological models (**?**).

Altogether, a first comprehensive molecular ODE-based model of the M_2 receptor-induced DMR in CHO cells is presented and analyzed in this work (Engelhardt et al., 2017). For this purpose, I developed a novel approach to investigate signal transduction based on static models (Engelhardt et al., 2017). Furthermore, a novel concept for detecting hidden influences as well as missed and wrong interactions using measured dynamics of system internal variables is presented (Engelhardt et al., 2016). Based on these ideas, a deterministic and Markov chain Monte Carlo (MCMC)-based algorithm was developed, which predicts unknown inputs of hidden as well as observed variables (**?**).

In Chapter 2, this thesis first provides a general overview of the field of systems biology and related topics. Then, Chapter 3 introduces the motivating G protein (GP)-induced signaling and the employed experimental techniques. Next, the fundamental methodology of the developed techniques is explained in detail in Chapter 4. Starting with Chapter 5, the results and developed methods are explained. In particular, Chapter 5 describes the developed CFS as a method to analyze the static model of the GP-induced signaling in CHO cells. In addition, first conclusions about the signal amplification mechanism in CHO cells are drawn. Chapter 6 details the DEN as a novel approach to learn from the errors of dynamical biological systems. Following this, the BDEN is discussed in Chapter 7 as a Bayesian extension of the DEN. In Chapter 8, the dynamic model of the GP-induced signaling in CHO cells and the drawn conclusions are presented. The final conclusions are then drawn in Chapter 9.

CHAPTER 2

Modeling Biological Systems

Biological processes are the result of a complex interplay of various components on different levels (Kestler et al., 2008; Bruggeman and Westerhoff, 2007; Noble, 2012). Consequently, a systematic interdisciplinary strategy is necessary to provide a global view on biological processes (Kestler et al., 2008; Bruggeman and Westerhoff, 2007; Noble, 2012). At the beginning, biological phenomena were largely considered separately and investigated by independent branches of bioscience (Bruggeman and Westerhoff, 2007). Molecular biosciences, for instance, focused on the function of single molecules and did not consider the interplay between individual molecules in a broader context (Bruggeman and Westerhoff, 2007). But with new and improved insights into biological systems, the involved processes were considered from a more global perspective and thus the focus shifted towards holistic approaches (Bruggeman and Westerhoff, 2007). This progress was facilitated by a dramatic improvement of measurement techniques in metabolics, genomics and proteomics, such as mass spectrometry, quantitative western blotting (WB), DNA-sequencing and other high-throughput technologies (Blankenburg et al., 2009; Robyt and White, 1987; Cravatt et al., 2007). The new interdisciplinary and holistic concept led to the development of the field of systems biology (Kestler et al., 2008; Bruggeman and Westerhoff, 2007; Noble, 2012). Systems biology aims for a global view on biological processes in well-defined systems (Kestler et al., 2008; Bruggeman and Westerhoff, 2007; Noble, 2012). Until 2004, when 99% of the euchromatic sequence of the human genome, including about 2.85 billion nucleotides, were successfully uncovered, systems biology mainly dealt with protein-protein interactions and the metabolism on different cellular levels (Lander et al., 2001; Consortium, 2004; Hood and Galas, 2003). With the decoding of the human genome, gene regulation and gene expression became important topics in systems biology as well (Lander et al., 2001; Consortium, 2004; Eisenberg et al., 2000). The complexity and diversity of biological systems requires an interdisciplinary and comprehensive approach (Aloy and Russell, 2006; Sauer et al., 2007). Thus, many concepts from physics and mathematics have been adapted and applied to connect and uncover cellular mechanisms (Kestler et al., 2008; Bruggeman and Westerhoff, 2007; Noble, 2012; Sauer et al., 2007).

In contrast to bioinformatics and other branches of computational biology, system biology attempts to uncover the mechanisms behind biological processes rather than to purely predict the behavior of biological systems (Bruggeman and Westerhoff, 2007). Although bioinformatics is able to reveal the players in complex biological processes, e.g., in the form of gene expression or protein concentrations, the detailed mechanisms remain largely unknown, which means that we are able to observe the effect but not the cause (Aloy and Russell, 2006; Sauer et al., 2007). Hence the fundamental challenge remains: to understand the underlying mechanisms and functions on a detailed and holistic level (Sauer et al., 2007). Therefore, the development of novel tools to predict the mode of action of the observed structures is highly desirable (Meier-Schellersheim et al., 2009; Hübner et al., 2011; Kahm et al., 2012). Commonly, systems biology translates biological processes into mathematical models to reflect their driving mechanisms (Bruggeman and Westerhoff, 2007; Hübner et al., 2011). Models in systems biology are usually composed of proteins, metabolites and genes which may interact on different cellular levels (Noble, 2012). Figure 2.1 illustrates the systematic hierarchy of biological systems (Noble, 2012). Although the questionable central dogma of molecular biology still persists, the main reason why genes are located at the bottom of this hierarchy is the fact that they represent the so-far smallest known unit (Noble, 2012; Crick, 1970). However, the assumption that the genome solely controls or determines the whole system has been proven obsolete (Noble, 2012; Crick, 1970; Alberts et al., 2014). Even though gene mutation may change the behavior of an organism, these mechanisms are still triggered by epigenetic controls or other factors on various levels of the organism (Noble, 2012; Crick, 1970; Alberts et al., 2014). On these grounds, the valid question arises whether genetic variations are the cause of biological phenomena or if they are the result of a change in the metabolism triggered by the environment (Noble, 2012; Crick, 1970; Alberts et al., 2014). Much effort has been devoted to the description of the mechanisms of these complex networks involving proteomics, metabolics and genomics (Meier-Schellersheim et al., 2009; Hübner et al., 2011; Chuang et al., 2010).

In systems biology, the common distinction is between top-down and bottom-up strategies (Bruggeman and Westerhoff, 2007). The latter starts with a comparatively huge amount of data and knowledge and tries to distill the driving mechanisms through reducing the complexity (Bruggeman and Westerhoff, 2007). In contrast, bottom-up strategies start with a comparatively small model and detailed knowledge and extent this model until the biological process in question is completely described by the model (Bruggeman and Westerhoff, 2007). Besides the modeling strategy, the kind of behavior that is observed is crucial for choosing the right approach (Kestler et al., 2008; Bruggeman and Westerhoff, 2007; Sauer et al., 2007). Even if a complete biological model includes all spatio-temporal interactions for a given task, it is often sufficient to model only the important part of the system (Bruggeman and Westerhoff, 2007; Sauer et al., 2007; Azeloglu and Iyengar, 2015). In many cases the model is restricted to distinguished states of the system because the complexity of the model increases with its dynamics (Bruggeman and Westerhoff, 2007; Sauer et al., 2007). Consequently, a clear distinction between static and dynamic biological processes is necessary (Bruggeman and Westerhoff, 2007; Sauer et al., 2007).

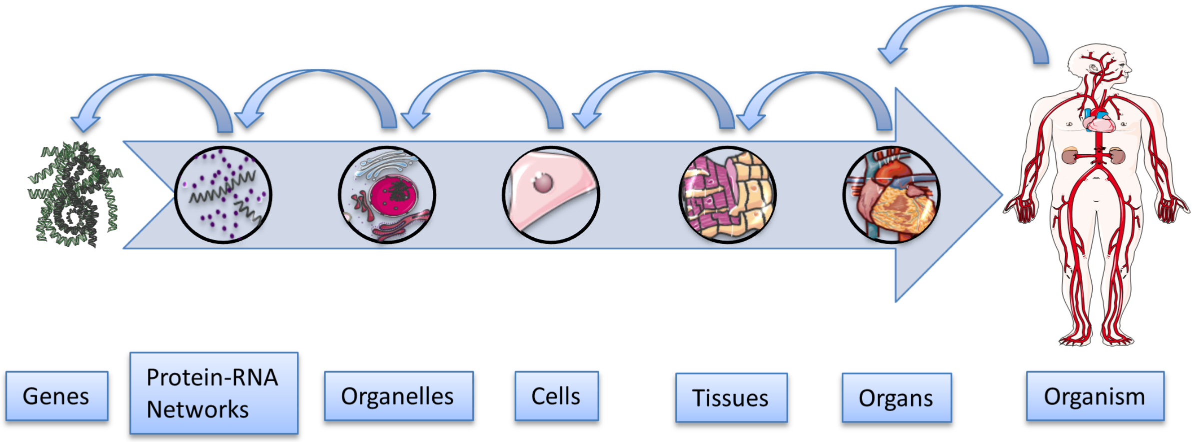


Figure 2.1: General hierarchy in systems biology. Protein-RNA networks, regulated by genes, are grouped in certain compartments or cell organelles which are the building blocks of cells. Interacting cells, such as cardiomyocytes in the heart, build tissues, e.g., the cardiac muscle. Different types of tissues then form certain organs such as the heart which is only a part of the entire organism. According to this natural hierarchy, initially, systems biology considered genes to be the main unilateral regulatory entity of living systems. Modern systems biology instead assumes that all components of the organism bilaterally interact on all levels. However, for modeling purposes, it makes sense to start either at the top or at the bottom of the illustrated hierarchy to successively reduce or increase the complexity of the system in a meaningful manner (Noble, 2012; Alberts et al., 2014; Crick, 1970).

This work focuses on static and dynamic mechanistic models of chemical reaction networks (CRNs), adopting a bottom-up approach. In contrast to, for instance, forecasting models, mechanistic models are not only capable to predict the behavior of biological systems but also to reveal their underlying mechanisms (Meier-Schellersheim et al., 2009; Hübner et al., 2011; Chuang et al., 2010). Strictly speaking, solely predictive models are not systems biology models because they only forecast the behavior of the system rather than reveal emergent properties (Kestler et al., 2008; Bruggeman and Westerhoff, 2007). However, the conceptual borders of systems biology are not clearly stated in literature (Aloy and Russell, 2006; Chuang et al., 2010; Hübner et al., 2011). Therefore, the term systems biology has different meanings in different communities (Aloy and Russell, 2006). Because of its origin, systems biology seeks to understand the mechanisms of living organic systems based on explanatory models and not to build forecasting models with large predictive power but without the ability to supply detailed mechanistic insights (Bruggeman and Westerhoff, 2007; Sauer et al., 2007; Noble, 2012; Azeloglu and Iyengar, 2015).

2.1 Modeling Strategies

The sheer complexity of biological systems and their interaction requires a practical solution strategy to uncover the underlying mechanisms (Azeloglu and Iyengar, 2015). A trade-off between complexity and reliability is essential to avoid oversimplification on the one hand

and overcomplication on the other (Azeloglu and Iyengar, 2015; Borkovich and Ebbole, 2010; Bruggeman and Westerhoff, 2007). Oversimplified models can deviate substantially from reality by ignoring important details or involving abstracted and thus unquantifiable parameters (Azeloglu and Iyengar, 2015). In contrast, the number of selectable variables in overcomplicated models is higher than the number of data points and thus they can insufficiently be derived (Ashyraliyev et al., 2009). Both cases result in unrealistic and unreliable models (Azeloglu and Iyengar, 2015). This can be addressed, at its extreme, from two different angles, i.e., the bottom-up and top-down perspective (Borkovich and Ebbole, 2010; Bruggeman and Westerhoff, 2007). As discussed, bottom-up approaches start with comprehensive models which then get extended and, vice versa, top-down approaches start with large settings, often based on a huge amount of data, which then get reduced (Borkovich and Ebbole, 2010; Bruggeman and Westerhoff, 2007). However, gene-regulatory networks are influenced by processes occurring at a proteomic level (Aloy and Russell, 2006). It is therefore important to have detailed knowledge about the mechanisms at the proteomic level to understand large-scale genomics data (Aloy and Russell, 2006). Mechanisms at the proteomic level are based on protein structures, and thus detailed knowledge about the underlying three-dimensional structure of proteins is necessary to fully understand biological processes (Aloy and Russell, 2006).

Although top-down approaches are commonly referred to as phenomenological models, the concept of reduction until the underlying mechanisms become clear can be transferred to pure mechanistic models based on proteomics (Bruggeman and Westerhoff, 2007). Here, a similar question arises: whether to start with complete but complex or comprehensive but simplified models (Hübner et al., 2011; Bruggeman and Westerhoff, 2007).

2.1.1 Top-down Strategies

Since the basis of top-down modeling is mostly phenomenological, the underlying mechanisms are unclear (Bruggeman and Westerhoff, 2007). The amount of prior knowledge is often very restricted such that not even relationships between proteins are known (Borkovich and Ebbole, 2010; Bruggeman and Westerhoff, 2007). However, large-scale mechanistic models or correlation-based models can be a starting point for top-down modeling (Borkovich and Ebbole, 2010; Bruggeman and Westerhoff, 2007). The common principle of such modeling approaches is the idea that similar changes of components indicate a functional relation (Borkovich and Ebbole, 2010). Thereby, system properties distilled from, for instance, proteom or transcriptome data, are mapped to a mechanistic model with a certain complexity (Borkovich and Ebbole, 2010; Bruggeman and Westerhoff, 2007). Top-down strategies take the entirety of large-scale networks into account and thus are able to predict completely unknown interactions (Heinrich and Schuster, 1996; Choffnes et al., 2011; Bruggeman and Westerhoff, 2007). Yet the basic mechanisms will remain hidden without further experimental observations (Heinrich and Schuster, 1996; Choffnes et al., 2011; Bruggeman and Westerhoff, 2007). In situations where the underlying causality is completely unclear or just too complex for detailed modeling, top-down approaches are often the method of choice (Heinrich and Schuster, 1996; Choffnes et al., 2011; Bruggeman and Westerhoff, 2007).

Top-down approaches require a clear biological question to choose the right experimental design in order to gain a large and information-rich data set (Borkovich and Ebbole, 2010). No prior knowledge regarding detailed interactions and the involved processes is necessary (Borkovich and Ebbole, 2010). The experimental data sets used for top-down approaches are usually very large and thus efficient statistical tools must be used to discover behavioral patterns and functional clusters (Bruggeman and Westerhoff, 2007). In consequence, the underlying biological process can be uncovered with increasing level of detail (Bruggeman and Westerhoff, 2007). Starting with potentially complete, often genome-wide data, the biological process can be successively revealed until the mechanistic details of the observed properties are fully known (Bruggeman and Westerhoff, 2007).

Besides their benefits, top-down approaches also have serious disadvantages, not the least because causality of the data has to be assumed (Choffnes et al., 2011; Bruggeman and Westerhoff, 2007). Hence, mechanisms which are well-described by correlations are often far away from reality (Bruggeman and Westerhoff, 2007). Thus, the findings have to be verified experimentally. As recently demonstrated, it is very likely to arbitrarily find spurious structures in big data sets as long as the data set is sufficiently large (Calude and Longo, 2016). In order to successfully uncover the functional and biochemical mechanisms of a biological system, it is essential to develop a detailed model of the phenomena under investigation (Meier-Schellersheim et al., 2009; Azeloglu and Iyengar, 2015; Aloy and Russell, 2006).

2.1.2 Bottom-up Strategies

As an alternative to top-down strategies, bottom-up strategies start with restricted but detailed mechanistic, mostly biochemical networks, with the intention to eventually obtain a complete model of the system (Bruggeman and Westerhoff, 2007). This requires prior knowledge and prior assumptions about the underlying mechanisms (Bruggeman and Westerhoff, 2007). Starting from a part of the entire model, e.g., a well-known subsystem, the model is iteratively extended through the incorporation of more and more details (Bruggeman and Westerhoff, 2007). In consequence, all components and mechanisms which are not or only partially known are neglected in this approach (Borkovich and Ebbole, 2010). The question arises how to extend the model at a certain level to avoid the inclusion of unnecessary details while at the same time including all important mechanisms (Bruggeman and Westerhoff, 2007). This is in stark contrast to top-down approaches, where the main challenge is to keep all important mechanisms while removing unnecessary details (Borkovich and Ebbole, 2010; Bruggeman and Westerhoff, 2007).

In bottom-up approaches, the required information about the involved kinetics and physicochemical properties of the components are composed in detailed models (Heinrich and Schuster, 1996; Choffnes et al., 2011; Bruggeman and Westerhoff, 2007). Gathering meaningful data reflecting the properties of the system and the involved mechanisms is crucial for such detailed models (Hübner et al., 2011; Bruggeman and Westerhoff, 2007). Today, information about specific mechanisms and interactions is stored in several databases which allows for the adaption of concrete models to the observed properties (Kanehisa et al., 2016; Juty et al., 2015). Information about the specific parameters is nevertheless rarely available because the conditions under which the data was collected may vary from the conditions under which the parameters are estimated (Azeloglu and Iyengar, 2015). It is therefore important to reach a compromise between complexity and observability while at the same time avoiding the neglection of fundamental biological mechanisms which might lead to wrong conclusions (Azeloglu and Iyengar, 2015). On the other hand, it is also vital to avoid overcomplication and therefore weak predictions (Azeloglu and Iyengar, 2015). Parameter estimation is often the only way to determine the parameters of such models (Bruggeman and Westerhoff, 2007; Azeloglu and Iyengar, 2015). In consequence, bottom-up studies might show different levels of accuracy (Bruggeman and Westerhoff, 2007; Azeloglu and Iyengar, 2015). In this context, it is questionable if biologically exact kinetics, including thermodynamically feasible and accurate parameters, are essential for a realistic model (Bruggeman and Westerhoff, 2007; Azeloglu and Iyengar, 2015). An abstraction of the exact kinetics, including an approximation of the involved parameters, is often sufficient to describe the properties of the biological system and to uncover new mechanisms (Ashyraliyev et al., 2009; Azeloglu and Iyengar, 2015).

To balance the advantages and disadvantages of the discussed approaches, a combination of both strategies has been proposed (Meier-Schellersheim et al., 2009). This involves combining a large-scale network with both correlation-based interactions and a detailed mechanistic model to end up with a holistic model of the whole organism (Meier-Schellersheim et al., 2009; Noble, 2012). Both the entire network and the driving mechanisms are considered when adopting this strategy (Meier-Schellersheim et al., 2009). As an extension of this concept, multi-scale modeling unifies models of different levels and scales (Meier-Schellersheim et al., 2009; Noble, 2012). For the human pathogen *Mycoplasma genitalium*, an organism with one of the smallest genomes, a model including all components and interactions has been proposed in 2012 (Karr et al., 2012). The complex model contains different levels, starting with the genome, and is able to simulate the whole life cycle of *Mycoplasma genitalium* (Karr et al., 2012). The idea behind the combination of both strategies is to combine different aspects of an organism, such as genome information, protein interaction, inter-cellular signaling and organelle mechanics, into one predictable detailed mechanistic model with maybe probabilistic components (Meier-Schellersheim et al., 2009). Such models then allow for more realistic in silico knock-out experiments and drug-target prediction (Meier-Schellersheim et al., 2009; Kitano, 2002; Scheidel et al., 2016). In addition, an alternative promising strategy of increasing popularity is the combined integration of various inherently different data sources, e.g., genomics, metabolics, proteomics and literature-based knowledge (Sauer et al., 2007; Praveen and Fröhlich, 2013; Bruggeman and Westerhoff, 2007).

2.2 Research Cycle

Mathematical modeling can be interpreted as a continuous process of model revision ideally starting with prior knowledge and data, as illustrated in Figure 2.2 (Azeloglu and Iyengar, 2015; Kitano, 2002). Depending on the quality and quantity of the available data and knowledge, a bottom-up or top-down strategy may be chosen (Meier-Schellersheim et al., 2009; Kitano, 2002). In any case, a prior model of the considered biological process is required to explain the given observations (Kitano, 2002). This can be a detailed biochemical reaction network, which is often the case for bottom-up approaches, or a more general whole-cell model or patient model which allows for the identification of the involved subnetworks and components (Kitano, 2002; Bruggeman and Westerhoff, 2007). The parameters of the model must then be adapted to the given data in order to gain reliable model predictions (Azeloglu and Iyengar, 2015; Kitano, 2002; Bruggeman and Westerhoff, 2007). The obtained in silico predictions are afterwards compared with further experimental observations for verification and falsification (Kitano, 2002; Bruggeman and Westerhoff, 2007). In consequence, the inclusion and exclusion of components, subnetworks or a certain level of detail leads to a refined model (Kitano, 2002; Bruggeman and Westerhoff, 2007). The refined model has to be adapted again to the given data in order to generate new in silico predictions (Kitano, 2002; Bruggeman and Westerhoff, 2007). This leads to a cycle of testing and model refinement which continues until the in silico model is in sufficient agreement with the experimental observations (Kitano, 2002; Bruggeman and Westerhoff, 2007).

Several challenges have to be addressed during such a research cycle (Kitano, 2002; Bruggeman and Westerhoff, 2007). Depending on the complexity of the model and the quality of the underlying data, for a given model, based on prior knowledge and a set of observations, the unknown parameters, e.g., the kinetic parameters or regression coefficients, of this model must be estimated (Jaqaman and Danuser, 2006; Kitano, 2002; Bruggeman and Westerhoff, 2007). The adoption of model parameters is a difficult task because the information regarding the involved actors is often limited, especially in case of top-down approaches (Azeloglu and Iyengar, 2015; Bruggeman and Westerhoff, 2007). Nevertheless, by comparing the model predictions with empirical data, the quality of the model can be evaluated (Azeloglu and Iyengar, 2015; Kitano, 2002; Bruggeman and Westerhoff, 2007). The evaluation of the model may be difficult in cases where the features or components of interest are not directly accessible (Azeloglu and Iyengar, 2015; Bruggeman and Westerhoff, 2007). In many situations, this requires the development of novel experimental in vivo or in vitro techniques, an adjustment of the research question or an adjustment of the experimental design (Azeloglu and Iyengar, 2015; Kitano, 2002; Borkovich and Ebbole, 2010).

Figure 2.2 illustrates the research cycle for bottom-up models (Azeloglu and Iyengar, 2015; Kitano, 2002). In contrast, for predictive modeling, large data-sets are used to train the algorithm, which is then able to predict a certain behavior based on fresh but similar data, as illustrated in Figure 2.2b (Bishop, 2007; Sauer et al., 2007; Gelman et al., 2013). Here, the initial data set is split into a training and a validation data set (Bishop, 2007). The latter is used to validate the predictive power of the model trained by the training data set (Bishop, 2007). Nowadays, machine learning methods such as support vector machines are able to identify potential drug targets and predict correlations between cellular processes (Bishop, 2007; Chuang et al., 2010; Murphy, 2011). This fundamentally differs from the understanding of systems biology as defined in this work (Bishop, 2007; Bruggeman and Westerhoff, 2007; Hübner et al., 2011). Unfortunately, for bottom-up approaches, the model of the related protein-protein interaction network, i.e., the biochemical reaction network, very frequently does not match the observations at first glance (Azeloglu and Iyengar, 2015; Kitano, 2002). Even after adaptation of the involved kinetic parameters, there sometimes remains a significant difference between observations and model predictions (Azeloglu and Iyengar, 2015). Consequently, the model must be improved by incorporating additional knowledge or further assumptions (Azeloglu and Iyengar, 2015). In these cases, the original model is replaced with a revised version and has to be evaluated again (Azeloglu and Iyengar, 2015). This process is repeated until the model is able to reproduce the given empirical data, as shown in Figure 2.2a (Azeloglu and Iyengar, 2015). A picture similar to the one in Figure 2.2 can be drawn for top-down approaches, but here the model is refined by reduction rather than extension (Azeloglu and Iyengar, 2015; Kitano, 2002; Bruggeman and Westerhoff, 2007).

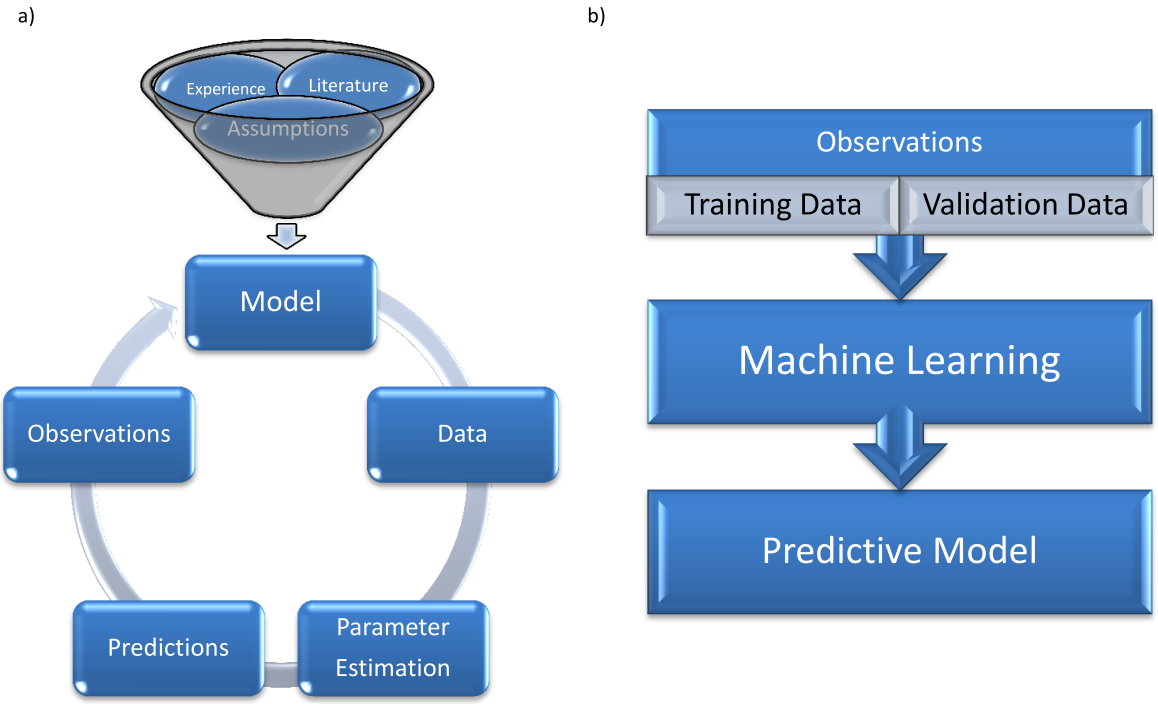


Figure 2.2: General principles of computational biology. (a) In systems biology, bottom-up modeling starts with an initial model based on literature or expert-based prior knowledge. By incorporating data obtained from wet lab experiments or extracted from databases, the model parameters can be estimated. The model then allows for predictions and comparison with real observations. These observations are not necessarily on the same level as the obtained data. An insufficient fit to the observed data indicates a mistake in the model and thus the model needs to be revised and the parameters must be fitted again. This process results in a predictive model with regard to the underlying mechanisms. (b) In contrast, machine learning approaches focus on correlations within the given data set based on a restricted amount of prior knowledge. The given data can be split into two individual and isolated data sets. One is used to train the model and the other to validate the findings. The aim of this approach is to derive a predictive model which allows to make accurate predictions and uncover the inherent features without knowledge of the underlying mechanisms.

2.3 Model Formalisms

As a holistic and interdisciplinary approach, modern systems biology utilizes mathematical theories and experimental tools from different branches of modern life sciences (Machado et al., 2011). Cellular components previously studied independently are nowadays analyzed in an integrated manner (Machado et al., 2011). The characterization of miscellaneous aspects of the methods used in systems biology can be done from different angles (Machado et al., 2011; Kitano, 2002; Bruggeman and Westerhoff, 2007). From a conceptual point of view, models in systems biology are divided into static and dynamic models, as illustrated in Figure 2.3 (Bruggeman and Westerhoff, 2007; Palsson, 2006). Both types of models are useful to describe biological processes, which are classically defined as metabolic and signaling networks (Klipp and Liebermeister, 2006; Palsson, 2006). Some authors consider gene regulatory networks as a special kind of biological model. However, because gene regulatory networks are bilaterally interacting with metabolic and signaling pathways, they are rather a part of these types than independent and autonomous (Machado et al., 2011; Noble, 2012). From a methodological point of view, a distinction can be made between deterministic and probabilistic approaches, while deterministic models still dominating current research in systems biology (Chuang et al., 2010; Murphy, 2011; Hübner et al., 2011).

2.3.1 Static and Dynamic Models

Static models reflect the topology of a biological system at a certain steady state and thus they can be considered as a special or simplified case of dynamic models (Palsson, 2006). A prominent static and hence structure-based approach is the so-called constraint-based modeling (Bordbar et al., 2014; Jerby et al., 2010; Lewis et al., 2010). Most structural models impose constraints on the biological system to predict the rate of turnover of metabolites in biological processes (Bordbar et al., 2014). The essential physico-chemical constraints of static models are often represented by stoichiometric matrices, composed of the stoichiometric coefficients of the involved reactions (Lewis et al., 2012). In contrast to dynamic systems, no kinetic information is required for structural modeling and thus predictions about the inherent dynamics or time-dependent regulations are not possible (Lewis et al., 2012; Klipp and Liebermeister, 2006). Particularly with regard to signaling pathways, time-dependent behavior is of importance because signal amplification is often caused by a change in enzyme activity (Klipp and Liebermeister, 2006; Heinrich et al., 2002). As stoichiometry-based models only deal with steady states, these phenomena are difficult to capture with this approach (Lewis et al., 2012). Under steady state conditions, all state variables are assumed to be constant (Palsson, 2006; Heinrich et al., 2002). However, this holds true only before and after stimulation but not for the transition between both states (Klipp and Liebermeister, 2006; Heinrich et al., 2002).

Studies of the detailed behavior of signaling processes or time-dependent metabolic phenomena require dynamic models (Klipp and Liebermeister, 2006; Klamt et al., 2006; Heinrich et al., 2002). Here, in contrast to top-down strategies, the comparatively small size of bottom-up models allows for dynamic and hence kinetic models (Bruggeman and Westerhoff, 2007). Be-

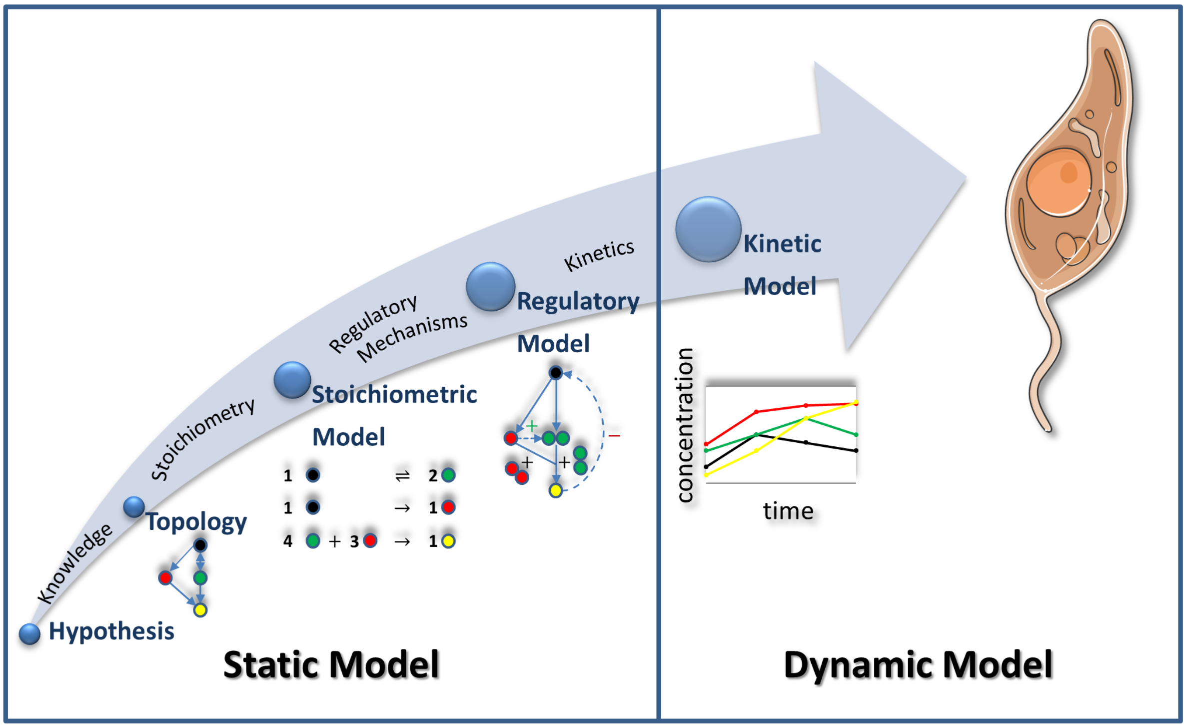


Figure 2.3: Process of model improvement with respect to the model's complexity, from the initial hypothesis to a complete model of the organism. Starting with a first hypothesis or biological question, prior knowledge can be used to define the general topology. Next, stoichiometric coefficients are incorporated to develop a network of biochemical reactions. The network of biochemical reactions can be extended by adding additional regulatory mechanisms, for instance, gene-regulatory components. Through incorporation of kinetics and an estimation of the involved parameters, the model is able to reflect time-resolved processes or even life cycles of complete organisms (Karr et al., 2012; Hübner et al., 2011; Bruggeman and Westerhoff, 2007).

sides the behavior of the system, in certain steady states, dynamic models reflect the dynamic features of the involved components, such as dynamic protein concentrations or time-dependent interactions (Bruggeman and Westerhoff, 2007). Time-resolved measurements are a prerequisite for the development of detailed dynamic models (Bruggeman and Westerhoff, 2007; Klipp and Liebermeister, 2006). As mentioned above, it is still impossible to directly observe chemical reactions in detail on a molecular level in real-time (Noble, 2012; Bruggeman and Westerhoff, 2007). Despite all these challenges, dynamic models are the most realistic type of biological models with the ability to reflect biological processes in real-time (Bruggeman and Westerhoff, 2007; Klipp and Liebermeister, 2006).

2.3.2 Signaling and Metabolic Networks

In literature, biological models are mainly divided into metabolic and signaling networks based on the fact that metabolic pathways are characterized by a flow of matter and signaling pathways are characterized by a flow of information (Klipp and Liebermeister, 2006; Palsson, 2006). Nevertheless, it is obvious that metabolic and signaling networks are connected (Klipp and Liebermeister, 2006). Therefore, attempts have been made to combine both types of networks (Engelhardt et al., 2017; Klamt et al., 2006; Behre and Schuster, 2009). Ultimately, both networks are based on the same fundamental mechanisms; only the point of view is different (Klamt et al., 2006; Behre and Schuster, 2009).

Metabolic systems can be described as the interplay between anabolic and catabolic processes within the cell and are mostly based on enzyme conversions of one chemical entity into another (Dale and Rang, 2011; Berg et al., 2013). Today, metabolic systems biology mainly focuses on the reconstruction of metabolic networks for a target organism on a genome-scale level (Choffnes et al., 2011). In this context, enzyme-catalyzed protein cascades and constrained steady state models are of high relevance (Lehninger et al., 2005; Palsson, 2006). To allow for the reconstruction of the metabolic system, the underlying network has to fulfill a number of assumptions, which were summarized by Bernhard Palsson (Choffnes et al., 2011). Following this, all cellular functions can be explained through chemistry and the cell always behaves in a context-specific manner (Choffnes et al., 2011). The sugar metabolism in *Saccharomyces cerevisiae* under aerobic and aerobic conditions might serve as an example (Rodrigues et al., 2006). It has been shown that *Saccharomyces cerevisiae* is able to switch its glucose catabolism between respiration and fermentation depending on the availability of receptive oxygen in the environment (Rodrigues et al., 2006). In addition, mass and energy must be conserved in metabolic systems (Choffnes et al., 2011). This intrinsic physical law enables the mathematical analysis of metabolic systems (Palsson, 2006).

Constraint-based modeling can be used to gain information on the turnover rates in metabolic networks (Bordbar et al., 2014; Jerby et al., 2010; Lewis et al., 2010). Most of these techniques impose constraints on the system and on the analysis method to predict the rate of turnover of metabolites in a reaction (Bordbar et al., 2014). Constrained-based models assume the metabolism as a system of biochemical reactions (Lewis et al., 2012). Therefore, substrates and enzymes must be present in or produced by the system and the system is considered to be approximately in-balance regarding to all transfers of matter and energy (Bordbar et al., 2014). The overall mass of the system thus remains constant over time and the direction of the involved reactions is under constraint of thermodynamic principles (Volkenshtein, 2009; Lewis et al., 2012). Nowadays, most of the required prior knowledge can be obtained from metabolic reconstruction databases (Lewis et al., 2012).

Signal transduction is often reduced to linear cascades involving protein-protein interactions and the interaction with second messengers, e.g., cAMP or Ca^{2++} (Kestler et al., 2008). Although linear systems are not able to reflect the real dynamics of signaling pathways, they are frequently employed because meaningful readouts of the pathway activation are available when time-resolved measurements of the involved components are not accessible (Kestler et al., 2008; Klipp and Liebermeister, 2006; Heinrich et al., 2002). An important improvement of linear signaling cascade models was the inclusion of positive and negative feedback loops (Kestler et al., 2008). It is well known that most signaling pathways contain several feedback loops to control the amplitude of the transduced signal (Kestler et al., 2008). Here, signal amplification allows for a fast adaptation to the environment (Kestler et al., 2008; Heinrich et al., 2002). The topology of signaling pathways and linear approximations enable significant insights into signaling pathways, for instance regarding the maximum amplitude or the stability of signals (Kestler et al., 2008; Heinrich et al., 2002). But in order to reflect the detailed mechanisms of the signaling pathways and their interaction, detailed models are required, including detailed knowledge of the kinetics and turnover rates (Kestler et al., 2008; Klipp and Liebermeister, 2006; Heinrich et al., 2002).

2.3.3 Probabilistic Approaches

Over the past decade, dynamic modeling has become a leading topic in systems biology (Chuang et al., 2010; Hübner et al., 2011). The majority of modeling approaches in systems biology applied to biochemistry are based on ODEs and partial differential equations followed by stoichiometry-based approaches (Machado et al., 2011; Hübner et al., 2011). ODE-based models are most suitable to describe small and medium-scale networks but often fail on a whole-cell level (Meier-Schellersheim et al., 2009; Hübner et al., 2011; Chuang et al., 2010). This is one of the main reasons why probabilistic methods become more and more important, in addition to deterministic, e.g., stoichiometry- and ODE-based methods (Chuang et al., 2010; Hübner et al., 2011; Klipp and Liebermeister, 2006). Besides other Boolean networks, Bayesian networks and Petri nets are also frequently used in systems biology (Machado et al., 2011; Chuang et al., 2010; Hübner et al., 2011; Klipp and Liebermeister, 2006).

Although Boolean networks are mostly used to model the regulation of genes based on logic rules, protein-protein interactions can be expressed via Boolean networks as well (Machado et al., 2011; Hübner et al., 2011). The components of Boolean networks are encoded as Boolean variables which can either be active or inactive (Machado et al., 2011). In that manner, at each discrete time point, the state of each Boolean variable is given by a function composed of the states of all related regulators, e.g., enzymes or transcription factors (Machado et al., 2011). The state of all components changes synchronously (Machado et al., 2011). Thus, for large-scale networks, the determination of all possible states is a cost-intensive process (Machado et al., 2011). However, it can be used to find attracted steady states and to analyze the robustness of the network (Machado et al., 2011). Probabilistic Boolean networks have been established to incorporate uncertainties (Machado et al., 2011).

In contrast to Boolean networks, Bayesian networks allow for more detailed dependencies and different states of the modeled components (Bishop, 2007; Machado et al., 2011). Bayesian networks are probabilistic graphs whereby their components are represented by random discrete or continuous variables with conditional dependencies, as mentioned in Section 4.3 (Bishop, 2007; Machado et al., 2011). This can be considered as a directed graph, and each component, i.e., node, contains a probabilistic function which depends on the values of all related influencing nodes (Bishop, 2007; Machado et al., 2011). Many different methods have been developed to infer the network structure and the corresponding probability parameters (Bishop, 2007; Machado et al., 2011). Due to the probabilistic nature of Bayesian networks, they allow to work with incomplete data sets and are frequently used for the modeling of gene-regulatory networks (Bishop, 2007; Machado et al., 2011; Chuang et al., 2010). Dynamic Bayesian networks are a common extension to simulate time-dependent biological systems (Machado et al., 2011).

Whereas Boolean and Bayesian networks only allow for one type of node representing a cluster of components or an individual component, Petri nets incorporate two different types of nodes, i.e., places and transitions, and the dependencies between the nodes are represented by directed and weighted arcs (Scheidel et al., 2016; Junker and Schreiber, 2008; Machado et al., 2011). The passive parts, i.e., components or clusters, are captured by the places and the chemical reactions are represented through transitions (Scheidel et al., 2016; Junker and Schreiber, 2008; Machado et al., 2011). Commonly, places are represented as circles and transitions as squares (Scheidel et al., 2016; Junker and Schreiber, 2008). Thereby the nodes of the resulting graph-structure represent either molecules or chemical reactions, whereas the directed edges represent the relation between components and reactions (Scheidel et al., 2016; Junker and Schreiber, 2008; Machado et al., 2011). Usually the edges are weighted according to their corresponding stoichiometric factor (Scheidel et al., 2016). This is in contrast to common graphical representations of biochemical networks, where components are represented by nodes and the corresponding reactions via directed edges (Scheidel et al., 2016; Palsson, 2006). The places are associated with tokens which can stand for the amount or concentration of a chemical substance or a fulfilled precondition of a reaction (Scheidel et al., 2016; Junker and Schreiber, 2008). Tokens are produced or consumed when the transitions fire (Machado et al., 2011). This means that the transitions modulate the tokens in a similar way as biological processes would change the state of the involved components (Scheidel et al., 2016; Junker and Schreiber, 2008; Machado et al., 2011). The state of the system is given by the distribution of the tokens, which thus describes the dynamics of the system (Junker and Schreiber, 2008; Machado et al., 2011).

2.3.4 Stoichiometric Matrices

Given sufficient prior information, each biological process based on protein-protein interactions can be described through a set of biochemical reactions (Heinrich et al., 2002; Palsson, 2006). These reaction processes are often an enzyme-catalyzed conversion of molecules or protein binding events (Berg et al., 2013).

Definition 1: Chemical Reaction Network

A chemical reaction network (CRN) is constituted by a set of M chemical net reactions R_m, N species X_n and K reaction velocities $r_k \geq 0$ of each single reaction, with $m = 1, ..., M$, $n = 1, ..., N$ and $k = 1, ..., K$. Each reaction R_m is clearly associated with a set of species and reaction velocities, e.g., $R_m = \{X_n, r_k\}$. The species partaking in a chemical reaction can be differentiated into educts and products. The educts are transformed into products as the reaction proceeds with a defined but not necessarily constant reaction velocity. The abundance of species X_n is denoted by x_n.

Unidirectional net reactions are called irreversible and composed of one reaction velocity. In contrast, bidirectional net reactions are called reversible and can be split into forward and backward unidirectional single reactions. For a system with purely irreversible reactions, the number of reactions M is equal to the number of single reactions and thus reaction velocities K. For reversible reactions, the net reaction of the related forward reaction velocity r_k and backward reaction velocity $r_{(k+1)}$ is given by their difference $r_k - r_{(k+1)}$.

Stationary reaction velocities in assigned states, e.g., steady states, are called fluxes.

The numbers of molecules of each educt and product occurring in a reaction are called stoichiometric coefficients $s_{n,k}$. One stoichiometric coefficient can be clearly assigned to each species and single reaction of the CRN. The coefficients corresponding to the reacting species or educts are negative, whereas those corresponding to the produced species or products are positive. The stoichiometric coefficient is zero for all other species not participating in the particular reaction. According to Definition 1, a simple CRN with one unidirectional chemical reaction:

$$2\mathrm{X}_1 + \mathrm{X}_2 \xrightarrow{r_1} \mathrm{X}_3 \tag{2.1}$$

consist of one irreversible net reaction R_1 with one reaction velocity r_1, two educts X_1 and X_2, and one product X_3. The reaction is defined as $R_1 = \{X_1, X_2, X_3, r_1\}$ with corresponding species concentrations x_1, x_2 and x_3. In this CRN, the stoichiometric coefficients are $s_{1,1} = -2$, $s_{2,1} = -1$ and $s_{3,1} = 1$.

This allows to describe each biological process that is based on chemical reactions, e.g., protein-protein interactions (Palsson, 2006; Berg et al., 2013). Because there is no one-to-one correspondence between genes associated with the metabolism and the chemical reactions that occur, it is not trivial to incorporate related interactions (Palsson, 2006; Berg et al., 2013). Often genes are associated with their encoding proteins, but this relation is not unique (Palsson, 2006; Berg et al., 2013; Lewis et al., 2012). It has been demonstrated that genes are able to encode different proteins and vice versa that genes can be regulated by different proteins (Palsson, 2006; Berg et al., 2013).

Definition 2: Stoichiometric Matrix

Let the vector $\boldsymbol{v}$ be the turnover rate of molecules through the system. Then, we can clearly define a linear transformation $S \in \mathbb{R}^{N \times K}$ which transforms the flux vector $\boldsymbol{v}$ into the derivative of a concentration vector $\boldsymbol{x}$ as follows:

$$\frac{d\boldsymbol{x}}{dt} = S\boldsymbol{v}.$$

For a CRN with K single reactions and N species, the dimension of S is $N \times K$. Each element $s_{n,k}$ represents the stoichiometric coefficient corresponding to species n in reaction k.

Every CRN can be split into external, exchange and internal reactions with respect to the boundaries of the network. In most situations, both exchange and internal reactions are considered whereas external reactions are frequently neglected, but, depending on the context, it may be useful to only consider internal reactions (Palsson, 2006). In general, the reaction velocities r_k are functions of the related species concentrations and a set of rate constants $\tilde{\boldsymbol{r}}_k$:

$$r_k = f_k\left(\boldsymbol{x}, \tilde{\boldsymbol{r}}_k\right), \tag{2.2}$$

with a general continuous and time-dependent function f_k. In general, the derivative of a concentration vector $\boldsymbol{x}$ is depend on the related reaction velocities and the corresponding stoichiometric coefficients, and the entities can be calculated through

$$\frac{dx_n}{dt} = \sum_{k=1}^{K} s_{n,k} f_k\left(\boldsymbol{x}, \tilde{\boldsymbol{r}}_k\right), \tag{2.3}$$

where $s_{n,k}$ is the stoichiometric coefficient of species n partaking in the single reaction k. At fixed time points e.g., under steady state conditions, $f_k\left(\boldsymbol{x}, \tilde{\boldsymbol{r}}_k\right) = v_k$ is obtained and Equation (2.3) yields

$$\frac{dx_n}{dt} = \sum_{k=1}^{K} s_{n,k} v_k. \tag{2.4}$$

The number of elements of the flux vector $\boldsymbol{v}$ is equal to the number of single-reaction velocities in the CRN. In most situations, the number of reactions in a CRNs is larger than the number of species and therefore $K > N$, implying that S may not be of full rank (Palsson, 2006).

The stoichiometric matrix S is commonly considered as a connectivity matrix or a network represented by a map (Palsson, 2006). The nodes in the map then correspond to the rows and the links to the columns of S, respectively (Palsson, 2006). Thus, S represents the general topology of the system and is typically used to investigate time-independent features of the CRN (Palsson, 2006).

These stoichiometry-based approaches have to rely on several assumptions (Palsson, 2006). First, the reaction component of interest is assumed to be uniformly distributed in the system (Palsson, 2006). Second, the concentration is assumed to be sufficiently high to define a real number (Palsson, 2006). In addition, the reaction velocities are considered to not explicitly depend on time (Palsson, 2006).

2.3.5 Differential Equations

In contrast to time-independent stoichiometry-based approaches, ODEs allow to capture the inherent dynamics of CRNs (Palsson, 2006; Azeloglu and Iyengar, 2015). Because the reaction dynamics are a key aspect of signaling processes, ODE-based approaches are frequently used to model this type of system (Hübner et al., 2011; Klipp and Liebermeister, 2006). ODEs are a widely discussed common topic in mathematics and interdisciplinary research (Mattheij and Molenaar, 2002).

Definition 3: Ordinary Differential Equation

Let $F(t)$ be a real-valued function $F : I \times \Omega \rightarrow \mathbb{R}$ with $\Omega \subset \mathbb{R}^{d+1}$, $I \subset \mathbb{R}$, $x(t)$ a real-valued function with $x : I \rightarrow \mathbb{R}$ and $x^{(d)}(t)$ the derivative of $x(t)$ to the order of d given by $\frac{d^d x(t)}{d^d t}$ with $t \in \mathbb{R}$. Then, the implicit ordinary differential equation to the order of d is given by

$$F\left(t, x(t), x^{(1)}(t), ..., x^{(d)}(t)\right) = 0.$$

In contrast, the explicit ordinary differential equation f with $f : I \times \tilde{\Omega} \rightarrow \mathbb{R}$ with $\tilde{\Omega} \subset \mathbb{R}^d$ is a special case of F and given by

$$x^{(d)}(t) = f\left(t, x(t), x^{(1)}(t), ..., x^{(d-1)}(t)\right).$$

The function $\Phi : \tilde{I} \rightarrow \mathbb{R}$ is named solution of the implicit ODE in the interval $\tilde{I} \subset \mathbb{R}$ if $\Phi\left(t, x(t), x^{(1)}(t), ..., x^{(d)}(t)\right) = 0$ and Φ is d times differentiable.

Accordingly, $\tilde{\Phi} : \tilde{I} \rightarrow \mathbb{R}$, $d-1$ times differentiable, is named solution of the explicit ODE in the interval $\tilde{I} \subset \mathbb{R}$ if $\tilde{\Phi}\left(t, x(t), x^{(1)}(t), ..., x^{(d-1)}(t)\right) = x^{(d)}$.

For example, the first-order ODE

$$\dot{x}(t) = -2x(t) \iff \dot{x}(t) + 2x(t) = 0, \tag{2.5}$$

with $\dot{x}(t) = x^{(1)}(t)$ has the solution $x(t) = c\exp(-2t)$ with $c \in \mathbb{R}$.

The above definition can be extended to a system of coupled ODEs incorporating several dependent variables (Walter, 2000). For systems biology, first-order systems of explicit ODEs are of interest.

Definition 4: First-Order System of Ordinary Differential Equations

Given n explicit ordinary differential equations $f_n : I \times \tilde{\Omega} \to \mathbb{R}$ with $\tilde{\Omega} \subset (\mathbb{R}^n)^1$, $I \subset \mathbb{R}$, $\boldsymbol{x} : I \to \mathbb{R}^n$ and $t \in \mathbb{R}$ the system

$$\boldsymbol{x}^{(1)}(t) = \boldsymbol{f}(t, \boldsymbol{x}(t))$$

is called system of ordinary differential equations of the order one. The system is said to be autonomous if $\boldsymbol{f}$ is not directly dependent on t.

In general, due to the fact that ODE systems are basically solved via integration, they do not necessarily have a unique solution, unless further constraints or initial conditions are given (Walter, 2000). For biological systems, however, the initial conditions are frequently known and hence in many cases the arising initial value problem (IVP) has a unique solution (Palsson, 2006; Kremling, 2012; Walter, 2000).

Definition 5: Initial Value Problem

Given a vector-valued function $\boldsymbol{f} : I \times \tilde{\Omega} \to \mathbb{R}^n$ with $\tilde{\Omega} \subset (\mathbb{R}^n)^d$, $I \subset \mathbb{R}$, $\boldsymbol{x} : I \to \mathbb{R}^n$ and $t \in \mathbb{R}$, the system

$$\begin{aligned} \boldsymbol{x}^{(d)}(t) &= \boldsymbol{f}\left(t, \boldsymbol{x}(t), \boldsymbol{x}^{(1)}(t), \ldots, \boldsymbol{x}^{(d-1)}(t)\right) \\ \boldsymbol{x}(0) &= \boldsymbol{x}_0, \end{aligned}$$

with a vector of initial conditions $\boldsymbol{x}_0 \in \mathbb{R}^n$ is called initial value problem.

The Picard-Lindelöf theorem guarantees a unique solution of the IVP given by Definition 5 for Lipschitz-continuous functions (Teschl, 2012). Lipschitz continuity is a stricter form of continuity ensuring that a given continuous function is limited in the rate of change over time (Teschl, 2012). For most high-dimensional IVPs, an analytical solution does not exist (Teschl, 2012; Walter, 2000). Instead, the IVP can be solved numerically for discrete time points and restricted due to its complexity (Hanke-Bourgeois, 2008).

Complex processes, e.g., those involving dynamics in time and space, have more than one independent variable (Kholodenko, 2006). These phenomena are usually addressed via partial differential equations (Kholodenko, 2006). Hybrid models in cancer research, where cell growth is often described through growth-consumption dynamics and cell dynamics on a lattice, are based on partial differential equations (Altrock et al., 2015). In contrast, some biological processes are of random nature, such as the synaptic activity of neurotransmitters released from synaptic vesicles (Saarinen et al., 2008). Here, stochastic differential equations incorporating random processes are commonly used (Saarinen et al., 2008). Unlike ODEs or partial differential equations, stochastic differential equations are defined via integrals (Bauer, 2002, 1992). This is equivalent to the derivative-based formulation but circumvents the lack of differentiability of stochastic processes, for instance, in case of Brownian motion (Bauer, 2002, 1992).

2.4 Properties of Protein Interactions

Biological processes are based on biochemical reactions which themselves are driven by chemical interactions (Berg et al., 2013; Choffnes et al., 2011). For general biological phenomena, it is not possible to postulate special laws of nature as in physics (Berg et al., 2013). However, all living organisms share a large number of biochemical features (Berg et al., 2013). Like all chemical reactions, biological processes are driven by the laws of thermodynamics (Berg et al., 2013). Besides the underlying chemical principles, the structure of the involved macromolecules is also important (Alberts et al., 2014; Berg et al., 2013). This becomes particularly clear in case of DNA, which stores all genetic information (Alberts et al., 2014; Berg et al., 2013). In general, two DNA strands form a double helix via hydrogen bridge linkages between two corresponding bases (Berg et al., 2013). In the DNA double helix, not the chemical interactions but the sequence of base pairs encode the genetic information (Berg et al., 2013). This very complex mechanism helps to ensure a stable storage of genetic information because even if one DNA strand is damaged it can be rebuild by the corresponding strand (Berg et al., 2013). As illustrated, both mechanisms are rather closely interlinked than independent (Berg et al., 2013). The structure has a strong impact on the binding behavior of proteins which, in turn, influences the binding capacity and therefore the structure itself (Berg et al., 2013; Palsson, 2006). However, the dominant driving force of protein-protein interactions are the laws of thermodynamics (Berg et al., 2013; Palsson, 2006).

As stated in the first law of thermodynamics, energy can neither be consumed nor produced, it can only be converted from one form into another (Kremling, 2012; Berg et al., 2013). For spontaneous processes, the disorder of the system, called entropy, increases (Kremling, 2012; Berg et al., 2013). The entropy denotes the ratio between the reversible heat and the system's absolute temperature (Kremling, 2012; Berg et al., 2013). The change in entropy can be directly measured by the so-called Gibbs free energy ΔG (Berg et al., 2013). In addition, ΔG is a measure for the work which can be performed by reversible chemical reactions under constant conditions (Berg et al., 2013). Consequently, the Gibbs free energy is a measure for the driving force of a reaction and thus enables a differentiation between exergonic spontaneous processes and endergonic processes requiring an absorption of energy from the surroundings (Berg et al., 2013). Biochemical reactions proceed spontaneous only if the entropy increases, i.e., if ΔG is negative (Berg et al., 2013). In thermodynamic equilibrium, the Gibbs free energy is equal to zero (Berg et al., 2013). ΔG is always negative for exergonic spontaneous processes and always positive for endergonic processes (Berg et al., 2013). In other words, the overall entropy increases in exergonic processes and decreases in endergonic processes (Berg et al., 2013). Only reactions increasing the entropy with $\Delta G < 0$ occur spontaneously and therefore the chemical equilibrium is shifted towards this direction (Berg et al., 2013). The reaction which decreases the entropy with $\Delta G > 0$ proceeds more slowly and the reaction rate inclines towards zero (Berg et al., 2013). Please note that ΔG alone provides no information regarding the reaction rate of the involved reactions (Berg et al., 2013). In practice, reactions are often accelerated by enzymes (Berg et al., 2013).

Enzymes are macromolecules able to catalyze biochemical reactions (Alberts et al., 2014; Berg et al., 2013). Enzymes do not change the chemical equilibrium but may dramatically increase the reaction rate (Alberts et al., 2014; Berg et al., 2013). In such a case, the products of reactions catalyzed by enzymes can be produced within seconds, whereas the same reaction might require hundreds of years to produce the same amount of products without the presence of the enzyme (Berg et al., 2013). At the same time, enzymes are highly specific and mostly activated by other small cofactors, e.g., molecules or metal ions (Alberts et al., 2014; Berg et al., 2013). After activation by a cofactor, the enzymes bind to the target substrate and thus build an intermediate complex, thereby reducing the activation energy of the chemical reaction (Kremling, 2012; Berg et al., 2013). Afterwards, the enzyme-substrate complex dissociates into the enzyme and the modified target (Berg et al., 2013). Enzymes are often inhibited reversibly or irreversibly by small molecules or ions in different ways (Alberts et al., 2014; Berg et al., 2013). In case of an irreversible inhibition, the molecule binds to the enzyme so that a dissociation process becomes very slow or almost impossible (Berg et al., 2013). In contrast, reversible inhibitors allow for a faster dissociation from the enzyme (Berg et al., 2013).

As shown, the timing of biochemical reactions, i.e., the kinetics, is driven by the laws of thermodynamics, but can also be manipulated by cellular mechanisms, for instance, via enzymatic catalysis (Kremling, 2012; Berg et al., 2013). Biochemical reactions are based on both elementary reactions which cannot be further taken apart and the law of mass action (Nic et al., 2009; Berline and Bricker, 1969; Waage and Gulberg, 1986). Thus the rate of any given chemical reaction is proportional to the product of the activities or concentrations of the reactants (Berline and Bricker, 1969; Waage and Gulberg, 1986). In other words, the reaction velocity is proportional to the probability of collision between the reactants (Berline and Bricker, 1969; Waage and Gulberg, 1986). For a CRN with K single reactions, N species X_n and the corresponding stoichiometric coefficients $s_{n,k}$ with $n = 1, ..., N$ and $k = 1, ..., K$, the reaction rate r_k for the law of mass action is given by

$$r_k = \tilde{r}_k \prod_{n=1}^{N} x_n^{s_{n,k}}, \tag{2.6}$$

with the rate constant $\tilde{r}_k \in \mathbb{R}$ (Palsson, 2006; Nic et al., 2009). Each CRN can be expressed as a set of elementary reactions based on the law of mass action (Nic et al., 2009; Berg et al., 2013). But the modeling of coupled reactions based on the law of mass action often results in a large set of intermediate reactions with a huge amount of additional parameters, especially for enzymatic reactions (Berg et al., 2013; Kremling, 2012).

Basic enzymatic reactions can help to describe the catalysis of substrate conversion by a single enzyme (Berg et al., 2013; Kremling, 2012). For a given biochemical reaction describing the conversion of a single substrate S into the single product P catalyzed by a single enzyme E, this basic enzymatic reaction is given by elementary reactions involving intermediate complexes E:S as follows:

$$\mathrm{E} + \mathrm{S} \underset{r_2}{\overset{r_1}{\rightleftarrows}} \mathrm{E{:}S} \underset{r_4}{\overset{r_3}{\rightleftarrows}} \mathrm{P} + \mathrm{E}, \tag{2.7}$$

with a set of reaction rates $\{r_1, r_2, r_3, r_4\}$ (Berg et al., 2013; Kremling, 2012).

According to Equation (2.7), the enzyme and the target protein first build an enzyme-substrate complex (Berg et al., 2013; Kremling, 2012). Then, the substrate is converted into the product, but is still bound to the enzyme (Berg et al., 2013; Kremling, 2012). This process is often not quantifiable and thus neglected (Berg et al., 2013; Kremling, 2012). In the last step, the enzyme-product complex dissociates (Berg et al., 2013). Because the substrate is modified immediately after the enzyme and the substrate meet and the reverse reaction does not occur, $r_4 \approx 0$ is assumed to simplify Equation (2.7) yielding

$$\mathrm{E} + \mathrm{S} \underset{r_2}{\overset{r_1}{\rightleftarrows}} \mathrm{E{:}S} \xrightarrow{r_3} \mathrm{P} + \mathrm{E}, \tag{2.8}$$

with a set of reaction rates $\{r_1, r_2, r_3\}$ (Berg et al., 2013; Kremling, 2012). As discussed in Section 2.3.4 an elementary-based CRN with K reaction velocities $f_k(\boldsymbol{x}, \tilde{\boldsymbol{r}}_k)$ and the corresponding stoichiometric matrix S can be expressed as

$$\frac{dx_n}{dt} = \sum_{k=1}^{K} s_{n,k} f_k(\boldsymbol{x}, \tilde{\boldsymbol{r}}_k). \tag{2.9}$$

For Michaelis-Menten kinetics, the elementary reactions of the network can be reformulated as

$$\mathrm{E} + \mathrm{S} \underset{r_2}{\overset{r_1}{\rightleftarrows}} \mathrm{E{:}S} \xrightarrow{r_3} \mathrm{P} + \mathrm{E} \Longrightarrow \mathrm{S} \xrightarrow{f_k(\boldsymbol{x}, \tilde{\boldsymbol{r}}_k)} \mathrm{P}. \tag{2.10}$$

The rate of formation of the enzyme-substrate complex and the corresponding dissociation rate from Equation (2.8) are constituted by $r_1\mathrm{es}$ and $(r_2 + r_3)\,\mathrm{e{:}s}$, where the protein concentrations are denoted as lower case letters (Berg et al., 2013; Kremling, 2012; Klonowski, 1983). It is known that enzymes are able to adapt their activation to the substrate concentration and thus, after an initial time period, e:s can be assumed as approximately constant (Berg et al., 2013; Kremling, 2012; Klonowski, 1983). Then, it is possible to determine the forward dissociation rate of the enzyme-substrate complex in a suitable manner (Berg et al., 2013; Kremling, 2012; Klonowski, 1983). In this way, the Michaelis-Menten kinetic from Equation (2.8) can be expressed as a set of coupled ODEs

$$\frac{d\mathrm{e}}{dt} = (r_2 + r_3)\,\mathrm{e{:}s} - r_1\mathrm{es} \tag{2.11a}$$

$$\frac{d\mathrm{s}}{dt} = r_2\mathrm{e{:}s} - r_1\mathrm{es} \tag{2.11b}$$

$$\frac{d\mathrm{e{:}s}}{dt} = r_1\mathrm{es} - (r_2 + r_3)\,\mathrm{e{:}s} \tag{2.11c}$$

$$\frac{d\mathrm{p}}{dt} = r_3\mathrm{e{:}s}. \tag{2.11d}$$

From $\mathrm{e{:}s} \approx const.$ follows $\frac{d\mathrm{e{:}s}}{dt} \approx 0$ after an initial time period and hence

$$r_1 \mathrm{es} = (r_2 + r_3)\,\mathrm{e{:}s} \Leftrightarrow \mathrm{e{:}s} = \frac{r_1 \mathrm{es}}{(r_2 + r_3)}. \tag{2.12}$$

Now, supposed that $\mathrm{e} + \mathrm{e{:}s} = \mathrm{e}_c = const.$,

$$\mathrm{e{:}s} = \frac{r_1\,(\mathrm{e}_c - \mathrm{e{:}s})\,\mathrm{s}}{(r_2 + r_3)}. \tag{2.13}$$

Combining Equations (2.11b), (2.11c) and (2.11d), and with $\frac{d\mathrm{e{:}s}}{dt} = 0$, it follows that

$$\frac{d\mathrm{s}}{dt} + \frac{d\mathrm{e{:}s}}{dt} + \frac{d\mathrm{p}}{dt} = 0 \Leftrightarrow \frac{d\mathrm{s}}{dt} = -\frac{d\mathrm{p}}{dt}. \tag{2.14}$$

Hence

$$-\frac{d\mathrm{s}}{dt} = \frac{d\mathrm{p}}{dt} = r_3 \frac{r_1\,(\mathrm{e}_c - \mathrm{e{:}s})\,s}{(r_2 + r_3)} = V_{max}\frac{\mathrm{s}}{K_m + s}, \tag{2.15}$$

with $V_{max} = r_3 e_c$, and $K_m = \frac{r_2 + r_3}{r_1}$ holds true. Finally, the reduced system can be derived:

$$\frac{d\mathrm{s}}{dt} = -V_{max}\frac{s}{K_m + s} \tag{2.16a}$$

$$\frac{d\mathrm{p}}{dt} = V_{max}\frac{s}{K_m + s}. \tag{2.16b}$$

The rate constant V_{max} illustrated in Figure 2.4 reflects the maximum rate which occurs when all binding sites of the enzyme are saturated (Berg et al., 2013). The Michaelis-Menten constant K_m is used to characterize enzyme substrate interactions (Berg et al., 2013).

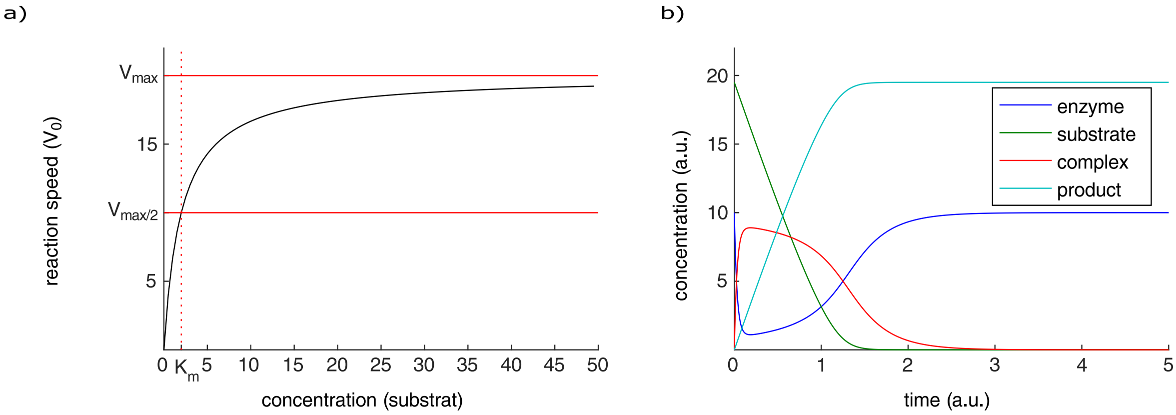


Figure 2.4: Illustration of the Michaelis-Menten kinetics. (a) Relationship between the substance concentration, the initial reaction rate V_0 at time $t \approx 0$ and the Michaelis-Menten constant K_m for a single substrate. The Michaelis-Menten constant K_m is defined as the substrate concentration at which the reaction rate is half of V_{max}. (b) Related time-concentration curve for each of the involved species.

In the past, the applicability of the quasi-steady-state approximation (QSSA)

$$\frac{d\mathrm{e{:}s}}{dt} \approx 0 \tag{2.17}$$

to biochemical networks has been controversially discussed (Klonowski, 1983; Turanyi et al., 1993; Schauer and Heinrich, 1983). The QSSA always causes an error, and the overall error of all QSSAs is crucial for the dynamics of the system (Turanyi et al., 1993). More importantly, the QSSA does not necessarily imply that $\frac{d\mathrm{e{:}s}}{dt} \approx 0$, and it can generally only be applied after an initial time period (Turanyi et al., 1993). For biochemical reactions, the formation of the enzyme-substrate complex is a very fast process compared to the product formation, which justifies the application of the QSSA to enzyme kinetics (Schauer and Heinrich, 1983; Klonowski, 1983). Hence, the QSSA can still be considered as a powerful tool to reduce the dimension of biochemical reactions (Klonowski, 1983; Schauer and Heinrich, 1983; Keener and Sneyd, 2009).

Altogether, it is necessary that the total amount of the free substrate must be close to the total substrate concentration in the system and thus that the enzyme association and dissociation must happen much faster than the production of the reaction product (Slyke and Cullen, 1914; Briggs and Haldane, 1925). If the amount of free substrate is not close to the total substrate concentration in the system, a significant fraction is bound to the enzyme (Slyke and Cullen, 1914; Briggs and Haldane, 1925). In this case, the quadratic velocity equation for tight-binding substrates can be used

$$\frac{d\mathrm{p}}{dt} = V_{max} \frac{(\mathrm{e}_c + s_c + K_m) - \sqrt{(\mathrm{e}_c + \mathrm{s}_c + K_m)^2 - 4\mathrm{e}_c\mathrm{s}_c}}{2\mathrm{e}_c}, \tag{2.18}$$

with the total substrate concentration given by $\mathrm{s}_c = \mathrm{s} + \mathrm{e{:}s}$ (Morrison, 1969). Besides the quadratic velocity equation for tight-binding substrates, there are many other kinetics that can be used to simplify biochemical reactions (Marangoni, 2003). The Hill equation, which is related to the logistic function, is an important kinetic model used to simplify enzymatic reactions with many binding sites (Kremling, 2012).

In a broader sense, gene regulation can also be described with the help of kinetics (Chen et al., 1999; Karlebach and Shamir, 2008). The dynamic of the messenger ribonucleic acid (mRNA) concentration $\frac{d[\mathrm{mRNA}]}{dt}$ is described as the transcription $f([\mathrm{protein}])$ with respect to the related proteins minus its degradation $r_1\,[\mathrm{mRNA}]$. The dynamic of the proteins is given by the translation $r_2\,[\mathrm{mRNA}]$ minus the degradation $r_3\,[\mathrm{protein}]$. Here, a unique relation between the mRNA and the encoded protein is frequently assumed (Chen et al., 1999). Given the basic model

$$\frac{d\,[\mathrm{mRNA}]}{dt} = f([\mathrm{protein}]) - r_1\,[\mathrm{mRNA}] \tag{2.19a}$$

$$\frac{d\,[\mathrm{protein}]}{dt} = r_2\,[\mathrm{mRNA}] - r_3\,[\mathrm{protein}], \tag{2.19b}$$

the transcription function $f\left([\text{protein}]\right)$ can be modeled by a linear function

$$f\left([\text{protein}]\right) = \sum_{n=1}^{N} \tilde{r}_n \left[\text{protein}_n\right] \tag{2.20}$$

describing a combined effect of activators and inhibitors (Chen et al., 1999). In more complex situations, $f\left([\text{protein}]\right)$ can be derived via a first order Taylor approximation (Chen et al., 1999).

Overall, ODEs allow for detailed models of biological processes but require a huge amount of prior knowledge regarding the general network topology, as well as high-quality experimental data to estimate the involved species concentrations (Azeloglu and Iyengar, 2015; Sauer et al., 2007).

CHAPTER 3

M_2 Receptor-induced Signaling

Mechanisms of drug action at various levels and from various points of view are of common pharmacological interest (Dale and Rang, 2011). For more than 60 years, pharmacologists have successfully discovered new drugs and therapies for many complex diseases, even though the costs of drug development have increased dramatically (Zhao and Iyengar, 2012). Often the use of correlations as starting points for drug discovery does not lead to desirable results (Zhao and Iyengar, 2012). Correlation-based and also other black box models are often not able to reflect the complex relation of interactions on a molecular level and their effects on an organ level (Zhao and Iyengar, 2012). As a consequence, processes discovered in silico, in vitro or even in model organisms differ significantly from the human metabolism (Zhao and Iyengar, 2012). Besides other pitfalls in drug development, the uncertainty of black box models is reflected by the high fraction of drugs which nowadays fail testing on patients to appraise their efficacy, effectiveness and safety (Wein, 2016; Zhao and Iyengar, 2012). This led to the development of a new promising field of systems pharmacology (Zhao and Iyengar, 2012). Systems pharmacology investigates drug actions on a genome-wide level and combines networks of different size and level (Zhao and Iyengar, 2012). In addition to genomics, well-defined drug-induced signaling pathways are also of interest (Zhao and Iyengar, 2012). In this spirit, this work aims to investigate the M_2 receptor-mediated ligand-induced change of the cell's optical density.

Drugs such as ligands are substances with measurable biological effects which often target receptors (Dale and Rang, 2011). Ligands are molecules forming complexes with specific targets without getting converted (Dale and Rang, 2011). Reversible binding to a target, e.g., a receptor, causes cellular responses mediated by signaling cascades (Dale and Rang, 2011). Due to their effect on receptors, ligands are distinguished into antagonists and agonists (Dale and Rang, 2011). Agonists are able to stimulate receptors whereas antagonists have inhibitory effects (Dale and Rang, 2011). Falsely, antagonists are often called inverse agonists (Dale and Rang, 2011). However, inverse agonists are ligands which reduce the receptor activity of their targets and thus induce a response opposite to the agonist (Dale and Rang, 2011). In case of the M_2 receptor, the ligand moves to the ligand-binding pocket of the receptor, which is deeply buried within the membrane, and engages in extensive hydrophobic contact with the receptor (Kruse et al., 2014). An elicited change of the three-dimensional shape orientation of the receptor then induces a signal transduction (Kruse et al., 2014).

3.1 Muscarinic Acetylcholine Receptors

Nowadays up to 50% of available drugs are related to G protein–coupled receptors (GPCRs), one of the largest protein families encoded in the human genome, approximately containing 865 receptors (Fang et al., 2015; Bylund, 2007; Lander et al., 2001). Unsurprisingly, the GPCR family has become one of the most important target classes of proteins for drug discovery (Zheng, 2006; Fang et al., 2015; Dale and Rang, 2011). Most drugs for cardiovascular diseases are targeting GPCRs, and nowadays GPCRs are also potent therapeutic targets for Alzheimer's disease (Salazar et al., 2007; Zhao et al., 2016). Due to its structure, GPCRs are also called 7-transmembrane receptors (Dale and Rang, 2011). This integral membrane proteins consist of seven transmembrane helices connected through three intra- and three extracellular loops (Dale and Rang, 2011). Interestingly enough, the five major families have only a few identical sequences in common (Katritch et al., 2013).

Members of the GPCR family vary widely in their function and distribution (Dale and Rang, 2011). The GPCR family includes receptors for many hormones and slow transmitters, such as adrenergic receptors or chemokine receptors, which are involved in the pathogenesis of many important diseases, e.g., Alzheimer's disease, asthma, cancer and acquired immunodeficiency syndrome (Russo-Neustadt and Cotman, 1997; Vela et al., 2015; Sears and Lötvall, 2005). Adrenergic receptors are one of the longest known group of GPCRs (Strosberg, 1993). They include nine subtypes and mediate (no-)adrenaline signaling (Strosberg, 1993). A prominent member of this group is the β_2-adrenergic receptor, which is the main drug target for asthma (Sears and Lötvall, 2005). In addition, it is strongly related to chronic heart failures and aging (Brodde et al., 2006; Dale and Rang, 2011).

As another important and widely studied class of GPCRs, muscarinic acetylcholine receptors (mACHRs) are divided into five subtypes regarding to their distribution and physiological function (see Table 3.1) (Dale and Rang, 2011; Bonner et al., 1987; Brodde and Michel, 1999). The M_2 receptor, an important subtype of acetylcholine receptors, is among other locations mainly expressed in presynaptic terminals of peripheral, central neurons and cardiomyocytes (Dale and Rang, 2011; Hernandez and Rathinavelu, 2006). Since the M_2 receptor influences the heart beat frequency, it is related to negative dromotropic and negative chronotropic events (Dale and Rang, 2011; Brodde and Michel, 1999). Its malfunctioning has been associated with a number of diseases, such as cardiomyopathies (Brodde and Michel, 1999). Like all mACHRs, the M_2 receptor has an allosteric and a classical orthosteric binding site (Kruse et al., 2014). Ligands bound to the allosteric binding site are able to modulate the receptor by changing its conformation (Kruse et al., 2014). Thereby either the activity of the classical orthosteric binding site gets modified or the binding affinity of the receptor is changed (Berg et al., 2013; Dale and Rang, 2011). A prominent allosteric modulator of

Subtype	Distribution	Function
M_1	• autonomic ganglia • central nervous system	• gastric secretion • excitation of the central nervous system
M_2	• heart • central nervous system • smooth muscles	• central muscarinic effects • cardiac inhibition • neural inhibition
M_3	• heart • glands • smooth muscles	• gastrointestinal smooth muscle contraction • ocular accommodation • salivary secretion
M_4	• central nervous system	• enhanced locomotion
M_5	• central nervous system	• vasodilation

Table 3.1: Classification of mACHR subtypes according to their distribution and function (Dale and Rang, 2011; Bonner et al., 1987; Brodde and Michel, 1999; Elhusseiny and Hamel, 2000).

the M_2 receptor is THRX-160209, a bitropic muscarinic ligand (Kruse et al., 2014). Bitropic ligands are rationally designed hybrid molecules combining the high binding affinity of orthosteric ligands with the receptor subtype selectivity of allosteric ligands by interacting with both topographically distinct binding sides (Kruse et al., 2014). Besides the big advantage of a highly affine and subtype-selective ligand, in contrast to orthosteric ligands, bitropic ligands have no well defined structure-activity relationships and they are not as signaling-selective as allosteric modulators (Kruse et al., 2014).

In contrast to the external ligand-based receptor modulations, there are also internal regulatory strategies (Gurevich and Gurevich, 2008). After the receptors are delivered either as monomers or oligomers, they bind to the membrane and can build dimers and higher oligomers or stay as monomers (Gurevich and Gurevich, 2008). The kind and size of the receptor complex might have an important effect on the cross-talk with other receptors and on general receptor regulation (Katritch et al., 2013; Gurevich and Gurevich, 2008). But it is still not clear how and why receptors build such complexes (Gurevich and Gurevich, 2008). Besides the ability of the receptors to build complexes or clusters on the cell surface, there are second messenger-mediated mechanisms that decrease the signaling via receptor phosphorylation, which are called desensitization (Pierce et al., 2002). There are several proteins that are able to directly down-regulate the receptor activity, such as protein kinase A (PKA) and protein kinase C (PKC) (Pierce et al., 2002). Afterwards the phosphorylated receptor directly dissociates from the bound GP (Pierce et al., 2002). A more agonist-specific regulation is mediated by G protein-coupled receptor kinases (GRK) only targeting the agonist-occupied receptor conformation controlled by several factors (Pierce et al., 2002). There are a lot of other mechanisms of receptor desensitization, such as the receptor degradation or regulation of the gene transcription, which consequentially all lead to a temporary decrease of the signal (Pierce et al., 2002). But there are also mechanisms leading to a redistribution of the receptor via internalization (Pierce et al., 2002). Receptor internalization can either lead to the degradation or the dephosphorylation of the receptor (Pierce et al., 2002). Hence a positive or negative feedback might occur where the receptor is internalized and transported to the endosome, e.g., via vesicles (Pierce et al., 2002). The vesicle-based transport is mediated via β-arrestin (Pierce et al., 2002). After the ligand has bound to the dephosphorylated receptor and hence the receptor got activated, the signal is transduced by heterotrimeric GPs (Oldham and Hamm, 2008; Dale and Rang, 2011; Kruse et al., 2014).

3.2 G Protein-mediated Signaling

Once activated, GPCRs bind to a small number of GPs (Oldham and Hamm, 2008; Dale and Rang, 2011). GPs are composed of α, βand γ subunits which are bound in the inactive state of the GP (Oldham and Hamm, 2008). There are 21 different α, 6 different β and 12 different γ subunits, as shown in Table 3.2 (Oldham and Hamm, 2008). All α subunits feature a GTPase binding side and a helical domain (Oldham and Hamm, 2008). GTPase represents a family of enzymes which hydrolyze guanosine-5'-triphosphate (GTP) (Alberts et al., 2014). The GTPase binding side is typical for all members of the GP superfamily allowing them to bind to the β\γ complex and is part of the domain responsible for binding and hydrolyzing GTP (Oldham and Hamm, 2008). The second domain of the α subunit is to regulate membrane localization and interaction with other proteins involved in the signaling (Oldham and Hamm, 2008). In contrast, the β and γ subunit build a functional complex which can only be dissociated under denaturing conditions (Oldham and Hamm, 2008). The interaction of the α subunit with the β\γ complex and the binding of the β subunit to the γ subunit is, in general, non-specific, but there is also evidence for a direct interaction of the α subunit with the γ subunit (Oldham and Hamm, 2008).

The β\γ-bound α subunit is not able to interact with other molecules (Oldham and Hamm, 2008). Hence, in the bound state, neither the β\γ subunit nor the α subunit are able to transmit any signal (Oldham and Hamm, 2008). As illustrated in Figure 3.1, in this inactive state, guanosine-5'-diphosphate (GDP) is bound to the α subunit (Dale and Rang, 2011). This causes a high binding affinity with the receptor (Oldham and Hamm, 2008). The binding of the ligand to the receptor leads to a structural change within the α subunit (Oldham and Hamm, 2008). This inter-domain movement then results in a release of GDP, and GTP can bind to the α

Subtypes	Main Targets	Distribution
α_s	• adenylyl cyclases • calcium channels • Src tyrosine kinases	• ubiquitous
$\alpha_{i\backslash o}$	• adenylyl cyclases • ERK/MAP kinase • calcium channels • potassium channels • Src tyrosine kinases	• neurons • heart • others
α_q	• Bruton's tyrosine kinase • phospholipase C • potassium channels	• ubiquitous
$\beta\backslash\gamma$	• as other subtypes • GPCR kinases • protein kinase D	• ubiquitous

Table 3.2: Function and distribution of the main GP subunits (Dale and Rang, 2011; Milligan and Kostenis, 2006).

subunit (Oldham and Hamm, 2008). Because the activation of GP is influenced by various factors, kinetic data is difficult to obtain and difficult to compare (Oldham and Hamm, 2008). The α-β\γ complex is mainly localized at the cellular membrane to decrease the spontaneous release of GDP and hence to reduce the basal activity of GPs (Kimple et al., 2011). The targets of GP-signaling are active until the α subunit-bound GTP is hydrolyzed (Kimple et al., 2011). This process is enhanced by several downstream proteins such as PLCβ1 and is necessary for controlling the ligand-induced signal (Kimple et al., 2011). The velocity of the α subunit deactivation via hydrolysis is drastically increased by RGS proteins, a large family of α subunit-directed GTPase-accelerating proteins (Kimple et al., 2011; Oldham and Hamm, 2008; Xie and Palmer, 2007). There are over 20 subtypes of mammalian RGS proteins which are

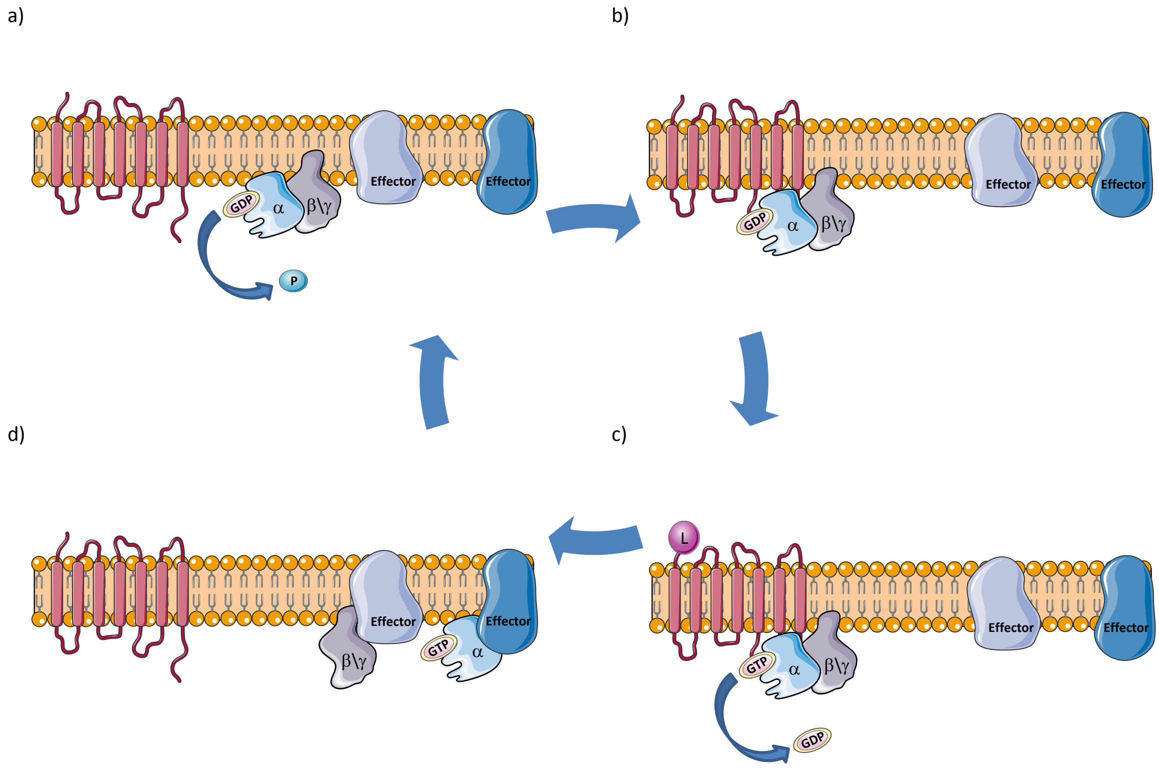


Figure 3.1: Illustration of the basic G protein activation cycle. (a) The GDP-bound α subunit shows a high affinity for binding with the β\γ subunit. The resulting α-β\γ complex is not able to interact with other downstream proteins. (b) The inactive G protein complex consisting of the α subunit and β\γ subunits spontaneously binds to the receptor with high affinity. (c) The ligand-induced structural change of the receptor catalyzes the release of GDP and the subsequent binding of GTP. (d) After the GDP is replaced with GTP, the GTP-bound α subunit and the β\γ dimer dissociate. In their free form, both are able to independently interact with downstream targets. After activation of the target protein, the hydrolysis of the bound GTP to GDP through elimination of a phosphate residue is initiated spontaneously or via other mechanisms. The GDP-bound α subunits again show a high affinity for binding with the β\γ subunit (Dale and Rang, 2011; Milligan and Kostenis, 2006; Kimple et al., 2011).

able to regulate GPCRs and other proteins very specifically and selectively but with varying effectiveness (Kimple et al., 2011; Xie and Palmer, 2007). Recent studies demonstrated that the selectivity of RGS proteins can be explained by their co-expression with the target at the right time and location (Xie and Palmer, 2007).

Once activated, GP subunits transduce the ligand-mediated signal to several downstream proteins (Dale and Rang, 2011). One of the most important targets for GP-induced signaling are various subtypes of adenylyl cyclase (AC) (Sunahara and Taussig, 2002; Pierce et al., 2002). For mammals, nine membrane-bound and one soluble isoform are known (Sunahara and Taussig, 2002; Pierce et al., 2002). The membrane-bound isoforms can be structured into two transmembrane clusters and two cytoplasmic loops (Cooper, 2003; Hurley, 1999). With the help of metal ions, AC catalyzes the conversion of adenosine triphosphate (ATP) to cAMP and pyrophosphate (Sunahara and Taussig, 2002; Pierce et al., 2002; Hurley, 1999). The AC activity is mainly regulated by α subunits, depending on the isoform and the GP subunit (Sunahara and Taussig, 2002). In general, α_s subunits have stimulatory and α_i subunits have inhibitory effects (Sunahara and Taussig, 2002). But there are also other regulators of the AC activity, e.g., the $\beta\backslash\gamma$ subunits, PKA and Ca^{2+} ions (Sunahara and Taussig, 2002). Since GDP-bound α subunits have a tenfold lower affinity for AC, the GP-mediated catalysis is timed by the hydrolysis of GTP, which terminates the GP-mediated signaling (Sunahara and Taussig, 2002). The hydrolysis and therefore the termination of the signal is enhanced by RGS proteins (Sunahara and Taussig, 2002). Besides the PKA-mediated inhibition of AC by phosphorylation, forskolin is a very potent stimulator which is widely used for in vitro experiments (Sunahara and Taussig, 2002; Pierce et al., 2002; Hurley, 1999).

As a consequence of the production of cAMP catalyzed by AC, the concentration of the second messenger cAMP within the cytoplasma increases (Sunahara and Taussig, 2002; Pierce et al., 2002). It is known that cAMP is not uniformly distributed in the cytoplasm, i.e., the concentration of cAMP next to AC and PKA differs from the overall cytosol concentration, which may lead to a compartmentalization of the cAMP concentration (Cooper, 2003; Kholodenko, 2006). The second messenger molecule cAMP directly activates the enzyme PKA (Sunahara and Taussig, 2002; Pierce et al., 2002). In the inactive state, PKA builds a complex comprised of two catalytic subunits and one regulatory subunit dimer (Taylor et al., 2012). There are four known isoforms of the regulatory subunit and three known isoforms of the catalytic subunits (Taylor et al., 2012). PKA downstream signaling is the main pathway induced by GPs (Taylor et al., 2012). The PKA activity is directly controlled by the cAMP level (Taylor et al., 2012). In the inactive state, the regulatory dimer blocks the binding sites of the catalytic subunits making a binding to other substrates impossible (Taylor et al., 2012). Hence the level of PKA activity depends on the local cAMP concentration, and the hydrolysis of cAMP catalyzed by 3',5'-cyclonukleotid-phosphodiesterase (PDE) decreases the activity of PKA (Maurice et al., 2014). There are eleven mammalian PDE families which widely differ in their function (Maurice et al., 2014). Hence, when PDE gets activated by PKA, it directly regulates the downstream cAMP-induced signaling and can be considered as a negative feedback loop (Omori and Kotera, 2007). PKA is then able to activate a large number of enzymes such as RGS proteins or GRKs (Alberts et al., 2014).

3.3 Experimental Tools

A wide range of techniques to measure responses mediated by external stimuli at cellular levels is coupled to very specific cellular events (Dale and Rang, 2011). Most of them are invasive and might cause significant changes in the cellular physiology (Fang et al., 2006). The second messenger cAMP is a prominent target for this kind of readouts (Engelhardt et al., 2017). However, measuring the concentration of cAMP only allows for conclusions about the receptor-mediated downstream cascade from the receptor above the production of cAMP (Dale and Rang, 2011). The limited transferability to native systems and the restricted conclusions are a drawback for drug discovery and development (Fang et al., 2006). To ensure medical applicability, it is important to understand the functions of cellular targets under native conditions (Fang et al., 2006). Today, several types of optical biosensors allow to monitor cellular signaling processes non-invasively in real time and thus lead to an improved understanding of cell biology and physiology (Schröder et al., 2010; Schröder et al., 2011; Fang et al., 2006). Over the last decades, the number and range of biosensors have drastically increased (Fang et al., 2006). In the past, several optical principles were used to investigate ligand-induced receptor responses in membrane homogenates (Fang et al., 2006). The resonant waveguide grating (RWG) technique has lead to the development of novel mass redistribution cell assay technologies (Fang et al., 2006). It allows to screen the DMR within the bottom portion of living cells (Schröder et al., 2010; Schröder et al., 2011; Fang et al., 2006).

As illustrated in Figure 3.2, broadband light enters the waveguide resonant grating (Rosenblatt et al., 1997; Fang et al., 2006). The intensity of the reflected light can be considered to be a function of the resonant wavelength (Rosenblatt et al., 1997; Fang et al., 2006). The native cell can be placed on the grating to measure the optical density. The waveguide resonant grating basically consists of a substrate layer and a waveguide film with an embedded grating structure (Rosenblatt et al., 1997; Fang et al., 2006). Because the height of the cells is mostly beyond the wavelength of the incident light and larger than the penetration depth, the sensing volume is restricted to the bottom of the cell not necessarily attached to the waveguide (Schröder et al., 2010; Schröder et al., 2011; Fang et al., 2006). In contemporary devices, the broadband light beam illuminates the whole waveguide, which allows for instantaneous measurements, whereas previously the beam only illuminated the waveguide punctually, resulting in delayed measurements (Fang et al., 2006). As demonstrated in several studies, the mass reorganization within the sensing volume changes the incident angle of the incoupled light and hence leads to a shift of the wavelength of the outcoupled light (Schröder et al., 2010; Schröder et al., 2011; Fang et al., 2006).

Although the underlying details of the DMR are not yet known, it allows for dividing the complex GPCR-mediated signal patterns into its separated subsequent GP downstream pathways (Schröder et al., 2010; Schröder et al., 2011). Through inhibition of concurrent GP-mediated downstream pathways with several inhibitors, a unique fingerprint of the concurrent pathways was found (Schröder et al., 2010; Schröder et al., 2011). Common inhibitors for GP signaling are pertussis toxin (PTX), YM-254890 (YM) and cholera toxin (CTX) (Schröder et al., 2010; Schröder et al., 2011). It is important to note that PTX, YM and CTX induce

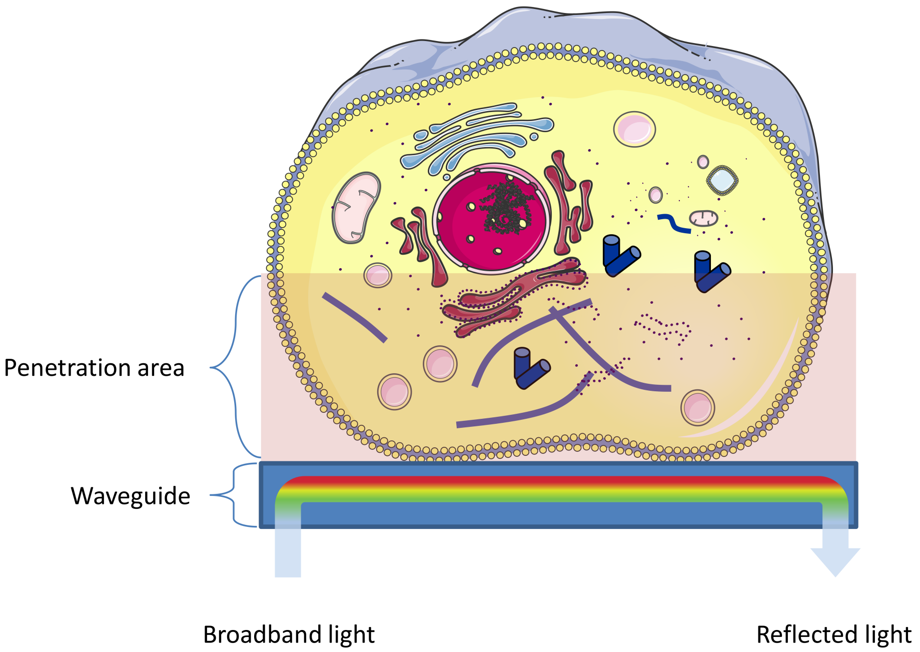


Figure 3.2: Illustration of the resonant waveguide grating of living cells. The broad band light enters the waveguide and penetrates the bottom area of the cells. Depending on spatial and temporal events related to protein trafficking and cytoskeletal rearrangements, the wavelength of the reflected light is changed. The difference in wavelength between the reflected light before and after the treatment is related to the shift in the optical density of the cell (Schröder et al., 2010; Schröder et al., 2011).

no measurable response on their own (Schröder et al., 2010; Schröder et al., 2011). Rather, the response varies with the cell and receptor subtype (Schröder et al., 2010; Schröder et al., 2011). The GPCR-induced DMR signatures are a result of highly regulated complex spatial and temporal events, such as protein trafficking and cytoskeletal rearrangements (Schröder et al., 2010; Schröder et al., 2011; Fang et al., 2006). Protein trafficking, Golgi-derived transport carriers and cytoskeletal rearrangements themselves are known to be closely related to the actin microfilaments and microtubules (Etienne-Manneville, 2004; Schmidt and Hall, 1998; Hammond et al., 2008; Egea et al., 2006).

Whereas DMR measurements reflect the overall cellular response, they provide no insight into the causal biochemical mechanisms on the protein level (Schröder et al., 2010; Schröder et al., 2011). In contrast, WB is a commonly used technique which allows to derive information about the dynamic behavior of single proteins (Towbin et al., 1979; Renart et al., 1979). Prior to the actual WB, i.e., the transfer of proteins from a gel to a membrane, the proteins within the sample, for instance, prepared lysates, have to be separated according to their electrophoretic mobility (Towbin et al., 1979; Smithies, 1955). Gel electrophoresis is a standard method used to separate macromolecules because the motion of molecules through the gel from the cathode to the anode varies in speed depending on the size and charge of the molecules (Schägger, 2006;

Robyt and White, 1987; Smithies, 1955). The gel used in sodium dodecyl sulfate polyacrylamide electrophoresis has a filter structure leading to a slow motion of big molecules and a separation of molecules relative to their size (Schägger, 2006; Robyt and White, 1987). The gel-trapped molecules can be transferred to a membrane via WB (Robyt and White, 1987; Renart et al., 1979). Therefore, a membrane is placed over the gel and, under application of an electric field, the molecules move towards the anode from the gel to the membrane (Robyt and White, 1987; Renart et al., 1979).

After transfer, the membrane-bound molecules are still separated and do not change their location compared to their absolute position on the gel (Towbin et al., 1979; Renart et al., 1979). Once the molecules, e.g., proteins, are transferred to the membrane, they can be marked and thus highlighted via specific antibodies (Robyt and White, 1987). The marked proteins are now ready for quantification with the help of a special camera systems (Robyt and White, 1987). Because gel electrophoresis and WB are highly sensitive techniques, it is possible to determine the protein concentration based on the intensity of the markers (Robyt and White, 1987). The light intensity of the marker and the protein concentration are related because the light intensity of the marker increases with the number of marked molecules (Robyt and White, 1987). The underlying idea is to wash the membrane after adding the antibodies so that only marked antibodies bound to the specific antigen, i.e., protein, will remain on the membrane (Robyt and White, 1987).

CHAPTER 4

Methodology

In order to reconstruct the M_2 receptor-induced signaling in CHO cells and thus the inherent signal amplification, well-known methods from metabolic pathway reconstruction are extended to signaling cascades in this work. The stoichiometric matrix provides information about the relevance of the involved species and subpathways (Engelhardt et al., 2017). As mentioned in Chapter 2, signaling is encoded in the dynamic behavior of a system and hence, in addition to steady state analyzes, a dynamic model is necessary for further investigations (Hübner et al., 2011; Klipp and Liebermeister, 2006).

Stoichiometric analyses of biological systems mainly based on algebraic tools have been employed for a long time (Wei and Prater, 1962; Palsson, 2006). As illustrated in Chapter 2, for model development, it is reasonable to start with the analysis of the network topology and the behavior of the system under steady state conditions based on elementary reactions (Wei and Prater, 1962; Palsson, 2006). Afterwards, the model should be complemented with advanced kinetics to incorporate the dynamic behavior (Azeloglu and Iyengar, 2015; Hübner et al., 2011). Here this strategy, which is similar to the bottom-up approach discussed in Section 2.1, is required because the amount of reliable knowledge and data is restricted (Bruggeman and Westerhoff, 2007; Azeloglu and Iyengar, 2015; Engelhardt et al., 2017).

4.1 Stoichiometric Network Analysis

A reconstruction of biochemical networks involves both the network topology and the kinetic rates (Azeloglu and Iyengar, 2015; Palsson, 2006). In case of small- or medium-sized networks, ODE-based dynamical systems are often employed. However, topology-based approaches can also yield valuable insights into the mechanisms and basic features of biochemical networks (Wei and Prater, 1962; Heinrich et al., 2002). Especially metabolic networks are frequently analyzed based on their stoichiometry (Palsson, 2006; Bordbar et al., 2014). Metabolic networks can be sufficiently characterized by certain steady states (Palsson, 2006). This assumption is justified by the high relaxation time of metabolic systems, which is typically in the order of minutes or seconds (Palsson, 2006). In contrast, a transient behavior is often observable for signaling networks (Klamt et al., 2006). Hence, as a method, structural analysis is well-established for metabolic, often genome-scale, networks but not for signaling networks (Klamt et al., 2006). The topology of a CRN, including the stoichiometric coefficients, is encoded in the stoichiometric matrix from Definition 2. A general definition of the stoichiometric matrix is given in Section 2.3.4. As detailed in Figure 4.1, the stoichiometric matrix as a function maps

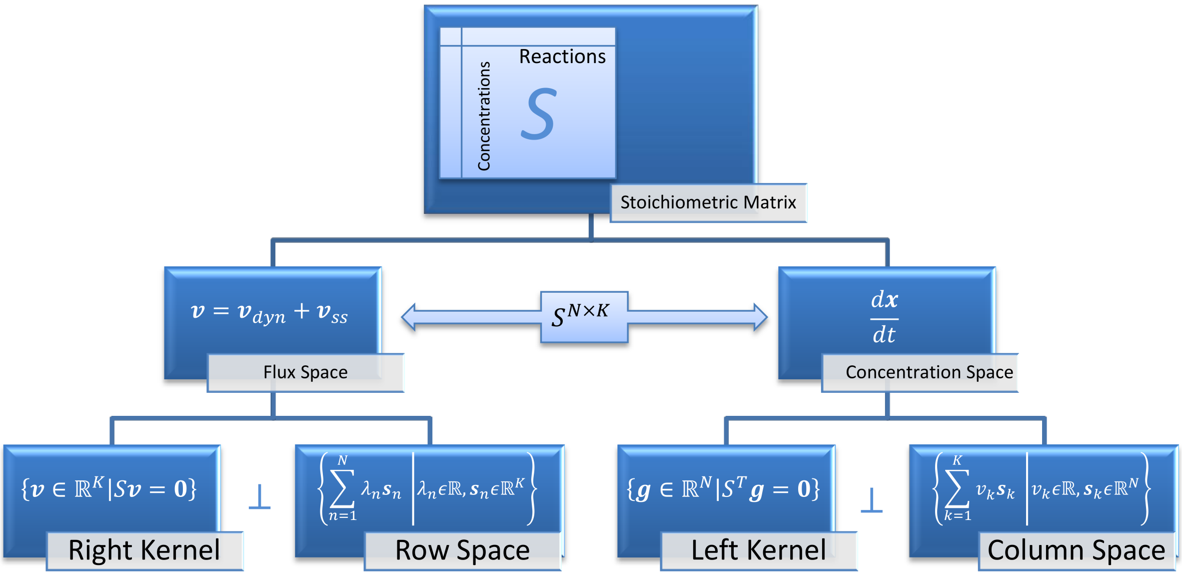


Figure 4.1: Illustration of the concept behind the stoichiometric analysis of chemical reaction networks. The stoichiometric matrix S can be considered as a function mapping the flux space to the space of concentration changes (here denoted as concentration space for reasons of simplicity). The kernel of this linear function includes the fluxes under steady state conditions because they are mapped to zero. All dynamic fluxes are described in the row space and hence they are mapped to the concentration space. The image of the linear mapping is equivalent to the concentration space, which can be divided into variable species and constant species relationships. The latter are represented by the left kernel, whereas all variable species relationships are part of the column space (Palsson, 2006).

the flux space to the space of concentration changes. The flux space is composed of all constant fluxes $\boldsymbol{v}_{ss}$ mapped to zero and all dynamic fluxes $\boldsymbol{v}_{dyn}$ which are not mapped to zero and thus not in steady state. The column space of the stoichiometric matrix or the span of $\dot{\boldsymbol{x}}$ can be decomposed into $\sum_{k=1}^{K} \boldsymbol{s}_k v_k$ with $\boldsymbol{v} \in \mathbb{R}^K$ and the column vectors $\boldsymbol{s}_k \in \mathbb{R}^N$ of the stoichiometric matrix $S \in \mathbb{R}^{N \times K}$ (Palsson, 2006). For every time point t with $\dot{\boldsymbol{x}} \approx 0$, each reaction velocity r_k of the CRN is represented by a constant integer value v_k. Consequently, each entity v_k of the flux vector corresponds to the overall reaction rate of the corresponding reaction at an assigned steady state (Palsson, 2006). More general, the flux vector $\boldsymbol{v}$ with $\dot{\boldsymbol{x}} = S\boldsymbol{v}$ is composed of a dynamic component $\boldsymbol{v}_{dyn}$ and a static component $\boldsymbol{v}_{ss}$ such that (Wei and Prater, 1962; Palsson, 2006)

$$\frac{d\boldsymbol{x}}{dt} = S\boldsymbol{v}_{dyn} + S\boldsymbol{v}_{ss}. \tag{4.1}$$

Due to the fast time-scale of metabolic systems, the right and left null space, and hence

$$\frac{d\boldsymbol{x}}{dt} = \boldsymbol{0}, \tag{4.2}$$

are in focus of the analysis. In the steady state, i.e., $\dot{\boldsymbol{x}} = \boldsymbol{0}$, the whole system is approximately in balance, though there is a continuous flow of the component materials (Wei and Prater, 1962; von Bertalanffy, 1950). Even if the whole system is not in a steady state, a subsystem with $S\boldsymbol{v}_{ss} \approx 0$ may exist (Wei and Prater, 1962; Palsson, 2006). This work demonstrates that the steady state analysis can also be extended to complex signaling networks (Engelhardt et al., 2017).

4.1.1 Left and Right Null Space

Since every column of the stoichiometric matrix corresponds to a reaction and every row corresponds to a species, investigating the left and right null space of the stoichiometric matrix allows to gain insights into the reaction and species dependencies, as illustrated in Figure 4.1 (Wei and Prater, 1962; Palsson, 2006).

Definition 6: Left and Right Null Space

Given a CRN with a stoichiometric matrix $S \in \mathbb{R}^{N\times K}$, the right null space or kernel $\ker(S)$ is given by

$$\ker(S) = \left\{\boldsymbol{v} \in \mathbb{R}^{K} \,|S\boldsymbol{v} = \boldsymbol{0}\right\}.$$

Accordingly, the left null space $\ker\left(S^{T}\right)$ is given by

$$\ker\left(S^{T}\right) = \left\{\boldsymbol{g} \in \mathbb{R}^{N} \,\middle|S^{T}\boldsymbol{g} = \boldsymbol{0}\right\}.$$

For a CRN with K single reactions and N species, using common definitions from linear algebra, the number of reactions K is equal to $\dim(\ker(S)) + \dim(\operatorname{im}(S))$ (Palsson, 2006). In addition, the image of the mapping S is defined as

$$\operatorname{im}(S) = \left\{\boldsymbol{x} \in \mathbb{R}^{N} \,\middle|\dot{\boldsymbol{x}} = S\boldsymbol{v} \quad \boldsymbol{v} \in \mathbb{R}^{K}\right\}. \tag{4.3}$$

Therefore, the dimension of $\ker(S)$ is equivalent to the number of independent fluxes and the dimension of $\operatorname{im}(S)$ is equivalent to the number of dependent fluxes (Palsson, 2006). In contrast, the number of species N depends on $\dim\left(\ker\left(S^{T}\right)\right)$ and $\dim\left(\operatorname{im}\left(S^{T}\right)\right)$. The column space of S is orthogonal to the kernel of S^{T} and hence the space of concentration changes can be partitioned into the column space and the left null space (Palsson, 2006).

Right Null Space

The behavior of the system under steady state conditions is characterized by the right null space (Palsson, 2006). It is always possible to find a set of basis vectors $\boldsymbol{b} \in \mathbb{R}^{K}$ for the right null space such that

$$\boldsymbol{v} = \sum_{j=1}^{(K-\dim(\operatorname{im}(S)))} \alpha_j \boldsymbol{b}_j \;\; \forall \boldsymbol{v} \colon S\boldsymbol{v} = \boldsymbol{0}, \tag{4.4}$$

with $\alpha_i \in \mathbb{R}$ (Llaneras and Picó, 2010). The basis vectors $\boldsymbol{b}_j$ are not necessarily positive and can therefore not be directly interpreted as fluxes in a biological manner (Wei and Prater, 1962; Palsson, 2006). In contrast, a convex representation of the kernel can often be directly mapped to pathways of the CRN (Llaneras and Picó, 2010). This is due to the fact that a flux, i.e., the rate of turnover of molecules, through a biological system is unidirectional and thus always non-negative (Llaneras and Picó, 2010). A set of non-negative basis vectors leads to concepts based on convex analysis (Palsson, 2006; Clarke, 1980). All possible steady state fluxes are enclosed by a polytope spanned by a set of conically independent generating vectors, where the number of generating vectors in this set can exceed the dimension of the right null space (Palsson, 2006; Clarke, 1980). In contrast to Equation (4.4), the space of feasible steady states is in general given by

$$P = \left\{ \boldsymbol{v} \in \mathbb{R}^K \, | S\boldsymbol{v} = \boldsymbol{0}, D\boldsymbol{v} \geq \boldsymbol{0} \right\}, \tag{4.5}$$

with a diagonal matrix $D \in \mathbb{R}^{K \times K}$, where $D_{ii} = 1$ if the flux is irreversible and $D_{ii} = 0$ otherwise (Llaneras and Picó, 2010). This is important because often two opposed irreversible reactions are combined to one reversible reaction in order to reduce the dimension of the stoichiometric matrix. The composition of the related polytope is commonly addressed by extreme currents or EFMs (Palsson, 2006; Clarke, 1980).

Left Null Space

Whereas $\ker(S)$ and the row space help to explain the flux distribution, the left kernel $\ker\left(S^T\right)$ helps to explain species dependencies (Wei and Prater, 1962; Palsson, 2006). The left kernel adds all time derivatives up to zero (Wei and Prater, 1962; Palsson, 2006). Hence, for a convex basis, the time-invariant quantities lead to conservation relationships (Wei and Prater, 1962; Palsson, 2006). Conservation relationships are all those vectors **g** with

$$\boldsymbol{g} \in \left\{ \boldsymbol{y} \in \mathbb{R}^N : S^T \boldsymbol{y} = \boldsymbol{0} \right\}. \tag{4.6}$$

This implies

$$\boldsymbol{g}^T \dot{\boldsymbol{x}} = \boldsymbol{g}^T S \boldsymbol{v} = \left(S^T \boldsymbol{g}\right)^T \boldsymbol{v} = \boldsymbol{0} \tag{4.7a}$$

$$\Rightarrow \frac{d}{dt}\left(\sum_{n=1}^{N} g_n x_n\right) = 0 \tag{4.7b}$$

$$\Rightarrow \boldsymbol{g}^T \boldsymbol{x} = \sum_{n=1}^{N} g_n x_n = const. \tag{4.7c}$$

Each conservation relationship $\boldsymbol{g}$ represents a linear combination of constant species (Wei and Prater, 1962; Palsson, 2006). They can be interpreted as a cycle of mass conservation, e.g., the conservation of phosphate groups within a certain subpathway (Wei and Prater, 1962; Palsson, 2006). A convex representation of $\ker\left(S^T\right)$ seems natural because the involved quantities are related to species concentrations which are always greater or equal to zero (Wei and Prater, 1962; Palsson, 2006). The conservation relationships hold true for all network states and thus

$$\boldsymbol{g}^T\boldsymbol{x} = \boldsymbol{g}^T\boldsymbol{x}_{ref} = const. \tag{4.8}$$

or

$$\boldsymbol{g}^T\left(\boldsymbol{x} - \boldsymbol{x}_{ref}\right) = 0. \tag{4.9}$$

In contrast to the original affine hyperspace resulting from the inhomogeneous linear system $\boldsymbol{g}^T\boldsymbol{x} = const.$, the introduced reference vector $\boldsymbol{x}_{ref}$ transforms the concentration space into a non-affine hyperspace. Consequently, the basis of this system is equivalent to $\ker\left(S^T\right)$ (Palsson, 2006).

4.1.2 Constraint-based Modeling

In addition to the left kernel, the image of the linear mapping related to the stoichiometric matrix is spanned by the column space of S. The decomposition $\dot{\boldsymbol{x}} = \sum_{k=1}^{K} \boldsymbol{s}_k v_k$ and the fact that the fluxes v_k are constrained by finite values $\boldsymbol{\alpha} \leq \boldsymbol{v} \leq \boldsymbol{\beta}$ lead to a constrained equation system (Palsson, 2006; Bordbar et al., 2014). The conservation relationships introduce additional constraints, resulting in a linear equation system

$$\frac{d\boldsymbol{x}}{dt} = S\boldsymbol{v} = \boldsymbol{b} \tag{4.10a}$$

$$\boldsymbol{\alpha} \leq \boldsymbol{v} \leq \boldsymbol{\beta} \tag{4.10b}$$

$$\frac{d\boldsymbol{g}^T\boldsymbol{x}}{dt} = 0, \tag{4.10c}$$

with real-valued vectors $\boldsymbol{\alpha} \leq \boldsymbol{\beta}$ (Orth et al., 2010; Bordbar et al., 2014). The interval $[\boldsymbol{\alpha},\boldsymbol{\beta}]$ is often used to define core reactions which are forced to be active and hence $0 \notin [\alpha_i,\beta_i]$. For some biological problems, this leads to an optimization task when including an objective function, e.g., $\max\left(\boldsymbol{c}^T\boldsymbol{v}\right)$ (Becker and Palsson, 2008). The objective function mostly describes the growth rate or the ATP production, which are optimized in some biological settings (Nam et al., 2014). For instance, in cancer cells, cell proliferation and ATP production are often assumed to be maximized (Nam et al., 2014). An example from biotechnology is the regulation of the yeast metabolism such that defined metabolites are optimized (Brochado et al., 2010; Vargas et al., 2011). These phenomena are typically addressed via a flux balance analysis (FBA) (Orth et al., 2010). Here, the system given through Equation (4.10) is optimized through

$$\max\left(\boldsymbol{c}^T \boldsymbol{v}\right) \tag{4.11a}$$

$$\frac{d\boldsymbol{x}}{dt} = \boldsymbol{0} = S\boldsymbol{v} \tag{4.11b}$$

$$\boldsymbol{\alpha} \leq \boldsymbol{v} \leq \boldsymbol{\beta}, \tag{4.11c}$$

with an objective $\boldsymbol{c}^T\boldsymbol{v}$ encoded by a real-valued vector $\boldsymbol{c}$ with $c_k \neq 0$ for all fluxes which are assumed to be maximized. Vice versa $\boldsymbol{c}^T\boldsymbol{v}$ can also be minimized if necessary.

The solution space of the system described through Equations (4.10) is large and of irregular shape (Schellenberger and Palsson, 2009; Palsson, 2006). In consequence, it can have an infinite number of solutions (Palsson, 2006; Fischer, 2013). A promising approach to address this issue is to describe the constraints a system has to satisfy rather than computing an explicit solution $\dot{\boldsymbol{x}} = \boldsymbol{b}$ (Schellenberger and Palsson, 2009; Palsson, 2006). Therefore, the solution space is reduced to feasible solutions (Schellenberger and Palsson, 2009; Palsson, 2006). As for the FBA, an important case described through Equations (4.10) is given by $\dot{\boldsymbol{x}} = \boldsymbol{0}$ (Palsson, 2006). The steady state is of polyhedral shape and therefore not as irregular as the general solution space for constraint-based models (Schellenberger and Palsson, 2009; Palsson, 2006). However, the calculation of the solution space for $\dot{\boldsymbol{x}} = \boldsymbol{0}$ and the left kernel are non-trivial tasks and commonly addressed by MCMC algorithms (Schellenberger and Palsson, 2009; Megchelenbrink et al., 2014). A brief description of the latter is provided in Section 4.3.2. MCMC methods allow to determine the size and shape of the steady state solution space and provide information about the probabilities of feasible solutions (Wiback et al., 2004). The aim of a steady state analysis is to explore all possible and biologically relevant paths through the system in the steady state and calculate their probabilities (Wiback et al., 2004; Schellenberger and Palsson, 2009; Palsson, 2006).

4.2 Optimal Control Theory

OCT is very well known in engineering and also frequently applied to biological systems, especially to optimize bioprocesses in biotechnology and bioengineering (Paiewonsky, 1965; Crassidis and Junkins, 2011; Banga et al., 2005). The main intention is to control the trajectories of the variables in a dynamical system (Luenberger, 1979). In contrast to control theory, where the dynamic behavior and the modification via feedback loops are in focus, OCT asks for an optimal modification with respect to a certain objective (Sontag, 2005). Thus OCT is closely related to optimization and the calculus of variations (Luenberger, 1979).

In more general terms, a dynamical system and a set of initial values is given according to Definition 5. For the purpose of optimal control, it is necessary to explicitly distinguish between the states, i.e., dynamic model variables $\boldsymbol{x}(t)$ and additional inputs $\boldsymbol{u}(t)$ (Luenberger, 1979; Martinelli, 2015). The desired objective functional $J(\boldsymbol{x}(t), \boldsymbol{u}(t))$, gives rise to a general first-order optimal control problem (OCP), as described below (Luenberger, 1979).

Definition 7: Optimal Control Problem

For a system of first-order differential equations, a vector-valued function $\boldsymbol{f}(\boldsymbol{x}(t), \boldsymbol{u}(t)) : \mathbb{R}^N \times \mathcal{U} \to \mathbb{R}^N$ with a set of admissible controls $\mathcal{U} \subset \mathbb{R}^P$, a time-dependent state-vector $\boldsymbol{x}(t) \in \mathbb{R}^N$, a set of initial values $\boldsymbol{x}(0) = \boldsymbol{x}_0$, known inputs $\boldsymbol{u}(t) \in \mathcal{U}$ and an objective functional $J(\boldsymbol{x}(t), \boldsymbol{u}(t)) : \mathbb{R}^N \times \mathcal{U} \to \mathbb{R}$ and $t = [0, F]$, the optimal control problem is given by

$$\begin{aligned} \frac{d\boldsymbol{x}}{dt} &= \boldsymbol{f}(\boldsymbol{x}(t), \boldsymbol{u}(t)) \\ \boldsymbol{x}(0) &= \boldsymbol{x}_0 \\ J(\boldsymbol{x}(t), \boldsymbol{u}(t)) &= \phi(\boldsymbol{x}(F)) + \int_0^F l(\boldsymbol{x}(t), \boldsymbol{u}(t))\, dt \\ &\max_{\mathbf{u} \in \mathcal{U}} \left[J(\boldsymbol{x}(t), \boldsymbol{u}(t))\right], \end{aligned}$$

with a terminal constraint $\phi(\boldsymbol{x}(F)) \in \mathbb{R}$ and a real-valued objective function $l(\boldsymbol{x}(t), \boldsymbol{u}(t)) : \mathbb{R}^N \times \mathcal{U} \to \mathbb{R}$.

For general biological systems with observations $\boldsymbol{y}(t) = \boldsymbol{h}(\boldsymbol{x}(t))$ and $\boldsymbol{y}(t) \in \mathbb{R}^I$, the objective functional can be stated as

$$J(\boldsymbol{x}(t), \boldsymbol{u}(t)) = \phi(\boldsymbol{h}(\boldsymbol{x}(F))) + \int_0^F l(\boldsymbol{h}(\boldsymbol{x}(t)), \boldsymbol{u}(t))\, dt, \tag{4.12}$$

and the corresponding OCP is

$$\min_{\mathbf{u} \in \mathcal{U}} \left[J(\boldsymbol{x}(t), \boldsymbol{u}(t))\right]. \tag{4.13}$$

Thus, the objective functional depends on a combination $\boldsymbol{h}(\boldsymbol{x}(t))$ of the state variables $\boldsymbol{x}(t)$. For biological systems, the observations and hence the function $\boldsymbol{h}(\boldsymbol{x}(t))$ often reflect a linear dependency of the state variables.

4.2.1 Derivation

Although a positive change in $J(\boldsymbol{x}(t), \boldsymbol{u}(t))$ caused by a small shift of $\boldsymbol{u}(t)$ is a sufficient condition for optimization, it is difficult to derive a direct method to address this problem because even an infinitely small change in $\boldsymbol{u}(t)$ may cause a change in $\boldsymbol{x}(t)$ (Luenberger, 1979). Vice versa, for the OCP given in Definition 7, a negative change in $J(\boldsymbol{x}(t), \boldsymbol{u}(t))$ indicates distancing from the optimal solution (Luenberger, 1979). To circumvent the direct determination of $J(\boldsymbol{x}(t), \boldsymbol{u}(t))$, an auxiliary variable $\boldsymbol{\lambda}(t) \in \mathbb{R}^N$ is introduced (Luenberger, 1979).

The auxiliary variable leads to the Hamiltonian function $H(\boldsymbol{\lambda}(t), \boldsymbol{x}(t), \boldsymbol{u}(t)) : \mathbb{R}^N \times \mathbb{R}^N \times \mathcal{U} \rightarrow \mathbb{R}$ with $\mathcal{U} \subset \mathbb{R}^P$ given by (Luenberger, 1979)

$$H(\boldsymbol{\lambda}(t), \boldsymbol{x}(t), \boldsymbol{u}(t)) = \boldsymbol{\lambda}^T(t) \boldsymbol{f}(\boldsymbol{x}(t), \boldsymbol{u}(t)) + l(\boldsymbol{x}(t), \boldsymbol{u}(t)). \tag{4.14}$$

Applying Equation (4.14), the functional $J(\boldsymbol{x}(t), \boldsymbol{u}(t))$ can be reformulated by adding zero in such a way that the modified functional

$$\tilde{J}(\boldsymbol{x}(t), \boldsymbol{u}(t), \boldsymbol{\lambda}(t)) = \phi(\boldsymbol{x}(F)) + \int_0^F \left[H(\boldsymbol{\lambda}(t), \boldsymbol{x}(t), \boldsymbol{u}(t)) - \boldsymbol{\lambda}^T(t) \frac{d\boldsymbol{x}(t)}{dt} \right] dt \tag{4.15}$$

becomes equivalent to $J(\boldsymbol{x}(t), \boldsymbol{u}(t))$ for each trajectory, but the freedom in the choice of $\boldsymbol{\lambda}(t)$ can be used to reduce the complexity of the OCP (Luenberger, 1979). For each trajectory, the identity $\dot{\boldsymbol{x}} - \boldsymbol{f}(\boldsymbol{x}(t), \boldsymbol{u}(t)) = \boldsymbol{0}$ holds true and hence it directly follows that

$$J(\boldsymbol{x}(t), \boldsymbol{u}(t)) = \tilde{J}(\boldsymbol{x}(t), \boldsymbol{u}(t), \boldsymbol{\lambda}(t)) \qquad \forall \boldsymbol{\lambda}(t) \in \mathbb{R}^N. \tag{4.16}$$

The obtained difference

$$\Delta\tilde{J} = \tilde{J}(\tilde{\boldsymbol{x}}(t), \tilde{\boldsymbol{u}}(t), \boldsymbol{\lambda}(t)) - \tilde{J}(\boldsymbol{x}(t), \boldsymbol{u}(t), \boldsymbol{\lambda}(t)), \tag{4.17}$$

where $\tilde{\boldsymbol{u}}(t) = \boldsymbol{u}(t) + \Delta\boldsymbol{u}(t)$ and $\tilde{\boldsymbol{x}}(t) = \boldsymbol{x}(t) + \Delta\boldsymbol{x}(t)$, serves as a measurement for the change in the objective functional caused by $\Delta\boldsymbol{u}(t)$. Trivially, $\Delta\boldsymbol{u}(t) = \boldsymbol{0}$ leads to $\Delta\tilde{J} = 0$. It is important to see that a change in $\tilde{\boldsymbol{u}}(t)$ causes a change in the states of the system, but even a large change of the input over a short interval of time may lead to a comparatively small change of the states for each time point (Luenberger, 1979). This is caused by the direct dependency on the integration of the objective function $l(\boldsymbol{x}(t), \boldsymbol{u}(t))$ (Luenberger, 1979). Please note that, when introducing additional state or control constraints, the explained variation is no longer directly solvable (Crassidis and Junkins, 2011; Luenberger, 1979).

In consequence, through Taylor approximation and considering the fact that the input has no influence on the initial values, i.e., $\Delta\boldsymbol{x}(0) = \boldsymbol{0}$, $\Delta\tilde{J}$ can be reformulated as

$$\begin{aligned} \Delta\tilde{J} = &\left(\frac{\partial\phi(\boldsymbol{x}(F))}{\partial\boldsymbol{x}} - \boldsymbol{\lambda}^T(F) \right) \Delta\boldsymbol{x}(F) \\ &+ \int_0^F \left[\frac{\partial H(\boldsymbol{\lambda}(t), \boldsymbol{x}(t), \boldsymbol{u}(t))}{\partial\boldsymbol{x}} + \frac{d\boldsymbol{\lambda}^T(t)}{dt} \right] \Delta\boldsymbol{x}(t)\, dt \\ &+ \int_0^F \left(H(\boldsymbol{\lambda}(t), \tilde{\boldsymbol{x}}(t), \tilde{\boldsymbol{u}}(t)) - H(\boldsymbol{\lambda}(t), \boldsymbol{x}(t), \boldsymbol{u}(t)) \right) dt + o(\Delta x), \end{aligned} \tag{4.18}$$

with the gradient $\frac{\partial}{\partial \mathbf{x}}$ and an additional error of the order $o(\Delta x)$, following the Bachmann–Landau notation, caused by the Taylor approximation and the application of the first-order differential quotient (Luenberger, 1979; Hanke-Bourgeois, 2008). The reformulated Equation (4.18) is still difficult to calculate because it depends on the final values of $\boldsymbol{x}(t)$ and $\boldsymbol{\lambda}(t)$ which change with $\Delta\boldsymbol{u}(t)$. The fact that $\boldsymbol{\lambda}(t)$ can be chosen arbitrarily can be used to eliminate all terms except for the last integral and thus to reduce the complexity of the system (Luenberger, 1979). Altogether, this leads to the adjoint differential equation

$$-\left[\frac{d\boldsymbol{\lambda}(t)}{dt}\right]^T = \frac{\partial H(\boldsymbol{\lambda}(t),\boldsymbol{x}(t),\boldsymbol{u}(t))}{\partial\boldsymbol{x}} \tag{4.19a}$$

$$\left[\frac{d\boldsymbol{\lambda}(F)}{dt}\right]^T = \frac{\partial\phi(\boldsymbol{x}(F))}{\partial\boldsymbol{x}}, \tag{4.19b}$$

which reduces Equation (4.18) to

$$\Delta\tilde{J} = \int_0^F \left(H(\boldsymbol{\lambda}(t),\tilde{\boldsymbol{x}}(t),\tilde{\boldsymbol{u}}(t)) - H(\boldsymbol{\lambda}(t),\boldsymbol{x}(t),\boldsymbol{u}(t))\right)dt + o(\Delta x)\,. \tag{4.20}$$

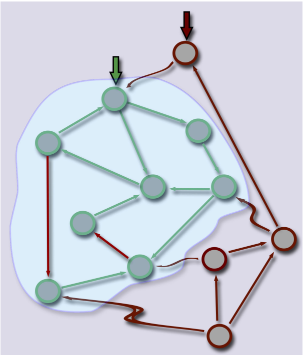
Equation (4.20) allows to establish a direct criterion for the improvement of $\Delta\tilde{J}$ because for any fixed $\boldsymbol{u} \in \mathcal{U}$ for $H(\boldsymbol{\lambda}(t),\tilde{\boldsymbol{x}}(t),\tilde{\boldsymbol{u}}(t)) \geq H(\boldsymbol{\lambda}(t),\boldsymbol{x}(t),\boldsymbol{u})$ the integral in Equation (4.20) is greater than or equal to zero and hence $\Delta\tilde{J}\big|_{\tilde{u}(t)} \geq \Delta\tilde{J}\,|_u$ due to the linearity of integrals and

$$H(\boldsymbol{\lambda}(t),\tilde{\boldsymbol{x}}(t),\tilde{\boldsymbol{u}}(t)) \geq H(\boldsymbol{\lambda}(t),\boldsymbol{x}(t),\boldsymbol{u}(t)) \qquad \forall t \in [0,F] \tag{4.21a}$$

$$\Rightarrow \int_0^F H(\boldsymbol{\lambda}(t),\tilde{\boldsymbol{x}}(t),\tilde{\boldsymbol{u}}(t))\,dt \geq \int_0^F H(\boldsymbol{\lambda}(t),\boldsymbol{x}(t),\boldsymbol{u})\,dt. \tag{4.21b}$$

In consequence, $\tilde{\boldsymbol{u}}(t)$ is optimal with respect to any fixed $\boldsymbol{u} \in \mathcal{U}$ (Luenberger, 1979). According to this, the optimal control $\tilde{\boldsymbol{u}}(t)$ can be determined at each time point to satisfy $\Delta\tilde{J}\big|_{\tilde{u}(t)} \geq \Delta\tilde{J}\,|_u$ with respect to all admissible control functions (Crassidis and Junkins, 2011). This gives rise to Pontryagin's maximum (or minimum) principle which provides necessary optimality conditions and thus is the basis of many numerical algorithms (Luenberger, 1979). Nevertheless, optimizing $\Delta\tilde{J}$ remains a difficult task. Several numerical methods have been developed to provide a non-improvable solution by circumventing Pontryagin's maximum principle (Banga et al., 2005; Gong et al., 2008). However, the uniqueness of the solution with respect to all possible trajectories is not always guaranteed (Banga et al., 2005; Gong et al., 2008).

Box 1: Hidden Inputs



In general, for biological systems, the inputs $\boldsymbol{u}(t)$ can be divided into known stimuli and unknown but influential stimuli. With respect to control theory, inputs are equivalent to either fixed or variable controls.
Here the known biological system (green) includes one stimuli. However, there may exist another unknown pathway (red) which influences the system at certain nodes or other erroneous interactions within the system. This hidden inputs can be considered as unmodeled control variables. Inspired by this, the dynamic elastic-net was developed (Engelhardt et al., 2016).

4.2.2 Numerical Solutions

Since OCPs appear in many fields, a variety of solution strategies have been developed (Banga et al., 2005; Gong et al., 2008; Ross and Karpenko, 2012). These methods are commonly divided into dynamic programming, indirect and direct approaches (Banga et al., 2005). Besides other methods, such as stochastic and hybrid techniques based on vector parametrization, pseudospectral optimal control is a commonly used method (Banga et al., 2005). In this work, a pseudospectral optimal control algorithm was chosen to solve the DEN problem detailed in Chapter 6 (Engelhardt et al., 2016). Interestingly enough, both OCPs and pseudospectral theory are based on Sobolev spaces (Banga et al., 2005). Sobolev spaces consist of p-times integrable functions or equivalence classes with p-times integrable weak derivatives (Walter, 1994). Weak derivatives are generalized derivative functions which are not differentiable but integrable (Walter, 1994). The fundamental property leading to pseudospectral optimal control methods is given by the Stone-Weierstrass approximation, i.e., there exists a polynomial sequence $\boldsymbol{x}^i(t)$ in the Sobolev space for each state converging against $\boldsymbol{x}(t)$ for $i \to \infty$ (Ross and Karpenko, 2012). Thus, the original problem becomes discretized under appropriate conditions and can be solved when a sufficiently large set of points is used for calculation, for instance, on a mesh based on Gaussian rules (Gong et al., 2008). In consequence, the complex problem can be solved by solving a sequence of computationally feasible subproblems (Ross and Karpenko, 2012; Banga et al., 2005). In more detail, the exact solution is obtained by combining domain transformation, interpolation, differentiation and integration to solve a sequence of inexact solutions (Ross, 2005; Ross and Karpenko, 2012). In contrast to other methods, pseudospectral optimal control methods try to approximate the solution rather than manipulate equations (Gong et al., 2008).

4.2.3 Objective Functions

In addition to numerical issues, the solvability of OCPs strongly depends on the objective function $l(\boldsymbol{x}(t), \boldsymbol{u}(t))$. For biological problems, the fit of the model to observations evaluated through the least-square-error (LSE) is mainly of interest. The $L_2(\Omega)$ norm defined on the space of square-integrable functions in $\Omega \in \mathbb{R}$ is given by

$$\|g(t)\|_2 = \left(\int_{\Omega} |g(t)|^2 \, dt \right)^{\frac{1}{2}}, \tag{4.22}$$

where $g(t) : \Omega \to \mathbb{R}$ is a real-valued function with $\Omega \subset \mathbb{R}$. The model error of a dynamical system is typically defined as the difference between the model predictions $\boldsymbol{y}(t)$ and observations $\tilde{\boldsymbol{y}}(t)$ (Kremling, 2012). For discrete time points $t_l = t_1, ..., t_F$, with respect to the inaccuracy of measurements $\boldsymbol{\sigma}_{t_l}^2$, the LSE yields

$$\mathrm{LSE}(\boldsymbol{y}, \tilde{\boldsymbol{y}}) = \sum_{i=1}^{I} \sum_{l=1}^{F} \frac{(y_i(t_l) - \tilde{y}_i(t_l))^2}{\sigma_{i,t_l}^2}. \tag{4.23}$$

In complex cases, which are typical for biological processes, a unique solution or even a solution at all is not ensured (Jarre and Stoer, 2004). In view of OCPs which include additional control variables, a general criteria for uniqueness and solvability cannot be given (Luenberger, 1979). In addition, as discussed earlier, even strong inputs may result in comparatively weak changes in states and observations of the system. As shown in Figure 4.2, in some cases the order of magnitude of the difference between states and the related input can reach any size (Engelhardt et al., 2016). In consequence, it is not possible to draw direct conclusions from the states or observations of a system to the control variables (Luenberger, 1979; Engelhardt et al., 2016).

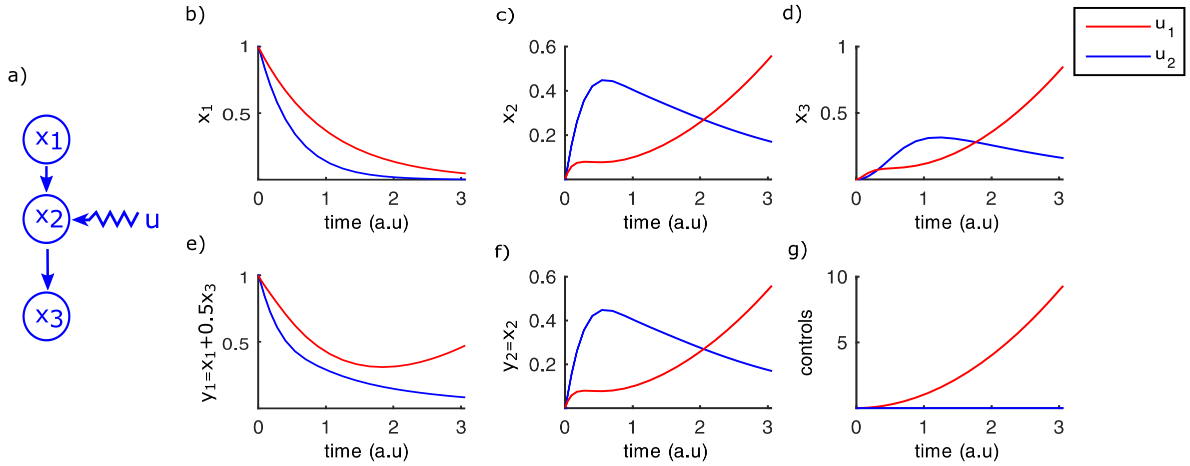


Figure 4.2: Illustration of the effect of even high changes in the control variable. (a) Network graph with observed states and one control variable. (b-d) State trajectories with respect to different admissible controls $\boldsymbol{u}_1(t)$ (blue) and $\boldsymbol{u}_2(t)$ (red). (e,f) Observed dynamics $\boldsymbol{y}_1(t)$ and $\boldsymbol{y}_2(t)$ with respect to $\boldsymbol{u}_1(t)$ (blue) and $\boldsymbol{u}_2(t)$ (red). (g) Time series of $\boldsymbol{u}_1(t)$ (blue) and $\boldsymbol{u}_2(t)$ (red). The change in the states and observations is much smaller than the variation of the control variable.

For some purposes, especially in cases where sparsity is required, the $L_1(\Omega)$ norm

$$\|g(t)\|_1 = \left(\int_{\Omega} |g(t)|\, dt\right) \tag{4.24}$$

is used, but this leads to a non-smooth objective functional (Vossen and Maurer, 2006). In many applications, also in other fields of research, e.g., robotics, the L_1 norm is a more realistic objective because it weighs all function values equally (Vossen and Maurer, 2006). It has been shown that the L_1 norm must be treated numerically different, in contrast to so-called smooth OCPs based on LSEs (Vossen and Maurer, 2006). In practice, non-smooth OCPs with added control variables are often solved via regularization and augmentation approaches (Vossen and Maurer, 2006). Regularization approaches solve the problem for general $L_p(\Omega)$ norms defined through

$$\|g(t)\|_p = \left(\int_{\Omega} |g(t)|^p\, dt\right)^{\frac{1}{p}}, \tag{4.25}$$

with $p \geq 1$. The OCP can then be solved by determining the limit

$$\lim_{\epsilon \to 0} \left[\left(\int_{\Omega} \sum_{i=1}^{I} |g_i(t)|^{(p+\epsilon)} dt\right)^{\frac{1}{p+\epsilon}}\right] \tag{4.26}$$

for $p = 1$ (Vossen and Maurer, 2006). Alternatively, an augmentation approach has been proposed where the control variable is split into its positive and negative fraction (Vossen and Maurer, 2006). This approach doubles the number of involved control variables. In both approaches, the control variables tend to behave like step functions taking only zero, the defined maximum and the defined minimum as values (Vossen and Maurer, 2006).

4.2.4 Properties

In practice, it is important to know whether the system can be clearly reconstructed based on the given data or if there are non-distinguishable parts. This is directly related to the question of observability (Luenberger, 1979). In contrast to non-linear systems, observability is well-investigated for linear dynamical systems given by

$$\frac{d\boldsymbol{x}(t)}{dt} = \boldsymbol{A}\boldsymbol{x}(t) + \boldsymbol{B}\boldsymbol{u}(t) \tag{4.27a}$$

$$\boldsymbol{y}(t) = \boldsymbol{C}\boldsymbol{x}(t) \tag{4.27b}$$

$$\boldsymbol{x}(0) = \boldsymbol{x}_0, \tag{4.27c}$$

where $\boldsymbol{x}_0 \in \mathbb{R}^N$ is the initial value (Luenberger, 1979; Adamy, 2014). The lack of observability is imputable to the linear mapping C caused by the fact that usually only parts of the system and combinations of species are experimentally accessible.

Definition 8: Observability

Given the linear system with $A \in \mathbb{R}^{N \times N}, B \in \mathbb{R}^{N \times P}, C \in \mathbb{R}^{I \times N}, \boldsymbol{x}(t) \in \mathbb{R}^{N}, \boldsymbol{y}(t) \in \mathbb{R}^{I}$ and $\boldsymbol{u}(t) \in \mathcal{U} \subset \mathbb{R}^{P}$, the system

$$\begin{aligned} \frac{d\boldsymbol{x}(t)}{dt} &= A\boldsymbol{x}(t) + B\boldsymbol{u}(t) \\ \boldsymbol{y}(t) &= C\boldsymbol{x}(t) \\ \boldsymbol{x}(0) &= \boldsymbol{x}_0, \end{aligned}$$

is called completely observable for $t \geq 0$ within the finite interval $[0, F]$ if the initial values $\boldsymbol{x}(0)$ can be clearly determined through $\boldsymbol{y}(t)$ and $\boldsymbol{u}(t)$ for all $\boldsymbol{u}(t)$ continuously differentiable.

Observability means that all states can be determined by observing the output structure (Luenberger, 1979). If the optimum is given by a clear trajectory related to one unique control for a given OCP, then the system is distinguishable in both states and control variables (Martinelli, 2015; Crassidis and Junkins, 2011; Luenberger, 1979). Thus observability is a necessary condition to achieve reliable models. In practice, one is often interested in particular mechanisms of the system under investigation and thus it is sufficient that at least these mechanisms are observable (Gao et al., 2014). In contrast to complex systems, for time-invariant linear systems with $\dot{\boldsymbol{x}} = A\boldsymbol{x}(t)$, there exists a simple criterion to test observability (Luenberger, 1979; Hermann and Krener, 1977). A time-invariant linear system is observable if and only if the rank of $M \in \mathbb{R}^{IN \times N}$ is N with (Luenberger, 1979)

$$M = \begin{pmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^{N-1} \end{pmatrix}. \tag{4.28}$$

The inverse of the observability matrix M in Equation (4.28) is used to transform a given system into the canonical form which is always observable and hence M has to be invertible, i.e., the column space of M has to be full (Hermann and Krener, 1977; Crassidis and Junkins, 2011). This means that all states of the system clearly contribute to the output and that no state is projected to the null space and thus vanishes in the output structure. Even though the concept of observability can be easily extended to non-linear systems, there exists no simple condition to prove observability for such systems (Martinelli, 2015).

A similar property of dynamical systems with additional input variables is called controllability (Hermann and Krener, 1977). For time-invariant linear systems, controllability is dual to observability and thus the theory and criteria are similar (Hermann and Krener, 1977; Adamy, 2014).

Definition 9: Controllability

Given the linear system with $A \in \mathbb{R}^{N \times N}, B \in \mathbb{R}^{N \times P}, C \in \mathbb{R}^{I \times N}, \boldsymbol{x}(t) \in \mathbb{R}^N, \boldsymbol{y}(t) \in \mathbb{R}^I$ and $\boldsymbol{u}(t) \in \mathcal{U} \subset \mathbb{R}^P$, the system

$$\begin{aligned} \frac{d\boldsymbol{x}(t)}{dt} &= A\boldsymbol{x}(t) + B\boldsymbol{u}(t) \\ \boldsymbol{y}(t) &= C\boldsymbol{x}(t) \\ \boldsymbol{x}(0) &= \boldsymbol{x}_0, \end{aligned}$$

is called controllable if there is a piecewise control variable $\boldsymbol{u}(t)$ for any given state $\boldsymbol{x}_{t_l}$ such that $\boldsymbol{x}(t_l) = \boldsymbol{x}_{t_l}$ and $\boldsymbol{x}(0) = \boldsymbol{x}_0$ for a finite time t with $0 \leq t \leq t_l$.

A linear system is controllable if and only if the rank of $M \in \mathbb{R}^{N \times PN}$ is N with (Luenberger, 1979)

$$M = \left(A, AB, A^2B, \ldots, A^{n-1}B\right). \tag{4.29}$$

The similarity between Equation (4.28) and Equation (4.29) clarifies the close connection between observability and controllability. For a CRN, it is usually not possible and mostly not necessary to control the entire network (Gao et al., 2014). Also, methods to efficiently control the dynamics of complex networks are rarely given (Gao et al., 2014). However, a recent study focusing on the identification of patterns within complex systems provided a possibility to identify the minimal set of necessary controls for a given objective (Gao et al., 2014). Besides criteria based on Lie derivations, which allow to investigate the derivative of the vector field generated by the state space corresponding to a constant control in direction to $\boldsymbol{f}(\cdot, \boldsymbol{u}(t))$, a raw approach to locally investigate non-linear systems is the first-order Taylor approximation around a local point or the use of other linearization methods (Hermann and Krener, 1977; Adamy, 2014; Crassidis and Junkins, 2011). In any case, the error caused by linearization has to be taken into account (Adamy, 2014; Crassidis and Junkins, 2011). Anyhow, this is useful to investigate a system with well-investigated linear methods (Adamy, 2014; Crassidis and Junkins, 2011).

So far, the states of the considered dynamical system are only restricted by their initial values $\boldsymbol{x}_0$ (Luenberger, 1979). In practice, also terminal constraints might be given, such as final data points or known final states. Without a constrained terminal state $\boldsymbol{x}(F)$, the final value of the state vector is completely arbitrary (Luenberger, 1979). OCPs therefore often lead to boundary value problems which are harder to solve than IVPs (Ross and Karpenko, 2012). Thus the inclusion of terminal constraints may lead to situations where neither a feasible control nor a feasible state trajectory can be found to satisfy both the initial and terminal constraints (Luenberger, 1979). On the other hand, there could be only one feasible solution which then leads to an unappreciation of the objective (Luenberger, 1979). It becomes clear that the adjoint variable $\lambda(F)$ at terminal constraints must not be specified in order to apply Pontryagin's maximum principle (Luenberger, 1979).
Thus it is only necessary to ensure that

$$\lambda(F) = \frac{\partial \phi(\boldsymbol{x}(F))}{\partial \boldsymbol{x}} \tag{4.30}$$

for all components of $\boldsymbol{x}(F)$ with $\Delta x_i(F) \neq 0$ (Luenberger, 1979). Consequently, for biological systems, it can be favorable to restrict the final state to a defined interval rather than setting fixed terminal constraints or removing final data points and thus terminal constraints from the system (Luenberger, 1979). This may help to satisfy the objective function in cases where it is not possible to find feasible control variables and state trajectories with feasible initial and terminal constraints (Luenberger, 1979). In contrast, in some experimental frameworks it might make sense to consider the last time point as an additional variable leading to free terminal time problems (Luenberger, 1979).

4.3 Bayesian Inference

In contrast to purely deterministic approaches, statistical methods allow for an intuitive handling of measurement noise, estimates of uncertain quantities and the analysis of random phenomena in general (Bishop, 2007; Gelman et al., 2013; Hoff, 2010). Such methods are often based on the inherent correlations in data (Bishop, 2007; Gelman et al., 2013; Hoff, 2010). It is clear that the necessity of introducing probabilities is caused by a lack of knowledge and that there are different ways to look at probabilities (Tschirk, 2014). The classical frequentism point of view, as the standard interpretation of probability, considers random variables as flexible values which change by multiple conduction of assigned random experiments (Tschirk, 2014). Consequently, probability is the relative frequency of an event (Tschirk, 2014). From the Bayesian point of view, variables with unknown values are random, too, but the probability is considered as the reasonable expectation of an event (Tschirk, 2014; Hoff, 2010).

Frequentism only allows knowledge described by relative frequencies whereas in Bayesian statistics there is no general restriction to a certain type of knowledge (Tschirk, 2014; Hoff, 2010). This is why frequentism is often considered a special case of Bayesianism (Tschirk, 2014). Besides theoretical aspects, this has practical consequences for the use of probabilities (Tschirk, 2014). From the classical frequentist point of view, it is not clear if there is any relation between the estimated parameter and the true parameter since the estimates are just based on random samples (Tschirk, 2014). Anyhow, there are some measurement variables, e.g., the efficiency, to characterize the relation between estimates and true values (Tschirk, 2014). In contrast, Bayesian statistics estimates parameter distributions by incorporating prior knowledge (Tschirk, 2014; Hoff, 2010). If there is only little prior knowledge, the results will be similar or equivalent to those obtained through frequentist approaches, with the deviation increasing with the amount of prior knowledge (Tschirk, 2014).

Since Bayesian analysis uses prior knowledge to determine the probability of certain events, its advantage over classical statistical methods increases with the amount of prior knowledge frequently given in biological problems (Tschirk, 2014; Hoff, 2010). In particular, Bayesian learning makes it possible to perform optimal updates of beliefs given new information (Hoff, 2010). Bayes' law, as detailed in Box 2, is the fundamental concept of Bayesian statistics and allows to deduce a posterior distribution based on the prior distribution and a sampling model (Hoff, 2010).

Box 2: Bayes' Law

Let θ be a parameter in the parameter space Θ and y a sample in the sample space Ω. Then the prior distribution for θ is given by $p(\theta)$ and the posterior distribution of the parameter given the samples can be obtained through

$$p(\theta | y) = \frac{p(y | \theta)\, p(\theta)}{\int\limits_{\Theta} p(y | \theta')\, p(\theta')\, d\theta'},$$

where $p(y | \theta)$ denotes the probability of the sample given the parameter.

Bayes' Law describes the distribution of the parameter based on a given sample through rational beliefs obtained from biological knowledge (Hoff, 2010). These beliefs are represented by prior knowledge about the parameter distribution and the sample distribution given this parameter. The nominator $\int_{\Theta} p(y | \theta')\, p(\theta')\, d\theta'$ is equivalent to the usually unknown unconditioned sample distribution $p(y)$ (Gelman et al., 2013). Based on Bayes' Law, the basic idea of Bayesian learning is to update the beliefs about parameter and sample distributions, where $p(y | \theta)$ is the belief of the sample if the parameter were known and $p(\theta | y)$ is the probability of the parameter based on the sample observations (Gelman et al., 2013). Bayes' Law describes how the beliefs change after gaining new information (Hoff, 2010).

From the Bayesian perspective, observations, e.g., protein concentrations, are nothing else than realizations of a random variable with a certain known or unknown measurement error (Tschirk, 2014; Gelman et al., 2013). In case of unknown measurement errors, priors based on assumptions or knowledge can be introduced (Gelman et al., 2013). If there is a lack of prior knowledge, it is also possible to add non-informative priors (Gelman et al., 2013; Hoff, 2010). In general, the more information is encoded in the data, the lower the influence of the prior on the posterior (Gelman et al., 2013; Hoff, 2010). In turn, if the amount of information obtained through the data or model is low, the prior dominates the posterior and, in consequence, the results become uncertain (Gelman et al., 2013; Hoff, 2010).

This work adopts common terms of probability theory, which are given below (Hoff, 2010; Gelman et al., 2013).

Definition 10: Random Variable

Given a random experiment with sample space Ω, then a random variable X is a mapping which assigns each $\omega \in \Omega$ a real value $X(\omega) = x$ according to

$$\begin{aligned} X : \Omega &\rightarrow \mathbb{R} \\ \omega &\mapsto X(\omega) = x, \end{aligned}$$

where x is called realization of the random variable X.

The distribution function $F(x) = P(X \leq x)$ assigns a probability $P(X \leq x)$ to each realization.

For continuous random variables and each interval $[a, b]$, there exists a density $f(x) \geq 0$ such that

$$P(a \leq X \leq b) = \int\limits_a^b f(x)\,dx,$$

with $\int\limits_{-\infty}^{\infty} f(x)\,dx = 1$ and $F(x) = \int\limits_{-\infty}^{x} f(t)\,dt$.

The variance is defined as

$$\sigma_f^2 = Var_f(X) = \int\limits_{-\infty}^{\infty} \left(x - \mu_f\right)^2 f(x)\,dx,$$

with an expected value $\mu_f = E_f(X) = \int\limits_{-\infty}^{\infty} x f(x)\,dx$. For simplicity of notation, in this work σ_f^2 is denoted as σ^2 and μ_f as μ.

For dynamic protein-protein interaction networks, the time series of the protein concentrations are often accessible (Azeloglu and Iyengar, 2015; Sauer et al., 2007). This means that several protein concentrations are measured at several well-defined time points. Consequentially, this work deals with multivariate random variables and therefore, for reasons of clarity, the following simplifications are used in this thesis. Given a set of observations $i = 1, ..., I$ at F discrete time points $t_l = t_1, ...t_F$, which is equivalent to $l = 1, ..., F$, a single observation i at time t is denoted as $y_{i,l}$. A set of observations is given by

$$\boldsymbol{y}_{1:F} = \left(\boldsymbol{y}_{1,1:F}, ..., \boldsymbol{y}_{I,1:F}\right)^T = \begin{pmatrix} y_{1,1} & \cdots & y_{1,F} \\ \vdots & \ddots & \vdots \\ y_{I,1} & \cdots & y_{I,F} \end{pmatrix}. \tag{4.31}$$

In this chapter a general multivariate random variable is denoted as $\boldsymbol{Y}$.

Box 3: Strict definition of multivariate random variables

It is important to note that Definition 10, as a common simplified definition, is not completely consistent from a mathematical point of view. For the sake of completeness, a stricter definition for multivariate random variables or random vectors is given below.

Let $(\Omega, \mathcal{G})$ be a measure where $\mathcal{G}$ is a σ-algebra over the sample space Ω and $(E, \mathcal{E})$ a measure where $\mathcal{E}$ is a σ-algebra over E. Then a mapping $X : \Omega \to E$ with

$$X^{-1}(B) = \{\omega \in \Omega : X(\omega) \in B\} \in \mathcal{G} \quad \forall B \in \mathcal{E}$$

is called random variable on $(\Omega, \mathcal{G})$ with values in $(E, \mathcal{E})$ and hence $X : (\Omega, \mathcal{G}) \to (E, \mathcal{E})$ is a measurable mapping. Then, for each probability P in $(\Omega, \mathcal{G})$, there exists a probability measure Q in $(E, \mathcal{E})$ with

$$Q(B) = P\left(X^{-1}(B)\right) \quad B \in \mathcal{E},$$

where Q is called distribution of X with respect to P.

Let Q be a probability measure in $\left(\mathbb{R}^N, \mathcal{B}\left(\mathbb{R}^N\right)\right)$ with a Borel set $\mathcal{B}\left(\mathbb{R}^N\right)$. Then Q is called probability distribution with

$$\left\{\begin{array}{l} Q\left(\bigtimes_{n=1}^{N} (-\infty, b_N]\right) \\ b = (b_1, ..., b_N) \in \mathbb{R}^N \text{abitrary} \end{array}\right. = \int_{-\infty}^{b_1} ... \int_{-\infty}^{b_N} f(X_1(x_1), ..., X_N(x_N))\, dx_1 ... dx_N,$$

and

$$\int_{-\infty}^{\infty} ... \int_{-\infty}^{\infty} f(X_1(x_1), ..., X_N(x_N))\, dx_1 ... dx_N = 1.$$

The function $f(X_1(x_1), ..., X_N(x_N))$ is called density with respect to Q.

Please note that, from a mathematical point of view, the integrals written in Definition 10 are vaguely defined. Because random variables are mappings, the integration variables are given by the realizations of those random variables and not by the random variables themselves. However, due to the fact that the domain and codomain are real-valued and for reasons of simplification the random variable is suppressed and thus

$$\int_{-\infty}^{b_1} ... \int_{-\infty}^{b_N} f(X_1(x_1), ..., X_N(x_N))\, dx_1 ... dx_N \Leftrightarrow \int_{-\infty}^{b_1} ... \int_{-\infty}^{b_N} f(x_1, ..., x_N)\, dx_1 ... dx_N.$$

Although $p(\boldsymbol{\theta}|\boldsymbol{Y})$ is the probability density of $\boldsymbol{\theta}$ conditioned on $\boldsymbol{Y}$, it is often equated with the probability of $\boldsymbol{\theta}$ given $\boldsymbol{Y}$ since the probability function can in general be derived from its distribution (Gelman et al., 2013; Hoff, 2010). In addition, discrete random variables are special cases of continuous random variables (Gelman et al., 2013; Hoff, 2010).

4.3.1 Markov Chain Monte Carlo Methods

The probability of the estimated parameter set given the sample is an important measure for its reliability (Brooks, 1998; Gelman et al., 2013; Hoff, 2010). As a direct consequence of Bayes' law (Box 2), the conditional posterior density of the parameter set can be approximated as

$$p\left(\boldsymbol{\theta}\,|\boldsymbol{Y}\right) \propto p\left(\boldsymbol{Y}\,|\boldsymbol{\theta}\right) p\left(\boldsymbol{\theta}\right), \tag{4.32}$$

where the constant proportion is given by $\int_{\Theta} \left(\boldsymbol{Y}\,|\boldsymbol{\theta}'\right) p\left(\boldsymbol{\theta}'\right) d\boldsymbol{\theta}'$ (Hoff, 2010). In most cases this integral is hard to solve and numerical integration or analytic approximation techniques must be used (Gelman et al., 2013; Hoff, 2010). Thus, neither the full posterior density nor the marginal posterior density are directly accessible (Hoff, 2010). MCMC methods allow for sampling directly from the posterior while avoiding analytical integration (Hoff, 2010). This way, sampling from the posterior based on prior knowledge about $\boldsymbol{\theta}$ and the distribution of the data $\boldsymbol{Y}$ given $\boldsymbol{\theta}$ is possible (Hoff, 2010). Although the parameter distribution for the given data is not available, information about $\boldsymbol{\theta}$ and the data given $\boldsymbol{\theta}$ can often be derived (Gelman et al., 2013; Hoff, 2010).

The fundamental concept of MCMC methods is based on the fact that the posterior distribution is known up to a certain multiplicative constant (Gelman et al., 2013; Hoff, 2010). Here, it is often possible to generate a chain of samples which converge against random samples from the posterior distribution (Brooks, 1998). More precisely, given a Markov chain $\left\{\boldsymbol{\theta}^{(s)} : s = 1, ..., S\right\}$ of realizations $\boldsymbol{\theta}^{(s)}$, then for any real-valued function $g\left(\boldsymbol{\theta}\right)$ the state distribution of the Markov chain converges against the posterior distribution with (Brooks, 1998)

$$P\left(\lim_{S\to\infty} \frac{1}{S} \sum_{s=1}^{S} g\left(\boldsymbol{\theta}^{(s)}\right) = \int_{-\infty}^{\infty} g\left(\boldsymbol{\theta}'\right) f\left(\boldsymbol{\theta}'\,|\boldsymbol{Y}\right) d\boldsymbol{\theta}'\right) = 1. \tag{4.33}$$

Markov chains are sequences of random variables $\boldsymbol{\theta}^{(s)}$ with

$$P\left(\boldsymbol{\theta}^{(s+1)}\,\middle|\boldsymbol{\theta}^{(s)}, ..., \boldsymbol{\theta}^{(1)}\right) = P\left(\boldsymbol{\theta}^{(s+1)}\,\middle|\boldsymbol{\theta}^{(s)}\right) \forall s. \tag{4.34}$$

The most important properties of Markov chains for MCMC transitions are irreducibility, recurrency and aperiodicity (Brooks, 1998). Irreducibility means that there exists a certain probability that a state can be transitioned into another state within one or more steps (Brooks, 1998; Gelman et al., 2013; Hoff, 2010). Recurrency guarantees convergence in a finite time, whereas aperiodicity means that a return to a certain state can occur at irregular times (Brooks, 1998; Gelman et al., 2013; Hoff, 2010).

It can be shown that an irreducible and aperiodic Markov chain converges to a unique stationary distribution (Brooks, 1998). This means that once the process behaves like a certain distribution, it will stay there and, in consequence, all samples are drawn from this unique stationary distribution (Brooks, 1998). In fact, a transition kernel $\mathcal{K}$ can be generated for Markov chains which describes the transition between two states (Brooks, 1998). Depending on the situation and the given amount of data and knowledge, different kernels have been developed in the past (Brooks, 1998). This thesis focuses on Metropolis-Hastings (MH) and Gibbs sampling. However, there are also other MCMC methods, such as approximate Bayesian computation (Toni et al., 2009).

Metropolis-Hastings Sampling

As discussed above, in situations where information about $p(\boldsymbol{Y}|\boldsymbol{\theta})$ and $p(\boldsymbol{\theta})$ is available but the posterior $p(\boldsymbol{\theta}|\boldsymbol{Y})$ is impossible to calculate, it is not possible to directly draw samples from the posterior and thus to characterize this distribution (Gelman et al., 2013). Consequently, by drawing samples from the posterior, it would be possible to achieve the expected value using Equation (4.33) (Gelman et al., 2013). A very general approach is to collect a large number of samples $\Theta = \left\{\boldsymbol{\theta}^{(1)}, ..., \boldsymbol{\theta}^{(S)}\right\}$ whose distribution approximates the prior and satisfies Equation (4.33) by drawing values from a sensitively chosen proposal distribution $\pi\left(\boldsymbol{\theta}^{(*)}\left|\boldsymbol{\theta}^{(s)}\right.\right)$ (Brooks, 1998). At first glance, a new sample $\boldsymbol{\theta}^{(*)}$ of the collection already including s samples should ideally satisfy $p\left(\boldsymbol{\theta}^{(*)}|\boldsymbol{Y}\right) \geq p\left(\boldsymbol{\theta}^{(s)}|\boldsymbol{Y}\right)$ (Brooks, 1998). This means, because $\boldsymbol{\theta}^{(s)}$ is already in the collection, $\boldsymbol{\theta}^{(*)}$ has to be included as well (Hoff, 2010). This concept can be quantified through a ratio, which yields

$$\tilde{r}\left(\boldsymbol{\theta}^{(s)}, \boldsymbol{\theta}^{(*)}\right) = \frac{p\left(\boldsymbol{\theta}^{(*)}|\boldsymbol{Y}\right)}{p\left(\boldsymbol{\theta}^{(s)}|\boldsymbol{Y}\right)}. \tag{4.35}$$

Now, two situations are possible: either $\tilde{r}\left(\boldsymbol{\theta}^{(s)}, \boldsymbol{\theta}^{(*)}\right) < 1$ or $\tilde{r}\left(\boldsymbol{\theta}^{(s)}, \boldsymbol{\theta}^{(*)}\right) \geq 1$ (Hoff, 2010). In the latter case, $p\left(\boldsymbol{\theta}^{(*)}|\boldsymbol{Y}\right) \geq p\left(\boldsymbol{\theta}^{(s)}|\boldsymbol{Y}\right)$ and thus the new sample has to be included in the collection (Hoff, 2010). However, if $\tilde{r}\left(\boldsymbol{\theta}^{(s)}, \boldsymbol{\theta}^{(*)}\right) < 1$, there should be only a fraction of samples $\boldsymbol{\theta}^{(*)}$ with $p\left(\boldsymbol{\theta}^{(*)}|\boldsymbol{Y}\right) < p\left(\boldsymbol{\theta}^{(\tilde{s})}|\boldsymbol{Y}\right)$ for each member $\boldsymbol{\theta}^{(\tilde{s})} \in \Theta$ of the collection (Hoff, 2010). Let

$$r\left(\boldsymbol{\theta}^{(s)}, \boldsymbol{\theta}^{(*)}\right) = \min\left[1, \frac{p\left(\boldsymbol{\theta}^{(*)}|\boldsymbol{Y}\right) p\left(\boldsymbol{\theta}^{(*)}\right)}{p\left(\boldsymbol{\theta}^{(s)}|\boldsymbol{Y}\right) p\left(\boldsymbol{\theta}^{(s)}\right)} \times \frac{p\left(\boldsymbol{\theta}^{(s)}\left|\boldsymbol{\theta}^{(*)}\right.\right)}{p\left(\boldsymbol{\theta}^{(*)}\left|\boldsymbol{\theta}^{(s)}\right.\right)}\right] \tag{4.36}$$

be the acceptance ratio, $p\left(\boldsymbol{\theta}^{(s)}|\boldsymbol{Y}\right)$ the posterior distribution with respect to $\boldsymbol{\theta}^{(s)}$ and $p\left(\boldsymbol{\theta}^{(*)}\left|\boldsymbol{\theta}^{(s)}\right.\right)$ the density used for the generation of the candidate $\boldsymbol{\theta}^{(*)}$ (Gelman et al., 2013).

Then the corresponding transition kernel $\mathcal{K}\left(\boldsymbol{\theta}^{(s)}, \Theta\right)$ of the MH algorithm, ensuring that the next sample is an element of Θ given $\boldsymbol{\theta}^{(s)} \in \tilde{\Theta}$, yields:

$$\begin{aligned}\mathcal{K}\left(\boldsymbol{\theta}^{(s)}, \Theta\right) = &\int_{\Theta} p\left(\boldsymbol{\theta}' \left|\boldsymbol{\theta}^{(s)}\right.\right) r\left(\boldsymbol{\theta}^{(s)}, \boldsymbol{\theta}'\right) d\boldsymbol{\theta}' \\ &+ \left(1 - \int_{\tilde{\Theta}} p\left(\boldsymbol{\theta}' \left|\boldsymbol{\theta}^{(s)}\right.\right) r\left(\boldsymbol{\theta}^{(s)}, \boldsymbol{\theta}'\right) d\boldsymbol{\theta}'\right) \mathbf{1}_{\Theta}\left(\boldsymbol{\theta}^{(s)}\right),\end{aligned} \tag{4.37}$$

with the indicator function or point mass $\mathbf{1}_{\Theta}\left(\boldsymbol{\theta}^{(s)}\right)$, which is one if $\boldsymbol{\theta}^{(s)}$ is already element of Θ and zero otherwise (Brooks, 1998). The last term in Equation (4.36) corrects for $\boldsymbol{\theta}^{(*)}$ because if $\boldsymbol{\theta}^{(*)}$ has a higher probability to be chosen than $\boldsymbol{\theta}^{(s)}$, it has to be considered in the acceptance ratio, otherwise $\boldsymbol{\theta}^{(*)}$ would be over-represented in the collection (Brooks, 1998).

As seen in Equation (4.37), the probability $\mathcal{K}\left(\boldsymbol{\theta}^{(s)}, \Theta\right)$ of the next state to be in Θ is composed of the marginal probability $p\left(\boldsymbol{\theta}' \left|\boldsymbol{\theta}^{(s)}\right.\right) r\left(\boldsymbol{\theta}^{(s)}, \boldsymbol{\theta}'\right)\Big|_{\boldsymbol{\theta}' \in \Theta}$ of selecting a state which is accepted and the possible non-zero probability of remaining in $\boldsymbol{\theta}^{(s)}$ given by the second term of Equation (4.37) (Brooks, 1998). The latter is one minus the probability that the new sample is accepted given $\boldsymbol{\theta}^{(s)}$ is already element of Θ (Brooks, 1998). Consequentially, the MH algorithm to update a vector-valued parameter $\boldsymbol{\theta}$ can be stated (Brooks, 1998; Gelman et al., 2013; Hoff, 2010).

Box 4: Metropolis-Hastings Algorithm

Given a posterior density $p\left(\boldsymbol{\theta} \left|\boldsymbol{Y}\right.\right)$ with a vector-valued parameter $\boldsymbol{\theta} = \left(\theta_1, ..., \theta_P\right)^T$ and samples $\boldsymbol{Y}$, a prior density $p\left(\boldsymbol{\theta}\right)$, a uniformly distributed random variable $u \sim U\left(0, 1\right)$ and a proposal distribution $\pi\left(\boldsymbol{\theta}^{(*)} \left|\boldsymbol{\theta}^{(s)}\right.\right)$ with a corresponding density $p\left(\boldsymbol{\theta}^{(*)} \left|\boldsymbol{\theta}^{(s)}\right.\right)$, the Metropolis-Hastings updates in step $s + 1$ are processed as follows:

a) Draw $\boldsymbol{\theta}^{(*)} \sim \pi\left(\boldsymbol{\theta}^{(*)} \left|\boldsymbol{\theta}^{(s)}\right.\right)$

b) Calculate $r\left(\boldsymbol{\theta}^{(s)}, \boldsymbol{\theta}^{(*)}\right)$

c) Draw $u \sim U\left(0, 1\right)$. If $u < r$ set $\boldsymbol{\theta}^{(s+1)} = \boldsymbol{\theta}^{(*)}$ and $\boldsymbol{\theta}^{(s+1)} = \boldsymbol{\theta}^{(s)}$ otherwise,

with

$$r\left(\boldsymbol{\theta}^{(s)}, \boldsymbol{\theta}^{(*)}\right) = \min\left[1, \frac{p\left(\boldsymbol{\theta}^{(*)} \left|\boldsymbol{Y}\right.\right) p\left(\boldsymbol{\theta}^{(*)}\right)}{p\left(\boldsymbol{\theta}^{(s)} \left|\boldsymbol{Y}\right.\right) p\left(\boldsymbol{\theta}^{(s)}\right)} \times \frac{p\left(\boldsymbol{\theta}^{(s)} \left|\boldsymbol{\theta}^{(*)}\right.\right)}{p\left(\boldsymbol{\theta}^{(*)} \left|\boldsymbol{\theta}^{(s)}\right.\right)}\right].$$

The only restriction on the proposal distribution in Box 4 is that it does not depend on any other elements θ_p of the parameter vector $\boldsymbol{\theta}$ previous to the most current values in order to retain the Markov process (Gelman et al., 2013; Hoff, 2010). A special case of Box 4 is the Metropolis algorithm, which is in fact the origin of the MH algorithm (Brooks, 1998). In contrast to the MH algorithm, the proposal distribution of the original Metropolis algorithm is symmetric (Gelman et al., 2013; Hoff, 2010). This means that $p\left(\boldsymbol{\theta}^{(*)}\left|\boldsymbol{\theta}^{(s)}\right.\right) = p\left(\boldsymbol{\theta}^{(s)}\left|\boldsymbol{\theta}^{(*)}\right.\right)$ and hence the correction term in Equation (4.36) vanishes (Gelman et al., 2013). A commonly used and more special case is the random walk Metropolis algorithm where the proposal density is generally given as a symmetric arbitrary density $f\left(\boldsymbol{\theta}^{(*)} - \boldsymbol{\theta}^{(s)}\right)$, i.e., a uniform or Gaussian distribution (Gelman et al., 2013). Another commonly used option is the independent proposal distribution which is independent of the current state $\boldsymbol{\theta}^{(s)}$ (Gelman et al., 2013; Hoff, 2010). This illustrates the universality of the MH algorithm and highlights the importance of the proposal distribution.

In fact, the behavior of the Markov chain heavily depends on the proposal distribution and thus an inadequate choice may result in a poor performance of the estimates (Andrieu and Thoms, 2008). Major drawbacks of such methods are the correlation of nearby samples and the fact that the initial samples have a different distribution than the target distribution (Gelman et al., 2013; Hoff, 2010). The latter can be circumvented by introducing a burn-in phase, i.e., by simply removing the first samples up to a certain number from the collection (Gelman et al., 2013). A convenient and conservative amount of omitted samples is half of the samples (Gelman et al., 2013). In contrast, the correlation of samples within the collection is much harder to avoid since the correlation corresponds to the distance between two successive samples (Gelman et al., 2013). An increase of the distance reduces the correlation but also increases the frequency of sample rejection, which in turn reduces the efficiency (Gelman et al., 2013). On the other hand, serial correlation is not necessarily an issue (Gelman et al., 2013). Hence, once convergence is reached, the samples are identically distributed and drawn from the posterior distribution anyway (Gelman et al., 2013). However, this still reduces the efficiency of the algorithm (Gelman et al., 2013). Here, thinning is a commonly used technique to reduce autocorrelation by only keeping every j-th sample (Brooks et al., 2011; Gelman et al., 2013). For more difficult tasks it is useful to monitor the convergence of samples and to estimate an effective sample size especially to include a good stopping criteria (Gelman et al., 2013).

Since the performance of MCMC algorithms depends on the proposal distribution, a lot of effort has been devoted towards the optimization of the step size of the proposal (Andrieu and Thoms, 2008; Gelman et al., 2013). Clearly, the variance of the proposal directly influences the estimations (Andrieu and Thoms, 2008). In addition, optimal estimations are supposed to have a small variance (Andrieu and Thoms, 2008; Gelman et al., 2013). But proposal variances that are too small or too large lead to inaccurate estimations (Andrieu and Thoms, 2008). Thus, it is often important to design optimal proposal variances to achieve efficient and reliable procedures (Andrieu and Thoms, 2008). Controlled MCMC methods address this issue with sequential updates of the proposal variance (Andrieu and Thoms, 2008). A major drawback of such algorithms is that they tend to loose the ability to pass from one state to another within finite time which is a minimum requirement for consistent estimators (Andrieu and Thoms, 2008). Accordingly, it is necessary to ensure that the algorithm does not produce poor proposal distributions and becomes unstable (Andrieu and Thoms, 2008). An automatic choice of the proposal variance and thus the step size between two successive samples is therefore highly desirable (Andrieu and Thoms, 2008). This leads to adaptive MCMC and MH procedures where the proposal distribution is automatically adapted and chosen (Andrieu and Thoms, 2008). For vector-valued parameters, these algorithms update and scale the parameters and the related proposal distributions either component-wise or globally (Andrieu and Thoms, 2008). However, in practice, standard algorithms, such as random walk Metropolis or normal MH algorithms, are often sufficient to explore the parameter space (Andrieu and Thoms, 2008; Gelman et al., 2013; Hoff, 2010).

Gibbs Sampling

Gibbs sampling can be viewed as a special case of MH sampling which makes the estimation more feasible and efficient especially for multidimensional problems (Casella, 2001; Brooks, 1998). In contrast to MH sampling, for Gibbs transitions, e.g., with two unknown parameters, namely $\theta_1 \in \boldsymbol{\theta}$ and $\theta_2 \in \boldsymbol{\theta}$, the proposal distributions $p\left(\theta_2^{(*)} \middle| \theta_1^{(s)}, \theta_2^{(s)}\right)$ and $p\left(\theta_1^{(*)} \middle| \theta_1^{(s)}, \theta_2^{(s)}\right)$, respectively, are set as the so-called full conditional distributions of θ_1 given θ_2 and vice versa (Gelman et al., 2013; Hoff, 2010). This means that

$$p\left(\theta_1^{(*)} \middle| \theta_1^{(s)}, \theta_2^{(s)}\right) = p\left(\theta_1^{(*)} \middle| \theta_2^{(s)}, \boldsymbol{Y}\right) \tag{4.38}$$

and

$$p\left(\theta_2^{(*)} \middle| \theta_1^{(s)}, \theta_2^{(s)}\right) = p\left(\theta_2^{(*)} \middle| \theta_1^{(s)}, \boldsymbol{Y}\right). \tag{4.39}$$

Gibbs samplers are often based on conjugate priors which ideally allow to deduce a closed form of the posterior distribution (Casella, 2001; Hoff, 2010). As a consequence, it is possible to draw samples directly from this closed form (Casella, 2001; Gelman et al., 2013). A prior distribution is conjugate to the posterior distribution and hence in the same family $\mathcal{P}$ if

$$p(\boldsymbol{\theta}) \in \mathcal{P} \Rightarrow p(\boldsymbol{\theta} | \boldsymbol{Y}) \in \mathcal{P} \tag{4.40}$$

holds true (Hoff, 2010). For Gibbs sampling-based algorithms, conditional conjugate prior distributions are often sufficient (Hoff, 2010). Similar to Equation (4.40), the conditional conjugate prior $p\left(\theta_p \middle| \boldsymbol{\theta}_{-p}\right)$ with $\boldsymbol{\theta}_{-p} = \left(\theta_1, ..., \theta_{p-1}, \theta_{p+1}, ..., \theta_P\right)$ is assumed to be conjugate to the posterior $p(\boldsymbol{\theta} | \boldsymbol{Y})$ (Gelman et al., 2013). This allows to draw iteratively from a set of unknown parameters (Gelman et al., 2013).

MH algorithms are also able to deal with models where more than one parameter has to be updated iteratively (Gelman et al., 2013; Hoff, 2010; Brooks, 1998). But for Gibbs transitions the acceptance ratio always becomes one (Gelman et al., 2013; Hoff, 2010). Accordingly, all produced samples are accepted, which leads to an optimal efficiency of the algorithm (Brooks, 1998; Gelman et al., 2013; Hoff, 2010). The fact that the acceptance ratio for updating θ_p equals one for each update follows from

$$\begin{aligned} p\left(\theta_p^{(*)}, \boldsymbol{\theta}_{-p}^{(s)} \middle| \boldsymbol{Y}\right) p\left(\theta_p^{(s)} \middle| \theta_p^{(*)}, \boldsymbol{\theta}_{-p}^{(s)}\right) &= p\left(\theta_p^{(*)}, \boldsymbol{\theta}_{-p}^{(s)} \middle| \boldsymbol{Y}\right) p\left(\theta_p^{(s)} \middle| \boldsymbol{\theta}_{-p}^{(s)}, \boldsymbol{Y}\right) && (4.41a) \\ &= p\left(\theta_p^{(*)} \middle| \boldsymbol{\theta}_{-p}^{(s)}, \boldsymbol{Y}\right) p\left(\boldsymbol{\theta}_{-p}^{(s)} | \boldsymbol{Y}\right) p\left(\theta_p^{(s)} \middle| \boldsymbol{\theta}_{-p}^{(s)}, \boldsymbol{Y}\right), && (4.41b) \end{aligned}$$

with the identity $p(x, y | z) = p(x | y, z)\, p(y | z)$ used in Equation (4.41b) (Hoff, 2010). For Gibbs sampling algorithms, it is again possible that the samples are located in the same region for many iterations which means that, once $\boldsymbol{\theta}^{(s)}$ is in a local pattern, the probability of $\boldsymbol{\theta}^{(s+1)}$ being again in the same local pattern is high (Gelman et al., 2013; Hoff, 2010). This is supported by the fact that Gibbs samplers tend to sample $\boldsymbol{\theta}^{(s+1)}$ nearby $\boldsymbol{\theta}^{(s)}$ and thus samples may remain in such regions for a long time, which increases the number of samples required for a sufficient posterior approximation (Hoff, 2010). Additionally, in many situations, it is not possible to derive a closed form of the prior distribution due to a lack of conjugate priors (Casella, 2001; Gelman et al., 2013; Hoff, 2010). In general, distributions belonging to the exponential family have natural conjugate priors, whereas conjugate distributions are rarely given for other families (Gelman et al., 2013).

Further Aspects

The superiority of Gibbs sampling becomes more clear for component-wise updates of parameter vectors and hierarchical models (Gelman et al., 2013; Hoff, 2010). In contrast to MH-based algorithms, Gibbs transitions guarantee acceptance in each draw and hence the components can be drawn successively (Gelman et al., 2013; Hoff, 2010). It is not necessary that each component is updated in each iteration; it is only important to update them periodically (Gelman et al., 2013; Hoff, 2010). On the contrary, MH transitions do not guarantee acceptance for each draw and thus it is not ensured that the whole parameter space can be covered within a reasonable time (Gelman et al., 2013; Hoff, 2010).

If the conditional posterior depends on additional hyperparameters, often the subset of these additional unknowns are estimated and their estimators are used to calculate the posterior based on Gibbs transitions (Casella, 2001). But this is not an analysis in a Bayesian sense since the posterior distribution still depends on all the estimated hyperparameters (Casella, 2001). In a full Bayesian framework, this leads to hierarchical or multi-level models where the hyperparameters are treated as additional random variables, as described in Definition 11 (Gelman et al., 2013; Hoff, 2010). In contrast to parameters which are part of the underlying model, hyperparameters are additional unknowns influencing the prior distribution. The distribution of a hyperparameter is called hyperdistribution.

Definition 11: Hierarchical Model

Let $p(\boldsymbol{\theta}|Y)$ be the posterior distribution with a vector-valued parameter $\boldsymbol{\theta}$ with $\boldsymbol{\theta} \in \Theta$, a set of vector-valued hyperparameters $\boldsymbol{\phi} = \{\phi_1, ..., \phi_P, \}$ with $\boldsymbol{\phi} \in \Phi$ and a prior distribution $p(\boldsymbol{\theta}, \boldsymbol{\phi})$. If the prior distribution is factorable in $p(\boldsymbol{\theta}|\phi_1), p(\phi_1|\phi_2), ..., p(\phi_{P-1}|\phi_P)$ and its marginal distribution $p\left(\phi_p\right)$ with

$$p(\boldsymbol{\theta}) = \int\limits_{\Phi} p(\boldsymbol{\theta}|\phi_1)\, p(\phi_1|\phi_2) \cdots (\phi_{P-1}|\phi_P)\, p(\phi_P)\, d\boldsymbol{\phi},$$

then the model

$$\begin{aligned}
\boldsymbol{Y}|\boldsymbol{\theta}, \boldsymbol{\phi} &\sim P(\boldsymbol{Y}|\boldsymbol{\theta}, \boldsymbol{\phi}) \\
\boldsymbol{\theta}|\phi_1, ..., \phi_P &\sim P(\boldsymbol{\theta}|\phi_1, ..., \theta_P) \\
\phi_1|\phi_2, ..., \phi_P &\sim P(\phi_1|\phi_2, ..., \phi_P) \\
\phi_2|\phi_3, ..., \phi_P &\sim P(\phi_2|\phi_3, ..., \phi_P) \\
&\vdots \\
\phi_P &\sim P(\phi_P)
\end{aligned}$$

is called a hierarchical model of order $P + 2$ with P hyperparameters.

As a subclass of graphical models, hierarchical models can be represented in an illustrative way (Gelman et al., 2013; Bishop, 2007). Graphical models are a simple way to illustrate even complex probabilistic models where each node represents a single random variable or a set of random variables and each edge represents a probabilistic relationship (Bishop, 2007). The general idea is to factorize the joint distribution into a product of factors which depend only on a subset of random variables (Bishop, 2007). In consequence, samples can be drawn iteratively (Bishop, 2007). Hierarchical models are in general more robust than non-hierarchical Bayesian models and they reflect the hierarchical structure of multi-level observations (Gelman et al., 2013; Hoff, 2010). Often only a part of the hyperparameters have conjugate prior distributions and hence a combination of Gibbs and MH transitions is necessary (Casella, 2001; Gelman et al., 2013; Hoff, 2010). However, hierarchical models allow for an efficient sampling and the incorporation of structural knowledge through a proper choice of the prior (Gelman et al., 2013; Hoff, 2010).

Still, the choice of an adequate prior is crucial for Bayesian analysis (Gelman et al., 2013; Hoff, 2010). Since models do not produce information, they are only able to extract the information from both the sample and the prior (Gelman et al., 2013; Tschirk, 2014). In relation to the information provided by the data, the prior contributes to the predictions and thus should be chosen carefully (Gelman et al., 2013; Tschirk, 2014). If no knowledge is available, one should select an uninformative prior where all parameter values have an identical plausibility and thus $p(\boldsymbol{\theta}) = 1$ (Gelman et al., 2013; Tschirk, 2014). The resulting uniform distribution is not transformation-invariant, which means that although the density is uniformly distributed for a certain parametrization, it may not be for another (Gelman et al., 2013; Tschirk, 2014). This leads to the so-called Jeffrey's prior which is in fact invariant against monotonous transformations (Gelman et al., 2013; Tschirk, 2014).

To gain some confidence regarding the convergence of the samples MCMC diagnostics are necessary (Gelman et al., 2013; Hoff, 2010). Mostly, it is more easy to determine whether the samples are not converging than to ensure convergence (Hoff, 2010; Brooks et al., 2011). Many different MCMC diagnostic tools and view-points can be found in literature, however, there is no perfect MCMC diagnostic tool due to the fact that the state space is usually poorly connected (Brooks et al., 2011). Poorly connected state spaces result in a huge number of iterations until the samples jump from one region to another (Hoff, 2010; Brooks et al., 2011). In those situations, the convergence is conditioned on the part where the chain was started (Hoff, 2010; Brooks et al., 2011). Although this is related to multimodality, i.e., the posterior distribution has multiple peaks, multimodality does not necessarily cause conditional convergence (Brooks et al., 2011). This all gives rise to several questions including, e.g., the number of MCMC runs, the sample size, the determination of the starting point or the introduction of additional variables to increase convergence (Gelman et al., 2013; Brooks et al., 2011). It is important that the sample size is large enough to avoid that the samples remain in a region with low probability and to enable transitions between all regions with high probability (Gelman et al., 2013; Hoff, 2010). The latter phenomena is called mixing, whereas the first is referred to as stationarity (Gelman et al., 2013; Hoff, 2010). A perfect Markov chain does not stick and is able to jump between different regions in one step (Hoff, 2010). A direct measurement for stickiness is autocorrelation, which measures the correlation between samples depending on their distance (Gelman et al., 2013; Hoff, 2010). In general, the higher the autocorrelation, the larger the sample size has to be in order to achieve a good precision (Gelman et al., 2013; Hoff, 2010).

The Hamiltonian Monte Carlo algorithm is a further improvement of MCMC methods and allows to explore the state space more quickly (Gelman et al., 2013). For high-dimensional parameter spaces, even after re-parametrization and adjusted jump rules, Gibbs and MH transition-based models retain their random walk behavior (Gelman et al., 2013). In contrast, the Hamiltonian Monte Carlo algorithm combines MCMC methods with deterministic aspects by introducing an additional variable (Gelman et al., 2013). This variable is correlated to the proposal distribution and hence its trajectory allows for efficient parameter proposals (Gelman et al., 2013). Even though the Hamiltonian Monte Carlo algorithm is very efficient, in practice, it is often not applicable since the gradient of the log posterior density must be analytically accessible (Gelman et al., 2013). A numerical approximation of the gradients is considered too costly since it has at least as many components as the dimension of the related parameters (Gelman et al., 2013).

In addition to the discussed MCMC algorithms particle filters are often used as an alternative (Särkkä, 2013; Gelman et al., 2013). This class of Bayesian filters is useful in the case of multimodal or partly discrete posterior distributions, where the commonly used Gaussian processes are not applicable (Särkkä, 2013; Gelman et al., 2013). Several such algorithms have been reported in literature, e.g., importance sampling or sequential importance re-sampling (Särkkä, 2013). Basically, in these algorithms, parallel sequences are generated and evaluated at defined points (Särkkä, 2013). The samples of each sequence form a Markov chain and can be drawn in such a way that they form an approximation to the posterior distribution (Särkkä, 2013). Due to the fact that the approximation is not exact, the samples must be weighted accordingly (Särkkä, 2013). This way, sequences with poor approximations tend to vanish and, to compensate, new sequences are generated (Särkkä, 2013). After a sufficient number of iterations, the sequences with a poor fit to the posterior distribution are completely removed and those providing a good approximation are enriched (Särkkä, 2013). Because the sequences are initialized at different areas of the parameter space, they are more likely to explore the whole state space than MCMC algorithms (Särkkä, 2013; Gelman et al., 2013).

Once the parameter values are estimated, it is desirable to characterize the quality of the estimates and to further describe the region of the parameter space which likely contains the true parameter values (Gelman et al., 2013; Hoff, 2010). From a Bayesian point of view, for instance, a 95% credible interval is the region which covers the true parameter with a probability of 95% after the data is observed (Gelman et al., 2013; Hoff, 2010). Even though this concept is very useful and commonly used to report confidences of parameter estimates, sometimes credible intervals are less restrictive, which means that in certain situations values outside the credible intervals have a higher probability than values inside the interval (Gelman et al., 2013; Hoff, 2010). To address this issue more strict so-called highest posterior density interval are used (Gelman et al., 2013; Hoff, 2010). This type of credible interval is preferable because, at least for unimodal distributions, it ensures that all points in the interval are more likely than values outside the interval (Gelman et al., 2013; Hoff, 2010).

4.3.2 Application: The Solution Space of Static CRNs

As discussed in Section 4.1.2 the determination of the solution space for static biological systems is frequently addressed via MCMC algorithms (Schellenberger and Palsson, 2009; Megchelenbrink et al., 2014). Especially for constrained-based models the feasibility of the solution has to be guaranteed in each draw (Schellenberger and Palsson, 2009; Megchelenbrink et al., 2014). An intuitive method to explore such solution spaces is to generate independent uniform deviates in the unit sphere; but with the dimension of the system, the size of the solution space growths exponentially (Megchelenbrink et al., 2014; Kaufman and Smith, 1998). The method of choice is hit-and-run sampling (HRS), which is based on uniform sampling, where steps in random length and directions are taken iteratively (Smith, 1996; Megchelenbrink et al., 2014; Kaufman and Smith, 1998). The HRS method allows to move to each point of the entire solution space in a single step (Smith, 1996). It initiates a starting point $\boldsymbol{x}^{(0)}$ within the solution space Ω and generates a direction $\boldsymbol{d}^{(s+1)}$ (Smith, 1996). Then, a feasible point $\boldsymbol{x}^{(s+1)} = \boldsymbol{x}^{(s)} + \lambda \boldsymbol{d}^{(s+1)}$ is generated, where the density $f_s(\lambda)$ of $\lambda \in \Lambda_s = \left\{\lambda \in \mathbb{R} \middle| \boldsymbol{x}^{(s)} + \lambda \boldsymbol{d}^{(s+1)} \in \Omega\right\}$ of the target distribution is given by (Smith, 1996)

$$f_s(\lambda) = \frac{f\left(\boldsymbol{x}^{(s)} + \lambda \boldsymbol{d}^{(s+1)}\right)}{\int_{\Lambda_s} f\left(\boldsymbol{x}^{(s)} + r\boldsymbol{d}^{(s+1)}\right) dr}. \tag{4.42}$$

For stoichiometric analyses as detailed in Section 4.1, acquiring a uniform target distribution $f_s(\lambda)$ is sufficient (Megchelenbrink et al., 2014). Therefore, λ and $\boldsymbol{d}^{(s+1)}$ can be chosen such that they are uniformly distributed within the solution space (Smith, 1996; Megchelenbrink et al., 2014). Consequently, the samples $\boldsymbol{x}^{(s)}$ are also uniformly distributed and build a Markov chain in Ω (Smith, 1996; Megchelenbrink et al., 2014). In practice, this leads to an accumulation of sample points close to the boundaries of the solution space because the sampler tends to draw successive samples which are close to each other (Smith, 1996; Megchelenbrink et al., 2014). This drawback is addressed by the artificial centering hit-and-run (ACHR) method, which basically normalizes the chosen direction with respect to an approximate center to the previous iterates (Megchelenbrink et al., 2014; Kaufman and Smith, 1998). In most cases, this leads to a faster convergence and makes larger steps possible (Megchelenbrink et al., 2014). On the other hand, due to the normalization based on the approximate center, each sample depends on prior iterations and directions (Megchelenbrink et al., 2014). Hence the sampled sequence is no Markov chain and thus the convergence to the target distribution is not guaranteed (Megchelenbrink et al., 2014). Nonetheless, as an heuristic approach, the ACHR is commonly used to sample the solution space of CRNs due to its capability to explore the whole solution space (Schellenberger and Palsson, 2009; Megchelenbrink et al., 2014; Kaufman and Smith, 1998).

4.4 Parameter Estimation

As discussed in Chapter 2, developing ODE-based mathematical models in biology is an interactive process with the overall aim to match the given experimental data (Azeloglu and Iyengar, 2015; Jaqaman and Danuser, 2006; Ashyraliyev et al., 2009). The inverse problem is the estimation of the model parameters (Jaqaman and Danuser, 2006; Ashyraliyev et al., 2009). ODE-based models of biological systems are composed of reactions and thus kinetic rate laws. Dependent on the complexity of the system and the number of involved reactions and kinetic rate laws, the number of unknown or partially known parameters can be very large. Hence, estimating all parameters of a biological model is usually impossible due to the complexity of biological processes (Azeloglu and Iyengar, 2015; Ashyraliyev et al., 2009). Even if error free data is assumed, the structure of biological models may lead to non-identifiable parameters because in order to estimate a certain number of unknowns at least the same number of data points is required (Jaqaman and Danuser, 2006; Ashyraliyev et al., 2009). Such structural non-identifiable models have multiple solutions and thus are not observable (cp. Section 4.2.4) (Ashyraliyev et al., 2009). However, it is often possible to detect the non-identifiable parameters and subnetworks of such models (Ashyraliyev et al., 2009; Raue et al., 2009). Consequently, methods for model reduction can be used to reduce the fraction of non-identifiable parameters (Raue et al., 2009).

Although all parameters are structurally identifiable, there may be parameters with low confidence, i.e., parameters which are practically non-identifiable (Ashyraliyev et al., 2009; Raue et al., 2009). In recent literature there exist different definitions of identifiability (Ashyraliyev et al., 2009). Commonly, a parameter is called non-identifiable if its confidence interval is infinite (Raue et al., 2009). Roughly speaking, this means that there is an infinite number of alternative parameter values which match the observations as well. A simple way to proof structural identifiability of non-linear systems is based on Taylor approximations at measured time points (Ashyraliyev et al., 2009). A unique solution of the resulting system in terms of parameters implies structural identifiability and thus a distinguishable model structure (Ashyraliyev et al., 2009). Hence, structural identifiability is strongly related to the observability of dynamical systems, which can also be assessed via local Taylor approximations, as explained in Section 4.2.4.

Due to a lack of data and the non-identifiability of the parameters, models in systems biology are frequently build from incomplete datasets with mainly unknown parameters, and the direct determination of the parameters from experimental data is often not possible (Azeloglu and Iyengar, 2015; Jaqaman and Danuser, 2006; Ashyraliyev et al., 2009). Although nowadays large databases of measured parameters, e.g., kinetic constants, are accessible, in practice, they are of little value for mathematical modeling (Azeloglu and Iyengar, 2015). One major drawback of such databases and in general of combining data from different sources and systems is the variety of conditions under which they were measured (Azeloglu and Iyengar, 2015). These conditions, e.g., the pH-value or temperature, often disagree with the experimental data, which the modeler wants to fit and are thus biased (Azeloglu and Iyengar, 2015). However, they are useful starting points for parameter estimation (Azeloglu and Iyengar, 2015; Ashyraliyev et al., 2009). Ideally, parameters are estimated based on absolute protein concentrations. But, in most cases, only variations in the protein abundance and thus relative protein concentrations are accessible (Azeloglu and Iyengar, 2015).

To address the above issues, different strategies for data-based parameter estimation on a whole system-level have been developed (Jaqaman and Danuser, 2006; Ashyraliyev et al., 2009). Basically, approaches for parameter estimation aim to optimize a well-defined objective function which represents the discrepancy between experimental data and model predictions (Jaqaman and Danuser, 2006; Ashyraliyev et al., 2009). To avoid getting stuck in local extrema, global search algorithms, such as simulated annealing or evolutionary algorithms, are mainly probabilistic methods with the theoretical ability to explore the whole parameter space (Jaqaman and Danuser, 2006; Ashyraliyev et al., 2009). The genetic algorithm is a common evolutionary algorithm used for parameter estimation which adapts the survival-of-the-fittest principle from nature (Ashyraliyev et al., 2009; Raol et al., 2004). Similar to a population of individuals, a set of randomly chosen parameter vectors is initialized to explore the parameter space (Raol et al., 2004). As in evolutionary processes, individuals are then basically selected, recombined and mutated based on their fitness or, in this context, feasibility (Ashyraliyev et al., 2009; Raol et al., 2004). This means those individuals which are close to the optimal solutions pass through this process and survive or are combined with other individuals close to the optimum, and individuals far away from the optimum vanish (Ashyraliyev et al., 2009; Raol et al., 2004). Although global search strategies are capable to explore the whole parameter space, convergence is not guaranteed (Ashyraliyev et al., 2009; Raol et al., 2004).

The second class of methods are local search algorithms (Ashyraliyev et al., 2009). Local search algorithms guarantee solutions with a certain accuracy based on theoretical proofs of convergence provided that the initial guess is sufficiently close to the optimum (Ashyraliyev et al., 2009). Gradient-based methods, which require first derivatives or their approximations with respect to the parameters, are frequently used in systems biology (Karr et al., 2015; van Riel, 2006). Although derivative-based methods are very efficient for a couple of systems, they are often trapped in local extrema (Ashyraliyev et al., 2009; Karr et al., 2015). In addition, gradient-based methods are sensitive to measurement errors and starting points due to their deterministic nature (Ashyraliyev et al., 2009; Karr et al., 2015).

Although global search algorithms are able to explore the whole parameter space, they hardly converge towards the exact optimum of the objective (Ashyraliyev et al., 2009). On the other hand, local search algorithms converge faster towards the optimum but tend to get stuck in local extrema if they start far away from the optimum (Ashyraliyev et al., 2009). Therefore, it makes sense to combine both approaches, i.e., to use the global exploration of the parameter space and increase the convergence speed in areas close to the optimum via local search algorithms (Ashyraliyev et al., 2009). Today, these hybrid methods are frequently used (Ashyraliyev et al., 2009; Karr et al., 2015).

Bayesian inference, as discussed in Section 4.3, is another frequently used approach to estimate parameters in biological systems (Jaqaman and Danuser, 2006; Coleman and Block, 2006). Here, the quantity of interest such as credible intervals of parameters or expected parameter values are estimated by the use of prior knowledge via MCMC methods (Jaqaman and Danuser, 2006; Coleman and Block, 2006). Bayesian methods provide reliability measurements such as correlations and variances (Jaqaman and Danuser, 2006; Coleman and Block, 2006). According to Section 4.3, the prior distribution $p(\boldsymbol{\theta})$ reflecting the prior knowledge about the parameters together with the density of data given parameters $p(\boldsymbol{Y}|\boldsymbol{\theta})$ allows to draw samples from the parameter distribution $p(\boldsymbol{\theta}|\boldsymbol{Y})$ based on Bayes' Law (Jaqaman and Danuser, 2006; Coleman and Block, 2006). The probability $P(\boldsymbol{Y}|\boldsymbol{\theta})$ reflects the discrepancy between predictions and data with respect to the measurement noise $\boldsymbol{\sigma}^2$ and the linear combination of the observable states $\boldsymbol{h}(\boldsymbol{x})$ of the given model. Hence, $P(\boldsymbol{Y}|\boldsymbol{\theta})$ is often given by $P(\boldsymbol{Y}|\boldsymbol{\theta}) \sim N\left(\boldsymbol{h}(\boldsymbol{x}), \boldsymbol{\sigma}^2\right)$ or similar distributions (Jaqaman and Danuser, 2006; Hug et al., 2013; Coleman and Block, 2006). Bayesian inference is especially suitable for high-dimensional systems where other methods fail due to the underlying complex parameter space (Jaqaman and Danuser, 2006; Coleman and Block, 2006).

A well-chosen objective function is highly important for parameter estimation (Ashyraliyev et al., 2009; Karr et al., 2015; Kreutz et al., 2013). As discussed in Section 4.2.3, the LSE is an intuitive measurement for discrepancies between data and model predictions given a set of model parameters $\tilde{y}_i(t_l)\big|_{\boldsymbol{\theta}}$ (Ashyraliyev et al., 2009; Kreutz et al., 2013). Under the assumption of Gaussian-distributed measurement errors, the LSE estimator

$$\hat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}} \left[\sum_{i=1}^{I} \sum_{l=1}^{F} \frac{\left(y_i\left(t_l\right) - \tilde{y}_i\left(t_l\right) \big|_{\boldsymbol{\theta}} \right)^2}{\sigma_{i,t_l}^2} \right], \tag{4.43}$$

is equivalent to the maximum likelihood estimator (MLE) (Kreutz et al., 2013). With the probability of the observations $\boldsymbol{y}$ given a set of model parameters $\boldsymbol{\theta}$, the MLE

$$\hat{\boldsymbol{\theta}} = \arg\max_{\boldsymbol{\theta}} \left[\prod_{i=1}^{I} p\left(\boldsymbol{y}_{i,1:F} \,|\boldsymbol{\theta} \right) \right] \tag{4.44}$$

maximizes the overall probability of the data given the estimated set of parameters (Kreutz et al., 2013; Raue et al., 2009). One advantage of the MLE is the derivable profile likelihood as a direct measure for the impact of each model parameter and thus the identifiability of those parameters (Kreutz et al., 2013; Raue et al., 2009; Swameye et al., 2003). The profile likelihood reflects the dependency of the probability of the data given the optimal parameter set conditioned on a single parameter and allows for accurate confidence or credible intervals, respectively (Kreutz et al., 2013; Raue et al., 2009).

Altogether, estimates based on a single set of experimental data do not imply reliable estimations (Ashyraliyev et al., 2009; Azeloglu and Iyengar, 2015). Identifiability checks, validation based on perturbed data or other procedures can help to proof the credibility of the model (Ashyraliyev et al., 2009; Azeloglu and Iyengar, 2015). Perturbed data and other validation data are often hardly accessible, and thus a parameter identifiability analysis remains the only option to gain confidence about the estimates (Ashyraliyev et al., 2009; Azeloglu and Iyengar, 2015). On the other hand, one may argue that exact parameter estimates are not required to draw meaningful biological conclusions (Azeloglu and Iyengar, 2015).

CHAPTER 5

Analysis of the M_2 Receptor-dependent Signaling in CHO Cells

In this chapter, a time-independent model of the M_2 receptor-dependent signaling in CHO cells is developed and analyzed on the basis of the novel CFS. The thus obtained insights into the amplification processes serve as a first attempt to describe the DMR-related signaling. To gain a clearer picture about the involved time-dependent amplification processes the model proposed in this chapter is further extended in Chapter 8 by including time-resolved kinetics. This work is accepted in a peer-reviewed journal (cp. Appendix F).

5.1 Introduction

As discussed in Chapter 3, the M_2 receptor (encoded by the *CHRM2* gene) belongs to the family of GPCRs and is related to negative dromotropic and negative chronotropic events. Since GPCRs represent one of the most important target classes of proteins for drug discovery, the development of specific agonists and antagonists for muscarinic receptors, including the M_2 receptor, is still of high interest (Zheng, 2006). Iperoxo is a highly affine and efficacious muscarinic agonist that has recently served to elucidate the crystal structure of the active state of the M_2 receptor (Schrage et al., 2013, 2014; Kruse et al., 2013; Hu et al., 2010). In traditional pharmacology the ligand-binding event, second messenger concentrations, ion channel function, as well as tissue, organ or body responses are recorded. As discussed in Chapter 3, whole cell techniques nowadays are used to dissect the signaling of intact cells into different components (Schröder et al., 2011). In addition, iperoxo and its derivatives turned out to be valuable tools for gaining deeper insight into structure-signal relationships (Bock et al., 2014; Antony et al., 2009).

Recent experimental work explored the cellular response to iperoxo-induced M_2 receptor stimulation in CHO cells (Schrage et al., 2013; Kruse et al., 2013; Schrage et al., 2016). The cellular response was measured by DMR (Schröder et al., 2011). Since the DMR response can be assumed to be dependent on the M_2 receptor-induced signaling our aim was to model and study the corresponding reaction system. The pathway consists of proteins as well as the secondary messenger cAMP. The respective biochemical reactions are principally well known (Pierce et al., 2002; Linderman, 2009; Sunahara and Taussig, 2002; Taylor et al., 2012), but to the best of our knowledge no effort has been taken so far to derive a mathematical model, especially for CHO cells, which are very important in pharmaceutical research and for the industrial production of recombinant protein therapeutics (De Jesus and Wurm, 2011; Walsh, 2010).

In this work we developed a mass action based mathematical description of the M_2 receptor-dependent signaling network. Our developed model consists of 79 reactions, altogether involving 64 relevant proteins and secondary messenger molecules described in literature. In our joint signaling and secondary messenger model, all binding and (de-)phosphorylation events are explicitly taken into account in order to enable subsequent stoichiometric matrix and flux distribution analysis (Wiback et al., 2004). Although this kind of analysis is usually only employed for metabolic networks, as detailed in Section 4.1, our explicit modeling of binding and phosphorylation events enables the adaption of these techniques to a mixed signaling and secondary messenger system. The usefulness of applying stoichiometric matrix analysis techniques to signaling pathways has for example been demonstrated by Behre and Schuster (2009), who adapted EFM analysis to this situation. This work demonstrates, how the known flux sampling technique can be extended to incorporate partially available experimental information (here: cAMP production, PDE4 activation) (Smith, 1996). We tested our combined modeling and data driven sampling method by predicting key signaling mechanisms known from literature, but not explicitly encoded into the model. Our proposed CFS technique allows for qualitative predictions of downstream stimulation effects on actin and tubulin levels, which here serve as our markers for the mass redistribution effect. These qualitative predictions are in agreement with the experimental observations, which suggests CFS as a technique for model checking. This is further underlined by the possibility to combine CFS and EFM analysis yielding a statistical ranking of EFMs according to their expected biological relevance.

5.2 Network Reconstruction

GPCR-induced signaling is well-known in common literature (Pierce et al., 2002; Linderman, 2009; Taylor et al., 2012; Sunahara and Taussig, 2002). Specifically the link to the cAMP-induced signaling is in the focus of current pharmaceutical research (Milligan and Kostenis, 2006; Hu et al., 2010). Figure 5.1 depicts a schematic representation of the whole set of relevant molecules and their interplay, which are considered in our model. In particular the process of receptor-induced GP activation is well studied, where the ligand-bound receptor changes its physical structure and the inactive associated GP interacts with the receptor and dissociates into its subunits (Pierce et al., 2002). Thereby the α_i, α_s and $\beta\backslash\gamma$ subunits are activated and are able to interact independently with other proteins like AC (Sunahara and Taussig, 2002; Milligan and Kostenis, 2006). The GP subunit α_o has no significant influence on AC but it has an influence on the DMR (Milligan and Kostenis, 2006). AC is one of the most important proteins within the GP-mediated pathway and responsible for the secondary messenger production. The large number of AC and GP subtypes causes a highly complex subnetwork with many cross-reactions (Milligan and Kostenis, 2006; Sunahara and Taussig, 2002). Also the receptor activation cycle itself is not trivial. This first step in the signaling cascade is highly interesting for pharmaceutical research and led to well developed models for receptor activation and inhibition (Woodroffe et al., 2009; Chen et al., 2003; Strange, 2008; Bornheimer et al., 2004).

Besides this completely membrane bound subnetwork the PKA-induced phosphorylation cascade, and the feedback loop causing cAMP degradation, is well studied (Taylor et al., 2012). cAMP binds to PKA and causes its activation. But an increase in PKA activity also leads to an increase in PDE activity, which inactivates cAMP by degrading it to adenosine monophosphate (AMP) (Boswell-Smith et al., 2006). Through this mechanism the cell prevents a continuous overstimulation by excessive cAMP levels. Stimulation of the receptor population via the muscarinic agonist iperoxo induced a cellular DMR response at concentrations that are far lower than the corresponding concentration-binding relationships (Schrage et al., 2013). The same authors reported this amplification phenomenon also for other ligands, including the natural ligand ACh (Schrage et al., 2013). The exact nature of the amplification process is not understood so far, but may at least be partially attributed to intracellular signaling events (Schrage et al., 2016). According to common literature we suppose RGSs and GRKs to be of relevance. These proteins are closely related to the deactivation of the receptor and the GP subunits (Pierce et al., 2002; De Vries et al., 2000; Hollinger et al., 2003). In this work, we chose RGS14, GRK6 and GRK2 as important representatives for each group.

The DMR, as an optical biosensor-based procedure, measures the shift in wavelength resulting from intracellular mass movement caused by the rearrangement of cell organelles and transportation processes (Strange, 2008). In this work we chose actin and tubulin as DMR markers. Actin and tubulin are closely related to the cellular movement and we assume a strong correlation between changes in both protein and secondary messenger concentrations and the relative wavelength shifts measured by DMR (Hammond et al., 2008; Schmidt and Hall, 1998). Therefore, we took all proteins directly linked to actin and tubulin into account and assumed their activation to be correlated with the wavelength shift (cp. Figure 5.1). Maximal DMR response induced by iperoxo occurs typically after approximately 10 minutes (Schrage et al., 2013). According to the timescale and common literature we did not consider transcriptional downstream responses (Mayr and Montminy, 2001; Shaywitz and Greenberg, 1999). Further references are given in Appendix B.

5.3 Mathematical Modeling

All interactions shown in Figure 5.1 are explicitly formalized as mass action based elementary reactions and all known proteins and their occurring complexes are included (Horn and Jackson, 1972). Hence biological information from the available biochemical knowledge is preserved. Let $x_1, \ldots, x_N$ denote the concentrations of all molecules in the system. Then, according to

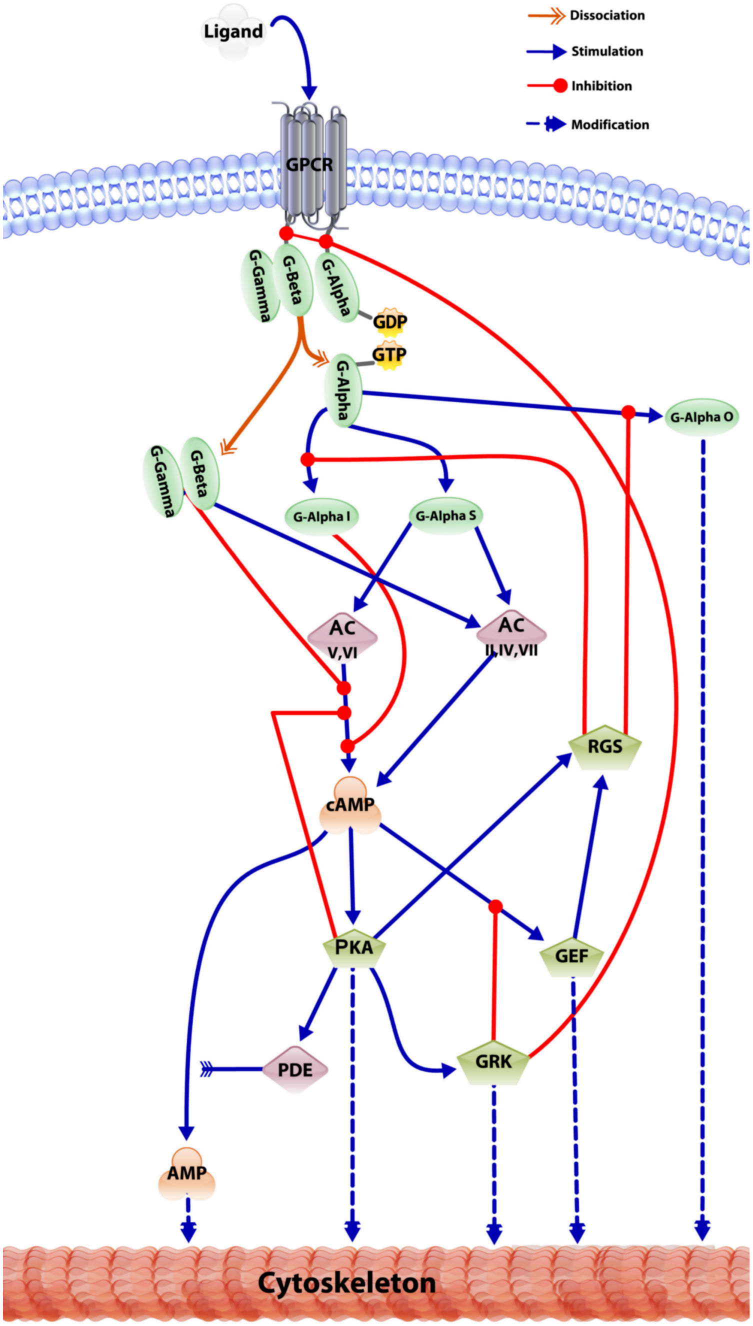


Figure 5.1: Schematic illustration of the M_2 receptor-dependent signaling and secondary messenger network in CHO-hM2 cells based on the known literature. The receptor is activated by a ligand (e.g. iperoxo) and induces the membrane-bound signaling cascade including G protein (here denoted as G) activation and the production of cAMP by AC. Via cAMP the signal is transferred to the PKA-induced phosphorylation cascade. The detailed reaction system can be found in Appendix B.

Section 2.3.4, the change of the concentration of molecule n can be written as

$$\frac{dx_n}{dt} = \sum_{k=1}^{K} s_{n,k} r_k, \tag{5.1}$$

where $s_{n,k}$ is the stoichiometric coefficient of molecule n in reaction k and r_k denotes the corresponding reaction velocity. As an example, here the PKA activation by cAMP is shown (Corbin et al., 1988):

$$PKA \rightarrow \overline{PKA} \tag{5.2a}$$
$$\overline{PKA} + 2cAMP \rightarrow PKA. \tag{5.2b}$$

Here, $\overline{PKA}$ denotes the inactive form of PKA. Let us denote the rate of both reactions by r_1 and r_2, respectively. The stoichiometric coefficients are $s_{1,1} = s_{2,2} = -1, s_{1,2} = s_{2,1} = 1$ and $s_{3,2} = -2$. We then obtain

$$\frac{d[PKA]}{dt} = s_{1,1} r_1 + s_{1,2} r_2 \tag{5.3a}$$
$$\frac{d[\overline{PKA}]}{dt} = s_{2,1} r_1 + s_{2,2} r_2 \tag{5.3b}$$
$$\frac{d[cAMP]}{dt} = s_{3,2} r_2. \tag{5.3c}$$

Altogether the modeled system contains 79 elementary reactions, which can be found in Appendix B. The full reaction system can be represented via a stoichiometric matrix $S \in \mathbb{R}^{N \times K}$ with entries $s_{n,k}$. In this matrix every molecule is represented by one row and every reaction is represented by one column.

The modeled system consists of several biochemical reaction types, namely binding, stimulation and inhibition. These biochemical events need to be represented appropriately in the reaction system and the stoichiometric matrix, respectively. This was done as follows: The protein activation via phosphorylation was modeled with the help of an intermediate molecule which represents the complex of the substrate and the related kinase. The kinase binds reversibly to the substrate and forms an intermediate complex which then dissociates irreversibly into the kinase and the modified substrate. For instance, GRK2 is phosphorylated by PKA (Cong et al., 2001). For this purpose we introduce the intermediate complex $PKA : GRK2$ and write the reaction system as

$$PKA + \overline{GRK2} \rightleftharpoons PKA : GRK2 \tag{5.4a}$$
$$PKA : GRK2 \rightarrow GRK2 + PKA. \tag{5.4b}$$

For every phosphorylation step we assumed a backward reaction $P^* \rightarrow P$, which dephosphorylates the phosphoprotein P^* with the help of an unknown phosphatase. In this example

GRK2 is dephosphorylated into $\overline{GRK2}$

$$GRK2 \rightarrow \overline{GRK2}. \tag{5.5}$$

Protein inhibition by kinases is modeled in a similar manner. A kinase binds reversibly to the target protein and forms an intermediate complex which then dissociates irreversibly into the kinase and the inactive protein. The inactive protein is now able to be activated again by another kinase. This process is illustrated using the inhibition of the guanine nucleotide exchange factor (GEF) by GRK2 (Eijkelkamp et al., 2010)

$$GRK2 + GEF \rightleftharpoons GRK2 : \overline{GEF} \tag{5.6a}$$

$$GRK2 : \overline{GEF} \rightarrow GRK2 + \overline{GEF}. \tag{5.6b}$$

As shown above, these reactions can be represented in the stoichiometric matrix S. The dimension of the stoichiometric matrix can be decreased by expressing the forward and backward direction of the same reversible reaction by one row where reaction rates can be both positive and negative. This is in contrast to strictly irreversible reactions where only positive reaction rates are allowed.

We also used the stoichiometric model to derive a system of ODEs based on the assumption of mass action kinetics, as shown in Appendix B.

5.4 Methods

5.4.1 Conservation Relationships

Since signaling events are relatively fast, we can assume that for each protein the overall total amount of phosphorylated, bound and unphosphorylated proteins is approximately constant, provided that the biological system is in steady state and the model was correct. Hence, checking conservation relationships is a means to check the consistency of our model.

According to Palsson (2006), conservation relationships under steady state conditions are mathematically identifiable from the null space of S^T. As detailed in Section 4.1.1, conservation relationships are all those vectors $\mathbf{g}$ for which

$$S^T\mathbf{g} = \mathbf{0}. \tag{5.7}$$

Each entry in $\mathbf{g}$ corresponds to exactly one molecule. The analysis of the entries of vectors $\mathbf{g}$ provides thus a means to verify whether the expected constant total concentration of each protein is fulfilled in reality. Moreover, we can also obtain insights into possibly existing constant protein concentrations within whole reaction cascades.

5.4.2 Stimulation of the System

We are interested in qualitative changes upon receptor stimulation. For the unstimulated system we assume a steady state characterized by constant concentrations of all molecules. Receptor stimulation causes a perturbation of this steady state resulting in dynamic changes of molecular concentrations. However, we assume that after some relaxation time the system will attain a – supposingly different – steady state, which is characterized by the molecular concentrations in the stimulated state. In reality, the stimulated state does not need to be a dynamic equilibrium in the strict sense, but we believe it to be a useful approximation for a situation of maximum response, where all concentrations are nearly constant over time. We believe that this working hypothesis is useful to analyze qualitative changes between the unstimulated and stimulated states, which is also supported by the fast – usually milliseconds – time scale of the signaling events in comparison to the observable duration of responses to receptor stimulation.

The unstimulated and the stimulated state correspond to different solutions of system under steady state conditions. i.e., $S\,\mathbf{v} = \mathbf{0}$. Our strategy will be to constrain the solution space of this system using experimental data. We will then use Monte Carlo sampling to compare possible stationary reaction rates $\boldsymbol{v}$ in the simulated and unstimulated state.

For a qualitative comparison we suppose the DMR response to be given as the sum of all fluxes with known influence on the wavelength shift

$$Response = \sum_{\tilde{k}=1}^{\tilde{K}} v_{\tilde{k}}. \tag{5.8}$$

Here $v_{\tilde{k}}$ denotes the activating flux related to molecule n with influence on the wavelength shift. The sum runs over all $\tilde{K}$ in-fluxes into tubulin and actin, which are considered as markers of the DMR response (Hammond et al., 2008; Schmidt and Hall, 1998; Schröder et al., 2011).

5.4.3 Sampling the Flux Polytope

Since we are interested in the general behavior of the system without the incorporation of additional rate parameters, steady state solutions of the system can in principle be found through MCMC HRS, as discussed in Section 4.3.2 (Smith, 1996; Brooks, 1998; Price et al., 2004). A single move in the HRS is performed by making a randomly chosen move within the unit sphere from a given feasible solution. Afterwards the step size is adjusted such that the new solution is also feasible (Smith, 1996; Megchelenbrink et al., 2014). A solution $\mathbf{v}^*$ is called feasible, if it satisfies

$$S\,\mathbf{v}^* = \mathbf{0} \tag{5.9a}$$

$$\alpha_k \leq v_k^* \leq \beta_k, \quad \forall k = 1, \dots, K \tag{5.9b}$$

with bounds α_k,β_k. Note that, without further constraints, fluxes could take any real value, but in reality fluxes are bounded. Hence, for all reversible reactions, we set $\alpha_k = -1000$ and

$\beta_k = 1000$ as loose bounds. For irreversible reactions, we set $\alpha_k = 0$. The flux bounds can in principle be used to incorporate experimental data. We will modify the flux bounds to qualitatively incorporate fold changes between stimulated and unstimulated cells, as explained in the following section.

5.4.4 Constrained Flux Sampling

We incorporate partially available data of experimentally measured relative (steady state) molecular concentrations into the above described flux sampling scheme in order to make qualitative predictions about flux changes upon stimulation. The approach thus does not require detailed knowledge of the kinetic rate constants.

Let $\tilde{v}_k$ denote the steady state flux of the k-th reaction in the case of an unstimulated receptor. According to the law of mass action (see Section 2.4) with rate parameters $\tilde{r}_k$ we have

$$\tilde{v}_k = \tilde{r}_k \prod_{n=1}^{N} \tilde{x}_n^{s_{n,k}}, \tag{5.10}$$

where $\{\tilde{x}_n\}$ is the set of molecules taking part in the particular reaction and $\tilde{x}_n$ their concentrations. Note that at this point we suppose the involved reversible reactions to be split into two irreversible ones. The flux $\hat{v}_k$ for the same reaction under stimulation can be defined accordingly now with concentrations $\hat{x}_n$. Usually, in experiments, relative concentration changes (fold changes) $f_n = \frac{\hat{x}_n}{\tilde{x}_n}$ are determined. Obviously,

$$\frac{\hat{v}_k}{\tilde{v}_k} = \prod_{n=1}^{N} f_n^{s_{n,k}} \tag{5.11}$$

which implies

$$\hat{v}_k = \tilde{v}_k \prod_{n=1}^{N} f_n^{s_{n,k}}. \tag{5.12}$$

Note that the stoichiometric coefficients $s_{n,k}$ in most cases are 1. The equation suggests a principal two-step procedure:

1. Perform conventional flux sampling for the unstimulated situation. This yields a set $\{\tilde{v}_k\}$.
2. Perform flux sampling for the stimulated situation by plugging the observed fold changes into Equation (5.12) in order to constrain the sampled fluxes.

In reality it may be more appropriate to consider confidence intervals $[f_n^{Min}, f_n^{Max}]$ for f_n because fold changes are subject to uncertainty. This can be addressed straightforwardly by replacing Equation (5.12) by an inequality

$$\tilde{v}_k \prod_{n=1}^{N} \left(f_n^{Min}\right)^{s_{n,k}} \leq \hat{v}_k \leq \tilde{v}_k \prod_{n=1}^{N} \left(f_n^{Max}\right)^{s_{n,k}}. \tag{5.13}$$

The quantity $\tilde{v}_k$ in practice needs to be estimated from the empirical flux distribution under steady state conditions. A reasonable choice is to take the mean or median of the sample distribution plus/minus the standard deviation for that purpose.

5.4.5 Elementary Flux Modes

According to Section 4.1.1, Schuster and Hilgetag (1994) introduced EFM analysis for characterizing the geometry of the solution polytope of the equation system $S\,\mathbf{v} = \mathbf{0}$ in a biologically interpretable manner. All solution vectors occur as linear combinations of EFMs. This leads to the polyhedral flux cone $P(S)$

$$P(S) = \left\{ \mathbf{v} \in \mathbb{R}^K \,\middle|\, \mathbf{v} = \sum_{\tilde{k}=1}^{Ke} w_k \mathbf{e}_k \quad w_k \geq 0 \right\}. \tag{5.14}$$

EFMs are then defined as the extreme rays or edges of the flux cone $P(S)$ (Schuster and Hilgetag, 1994; Llaneras and Picó, 2010). A formal assumption made in this equation is that reversible reactions are split into irreversible ones. Each EFM can be characterized as the minimal set of reactions which are required for a subsystem to exist as a functional unit (Papin et al., 2004). These subsystems either reflect fluxes through the whole reaction system or functional cycles within the system. Thus, the analysis of EFMs allows to identify biologically functional and interpretable “building blocks” of the biological reaction system. In case of signaling, this also implies that, without stimulation, there exists no EFM representing the whole network and no EFM describing the signaling flow through it.

In this work EFM analysis is combined with CFS: After having determined the flux distributions of the overall system under stimulated and unstimulated conditions we map the fluxes to each of the calculated EFMs. This is possible because each flux corresponds uniquely to one reaction. We then compute the median of all fluxes related to a specific EFM. Since we generated a large (here: $100,000$) sample of flux vectors we obtained an empirical distribution of these medians for each EFM. The significance of the difference in these distributions between stimulated and unstimulated conditions can be assessed via a Wilcoxon rank test, yielding a p-value. Because we do not only compare one but several EFMs, multiple testing correction of p-values is performed via control of the false discovery rate (FDR) (Benjamini and Hochberg, 1995). Moreover, we estimated the median fold change between stimulated and unstimulated conditions.

5.5 Data

Parts of the experimental data (i.e., the dose-response relationships) were taken from Schrage et al. (2013): In that paper, the DMR was measured at 13 different concentrations of the M_2 receptor-specific activator iperoxo, giving rise to dose-response curves (cp. Figure 5.2a).

In addition to these data obtained by Schrage *et al.*, the cAMP response to iperoxo treatment was measured here (cp. Figure 5.2b). This was done after 30 minutes of iperoxo incubation with a concentration of 0.1 μM = 10^{-7} M, which corresponds to a full DMR response (Schrage et al., 2013). The induced cAMP fold change, calculated as the ratio between the cAMP level related to the iperoxo concentration of 0.1 μM and the 95% confidence interval of the basal cAMP level (see Figure 5.2b), is given by [2.22, 2.71].

In addition, we measured the activation of PDE4 after 30 minutes for the iperoxo concentration of 0.1 μM. The 95% confidence interval of the active PDE4 level (see Figure 5.2d) is [3.21, 3.46]. Further information regarding the conducted experiments can be found in Appendix B.

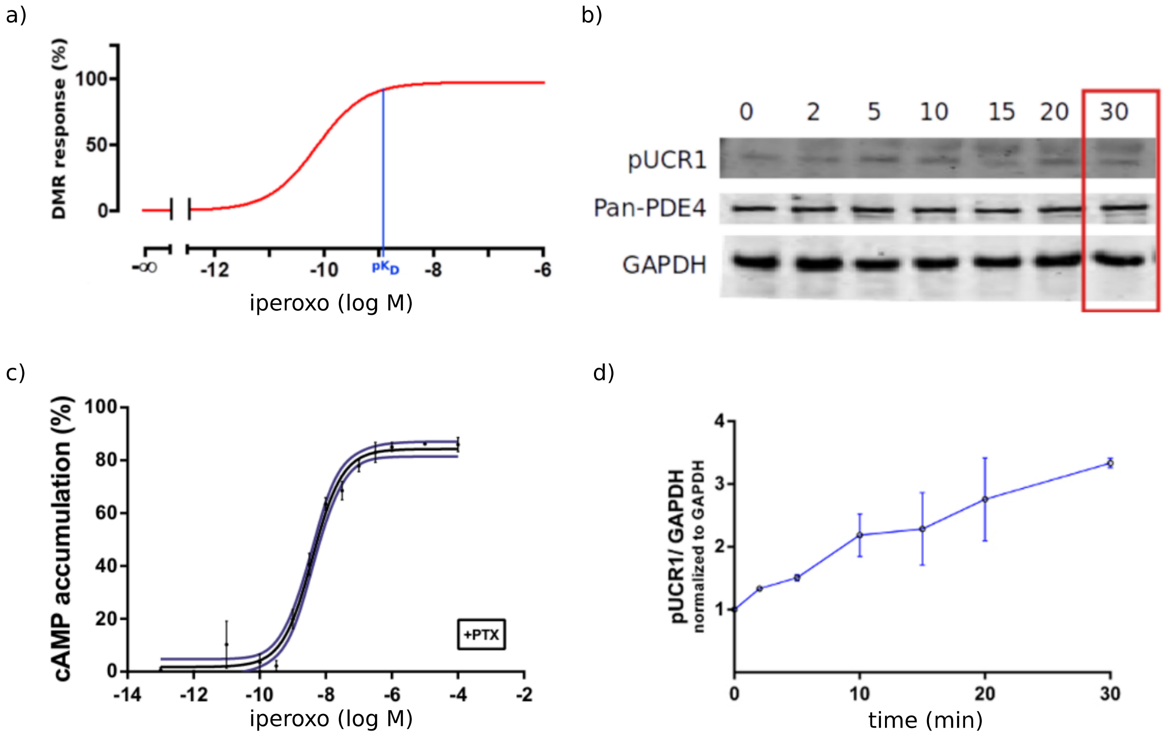


Figure 5.2: Observed DMR, cAMP and PDE4 levels. (a) DMR concentration-response curve of iperoxo modified from (Schrage et al., 2013). The affinity between iperoxo and the receptor (pK_D) of the iperoxo-induced ligand-binding curve for intact cells (obtained from (Schrage et al., 2014)) is marked by the blue line. (b) Western blot for the total amount of PDE4 (Pan-PDE4) and active PDE4 (pUCR1) under stimulation with 0.1 μM iperoxo normalized against GAPDH. (c) Concentration effect curve for the measured iperoxo-induced G protein subtype α_s mediated accumulation of cAMP with standard deviations and estimated confidence intervals marked by the blue line. The inactivation of G protein subtype α_i was induced via a pretreatment with PTX. The inactivation of the cAMP-inhibiting G protein subtype α_i allows for matching the measurements with their corresponding G protein subtype α_s mediated network fluxes. The accumulation of cAMP in the absence of test compounds was set to 0%, and maximum forskolin-induced binding was set to 100%. (d) Fold change for normalized pUCR1.

5.6 Results

5.6.1 Resulting Conservation Relationships

We uncovered 14 conservation relationships within the modeled biological system under steady state conditions. Figure 5.3 illustrates which sets of proteins have to maintain a constant total concentration. As illustrated in Figure 5.3, we found at least one conservation relationship for each protein which reflects the main endogenous signaling cycles. Thus our expectations coming from the signaling character of our modeled system are verified.

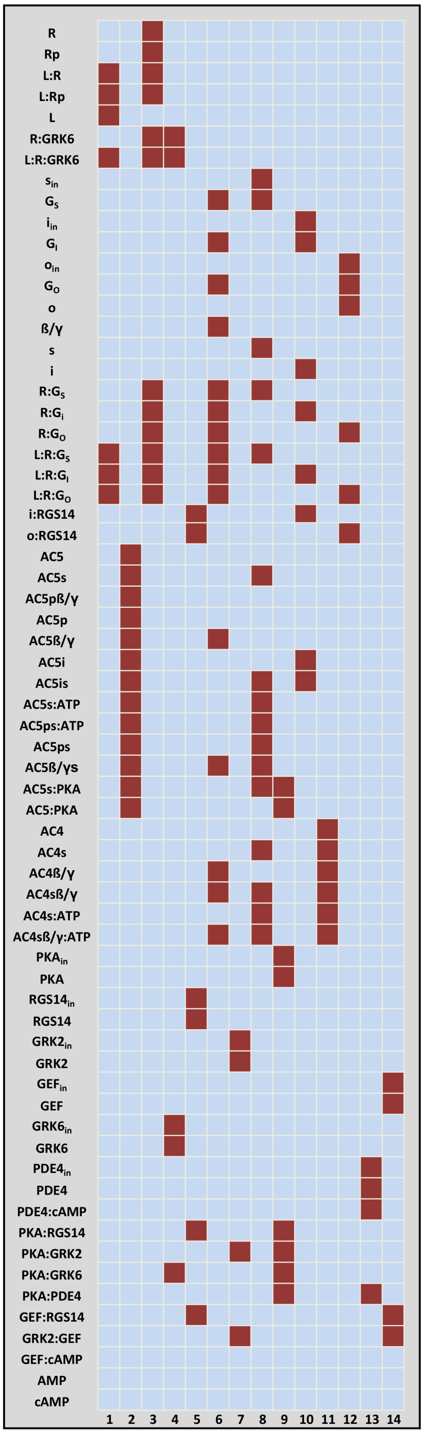
5.6.2 Constraint Flux Sampling Correctly Predicts DMR Response under Receptor Activation

We applied the CFS framework described above incorporating cAMP as well as PDE4 fold changes into flux constraints. DMR response measurements were not taken into consideration at this point, but left out for independent validation. Figure 5.4 depicts the distributions of those selected fluxes, which according to our CFS analysis are predicted to show a statistically significant shift under stimulation (FDR <1%, Wilcoxon signed rank test with Benjamini & Hochberg's FDR control of p-values for multiple testing (Hollander, 1999; Benjamini and Hochberg, 1995)). The CFS predicts a significant change in RGS14. A slight decrease of the receptor deactivation and increase of the GP subtype α_i deactivation via RGS14 can be expected under stimulation according to our simulation. The receptor deactivation is compensated by an increasing receptor recycling. This phenomenon of signal regulation by RGS14 and GRK6 is well described in literature where both proteins are known as important signal regulators (Pierce et al., 2002; Dale and Rang, 2011; Berridge, 2012). The RGS family is involved in the extinction of the GP-dependent signaling (Zhang and Mende, 2011), e.g., via receptor desensitization or endocytosis (Reiter and Lefkowitz, 2006). This causes the downregulation of the GP subtype α_i downstream and together with cAMP forms a positive feedback loop. More specifically, the inhibition of the cAMP inhibitor GP subtype α_i leads to an increase of the cAMP production. In addition, significant GP subtype α_i deactivation causes a significant increase of the GRK6 activation.

The box plots clearly highlight that – besides cAMP – increasing levels of AMP production and cAMP degradation are expected under stimulation, which is in agreement with current literature (Berridge, 2012; Pierce et al., 2002; Sunahara and Taussig, 2002; Strange, 2008). CFS is able to correctly predict a significant positive wavelength shift, i.e., DMR response, under receptor stimulation which is in agreement with our experimental validation data (cp. Figure 5.2a). Hence, CFS allows for a qualitative check of our pathway model.

5.6.3 Knock-out Simulations

To further check the hypothesized relevance of RGS14 for the observed DMR response we conducted an in silico knock-out simulation. This means we restricted all fluxes going through



	1	2	3	4	5	6	7	8	9	10	11	12	13	14
R			🟥											
Rp			🟥											
L:R	🟥		🟥											
L:Rp	🟥		🟥											
L	🟥													
R:GRK6			🟥	🟥										
L:R:GRK6	🟥		🟥	🟥										
s_{in}								🟥						
G_S						🟥		🟥						
i_{in}										🟥				
G_I						🟥				🟥				
o_{in}												🟥		
G_O						🟥						🟥		
o												🟥		
ß/γ						🟥								
s								🟥						
i										🟥				
R:G_S			🟥			🟥		🟥						
R:G_I			🟥			🟥				🟥				
R:G_O			🟥			🟥						🟥		
L:R:G_S	🟥		🟥			🟥		🟥						
L:R:G_I	🟥		🟥			🟥				🟥				
L:R:G_O	🟥		🟥			🟥						🟥		
i:RGS14					🟥					🟥				
o:RGS14					🟥							🟥		
AC5		🟥												
AC5s		🟥						🟥						
AC5pß/γ		🟥												
AC5p		🟥												
AC5ß/γ		🟥				🟥								
AC5i		🟥								🟥				
AC5is		🟥						🟥		🟥				
AC5s:ATP		🟥						🟥						
AC5ps:ATP		🟥						🟥						
AC5ps		🟥						🟥						
AC5ß/γs		🟥				🟥		🟥						
AC5s:PKA		🟥						🟥	🟥					
AC5:PKA		🟥							🟥					
AC4											🟥			
AC4s								🟥			🟥			
AC4ß/γ						🟥					🟥			
AC4sß/γ						🟥		🟥			🟥			
AC4s:ATP								🟥			🟥			
AC4sß/γ:ATP						🟥		🟥			🟥			
PKA_{in}									🟥					
PKA									🟥					
$RGS14_{in}$					🟥									
RGS14					🟥									
$GRK2_{in}$							🟥							
GRK2							🟥							
GEF_{in}														🟥
GEF														🟥
$GRK6_{in}$				🟥										
GRK6				🟥										
$PDE4_{in}$													🟥	
PDE4													🟥	
PDE4:cAMP													🟥	
PKA:RGS14					🟥				🟥					
PKA:GRK2							🟥		🟥					
PKA:GRK6				🟥					🟥					
PKA:PDE4									🟥				🟥	
GEF:RGS14					🟥									🟥
GRK2:GEF							🟥							🟥
GEF:cAMP														
AMP														
cAMP														

Figure 5.3: Calculated conservation relationships: Each column represents one conservation relationship and each row a protein. Red cells indicate proteins involved in a concentration relationship. The sum over all marked protein concentrations per column is constant. The inactive form of each protein is indicated by the subscript "*in*".

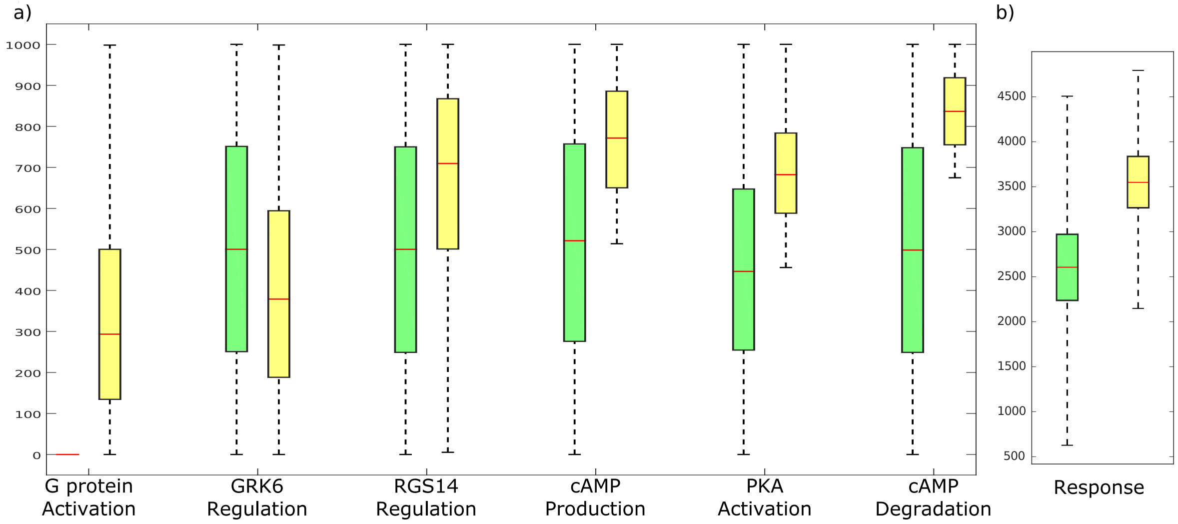


Figure 5.4: Predicted fluxes without stimulation (green) and under stimulation (yellow). (a) Box plots illustrating the distribution of selected fluxes under different conditions using cAMP measurements. Green indicates the unstimulated and yellow the stimulated condition. Ligand-induced G protein activation is shown for the α_s subtype here. Box plots for all fluxes can be found in Appendix B. (b) Overall response given as the sum of fluxes into tubulin and actin, cp. Equation (5.8). Related median fold changes are shown in Table 5.1.

	GRK6 Regulation	RGS14 Regulation	cAMP Production	PKA Activation	cAMP Degradation	Response
FC (median)	1.27	1.33	1.47	1.54	1.68	1.36

Table 5.1: Median fold changes (FC) related to Figure 5.4.

this molecule to zero while repeating our CFS. To investigate the effect of the different molecules on the DMR response, we performed knock-out simulations for all molecules except for GP subtype α_s, AMP, AC, cAMP, PKA, the receptor and the ligand. We then ranked the molecules according to their statistical significance of the influence on the DMR-related fluxes because no steady state solutions are possible when constraining the fluxes through these important signaling molecules to zero. AMP, AC, cAMP, PKA, the receptor and the ligand are not considered because these are characteristic molecules for signaling, and removing these molecules is unphysiological. Table 5.2 shows the proteins with influence on the response under stimulation.

Altogether our simulations underline our findings from Section 5.6.2. The highest impact was found for RGS14 and GP subtype α_o, followed by PDE4. Furthermore, GRK6 has a significant influence.

Rank	FDR	Median Fold Change (simulated)	Median Flux Difference (simulated)	Protein
1	$< 1E-6$	1.24	680	GP α_o
2	$< 1E-6$	1.24	675	RGS14
3	$< 1E-6$	1.18	545	PDE4
4	$< 1E-6$	1.17	535	GRK2
5	$< 1E-4$	1.12	395	GP α_i
6	$< 1E-4$	1.12	394	GRK6

Table 5.2: Proteins ranked with respect to their predicted influence on the DMR response. The influence on the response was estimated by the median fold change expected by a knock-out simulation of each protein. The statistical significance of each simulated fold change is shown in terms of FDR. A high fold change implies a strong influence of the particular protein.

5.6.4 Combining EFMs and CFS Reveals Important Subnetworks and Regulatory Mechanisms

As a last step we applied EFM analysis to the system with ligand stimulation, resulting in 63 EFMs (see Appendix B). Notably, many of these EFMs represent similar biological mechanisms. We ranked all EFMs with respect to their predicted change under stimulation by the method described in Section 5.4.4. Table 5.3 shows all EFMs with an FDR lower than 0.001 and a median fold change greater than 1. Interestingly enough, four of the most significant EFMs describe the GP subtype α regulation via RGS14, and also the receptor regulation via GRK6 is among the most significant EFMs (Figures 5.5 and 5.6). This is in full agreement with our previous findings and provides a possible explanation for the relevance of these molecules. Further significant EFMs are related to PDE activation and cAMP/GEF production (in agreement with our experimental data).

5.7 Conclusion

In this work we presented a comprehensive mathematical model of the M_2 receptor-dependent joint signaling and secondary messenger network. The motivation for this work comes from the pharmacological relevance of the M_2 receptor and the induced cellular responses. Whereas in principle the individual parts of our studied system are well described in the biological literature, to our knowledge there have been no attempts so far to combine these information into a mathematical model.

A quite specific property of our model is the combination of pure signaling events on the protein level with secondary messenger molecule production and degradation. Following Behre and Schuster (2009), we here adapted methods from stoichiometric matrix analysis (EFM analysis, conservation laws, flux sampling) that are usually known in the field of metabolic network analysis and do not require information about kinetic rate constants. These methods

Rank	Fold Change	Median Flux (unstimulated)	Median Flux (stimulated)	Related to
1	*NA*	0	619	GRK6
2	*NA*	0	539	GRK6:R
3	*NA*	0	499	Receptor
4	2.00	250	501	GP α_s
5	1.68	499	837	PDE
6	1.58	506	816	GEF
7	1.57	507	815	GEF
8	1.57	507	816	GEF
9	1.57	507	815	GEF
10	1.54	446	682	cAMP
11	1.54	446	682	cAMP
12	1.54	446	682	cAMP
13	1.54	446	682	GEF
14	1.53	513	789	GEF
15	1.52	513	789	GEF
16	1.52	514	788	GEF
17	1.52	514	788	cAMP
18	1.48	481	709	cAMP
19	1.48	481	709	cAMP
20	1.48	481	709	cAMP
21	1.48	481	709	cAMP
22	1.34	500	709	RGS14
23	1.34	500	709	RGS14
24	1.33	500	707	RGS14
25	1.33	500	707	RGS14
26	1.27	500	379	GRK6

Table 5.3: Significant EFMs ($FDR < 1E - 6$) ranked by their median predicted fold change induced by stimulation. Note that the first three fold changes are not computable because without stimulation there is no flux through these EFMs. The complete table can be found in Appendix B.

operate mainly on the level of a biochemical reaction network. We demonstrated that in this way it is possible to draw biologically meaningful conclusions about the principle behavior of our studied system, which are in agreement with the current knowledge. In that context our proposed constrained flux sampling method allowed to include measured cAMP and PDE fold changes in order to make qualitative predictions about receptor stimulation effects on network level that were verifiable via experimental data (DMR measurement) not been used by CFS. This work demonstrates that our CFS method can be used in the context of in silico knock-out simulations in order to identify relevant features of our studied system. More specifically, we found RGS14 of major relevance, which is in agreement with the current literature, but certainly

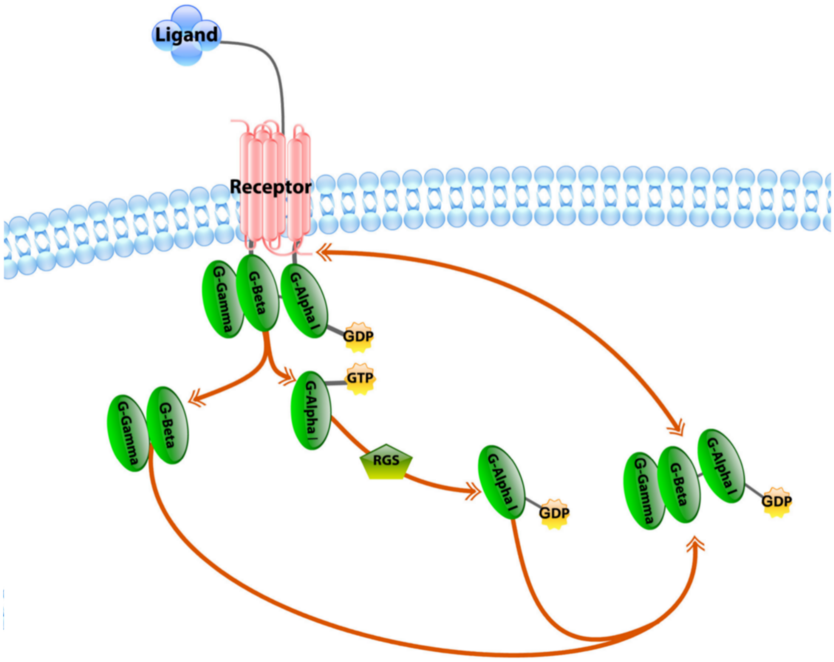


Figure 5.5: Elementary flux mode for G protein (here denoted as G) regulation via RGS14. First the inactive G protein complex consisting of the subunits α_i and $\beta\backslash\gamma$ binds to the active receptor and the bound GDP is replaced by GTP while the G protein dissociates from the receptor and splits into its subunits $\beta\backslash\gamma$ and α_i. Afterwards the activated α subunit is deactivated by replacing GTP with GDP mediated by RGS14. In the last step, the deactivated GDP-bound subunit α again associates with the subunit $\beta\backslash\gamma$ and forms the inactive G protein.

requires further experimental investigations of the biological system under consideration. The combination of CFS and EFM suggests that this is specifically due to the regulation of the GP subunit α_i via RGS14. The dynamical system developed in Chapter 8 further underlines the importance of RGS14 for the M_2 receptor-dependent whole cell response.

Altogether this work demonstrates that CFS, as well as other established methods for stoichiometric matrix analysis, could be valuable tools for model checking for mixed signaling and non-signaling networks. Such a model checking procedure would also simplify and speed up possible follow up model refinements driven by quantitative time series measurements.

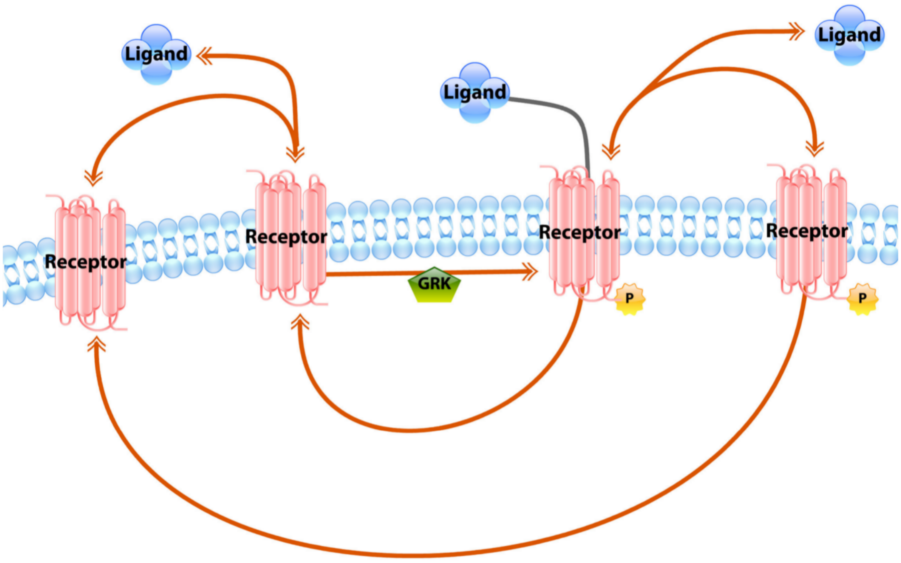


Figure 5.6: Elementary flux mode for the GRK-mediated receptor inactivation via phosphorylation. The ligand-bound receptor gets phosphorylated by GRK. The phosphorylated and hence inactive receptor is no longer able to mediate the G protein activation. After ligand-dissociation the receptor gets dephosphorylated and is now again able to mediate the G protein activation.

CHAPTER 6

The Dynamic Elastic-Net

Even though general ability of the derived M_2 receptor-dependent signaling pathway to reflect the observed properties has been demonstrated in Chapter 5, it is not fully ensured that all driving mechanisms are included in a refined time-resolved system. As discussed in Chapter 2, dynamical systems are more complex than static systems and thus a method to detect and correct erroneous models is highly desirable. The DEN is a deterministic approach based on OCT to address this issue, and a first attempt to towards automatic model correction (Engelhardt et al., 2016). This is important in order to refine the developed model for the M_2 receptor-dependent signaling in CHO cells in a reliable manner. This work has been published in a peer-reviewed journal (cp. Appendix F).

6.1 Introduction

Ideally, a good model covers the essential features of the system whilst still being simple enough for interpretation and mechanistic understanding. Developing a good model is usually a labor-intensive manual effort. As discussed in Chapter 2, in biology the system to be modeled is often only partially known and the distinction of relevant and irrelevant features and variables can be difficult (Swameye et al., 2003; Sunnåker et al., 2013; Kahm et al., 2012). But, even if the major components of a biological system are well known, the sheer complexity of the system might prevent the development of an accurate mathematical model, either because the quantitative data necessary for modeling is not available or because the model is itself too complex to be useful. Thus, researchers in systems biology are frequently confronted with a paradoxical situation: A model is needed to better understand the system and to design informative experiments, but the system is too large and complex for mathematical modeling given the limited amount of knowledge, data and time.

One strategy, similar to bottom-up approaches (see Chapter 2), for modeling is to start with a simple model, which incorporates the most interesting variables and interactions as well as the known input stimuli to the system (Figure 6.1a). For example, to model a CRN, we might incorporate the concentrations of a few interesting proteins as dynamic state variables and integrate the knowledge about the reactions into simplified assumptions about the interactions between these states. Here, this simple draft model is referred to as the nominal system.

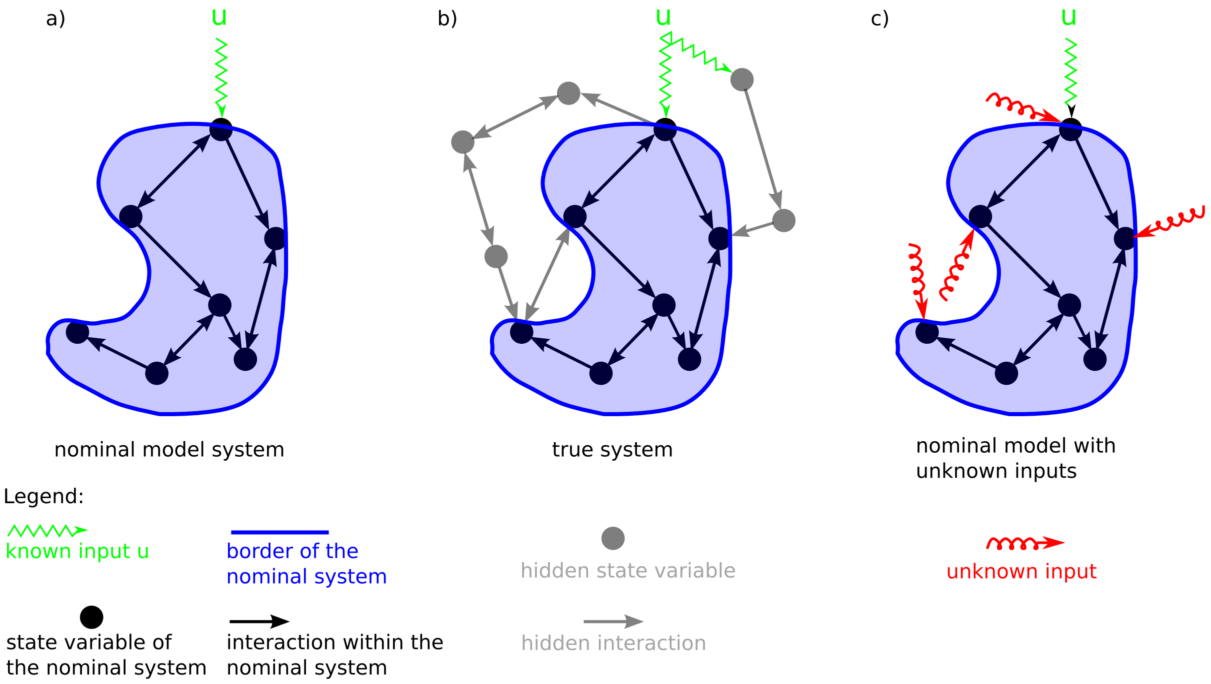


Figure 6.1: (a) The nominal model represents the current assumptions about the true system. The system is specified by its dynamic state variables and their interactions, here represented as vertices and edges of a graph. The system border defines the distinction between internal states and exogenous inputs. The exogenous inputs u are assumed to be known. (b) In reality, the nominal model is embedded in a larger network outside the nominal system border. The hidden dynamics of the exosystem interact with the nominal system. In addition, some interactions between nominal state variables might be missing or misspecified in the nominal model. These model errors can potentially lead to discrepancies between model and experimental data. (c) Representation of model errors as hidden inputs to the nominal model. The DEN approach infers the hidden inputs from data and thereby corrects for the bias in the nominal state variables induced by model errors.

There are two reasons why the nominal model might not be in sufficient agreement with the experimental data (Figure 6.1b): First, some interactions between the nominal state variables could be missing or misspecified. For a reaction network, this would mean that there are missing biochemical reactions, incorrect assumptions about the reaction kinetics or inaccurate parameter estimates. Second, the nominal system is in fact – opposed to the typical situation in many areas of physics – open and embedded into a larger dynamical system (von Bertalanffy, 1950). Exogenous variables, which are not incorporated, but interact with the nominal model, might act as hidden inputs and thereby alter the dynamics of the nominal system. It is the task of the modeler to first identify the most relevant errors in the nominal model and then compare different model versions in order to achieve a better fit to the available experimental data. This process is labor-intensive, and in many cases a trial-and-error exercise, even with the help of innovative software and algorithms assisting modeling and model comparison (Sunnåker et al., 2013; Rodriguez-Fernandez et al., 2013; Babtie et al., 2014).

We introduced a computational method for ODEs, which automatically estimates the model error from the data. As discussed, ODEs are frequently used in different areas of biology including biochemical reaction networks, pharmacokinetics, pharmacodynamics and population dynamics.

The basic idea of the method is to represent errors in the nominal model as hidden inputs to the state variables (Figure 6.1c) and to estimate these inputs from the experimental data (Mook and Junkins, 1988; Kolodziej and Mook, 2005; Kahm et al., 2012). Since this is an inverse problem with potentially many different solutions, we proposed a regularized method which provides parsimonious error estimates. Due to its formal similarity to the elastic-net regression approach (Zou and Hastie, 2005), we termed our algorithm the dynamic elastic-net (DEN).

The dynamic elastic-net provides important information about the variables in the nominal model, which are targeted by model errors. In addition, the DEN removes the bias in the nominal state variables induced by the model error. This is important for the frequent situation where not all nominal states (e.g. protein concentrations) can directly be measured. The utility of the DEN is demonstrated here for two established models of the erythropoietin receptor (EpoR) and of the photomorphogenic ultraviolet B (UV-B) signaling network (Swameye et al., 2003; Ouyang et al., 2014). Further examples, including a model for G protein signaling and models for several network motifs as well as some technical details, are given in Appendix C.

6.2 Methods

6.2.1 The Nominal Model

We assume that a nominal ODE model

$$\begin{aligned} \dot{\tilde{\boldsymbol{x}}}(t) &= \tilde{\boldsymbol{f}}(\tilde{\boldsymbol{x}}(t), \boldsymbol{u}(t)) && (6.1a) \\ \tilde{\boldsymbol{y}}(t) &= \boldsymbol{h}(\tilde{\boldsymbol{x}}(t)) && (6.1b) \\ \tilde{\boldsymbol{x}}(0) &= \tilde{\boldsymbol{x}}_0 && (6.1c) \end{aligned}$$

has been proposed to describe the dynamics of the system under consideration. The state vector $\tilde{\boldsymbol{x}}(t) = (\tilde{x}_1(t), \ldots, \tilde{x}_N(t))^T$ contains the N dynamic variables $\tilde{x}_n(t)$, and $\dot{\tilde{\boldsymbol{x}}}(t)$ is the derivative with respect to time t. The initial value of the state vector is $\tilde{\boldsymbol{x}}_0$. For a biochemical reaction network, $\tilde{x}_n$ is often the concentration or abundance of the n-th species. The function $\boldsymbol{u}(t) = (u_1(t), \ldots, u_P(t))^T$ represents a known external input to the system. The dynamics of the state variables is determined by the function $\tilde{\boldsymbol{f}} = (\tilde{f}_1, \ldots, \tilde{f}_N)^T$ and encodes the model assumptions made in the nominal model. This can be represented as a graph, where each node corresponds to one variable and a directed edge from n to j indicates that the time-derivative of $\tilde{x}_n$ depends on $\tilde{x}_j$ (Figure 6.1a) (Liu et al., 2013). A green zigzag arrow illustrates that $\dot{\tilde{x}}_n$ is directly influenced by a known input. Typically, not all state variables $\tilde{\boldsymbol{x}}$ can directly be measured. The variables $\tilde{\boldsymbol{y}}(t) = (\tilde{y}_1(t), \ldots, \tilde{y}_l(t))^T$ represent all outputs which are experimentally accessible. In Equation (6.1b), we assume that the mapping $\boldsymbol{h}$ from the state $\boldsymbol{x}$ to the output $\boldsymbol{y}$ is known. Tilde is used to highlight that $\tilde{\boldsymbol{f}}$ and thus $\tilde{\boldsymbol{x}}(t)$ are usually not perfectly known due to limited or uncertain knowledge about the true underlying dynamics.

6.2.2 Representation of the Model Error

The response of the real natural system to a known input stimulus $\boldsymbol{u}(t)$ is usually measured at discrete time points $t_1 \leq \cdots \leq t_F$ and provides experimental observations for the output $\boldsymbol{y}(t_l)$. A part of these data is usually used to estimate the parameters of the model. We consider the initial parameter estimates as part of the nominal model specification $\tilde{\boldsymbol{f}}$ in Equation (6.1a).

The nominal model is deemed unsatisfactory, when its output $\tilde{\boldsymbol{y}}(t_1), \ldots, \tilde{\boldsymbol{y}}(t_F)$ is not in sufficient agreement with the data $\boldsymbol{y}(t_1), \ldots, \boldsymbol{y}(t_F)$. One source of model error comes from hidden inputs to the nominal system, which are caused by dynamical processes exogenous to the nominal system (Figure 6.1b). In addition, there might be missing or erroneous interactions between the state variables $\tilde{\boldsymbol{x}}$ in the nominal model itself. Both types of model error can be represented by hidden inputs $\boldsymbol{w}(t) = (w_1(t), \ldots, w_N(t))^T$ acting on the nodes of the nominal model (Figure 6.1c). The "true" dynamics $\boldsymbol{x}(t)$ of the real system can be described by

$$\begin{aligned} \dot{\boldsymbol{x}}(t) &= \tilde{\boldsymbol{f}}\,(\boldsymbol{x}(t), \boldsymbol{u}(t)) + \boldsymbol{w}(t) && (6.2a) \\ \boldsymbol{y}(t) &= \boldsymbol{h}\,(\boldsymbol{x}(t)) && (6.2b) \\ \boldsymbol{x}(0) &= \boldsymbol{x}_0. && (6.2c) \end{aligned}$$

Here, the state $\boldsymbol{x}(t) = (x_1(t), \ldots, x_N(t))^T$ represents the same variables as the nominal state $\tilde{\boldsymbol{x}}(t)$, but tilde is suppressed to distinguish solutions of model (6.2) from that of the nominal model. The model error is the difference $\boldsymbol{w}(t) = \dot{\boldsymbol{x}}(t) - \tilde{\boldsymbol{f}}\,(\boldsymbol{x}(t), \boldsymbol{u}(t))$ between the rate of change of the true system $\dot{\boldsymbol{x}}(t)$ and the nominal system $\tilde{\boldsymbol{f}}\,(\boldsymbol{x}(t), \boldsymbol{u}(t))$, evaluated along the true state trajectory $\boldsymbol{x}(t)$. Thus it incorporates any discrepancy between the true system and the nominal system. The known input $\boldsymbol{u}$ and the output function $\boldsymbol{h}$ are assumed to be identical to the nominal model (6.1). However, the impact of measurement noise is discussed below.

The typical approach to model improvement is to compensate for the model error $\boldsymbol{w}(t)$ by explicit mathematical expressions, often additional differential equations. This increases the number of variables and parameters in the model. Here, we proceed differently by estimating the model error $\boldsymbol{w}$ from the data, which also enables us to correct for the bias $\boldsymbol{x}(t) - \tilde{\boldsymbol{x}}(t)$ of the state estimate incurred by the nominal model.

6.2.3 Estimating the Unmodeled Dynamics

To estimate the model error $\boldsymbol{w}(t)$, we use the observer system

$$\begin{aligned} \dot{\hat{\boldsymbol{x}}} &= \tilde{\boldsymbol{f}}(\hat{\boldsymbol{x}}(t), \boldsymbol{u}(t)) + \hat{\boldsymbol{w}}(t) && (6.3a) \\ \hat{\boldsymbol{y}}(t) &= \boldsymbol{h}(\hat{\boldsymbol{x}}(t)), && (6.3b) \end{aligned}$$

which is a copy of Equations (6.2a) and (6.2b). The hat marks estimates of the state $\hat{\boldsymbol{x}}(t)$, of the output $\hat{\boldsymbol{y}}(t)$ and of the model error $\hat{\boldsymbol{w}}(t)$. The latter is obtained by minimizing the error functional

$$J\left(\hat{\boldsymbol{x}}\left(t\right),\hat{\boldsymbol{w}}\left(t\right)\right)=\sum_{l=1}^{F}||\boldsymbol{y}(t_l)-\hat{\boldsymbol{y}}(t_l)||^2_{Q(t_l)}+\mathcal{R}[\hat{\boldsymbol{w}}]\,. \tag{6.4}$$

The first term in Equation (6.4) is the weighted mean square error between the measured outputs $\boldsymbol{y}(t_l)$ and the outputs $\hat{\boldsymbol{y}}(t_l)$ of the observer system in Equations (6.3a) and (6.3b). The weighted square norm

$$||\boldsymbol{y}(t_l)-\hat{\boldsymbol{y}}(t_l)||^2_{Q(t_l)}=\left(\boldsymbol{y}(t_l)-\hat{\boldsymbol{y}}(t_l)\right)^T Q(t_l)\left(\boldsymbol{y}(t_l)-\hat{\boldsymbol{y}}(t_l)\right) \tag{6.5}$$

contains the symmetric weighting matrix $Q(t_l)$, which is often chosen to be diagonal and can be used to transform outputs of very different magnitude to a common scale or to incorporate precision estimates of the measurements at the different time points t_l. The regularization term

$$\mathcal{R}[\hat{\boldsymbol{w}}] = \alpha_1||\hat{\boldsymbol{w}}||_1+\frac{\alpha_2}{2}||\hat{\boldsymbol{w}}||_2^2 \tag{6.6a}$$

$$||\hat{\boldsymbol{w}}||_1 = \int_{t_1}^{t_F}\sum_{n=1}^{N}|\hat{w}_n(t)|dt \tag{6.6b}$$

$$||\hat{\boldsymbol{w}}||_2^2 = \int_{t_1}^{t_F}\sum_{n=1}^{N}|\hat{w}_n(t)|^2\,dt. \tag{6.6c}$$

is necessary to avoid overfitting of the data $\boldsymbol{y}(t_l)$ by overly complex estimates $\hat{\boldsymbol{w}}(t)$. The nonnegative parameters α_1 and α_2 determine the relative contributions of the L_1 norm in Equation (6.6b) and of the L_2 norm in Equation (6.6c). Minimization of Equation (6.4) under the constraints in Equations (6.3a) and (6.3b) is an OCP, which needs to be solved numerically (see Appendix C) (Pontryagin et al., 1986; Fleming and Rishel, 1975; Gerdts, 2012). As detailed in Section 4.2, in contrast to common OCPs, the controls are split into known and unknown variables.

The combined L_1 - L_2 regularization in Equation (6.6a) is reminiscent of the elastic-net penalty used in regression (Zou and Hastie, 2005). Therefore, we termed our approach the dynamic elastic-net. In analogy to regression, the L_1 term causes some components $\hat{w}_n(t)$ of the estimated model error to shrink to zero (cp. Appendix C). The amount of shrinkage is determined by α_1, which can be chosen to suppress small error signals or noise distributed over many components of the estimate $\hat{\boldsymbol{w}}$. The resulting sparse estimate is useful because it provides information about the states of the system which are targeted by systematic model errors, as represented by hidden inputs.

In contrast to regression, a pure L_1 or least absolute shrinkage and selection operator (LASSO) type regularization is not useful in the dynamic setting because the solution for $\alpha_2 = 0$ can result in unbounded estimates of $\hat{\boldsymbol{w}}(t)$ (Tibshirani, 2011). Even when additional constraints on $\hat{\boldsymbol{w}}(t)$ are imposed, the resulting solution is not smooth and either zero or at the boundaries of the constraints (Vossen and Maurer, 2006). These insights about the OCP can be obtained from

Pontryagin's minimum principle, as detailed in Appendix C together with some strategies to chose suitable regularization parameters α_1 and α_2 (Pontryagin et al., 1986; Fleming and Rishel, 1975). In addition to sparse but smooth estimates of the model error, the DEN automatically provides a state estimate $\hat{\boldsymbol{x}}(t)$. Often this is very interesting information, especially when not all state variables are experimentally accessible.

The OCP in Equations (6.4) and (6.3) for $\hat{\boldsymbol{w}}(t)$ requires the specification of an initial condition $\hat{\boldsymbol{x}}(0) = \boldsymbol{x}_0$, which is often not known or uncertain. Alternatively and according to Section 4.2.4, one can add the additional constraint

$$\|\boldsymbol{y}(t_1) - \hat{\boldsymbol{y}}(t_1)\|_{Q(t_1)} \leq \Delta_{t_1} \tag{6.7}$$

to the OCP, where Δ_{t_1} is a preset tolerance given for the fit of $\hat{\boldsymbol{y}}(t_1)$ to $\boldsymbol{y}(t_1)$ at time t_1. Similarly, a tolerance Δ_1 can be prescribed to the fit at the last data point through

$$\|\boldsymbol{y}(t_F) - \hat{\boldsymbol{y}}(t_F)\|_{Q(t_F)} \leq \Delta_{t_F}. \tag{6.8}$$

The tolerance parameters Δ_{t_1} and Δ_{t_F} of these optional constraints can often be obtained from the error bars of the measurements.

6.3 Results

6.3.1 JAK-STAT Signaling

To illustrate the DEN estimator for a small and comprehensible model we used established experimental data for the JAK-STAT signal transduction pathway (Swameye et al., 2003). The four state variables of the system represent the unphosphorylated cytoplasmic signal transducer and activator of transcription 5 (STAT5) (x_1), the phosphorylated monomeric STAT5 (x_2), the phosphorylated dimeric STAT5 (x_3) and the nuclear dimeric STAT5 (x_4). The nominal model

$$\begin{pmatrix} \dot{\tilde{x}}_1 \\ \dot{\tilde{x}}_2 \\ \dot{\tilde{x}}_3 \\ \dot{\tilde{x}}_4 \end{pmatrix} = \underbrace{\begin{pmatrix} -\theta_1 \tilde{x}_1 u \\ \theta_1 \tilde{x}_1 u - 2\theta_2 \tilde{x}_2^2 \\ \theta_2 \tilde{x}_2^2 - \theta_3 \tilde{x}_3 \\ \theta_3 \tilde{x}_3 \end{pmatrix}}_{\tilde{\boldsymbol{f}}(\tilde{\boldsymbol{x}},u)} \tag{6.9a}$$

$$\begin{pmatrix} \tilde{y}_1 \\ \tilde{y}_2 \end{pmatrix} = \underbrace{\begin{pmatrix} \theta_4(\tilde{x}_2 + 2\tilde{x}_3) \\ \theta_5(\tilde{x}_1 + \tilde{x}_2 + 2\tilde{x}_3) \end{pmatrix}}_{\boldsymbol{h}(\tilde{\boldsymbol{x}})} \tag{6.9b}$$

describes the phosphorylation of cytoplasmic STAT5 upon activation of the EpoR (the known input u), the dimerization of phosphorylated STAT5 and the export to the nucleus (Figure 6.2) (Swameye et al., 2003). Time course data for the amount of cytoplasmic phosphorylated STAT5 (y_1) and total cytoplasmic STAT5 (y_2) were used to calibrate the parameters $\theta_1, \ldots, \theta_5$ (Swameye et al., 2003). However, the presence of a systematic model error is apparent from the inalterable discrepancy between the experimental data and the nominal model incorporating optimized parameter values (Figure 6.2b,c).

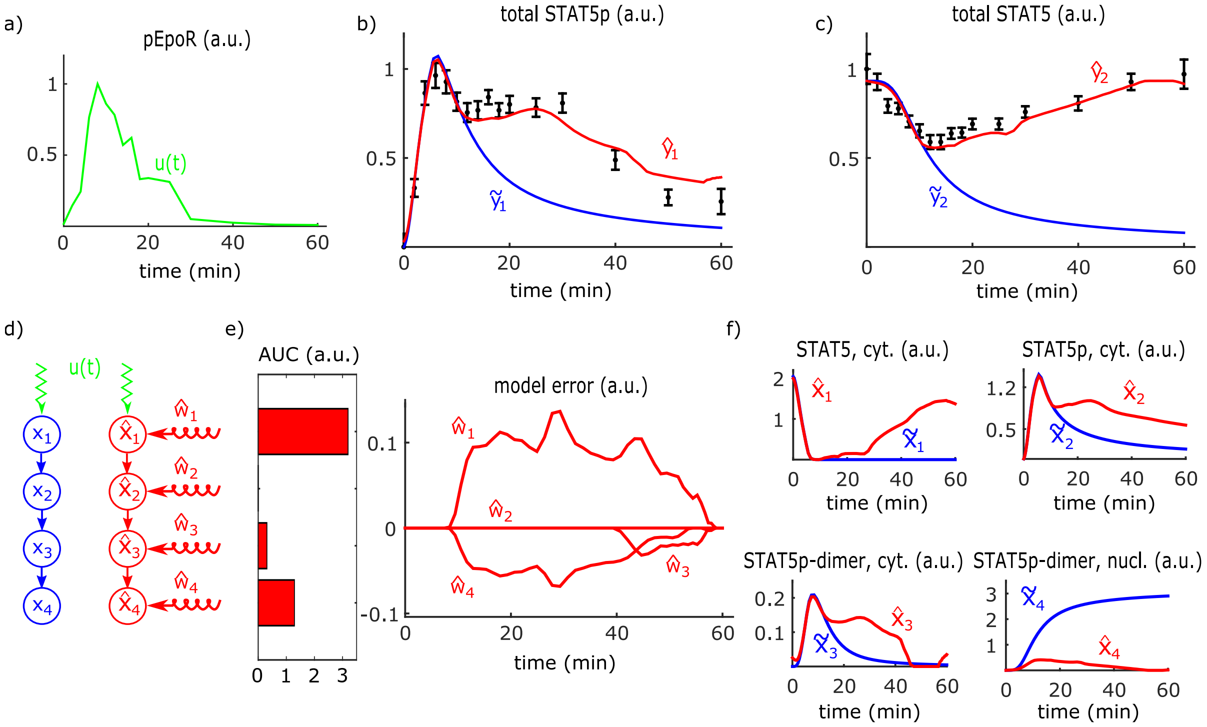


Figure 6.2: Estimating the model error for the JAK-STAT pathway. (a) The known input $u(t)$ is given by linearly interpolated phosphorylation measurements for the EpoR (Swameye et al., 2003). (b,c) Output measurements (Swameye et al., 2003) (black) for phosphorylated STAT5 (y_1) and total STAT5 (y_2) in the cytoplasm compared to the outputs of the nominal model (blue) and the fit of the DEN (red). (d) Graph of the nominal model (blue) and of the observer system (red) with the state variables cytoplasmic STAT5 (x_1), phosphorylated monomeric STAT5 (x_2), phosphorylated dimeric STAT5 (x_3) and nuclear STAT5 (x_4). (e) DEN estimates $\hat{\boldsymbol{w}}(t) = (\hat{w}_1(t), \ldots, \hat{w}_4(t))^T$ of the model error and the AUC for the magnitude of each component $|\hat{w}_n(t)|$. (f) State estimates $\tilde{x}_1, \ldots, \tilde{x}_4$ obtained from the nominal model (blue) and the DEN observer (see $\hat{x}_1, \ldots, \hat{x}_4$ in red).

To estimate this model error $\boldsymbol{w}(t)$, we numerically fitted the DEN with the nominal model (6.9) to the output measurements. To quantify the magnitude of the different components, we numerically computed the area under the curve (AUC) of each $|\hat{w}_n(t)|$, i.e., $\mathrm{AUC}_n = \int_{t_1}^{t_F} |\hat{w}_n(t)| dt$. The AUC and the estimated time course $\hat{\boldsymbol{w}}(t) = (\hat{w}_1(t), \ldots, \hat{w}_N(t))^T$ of the model error indicate (Figure 6.2e) that the dominant contributions $\hat{w}_1(t)$ and $\hat{w}_4(t)$ of the model error target the states x_1 and x_4, representing the amount of unphosphorylated cytoplasmic STAT5 and nuclear STAT5, respectively. The second component $\hat{w}_2(t)$ of the DEN estimate is identically zero for the whole time interval (Figure 6.2e). Apart from the small signal $\hat{w}_3(t)$ initiated after approximately 40 minutes, this is consistent with the improved nucleocytoplasmic cycling model, which is based on the same data and incorporates the relocation of orthophosphate nuclear STAT molecules into the cytoplasm (Swameye et al., 2003; Raue et al., 2009). Importantly, the DEN also provides modified estimates for the four STAT5 state variables (Figure 6.2f), which are also in good agreement with the nucleocytoplasmic cycling model (Appendix C).

An important problem with regularization approaches is the choice of the regularization parameters α_1 and α_2. We used $\alpha_1 = 10$ and $\alpha_2 = 0.2$ in Figure 6.2, but we found empirically that the AUC clearly indicate the target points of the model error for a wide range of α_1-values (Appendix C, Figure C.2). The L_2 parameter α_2 was chosen to balance the smoothness of $\hat{\boldsymbol{w}}$ and the accuracy of the fit to the output measurements. In addition, the bias induced by the double regularization can be compensated by a simple thresholding strategy (Zou and Hastie, 2005): Given an initial estimate $\hat{\boldsymbol{w}}(t) = (\hat{w}_1(t), \ldots, \hat{w}_N(t))^T$ of the model error, we refit the DEN by constraining all the components with a small AUC to zero. Thresholding is known in the regression context and we found it to improve the state estimates as well as the time course estimates of the remaining model errors (Appendix C, Figure C.3) (van de Geer et al., 2011).

To explore the robustness of the DEN against measurement noise, we added random perturbations to the experimental data (Swameye et al., 2003). For a given noise level, we generated 500 perturbed data sets by adding Gaussian random numbers with mean zero and standard deviation scaled by a multiple of the empirical standard deviation (see the error bars in Figure 6.2b,c) to each experimental data point. Thus, the noise level is defined as a multiple of the empirical standard deviation. The DEN was then fitted to each output sample, and the corresponding area under the curve for each component of the estimated model error $\hat{\boldsymbol{w}}(t)$ was computed. The plots for these AUC values curves versus the noise level are shown in Figure 6.3a. The median values of the AUC for the components $\hat{w}_1, \ldots, \hat{w}_4$ are largely independent of the noise level, but the variability of the AUC estimates increases with the measurement noise. Nevertheless, the AUC values for $\hat{w}_1$ and $\hat{w}_4$ are always much larger than zero, whereas the AUC of $\hat{w}_2$ and $\hat{w}_3$ is close or even equal to zero for many samples with higher noise level. This increases the confidence that the nodes $\tilde{x}_1$ and $\tilde{x}_4$ (Figure 6.2d) of the nominal JAK-STAT model are the main target points of the model error.

The impact of parameter uncertainty on the nominal model was assessed in a similar way. Parameter estimation algorithms applied to the nominal model using the experimental data (Figure 6.2b,c) provide point estimates and confidence intervals for each component of the parameter vector (Swameye et al., 2003; Raue et al., 2009; Balsa-Canto et al., 2010). These confidence intervals were again scaled by the noise level, yielding an interval for each parameter from which uniform random samples were drawn. Again, we generated 500 modified parameter vectors per noise level. For each parameter sample, the system was taken as the nominal model and the AUC of the resulting estimates $\hat{w}_1, \ldots, \hat{w}_4$ was recorded (Figure 6.3b). Again, there is no systematic trend for the AUC of the different components of the estimated error $\hat{\boldsymbol{w}}$. However, the variation of the AUC increases much faster than in Figure 6.3a. Apart from the different sampling distributions used, this effect is related to the definition of the model error $\boldsymbol{w}$, which is always defined with respect to the nominal model (confer Equation 6.2a). Hence, the estimated model error $\hat{\boldsymbol{w}}$ contains contributions from both structural and parameter misspecifications in the nominal model. Nevertheless, it is still possible to infer the dominant components $\hat{w}_1$ and $\hat{w}_4$ with high confidence. Similar results were found for the sensitivity against the number of measurement time points (Appendix C, Figure C.6).

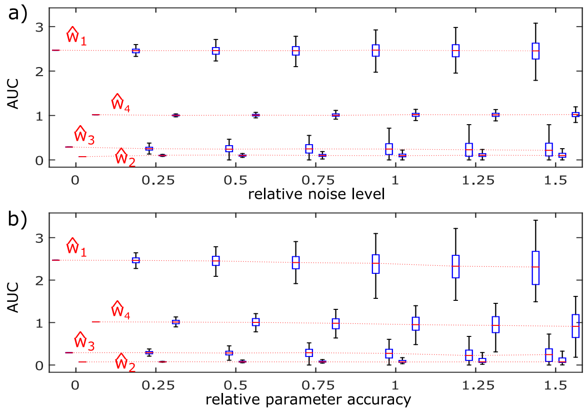


Figure 6.3: Impact of simulated measurement noise and parameter uncertainty on the DEN estimate in the JAK-STAT model. (a) Box plots visualizing the variation of the AUC of $|\hat{w}_1(t)|, \ldots, |\hat{w}_4(t)|$ for the DEN estimates caused by measurement noise (see main text for details). For better visualization, the box plots at a given noise level are slightly offset. (b) Variation of the AUC caused by parameter uncertainty.

6.3.2 Photomorphogenic UV-B Signaling

As a test case for a larger system, we used a recent model for the coordination of photomorphogenic UV-B signaling in plants (Ouyang et al., 2014). The model consists of 11 ODEs describing the dynamics of protein concentrations $\tilde{\boldsymbol{x}} = (\tilde{x}_1, \ldots, \tilde{x}_{11})^T$ coupled by 10 chemical reactions (Figure 6.4). We considered this model as the nominal model in order to test the dynamic elastic-net method for a situation where the ground truth is known. The model error was simulated by adding the hidden inputs $w_3(t) = w_9(t) = 1000 \cdot [1 - 1/(1 + t)]$ to the nodes $\tilde{x}_3$ and $\tilde{x}_9$. The output function $\boldsymbol{h}(\tilde{\boldsymbol{x}})$ is a linear combination of 7 different state variables (see Appendix C for all equations). The synthetic data was sampled at discrete time points from the outputs of the true model and Gaussian random perturbations were added to simulate measurement noise (Figure 6.4b-f). The DEN with the nominal model was used to reconstruct the model error $\boldsymbol{w}(t)$ and the true state $\boldsymbol{x}(t)$ from the simulated data. The absolute area under the curve for each component of the model error estimate $\hat{\boldsymbol{w}}(t)$ clearly indicates that the states $\tilde{x}_3$ and $\tilde{x}_9$ are targeted by hidden inputs (Figure 6.4g), whereas all other components are either very small ($|\hat{w}_6|$) or even zero. This illustrates the sparsity of the DEN estimate, which is a clear advantage over pure L_2 regularization The discrepancy $\boldsymbol{w}(t) - \hat{\boldsymbol{w}}(t)$ between the model error and the corresponding estimate relative to the amplitude $A = \max_{t_l \in (t_1, \ldots, t_F)} w_{3,9}(t_l)$ of the true model error is at most 10% (Figure 6.4h) and mainly caused by numerical inaccuracies. Most importantly, the discrepancy $\boldsymbol{x}(t) - \hat{\boldsymbol{x}}(t)$ between the true and the estimated state trajectory is almost zero (Figure 6.4i), indicating the excellent performance of the DEN as a state observer.

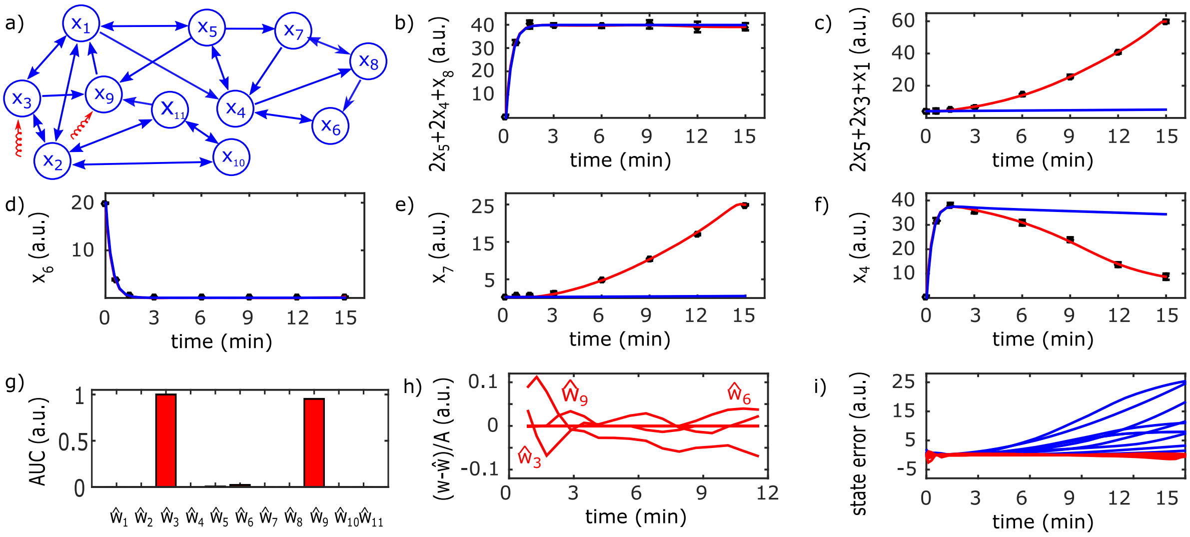
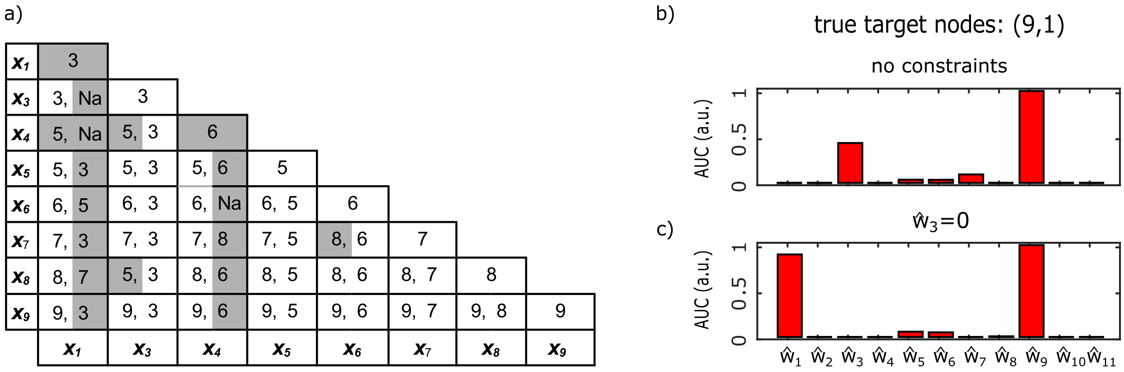


Figure 6.4: Application of the DEN to the modeling of the photomorphogenic UV-B signaling. (a) Graph (without self loops) of the model states (Ouyang et al., 2014). The target points of the simulated model errors are indicated by the red arrows. (b-f) Simulated output $y_1, \ldots, y_5$ with error bars (black), output of the nominal model (blue) and output of the DEN (red). (g) AUC of the absolute model errors $|\hat{w}_1(t)|, \ldots, |\hat{w}_{11}(t)|$. (h) Components of $\boldsymbol{w} - \hat{\boldsymbol{w}}$ relative to the amplitude A of the true model error. (i) Discrepancy $\boldsymbol{x}(t) - \tilde{\boldsymbol{x}}(t)$ between the true state and the nominal model state (blue) compared to the discrepancy $\boldsymbol{x}(t) - \hat{\boldsymbol{x}}(t)$ of the DEN (red).

As for any inverse method, there are limitations to the dynamic elastic-net method. Some model errors $\boldsymbol{w}(t)$ are unobservable because there exists a different hidden input function $\boldsymbol{w}^{\dagger}(t)$ which generates an output $\boldsymbol{y}(t)$ identical to the output obtained for $\boldsymbol{w}(t)$, see Appendix C for a simple example. Other model errors might be practically unobservable because the output for another hidden input function might not be distinguishable within the measurement errors. A special case are model errors which have no or almost no effect on the output at all. These will not be noticed during modeling and the nominal model will be accepted.

To further test the ability of the DEN to infer the states targeted by the model error, i.e., the non-zero components of the true model error $\boldsymbol{w}(t)$, we systematically simulated perturbations to different nodes and node pairs. First, we simulated model errors $w_n(t) = 1000 \cdot [1 - 1/(1+t)]$, $w_j(t) = 0$ for $j \neq n$ targeting a single node n in the same way as before. For the nodes $n = 2, n = 10$ and $n = 11$, there was no effect on the output (cp. Figure 6.4b-f) and thus these nodes were omitted from further analysis. In addition, we simulated hidden inputs for all remaining two-node combinations. For each of these 36 simulated true models we tested the ability of the DEN to recover the correct target nodes from the AUC of the estimated $|\hat{w}_n(t)|$. We considered a node or a node pair to be correctly recovered if their AUC was at least 85% of the total AUC over all nodes. Using this stringent criterion, we found that two single node errors targeting $\tilde{x}_1$ or $\tilde{x}_4$ were not correctly detected and another single node was predicted to be the target of the model error (Figure 6.5a). This indicates that these model errors are unobservable and that the observed output data can be explained by different inputs to different nodes. With two exceptions ((8, 3) and (7, 6)), the mistakes made by the algorithm for simulated pairwise model errors involve these two state nodes 1 and 4. However, with exception of the combination (1, 4), at least one node is correctly predicted.



	x_1	x_3	x_4	x_5	x_6	x_7	x_8	x_9
x_1	3							
x_3	3, Na	3						
x_4	5, Na	5, 3	6					
x_5	5, 3	5, 3	5, 6	5				
x_6	6, 5	6, 3	6, Na	6, 5	6			
x_7	7, 3	7, 3	7, 8	7, 5	8, 6	7		
x_8	8, 7	5, 3	8, 6	8, 5	8, 6	8, 7	8	
x_9	9, 3	9, 3	9, 6	9, 5	9, 6	9, 7	9, 8	9

Figure 6.5: Detection of the target nodes of the simulated model errors in the UV-B signaling network. (a) All nodes and all pairs of nodes were perturbed by a simulated model error. Nodes x_2, x_{10} and x_{11} are omitted, since the simulated error signal had no effect on the output. The rows and the columns correspond to the true target nodes of the model error and the numbers in the cells are the nodes found by the DEN (NA means that no second node was assigned). Gray cells indicate errors made by the DEN for unobservable model errors. (b) An example for an unobservable model error. The true target nodes of the model error are (9, 1), but the DEN predicts the target nodes (9, 3). (c) Refitting the DEN under the constraint $\hat{w}_3 = 0$ provides an alternative solution. The other two combinations $\hat{w}_9 = 0$ and $\hat{w}_3 = \hat{w}_9 = 0$ of the nodes (9, 3) did not fit the output data.

These results demonstrate the inherent limitations of any attempt to recover the model error from observed outputs. For an unobservable model error, the true model error $\boldsymbol{w}(t)$ might correspond to a slightly larger value of the error functional given through Equation (6.4) than the minimum $\hat{\boldsymbol{w}}(t)$ obtained through the DEN. A heuristic approach to explore some of these slightly suboptimal solutions is to rerun the dynamic elastic-net method with some of the estimated target nodes (from the first run) excluded and to check whether the output data can satisfactorily be fitted with the same level of sparsity. This is illustrated in Figure 6.5b for the node pair (9, 1), which was predicted to be (9, 3) by our criterion. Refitting the DEN under the constraint $\hat{w}_3 = 0$ identifies the correct nodes (9, 1), as illustrated in Figure 6.5c. The two other combinations $\hat{w}_9 = 0$ and $\hat{w}_3 = \hat{w}_9 = 0$ do not provide a satisfactory fit to the data (Appendix C, Figure C.8 and Figure C.9). For the UV-B signaling network, we find that the slightly suboptimal solutions identified by this heuristics always contain the correct target node configuration. The combinatorial explosion of this strategy should typically not be a problem, thanks to the sparsity of the DEN predictions. The decision, which of the predicted target node sets, (9, 3) or (9, 1), is the correct one can in practice only be made when additional states are measured. However, this example shows, how the DEN provides useful information to select further states for experimental observation (Raue et al., 2010; Liu et al., 2013).

6.4 Conclusion

Efficient computational methods to learn from incomplete model drafts and to direct model improvement are urgently needed. The proposed dynamic elastic-net approach provides suggestions for the location of these model errors in the network and estimates their dynamic time courses from measured output data. The sparsity of the proposed target points for the model error promotes model improvements in the most parsimonious way. Even for an incomplete nominal model the algorithm can provide estimates for the system states which are not experimentally accessible. This is in stark contrast to many other state estimators including the Kalman filter for linear systems and its various extensions for nonlinear systems, which usually require a correctly specified model (Kalman, 1960; Julier and Uhlmann, 1997; Crassidis and Junkins, 2011).

Not all model errors can uniquely be determined from the output. For such unobservable model errors, our strategy to explore alternative, slightly suboptimal solutions might indicate alternative explanations for observed discrepancies between the data and the nominal model. In addition, this approach can also be informative for selecting additional nodes required for observing the state from output measurements (Raue et al., 2010; Liu et al., 2013). Further research is needed to establish the relationship between the network topology and the observability of a model error.

Model errors arising in kinetic reaction systems can originate from erroneous rate equations or lacking reactions. The DEN can detect both types of errors as hidden inputs to the corresponding nodes of the network, but it can not discriminate between these errors. However, knowing the nodes affected by a model error might already be very informative for systematic model improvement. This work raises fundamental questions regarding successful modeling strategies. The approach to manually include more and more details into the model to compensate for the initial model errors is often not practical or at least very time consuming. The dynamic elastic-net hence paves the way towards a more principled and systematic way, in which models could be adapted based on experimental data.

CHAPTER 7

The Bayesian Dynamic Elastic-Net

In Chapter 6 I proposed the DEN as a promising approach to make ODE-based biological models, e.g., the developed model of the M_2 receptor-induced signaling, complete and consistent. The BDEN, as a Bayesian extension of the DEN, incorporates measurement uncertainties and allows for automatic correction of the nominal model by adding or removing single interactions. This results in more robust and reliable estimations of the involved hidden influences. This work is under review in a peer-reviewed journal (cp. Appendix F).

7.1 Introduction

The proposed DEN aims for estimating the dynamics of exogenous hidden influences, i.e., exogenous hidden inputs, and misspecified molecular interactions, i.e., endogenous hidden inputs to the nominal model via a penalized estimation procedure resembling elastic-net regression (Zou and Hastie, 2005; Engelhardt et al., 2016). This is strongly required since in most situations researchers have partial knowledge and preliminary hypotheses about their system, which needs to be integrated into a restricted but still predictive and experimentally validatable model (Swameye et al., 2003). Even, if the biological system is partially known and the data is given for almost all molecular species it is not clear how to deal with insufficient predictions (Azeloglu and Iyengar, 2015).

As discussed in Chapter 6, the DEN addresses these issues in a suitable manner. Although the method was tested successfully on several applications, such as the EpoR-dependent signaling network (Swameye et al., 2003), it has still several shortcomings, which are addressed in this chapter. More specifically, the DEN is not a probabilistic approach and thus does not fully address the unavoidable uncertainty about estimates. Moreover, the DEN does not answer the question whether estimated hidden inputs could be attributed to missed or wrongly modeled interactions among the known molecular species. Hence, here the Bayesian DEN (BDEN) is introduced as a new and fully probabilistic approach, which deals with all these aspects.

In contrast to the DEN, the BDEN method does not require prespecified hyperparameters. The predictive power of the BDEN compared to the DEN is illustrated in several real biological models and test cases. The BDEN thus provides a systematic Bayesian computational method to identify target nodes and reconstruct the corresponding error signal including detection of missing and wrong molecular interactions within the assumed model. The method works for ODE-based systems even with uncertain knowledge and noisy data. In contrast to approaches based on point estimates the Bayesian framework incorporates the given uncertainty and circumvents numerical pitfalls which frequently arise from optimization methods (Betts, 2009; Rao, 2010).

7.2 Methods

7.2.1 Motivation

The modeling process is assumed to start with an initial, potentially incomplete or partially misspecified nominal model including all known but not necessarily observable molecules (Liu et al., 2013). Figure 7.1 illustrates the general idea of automatic model correction. As discussed in Chapter 6, in most situations the real system differs from the initially modeled nominal system, which is reflected by an insufficient fit to the given data caused by (i) hidden influences and (ii) erroneous molecular interactions. Exogenous hidden influences could, for example, be stimulatory (e.g., phosphorylation) or inhibitory (e.g., dephosphorylation) events affecting the modeled system from outside. In addition, there could exist stimulatory or inhibitory influences within the system, which are not included in the model due to lack of knowledge, i.e., missing molecular interactions. Similarly, wrongly included molecular interactions could exist. In biochemical systems molecular interactions are based on biochemical reactions, e.g., phosphorylation and binding events.

Due to the fact that biological systems are open the number of potential erroneous nodes (e.g., proteins or other molecules) within the nominal model is huge (Babtie et al., 2014). Without further knowledge, independent error terms have to be assigned to each node. If the respective node is in reality not directly targeted by a hidden influence, the hidden input takes the value zero. Only nodes directly affected by hidden influences have non-zero errors. Wrongly modeled or missing interactions between two nodes can be represented by two error terms, one for each of the respective nodes, which will be correlated over time. This idea is exploited to detect missing or erroneous interactions in a given ODE-based nominal model.

7.2.2 Approach

The assumed dynamical model is given by

$$\dot{\boldsymbol{x}}(t) = \boldsymbol{f}(\boldsymbol{x}(t), \boldsymbol{u}(t)) + \boldsymbol{w}(t) \tag{7.1a}$$

$$\boldsymbol{y}(t) = \boldsymbol{h}(\boldsymbol{x}(t)) + \boldsymbol{\epsilon}(t) \tag{7.1b}$$

$$\boldsymbol{x}(0) = \boldsymbol{\eta}, \tag{7.1c}$$

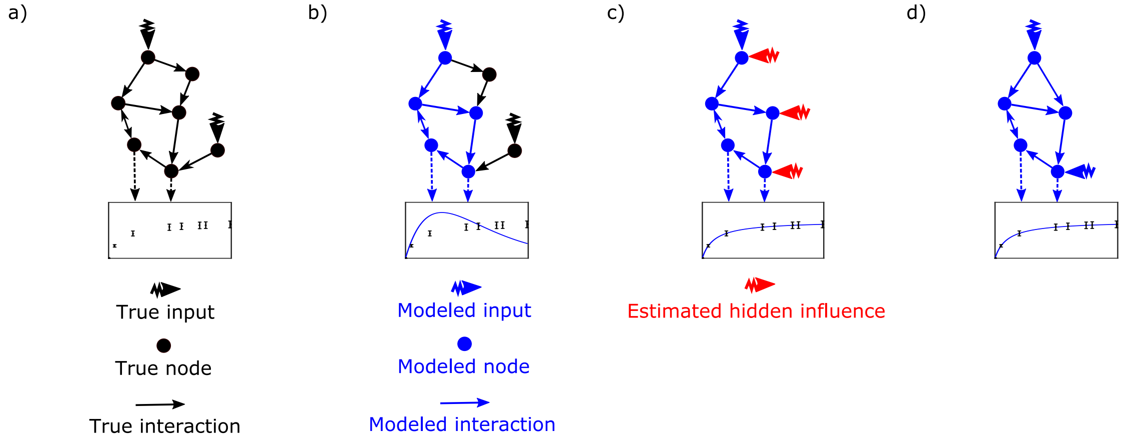


Figure 7.1: Illustration of the Bayesian approach to estimate hidden influence variables in ODE-based models. Here, a fictional interaction network is shown. (a) True system (black) with two inputs and one fictive observable measurement as a combination of two nodes (box). (b) In reality, the available knowledge represented by an a priori model (blue) does not necessarily cover the whole system but only a part of the true system. Hence the nominal model leads to an unsatisfactory fit with the observable measurement. This may be caused by exogenous influences or by misspecified molecular interactions (i.e., missing or wrong edges in the interaction network). (c) Our approach aims for estimating these hidden influences (red) and the directly involved molecular species. (d) Some of the estimated hidden influences may correspond to missing or wrong molecular interactions within the system. Hence, in a last step our method tries to further distinguish between intrinsic and exogenous hidden influences. We therefore identify erroneous parts of the nominal ODE system and give detailed hints for their correction.

according to the proposed observer system described through Equations (6.3). As discussed in Chapter 6.2.3, $\dot{\boldsymbol{x}}$ denotes the time derivative of the state vector $\boldsymbol{x}(t) = (x_1(t), \ldots, x_N(t))^T$ with initial value $\boldsymbol{\eta}$. The not necessarily linear function $\boldsymbol{f}$ represents the nominal model, which describes the current assumptions about the dynamic interactions between the state variables and in addition the effect of the known input function $\boldsymbol{u}(t)$. The additive dynamic hidden influence $\boldsymbol{w}(t)$ subsumes missing or wrong interactions between the state variables as well as exogenous influences caused by crosstalk with other biological processes (cp. Appendix D). This is equivalent to the hidden input signals as defined in Chapter 6. Of course, $\boldsymbol{w}(t)$ is unknown and intended to be estimated from the data (Engelhardt et al., 2016). Notably, the hidden influence $\boldsymbol{w}(t)$ is not restricted to be constant or linear and thus can be any arbitrary function of time.

The map from the state to the measurable output $\boldsymbol{y}(t) = (y_1(t), \ldots, y_I(t))^T$, with I not necessarily equal to N, is given by the measurement function $\boldsymbol{h}$, which is assumed to be known (Equation (7.1b)). In addition white Gaussian measurement noise $\boldsymbol{\epsilon}(t)$ with expectation zero and a noise covariance matrix $\Xi_l \in \mathbb{R}^{I \times I}$ is assumed, see below. According to Section 4.3, in practice the data is given as measurements $\boldsymbol{y}(t_l)$ at discrete time points t_l with $l \in \{1, \ldots, F\}$. The notation $y_{i,l} = y_i(t_l)$ is used for the measured output $i \in \{1, \ldots, I\}$ at measurement time t_l and the analogous notation for the other variables, i.e., $x_n(t_l) = x_{n,l}$ and $h_i(\boldsymbol{x}(t_l)) = h_i(\boldsymbol{x}_l)$. For sake of simplicity in the following the matrix of observed measurements is represented by $\boldsymbol{y}_{1:F} = \{\boldsymbol{y}_1, \ldots, \boldsymbol{y}_F\} \in \mathbb{R}^{I \times F}$ and the corresponding state and hidden influence matrices by

$\boldsymbol{x}_{1:F} \in \mathbb{R}^{N\times F}$ and $\boldsymbol{w}_{1:F} \in \mathbb{R}^{N\times F}$ with initial values $\boldsymbol{x}_0 = \boldsymbol{\eta}$ and $\boldsymbol{w}_0 = \boldsymbol{0}$.

From now on we are interested in hidden influences at discrete time points. Under this assumption Equations (7.1) can be rewritten as:

$$\boldsymbol{x}_l = \boldsymbol{x}_{l-1} + \int_{t_{l-1}}^{t_l} \boldsymbol{f}(\boldsymbol{x}(\tilde{t}), \boldsymbol{u}(\tilde{t})) + \hat{\boldsymbol{w}}(\tilde{t})\, d\tilde{t} \tag{7.2a}$$

$$\boldsymbol{y}_l = N(\boldsymbol{h}(\boldsymbol{x}_l), \Xi_l) \tag{7.2b}$$

$$\boldsymbol{x}_0 = \boldsymbol{\eta}. \tag{7.2c}$$

Consequently, we obtain a first order Markov process over the state variables $\boldsymbol{x}$. The function $\hat{\boldsymbol{w}}(t)$ is obtained by fitting a cubic smoothing spline through each of the N discrete time series of hidden influence signals $\boldsymbol{w}_F$ with an initial value $\boldsymbol{w}_0$ (Fritsch and Carlson, 1980). The assumption of Gaussian measurement noise can, if necessary, approximately be fulfilled after a variance-stabilizing transformation (Atkinson, 1987). In addition, $\Xi_l = \mathrm{diag}\left(\xi_{1,l}^2, \ldots, \xi_{I,l}^2\right)$ is supposed.

7.2.3 Marginal Likelihood of the Data

The likelihood of the observed data

$$p(\boldsymbol{y}_{1:F} | \boldsymbol{x}_{0:F}, \boldsymbol{w}_{0:F}, \Xi_{1:F}) = \prod_{l=1}^{F} \prod_{i=1}^{I} p\left(y_{i,l} \middle| \boldsymbol{x}_l, \xi_{i,l}^2\right) \times p(\boldsymbol{x}_l | \boldsymbol{x}_{l-1}, \boldsymbol{w}_{l-1:l}). \tag{7.3}$$

can be factorized due to the independence of the measurement noise with respect to time and observables. Note that $p(\boldsymbol{x}_l | \boldsymbol{x}_{l-1}, \boldsymbol{w}_{l-1:l})$ is defined by Equation (7.2a). In addition $\boldsymbol{x}_{l-1:l}$ and $\boldsymbol{w}_{l-1:l}$ are conditionally independent from $\Xi_{1:F}$. Since typically the number of replicate measurements per time point is small, the empirical variance is not a reliable estimator of the true measurement noise. Therefore, we impose an inverse gamma prior on the variance of the measurement noise

$$\xi_{i,l}^2 \sim IG(\alpha, \beta). \tag{7.4}$$

The marginal likelihood of the data is obtained by marginalizing over the variance of the measurement noise variable

$$p(\boldsymbol{y}_{1:F} | \boldsymbol{x}_{0:F}, \boldsymbol{w}_{0:F}, \alpha, \beta) \propto \prod_{l=1}^{F} \prod_{i=1}^{I} \int p\left(y_{i,l} \middle| \boldsymbol{x}_l, \xi_{i,l}^2\right) \times p\left(\xi_{i,l}^2 \middle| \alpha, \beta\right) d\xi_{i,l}^2 \times p(\boldsymbol{x}_l | \boldsymbol{x}_{l-1}, \boldsymbol{w}_{l-1:l}). \tag{7.5}$$

This integral can analytically be calculated to yield (Zacher et al., 2012)

$$p(\boldsymbol{y}_{1:F} | \boldsymbol{x}_{0:F}, \boldsymbol{w}_{0:F}, \alpha, \beta) \propto \prod_{l=1}^{F} \prod_{i=1}^{I} \underbrace{\frac{\Gamma\left(\alpha + \frac{1}{2}\right)}{\Gamma(\alpha)(2\pi\beta)^{\frac{1}{2}}} \frac{1}{\left(1 + \frac{1}{2\beta}(y_{i,l} - h_i(\boldsymbol{x}_l))^2\right)^{\alpha + \frac{1}{2}}}}_{p(\mathbf{y}_l|\mathbf{x}_l,\alpha,\beta) \times p(\mathbf{x}_l|\mathbf{x}_{l-1},\mathbf{w}_{l-1:l})}. \tag{7.6}$$

A detailed derivation of Equation (7.6) is provided in Appendix D.

According to Bayes' theorem the posterior density over the hidden input signals $\boldsymbol{w}_{1:F}$ is given by:

$$p\left(\boldsymbol{w}_{1:F} \left| \boldsymbol{x}_{0:F}, \boldsymbol{y}_{1:F}, \Xi_{1:F}\right.\right) \propto p\left(\boldsymbol{y}_{1:F} \left| \boldsymbol{x}_{0:F}, \boldsymbol{w}_{0:F}, \Xi_{1:F}\right.\right) \times p\left(\boldsymbol{w}_{1:F}\right). \tag{7.7}$$

Using Equation (7.6) we can directly draw samples from the posterior density of the hidden influence. For this purpose we propose a Bayesian elastic-net prior as detailed in the following sections.

7.2.4 Smoothness and Sparsity via a Bayesian Elastic Net-Prior

The hidden input signals $\boldsymbol{w}_{l-1:l}$ can be understood as the statistical residuals of the nominal system, and every deviation of observations from the nominal system could thus be explained by non-zero components in $\boldsymbol{w}_{l-1:l}$. However, we are only interested in hidden input signals, which are far stronger than measurement noise. Therefore, we assume that the hidden input signal is smooth and sparse. Sparsity corresponds to the a priori belief that only a small subset of state variables is truly affected by unknown external or internal input signals. In addition we assume the hidden input signal to be smooth over time. Smoothness and sparsity are encoded by a prior distribution inspired by the Bayesian elastic-net, which is here combined with a first-order Markov process over $\boldsymbol{w}_{1:F}$ (Li and Lin, 2010; Friedman et al., 1998). Overall our proposed approach is thus a hierarchical graphical model shown in Figure 7.2. Details can be found in Appendix D.

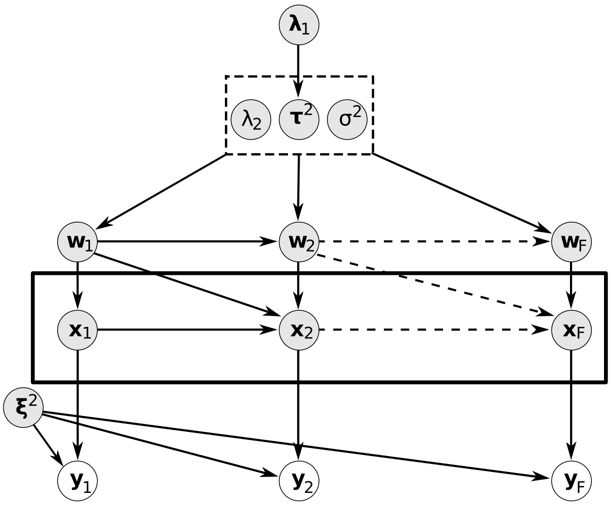


Figure 7.2: Representation of the proposed Bayesian Dynamic Elastic-Net approach as a probabilistic graphical model. The hidden influences $\boldsymbol{w}_l$ form a Markov chain over all time points $l = 1, \ldots, F$ and are directly dependent on the shared parameters $\boldsymbol{\lambda}_1$ and λ_2. Since the outcome of one integration step represents the initial value for the next integration step the system state variables $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_F$ are also successively dependent.

Briefly, the Bayesian elastic-net defines a conditional Gaussian prior over each $w_{n,l} \left| w_{n,l-1} \right.$ ($n = 1, \ldots, N$, $l = 1, \ldots, F$). The scale of the variance of the Gaussian prior is a strongly decaying and smooth distribution peaking at zero, which depends on parameters λ_2, $\boldsymbol{\tau}^2$ and σ^2. The parameter $\boldsymbol{\tau}^2$ is itself given an by exponential distribution (one for each component of vector $\boldsymbol{w}_n$) with parameters $\boldsymbol{\lambda}_1$. In consequence, sparsity is dependent on the parameter vector $\boldsymbol{\lambda}_1$, whereas smoothness is mainly controlled by λ_2 (Zou and Zhang, 2009; Li and Lin, 2010). These parameters are drawn from hyperpriors, which can be set in a non-informative manner or with respect to prior knowledge about the degree of shrinkage and smoothness of the hidden influences (Kyung et al., 2010). Further details are given in Appendix D.

7.2.5 Estimating Hidden Influences from Data

To estimate the hidden input and the parameters in the hierarchical model, we devised a Metropolis-Hastings algorithm with Gibbs updates of the Bayesian elastic-net hyperparameters. The algorithm proceeds sequentially between the different time points $1, \ldots, F$ by drawing different samples at each supporting point. At sampling step $s + 1$ and time point l a random component $w_{n,l}$ is selected of the hidden input vector (a node in the network) at the previous time point, which is modified by a sample from a univariate Gaussian transition kernel π (Brooks, 1998). The resulting vector $\boldsymbol{w}_l^*$ is accepted with probability

$$\boldsymbol{r}\left(\boldsymbol{w}^{(s)}{}_l \left| \boldsymbol{w}_l^* \right.\right) = \min\left\{1, \frac{p\left(\boldsymbol{y}_l \left| \boldsymbol{x}_l, \alpha, \beta\right.\right)}{p\left(\boldsymbol{y}_l \left| \boldsymbol{x}_l, \alpha, \beta\right.\right)} \times \frac{p\left(\boldsymbol{x}_l \left| \boldsymbol{x}_{l-1}, \boldsymbol{w}_l^*, \boldsymbol{w}_{l-1}\right.\right)}{p\left(\boldsymbol{x}_l \left| \boldsymbol{x}_{l-1}, \boldsymbol{w}_l^s, \boldsymbol{w}_{l-1}\right.\right)} \times \frac{p\left(\boldsymbol{w}_l^* \left| \boldsymbol{\theta}\right.\right)}{p\left(\boldsymbol{w}_l^s \left| \boldsymbol{\theta}\right.\right)}\right\}, \tag{7.8}$$

where $p\left(\boldsymbol{w}_l \left| \theta\right.\right)$ is the Bayesian elastic-net prior over the hidden influences conditioned by hyperparameters $\boldsymbol{\theta} = \left\{\boldsymbol{\lambda}_1, \lambda_2, \boldsymbol{\tau}^2, \sigma^2\right\}$. Because of the Gaussian measurement errors, the discrepancy between data component $y_{i,l}$ and the corresponding model output $h_i(\boldsymbol{x}_l)$ in Equation (7.6) is given by the quadratic error $(y_{i,l} - h_i(\boldsymbol{x}_l))^2$. Note that $\boldsymbol{x}_l$ is obtained by numerically integrating the ODE system from time point t_{l-1} using $\boldsymbol{x}_{l-1}$ as initial value according to Equation (7.2a). The code for the sampling algorithm and further details are provided in Appendix D.

7.2.6 Estimating Endogenous Hidden Influences

After having estimated the hidden influences on the state components in the nominal ODE system, the question arises whether these hidden variables could in fact correspond to missing or wrongly specified interactions within the nominal system. A simple strategy which we followed here, is to rank all state variables in the nominal system by their temporal correlation with the estimated hidden influences. The essential idea is that in case of a wrong or missing reaction the estimated hidden time courses should "compensate" erroneous predictions by the nominal system (Figure 7.3). In general, wrong or missing interactions can either have an increasing ("stimulatory") or decreasing ("inhibitory") influence on the target nodes. This results in a negative hidden influence with $\boldsymbol{w}_{n,1:F} < 0$ in the case of inhibition and a positive hidden influence with $\boldsymbol{w}_{n,1:F} > 0$ in the case of stimulatory events. In general, wrong or missing interactions can either have an increasing ("stimulatory") or decreasing ("inhibitory") influences on the target nodes. This results in a negative hidden influence with $\boldsymbol{w}_{n,1:F} < 0$ in the case of inhibition and to a positive hidden influence with $\boldsymbol{w}_{n,1:F} > 0$ in the case of stimulatory events. More specifically, several cases are distinguished as listed in Table 7.1 and further illustrated in Figure 7.3. Briefly, the idea is that an unmodeled stimulation between two state variables x_1, x_4 in the ODE system yields two error signals w_1 (influencing x_1) and w_4 (influencing x_4), which are anticorrelated. This is because w_1 and w_4 capture the unmodeled dynamics of the ODE system. Similarly, an unmodeled inhibition yields $\boldsymbol{w}_{1,1:F}$, $\boldsymbol{w}_{4,1:F} < 0$ and a positive correlation of w_1 and w_4. A wrongly modeled stimulation results in $\boldsymbol{w}_{1,1:F} < 0$ and $\boldsymbol{w}_{4,1:F} > 0$ which are anticorrelated. A wrongly modeled inhibition yields $\boldsymbol{w}_{1,1:F}$, $\boldsymbol{w}_{4,1:F} > 0$ and a positive correlation.

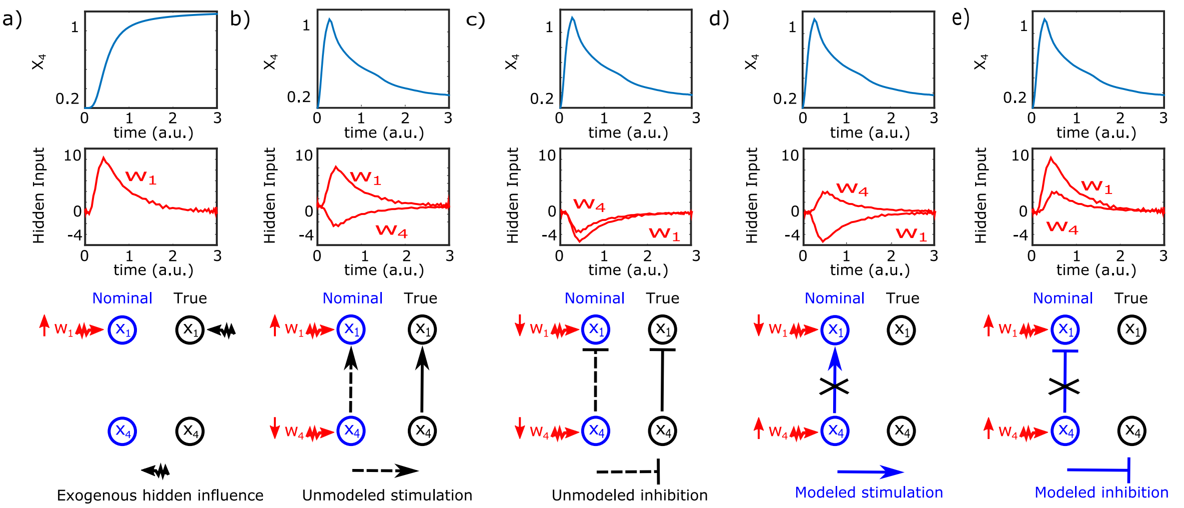


Figure 7.3: Illustration of hidden exogenous and endogenous influences by an arbitrary example system. (a) Hidden exogenous influence. No significant temporal correlation between x_4 and w_1 is expected. (b) Hidden endogenous influence as a missing stimulatory interaction (arrow) from x_4 to x_1. Here, hidden influences w_4 and w_1 are highly negatively correlated. This is caused by a missing stimulating effect of x_4 on x_1. The decrease of x_4 is correlated with an increase of w_1. (c) Hidden endogenous influence as a missing inhibitory interaction. Here, hidden influences w_4 and w_1 have a strong positive correlation and compensate a missing inhibitory effect of x_4 on x_1. The increase of x_4 is correlated with an decrease of w_1. (d) Hidden endogenous influence as an erroneous stimulation. Here, hidden influences w_4 and w_1 have a strong negative correlation. The increase of x_4 goes along with a decrease of w_1. (e) Hidden endogenous influence as an erroneous inhibition. The hidden influences w_1 and w_4 are concordant and correlate strongly with the state component x_4.

As exemplified in Figure 7.3, due to differences in the expected correlations the analysis should also allow for distinguishing missing inhibitory versus stimulating effects of missing interactions given monotonous interactions.

	Event	Corr. HI	Corr. State
Missing	stim.	-	+
	inh.	+	-
Wrong	stim.	-	-
	inh.	+	+

Table 7.1: Characterization of endogenous hidden influences. An endogenous hidden influence is expected to yield different types of (cross-)correlations with other hidden influence (Corr. HI) and state variable dynamics (Corr. State), depending on whether a molecular interaction is missing or wrongly specified in the nominal system. Missing and wrong molecular interactions can be further distinguished depending on whether the true molecular interaction is of stimulatory (stim.) or inhibitory (inh.) nature. Furthermore, molecular interactions between two species can either be modeled in the nominal system (wrong) or not (missing). Here the correlation of related hidden influence dynamics (Corr. HI) and the correlation between the target hidden influence dynamics and the related estimated model variables (Corr. State) is in focus. In case of endogenous hidden influences these correlations are expected to be either strongly positive (+) or strongly negative (-) if they reflect true molecular interactions. Figure 7.3 illustrates all four cases in detail.

Different measures exist to capture the strength of correlation between time courses. Apart from the Pearson correlation, we here used the cross-correlation coefficient

$$R_{XY}(\tau) = \sum_{(l=1)}^{(F-\tau-1)} X_{(l+\tau)} Y_l \tag{7.9}$$

to quantify temporal associations between two time series X and Y (Stoica and Moses, 2005). The cross-correlation $R_{XY}(\tau)$ depends on the time lag τ which was chosen as $\underset{\tau}{\operatorname{argmax}}\,[R_{XY}(\tau)]$.

7.3 Results

7.3.1 Tested Mathematical Models

The EpoR-induced JAK-STAT signaling pathway mediates a rapid signal transduction from the receptor to the nucleus related to cell proliferation and differentiation (Swameye et al., 2003). This pathway involves a rapid nucleocytoplasmic cycling of the STAT5 molecules which is not directly measurable (Swameye et al., 2003).

The GP cycling model quantitatively characterizes the heterotrimeric GP activation and deactivation in yeast (Yi et al., 2003). This model serves as a fully observed but complex test case where all states are measured.

In contrast, the model of the UV-B signaling in plants systematically links several signaling events induced by UV-B light to a comprehensive informational signaling pathway (Ouyang et al., 2014). Only combinations of small amounts of the involved molecules are accessible and thus it serves as a complex and not fully observed test case.

Network motifs are thought to represent building blocks of larger biological systems (Milo, 2002). It is thus informative to test the BDEN with respect to these motifs to better understand the possible dependency of the performance of our models on different basic network topologies.

The dynamic EpoR model reflects the information processing at the EpoR including turnover, recycling and mobilization of the EpoR after stimulation with erythropoietin (Epo) at the cell membrane (Becker et al., 2010). Consequently, it details the first part of the JAK-STAT signaling pathway. Only combinations of Epo concentrations in the medium, on the surface and in the cells are accessible and thus it represents a complex model with limited experimental data (Becker et al., 2010).

In contrast, the thermal isomerization of α-Pinene (αP) in the liquid phase has the purpose to investigate the applicability of the BDEN to small compound reaction networks (Fuguitt and Hawkins, 1947). The model details the racemization of αP and its simultaneous isomerization to dipentene (dP) and allo-ocimene (aO).

To further investigate the utility of the BDEN for complex systems we used a gene regulatory network composed of 6 genes and related proteins obtained from the DREAM6 challenge (Meyer et al., 2014).

Further details about the described models are given in Appendix D.

7.3.2 Simulation Study

First, the performances of the BDEN as well as the DEN approach to correctly predict the location of single hidden influence for the JAK-STAT signaling model, the heterotrimeric GP cycling, the UV-B signaling in plants and the aforementioned network motifs are compared (Swameye et al., 2003; Yi et al., 2003; Ouyang et al., 2014; Milo, 2002). This was done on the basis of simulated data for each system. Details about the simulations are given in Appendix D. For the BDEN we computed for each $\overline{\boldsymbol{w}}_n$ (where $\overline{\boldsymbol{w}}_n$ denotes the posterior mean taken over the MCMC samples) the area under the predicted hidden influence curve is computed by trapezoidal numerical integration (Atkinson, 1989). For the DEN the same method based on the provided point estimates of hidden influence curves is applied. The area under the predicted hidden influence curve was compared against the simulated existence and non-existence of a hidden signal at that node. Consequently, it is possible to compute an area under the ROC curve (AUROC) value and a corresponding Brier score (BS), i.e., the squared difference between the prediction score and the Boolean indicator of a true hidden influence (Fawcett, 2006; Murphy, 1973). Table 7.2 and Table D.1 show a favorable overall performance of our new method for different levels of measurement noise and simulated errors of kinetic parameter estimates.

Model	Noise Level	Method	AUROC	BS
JAK-STAT	2.5%	BDEN	0.90 (0.15)	0.11 (0.11)
		DEN	0.60 (0.40)	0.16 (0.06)
	7.5%	BDEN	0.83 (0.18)	0.21 (0.19)
		DEN	0.43 (0.29)	0.30 (0.14)
	12.5%	BDEN	0.75 (0.25)	0.26 (0.16)
		DEN	0.42 (0.31)	0.41 (0.12)
G protein	2.5%	BDEN	0.99 (0.02)	0.04 (0.03)
		DEN	1.00 (0.00)	0.09 (0.02)
	7.5%	BDEN	0.88 (0.13)	0.17 (0.09)
		DEN	0.80 (0.13)	0.16 (0.09)
	12.5%	BDEN	0.80 (0.16)	0.22 (0.10)
		DEN	0.71 (0.16)	0.20 (0.11)
UV-B	2.5%	BDEN	0.91 (0.11)	0.19 (0.06)
		DEN	0.80 (0.19)	0.22 (0.08)
	7.5%	BDEN	0.88 (0.14)	0.19 (0.04)
		DEN	0.80 (0.19)	0.20 (0.06)
	12.5%	BDEN	0.81 (0.15)	0.19 (0.05)
		DEN	0.71 (0.15)	0.19 (0.05)
Motifs	2.5%	BDEN	1.00 (0.00)	0.00 (0.00)
		DEN	0.90 (0.14)	0.11 (0.09)
	7.5%	BDEN	1.00 (0.00)	0.00 (0.00)
		DEN	0.81 (0.19)	0.19 (0.04)
	12.5%	BDEN	1.00 (0.00)	0.00 (0.00)
		DEN	0.80 (0.15)	0.19 (0.05)

Table 7.2: The performance of the BDEN and the DEN regarding the dependence on the measurement noise (median). The median absolute deviation for the AUROC and Brier Scores (BS) are given in brackets.

Next we investigated the performance of the BDEN to correctly detect more than one hidden influence. The comparatively large GP cycle model is used for this purpose. The results can be found in Table D.2. Notably, in this simulation, hidden influences for up to 50% of the nodes are added randomly and still a good prediction performance can be observed.

In a similar manner the performance of the BDEN to detect wrong and missing interactions is investigated (Table 7.3). Therefore wrong model specifications of the heterotrimeric GP cycling, the U-VB signaling in plants and the synthetic JAK-STAT signaling are simulated by randomly removing and adding interactions. As described above, the quantitative predictions of the BDEN are given in terms of (cross-)correlations. By comparing these correlation values against the true existence and non-existence of a particular interaction, it is possible to compute an AUROC value. Notably, missing and wrong interaction detection is only possible with the BDEN approach. Again a very good performance for all systems under investigation is archived. On average, 80% of the missing interactions are correctly detected by the BDEN. Among the correctly identified missing interactions, on average 90% were correctly classified as "stimulating" and "inhibiting", respectively (cp. Appendix D). Details regarding the dependency on the measurement noise are given in Table 7.3 and results in dependency of deviance of the parameter estimates are given in Tables D.4 and D.5.

	Model	Noise Level	AUROC		Model	Noise Level	AUROC
Missing Interaction	JS	2.5%	1.00 (0.00)	Wrong Interaction	JS	2.5%	1.00 (0.00)
		7.5%	0.83 (0.28)			7.5%	0.87 (0.32)
		12.5%	0.80 (0.32)			12.5%	0.80 (0.23)
	GP	2.5%	0.81 (0.19)		GP	2.5%	1.00 (0.00)
		7.5%	0.78 (0.22)			7.5%	0.81 (0.32)
		12.5%	0.62 (0.47)			12.5%	0.73 (0.40)
	UV-B	2.5%	1.00 (0.00)		UV-B	2.5%	0.81 (0.20)
		7.5%	0.91 (0.11)			7.5%	0.70 (0.24)
		12.5%	0.76 (0.16)			12.5%	0.68 (0.30)

Table 7.3: Performance of the BDEN to detect wrong and missing interactions depending on the measurement noise (median) evaluated for the JAK-STAT (JS), G protein (GP) and UV-B network. The median absolute deviation for the AUROC is given in brackets.

7.3.3 Illustrative Examples With Real Data

In the following we further illustrate the results obtained with the BDEN for the JAK-STAT signaling model, the information processing at EpoR and the isomerization of αP using experimental data.

JAK-STAT Signaling

The JAK-STAT signaling pathway model (cp. Section 6.3.1) consists of four molecular species: Unbound STAT5 molecules become phosphorylated ($\mathrm{STAT5_p}$) catalyzed by the erythropoietin receptor. Two $\mathrm{STAT5_p}$ molecules can form a dimer ($\mathrm{STAT5_{di}}$) and thus are able to enter the nucleus ($\mathrm{STAT5_n}$). Only the amount of phosphorylated STAT5 molecules, the total amount of STAT5 and the erythropoietin receptor are directly accessible. Experimental measurements are available at 16 time points (Swameye et al., 2003).

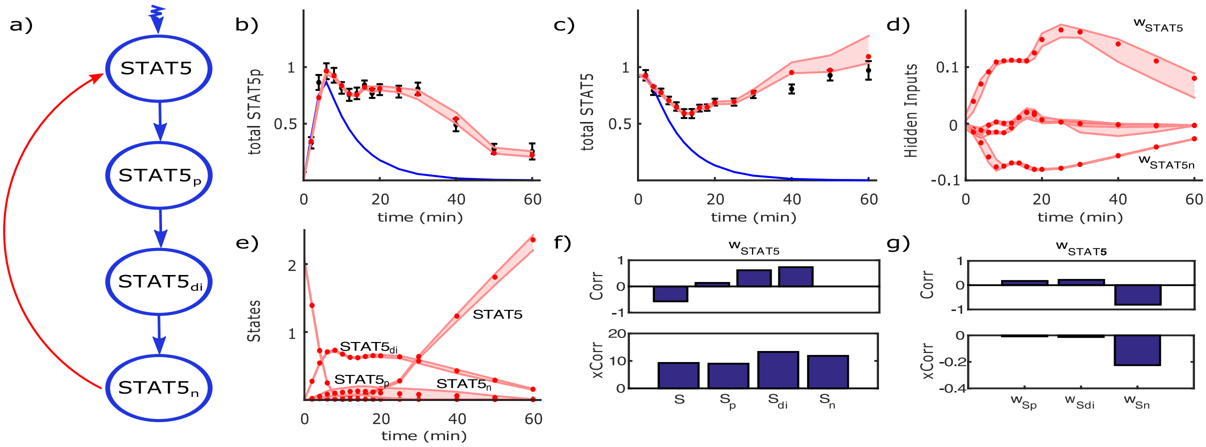


Figure 7.4: Reconstructing the model error for the JAK-STAT signaling pathway (Swameye et al., 2003). (a) Nominal model of the JAK-STAT signaling (blue) including its nucleocytoplasmic cycling (red). (b,c) Measurements (black) for phosphorylated cytoplasmic STAT5 and total cytoplasmic STAT5 and posterior BDEN predictions including the 95% credible intervals (red) based on the nominal system (blue). (d) Estimated hidden inputs (posterior means) and 95% credible intervals. There is a clear input located at STAT5 and STAT5 in the nucleolus. (e) Estimated time series posterior means for all modeled variables including 95% credible intervals. (f) Estimated correlations (Corr) and cross-correlations (xCorr) of the estimated hidden input w_{STAT5} located at STAT5 with all estimated state variables. (g) Estimated correlations and cross-correlations of w_{STAT5} with all remaining hidden influences. Here w_{STAT5} is clearly correlated to the hidden input located at $\mathrm{STAT5_n}$ and in addition it is correlated to the time series of $\mathrm{STAT5_n}$. High cross-correlation is a necessary but not sufficient condition.

Figure 7.4 illustrates the application of our method when ignoring the back-translocation of $\mathrm{STAT5_n}$ into the cytoplasm, which was hypothesized by the authors (Swameye et al., 2003). After parameter fitting the nominal system is not in sufficient agreement with the data. The introduction of hidden influence terms w_n leads to good agreement with the observations. The BDEN clearly localized two hidden influences w_{STAT5} and $w_{\mathrm{STAT5_n}}$ at STAT5 and $\mathrm{STAT5_n}$. The subsequent analysis shows a high positive (cross-)correlation of w_{STAT5} with $\mathrm{STAT5_n}$ and a negative one with $w_{\mathrm{STAT5_n}}$. Exactly the opposite pattern can be observed for $w_{\mathrm{STAT5_n}}$. According to the above explained procedure a stimulatory influence of $w_{\mathrm{STAT5_n}}$ on STAT5 is predicted. This is in agreement with the claimed nucleocytoplasmic cycling of the phosphorylated STAT5 dimer (Swameye et al., 2003; Raue et al., 2009).

EpoR Model

The complex core model of the EpoR regulation via receptor mobilization, turnover and recycling involves 6 species and 8 time points. Here the ligand Epo binds to EpoR on the surface and builds a ligand-receptor complex (Epo-EpoR). In consequence, Epo-EpoR triggers the phosphorylation of the cytoplasmic EpoR and thus induces the JAK-STAT signaling pathway (Becker et al., 2010). Several mechanisms affect the amount of active EpoR. This model covers the ligand-induced receptor endocytosis and thus the internalization of the ligand-bound receptor ($\mathrm{Epo\text{-}EpoR_i}$), receptor recycling and degradation of the internalized ligand-bound receptor.

Location-dependent degradation results in degraded Epo in the cytoplasm (dEpo$_\mathrm{i}$) and in the medium (dEpo$_\mathrm{e}$).

As a test case for our method we wrongly specified a receptor-induced feedback on the amount of available Epo. In consequence, we expect to detect this wrong interaction. As shown in Figure 7.5, the BDEN allows to correctly localize and characterize this erroneous interaction in the nominal model.

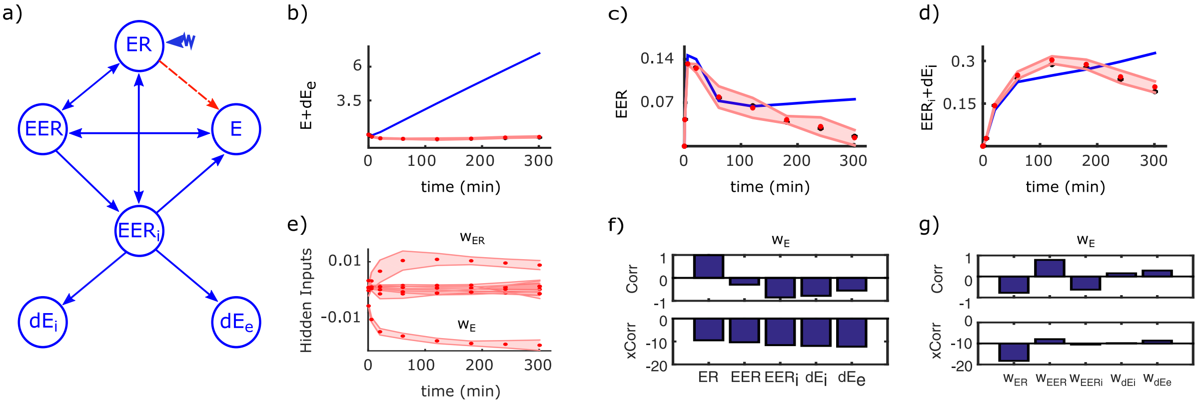


Figure 7.5: Reconstructing the hidden influence in a model of the EpoR regulation (Becker et al., 2010). (a) Reaction graph of the model (E=Epo, ER=EpoR, EER=Epo-EpoR, EER$_\mathrm{i}$=Epo-EpoR$_\mathrm{i}$, dE$_\mathrm{i}$=dEpo$_\mathrm{i}$, dE$_\mathrm{e}$=dEpo$_\mathrm{e}$). The red arrow is a wrong reaction. (b,c,d) Output measurements (black) compared to posterior BDEN predictions (red) including 95% credible intervals and the nominal model (blue). (e) Estimates of the hidden influences (posterior mean) including 95% credible intervals. (f) Estimated correlations (Corr) and cross-correlations (xCorr) of the hidden influence related to w_E with all estimated state variables. A clear correlation with EpoR is observable. (g) Estimated correlations (Corr) and cross-correlations (xCorr) of the hidden influence related to w_E with all estimated remaining hidden influences. Again w_E is clearly correlated with w_{ER}. In consequence, the direct interaction between ER and E is correctly classified as wrong and has to be removed.

α-Pinene Isomerization

The model of the dynamic isomerization of αP is composed of 4 molecular species. Measurements of αP, dP, aO and the dimer (Di) are available (Fuguitt and Hawkins, 1947). After heating, αP reacts either to dP or builds a dimer by reacting with aO. Furthermore, Di can react to aO.

To test the BDEN the dimerization step was wrongly replaced with a simple reaction involving only aO. In consequence the interaction between αP and dP is completely independent from the interaction between aO and Di. The erroneous nominal system can be corrected by using our method as illustrated in Figure 7.6. The BDEN is able to correctly locate and add the falsely removed reaction.

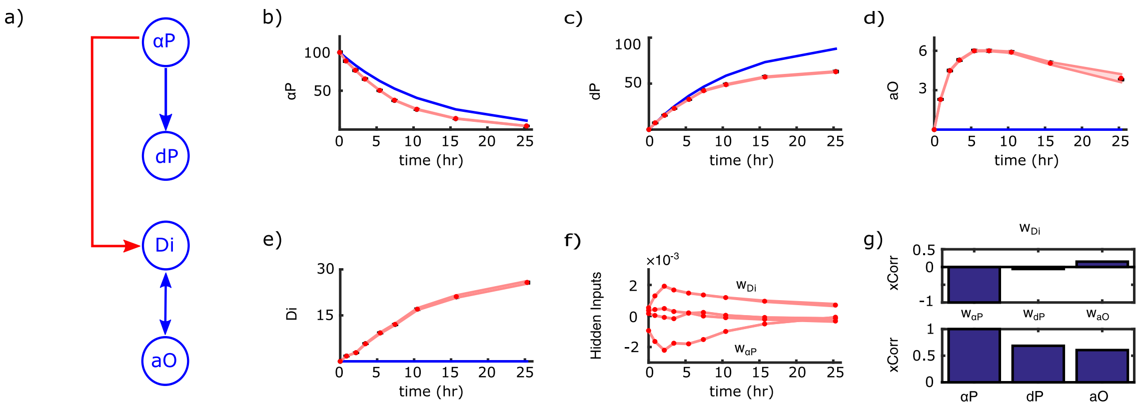


Figure 7.6: Reconstructing the hidden influence in a model of the dynamic αP isomerization (Fuguitt and Hawkins, 1947). (a) Reaction graph. The red arrow indicates a missing reaction. (b,c,d,e) Output measurements (black) compared to the posterior BDEN predictions (red) including 95% credible intervals and the nominal model (blue). (f) Estimates of the hidden influences (posterior mean) including 95% credible intervals. (g) Estimated relative cross-correlations (xCorr) of the hidden influence of w_{Di} to all estimated remaining hidden influences and state variables. The BDEN is able to correctly detect the missing reaction.

7.3.4 Further Examples With Simulated Data

In the following we further illustrate the results obtained with the BDEN for the GP cycle in yeast, the UV-B signaling model and a generic gene regulatory network using simulated data (2.5% noise level).

G protein Cycle in Yeast

The heterotrimeric GP cycle in yeast involves 6 species which are directly observable and coupled by several types of kinetics (for details please see the Appendix D) (Yi et al., 2003). Experimental data at 8 time points were simulated by adding Gaussian distributed noise to the predicted values of the observable variables. A noise intensity of 2.5% relative to the mean of the related time series for each observable variable is assumed. The nominal system was generated by adding one additional "wrong" interaction (between the receptor-ligand complex (LR) and the GP_α-inactive ($GP_{\alpha i}$)). Figure 7.7 illustrates the ability of the BDEN to localize and recover the wrong interaction within the nominal system.

UV-B Signaling

As a more complex example we simulated 7 data points of the photomorphogenic UV-B signaling in plants (cp. Section 6.3.2) (Ouyang et al., 2014). The model of the photomorphogenic UV-B signaling in the model plant *Arabidopsis thaliana* consists of 11 species coupled by several different kinetic rate expressions and five observable variables as a combination of 7 different species (for details please see Appendix D). As the nominal system the literature given model is

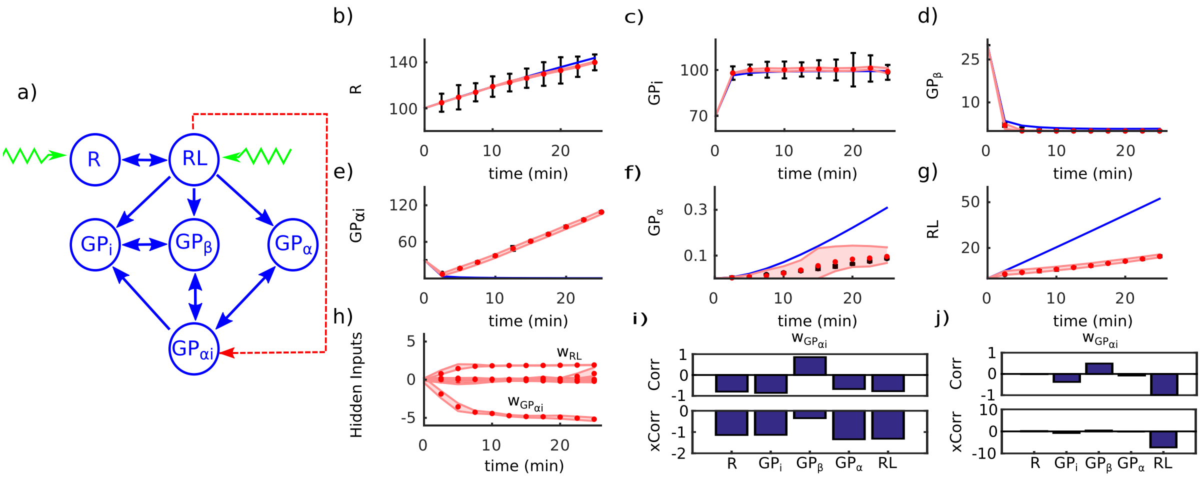


Figure 7.7: Reconstructing the hidden influence of the heterotrimeric G protein cycle in yeast (Yi et al., 2003). (a) The reaction graph. (b,c,d,e,f,g) Simulated output measurements (black) compared to the posterior mean of the output variables (red) including 95% credible intervals and the nominal model (blue). (h) Estimates of the hidden influences (posterior mean) including 95% credible intervals. (i) Estimated correlations (Corr) and cross-correlations (xCorr) of the hidden influence related to GP$_\alpha$-inactive with all estimated state variables. (j) Estimated correlations and cross-correlations of the $w_{GP_{\alpha i}}$ related to GP$_\alpha$-inactive with all remaining hidden influences.

used and a missing link is included by adding one interaction which influences two different species. Observed data were simulated by adding Gaussian distributed noise to the predicted values of the observable variables. Further a noise intensity of 2.5% with respect to the mean of the related time series is assumed for each observable variable. As shown in Figure 7.8 the BDEN is able to detect the missing molecular interaction and correctly identifies the corresponding proteins.

DREAM6 Challenge Network

To investigate the applicability of the BDEN to gene regulatory networks we took a model from the DREAM6 challenge (Meyer et al., 2014). The model consists of 6 genes and 6 proteins coupled by mass action and hill kinetics. In this model all proteins and one mRNA species are assumed to be directly observable (for details please see Appendix D) (Meyer et al., 2014). As the nominal system we used the provided model and included one inhibitory mechanism. Observed data were simulated at 5 time points by adding Gaussian distributed noise to the predicted values of the observable variables according to the original challenge (Meyer et al., 2014). The BDEN is able to detect and correct the spurious interaction, as illustrated in Figure 7.9.

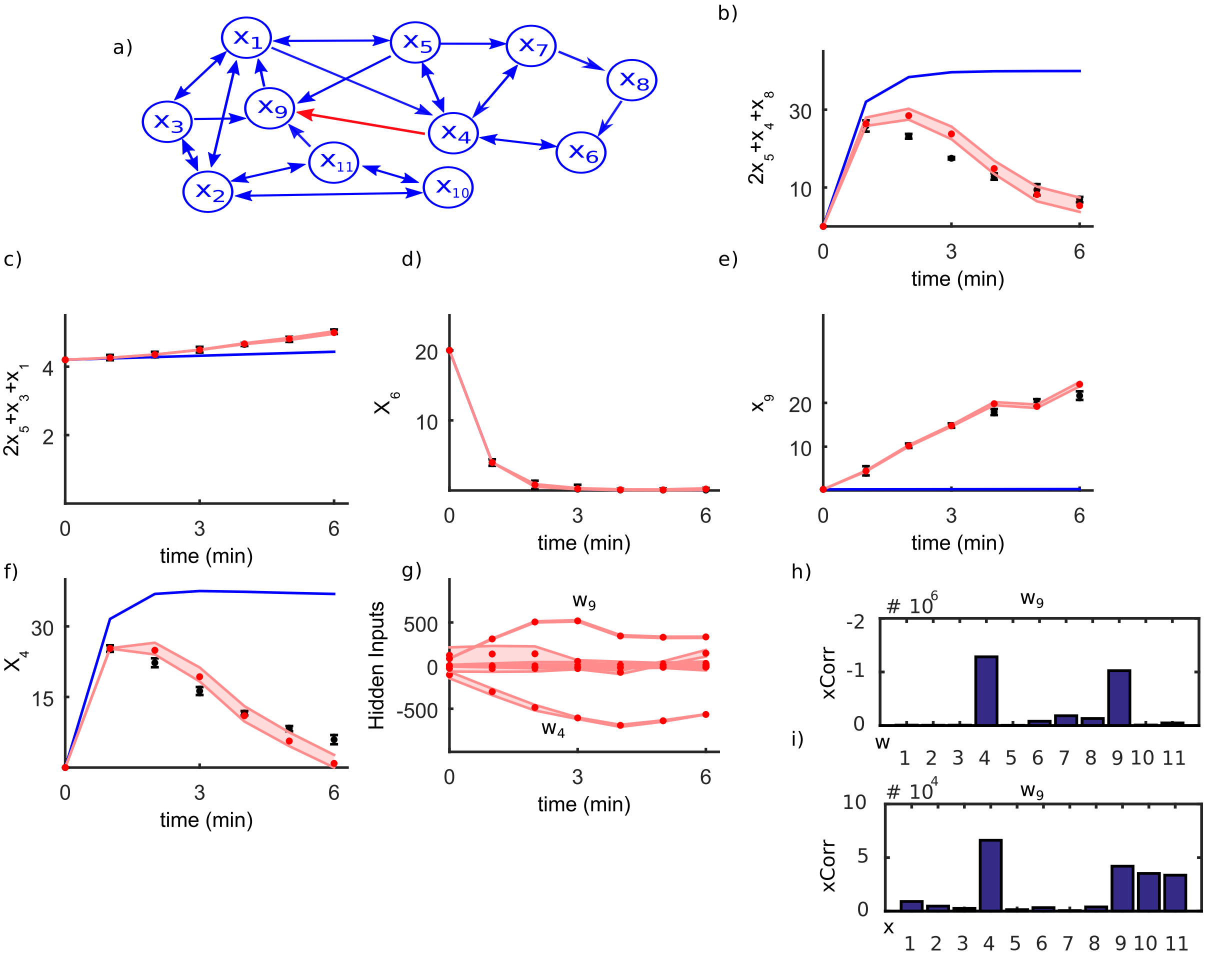


Figure 7.8: Reconstructing the model error for the photomorphogenic UV-B signaling (Ouyang et al., 2014) in plants. (a) Reaction graph. (b,c,d,e) Synthetic measurements (black) compared to the predicted (posterior mean) outputs including 95% credible intervals and the nominal model (blue). (f) Posterior means of the hidden influences including 95% credible intervals. (g) Posterior means of the model variables including 95% credible intervals. (h) Estimated cross-correlations of all involved hidden influences with respect to w_9. (i) Estimated cross-correlations of all state variables with respect to w_9.

7.4 Conclusion

Mathematical modelers in systems biology are frequently confronted with incomplete knowledge and limited understanding of a complex biochemical system (Babtie et al., 2014; Barzel et al., 2015; Gao et al., 2014). Consequently, there is a non-negligible chance that relevant molecular species are missed or interactions are misspecified (von Bertalanffy, 1950; Babtie et al., 2014). Commonly used statistical model selection and related methods require a strong knowledge of the system and its alternatives which is rarely given in practice (Azeloglu and Iyengar, 2015). Thus model selection can be very difficult, specifically if nothing is known about missing variables and their possible mechanisms. The question is, how to detect so far unknown molecules and their interactions in a more data-driven manner. This could guide

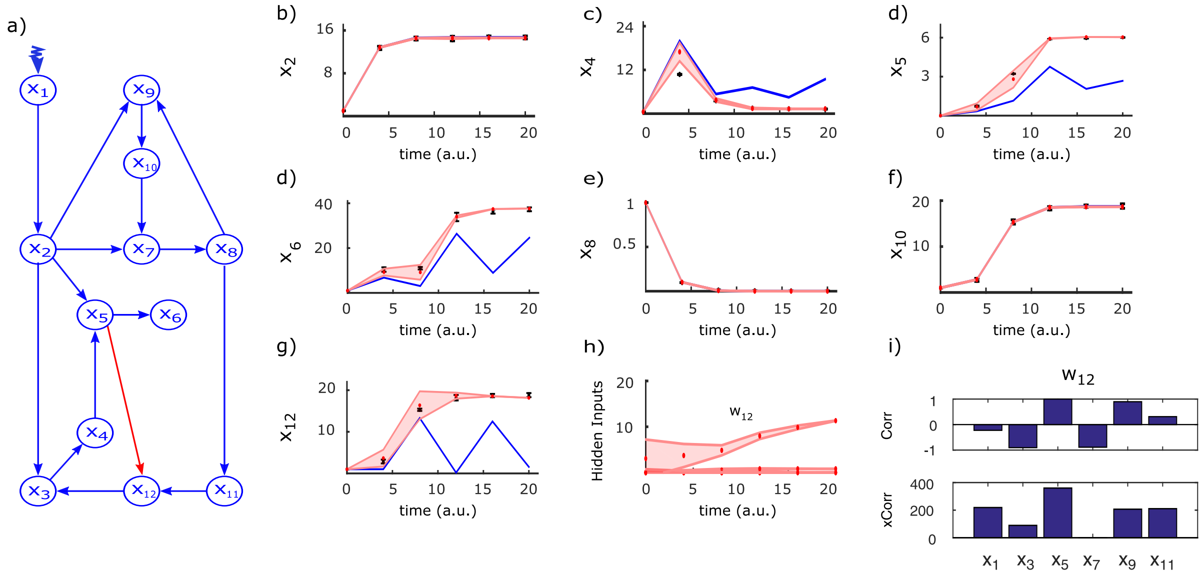


Figure 7.9: Reconstructing the hidden influence of the gene regulatory network obtained from the DREAM6 challenge (Meyer et al., 2014). (a) Interaction graph of the gene regulatory network including six proteins (even numbers) and related mRNAs (odd numbers). (b,c) Synthetic measurements of the protein concentrations (black) compared to the posterior mean of BDEN predictions (red) including 95% credible intervals and the nominal model (blue). (d) Synthetic measurements of the mRNA level (black) compared to the posterior mean of BDEN predictions (red) including 95% credible intervals and the nominal model (blue). (e,f,g) Synthetic measurements of the protein concentrations (black) compared to the posterior mean of BDEN predictions (red) including 95% credible intervals and the nominal model (blue). (h) Estimates of the hidden influences (posterior mean) including 95% credible intervals. (i) Estimated correlations (Corr) and cross-correlations (xCorr) of the hidden influence related to protein$_6$ (w_{12}) with all mRNAs ($x_1, x_3, x_5, x_7, x_9, x_{11}$) because in the gene regulatory network direct protein-protein interactions are not considered (Meyer et al., 2014). The highest correlation was found between protein$_6$ (x_{12}) and mRNA$_3$ (x_5). Hence, the spurious inhibition of protein$_6$ by mRNA$_3$ is correctly detected.

the modeler towards points in the given model, where the model is likely erroneous. In a second step the modeler can then try to link these erroneous points to known mechanisms. The BDEN is not designed to learn ODE systems purely from data and should thus not be confused with network reverse engineering methods (Sachs et al., 2005). Much more, the utility of the proposed BDEN and DEN are to ease identification of sources of errors in mechanism-based mathematical models.

The Bayesian dynamic elastic-net method addresses this issue by adopting a Bayesian framework which allows for inferring hidden influence variables, as well as estimating missing and wrong molecular interactions. It has been successfully validated in several real models as well as common network motifs. This was done with simulated as well as experimental data. Due to the fully Bayesian formulation all model parameters are sampled. Furthermore, the Bayesian approach allows to assign confidence levels to predictions. Besides these general features of a fully probabilistic framework, the BDEN method seems to be more stable and more robust because within the Bayesian framework over a large number of parameters is averaged and the BDEN does not rely on stiff integration methods.

A unique feature of our new approach is the distinction between exogenous and endogenous hidden influences in the biological system, allowing for the detection of missing and misspecified equations in the ODE system. Altogether, the BDEN is a further step towards a better automated and more objective revision of ODE-based models in systems biology. In view of the rapid progress of technologies to monitor biological dynamics, our approach could have implications for many fields including metabolic engineering, synthetic biology and pharmacokinetics/pharmacodynamics. As the DEN and the BDEN are designed for generic ODE models, they can also be applied to challenging modeling tasks in engineering, robotics and in the earth sciences (Engelhardt et al., 2016; Abarbanel, 2013).

CHAPTER 8

Refined Model of the M_2 Receptor-dependent Signaling

8.1 Introduction

As a further improvement, the structural model derived in Chapter 5 has been extended to include dynamic rate laws. The thus obtained time-resolved system allows to gain further insights into the behavior of the M_2 receptor-dependent signaling in CHO cells. The modified model now reflects the dynamic properties of the underlying signaling cascade. The involved kinetic parameters were estimated based on quantitative measurements of the related protein activities reflected by the amount of phosphorylated or associated protein structures. Therefore, the amount of phosphorylated proteins and thus active proteins can be obtained by performing repeated quantitative WB at defined time points. As discussed in Section 3.3, quantitative WBs are a common technique to measure protein concentrations. For this purpose the concentration of the phosphorylated form and the overall protein concentration are measured for each protein. As a result, the obtained phosphorylated protein concentrations are normalized with respect to the initial time point and the related amount of overall protein concentrations. This serves as a measure for the relative protein activation and reduces the measurement error typically arising from WBs.

8.2 Experimental Procedures

The protocols and techniques used in this study allow for an exact protein quantification, even if the protein concentration is low. Further experimental details can be found in Appendix E. The stable expression of the M_2 receptor was cross-checked through exposure to hygromycin. Successful and stable transfected CHO-Flip-In-hM2 cells include an additional gene which makes them resistant to hygromycin and thus populations with low expression levels can be sorted out. Furthermore, all antibodies were checked regarding their specificity and affinity to CHO-Flip-In-hM2 cells to ensure the correct targeting of all proteins. In contrast to all other

Protein	Active Form	Technique	Time (min)
AKT1	yes	WB	0, 2, 5, 10, 15, 20, 30
PKA	yes	WB	0, 2, 5, 10, 15, 20, 30
PDE	yes	WB	0, 2, 5, 10, 15, 20, 30
GRK2	yes	WB	0, 2, 5, 10, 15, 20, 30
cAMP	yes	WB	0, 2, 5, 10, 15, 20, 30
GEF	no	WB	0, 2, 5, 10, 15, 20, 30
M2	no	WB	0, 2, 5, 10, 15, 20, 30
RGS14	no	WB	0, 2, 5, 10, 15, 20, 30
GRK6	no	WB	0, 2, 5, 10, 15, 20, 30

Table 8.1: List of measured proteins and secondary messenger molecules.

measured proteins, the activity of PKA is not characterized by phosphorylation but by binding to cAMP. Because cAMP is a secondary messenger molecule, it is not possible to directly measure its binding to PKA based on WBs. The phosphorylation of several PKA specific motifs served as a marker of the overall PKA activation. This was addressed by a phospho-PKA substrate which detects peptides and proteins containing a phospho-Ser/Thr residue with arginine at the -3 and -2 positions. Apart from the involved GPs, the overall amount was measured for all modeled proteins to exclude a change in the expression levels of those proteins. This was considered evidence for the absence of gene-regulatory events within the observed time frame of 30 minutes.

The cAMP concentration was measured using a complete enzyme-linked immunosorbent assay (ELISA) kit which allows for the detection of extremely low cAMP levels. As summarized in Table 8.1, the concentrations of cAMP, 4 phosphoproteins, the related non-phosphoproteins and 4 additional non-phosphoproteins were measured at seven well-defined time points under stimulation with 1 µM iperoxo.

8.3 Methods

The reduced but comprehensive dynamic model was derived on the basis of the structural model illustrated in Figure 5.1 . For this purpose, the number of reactions and the number of species of the refined model were reduced by modeling the signaling cascades via the Michaelis-Menten kinetics (cp. Section 2.4). For CHO-Flip-In-hM2 cells, only a fraction of proteins is targetable through commercially available antibodies, and thus intermediate complexes were removed to avoid overcomplication. Thus, the resulting model is both reliable and realistic. This strategy is in contrast to Chapter 5, where a detailed model based on elementary reactions and detailed binding events was necessary to investigate the general topological properties of the system.

Figure 8.1 illustrates the included molecules and the involved kinetic laws. Due to a lack of experimental techniques, the detailed kinetics of the GP activation are not available. Therefore,

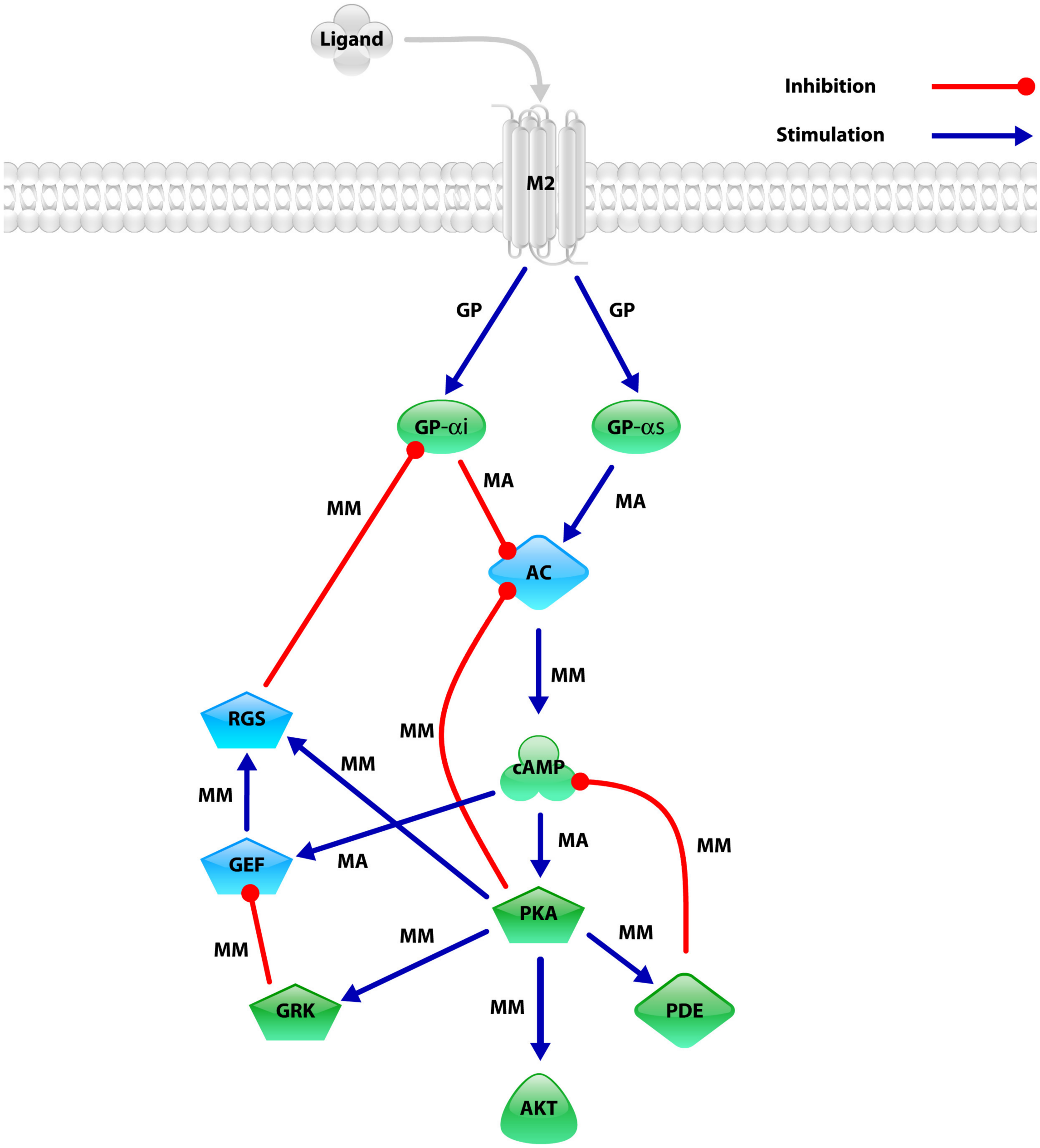


Figure 8.1: Schematic illustration of the redefined M_2 receptor-dependent signaling pathway including the rate-laws that were used (GP = GP Activation, MA = Mass Action, MM = Michaelis-Menten). The measured proteins are highlighted in green, modeled but not accessible proteins are colored in blue and not explicitly modeled proteins are colored in gray. As discussed, the kinetics of the GPs were approximated based on given data and were thus partially accessible. A detailed list of reactions can be found in Appendix E.

the irreversible GP activation had to be approximated by

$$GP(t) = GP\left(t_{eq}\right)\left(1 - \exp\left(-k_{obs} \times t\right)\right), \tag{8.1}$$

with the relative amount of active GP at time t given by $GP(t)$, $GP\left(t_{eq}\right)$ being the relative binding at equilibrium and k_{obs} the association rate constant (Jakubík et al., 1998). Even though Equation (8.1) was originally derived for GPs in liposomes, former studies have demonstrated that the equation can also be used to describe CHO-Flip-In-hM2 cells (Klemt, 2005). The rate constants for the iperoxo-induced GP activation in CHO-Flip-In-hM2 cells are based on [^{35}S]GTPγS assays and were taken from literature (Kebig, 2010; Strange, 2010). In agreement with all other measures, the GP activation is also normalized with respect to the initial time point.

Since Equation (8.1) approximates the total amount of relative GP activation, the fraction of α_i and α_s subunits was introduced as an additional parameter. Similarly, the amount of GP-bound and thus active AC is not directly accessible. Hence the complex of AC subunits is reduced to one comprehensive species. This does not affect the overall AC activation and inhibition behavior because the secondary messenger cAMP, which is a direct product, was measured. Notably, the GRK-induced receptor phosphorylation which leads to a deactivation of the receptor is covered by Equation (8.1) and thus not modeled explicitly.

Here, similar to the model presented in Chapter 5, actin and tubulin are used as markers for the DMR response. Especially the AKT1 phosphorylated at serine 473 serves as an indicator for changes in the actin cytoskeleton (Schmidt and Hall, 1998). In literature, a PKA-induced AKT dephosphorylation has been observed, which directly leads to changes in the cytoskeleton (Howe, 2004; Schmidt and Hall, 1998; Lou, 2002). As discussed in more extent in the following section, the measured data is in agreement with these results, as shown in Figure 8.2. Due to a lack of suitable antibodies, the activation of tubulin is not measurable in CHO cells. Therefore, GRK is used instead in the presented model. The full ODE system of the refined model is given in Appendix E.

8.4 Results

The data was normalized against the control value $\boldsymbol{x}(0)$ and the overall protein concentrations to obtain the dimensionless fold changes as a reliable measure for the protein activity:

$$\tilde{\boldsymbol{x}}(t) = \frac{\boldsymbol{x}_{overall}(0)}{\boldsymbol{x}_{phospho}(0)} \frac{\boldsymbol{x}_{phospho}(t)}{\boldsymbol{x}_{overall}(t)} \quad \forall t = \{0, 2, 5, 10, 15, 20, 30\}. \tag{8.2}$$

After parameter estimation (cp. Section 4.4), a good fit of the predictions to the data was archived (Kaschek et al., 2016). As shown in Figure 8.2, the predictions are within the measurement tolerances. The peak at about 9 minutes is characteristic for cAMP dynamics and has been observed elsewhere (Williamson et al., 2009). The observed deactivation of phosphorylated AKT at serin 473 is a well-known characteristic of the Rap1b-mediated cAMP inhibition of

AKT and has also been previously reported in literature (Lou, 2002; Howe, 2004).

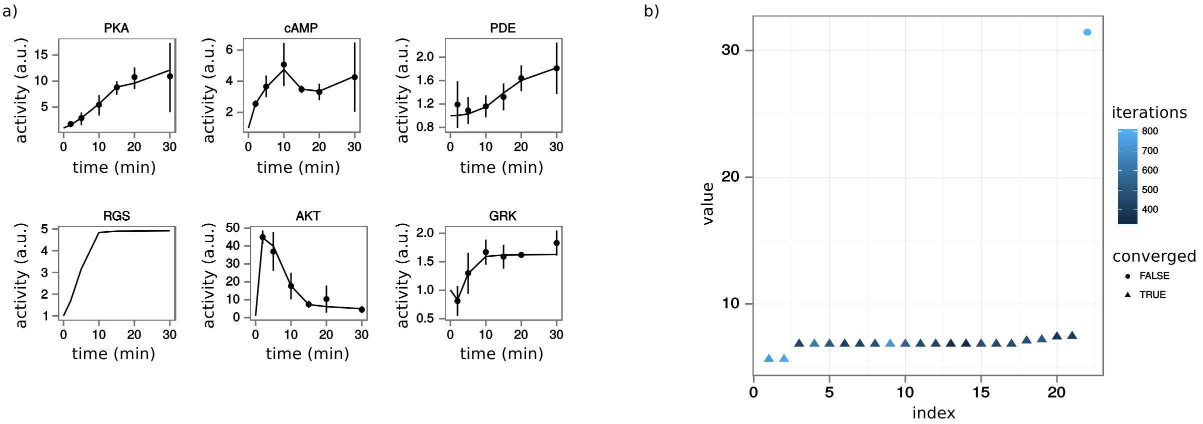


Figure 8.2: Model output. (a) Predictions (lines) and measured values including measurement errors (dots). All predictions are within the measurement tolerances. RGS was not measurable, and thus only the predicted dynamics are shown. (b) Achieved sets of parameter estimates for different runs (index) sorted by increasing objective values (y-axis). Low values indicate small errors and thus a good match with the data. The fits with the lowest values clearly separate from the second plateau. The complete plot including all modeled proteins can be found in Appendix E.

In addition, the clear increase of RGS is in agreement with the initial hypothesis stated in Chapter 5. To further investigate the role of RGS as an important regulator of the M_2 receptor-dependent signaling dynamics, in silico knock-out experiments were performed, as shown in Figure 8.3. For the experiments, the concentration of RGS and all RGS-related reactions were set to be constant zero. This mimics the absence of RGS from the system and thus allows to investigate the influence of RGS on the modeled signaling pathway. Accordingly, for verification, the concentration of cAMP and all cAMP-related reactions were set to zero as well. A clear decrease of the signaling strength was observed after RGS knock-out. In agreement with recent literature, a clear change in AKT was observed (Zhang and Mende, 2011; Dale and Rang, 2011; Howe, 2004; Pierce et al., 2002). This underlines the importance of the RGS-related feedback for the DMR response. The tubulin-related change of the DMR response also depends on the RGS regulation. The completeness and consistency of the refined model was tested both with the DEN and the BDEN. As shown in Figures 8.4 and 8.5, no clear model error was found. This underlines that the model predictions and the experimental data are in good agreement. Taken together, this provides evidence for the reliability of the obtained dynamic model.

8.5 Conclusion

A careful model revision and the incorporation of kinetics, e.g., Michaelis-Menten kinetics and GP activation, allowed for a time-resolved model of the M_2 receptor-dependent signaling in CHO cells. As shown in Figure 8.2, the model is in full agreement with the measurement data and common literature. It is also in agreement with the former derived static model described

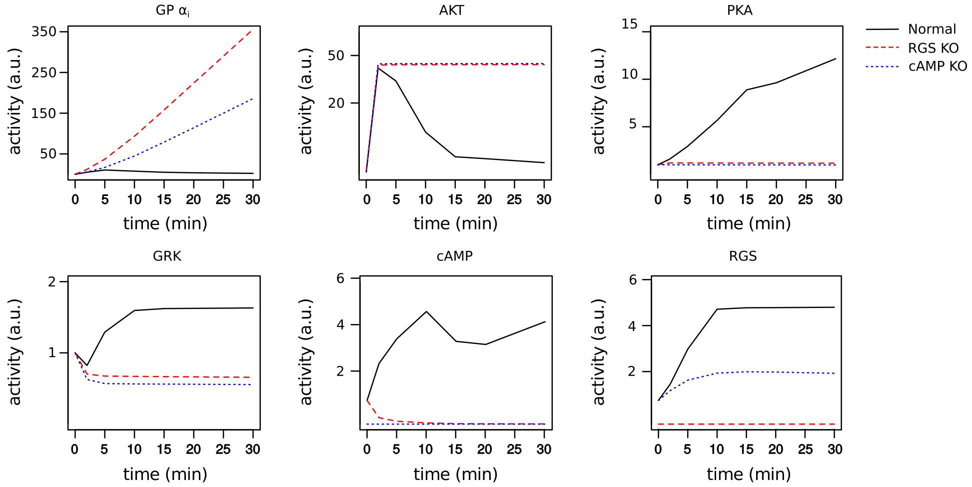


Figure 8.3: In silico knock-out studies. Model predictions without knock-out (black) compared with predictions with in silico RGS knock-out (red) and in silico cAMP knock-out (blue). The cAMP knock-out serves as control for the model reliability. As expected, after knock-out of cAMP, the downstream signaling is discontinued. In addition, the dephosphorylation of AKT vanishes as well. After RGS knock-out, again a clear shift in the cAMP-related downstream caused by an over-activity of GP α_i can be observed.

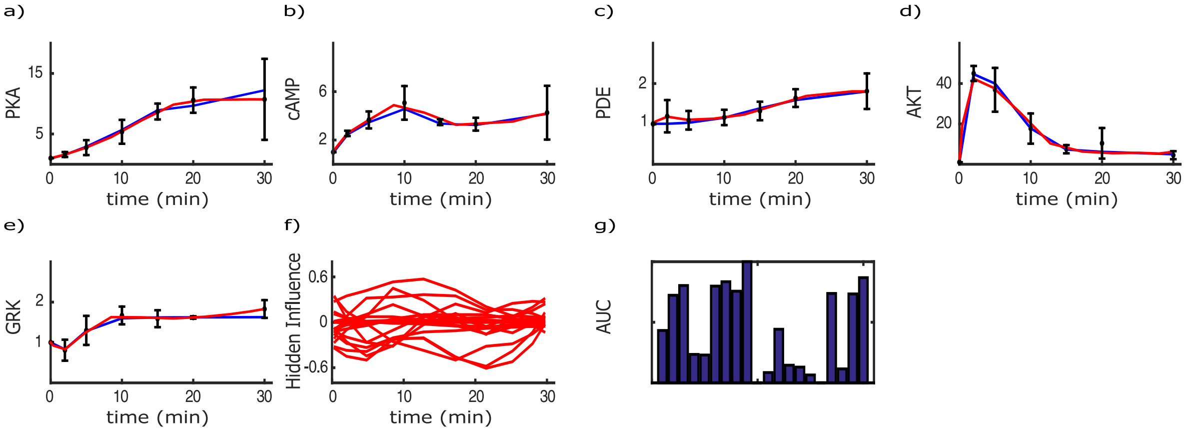


Figure 8.4: Model verification using the DEN. (a,b,c,d,e) Measurements (black) compared to the posterior mean of the output variables (red) including 95% credible intervals and the nominal model (blue). (f) Estimates of the hidden influences. (g) Absolute area under the hidden influence curves. Only adaption to the measurement noise can be observed which provides further evidence for the completeness and consistency of the obtained dynamic model.

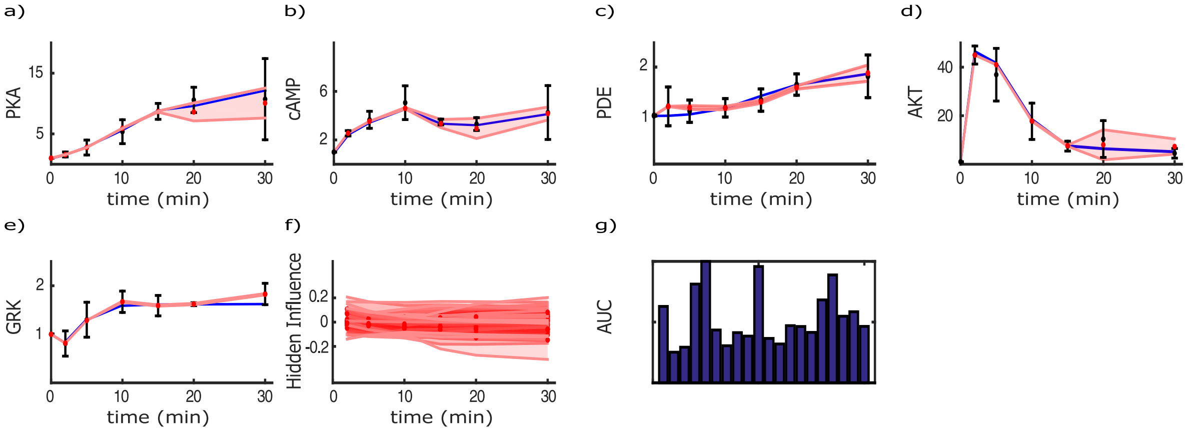


Figure 8.5: Model verification using the BDEN. (a,b,c,d,e) Measurements (black) compared to the posterior mean of the output variables (red) including 95% credible intervals and the nominal model (blue). (f) Estimates of the hidden influences (posterior mean) including 95% credible intervals. (g) Absolute area under the hidden influence curves (median). Only adaption to the measurement noise can be observed which provides further evidence for the completeness and consistency of the obtained dynamic model.

in Chapter 5. The dynamic model underlines the role of RGS as an important regulator of the M_2 receptor-induced DMR response, as illustrated in Figure 8.3.

Although the model proposed in this work is in full agreement with the obtained data, it is still restricted to the approximated GP activation and by the lack of available tubulin measurements. However, the approximated GP activation is in agreement with accepted conventions and thus is assumed to reflect the general behavior of the GP activation (Jakubík et al., 1998). It could be shown that this approximation holds true in CHO-Flip-In-hM2, but further time-resolved $[^{35}S]GTP\gamma S$ assays are necessary to verify this assumption (Strange, 2010). In addition, the PKA and AC activities are not directly measurable, which is covered by Equation (8.2). In consequence, the derived model parameters are also dimensionless and thus not directly interpretable. However, for the purpose of this work, i.e., uncovering the important players in the considered signaling cascade, this is no limitation.

To verify the reliability of the obtained model the developed DEN and BDEN were used. As shown in Figures 8.4 and 8.5 no further mechanisms were found. This provides evidence for the reliability of the refined model. However, further perturbation experiments are desirable to verify the role of RGS. Additional experimental evidence for the uncovered RGS-induced signal amplification could be obtained through in vitro knockdown of the related mRNAs, e.g. via RNA interference (Zhang and Mende, 2011). In recent literature, a RGS-independent modulation of the GP α_i subunits by PKA is described for yeast (Stefan et al., 2011). However, a direct modulation via PKA has so far not been described for other tissues, e.g., CHO cells, and the proposed findings give no evidence for a direct modulation via PKA. Again in vitro knockdown of the RGS-related mRNAs could give further evidence for the absence of such an

RGS-independent signal amplification.

Altogether, the derived model provides evidence for RGS as an important modulator of the M_2 receptor-induced DMR response. As illustrated in Figure 8.6, the RGS-dependent signal amplification can be considered as a positive feedback loop. Here, the AC-induced activation of PKA leads to an RGS-mediated inhibition of the α_i subunit. Consequently, the inhibitory effect of α_i on AC decreases. The therapeutic potential of RGS in the context of heart diseases based on RGS-induced signal amplification is well known for other systems and described in literature (Kimple et al., 2011; Engelhardt et al., 2017). This underlines the importance of RGS for the observed M_2 receptor-induced signal amplification. Thus the proposed model serves as a starting point for further in vitro investigations of the M_2 receptor-induced signaling in CHO cells.

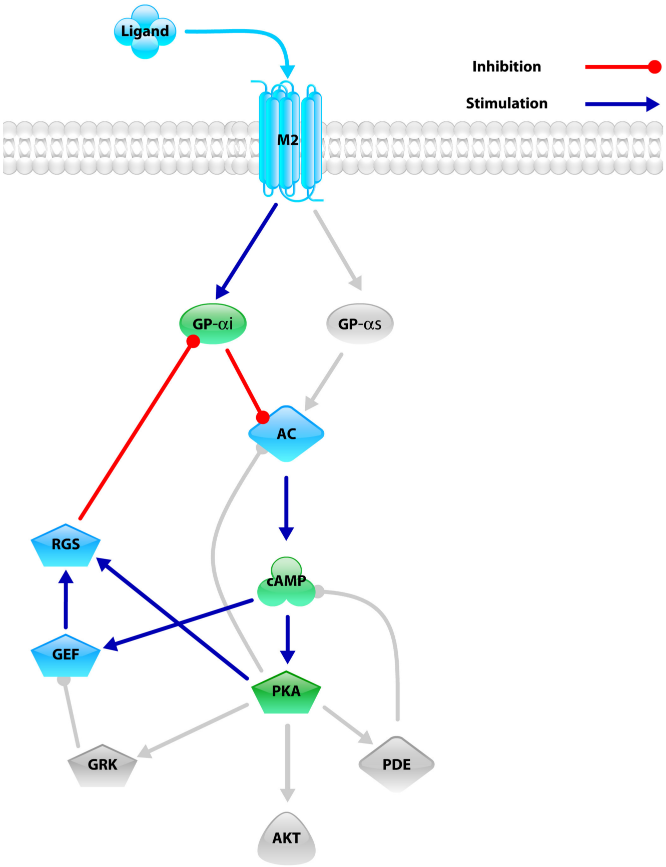


Figure 8.6: Schematic illustration of the RGS-induced positive feedback in the M_2 receptor-dependent signaling pathway. A clear RGS-mediated feedback loop was uncovered in this work. Related proteins are highlighted in green if measurable and blue if not. All unrelated proteins and interactions are gray-colored.

CHAPTER 9

Conclusions

This work demonstrates the importance of theoretical mathematical methods as a tool for solving concrete biological research tasks. Although there is already a huge range of mathematical and physical approaches in the field of life sciences, the present thesis underlines the significance of individually tailored methods. It is motivated by the strong signal amplification of the M_2 receptor-induced whole-cell response in CHO cells (Schrage et al., 2013). To address this phenomena in static models, the novel CFS is proposed in this work (Engelhardt et al., 2017). As part of the model development process, in addition, a novel and promising approach to learn from the errors of systems biological models was developed. This approach is not restricted to biological systems but can be applied to any ODE-based erroneous model (Engelhardt et al., 2016).

The proposed methods are inspired by modern systems biology. Thus the intention was to develop a mechanistic model of the underlying biological properties. However, this is just one possible approach to deal with such tasks in the field of mathematical biology and bioinformatics (Machado et al., 2011; Kitano, 2002). Hence, besides the algorithms developed in this work, there are also other possible routes (Machado et al., 2011; Kitano, 2002). For instance, a fully probabilistic strategy based on machine learning could help to uncover correlations between the modeled molecules and structures (Machado et al., 2011; Kitano, 2002; Bishop, 2007). In a further step, a predictive model of the M_2 receptor-induced whole-cell response could be established. This approach is preferable if the underlying detailed mechanisms are not completely understood. In contrast, in the present case, the aim was to derive a detailed explanation of the underlying mechanisms in the light of comparably rich and detailed mechanistic insights from literature. This is of essential interest because the whole signaling pathway has yet to be fully investigated, even though the M_2 receptor-dependent signaling in CHO cells is of high relevance for pharmacological research (Engelhardt et al., 2016; Schröder et al., 2011; Schrage et al., 2013). More precisely, the system investigated in this work could serve as a pharmacological model for the evaluation of drug affinities and effectiveness (Engelhardt et al., 2016; Schröder et al., 2011; Schrage et al., 2013). Other methods designed for the analysis of signaling cascades more strongly focus on signal duration and signal strength (Klipp and Liebermeister, 2006;

Heinrich et al., 2002). Therefore, this work provides the foundation for future investigations of the system (Engelhardt et al., 2017).

Starting with a static model of the M_2 receptor-dependent signaling cascade, this work serves as an example for integrating methods from different areas. In literature, metabolic models and signaling cascades are clearly separated (Palsson, 2006; Klamt et al., 2006). Often this distinction holds true because the points of view are quite different. Whereas network analysis is mainly interested in the turnover of well-defined molecules and the subsequent optimization of these processes (e.g., in biotechnology), the analysis of signaling cascades deals with the temporal and quantitative aspects of protein pathways. However, apart from the velocity of metabolic processes (extremely fast) vs. signaling (slower), the underling biological mechanisms are equivalent and the biological properties are identical (Klipp and Liebermeister, 2006). Thus it seems obvious to combine both areas to develop more comprehensive approaches. The proposed CFS adopts and extends methods known from metabolic network analysis to address the amplification mechanism in signaling cascades (Engelhardt et al., 2017). CFS provides first insights into the behavior of static systems and allows for an initial verification of the model structure (Engelhardt et al., 2017). This may help to speed up the research cycle (cp. Section 2.2) and allows for ranking the modeled molecules and functional subnetworks with respect to their importance for signaling (Engelhardt et al., 2017). In addition, this method can be optionally enhanced through the incorporation of (partially) available data (Engelhardt et al., 2017).

RGS was localized as the main driver of the investigated signaling cascade in CHO cells (Engelhardt et al., 2017). If signaling is considered as a directed flow of information from the stimulus downstream to the whole-cell response, RGS together with GP constitute a feedback loop. Such positive feedback loops are known to function as signal amplifiers (Alon, 2007). As discussed in Chapter 8, further mRNA knockdown experiments would be needed to provide full experimental evidence for the hypotheses presented in this work. Despite the gained insights, CFS is still restricted to topology-based phenomena. As demonstrated, this is a good starting point, especially for bottom-up strategies, but in general biological phenomena are also time-dependent. Thus a realistic dynamic model of the observed biological system was established and time-resolved protein measurements via WBs were conducted. Based on the obtained data, the fully dynamic model was fitted and validated.

The data was normalized to reduce measurement uncertainties and thus served as a measure of the relative protein activities. For the given task, this is a sufficient way of modeling. However, one drawback is the non-interpretability of the parameter estimates. This issue can be addressed by more detailed measurements of the involved parameters. For instance, mass spectrometry or specific assays could be used for more precise measurements. On the other hand, such experiments are very costly and time-intensive in cases where the experimental pipeline has to be established with respect to the underlying system, e.g., the CHO cell.

Fitting models to data is still a challenging issue because parameter estimation remains an ill-conditioned inverse problem. In consequence, even after parameter estimation hidden confounders may cause significant model errors (Engelhardt et al., 2016). The novel DEN and

BDEN approaches address this issue through the estimation of hidden variables as well as missing and wrong molecular interactions in ODE-based systems biology models. At first glance, the underlying hidden influence detection problem described in Chapters 6 and 7 could be solved a priori by adding minimal error terms to each node, e.g., protein, of the model. This would lead to a perfect match with the data. However, this is neither a meaningful nor realistic solution (Azeloglu and Iyengar, 2015; Engelhardt et al., 2016). In the past, regularization techniques such as LASSO have been invented to shrink the solution space in order to obtain more robust model fits (Tibshirani, 1996). In addition, suitable regularization techniques can be used to force biological meaningful results. In contrast, unrestricted systems tend to have diverse solutions due to their inverse nature. A priori, it is not clear which solution to choose. In the spirit of LASSO, the developed DEN is based on an elastic-net regularization – a method similar to LASSO – in order to obtain biological meaningful estimates (Engelhardt et al., 2016). This is done by forcing the estimated hidden input signals towards a sparse and smooth solution. In consequence, the most important errors are located and thus unphysiological solutions are avoided (Azeloglu and Iyengar, 2015; Engelhardt et al., 2016). As a consequent extension of the DEN, the uncertainties of the estimated hidden input signals were taken into account. This resulted in the development of the BDEN. The BDEN is a fully Bayesian method as an alternative to the deterministic DEN. In contrast to the DEN, the BDEN has a higher robustness regarding the dependency on the measurement noise and rate parameter estimates and allows for the detection of missing and wrong reactions. On the other hand, due to the fact that the heart of the BDEN is based on MH sampling, the computational effort is much higher compared to the DEN (cp. Figure 9.1).

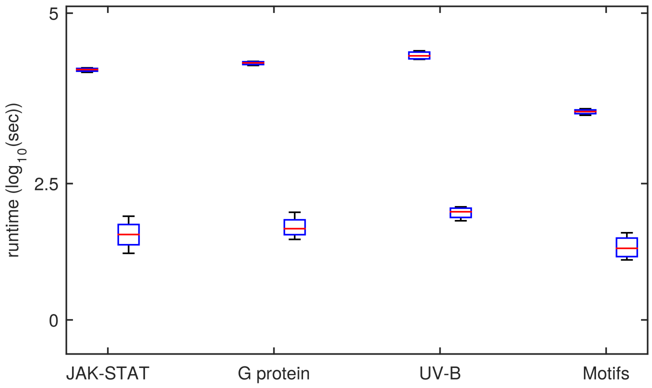


Figure 9.1: Runtime of the BDEN (left) in comparison with the DEN (right) for the JAK-STAT, UV-B and G protein network in seconds on a logarithmic scale ($\log_{10}$). Runtimes of the are on average 1 minute (DEN) and 254 minutes (BDEN) evaluated on a dual core laptop (Intel® Core™ I5-4200M CPU with 4x2.50 GHz and 16 GB RAM).

The DEN and the BDEN are in contrast to existing model selection approaches which need strong knowledge about the system and its alternatives (Azeloglu and Iyengar, 2015). However, model misspecifications are frequent at this point and will yield a lack of fit to the experimental data. Similar to existing approaches which are dealing with optimal experimental design, the proposed methods guide towards points in the given model, where the model is likely erroneous. These erroneous points can then be linked to known mechanisms and thus lead to follow-up experiments. Taking into account information about the accessibility of the involved molecules and their combinations, the proposed methods could be extended towards a full automatic ranking algorithm for optimal follow-up experiments.

As demonstrated in Chapters 6 and 7, both the DEN and the BDEN have limitations. These are due to structurally and practical unidentifiable structures in dynamical systems, as discussed in Sections 4.2 and 4.4. Linearization techniques or model reduction are a common starting point to address these issues with the final goal to determine the targetable hidden influences in the given biological system (Gao et al., 2014). Due to the fact that linear controllability and observability is a prerequisite of nonlinear control and observability, linearization of dynamical systems is commonly used for this purpose (Gao et al., 2014). Furthermore, the controllability of nonlinear systems is often structurally similar to linearized systems (Gao et al., 2014). As a deterministic approach, the DEN may also suffer from multiple solutions caused by non-identifiable hidden input signals. The BDEN addresses this issue by estimating distributions rather than point estimates. In case of the DEN, this can be addressed by the adaption of multiple shooting approaches. In terms of optimization, the proposed objective function may lead to several minima which are almost equivalent and thus the DEN may get stuck in one of those extrema. In consequence, it is not clear whether the solution represents a unique global minimum. In addition, the DEN is sensitive to the regularization parameters for smoothness and sparsity. In contrast, due to its full Bayesian nature, the BDEN also optimizes the involved regularization parameters. In general, a statement about the convexity of the problem cannot be given. However, convexity would imply the existence of a global minimum. To circumvent these issues, similar to multiple shooting approaches, the DEN could be initialized with different starting values. Here, several minima would lead to different solutions. In the final analysis, this could serve as an indicator for multiple solutions and in consequence to rank possible hidden influences with respect to optimality. Here the fit to the data, i.e., the LSE, can be used as an adequate measure for optimality.

Ultimately, an holistic approach of parameter estimation with respect to model correction in the flavor of respectful modeling is desirable. In contrast to existing state-of-the-art methods, overcomplicated models and thus excessive parameter fitting could be circumvented in that way. Although the BDEN, as well as the DEN, are robust against a certain level of errors of kinetic parameter estimates, parameters and the estimated hidden input signals are strongly dependent. Together with the task of avoiding structurally unidentifiable parameters, this is a non-trivial and mathematically complex issue and thus further research is required. In that context it is important to emphasize again that BDEN is not designed to learn ODE systems purely from data and should thus not be confused with network reverse engineering methods. However, one possible approach could be a combination of the BDEN and parameter re-estimation. But a priori it is not clear how to choose a regularization term in order to prefer either additional hidden influences or parameter re-estimation.

In Chapters 6 and 7, further applications of the DEN and the BDEN to other fields of research are already discussed. In addition, the proposed DEN, as well as the BDEN, are highly related to data assimilation approaches (Leeuwen et al., 2015). Data assimilation is commonly used in meteorology to archive reliable weather forecasts (Abarbanel, 2013). It aims for making forecasts based on given observations and mostly ODE-based model structures (Leeuwen et al., 2015). To obtain the current state of the system, a given model with several uncertainties has to be integrated (Leeuwen et al., 2015). Similar to the methods presented in this work, this yields an inverse problem because typically only a fraction of the current states are observable (Leeuwen et al., 2015). In consequence, the system is divided into known and unknown parts (Leeuwen et al., 2015). This could be a further promising application of the presented methods. Basically, one would try to determine the dynamics of the unknown part of these models to finally model them explicitly. The explicit modeling of the unknown system dynamics allows for a more effective integration of the system as a first step towards a holistic investigation of the involved black box models.

Altogether, this work represents a further step towards the knowledge- and data-driven modeling of realistic biological systems. CFS serves as a method for dealing with static models in cases were data is only partially available. In contrast, both the DEN and the BDEN are promising mathematical approaches with strong theoretical background to support the data-driven reliable modeling of dynamical systems. The intention of this work was to present supporting methods for practical research rather than pure in silico modeling tools. Especially the aspect of learning from the errors of models of biological systems is an alternative to existing approaches for model development and correction across a wide range of scientific areas.

Bibliography

H. D. I. Abarbanel. *Predicting the future: completing models of observed complex systems*. Understanding complex systems. Springer, New York; Heidelberg; Dordrecht; London, 1st edition, 2013. ISBN 978-1-4614-7217-9 978-1-4614-7218-6.

Jürgen Adamy. *Nichtlineare Systeme und Regelungen*. Springer, Berlin; Heidelberg, 2nd edition, July 2014. ISBN 978-3-642-45012-9.

Bruce Alberts, Alexander Johnson, Julian Lewis, David Morgan, Martin Raff, Keith Roberts, and Peter Walter. *Molecular Biology of the Cell*. Garland Pub, New York, 6th edition, November 2014. ISBN 978-0-8153-4524-4.

Uri Alon. *An introduction to systems biology: design principles of biological circuits*. Number 10 in Chapman & Hall/CRC mathematical and computational biology series. Chapman & Hall/CRC, Boca Raton, FL, 1st edition, 2007. ISBN 978-1-58488-642-6.

Patrick Aloy and Robert B. Russell. Structural systems biology: modelling protein interactions. *Nature Reviews Molecular Cell Biology*, 7(3):188–197, March 2006. doi: 10.1038/nrm1859.

Philipp M. Altrock, Lin L. Liu, and Franziska Michor. The mathematics of cancer: integrating quantitative models. *Nature Reviews Cancer*, 15(12):730–745, November 2015. doi: 10.1038/nrc4029.

Christophe Andrieu and Johannes Thoms. A tutorial on adaptive mcmc. *Statistics and Computing*, 18(4):343–373, December 2008. doi: 10.1007/s11222-008-9110-y.

Johannes Antony, Kerstin Kellershohn, Marion Mohr-Andrä, Anna Kebig, Stefanie Prilla, Mathias Muth, Eberhard Heller, Teresa Disingrini, Clelia Dallanoce, Simona Bertoni, Jasmin Schrobang, Christian Tränkle, Evi Kostenis, Arthur Christopoulos, Hans Dieter Höltje, Elisabetta Barocelli, Marco De Amici, Ulrike Holzgrabe, and Klaus Mohr. Dualsteric gpcr targeting: a novel route to binding and signaling pathway selectivity. *FASEB J*, 23(2):442–450, February 2009. doi: 10.1096/fj.08-114751.

Maksat Ashyraliyev, Yves Fomekong-Nanfack, Jaap A. Kaandorp, and Joke G. Blom. Systems biology: parameter estimation for biochemical models. *FEBS Journal*, 276(4):886–902, February 2009. doi: 10.1111/j.1742-4658.2008.06844.x.

A. C. Atkinson. *Plots, Transformations, And Regression: An Introduction to Graphical Methods of Diagnostic Regression Analysis*. Oxford Statistical Science Series, 1. Oxford University Press, New York, 1st edition, 1987. ISBN 978-0198533719.

Kendall E. Atkinson. *An introduction to numerical analysis*. Wiley, New York, 2nd edition, 1989. ISBN 978-0-471-62489-9.

Evren U. Azeloglu and Ravi Iyengar. Good practices for building dynamical models in systems biology. *Science Signaling*, 8(371):fs8, April 2015. doi: 10.1126/scisignal.aab0880.

Ann C. Babtie, Paul Kirk, and Michael P. H. Stumpf. Topological sensitivity analysis for systems biology. *Proceedings of the National Academy of Sciences*, 111(52):18507–18512, December 2014. doi: 10.1073/pnas.1414026112.

Eva Balsa-Canto, Antonio A. Alonso, and Julio R. Banga. An iterative identification procedure for dynamic modeling of biochemical networks. *BMC Systems Biology*, 4(11):1–18, February 2010. doi: 10.1186/1752-0509-4-11.

Julio R. Banga, Eva Balsa-Canto, Carmen G. Moles, and Antonio A. Alonso. Dynamic optimization of bioprocesses: Efficient and robust numerical strategies. *Journal of Biotechnology*, 117(4):407–419, June 2005. doi: 10.1016/j.jbiotec.2005.02.013.

Baruch Barzel, Yang Yu Liu, and Albert Lászlá Barabási. Constructing minimal models for complex system dynamics. *Nature Communications*, 6(7186):1–8, May 2015. doi: 10.1038/ncomms8186.

Heinz Bauer. *Maß- und Integrationstheorie*. de Gruyter, Berlin; New York, 2nd edition, January 1992. ISBN 978-3-11-087173-9.

Heinz Bauer. *Wahrscheinlichkeitstheorie*. de Gruyter, Berlin; New York, 5th edition, 2002. ISBN 978-3-11-017236-2.

Scott A. Becker and Bernhard Ø. Palsson. Context-specific metabolic networks are consistent with experiments. *PLoS Computational Biology*, 4(5):1–10, May 2008. doi: 10.1371/journal.pcbi.1000082.

V. Becker, M. Schilling, J. Bachmann, U. Baumann, A. Raue, T. Maiwald, J. Timmer, and U. Klingmüller. Covering a broad dynamic range: Information processing at the erythropoietin receptor. *Science*, 328(5984):1404–1408, June 2010. doi: 10.1126/science.1184913.

Jörn Behre and Stefan Schuster. Modeling signal transduction in enzyme cascades with the concept of elementary flux modes. *Journal of Computational Biology*, 16(6):829–844, June 2009. doi: 10.1089/cmb.2008.0177.

Yoav Benjamini and Yosef Hochberg. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B*, 57 (1):289–300, March 1995.

Jeremy M. Berg, John L. Tymoczko, Lubert Stryer, and Gregory J. Gatto. *Biochemie*. Springer Spektrum, Heidelberg, 7th edition, 2013. ISBN 978-3-8274-2988-9 3-8274-2988-9.

Steven Berline and Clark Bricker. The law of mass action. *Journal of Chemical Education*, 46 (8):499–501, August 1969. doi: 10.1021/ed046p499.

Michael J. Berridge. Module 2: Cell signalling pathway. *Cell Signalling Biology*, 2:csb0001002, April 2012. doi: 10.1042/csb0001002.

John T. Betts. *Practical Methods for Optimal Control and Estimation Using Nonlinear Programming*. Society for Industrial and Applied Mathematics, Philadelphia, 2nd edition, 2009. ISBN 978-0-89871-688-7.

Christopher Bishop. *Pattern Recognition and Machine Learning*. Springer, New York, 1st edition, 2007. ISBN 978-0-387-31073-2.

M. Blankenburg, L. Haberland, H. D. Elvers, C. Tannert, and B. Jandrig. High-throughput omics technologies: Potential tools for the investigation of influences of emf on biological systems. *Current Genomics*, 10(2):86–92, April 2009. doi: 10.2174/138920209787847050.

Andreas Bock, Brian Chirinda, Fabian Krebs, Regina Messerer, Julia Bätz, Mathias Muth, Clelia Dallanoce, Dominika Klingenthal, Christian Tränkle, Carsten Hoffmann, Marco De Amici, Ulrike Holzgrabe, Evi Kostenis, and Klaus Mohr. Dynamic ligand binding dictates partial agonism at a g protein-coupled receptor. *Nature Chemical Biology*, 10(1):18–20, January 2014. doi: 10.1038/nchembio.1384.

T. I. Bonner, N. J. Buckley, A. C. Young, and M. R. Brann. Identification of a family of muscarinic acetylcholine receptor genes. *Science*, 237(4814):527–532, July 1987. doi: 10.1126/science.3037705.

Aarash Bordbar, Jonathan M. Monk, Zachary A. King, and Bernhard Ø. Palsson. Constraint-based models predict metabolic and associated cellular functions. *Nature Reviews Genetics*, 15(2):107–120, January 2014. doi: 10.1038/nrg3643.

Katherine Borkovich and Daniel Ebbole. *Cellular and Molecular Biology of Filamentous Fungi*. ASM Press, Washington, DC, 1st edition, July 2010. ISBN 978-1-55581-473-1.

S. J. Bornheimer, M. R. Maurya, M. G. Farquhar, and S. Subramaniam. Computational modeling reveals how interplay between components of a gtpase-cycle module regulates signal transduction. *Proceedings of the National Academy of Sciences*, 101(45):15899–15904, November 2004. doi: 10.1073/pnas.0407009101.

Victoria Boswell-Smith, Domenico Spina, and Clive P. Page. Phosphodiesterase inhibitors. *British Journal of Clinical Pharmacology*, 147(Suppl 1):S252–S257, January 2006. doi: 10.1038/sj.bjp.0706495.

George Edward Briggs and John Burdon Sanderson Haldane. A note on the kinetics of enzyme action. *Biochemical Journal*, 19(2):338–339, January 1925. doi: 10.1042/bj0190338.

Ana Rita Brochado, Claudia Matos, Birger L. Møller, Jørgen Hansen, Uffe H. Mortensen, and Kiran Raosaheb Patil. Improved vanillin production in baker's yeast through in silico design. *Microbial Cell Factories*, 9(84):1–15, November 2010. doi: 10.1186/1475-2859-9-84.

Otto Erich Brodde and Martin C. Michel. Adrenergic and muscarinic receptors in the human heart. *Pharmacological reviews*, 51(4):651–690, December 1999.

Otto Erich Brodde, Heike Bruck, and Kirsten Leineweber. Cardiac adrenoceptors: physiological and pathophysiological relevance. *Journal of Pharmacological Sciences*, 100(5):323–337, April 2006. doi: 10.1254/jphs.CRJ06001X.

Stephen Brooks. Markov chain monte carlo method and its application. *Journal of the Royal Statistical Society: Series D*, 47(1):69–100, March 1998. doi: 10.1111/1467-9884.00117.

Steve Brooks, Andrew Gelman, Galin Jones, and Xiao Li Meng. *Handbook of Markov Chain Monte Carlo*. Chapman & Hall/CRC, Boca Raton; London; New York, 1st edition, 2011. ISBN 978-1-4200-7941-8.

Frank J. Bruggeman and Hans V. Westerhoff. The nature of systems biology. *Trends in Microbiology*, 15(1):45–50, January 2007. doi: 10.1016/j.tim.2006.11.003.

David B. Bylund. 7-transmembrane receptors. In *xPharm: The Comprehensive Pharmacology Reference*, pages 1–4. Elsevier, New York, 2007. ISBN 978-0-08-055232-3.

Cristian S. Calude and Giuseppe Longo. The deluge of spurious correlations in big data. *Foundations of Science*, pages 1–18, March 2016. doi: 10.1007/s10699-016-9489-4.

G. Casella. Empirical bayes gibbs sampling. *Biostatistics*, 2(4):485–500, December 2001. doi: 10.1093/biostatistics/2.4.485.

C. Chen, Y. Cordeaux, S. J. Hill, and J. R. King. Modelling of signalling via g-protein coupled receptors: pathway-dependent agonist potency and efficacy. *Bulletin of Mathematical Biology*, 65(5):933–958, September 2003. doi: 10.1016/S0092-8240(03)00055-7.

T. Chen, H. L. He, and G. M. Church. Modeling gene expression with differential equations. *Pacific Symposium on Biocomputing. Pacific Symposium on Biocomputing*, pages 29–40, 1999.

Eileen R. Choffnes, David A. Relman, Leslie A. Pray, and Institute of Medicine (U.S.), editors. *The science and applications of synthetic and systems biology: workshop summary*. National Academies Press, Washington, 2011. ISBN 978-0-309-21939-6.

Han Yu Chuang, Matan Hofree, and Trey Ideker. A decade of systems biology. *Annual Review of Cell and Developmental Biology*, 26:721–744, November 2010. doi: 10.1146/annurev-cellbio-100109-104122.

Bruce L. Clarke. Stability of complex reaction networks. In I. Prigogine and Stuart A. Rice, editors, *Advances in Chemical Physics*, volume 43, pages 1–215. John Wiley & Sons, Inc., Hoboken, January 1980. ISBN 978-0-470-14262-2 978-0-471-05741-3.

Matthew C. Coleman and David E. Block. Bayesian parameter estimation with informative priors for nonlinear systems. *AIChE Journal*, 52(2):651–667, February 2006. doi: 10.1002/aic.10667.

M. Cong, S. J. Perry, F. T. Lin, I. D. Fraser, L. A. Hu, W. Chen, J. A. Pitcher, J. D. Scott, and R. J. Lefkowitz. Regulation of membrane targeting of the g protein-coupled receptor kinase 2 by protein kinase a and its anchoring protein akap79. *The Journal of Biological Chemistry*, 276(18):15192–15199, May 2001. doi: 10.1074/jbc.M009130200.

International Human Genome Sequencing Consortium. Finishing the euchromatic sequence of the human genome. *Nature*, 431(7011):931–945, October 2004. doi: 10.1038/nature03001.

Dermot M. F. Cooper. Regulation and organization of adenylyl cyclases and camp. *Biochemical Journal*, 375(Pt 3):517–529, November 2003. doi: 10.1042/BJ20031061.

J. D. Corbin, C. E. Cobb, S. J. Beebe, D. K. Granner, S. R. Koch, T. W. Gettys, P. F. Blackmore, S. H. Francis, and J. N. Wells. Mechanism and function of camp - and cgmp-dependent protein kinases. *Advances in Second Messenger and Phosphoprotein Research*, 21:75–86, 1988. ISSN 1040-7952.

John L. Crassidis and John L. Junkins. *Optimal Estimation of Dynamic Systems*. Chapman & Hall/CRC, Boca Raton; London; New York, 2nd edition, 2011. ISBN 978-1-4398-3985-0.

Benjamin F. Cravatt, Gabriel M. Simon, and John R. Yates III. The biological impact of mass-spectrometry-based proteomics. *Nature*, 450(7172):991–1000, December 2007. doi: 10.1038/nature06525.

Francis Crick. Central dogma of molecular biology. *Nature*, 227(5258):561–563, August 1970. doi: 10.1038/227561a0.

Maureen M. Dale and Humphrey P. Rang. *Rang and Dale's pharmacology*. Elsevier, Churchill Livingstone, Edinburgh, 7th edition, 2011. ISBN 978-0-7020-3471-8 0-7020-3471-1.

Maria De Jesus and Florian M. Wurm. Manufacturing recombinant proteins in kg-ton quantities using animal cells in bioreactors. *European Journal of Pharmaceutics and Biopharmaceutics*, 78(2):184–188, June 2011. doi: 10.1016/j.ejpb.2011.01.005.

L. De Vries, B. Zheng, T. Fischer, E. Elenko, and M. G. Farquhar. The regulator of g protein signaling family. *Annual Review of Pharmacology and Toxicology*, 40:235–271, April 2000. doi: 10.1146/annurev.pharmtox.40.1.235.

Gustavo Egea, Francisco Lázaro-Diéguez, and Montserrat Vilella. Actin dynamics at the golgi complex in mammalian cells. *Current Opinion in Cell Biology*, 18(2):168–178, April 2006. doi: 10.1016/j.ceb.2006.02.007.

Niels Eijkelkamp, Huijing Wang, Anibal Garza-Carbajal, Hanneke L. D. M. Willemen, Fried J. Zwartkruis, John N. Wood, Robert Dantzer, Keith W. Kelley, Cobi J. Heijnen, and Annemieke Kavelaars. Low nociceptor grk2 prolongs prostaglandin e2 hyperalgesia via biased camp signaling to epac\rap1, protein kinase cepsilon, and mek \erk. *Journal of Neuroscience*, 30 (38):12806–12815, September 2010. doi: 10.1523/JNEUROSCI.3142-10.2010.

D. Eisenberg, E. M. Marcotte, I. Xenarios, and T. O. Yeates. Protein function in the post-genomic era. *Nature*, 405(6788):823–826, June 2000. doi: 10.1038/35015694.

Ahmed Elhusseiny and Edith Hamel. Muscarinic-but not nicotinic-acetylcholine receptors mediate a nitric oxide-dependent dilation in brain cortical arterioles: A possible role for the m5 receptor subtype. *Journal of Cerebral Blood Flow & Metabolism*, 20(2):298–305, February 2000. doi: 10.1097/00004647-200002000-00011.

Benjamin Engelhardt, Holger Fröhlich, and Maik Kschischo. Learning (from) the errors of a systems biology model. *Natur3 Scientific Reports*, 6(20772):1–9, February 2016. doi: 10.1038/srep20772.

Benjamin Engelhardt, Janine Holze, Christina Elliott, George S. Baillie, Maik Kschischo, and Holger Fröhlich. Modeling and mathematical analysis of the m2 receptor-dependent joint signaling and secondary messenger network in cho cells. *Mathematics Applied in Medicine and Biology*, accepted, 2017.

Sandrine Etienne-Manneville. Actin and microtubules in cell motility: Which one is in control? *Traffic*, 5(7):470–477, July 2004. doi: 10.1111/j.1600-0854.2004.00196.x.

Ye Fang, Ann M. Ferrie, Norman H. Fontaine, John Mauro, and Jitendra Balakrishnan. Resonant waveguide grating biosensor for living cell sensing. *Biophysical Journal*, 91(5):1925–1940, September 2006. doi: 10.1529/biophysj.105.077818.

Ye Fang, Terry Kenakin, and Changlu Liu. Editorial: Orphan gpcrs as emerging drug targets. *Frontiers in Pharmacology*, 6(295):1–2, December 2015. doi: 10.3389/fphar.2015.00295.

T. Fawcett. An introduction to roc analysis. *Pattern Recognition Letters*, 27(8):861–874, February 2006. doi: 10.1016/j.patrec.2005.10.010.

Gerd Fischer. *Lineare Algebra: Eine Einführung für Studienanfänger*. Springer Spektrum, Wiesbaden, 18th edition, 2013. ISBN 978-3-658-03944-8.

Wendell H. Fleming and Raymond W. Rishel. *Deterministic and stochastic optimal control*. Applications of mathematics; 1. Springer-Verlag, Berlin; New York, 1st edition, 1975. ISBN 978-0-387-90155-8.

Nir Friedman, Kevin Murphy, and Stuart Russell. Learning the structure of dynamic probabilistic networks. In *Proceedings of the Fourteenth Conference on Uncertainty in Artificial Intelligence*, UAI'98, pages 139–147, San Francisco, 1998. Morgan Kaufmann Publishers Inc. ISBN 1-55860-555-X.

F. N. Fritsch and R. E. Carlson. Monotone piecewise cubic interpolation. *SIAM Journal on Numerical Analysis*, 17(2):238–246, April 1980. doi: 10.1137/0717021.

Robert E. Fuguitt and J. Erskine Hawkins. Rate of the Thermal Isomerization of α-pinene in the liquid phase. *Journal of the American Chemical Society*, 69(2):319–322, February 1947. doi: 10.1021/ja01194a047.

Jianxi Gao, Yang Yu Liu, Raissa M. D'Souza, and Albert László Barabási. Target control of complex networks. *Nature Communications*, 5(5415):1–8, November 2014. doi: 10.1038/ncomms6415.

Andrew Gelman, John B. Carlin, Hal S. Stern, David B. Dunson, Aki Vehtari, and Donald B. Rubin. *Bayesian Data Analysis*. Chapman & Hall/CRC, Boca Raton; London; New York, 3rd edition, November 2013. ISBN 978-1-4398-4095-5.

Matthias Gerdts. *Optimal control of ODEs and DAEs*. De Gruyter textbook. De Gruyter, Berlin; Boston, 1st edition, 2012. ISBN 978-3-11-024995-8.

Qi Gong, Fariba Fahroo, and I. Michael Ross. Spectral algorithm for pseudospectral methods in optimal control. *Journal of Guidance, Control, and Dynamics*, 31(3):460–471, May 2008. doi: 10.2514/1.32908.

Vsevolod V. Gurevich and Eugenia V. Gurevich. Gpcr monomers and oligomers: it takes all kinds. *Trends in Neurosciences*, 31(2):74–81, February 2008. doi: 10.1016/j.tins.2007.11.007.

Jennetta W. Hammond, Dawen Cai, and Kristen J. Verhey. Tubulin modifications and their cellular functions. *Current Opinion in Cell Biology*, 20(1):71–76, February 2008. doi: 10.1016/j.ceb.2007.11.010.

Martin Hanke-Bourgeois. *Grundlagen der Numerischen Mathematik und des Wissenschaftlichen Rechnens*. Vieweg+Teubner Verlag, Wiesbaden, 3rd edition, 2008. ISBN 978-3-8348-0708-3.

R. Heinrich and S. Schuster. *The Regulation of Cellular System*. Chapman & Hall, New York, 1st edition, 1996. ISBN 978-0412032615.

Reinhart Heinrich, Benjamin G. Neel, and Tom A. Rapoport. Mathematical models of protein kinase signal transduction. *Molecular Cell*, 9(5):957–970, May 2002. doi: 10.1016/S1097-2765(02)00528-2.

R. Hermann and A. Krener. Nonlinear controllability and observability. *IEEE Transactions on Automatic Control*, 22(5):728–740, October 1977. doi: 10.1109/TAC.1977.1101601.

Maria A. Hernandez and Appu Rathinavelu. *Basic Pharmacology: Understanding Drug Actions and Reactions*. CRC Press, Boca Raton; London; New York, 1st edition, February 2006. ISBN 978-1-58716-160-5.

Peter D. Hoff. *A First Course in Bayesian Statistical Methods*. Springer, Heidelberg; London; New York, 1st edition, 2010. ISBN 978-0-387-92299-7.

Myles Hollander. *Nonparametric statistical methods*. Wiley series in probability and statistics. Wiley-Interscience, New York, 2nd edition, 1999. ISBN 978-0471190455.

Susanne Hollinger, Suneela Ramineni, and John R. Hepler. Phosphorylation of rgs14 by protein kinase a potentiates its activity toward g alpha i. *Biochemistry*, 42(3):811–819, January 2003.

Josef Honerkamp and Björn Olaf Schelter. *Statistical physics: an advanced approach with applications*. Graduate texts in physics. Springer, Berlin, 3rd edition, 2014. ISBN 978-3-642-28684-1.

Leroy Hood and David Galas. The digital code of dna. *Nature*, 421(6921):444–448, January 2003. doi: 10.1038/nature01410.

Robert Hooke. *Micrographia, or, Some physiological descriptions of minute bodies made by magnifying glasses*. Martyn, J. and Allestry, J., London, 1st edition, 1665. doi: 10.5962/bhl.title.904.

F. Horn and R. Jackson. General mass action kinetics. *Archive for Rational Mechanics and Analysis*, 47(2):81–116, January 1972. doi: 10.1007/BF00251225.

Alan K. Howe. Regulation of actin-based cell migration by camp/pka. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research*, 1692(2-3):159–174, July 2004. doi: 10.1016/j.bbamcr.2004.03.005.

Jianxin Hu, Yan Wang, Xiaohong Zhang, John R. Lloyd, Jian Hua Li, Joel Karpiak, Stefano Costanzi, and Jürgen Wess. Structural basis of g protein-coupled receptor-g protein interactions. *Nature Chemical Biology*, 6(7):541–548, July 2010. doi: 10.1038/nchembio.385.

Katrin Hübner, Sven Sahle, and Ursula Kummer. Applications and trends in systems biology in biochemistry: Systems biology in biochemical research. *FEBS Journal*, 278(16):2767–2857, August 2011. doi: 10.1111/j.1742-4658.2011.08217.x.

S. Hug, A. Raue, J. Hasenauer, J. Bachmann, U. Klingmüller, J. Timmer, and F.J. Theis. High-dimensional bayesian parameter estimation: Case study for a model of jak2/stat5 signaling. *Mathematical Biosciences*, 246(2):293–304, December 2013. doi: 10.1016/j.mbs.2013.04.002.

James H. Hurley. Structure, mechanism, and regulation of mammalian adenylyl cyclase. *Journal of Biological Chemistry*, 274(12):7599–7602, March 1999. doi: 10.1074/jbc.274.12.7599.

J. Jakubík, T. Haga, and S. Tucek. Effects of an agonist, allosteric modulator, and antagonist on guanosine-gamma-[35s]thiotriphosphate binding to liposomes with varying muscarinic receptor/go protein stoichiometry. *Molecular Pharmacology*, 54(5):899–906, November 1998. doi: 10.1124/mol.54.5.899.

Khuloud Jaqaman and Gaudenz Danuser. Linking data to models: data regression. *Nature Reviews Molecular Cell Biology*, 7(11):813–819, November 2006. doi: 10.1038/nrm2030.

Florian Jarre and Josef Stoer. *Optimierung*. Springer-Lehrbuch. Springer, Berlin; Heidelberg; New York, 1st edition, 2004. ISBN 978-3-540-43575-4.

Livnat Jerby, Tomer Shlomi, and Eytan Ruppin. Computational reconstruction of tissue-specific metabolic models: application to human liver metabolism. *Molecular Systems Biology*, 6 (401):1–9, September 2010. doi: 10.1038/msb.2010.56.

Andrew R. Joyce and Bernhard Ø. Palsson. The model organism as a system: integrating ’omics’ data sets. *Nature Reviews Molecular Cell Biology*, 7(3):198–210, March 2006. doi: 10.1038/nrm1857.

Simon J. Julier and Jeffrey K. Uhlmann. New extension of the kalman filter to nonlinear systems. In *Signal Processing, Sensor Fusion, and Target Recognition VI*, volume 3068, pages 182–193, July 1997. doi: 10.1117/12.280797.

Björn H. Junker and Falk Schreiber. *Analysis of Biological Networks*. Wiley-Interscience, Hoboken, 1st edition, March 2008. ISBN 978-0-470-04144-4.

N. Juty, R. Ali, M. Glont, S. Keating, N. Rodriguez, M. J. Swat, Sm Wimalaratne, H. Hermjakob, N. Le Novère, C. Laibe, and V. Chelliah. Biomodels: Content, features, functionality, and use. *CPT: Pharmacometrics & Systems Pharmacology*, 4(2):55–68, February 2015. doi: 10.1002/psp4.3.

Matthias Kahm, Clara Navarrete, Vicent Llopis-Torregrosa, Rito Herrera, Lina Barreto, Lynne Yenush, Joaquin Ariño, Jose Ramos, and Maik Kschischo. Potassium starvation in yeast: Mechanisms of homeostasis revealed by mathematical modeling. *PLoS Computational Biology*, 8(6):1–11, June 2012. doi: 10.1371/journal.pcbi.1002548.

R. E. Kalman. A new approach to linear filtering and prediction problems. *Transactions of the AMSE, Part D, Journal of Basic Engineering*, 82(1):35–45, March 1960.

Minoru Kanehisa, Yoko Sato, Masayuki Kawashima, Miho Furumichi, and Mao Tanabe. Kegg as a reference resource for gene and protein annotation. *Nucleic Acids Research*, 44(D1): D457–D462, January 2016. doi: 10.1093/nar/gkv1070.

Guy Karlebach and Ron Shamir. Modelling and analysis of gene regulatory networks. *Nature Reviews Molecular Cell Biology*, 9(10):770–780, October 2008. doi: 10.1038/nrm2503.

Jonathan R. Karr, Jayodita C. Sanghvi, Derek N. Macklin, Miriam V. Gutschow, Jared M. Jacobs, Benjamin Bolival, Nacyra Assad-Garcia, John I. Glass, and Markus W. Covert. A whole-cell computational model predicts phenotype from genotype. *Cell*, 150(2):389–401, July 2012. doi: 10.1016/j.cell.2012.05.044.

Jonathan R. Karr, Alex H. Williams, Jeremy D. Zucker, Andreas Raue, Bernhard Steiert, Jens Timmer, Clemens Kreutz, DREAM8 Parameter Estimation Challenge Consortium, Simon Wilkinson, Brandon A. Allgood, Brian M. Bot, Bruce R. Hoff, Michael R. Kellen, Markus W. Covert, Gustavo A. Stolovitzky, and Pablo Meyer. Summary of the dream8 parameter estimation challenge: Toward parameter identification for whole-cell models. *PLOS Computational Biology*, 11(5):1–21, May 2015. doi: 10.1371/journal.pcbi.1004096.

Daniel Kaschek, Wolfgang Mader, Mirjam Fehling-Kaschek, Marcus Rosenblatt, and Jens Timmer. Dynamic modeling, parameter estimation and uncertainty analysis in r. *bioRxiv*, pages 1–31, November 2016. doi: 10.1101/085001.

Vsevolod Katritch, Vadim Cherezov, and Raymond C. Stevens. Structure-function of the g-protein-coupled receptor superfamily. *Annual Review of Pharmacology and Toxicology*, 53 (1):531–556, January 2013. doi: 10.1146/annurev-pharmtox-032112-135923.

David E. Kaufman and Robert L. Smith. Direction choice for accelerated convergence in hit-and-run sampling. *Operations Research*, 46(1):84–95, February 1998. doi: 10.1287/opre.46.1.84.

Anna Elisabeth Kebig. *Signalwegsaktivierung an muskarinischen M2-Rezeptoren: Einfluss der allosterischen Bindungsstelle*. PhD thesis, Rheinische Friedrich-Wilhelms-Universität Bonn, 2010.

James P. Keener and James Sneyd. *Mathematical physiology*. Number 8 in Interdisciplinary applied mathematics. Springer, New York, 2nd edition, 2009. ISBN 978-0-387-09419-9.

Hans A. Kestler, Christian Wawra, Barbara Kracher, and Michael Kühl. Network modeling of signal transduction: establishing the global view. *BioEssays*, 30(11-12):1110–1125, November 2008. doi: 10.1002/bies.20834.

Boris N. Kholodenko. Cell-signalling dynamics in time and space. *Nature Reviews Molecular Cell Biology*, 7(3):165–176, February 2006. doi: 10.1038/nrm1838.

Adam J. Kimple, Dustin E. Bosch, Patrick M. Giguòre, and David P. Siderovski. Regulators of g-protein signaling and their gαsubstrates: Promises and challenges in their use as drug discovery targets. *Pharmacological Reviews*, 63(3):728–749, September 2011. doi: 10.1124/pr.110.003038.

Hiroaki Kitano. Computational systems biology. *Nature*, 420(6912):206–210, November 2002. doi: 10.1038/nature01254.

Steffen Klamt, Julio Saez-Rodriguez, Jonathan A. Lindquist, Luca Simeoni, and Ernst D. Gilles. A methodology for the structural and functional analysis of signaling and regulatory networks. *BMC Bioinformatics*, 7(56):1–26, February 2006. doi: 10.1186/1471-2105-7-56.

Dorothea Klemt. *Allosterische Modulation der Pilocarpin-induzierten G Protein-Aktivierung am muskarinischen M2- und M4-Acetylcholin-Rezeptor*. PhD thesis, Rheinische Friedrich-Wilhelms-Universität Bonn, 2005.

Edda Klipp and Wolfram Liebermeister. Mathematical modeling of intracellular signaling pathways. *BMC Neuroscience*, 7(Suppl 1):1–16, October 2006. doi: 10.1186/1471-2202-7-S1-S10.

Wlodzimierz Klonowski. Simplifying principles for chemical and enzyme reaction kinetics. *Biophysical Chemistry*, 18(2):73–87, September 1983. doi: 10.1016/0301-4622(83)85001-7.

Jason R. Kolodziej and Joseph D. Mook. A novel approach to model determination using the minimum model error estimation. In *Proceedings of the American Control Conference, 2005*, volume 7, Portland, June 2005. IEEE Service Center. doi: 10.1109/ACC.2005.1470812.

Andreas Kremling. *Kompendium Systembiologie: mathematische Modellierung und Modellanalyse*. Studium. Vieweg + Teubner, Wiesbaden, 1st edition, 2012. ISBN 978-3-8348-1907-9.

Clemens Kreutz, Andreas Raue, Daniel Kaschek, and Jens Timmer. Profile likelihood in systems biology. *FEBS Journal*, 280(11):2564–2571, June 2013. doi: 10.1111/febs.12276.

Andrew C. Kruse, Aaron M. Ring, Aashish Manglik, Jianxin Hu, Kelly Hu, Katrin Eitel, Harald Hübner, Els Pardon, Celine Valant, Patrick M. Sexton, Arthur Christopoulos, Christian C. Felder, Peter Gmeiner, Jan Steyaert, William I. Weis, K. Christopher Garcia, Jürgen Wess, and Brian K. Kobilka. Activation and allosteric modulation of a muscarinic acetylcholine receptor. *Nature*, 507(7478):101 – 106, November 2013. doi: 10.1038/nature12735.

Andrew C. Kruse, Brian K. Kobilka, Dinesh Gautam, Patrick M. Sexton, Arthur Christopoulos, and Jürgen Wess. Muscarinic acetylcholine receptors: novel opportunities for drug development. *Nature Reviews Drug Discovery*, 13(7):549–560, June 2014. doi: 10.1038/nrd4295.

Minjung Kyung, Jeff Gill, Malay Ghosh, and George Casella. Penalized regression, standard errors, and bayesian lassos. *Bayesian Analysis*, 5(2):369–411, June 2010. doi: 10.1214/10-BA607.

Eric S. Lander et al. Initial sequencing and analysis of the human genome. *Nature*, 409(6822): 860–921, February 2001. doi: 10.1038/35057062.

Peter Jan Van Leeuwen, Yuan Cheng, and Sebastian Reich. *Nonlinear Data Assimilation*. Springer, Heidelberg; New York; London, 1st edition, July 2015. ISBN 978-3-319-18346-6.

Albert L. Lehninger, David L. Nelson, and Michael M. Cox. *Principles of biochemistry*. W. H. Freeman, New York, 4th edition, 2005. ISBN 978-0-7167-4339-2.

Nathan E. Lewis, Gunnar Schramm, Aarash Bordbar, Jan Schellenberger, Michael P. Andersen, Jeffrey K. Cheng, Nilam Patel, Alex Yee, Randall A. Lewis, Roland Eils, Rainer König, and Bernhard Ø. Palsson. Large-scale in silico modeling of metabolic interactions between cell types in the human brain. *Nature Biotechnology*, 28(12):1279–1285, December 2010. doi: 10.1038/nbt.1711.

Nathan E. Lewis, Harish Nagarajan, and Bernhard Ø. Palsson. Constraining the metabolic genotype-phenotype relationship using a phylogeny of in silico methods. *Nature Reviews Microbiology*, 10(4):291–305, February 2012. doi: 10.1038/nrmicro2737.

Chen Li, Marco Donizelli, Nicolas Rodriguez, Harish Dharuri, Lukas Endler, Vijayalakshmi Chelliah, Lu Li, Enuo He, Arnaud Henry, Melanie I Stefan, Jacky L Snoep, Michael Hucka, Nicolas Le Novère, and Camille Laibe. Biomodels database: An enhanced, curated and annotated resource for published quantitative kinetic models. *BMC Systems Biology*, 4(92): 1–14, June 2010. doi: 10.1186/1752-0509-4-92.

Qing Li and Nan Lin. The bayesian elastic net. *Bayesian Analysis*, 5(1):151–170, March 2010. doi: 10.1214/10-BA506.

J. J. Linderman. Modeling of g-protein-coupled receptor signaling pathways. *Journal of Biological Chemistry*, 284(9):5427–5431, February 2009. doi: 10.1074/jbc.R800028200.

Y. Y. Liu, J. J. Slotine, and Albert Lászlá Barabási. Observability of complex systems. *Proceedings of the National Academy of Sciences*, 110(7):2460–2465, February 2013. doi: 10.1073/pnas.1215508110.

Francisco Llaneras and Jesús Picó. Which metabolic pathways generate and characterize the flux space? a comparison among elementary modes, extreme pathways and minimal generators. *BioMed Research International*, 2010(753904):1–14, May 2010. doi: 10.1155/2010/753904.

L. Lou. camp inhibition of akt is mediated by activated and phosphorylated rap1b. *Journal of Biological Chemistry*, 277(36):32799–32806, August 2002. doi: 10.1074/jbc.M201491200.

David G. Luenberger. *Introduction to Dynamic Systems: Theory, Models, and Applications*. Wiley, New York, 1st edition, May 1979. ISBN 978-0-471-02594-8.

Daniel Machado, Rafael S. Costa, Miguel Rocha, Eugénio C. Ferreira, Bruce Tidor, and Isabel Rocha. Modeling formalisms in systems biology. *AMB Express*, 1(1):45, December 2011. doi: 10.1186/2191-0855-1-45.

Alejandro G. Marangoni. *Enzyme kinetics: a modern approach*. Wiley-Interscience, Hoboken, 1st edition, 2003. ISBN 978-0-471-15985-8.

Agostino Martinelli. Nonlinear unknown input observability: Analytical expression of the observable codistribution in the case of a single unknown input. In *2015 Proceedings of the Conference on Control and its Applications*, pages 9–15. Society for Industrial and Applied Mathematics, Philadelphia, PA, July 2015. ISBN 978-1-61197-407-2.

Robert Mattheij and Jaap Molenaar. *Ordinary differential equations in theory and practice*. Number 43 in Classics in applied mathematics. SIAM, c2002, Philadelphia, 1st edition, 2002. ISBN 978-0-89871-531-6.

Donald H. Maurice, Hengming Ke, Faiyaz Ahmad, Yousheng Wang, Jay Chung, and Vincent C. Manganiello. Advances in targeting cyclic nucleotide phosphodiesterases. *Nature Reviews Drug Discovery*, 13(4):290–314, April 2014. doi: 10.1038/nrd4228.

Bernhard Mayr and Marc Montminy. Transcriptional regulation by the phosphorylation-dependent factor creb. *Nature Reviews Molecular Cell Biology*, 2(8):599–609, August 2001. doi: 10.1038/35085068.

Wout Megchelenbrink, Martijn Huynen, and Elena Marchiori. optgpsampler: An improved tool for uniformly sampling the solution-space of genome-scale metabolic networks. *PLOS ONE*, 9(2):1–8, February 2014. doi: 10.1371/journal.pone.0086587.

Martin Meier-Schellersheim, Iain D. C. Fraser, and Frederick Klauschen. Multiscale modeling for biologists. *Wiley Interdisciplinary Reviews. Systems Biology and Medicine*, 1(1):4–14, August 2009. doi: 10.1002/wsbm.33.

Pablo Meyer, Thomas Cokelaer, Deepak Chandran, Kyung Kim, Po-Ru Loh, George Tucker, Mark Lipson, Bonnie Berger, Clemens Kreutz, Andreas Raue, Bernhard Steiert, Jens Timmer, Erhan Bilal, DREAMVI-VII Parameter Estimation Consortium, Herbert M. Sauro, Gustavo Stolovitzky, and Julio Saez-Rodriguez. Network topology and parameter estimation: from experimental design methods to gene regulatory network kinetics using a community based approach. *BMC Systems Biology*, 8(13):1–18, February 2014. doi: 10.1186/1752-0509-8-13.

Graeme Milligan and Evi Kostenis. Heterotrimeric g-proteins: a short history. *British Journal of Pharmacology*, 147(Suppl 1):S46–S55, January 2006. doi: 10.1038/sj.bjp.0706405.

R. Milo. Network motifs: Simple building blocks of complex networks. *Science*, 298(5594): 824–827, October 2002. doi: 10.1126/science.298.5594.824.

D. Joseph Mook and John L. Junkins. Minimum model error estimation for poorly modeled dynamic systems. *Journal of Guidance, Control, and Dynamics*, 11(3):256–261, April 1988. doi: 10.2514/3.20302.

J. F. Morrison. Kinetics of the reversible inhibition of enzyme-catalysed reactions by tight-binding inhibitors. *Biochimica et Biophysica Acta (BBA) - Enzymology*, 185(2):269–286, August 1969. doi: 10.1016/0005-2744(69)90420-3.

Allan H. Murphy. A new vector partition of the probability score. *Journal of Applied Meteorology*, 12(4):595–600, June 1973. doi: 10.1175/1520-0450(1973)012<0595:ANVPOT>2.0.CO;2.

Robert F. Murphy. An active role for machine learning in drug development. *Nature Chemical Biology*, 7(6):327–330, June 2011. doi: 10.1038/nchembio.576.

Hojung Nam, Miguel Campodonico, Aarash Bordbar, Daniel R. Hyduke, Sangwoo Kim, Daniel C. Zielinski, and Bernhard Ø. Palsson. A systems approach to predict oncometabolites via context-specific genome-scale metabolic networks. *PLoS Computational Biology*, 10(9): 1–13, September 2014. doi: 10.1371/journal.pcbi.1003837.

Miloslav Nic, Jirí Jirat, Bedrich Kosata, Aubrey Jenkins, and Alan McNaught. Elementary reaction. In *IUPAC Compendium of Chemical Terminology*. IUPAC, Research Triagle Park, NC, 2.1.0 edition, June 2009.

D. Noble. A theory of biological relativity: no privileged level of causation. *Interface Focus*, 2 (1):55–64, February 2012. doi: 10.1098/rsfs.2011.0067.

William M. Oldham and Heidi E. Hamm. Heterotrimeric g protein activation by g protein-coupled receptors. *Nature Reviews Molecular Cell Biology*, 9(1):60–71, January 2008. doi: 10.1038/nrm2299.

Kenji Omori and Jun Kotera. Overview of pdes and their regulation. *Circulation Research*, 100 (3):309–327, February 2007. doi: 10.1161/01.RES.0000256354.95791.f1.

Jeffrey D. Orth, Ines Thiele, and Bernhard Ø. Palsson. What is flux balance analysis? *Nature Biotechnology*, 28(3):245–248, March 2010. doi: 10.1038/nbt.1614.

X. Ouyang, X. Huang, X. Jin, Z. Chen, P. Yang, H. Ge, S. Li, and X. W. Deng. Coordinated photomorphogenic uv-b signaling network captured by mathematical modeling. *Proceedings of the National Academy of Sciences of the United States of America*, 111(31):11539–11544, August 2014. doi: 10.1073/pnas.1412050111.

B. Paiewonsky. Optimal control - a review of theory and practice. *AIAA Journal*, 3(11): 1985–2006, November 1965. doi: 10.2514/3.3307.

Bernhard Ø. Palsson. *Systems biology: properties of reconstructed networks*. Cambridge University Press, Cambridge; New York, 1st edition, 2006. ISBN 978-0-521-85903-5.

Jason A. Papin, Joerg Stelling, Nathan D. Price, Steffen Klamt, Stefan Schuster, and Bernhard Ø. Palsson. Comparison of network-based pathway analysis methods. *Trends in Biotechnology*, 22(8):400–405, August 2004. doi: 10.1016/j.tibtech.2004.06.010.

Trevor Park and George Casella. The bayesian lasso. *Journal of the American Statistical Association*, 103(482):681–686, June 2008. doi: 10.1198/016214508000000337.

Kristen L. Pierce, Richard T. Premont, and Robert J. Lefkowitz. Seven-transmembrane receptors. *Nature Reviews Molecular Cell Biology*, 3(9):639–650, September 2002. doi: 10.1038/nrm908.

L. S. Pontryagin, V. G Boltyanskii, R. V. Gamkrelidze, and E. F. Mishchenko. *The mathematical theory of optimal processes*. Classics of Soviet mathematics. Gordon and Breach Science Publishers, New York, 1st edition, 1986.

Paurush Praveen and Holger Fröhlich. Boosting probabilistic graphical model inference by incorporating prior knowledge from multiple sources. *PLoS ONE*, 8(6):1–10, June 2013. doi: 10.1371/journal.pone.0067410.

Nathan D. Price, Jennifer L. Reed, and Bernhard Ø. Palsson. Genome-scale models of microbial cells: evaluating the consequences of constraints. *Nature Reviews Microbiology*, 2(11): 886–897, November 2004. doi: 10.1038/nrmicro1023.

Anil V. Rao. A survey of numerical methods for optimal control. *Advances in the Astronautical Sciences*, 135(1), January 2010.

J. R. Raol, G. Girija, and J. Singh. *Modelling and parameter estimation of dynamic systems*, volume 65 of *IEE control engineering series*. Institution of Electrical Engineers, London, 1st edition, 2004. ISBN 978-0-86341-363-6.

A. Raue, C. Kreutz, T. Maiwald, J. Bachmann, M. Schilling, U. Klingmüller, and J. Timmer. Structural and practical identifiability analysis of partially observed dynamical models by exploiting the profile likelihood. *Bioinformatics*, 25(15):1923–1929, August 2009. doi: 10.1093/bioinformatics/btp358.

A. Raue, V. Becker, U. Klingmüller, and J. Timmer. Identifiability and observability analysis for experimental design in nonlinear dynamical models. *Chaos*, 20(4):1–8, December 2010. doi: 10.1063/1.3528102.

Eric Reiter and Robert J. Lefkowitz. Grks and β-arrestins: roles in receptor silencing, trafficking and signaling. *Trends in Endocrinology and Metabolism*, 17(4):159–165, May 2006. doi: 10.1016/j.tem.2006.03.008.

J. Renart, J. Reiser, and G. R. Stark. Transfer of proteins from gels to diazobenzyloxymethyl-paper and detection with antisera: a method for studying antibody specificity and antigen structure. *Proceedings of the National Academy of Sciences*, 76(7):3116–3120, July 1979. doi: 10.1073/pnas.76.7.3116.

John F. Robyt and Bernard J. White. *Biochemical techniques: theory and practice*. Waveland Press, Prospect Heights, 1st edition, 1987. ISBN 978-0-88133-556-9.

Fernando Rodrigues, Paula Ludovico, and Cecilia Leão. Sugar metabolism in yeasts: an overview of aerobic and anaerobic glucose catabolism. In Gabor Peter and Carlos Rosa, editors, *Biodiversity and Ecophysiology of Yeasts*, pages 101–121. Springer, Berlin; Heidelberg, 2006. ISBN 978-3-540-30985-7.

Maria Rodriguez-Fernandez, Markus Rehberg, Andreas Kremling, and Julio R. Banga. Simultaneous model discrimination and parameter estimation in dynamic models of cellular systems. *BMC Systems Biology*, 7(76):1–14, August 2013. doi: 10.1186/1752-0509-7-76.

D. Rosenblatt, A. Sharon, and A. A. Friesem. Resonant grating waveguide structures. *IEEE Journal of Quantum Electronics*, 33(11):2038 –2059, November 1997. doi: 10.1109/3.641320.

I. Michael Ross. A historical introduction to the convector mapping principle. In *Proceedings of Astrodynamics Specialists Conference*. Calhoun, the NPS Institutional Archive, 2005.

I. Michael Ross and Mark Karpenko. A review of pseudospectral optimal control: From theory to flight. *Annual Reviews in Control*, 36(2):182–197, December 2012. doi: 10.1016/j.arcontrol.2012.09.002.

Amelia Russo-Neustadt and Carl W. Cotman. Adrenergic receptors in alzheimer's disease brain: Selective increases in the cerebella of aggressive patients. *Journal of Neuroscience*, 17(14):5573–5580, July 1997.

Antti Saarinen, Marja Leena Linne, and Olli Yli-Harja. Stochastic differential equation model for cerebellar granule cell excitability. *PLoS Computational Biology*, 4(2):1–11, February 2008. doi: 10.1371/journal.pcbi.1000004.

K. Sachs, O. Perez, Pe'er D., D. A. Lauffenburger, and G. P. Nolan. Causal protein-signaling networks derived from multiparameter single-cell data. *Science*, 308(5721):523–529, April 2005. doi: 10.1126/science.1105809.

Natasha C. Salazar, Juhsien Chen, and Howard A. Rockman. Cardiac gpcrs: Gpcr signaling in healthy and failing hearts. *Biochimica et biophysica acta*, 1768(4):1006–1018, April 2007. doi: 10.1016/j.bbamem.2007.02.010.

Simo Särkkä. *Bayesian Filtering and Smoothing*. Cambridge University Press, Cambridge; New York, 1st edition, September 2013. ISBN 978-1-107-61928-9.

U. Sauer, M. Heinemann, and N. Zamboni. Genetics: Getting closer to the whole picture. *Science*, 316(5824):550–551, April 2007. doi: 10.1126/science.1142502.

Hermann Schägger. Tricine-sds-page. *Nature Protocols*, 1(1):16–22, June 2006. doi: 10.1038/nprot.2006.4.

M. Schauer and R. Heinrich. Quasi-steady-state approximation in the mathematical modeling of biochemical reaction networks. *Mathematical Biosciences*, 65(2):155–170, August 1983. doi: 10.1016/0025-5564(83)90058-5.

Jennifer Scheidel, Leonie Amstein, Jörg Ackermann, Ivan Dikic, and Ina Koch. In silico knockout studies of xenophagic capturing of salmonella. *PLOS Computational Biology*, 12 (12):1–18, December 2016. doi: 10.1371/journal.pcbi.1005200.

J. Schellenberger and Bernhard Ø. Palsson. Use of randomized sampling for analysis of metabolic networks. *Journal of Biological Chemistry*, 284(9):5457–5461, February 2009. doi: 10.1074/jbc.R800048200.

Anja Schmidt and Michael N. Hall. Signaling to the actin cytoskeleton. *Annual Review of Cell and Developmental Biology*, 14(1):305–338, November 1998. doi: 10.1146/annurev.cellbio.14.1.305.

R. Schrage, W. K. Seemann, J. Klöckner, C. Dallanoce, K. Racké, E. Kostenis, M. De Amici, U. Holzgrabe, and K. Mohr. Agonists with supraphysiological efficacy at the muscarinic m2 ach receptor: Supraphysiological m2 receptor activation. *British Journal of Pharmacology*, 169(2):357–370, May 2013. doi: 10.1111/bph.12003.

R. Schrage, A. De Min, K. Hochheiser, E. Kostenis, and K. Mohr. Superagonism at g protein-coupled receptors and beyond. *British Journal of Pharmacology*, 173(20):3018–3027, October 2016. doi: 10.1111/bph.13278.

Ramona Schrage, Janine Holze, Jessica Klöckner, Aileen Balkow, Anne S. Klause, Anna-Lena Schmitz, Marco De Amici, Evi Kostenis, Christian Tränkle, Ulrike Holzgrabe, and Klaus Mohr. New insight into active muscarinic receptors with the novel radioagonist [3h]iperoxo. *Biochemical Pharmacology*, 90(3):307–319, August 2014. doi: 10.1016/j.bcp.2014.05.012.

Ralf Schröder, Nicole Merten, Jesper Mosolff Mathiesen, Lene Martini, Anamarija Kruljac-Letunic, Friederike Krop, Andree Blaukat, Ye Fang, Elizabeth Tran, Trond Ulven, Christel Drewke, Jennifer Whistler, Leonardo Pardo, Jesús Gomeza, and Evi Kostenis. The c-terminal tail of crth2 is a key molecular determinant that constrains gαi and downstream signaling cascade activation. *Journal of Biological Chemistry*, 284(2):1324–1336, September 2009. doi: 10.1074/jbc.M806867200.

Ralf Schröder, Johannes Schmidt, Stefanie Blättermann, Lucas Peters, Nicole Janssen, Manuel Grundmann, Wiebke Seemann, Dorina Kaufel, Nicole Merten, Christel Drewke, Jesús Gomeza, Graeme Milligan, Klaus Mohr, and Evi Kostenis. Applying label-free dynamic mass redistribution technology to frame signaling of g protein-coupled receptors noninvasively in living cells. *Nature Protocols*, 6(11):1748–1760, October 2011. doi: 10.1038/nprot.2011.386.

Ralf Schröder, Nicole Janssen, Johannes Schmidt, Anna Kebig, Nicole Merten, Stephanie Hennen, Anke Müller, Stefanie Blättermann, Marion Mohr-Andrä, Sabine Zahn, Jörg Wenzel, Nicola J Smith, Jesús Gomeza, Christel Drewke, Graeme Milligan, Klaus Mohr, and Evi Kostenis. Deconvolution of complex g protein-coupled receptor signaling in live cells using dynamic mass redistribution measurements. *Nature Biotechnology*, 28(9):943–949, September 2010. doi: 10.1038/nbt.1671.

Stefan Schuster and Claus Hilgetag. On elementary flux modes in biochemical reaction systems at steady state. *Journal of Biological Systems*, 2(2):165–182, June 1994. doi: 10.1142/S0218339094000131.

Malcolm R. Sears and Jan Lötvall. Past, present and future–2-adrenoceptor agonists in asthma management. *Respiratory Medicine*, 99(2):152–170, February 2005. doi: 10.1016/j.rmed.2004.07.003.

Adam J. Shaywitz and Michael E. Greenberg. Creb: A stimulus-induced transcription factor activated by a diverse array of extracellular signals. *Annual Review of Biochemistry*, 68(1): 821–861, July 1999. doi: 10.1146/annurev.biochem.68.1.821.

Donald D. Van Slyke and Glenn E. Cullen. The mode of action of urease and of enzymes in general. *Journal of Biological Chemistry*, 19(2):141–180, January 1914.

Robert L. Smith. The hit-and-run sampler: a globally reaching markov chain sampler for generating arbitrary multivariate distributions. In *Winter Simulation Conference*, pages 260–264. ACM Press, December 1996. doi: 10.1145/256562.256619.

O. Smithies. Zone electrophoresis in starch gels: group variations in the serum proteins of normal human adults. *The Biochemical Journal*, 61(4):629–641, December 1955. doi: 10.1042/bj0610629.

Eduardo D. Sontag. Molecular systems biology and control. *European Journal of Control*, 11 (4-5):396–435, January 2005. doi: 10.3166/ejc.11.396-435.

Eduard Stefan, Mohan K. Malleshaiah, Billy Breton, Po Hien Ear, Verena Bachmann, Michael Beyermann, Michel Bouvier, and Stephen W. Michnick. Pka regulatory subunits mediate synergy among conserved g-protein-coupled receptor cascades. *Nature Communications*, 2 (598):1–10, December 2011. doi: 10.1038/ncomms1605.

Petre Stoica and Randolph L. Moses. *Spectral analysis of signals*. Pearson, Prentice Hall, Upper Saddle River, 1st edition, 2005. ISBN 978-0-13-113956-5.

P. G. Strange. Agonist binding, agonist affinity and agonist efficacy at g protein-coupled receptors: Agonist binding, affinity and efficacy. *British Journal of Pharmacology*, 153(7): 1353–1363, April 2008. doi: 10.1038/sj.bjp.0707672.

Philip G. Strange. Use of the gtpγs ([35s]gtpγs and eu-gtpγs) binding assay for analysis of ligand potency and efficacy at g protein-coupled receptors: Gtpγs binding assays. *British Journal of Pharmacology*, 161(6):1238–1249, November 2010. doi: 10.1111/j.1476-5381.2010.00963.x.

A. Donny Strosberg. Structure, function, and regulation of adrenergic receptors. *Protein Science*, 2(8):1198–1209, August 1993. doi: 10.1002/pro.5560020802.

Roger K Sunahara and Ron Taussig. Isoforms of mammalian adenylyl cyclase: multiplicities of signaling. *Molecular interventions*, 2(3):168–184, June 2002. doi: 10.1124/mi.2.3.168.

M. Sunnåker, E. Zamora-Sillero, R. Dechant, C. Ludwig, A. G. Busetto, A. Wagner, and J. Stelling. Automatic generation of predictive dynamic models reveals nuclear phosphorylation as the key msn2 control mechanism. *Science Signaling*, 6(277):1–15, May 2013. doi: 10.1126/scisignal.2003621.

I. Swameye, T. G. Müller, J. Timmer, O. Sandra, and U. Klingmüller. Identification of nucleocytoplasmic cycling as a remote sensor in cellular signaling by databased modeling. *Proceedings of the National Academy of Sciences*, 100(3):1028–1033, April 2003. doi: 10.1073/pnas.0237333100.

Susan S. Taylor, Ronit Ilouz, Ping Zhang, and Alexandr P. Kornev. Assembly of allosteric macromolecular switches: lessons from pka. *Nature Reviews Molecular Cell Biology*, 13(10): 646–658, October 2012. doi: 10.1038/nrm3432.

Gerald Teschl. *Ordinary differential equations and dynamical systems*. Number v. 140 in Graduate studies in mathematics. American Mathematical Society, Providence, 1st edition, 2012. ISBN 978-0-8218-8328-0.

Robert Tibshirani. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society, Series B*, 58(1):267–288, February 1996. doi: 10.1.1.35.7574.

Robert Tibshirani. Regression shrinkage and selection via the lasso: a retrospective. *Journal of the Royal Statistical Society: Series B*, 73(3):273–282, June 2011. doi: 10.1111/j.1467-9868.2011.00771.x.

T. Toni, D. Welch, N. Strelkowa, A. Ipsen, and M. P. H. Stumpf. Approximate bayesian computation scheme for parameter inference and model selection in dynamical systems. *Journal of The Royal Society Interface*, 6(31):187–202, February 2009. doi: 10.1098/rsif.2008.0172.

H. Towbin, T. Staehelin, and J. Gordon. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proceedings of the National Academy of Sciences*, 76(9):4350–4354, September 1979.

Wolfgang Tschirk. *Statistik: Klassisch oder Bayes*. Springer-Lehrbuch. Springer, Berlin; Heidelberg, 1st edition, 2014. ISBN 978-3-642-54384-5 978-3-642-54385-2.

T. Turanyi, A. S. Tomlin, and M. J. Pilling. On the error of the quasi-steady-state approximation. *The Journal of Physical Chemistry*, 97(1):163–172, January 1993. doi: 10.1021/j100103a028.

Sara van de Geer, Peter Bühlmann, and Shuheng Zhou. The adaptive and the thresholded Lasso for potentially misspecified models (and a lower bound for the lasso). *Electronic Journal of Statistics*, 5:688–749, July 2011. doi: 10.1214/11-EJS624.

Natal A. W. van Riel. Dynamic modelling and analysis of biochemical networks: mechanism-based models and model-based experiments. *Briefings in Bioinformatics*, 7(4):364–374, September 2006. doi: 10.1093/bib/bbl040.

Felipe A. Vargas, Francisco Pizarro, J. Ricardo Pérez-Correa, and Eduardo Agosin. Expanding a dynamic flux balance model of yeast fermentation to genome-scale. *BMC Systems Biology*, 5(75):1–12, May 2011. doi: 10.1186/1752-0509-5-75.

Maria Vela, Mariana Aris, Mercedes Llorente, Jose A. Garcia-Sanz, and Leonor Kremer. Chemokine receptor-specific antibodies in cancer immunotherapy: Achievements and challenges. *Frontiers in Immunology*, 6(12):1–15, January 2015. doi: 10.3389/fimmu.2015.00012.

Mikhail V. Volkenshtein. *Entropy and information*. Birkhäuser Verlag AG, Basel; Boston, 1st edition, 2009. ISBN 978-3-0346-0078-1.

Ludwig von Bertalanffy. The theory of open systems in physics and biology. *Science*, 111 (2872):23–29, January 1950. doi: 10.1126/science.111.2872.23.

G. Vossen and H. Maurer. On l1-minimization in optimal control and applications to robotics. *Optimal Control Applications and Methods*, 27(6):301–321, November 2006. doi: 10.1002/oca.781.

P. Waage and Cato M. Gulberg. Studies concerning affinity. *Journal of Chemical Education*, 63 (12):1044–1047, December 1986. doi: 10.1021/ed063p1044.

Gary Walsh. Biopharmaceutical benchmarks 2010. *Nature Biotechnology*, 28(9):917–924, September 2010. doi: doi:10.1038/nbt0910-917.

Wolfgang Walter. *Einführung in die Theorie der Distributionen*. B I Wissenschaftsverlag, Mannheim; Leipzig; Wien; Zürich, 3rd edition, 1994. ISBN 978-3-86025-472-1.

Wolfgang Walter. *Gewöhnliche Differentialgleichungen: eine Einführung*. Springer-Lehrbuch. Springer, Berlin, 7th edition, 2000. ISBN 978-3-540-67642-3.

James Wei and Charles D. Prater. The structure and analysis of complex reaction systems. In *Advances in Catalysis*, volume 13, pages 203–392. Elsevier, 1962.

Wolfgang Wein. Drug development: successes, problems and pitfalls-the industry perspective. *ESMO Open*, 1(1):1–3, January 2016. doi: 10.1136/esmoopen-2016-000033.

Sharon J. Wiback, Iman Famili, Harvey J. Greenberg, and Bernhard Ø. Palsson. Monte carlo sampling can be used to determine the size and shape of the steady-state flux space. *Journal of Theoretical Biology*, 228(4):437–447, June 2004. doi: 10.1016/j.jtbi.2004.02.006.

Thomas Williamson, Jean-Marc Schwartz, Douglas B. Kell, and Lubomira Stateva. Deterministic mathematical models of the camp pathway in saccharomyces cerevisiae. *BMC Systems Biology*, 3(1):1–15, July 2009. doi: 10.1186/1752-0509-3-70.

P. J. Woodroffe, L. J. Bridge, J. R. King, and S. J. Hill. Modelling the activation of g-protein coupled receptors by a single drug. *Mathematical Biosciences*, 219(1):32–55, May 2009. doi: 10.1016/j.mbs.2009.02.003.

Guo-xi Xie and Pamela Pierce Palmer. How regulators of g protein signaling achieve selective regulation. *Journal of Molecular Biology*, 366(2):349–365, February 2007. doi: 10.1016/j.jmb.2006.11.045.

Fumihiko Yasuma and Jun Ichiro Hayano. Respiratory sinus arrhythmia: why does the heartbeat synchronize with respiratory rhythm? *Chest*, 125(2):683–690, February 2004. doi: 10.1378/chest.125.2.683.

T. M. Yi, H. Kitano, and M. I. Simon. A quantitative characterization of the yeast heterotrimeric g protein cycle. *Proceedings of the National Academy of Sciences*, 100(19):10764–10769, September 2003. doi: 10.1073/pnas.1834247100.

Benedikt Zacher, Khalid Abnaof, Stephan Gade, Erfan Younesi, Achim Tresch, and Holger Fröhlich. Joint bayesian inference of condition-specific mirna and transcription factor activities from combined gene and microrna expression data. *Bioinformatics*, 28(13):1714–1720, July 2012. doi: 10.1093/bioinformatics/bts257.

P. Zhang and U. Mende. Regulators of g-protein signaling in the heart and their potential as therapeutic targets. *Circulation Research*, 109(3):320–333, July 2011. doi: 10.1161/CIRCRESAHA.110.231423.

Juan Zhao, Yulin Deng, Zhaotan Jiang, and Hong Qing. G protein-coupled receptors gpcrs in alzheimer's disease: A focus on bace1 related gpcrs. *Frontiers in Aging Neuroscience*, 8(58):1–15, March 2016. doi: 10.3389/fnagi.2016.00058.

Shan Zhao and Ravi Iyengar. Systems pharmacology: Network analysis to identify multiscale mechanisms of drug action. *Annual Review of Pharmacology and Toxicology*, 52(1):505–521, February 2012. doi: 10.1146/annurev-pharmtox-010611-134520.

C. J. Zheng. Therapeutic targets: Progress of their exploration and investigation of their characteristics. *Pharmacological Reviews*, 58(2):259–279, June 2006. doi: 10.1124/pr.58.2.4.

Hui Zou and Trevor Hastie. Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society: Series B*, 67(2):301–320, April 2005. doi: 10.1111/j.1467-9868.2005.00503.x.

Hui Zou and Hao Helen Zhang. On the adaptive elastic-net with a diverging number of parameters. *The Annals of Statistics*, 37(4):1733–1751, August 2009. doi: 10.1214/08-AOS625.

Appendices

APPENDIX A

List of Acronyms

AC	adenylyl cyclase
ACHR	artificial centering hit-and-run
AMP	adenosine monophosphate
aO	allo-ocimene
αP	α-Pinene
ATP	adenosine triphosphate
AUC	area under the curve
AUROC	area under the ROC curve
BDEN	Bayesian dynamic elastic-net
cAMP	cyclic adenosine monophosphate
CFS	constrained flux sampling
CHO	Chinese hamster ovary
CRN	chemical reaction network
CTX	cholera toxin
DEN	dynamic elastic-net
DMR	dynamic mass redistribution
DNA	deoxyribonucleic acid
dP	dipentene

EFM	elementary flux mode
ELISA	enzyme-linked immunosorbent assay
EpoR	erythropoietin receptor
FBA	flux balance analysis
FDR	false discovery rate
GAPDH	glycerinaldehyd-3-phosphat-dehydrogenase
GDP	guanosine-5'-diphosphate
GEF	guanine nucleotide exchange factor
GP	G protein
GPCR	G protein–coupled receptor
GRK	G protein-coupled receptor kinases
GTP	guanosine-5'-triphosphate
HRS	hit-and-run sampling
IVP	initial value problem
LASSO	least absolute shrinkage and selection operator
LSE	least-square-error
mACHR	muscarinic acetylcholine receptor
MCMC	Markov chain Monte Carlo
MH	Metropolis-Hastings
MLE	maximum likelihood estimator
mRNA	messenger ribonucleic acid
M_2 receptor	muscarinic acetylcholine receptor subtype 2
OCP	optimal control problem
OCT	optimal control theory
ODE	ordinary differential equation
PDE	3',5'-cyclonukleotid-phosphodiesterase

PKA	protein kinase A
PKC	protein kinase C
PTX	pertussis toxin
pUCR1	phosphorylated upstream conserved region 1
QSSA	quasi-steady-state approximation
RGS	regulators of G protein signaling
RWG	resonant waveguide grating
STAT5	signal transducer and activator of transcription 5
UV-B	ultraviolet B
WB	western blotting
YM	YM-254890

APPENDIX B

M_2 Receptor-dependent Signaling in CHO Cells

Material and Methods

Cell Culture

Flp-In™-CHO cells stably expressing the hM2 receptor (CHO-hM2 cells) were cultured as described previously in Ham's nutrient mixture F-12 (Ham's F-12) supplemented with 10% (v/v) fetal calf serum (FCS), 100 U mL^{-1} penicillin, 100 mg mL^{-1} streptomycin and 2 mM l-glutamine (Sigma Aldrich, UK) (Schrage et al., 2013) . The cells were grown in a humidified incubator at 37 °C and 5% CO_2, and passaged by trypsinization at nearly confluence.

cAMP Assay

The quantification of the agonist-induced rise of the intracellular cAMP was performed using CHO-hM2 cells pretreated with 50 ng/ml^{-1} PTX for 16-22 h as described previously using the HTRF-cAMP dynamic kit (Cisbio, Bagnols-sur-Cèze, France) according to the manufacturer's introductions (Schröder et al., 2009). The cells were incubated for 30 minutes with iperoxo. The fluorescence was quantified on a Mithras LB 940 reader (Berthold Technologies, Bad Wildbad, Germany).

Western Blots

Cells were seeded into 6 well plates and stimulated with 0.1 µM iperoxo for 2, 5, 10, 15, 20, 30 minutes. After specified treatments, cellular lysates were prepared in lysis buffer [25 mM Hepes, 2.5 mM EDTA, 50 mM NaCl, 50 mM NaF, 30 mM sodium pyrophosphate, 10% (v/v) glycerol, 1% (v/v) Triton X-100, pH 7.5, containing Complete™ EDTA-free protease inhibitor cocktail tablets (Roche)].

Proteins were separated by SDS/PAGE (4–12% Bis-Tris gels) and transferred onto nitrocellulose membranes for western blotting using the NUPAGE system (Invitrogen). The membranes were blocked in 5% phosphoblocker. The blots were probed with phospho-UCR1 and pan-PDE4 (Millipore) diluted to 1/5000 in 1% phosphoblocker diluted in TBS-T (150 mM NaCl, 20 mM Tris–HCl, pH 7.6, 0.1% Tween-20). The bands were visualized using the appropriate secondary antibody and visualized using an Odyssey scanner (Licor Biosciences) and the band intensity was quantified using Image Studio (Licor Biosciences, v. 5.2).

Detailed Reaction Graph

A detailed reaction graph is illustrated in Figure B.1 and detailed information on each step can be obtained from Table B.1. For the full elementary network please refer to the Full List of Elementary Reactions. All proteins linked to tubulin and actin are considered in the response function (namely the GP subunit α_o production, PKA production, GRK2 production, cAMP-GEF1 production and AMP production). The elementary reactions mapped to Figure 5.4 (including the fold changes) are as follows:

Ligand-dependent GP activation (fold change: NA): $L : R : G_i \rightarrow L : R + i + \beta$
GRK2-mediated regulation (fold change: 1.27): $L : R : GRK6 \rightarrow \overline{L : R} + GRK6$
RGS14-mediated regulation (fold change: 1.34): $i : RGS14 \rightarrow \bar{i} + RGS14$ CI:
cAMP-conversion (fold change: 1, 47): $AC_{5s} : ATP \rightarrow AC_{5s} + cAMP$
PKA activation (fold change: 1.54): $\overline{PKA} + 2cAMP \rightarrow PKA$
cAMP degradation (fold change: 1.68): $PDE4 : cAMP \rightarrow PDE4 + AMP$

Reaction Graph

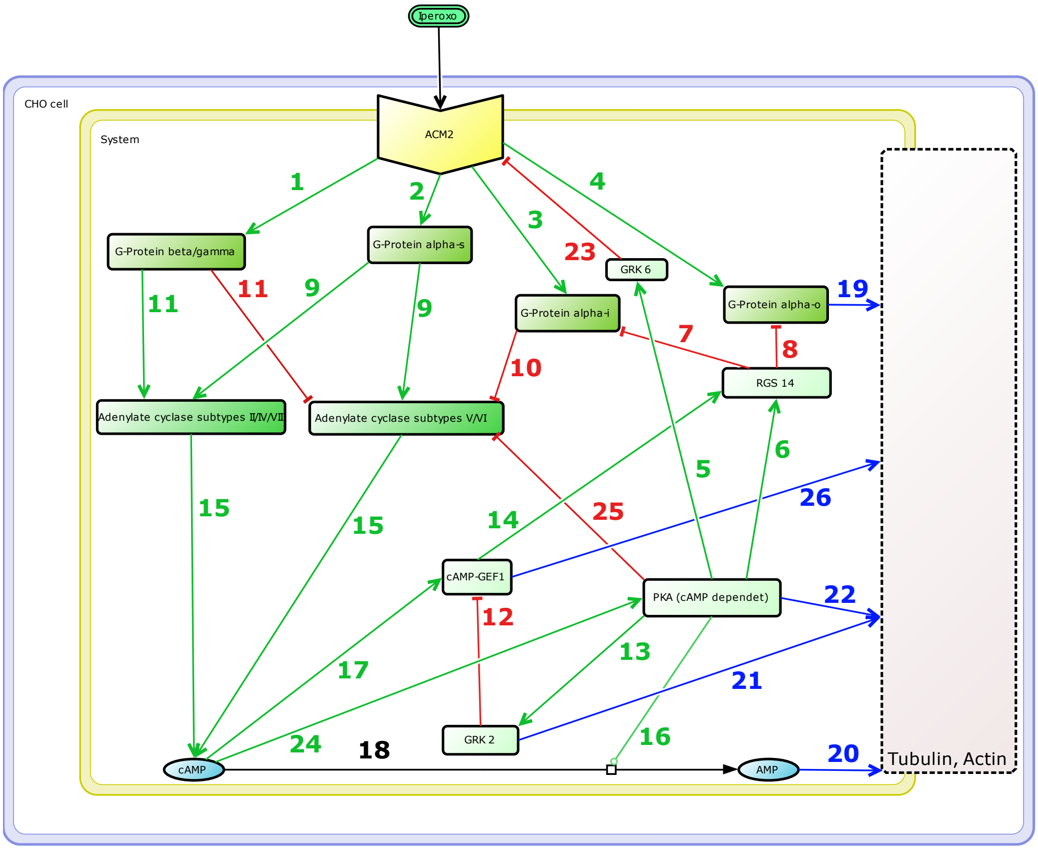


Figure B.1: Illustrated overall reaction network. Stimulation events are represented as green lines and inhibitory events represented as red lines accordingly. Details regarding specific reactions are given in Table B.1. A complete list of all elementary reactions is given in the Full List of Elementary Reactions.

Table B.1: Basic reactions in the modeled biological system: Species shown in the column “From” are biochemically converted into those shown in the column “To”. The underlying biological mechanism and its effect on the species in column “To” are shown in the fourth and fifth columns, respectively. We distinguish between binding, phosporylation, transformation, catalysis and unspecific chemical reactions. The effect of unspecific chemical reactions is denoted as “technical” because they are neither activating nor inhibiting. The derived elementary reactions can be found in the Full List of Elementary Reactions.

Nr.	From	To	Mechanism	Effect	Lead Author, Year, Journal[Impact]
1	ACM2	G-Protein beta/gamma	binding	activation	Wettschurek N, 2005, Physiol Rev.[30.2]
2	ACM2	G-Protein alpha-s	transformation	activation	Cao J, 2003, Bioinformatics.[5.3]
3	ACM2	G-Protein alpha-i family	transformation	activation	Wettschurek N, 2005, Physiol Rev.[30.2]
4	ACM2	G-Protein alpha-o	transformation	activation	Azpiazu I, 2004, J Biol Chem.[4.56]
5	PKA-cat (cAMP-dependent)	GRK6	phosphorylation	inhibition	Rapacciulolo A, 2003, J Biol Chem.[4.56]
6	PKA-cat (cAMP-dependent)	RGS14	phosphorylation	activation	Riddle EL, 2005, Circ Res.[11.86]
7	RGS14	G-Protein alpha-i family	transformation	inhibition	Kimple RJ, 2002, Nature.[38.58]
8	RGS14	G-Protein alpha-o	transformation	inhibition	Rual JF, 2005, Nature.[38.58]
9	G-Protein alpha-s	Adenylate cyclase	binding	activation	Pourquie O, 2005, Nature[38.58]
10	G-Protein alpha-i	Adenylate cyclase subtype 5	binding	inhibition	Taussig R, 1993, Science.[31.03]
10	G-Protein alpha-i	Adenylate cyclase subtype 6	binding	inhibition	Taussig R, 1993, Science.[31,03]
11	G-Protein beta/gamma	Adenylate cyclase subtype 2	binding	activation	Sunahara RK, 2002, Mol Interv[6.48]
11	G-Protein beta/gamma	Adenylate cyclase subtype 4	binding	activation	Sunahara RK, 2002, Mol Interv[6.48]
11	G-Protein beta/gamma	Adenylate cyclase subtype 5	binding	inhibition	Sunahara RK, 2002, Mol Interv[6.48]
11	G-Protein beta/gamma	Adenylate cyclase subtype 6	binding	inhibition	Sunahara RK, 2002, Mol Interv[6.48]
11	G-Protein beta/gamma	Adenylate cyclase subtype 7	binding	activation	Sunahara RK, 2002, Mol Interv[6.48]
12	GRK2	cAMP-GEFI	binding	inhibition	Eijkelkamp N, 2010, J Neurosci.[7.3]
13	PKA-cat (cAMP-dependent)	GRK2	phosphorylation	activation	Cong M, 2001, J Biol Chem.[4.65]
14	cAMP-GEFI	RAP-2A	transformation	activation	Ohba Y, 2000, Mol Cell Biol.[5.37]
14	RAP-2A	RGS14	binding	activation	Bandyopadhyay S, 2010, Nat Methods.[23.57]
15	Adenylate cyclase	ATP = cyclic AMP + Pyrophosphate	catalysis	activation	Wong ST, 2000, Neuron.[15.77]
16	PKA-cat (cAMP-dependent)	PDE4B	phosphorylation	activation	Ghigo A, 2012, Circulation.[15.20]
16	PDE4B	cyclic AMP + H(,2)O = AMP	catalysis	activation	Huston E, 1997, Biochem J.[5.37]
17	cyclic AMP intracellular	cAMP-GEFI	binding	activation	Do Rooij Z, 1998, Nature.[38.58]
18	ATP in cytoplasm	ATP = cyclic AMP + Pyrophosphate	reaction	technical	
18	ATP = cyclic AMP + Pyrophosphate	cyclic AMP intracellular	reaction	technical	
18	cyclic AMP intracellular	cyclic AMP + H(,2)O = AMP	reaction	technical	
18	cyclic AMP + H(,2)O = AMP	AMP in cytoplasm	reaction	technical	
19	G-Protein alpha-o	Tubulin (in microtubules)	transformation	inhibition	Roychowdhury S, 1999, J Biol Chem.[4.65]
20	AMP	Tubulin (in microtubules)	transformation	inhibition	Lee JH, 2007, Nat. [41.46]
21	GRK2	Tubulin (in microtubules)	phosphorylation	activation	Storti B, 2012, J Biol Chem.[4.65]
22	PKA-cat (cAMP-dependent)	Actin cytoskeletal	phosphorylation	inhibition	Howe AK, 2004 Biochim Biophys Acta.[5.54]
23	GRK6	ACM2	phosphorylation	inhibition	Loudon RP, 1994, J Biol Chem.[4.65]
24	cyclic AMP intracellular	PKA-reg (cAMP-dependent)	binding	activation	Lin JY, 2007, Nat.[38.58]
24	PKA-reg (cAMP-dependent)	PKA-cat (cAMP-dependent)	binding	activation	Ghigo A, 2012, Circulation.[15.20]
25	PKA-cat (cAMP-dependent)	Adenylate cyclase subtype 5	phosphorylation	inhibition	Iwami G, 1995, J Biol Chem.[4.65]
25	PKA-cat (cAMP-dependent)	Adenylate cyclase subtype 6	phosphorylation	inhibition	Defer N, 2000, Am J Physiol Renal Physiol.[9.73]
26	cAMP-GEFI	AKT1	binding	activation	Misra UK, 2005, J Biol Chem.[4.65]
A	ACM2	ACM2	several	regulation	Pierce KL, 2002, Nat Rev Mol Cell Biol.[37.16]
B	ACM2	G protein	spontaneous	activation	Trzaskowski B, 2012, Curr Med Chem.[3.85]

Full List of Elementary Reactions

The first number (red) counts the reaction, the second number is linked to Table B.1. Fluxes involved in the cellular response as defined in Section 5.4.2 are marked by “*“.

$L + R \rightleftharpoons L : R$ (1|A)

$\overline{R} \rightarrow R$ (2|A)

$\overline{L : R} \rightarrow L : R$ (3|A)

$R + GRK6 \rightleftharpoons GRK6 : R$ (4|23)

$GRK6 : R \rightarrow \overline{R} + GRK6$ (5|23)

$L : R + GRK6 \rightleftharpoons L : R : GRK6$ (6|23)

$L : R : GRK6 \rightarrow \overline{L : R} + GRK6$ (7|23)

$\overline{L : R} \rightleftharpoons \overline{R} + L$ (8|A)

$R + G_s \rightleftharpoons R : G_s$ (9|1; 2; B)

$R : G_s \rightarrow R + s + \beta/\gamma$ (10|1; 2; B)

$R + G_i \rightleftharpoons R : G_i$ (11|1; 3; B)

$R : G_i \rightarrow R + i + \beta/\gamma$ (12|1; 3; B)

$R + G_o \rightleftharpoons R : G_o$ (13|1; 4; B)

$R : G_o \rightarrow R + o + \beta/\gamma$ * (14|1; 4; B)

$L : R + G_s \rightleftharpoons L : R : G_s$ (15|1; 2)

$L : R : G_s \rightarrow L : R + s + \beta/\gamma$ (16|1; 2)

$s \rightarrow \overline{s}$ (17|2)

$L : R + G_i \rightleftharpoons LR : G_i$ (18|1; 3)

$L : R : G_i \rightarrow L : R + i + \beta/\gamma$ (19|1; 3)

$i + RGS14 \rightleftharpoons i : RGS14$ (20|7)

$i : RGS14 \rightarrow \overline{i} + RGS14$ (21|7)

$\beta/\gamma + \overline{i} \rightleftharpoons G_i$ (22|1; 3)

$\beta/\gamma + \overline{s} \rightleftharpoons G_s$ (23|2; 3)

$L : R + G_o \rightleftharpoons L : R : G_o$ (24|1; 4)

$L : R : G_o \rightarrow L : R + o + \beta/\gamma$ * (25|1; 4)

$\beta/\gamma + \overline{o} \rightleftharpoons G_o$ (26|1; 4)

$o + RGS14 \rightleftharpoons o : RGS14$ (27|8)

$o : RGS14 \rightarrow \overline{o} + RGS14$ (28|8)

$\beta/\gamma + AC_4 \rightleftharpoons AC_{4\beta}$ (29|11)

$s + AC_4 \rightleftharpoons AC_{4s}$ (30|9)

$s + AC_{4\beta} \rightleftharpoons AC_{4\beta/\gamma s}$ (31|9; 11)

$\beta/\gamma + AC_{4s} \rightleftharpoons AC_{4\beta/\gamma s}$ (32|9; 11)

$i + AC_5 \rightleftharpoons AC_{5i}$ (33|10)

$s + AC_5 \rightleftharpoons AC_{5s}$ (34|9)

$\beta/\gamma + AC_5 \rightleftharpoons AC_{5\beta/\gamma}$ (35|11)

$\beta/\gamma + AC_{5s} \rightleftharpoons AC_{5\beta/\gamma s}$ (36|9)

$s + AC_{5\beta/\gamma} \rightleftharpoons AC_{5\beta/\gamma s}$ (37|9; 11)

$i + AC_{5s} \rightleftharpoons AC_{5si}$ (38|9; 10)

$s + AC_{5i} \rightleftharpoons AC_{5si}$ (39|9; 10)

$AC_{4s} \rightleftharpoons AC_{4s} : ATP$ (40|15)

$AC_{4s} : ATP \rightarrow AC_{4s} + cAMP$ (41|15)

$AC_{4\beta s} \rightleftharpoons AC_{4\beta/\gamma s} : ATP$ (42|15)

$AC_{4\beta/\gamma s} : ATP \rightarrow AC_{4\beta/\gamma s} + cAMP$ (43|15)

$\beta/\gamma + s + AC_4 \rightleftharpoons AC_{4\beta/\gamma s}$ (44|9; 11)

$AC_5 + i + s \rightleftharpoons AC_{5is}$ (45|10; 11)

$AC_5 + \beta/\gamma + s \rightleftharpoons AC_{5\beta/\gamma s}$ (46|9; 11)

$AC_{5s} \rightleftharpoons AC_{5s} : ATP$ (47|15)

$AC_{5s} : ATP \rightarrow AC_{5s} + cAMP$ (48|15)

$AC_5^p + s \rightleftharpoons AC_{5s}^p$ (49|9)

$AC_{5s}^p \rightleftharpoons AC_{5s}^p : ATP$ (50|15)

$AC_{5s}^p : ATP \rightarrow AC_{5s}^p + cAMP$ (51|15)

$AC_5 + PKA \rightleftharpoons AC_5 : PKA$ (52|25)

$AC_5 : PKA \rightarrow AC_5^p + PKA$ (53|25)

$AC_5^p \rightarrow AC_5$ (54|25)

$AC_{5s} + PKA \rightleftharpoons AC_{5s} : PKA$ (55|25)

$AC_{5s} : PKA \rightarrow AC_{5s}^p + PKA$ (56|25)

$PKA \rightarrow \overline{PKA}$ (57|24)

$\overline{PKA} + 2cAMP \rightarrow PKA$ * (58|24)

$PKA + \overline{RGS14} \rightleftharpoons PKA : RGS14$ (59|6)

$PKA : RGS14 \rightarrow RGS14 + PKA$ (60|6)

$RGS14 \rightarrow \overline{RGS14}$ (61|6)

$PKA + \overline{GRK2} \rightleftharpoons PKA : GRK2$ (62|13)

$PKA : GRK2 \rightarrow GRK2 + PKA$ * (63|13)

$GRK2 \rightarrow \overline{GRK2}$ (64|13)

$PKA + \overline{GRK6} \rightleftharpoons PKA : GRK6$ (65|5)

$PKA : GRK6 \rightarrow GRK6 + PKA$ (66|5)

$GRK6 \rightarrow \overline{GRK6}$ (67|5)

$PKA + \overline{PDE4} \rightleftharpoons PKA : PDE4$ (68|16)

$PKA : PDE4 \rightarrow PDE4 + PKA$ (69|16)

$PDE4 \rightarrow \overline{PDE4}$ (70|16)

$PDE4 + cAMP \rightleftharpoons PDE4 : cAMP$ (71|16)

$PDE4 : cAMP \rightarrow PDE4 + AMP$ (72|16)

$AMP \rightarrow \emptyset$ (73|18)

$GEF + \overline{RGS14} \rightleftharpoons GEF : RGS14$ (74|14)

$GEF : RGS14 \rightarrow RGS14 + GEF$ (75|14)

$GRK2 + GEF \rightleftharpoons GRK2 : \overline{GEF}$ (76|12)

$GRK2 : \overline{GEF} \rightarrow GRK2 + \overline{GEF}$ (77|12)

$GEF \rightarrow \overline{GEF}$ (78|17, 12)

$cAMP + \overline{GEF} \rightleftharpoons GEF$ * (79|17)

$X : Y$ = Complex of X and Y

$\overline{X}$ = Inactive form of protein X

X^p = Phosphorylated protein

Ranked EFMs

Results

Table B.2: EFMs ranked by their median fold change value.

Rank	FDR	Fold Change	Related to
1	$< 1E-6$	*Na*	GRK6/LR
2	$< 1E-6$	*Na*	GRK6/R
3	$< 1E-6$	*Na*	Receptor
4	$< 1E-6$	2.00	GP α_s
5	$< 1E-6$	1.68	PDE
6	$< 1E-6$	1.58	GEF/cAMP/GRK2
7	$< 1E-6$	1.57	GEF/cAMP/GRK2
8	$< 1E-6$	1.57	GEF/cAMP/GRK2
9	$< 1E-6$	1.57	GEF/cAMP/GRK2
10	$< 1E-6$	1.54	cAMP/PKA
11	$< 1E-6$	1.54	cAMP/PKA
12	$< 1E-6$	1.54	cAMP/PKA
13	$< 1E-6$	1.54	GEF/cAMP
14	$< 1E-6$	1.53	GEF/cAMP
15	$< 1E-6$	1.52	GEF/cAMP
16	$< 1E-6$	1.52	GEF/cAMP
17	$< 1E-6$	1.52	cAMP
18	$< 1E-6$	1.48	cAMP
19	$< 1E-6$	1.48	cAMP
20	$< 1E-6$	1.48	cAMP
21	$< 1E-6$	1.48	cAMP
22	$< 1E-6$	1.34	RGS14/GP α_i
23	$< 1E-6$	1.34	RGS14/GP α_i
24	$< 1E-6$	1.33	RGS14/GP α_o
25	$< 1E-6$	1.33	RGS14 /GP α_o
26	$< 1E-6$	1.27	GRK6/R
27	$< 1E-6$	1.01	AC5
28	0.01	1.01	AC5
29	0.01	1.01	AC5
30	0.01	1.01	AC5
31	0.02	1.01	AC5
32	> 0.10	1.00	RGS14
33	> 0.10	1.00	GP α_s
34	> 0.10	1.00	GRK6
35	> 0.10	1.00	GRK2
36	> 0.10	1.00	AC5
37	> 0.10	1.00	AC5
38	> 0.10	1.01	AC5
39	> 0.10	1.01	AC5
40	> 0.10	1.01	AC5
41	> 0.10	1.01	AC5
42	> 0.10	1.01	AC5
43	> 0.10	1.01	AC5
44	> 0.10	1.01	AC5
45	> 0.10	1.01	AC5
46	> 0.10	1.01	AC5
47	> 0.10	1.01	AC5
48	> 0.10	1.01	AC5
49	> 0.10	1.01	AC5
50	> 0.10	1.0	AC5
51	> 0.10	1.01	AC5
52	> 0.10	1.01	AC5
53	> 0.10	1.01	AC5
54	> 0.10	1.01	AC4
55	> 0.10	1.00	AC4
56	> 0.10	1.00	AC5
57	> 0.10	1.00	AC5
58	> 0.10	1.00	AC4
59	> 0.10	1.00	AC4
60	> 0.10	1.00	AC5
61	> 0.10	1.00	AC5
62	> 0.10	1.00	AC5
63	> 0.10	1.00	AC5

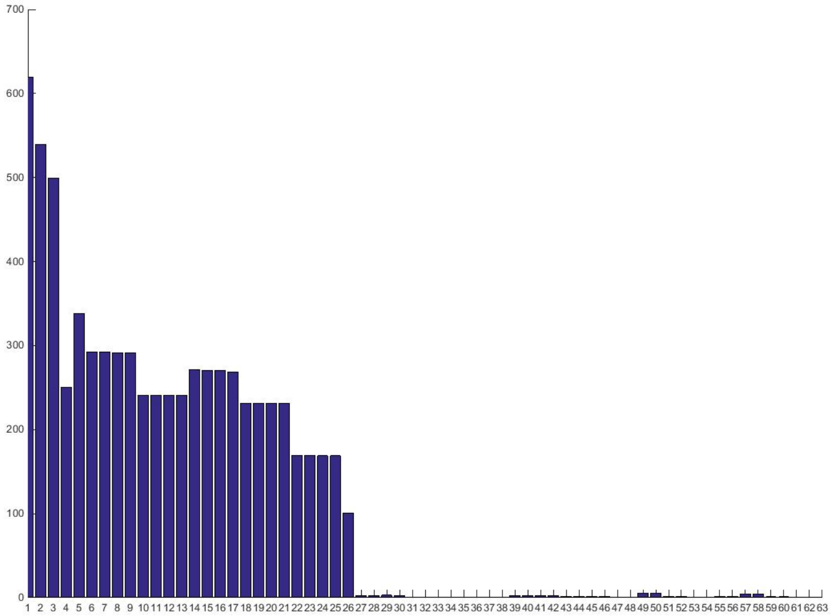
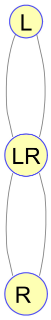
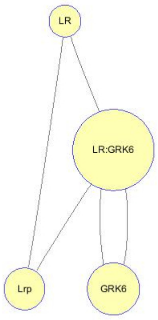
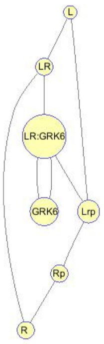


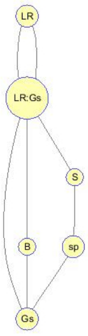
Figure B.2: Median differences of the EFMs ranked by their fold change value.

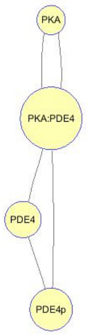
Top 26 EFMs

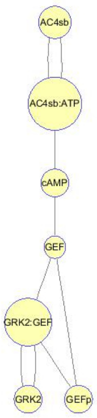


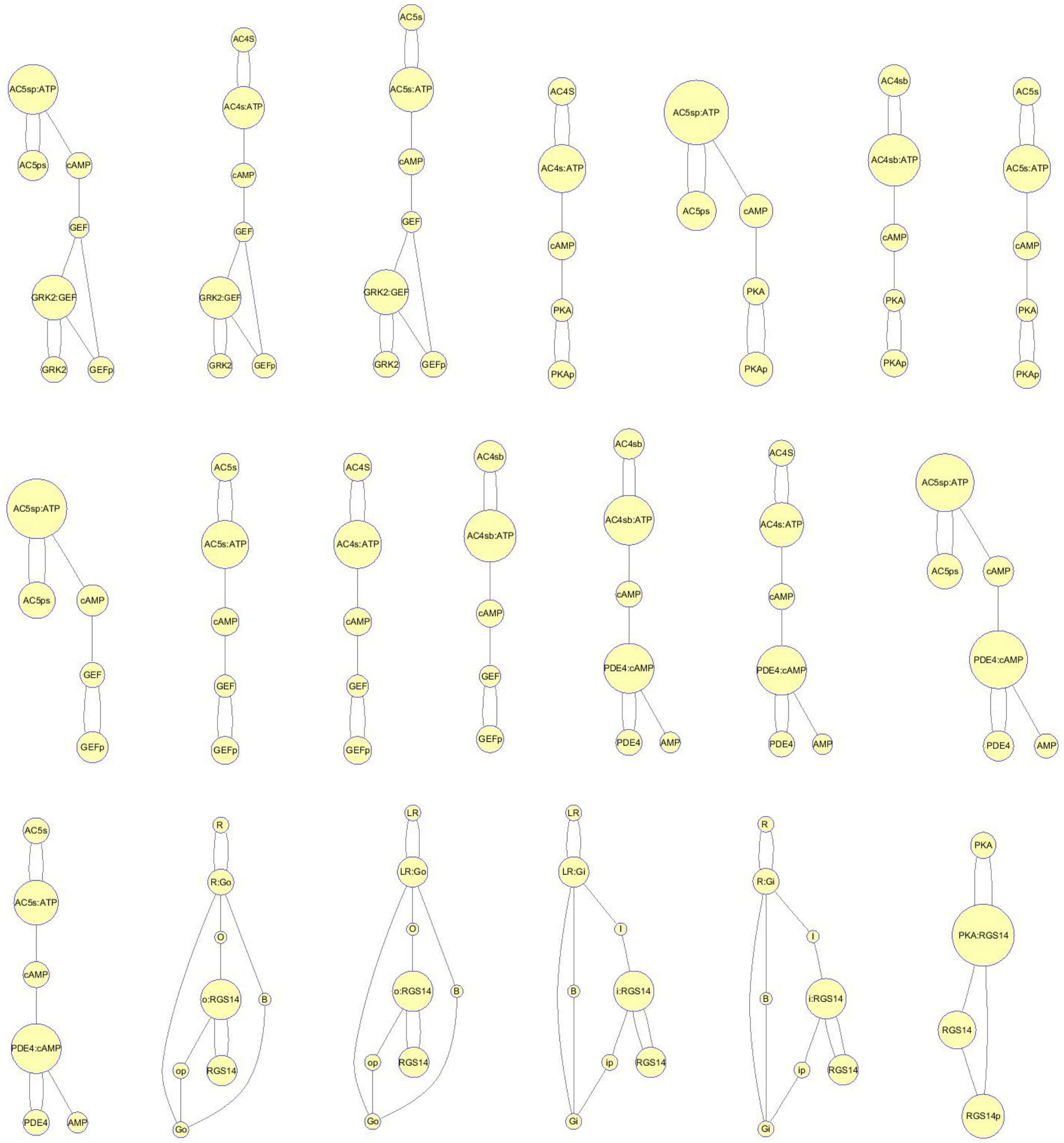











AC5sp:ATP
AC5ps
cAMP
GEF
GRK2:GEF
GRK2
GEFp
AC4S
AC4s:ATP
cAMP
GEF
GRK2:GEF
GRK2
GEFp
AC5s
AC5s:ATP
cAMP
GEF
GRK2:GEF
GRK2
GEFp
AC4S
AC4s:ATP
cAMP
PKA
PKAp
AC5sp:ATP
AC5ps
cAMP
PKA
PKAp
AC4sb
AC4sb:ATP
cAMP
PKA
PKAp
AC5s
AC5s:ATP
cAMP
PKA
PKAp
AC5sp:ATP
AC5ps
cAMP
GEF
GEFp
AC5s
AC5s:ATP
cAMP
GEF
GEFp
AC4S
AC4s:ATP
cAMP
GEF
GEFp
AC4sb
AC4sb:ATP
cAMP
GEF
GEFp
AC4sb
AC4sb:ATP
cAMP
PDE4:cAMP
PDE4
AMP
AC4S
AC4s:ATP
cAMP
PDE4:cAMP
PDE4
AMP
AC5sp:ATP
AC5ps
cAMP
PDE4:cAMP
PDE4
AMP
AC5s
AC5s:ATP
cAMP
PDE4:cAMP
PDE4
AMP
R
R:Go
O
o:RGS14
B
op
RGS14
Go
LR
LR:Go
O
o:RGS14
B
op
RGS14
Go
LR
LR:Gi
I
B
i:RGS14
ip
RGS14
Gi
R
R:Gi
I
B
i:RGS14
ip
RGS14
Gi
PKA
PKA:RGS14
RGS14
RGS14p

Flux Sampling Results

Unstimulated

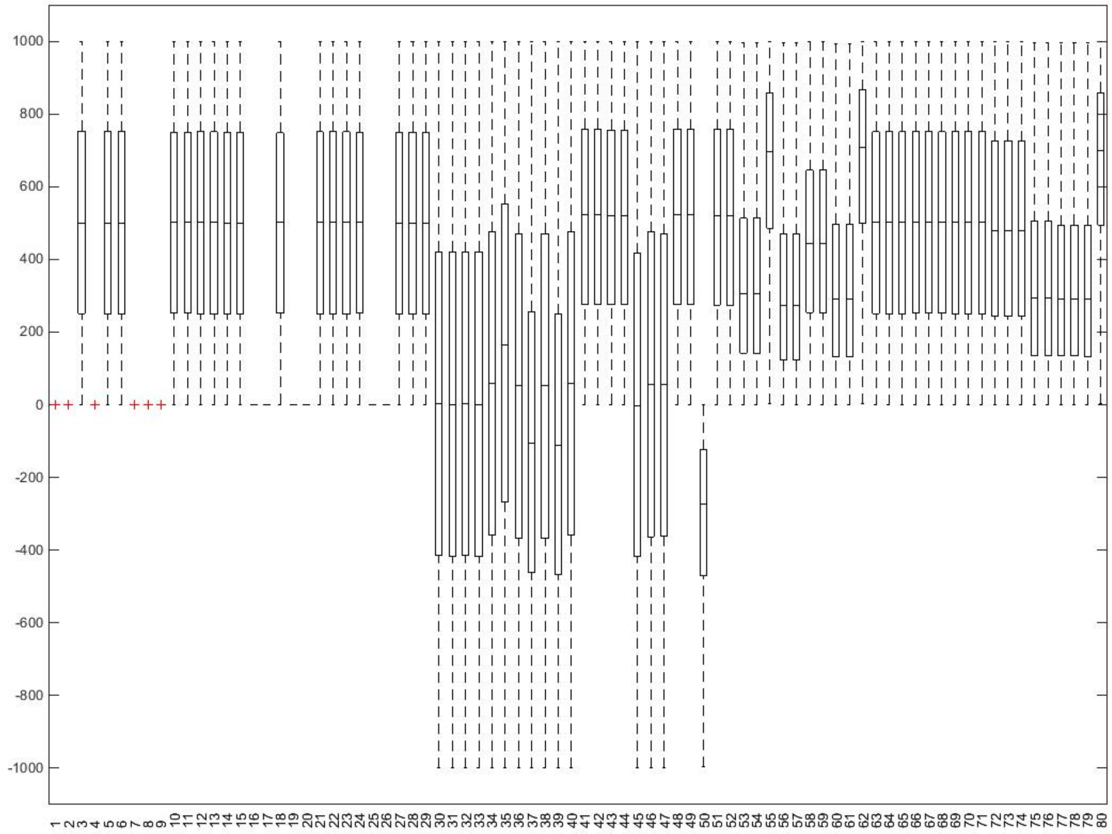


Figure B.3: System without stimulation sorted by reaction number. For detailed information regarding the corresponding reaction please refer to the Full List of Elementary Reactions.

Stimulated

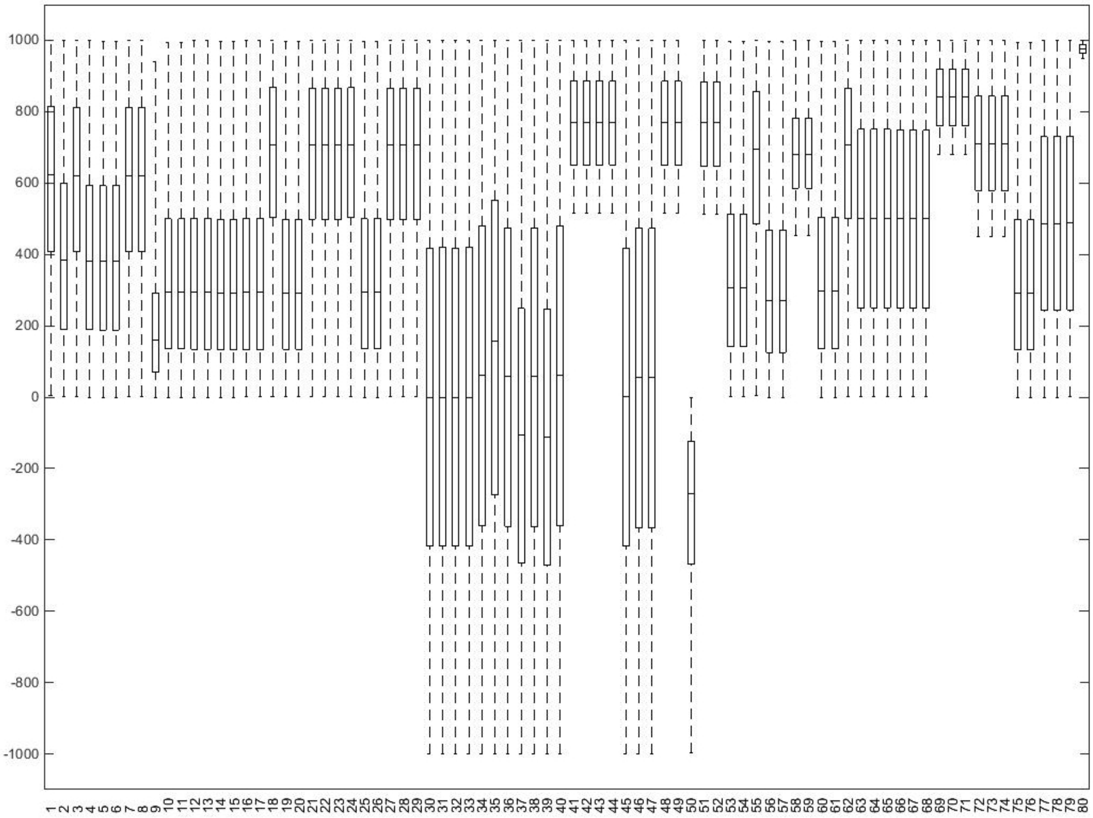


Figure B.4: System with stimulation under steady state conditions with additional data sorted by reaction number. For detailed information regarding the corresponding reaction please refer to the Full List of Elementary Reactions.

APPENDIX C

The Dynamic Elastic-Net

Software

The simulations were performed in MATLAB (R2014a, The MathWorks, Inc.) using TOMLAB v8.0 with SQOPT 7.2-5 QP and SNOPT 7.2-5 NLP (Tomlab Optimization AB) for solving the optimal control problems. The computing time for a single run of the DEN on a laptop (Intel® Core™ i5-4200M CPU with 4× 2.50GHz and 16 GB RAM) was on average 1 min.

Analysis of the Dynamic Elastic-Net Optimal Control Problem

To gain a deeper understanding of the DEN we assume that the output $\boldsymbol{y}(t)$ of the underlying true system

$$\dot{\boldsymbol{x}}(t) = \tilde{\boldsymbol{f}}(\boldsymbol{x}(t), \boldsymbol{u}(t)) + \boldsymbol{w}(t) \tag{C.1a}$$

$$\boldsymbol{y}(t) = \boldsymbol{h}(\boldsymbol{x}(t)) \tag{C.1b}$$

with initial conditions $\boldsymbol{x}(0) = \boldsymbol{x}_0$ can be measured for all time points t in the interval $[0, F]$. Accordingly, the error functional (Equation (6.4)) is replaced by the continuous time error functional

$$J\left(\hat{\boldsymbol{x}}\left(t\right), \hat{\boldsymbol{w}}\left(t\right)\right) = \int_0^F \|\boldsymbol{y}(t) - \hat{\boldsymbol{y}}(t)\|^2_{Q(t)} + \alpha_1 \|\hat{\boldsymbol{w}}(t)\|_1 + \frac{\alpha_2}{2}\|\hat{\boldsymbol{w}}(t)\|_2^2 dt \tag{C.2a}$$

$$= \int_0^F \sum_{i=1}^{I}\sum_{j=1}^{I}\left[q_{ij}(t)\left(y_i(t) - \hat{y}_i(t)\right)\left(y_j(t) - \hat{y}_j(t)\right)\right] + \sum_{n=1}^{N}\left[\alpha_1|\hat{w}_n(t)| + \frac{\alpha_2}{2}\hat{w}_n^2(t)\right] dt \tag{C.2b}$$

with the symmetric weighting matrix $Q(t) = (q_{ij}(t))$. The continuous time version will make the subsequent analysis more transparent without altering the main conclusions.

The DEN for continuous time measurements is then given by

$$\min_{\hat{\mathbf{w}}(t)} \left[J\left(\hat{\boldsymbol{x}}(t), \hat{\boldsymbol{w}}(t)\right)\right] \qquad \text{subject to} \tag{C.2c}$$

$$\dot{\hat{\boldsymbol{x}}} = \tilde{\boldsymbol{f}}(\hat{\boldsymbol{x}}(t), \boldsymbol{u}(t)) + \hat{\boldsymbol{w}}(t) \tag{C.2d}$$

$$\hat{\boldsymbol{y}}(t) = \boldsymbol{h}(\hat{\boldsymbol{x}}(t)). \tag{C.2e}$$

For most practical cases, this OCP can only be solved numerically (Gerdts, 2012; Pontryagin et al., 1986; Fleming and Rishel, 1975). However, the analysis of the necessary optimality conditions provides interesting insights. These conditions are formulated in terms of the Hamiltonian

$$\begin{aligned} H\left(\boldsymbol{\lambda}(t), \hat{\boldsymbol{x}}(t), \hat{\boldsymbol{w}}(t)\right) =& \boldsymbol{\lambda}^T(t)\left(\hat{\boldsymbol{f}}\left(\hat{\boldsymbol{x}}(t), \boldsymbol{u}(t)\right) + \hat{\boldsymbol{w}}(t)\right) \\ &+ \|\boldsymbol{y}(t) - \boldsymbol{h}\left(\hat{\boldsymbol{x}}(t)\right)\|_Q^2 + \alpha_1\|\hat{\boldsymbol{w}}(t)\|_1 + \frac{\alpha_2}{2}\|\hat{\boldsymbol{w}}(t)\|_2^2. \end{aligned} \tag{C.3}$$

According to Section 4.2, the new dynamic variable $\boldsymbol{\lambda} : [0, F] \mapsto \mathbb{R}^N$ is called the co-state or the adjoint state. It fulfills the adjoint differential equation (Fleming and Rishel, 1975; Gerdts, 2012)

$$\dot{\boldsymbol{\lambda}}(t) = -\left[\frac{\partial H\left(\boldsymbol{\lambda}(t), \hat{\boldsymbol{x}}(t), \hat{\boldsymbol{w}}(t)\right)}{\partial \hat{\boldsymbol{x}}}\right]^T \tag{C.4a}$$

$$= -\left[\frac{\partial \boldsymbol{f}\left(\hat{\boldsymbol{x}}(t)\right)}{\partial \hat{\boldsymbol{x}}}\right]^T \boldsymbol{\lambda}(t) - 2\left[\frac{\partial \boldsymbol{h}\left(\hat{\boldsymbol{x}}(t)\right)}{\partial \hat{\boldsymbol{x}}}\right]^T Q(t)\left(\boldsymbol{y}(t) - \boldsymbol{h}(\hat{\boldsymbol{x}}(t))\right). \tag{C.4b}$$

Without further conditions, the final co-state $\boldsymbol{\lambda}(F)$ equals zero. Alternatively, one can impose the condition that the fit of the DEN output $\hat{\boldsymbol{y}}(F)$ at final time F is not too far from the observed output $\boldsymbol{y}(F)$ by

$$(\boldsymbol{y}(F)) - \hat{\boldsymbol{y}}(F))^T\, Q(F)\,(\boldsymbol{y}(F)) - \hat{\boldsymbol{y}}(F)) - \Delta_F < 0, \tag{C.5}$$

where Δ_F is a given tolerance. For this terminal constraint, we obtain the boundary condition for the co-state

$$\boldsymbol{\lambda}(F) = \boldsymbol{\mu}^T \frac{\partial}{\partial \hat{\boldsymbol{x}}}\,(\boldsymbol{y}(F)) - \hat{\boldsymbol{y}}(F))^T\, Q(F)\,(\boldsymbol{y}(F)) - \hat{\boldsymbol{y}}(F)) \tag{C.6a}$$

$$= 2\boldsymbol{\mu}^T\left[\frac{\partial \boldsymbol{h}\left(\hat{\boldsymbol{x}}(F)\right)}{\partial \hat{\boldsymbol{x}}}\right]^T Q(T)\left(\boldsymbol{y}(F) - \boldsymbol{h}(\hat{\boldsymbol{x}}(F))\right), \tag{C.6b}$$

which involves additional Lagrange-parameters $\boldsymbol{\mu}$ to be determined. If the initial condition $\boldsymbol{x}_0$ is uncertain, then one can add a condition analogous to Equation (C.5).

Insights regarding the effect of the elastic-net regularization stem from Pontryagin's minimum principle (Pontryagin et al., 1986; Fleming and Rishel, 1975; Gerdts, 2012), which requires a minimization of the Hamiltonian (C.3) with respect to $\hat{\boldsymbol{w}}$. To this end, we rewrite the Hamiltonian as

$$H\left(\boldsymbol{\lambda}(t), \hat{\boldsymbol{x}}(t), \hat{\boldsymbol{w}}(t)\right) = \boldsymbol{\lambda}^T \hat{\boldsymbol{f}}\left(\hat{\boldsymbol{x}}(t), \boldsymbol{u}(t)\right) + \|\boldsymbol{y}(t) - \boldsymbol{h}\left(\hat{\boldsymbol{x}}\right)\|^2_{Q(t)} + \sum_{n=1}^{N} \mathfrak{h}_n(\lambda_n, w_n) \tag{C.7a}$$

$$\mathfrak{h}_n(\lambda_n, w_n) = \lambda_n \hat{w}_n + \alpha_1 |\hat{w}_n| + \frac{\alpha_2}{2} \hat{w}_n^2. \tag{C.7b}$$

Minimization of the Hamiltonian requires $\mathfrak{h}_n$ to be either zero or negative. From the derivative

$$\frac{\partial \mathfrak{h}_n}{\partial \hat{w}_n} = \lambda_n + \alpha_1 \operatorname{sign}(\hat{w}_n) + \alpha_2 \hat{w} \tag{C.8}$$

we find that the Hamiltonian $H\left(\boldsymbol{\lambda}(t), \hat{\boldsymbol{x}}(t), \hat{\boldsymbol{w}}(t)\right)$ is minimized for

$$\hat{w}_n^* = \begin{cases} -\frac{\lambda_n - \alpha_1}{\alpha_2} & \text{if } \lambda_n > \alpha_1 \\ 0 & \text{if } \lambda_n \in [-\alpha_1, \alpha_1] \\ -\frac{\lambda_n + \alpha_1}{\alpha_2} & \text{if } \lambda_n < -\alpha_1 \, . \end{cases} \tag{C.9}$$

In Figure C.1 we compare the function $\mathfrak{h}_n$ and the location of the minimum w_n^* for the DEN to pure L_1 or L_2 regularization. It can be seen that a pure LASSO-type estimate ($\alpha_2 = 0$) is not useful because the minimum will either be located at zero or at plus or minus infinity. Even when additional box constraints of $a \le \hat{w}_n \le b$ are imposed, the minimum can only jump between zero and the upper bound or zero and the lower bound (Vossen and Maurer, 2006).

Example for an Unidentifiable Model Error

As an example of an unidentifiable model error, we consider the following pair of linear systems

$$\begin{aligned} \dot{x}_1 &= -\alpha x_1 + x_2 + w_1(t) \\ \dot{x}_2 &= -\beta x_2 \\ y &= x_1 \end{aligned} \tag{C.10}$$

and

$$\begin{aligned} \dot{x}_1 &= -\alpha x_1 + x_2 \\ \dot{x}_2 &= -\beta x_2 + w_2(t) \\ y &= x_1 \end{aligned} \tag{C.11}$$

with the target points x_1 and x_2, respectively. Both systems are observable, as can be tested by the Kalman-condition. However, for any given input $w_2(t)$ and for

$$w_1(t) = \int_0^t e^{\beta\tau} w_2(\tau)\, d\tau \tag{C.12}$$

the output $y(t) = x_1(t)$ of both systems is identical. Thus, the model error is not observable.

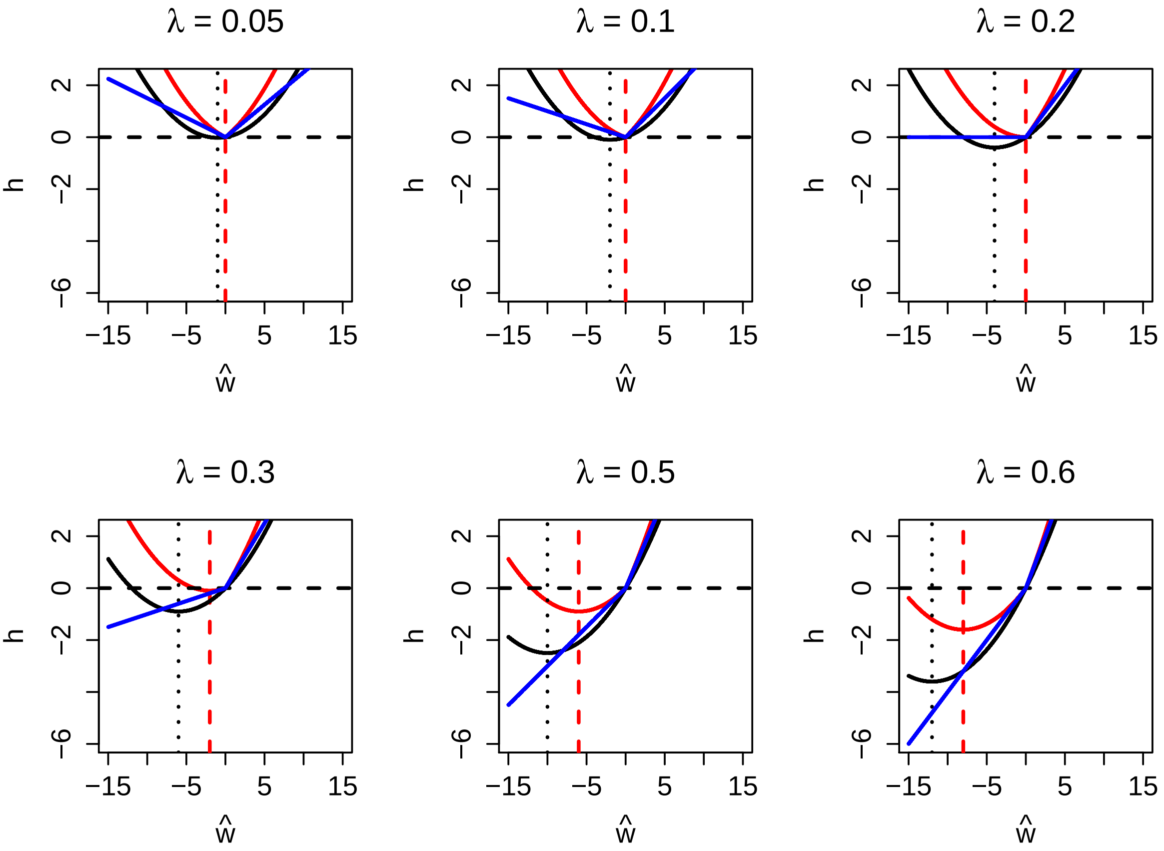


Figure C.1: The function $\mathfrak{h}(\hat{w}) = \lambda\hat{w} + \alpha_1|\hat{w}| + \frac{\alpha_2}{2}\hat{w}^2$ (red) for different values of λ and $(\alpha_1, \alpha_2) = (0.2, 0.05)$. For comparison, the case of pure L_1 regularization ($\alpha_2 = 0$, blue) and pure L_2 regularization ($\alpha_1 = 0$, black) is shown. The dashed lines indicate the location of the minimum for the DEN (red) and for pure L_2 regularization. The minimum of the pure L_1-case is always at zero or plus/minus infinity, if no further bounds are imposed on $\hat{w}$.

Tuning the Regularization Parameters

An important problem for regularized estimates is the choice of the tuning parameters, here α_1 and α_2 for L_1 and L_2 regularization. In regression, these parameters are often chosen by cross validation or bootstrap (Zou and Hastie, 2005). The idea is to estimate the prediction error and to minimize this over a grid of α_1, α_2 values. This could in principle also be done for the DEN. However, the number of replicates of the time course measurements in systems biology is typically too small for resampling from the original data.

Another approach is given by the discrepancy method (Honerkamp and Schelter, 2014). If the DEN output would not incur any bias, then the expectation $E(y_i(t) - \hat{y}_i(t))^2$ could be estimated by the variance σ_i^2 of the measurements. For a diagonal weighting matrix $Q = \mathrm{diag}(q_1, \ldots, q_I)$, we obtain

$$E\left(\sum_{i=1}^{I} q_i(y_i(t) - \hat{y}_i(t))^2\right) \approx \sum_{i=1}^{I} q_i\sigma_i^2 = d \tag{C.13}$$

for the variance of the square error term in Equation (C.2a) provided that the residual measurement errors of the different output components are statistically independent and independent of time. Note, that for $q_i = \sigma_i^{-2}$ we have $d = I$. Thus, the right hand side d provides a first rough criterion to tune the regularization parameters: Adjust α_1 and α_2 in such a way that the squared error term under the integral equals d. This can easily be extended to time dependent errors.

The sparsity of the DEN solution is controlled by α_1. We found empirically that a good approach is to monitor the solution $\hat{\boldsymbol{w}} = (\hat{w}_1, \ldots, \hat{w}_n)^T$ or the area under the hidden influence curves as a function of α_1. Typically, as α_1 is increased from zero, a sparse solution appears, which contains only a few dominant components. Further increasing α_1 does not change these components for a wide range of values, until the regularization dominates the whole error functional (Figure C.2). We thus chose α_1 large enough to obtain a sparse solution and then tuned the L_2-parameter α_2 to obtain a good data fit using the discrepancy (C.13).

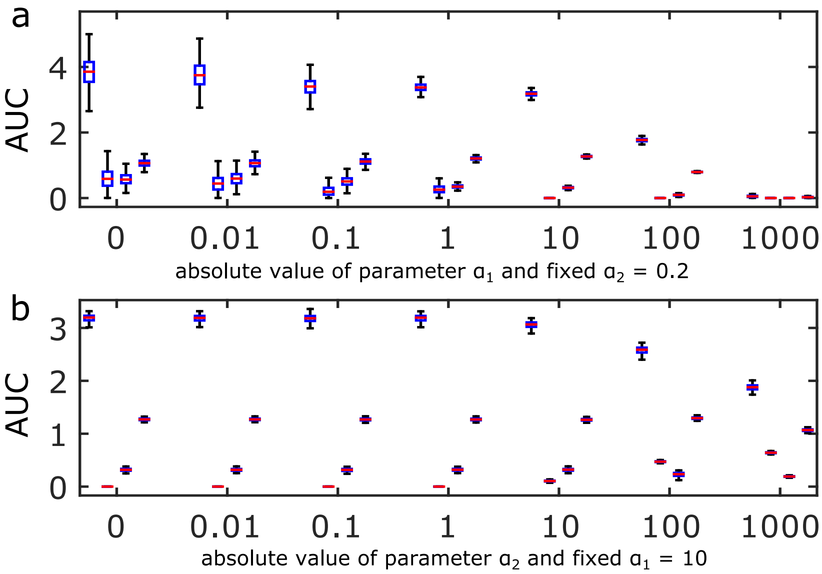


Figure C.2: Influence of the regulation parameters for the JAK-STAT system. (a) The box plots show the variation of the AUC of $|\hat{w}_1(t)|, \ldots, |\hat{w}_4(t)|$ for the DEN estimates caused by different values of the regularization parameter α_1 for $\alpha_2 = 0.2$. To ease visualization, the box plots for a given parameter value are slightly offset. (b) Influence of α_2 for $\alpha_1 = 10$.

The regularization parameters α_1 and α_2 as well as the weighting matrix Q used for the figures of the main text and for the additional examples below are collected in Table C.1.

System	α_1	α_2	$Q \times F$	Figure
JAK-STAT	10	0.2	$\frac{1}{i}\mathrm{diag}\left((\sigma_1^{-2}(t_l), \sigma_2^{-2}(t_l)\right)$	6.2, 6.3, C.4, C.5
JAK-STAT	1	0.2	$\frac{1}{i}\mathrm{diag}\left(\sigma_1^{-2}(t_l), \sigma_2^{-2}(t_l)\right)$	C.3
UV-B	0.1	0.0002	I_n	6.4, 6.5, C.7, C.8, C.9
G protein	0.1	0.002	I_n	C.10
Feed-Forward Loop	0.01	0.002	I_n	C.11
Diamond	0.1	0.002	I_n	C.12
Bi-Fan	0.01	0.002	I_n	C.13

Table C.1: Regularization parameters used in the main text. The standard deviation for output $y_i(t_l)$ at time point t_l is denoted as $\sigma_i(t_l)$. The number of measurement time points is F, the number of output components is I and I_n is a unit matrix.

Parameter Values for the JAK-STAT Model

The estimated parameters and their confidence intervals for the JAK-STAT system were obtained from `http://webber.physik.uni-freiburg.de/~jeti/PNAS_Swameye_Data` and are shown in Table C.2 (Swameye et al., 2003) . Because of the structural non-identifiability of the parameters (Raue et al., 2009), the confidence interval of θ_2 was set to the range $[-3; 1]$. For the same reason, the confidence intervals of θ_5, θ_6 and x_1 (0) were not taken into account. Mass conservation was accounted for by the constraint $2x_4(t) + 2x_3(t) + x_1(t) + x_2(t) = const.$ for all t (Swameye et al., 2003; Raue et al., 2009).

	Value	CI
θ_1	+0.31	$[14; 48]$
θ_2	-1.00	$[-33; 31]$
θ_3	-0.49	$[-1.14; 0.15]$
θ_4	+0.42	$[0.28; 0.56]$
θ_5	-0.21	$[-32; 31]$
θ_6	-0.34	$[-32; 31]$
x_1 (0)	+31	$[-32; 31]$

Table C.2: Parameter values and confidence intervals used for the JAK-STAT model (logarithmic scale).

Additional Analysis of the JAK-STAT Model

Comparison to the Model proposed by Raue *et al.* and to the Thresholded Dynamic Elastic-Net Estimator

In addition to the analysis in the main text, we compared the DEN to the following model published in (Raue et al., 2009):

$$\begin{aligned}\dot{x}_1 &= -\theta_1 x_1 u + 2\theta_4 x_4^\tau \\ \dot{x}_2 &= \theta_1 x_1 u - 2\theta_2 x_2^2 \\ \dot{x}_3 &= \theta_2 x_2^2 - \theta_3 x_3 \\ \dot{x}_4 &= \theta_3 x_3 - \theta_4 x_4^\tau \\ \\ y_1 &= \theta_4 (x_2 + 2x_3) \\ y_2 &= \theta_5 (x_1 + x_2 + 2x_3).\end{aligned} \tag{C.14}$$

This model incorporates cytoplasmic cycling by using the delay term $x_4^\tau = x_4\,(t - \tilde{\tau})$ with $\tilde{\tau} \approx 5$.

The estimates $\hat{\boldsymbol{w}}(t)$ described in the main text are dominated by $\hat{w}_1$ and $\hat{w}_4$. Thus, we fitted another DEN estimate enforcing $\hat{w}_2 = \hat{w}_3 = 0$. This thresholding procedure can be used to decrease the bias caused by regularized least squares estimation (van de Geer et al., 2011) . The delay term x_4^τ in model (C.14) is well represented by the scaled model error $\frac{K}{\theta_4}\hat{w}_4(t)$ of the thresholded estimate (Figure C.3a). The scaling factor $K = \frac{1400\mu m^3}{450\mu m^3}$ accounts for the volume ratio of the cytoplasmic and nuclear compartments.

Taking the Delay Model as Ground Truth

In another test, we used the model (C.14) as ground truth and simulated data using the values of the standard deviation reported in (Swameye et al., 2003). We used the same nominal model as in the main text, which corresponds to a zero delay term $x_4^\tau = 0$. The DEN based on this nominal model correctly identifies the target points x_1 and x_4 (Figure C.4) and reconstructs the delay term $x_4^\tau(t)$ (Figure C.4g).

These results are also robust against measurement noise and parameter uncertainties (Figure C.5). The robustness simulations were performed as in the main text, but now we used the data simulated from the delay model (C.14) as observations.

Sensitivity Against the Number of Measurements

To investigate the sensitivity of the DEN against the number of measurement time points we randomly chose subsets of the original data sets and run the optimization using less input data. As can be seen from Figure C.6, the area under the hidden input curves diagrams indicating the

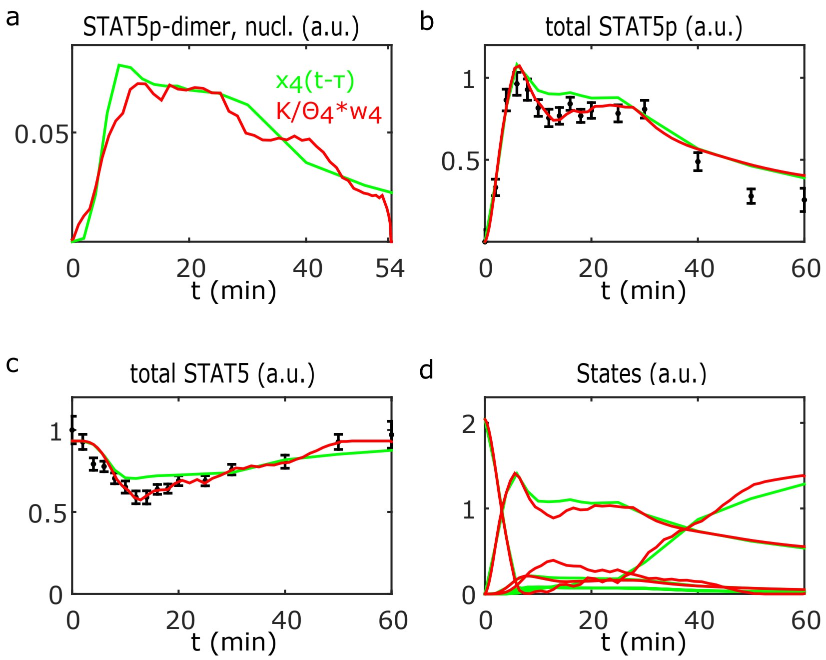


Figure C.3: Comparison of the thresholded DEN estimates (red) and the delay model (C.14) from (Raue et al., 2009) (green) with the real data (black). (a) The delay term x_4^τ representing the cytoplasmic cycling compared to the model error component $\hat{w}_4(t)$ of the thresholded DEN. The scaling factor K accounts for the volume ratio of the cytoplasmic and nuclear compartments. The output of the delay model and the thresholded DEN were compared to the measurements of total STAT5 (b) and total STAT5p (c). (d) Comparison of the state estimates.

nodes targeted by the model error are quite insensitive against a reduction of the number of data points, as long as the main dynamic features of the time courses are still covered by the data.

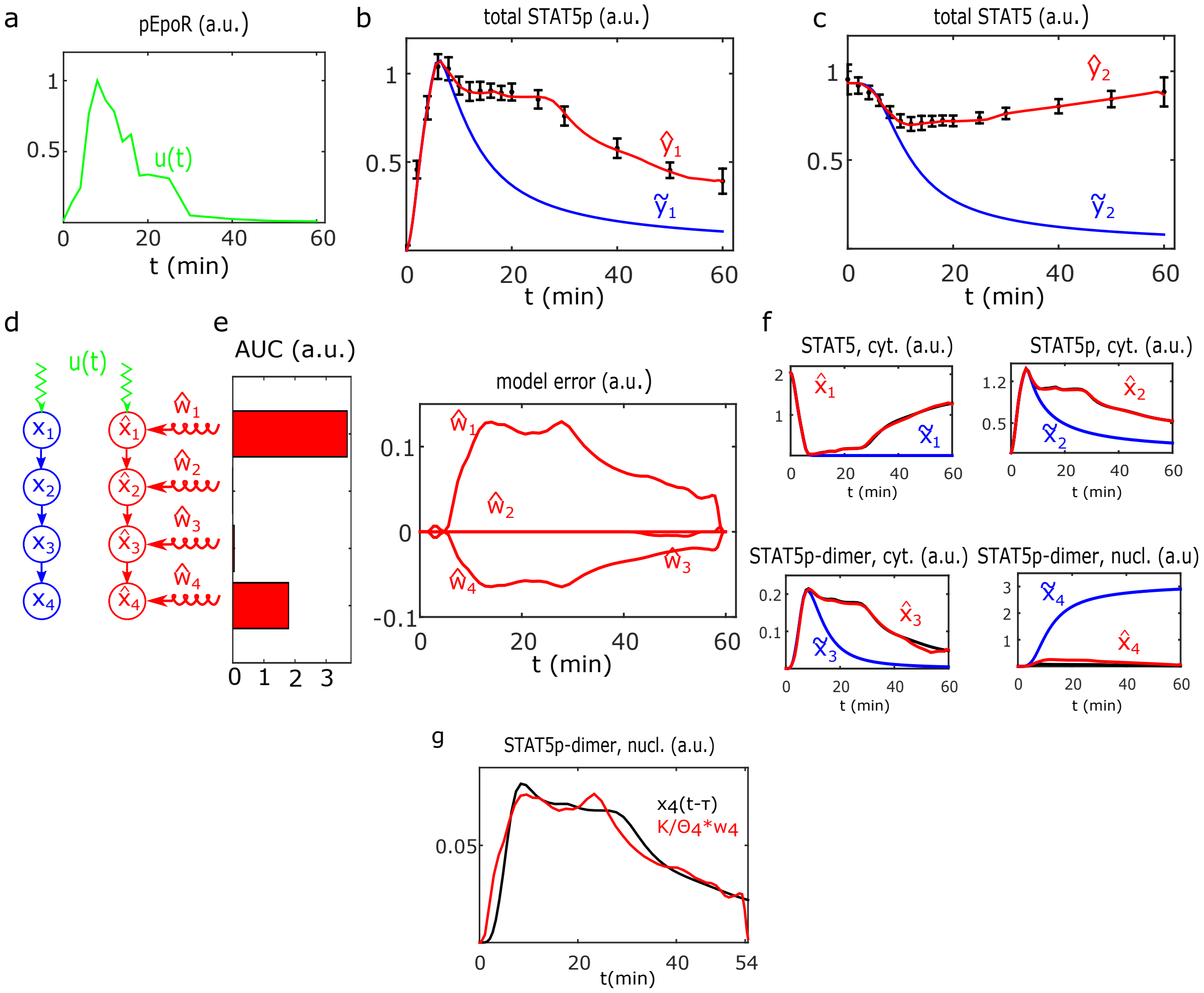


Figure C.4: Reconstructing the model error for data simulated from the delay model (C.14) for the JAK-STAT pathway (Raue et al., 2009). (a) As before, the known input $u(t)$ is given by linearly interpolated phosphorylation measurements for the erythropoietin receptor (Swameye et al., 2003). (b,c) Simulated output measurements (black) for phosphorylated cytoplasmic STAT5 (y_1) and total cytoplasmic STAT5 (y_2) compared to the outputs of the nominal model (blue) and the fit of the DEN (red). (d) Graph of the nominal model (blue) and of the observer system (red) with the state variables cytoplasmic STAT5 (x_1), phosphorylated monomeric STAT5 (x_2), phosphorylated dimeric STAT5 (x_3) and nuclear STAT5 (x_4). (e) DEN estimates of the model error. The AUC indicating the target points of the model error estimates $\hat{w}_1(t), \ldots, \hat{w}_4(t)$. (f) State estimates $\tilde{x}_1, \ldots, \tilde{x}_4$ obtained from the nominal model (blue) and the DEN observer ($\hat{x}_1, \ldots, \hat{x}_4$ in red) compared to the true state (black). (g) Comparison of the delay term (C.14) and the reconstruction by the DEN as already shown in (e).

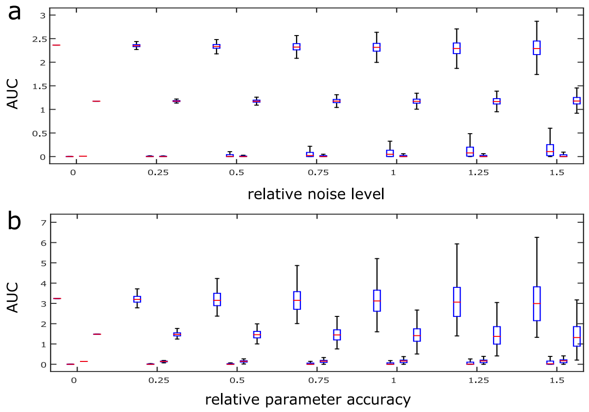


Figure C.5: Robustness against measurement noise and parameter uncertainty using data simulated from the delay model (C.14). (a) Box plots visualizing the variation of the AUC of the DEN estimates caused by measurement noise. To ease visualization, the box plots at a given noise level are slightly offset. (b) Variation of the AUC caused caused by parameter uncertainty.

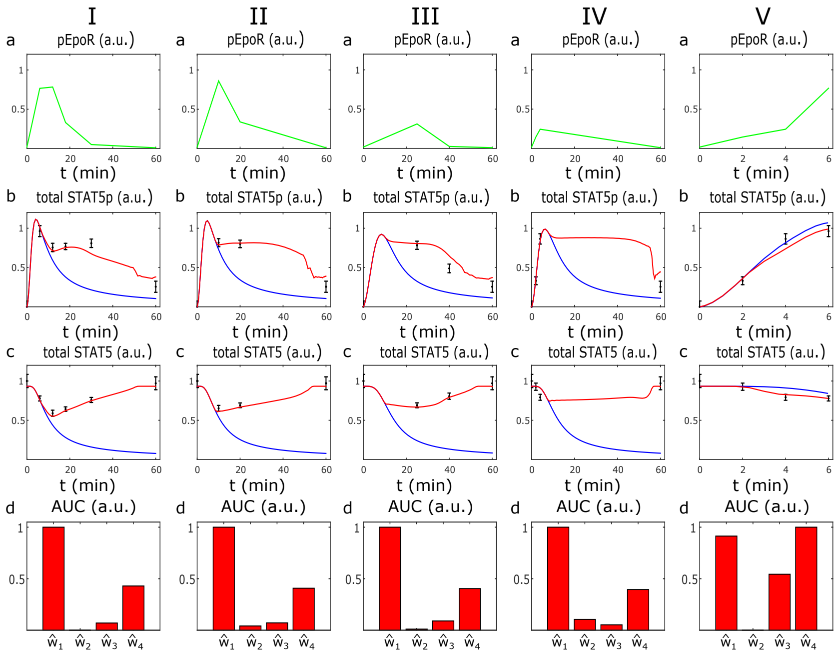


Figure C.6: Sensitivity against the number and location of the measurement time points for the JAK-STAT model (Swameye et al., 2003; Raue et al., 2009). Four different subsets of data were used to fit the DEN (columns I to V). (a) The known input $u(t)$ is given by linearly interpolated phosphorylation measurements for the erythropoietin receptor (Swameye et al., 2003; Raue et al., 2009). (b,c) The simulated output measurements (black) for the phosphorylated cytoplasmic STAT5 (y_1) and total cytoplasmic STAT5 (y_2) compared to the outputs of the nominal model (blue) and the fit of the DEN (red). (d) The AUC indicating the target points of the model error estimates $\hat{w}_1(t), \dots, \hat{w}_4(t)$.

The Photomorphogenic UV-B Signaling Network

The model equations for the UV-B signaling network (Ouyang et al., 2014) were obtained from the BioModels Database (Li et al., 2010), see BIOMD0000000545. The measurement noise of the simulated data in the main text was drawn from a Gaussian distribution with zero mean and standard deviation equal to 5% of the maximum value of the respective output signal.

Suboptimal Solutions for the UV-B Signaling Network

Figures C.7,C.9 and C.8 complement Figure 6.5 in the main text to show that, in cases where the model error is not observable, a perfect fit of the output data can be obtained for different sets of target nodes.

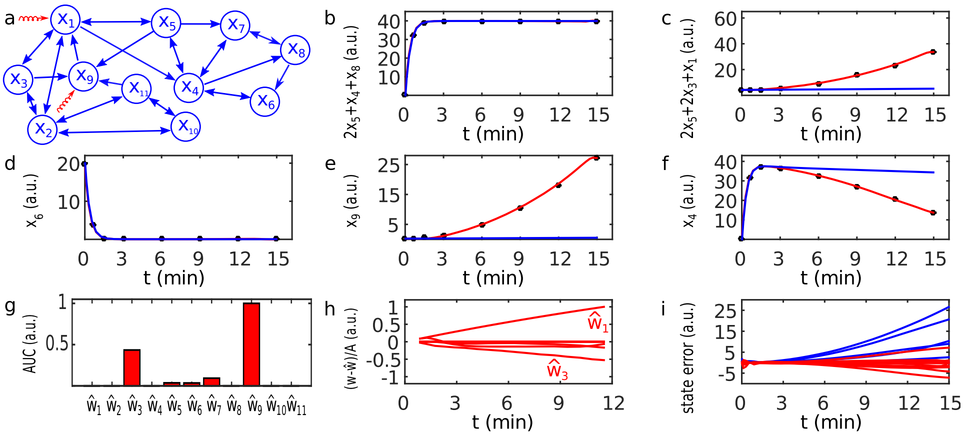


Figure C.7: DEN estimates for the photomorphogenic UV-B signaling network with model error targeting x_1 and x_9.

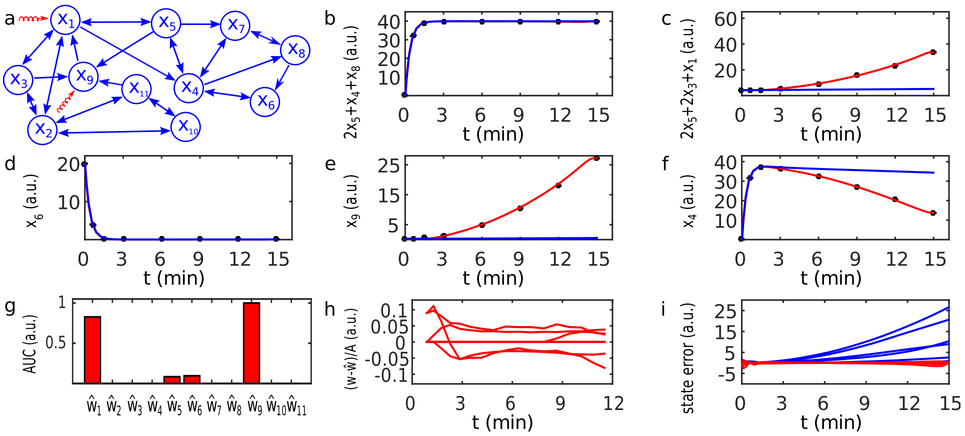


Figure C.8: The corresponding estimate when $\hat{w}_3$ is constrained to zero.

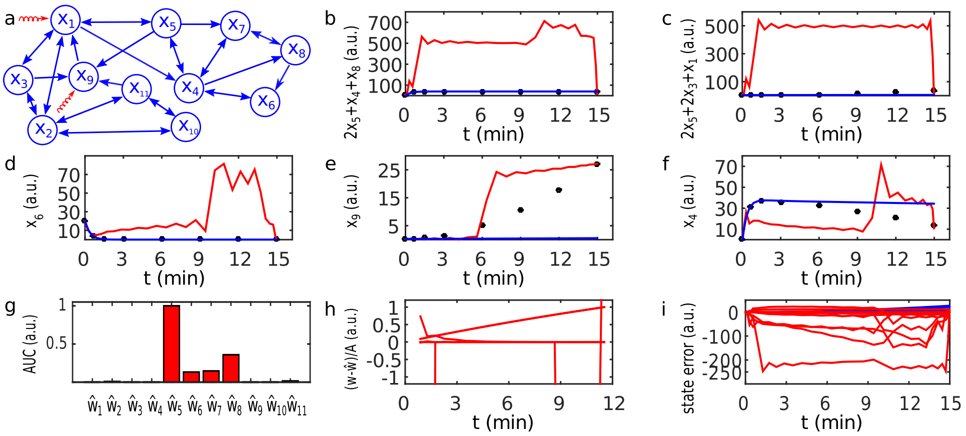


Figure C.9: The corresponding estimate when $\hat{w}_3$ and $\hat{w}_9$ are constrained to zero. The output data can not be fitted with this combination.

Additional Examples

G protein Signaling

A model for the heterotrimeric G protein cycle in yeast was downloaded from the BioModels Database (BIOMD0000000072) (Yi et al., 2003; Li et al., 2010). The model has six state variables: The amount of receptor R, the ligand-bound receptor LR the inactive associated G protein $\overline{\mathrm{GP}}$ and its active subunits GP_{α} and $GP_{\beta\setminus\gamma}$ and the inactive subunit $\overline{GP}_{\alpha}$. We used this model as a nominal model; the equations and parameters can be found in Appendix C.

The true model is assumed to have an additional input to x_4 representing $\overline{\mathrm{GP}}_{\alpha}$, which mimics a stronger regulation of GP_{α}. Synthetic measurement noise was added and the DEN was applied to reconstruct this model error (Figure C.10). All model states were assumed to be directly measurable.

Network Motifs

Network motifs are often considered to be building blocks of different biological networks (Milo, 2002). Their dynamics will of course be influenced by inputs from their exogenous networks. We used three well studied network motifs (the feed-forward loop, the diamond and the bi-fan) to test whether the DEN can reconstruct these hidden inputs from simulated data. We also investigated a phosphorylation cascade receptively feedback loop. Here, we assume that all state variables of the motifs are directly accessible to measurements. The model equations are given in Appendix C. For all networks, we used the unit step function

$$u(t) = 1 - \frac{1}{1+t}$$

as known input stimulus and the initial condition $\boldsymbol{x}(0) = \mathbf{0}$ for the state variables.

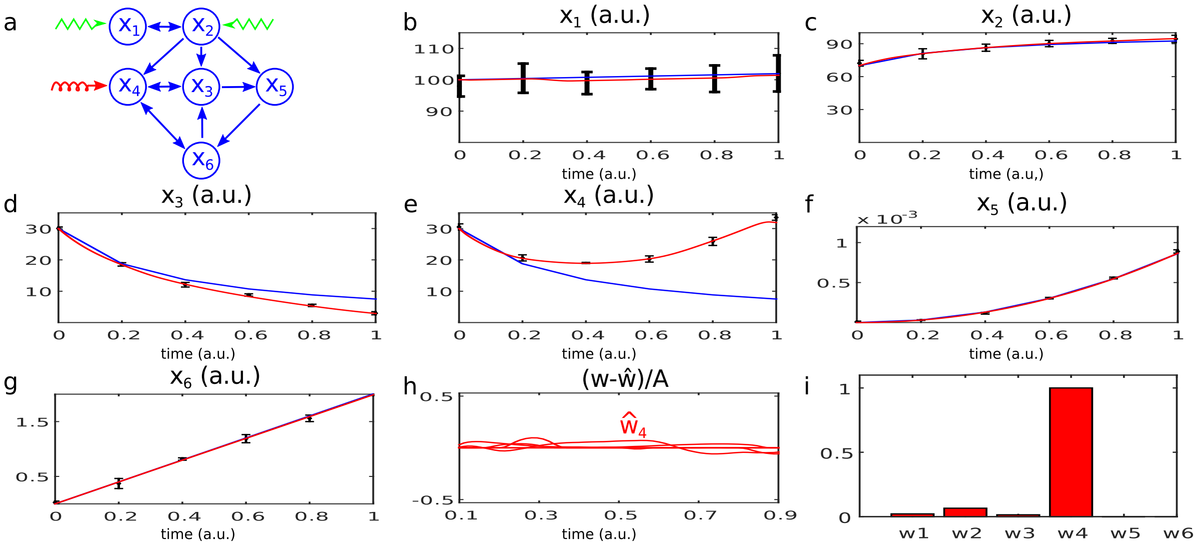


Figure C.10: The heterotrimeric G protein cycle. (a) Graph of the nominal model with known inputs (green) and the additional hidden influence (red) (b-g) Simulated data for species x_i with its standard errors (black) compared to the nominal model (blue) and to the fit of the DEN (red). (h) Discrepancy between the true model error and the DEN estimate, scaled by the maximum value of the true model error. (i) AUC bars for the estimates of the model errors (in percent of total AUC).

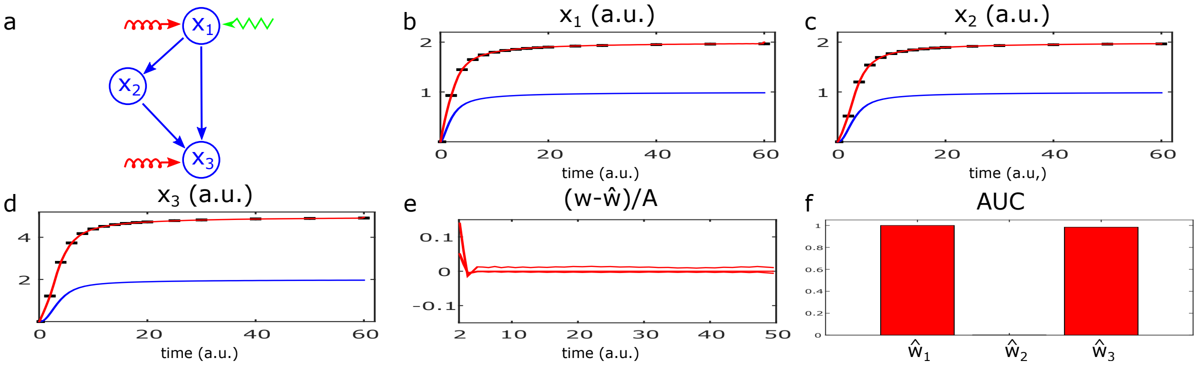


Figure C.11: The feed-forward loop. (a) Graph of the nominal model with known inputs (green) and additional hidden influences (red) (b-d) Simulated data for variables x_i (black) compared to the nominal model (blue) and to the fit of the DEN (red). (e) Discrepancy between the DEN estimate $\hat{\boldsymbol{w}}$ and the true model error caused by hidden inputs from the environment, scaled by the maximum value of the true hidden input. (f) AUC for the DEN estimates.

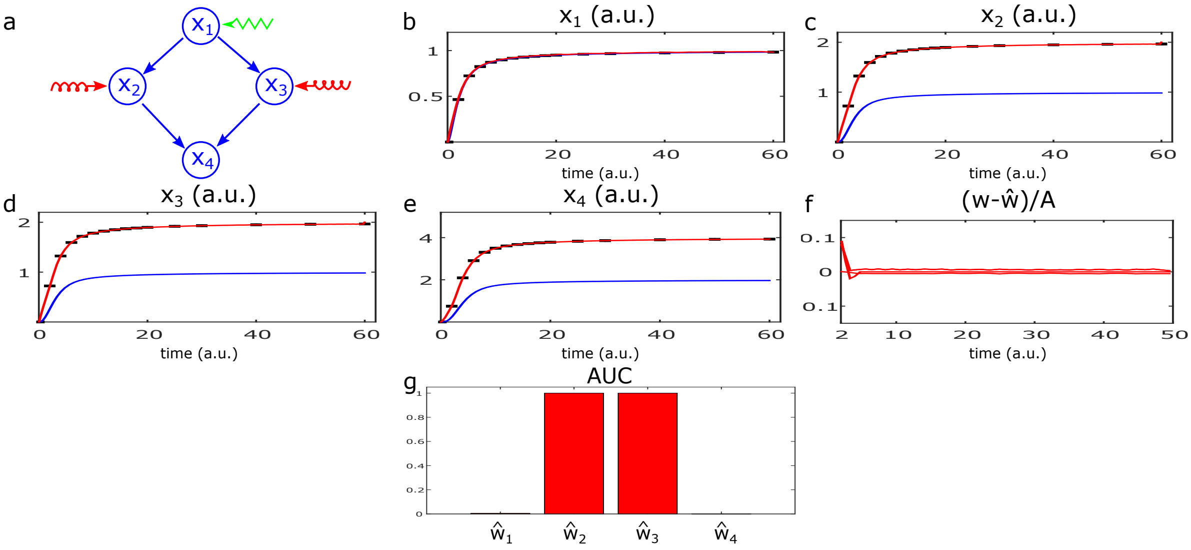


Figure C.12: The diamond motif. (a) Graph of the nominal model with known inputs (green) and additional hidden inputs (red). (b-e) Simulated data for the variables x_i (black) compared to the nominal model (blue) and to the fit of the DEN (red). (f) Discrepancy between the true error signal and the DEN estimate $\hat{\boldsymbol{w}}$, scaled by the maximum value of the true hidden input. (g) AUC for the DEN estimates.

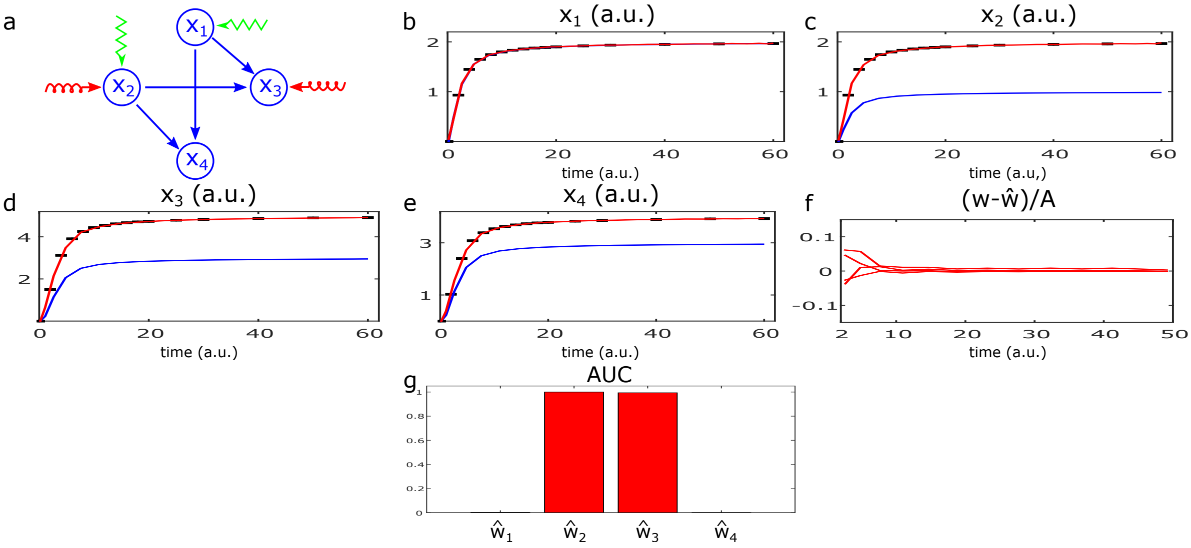


Figure C.13: The bi-fan motif. (a) Graph of the nominal model with known inputs (green) and additional hidden influences (red) (b-e) Simulated data for the variables x_i (black) compared to the nominal model (blue) and the fit of the DEN (red). (f) Discrepancy between the true error signal and the DEN estimate $\hat{\boldsymbol{w}}$, scaled by the maximum value of the true hidden input. (g) AUC for the DEN estimates.

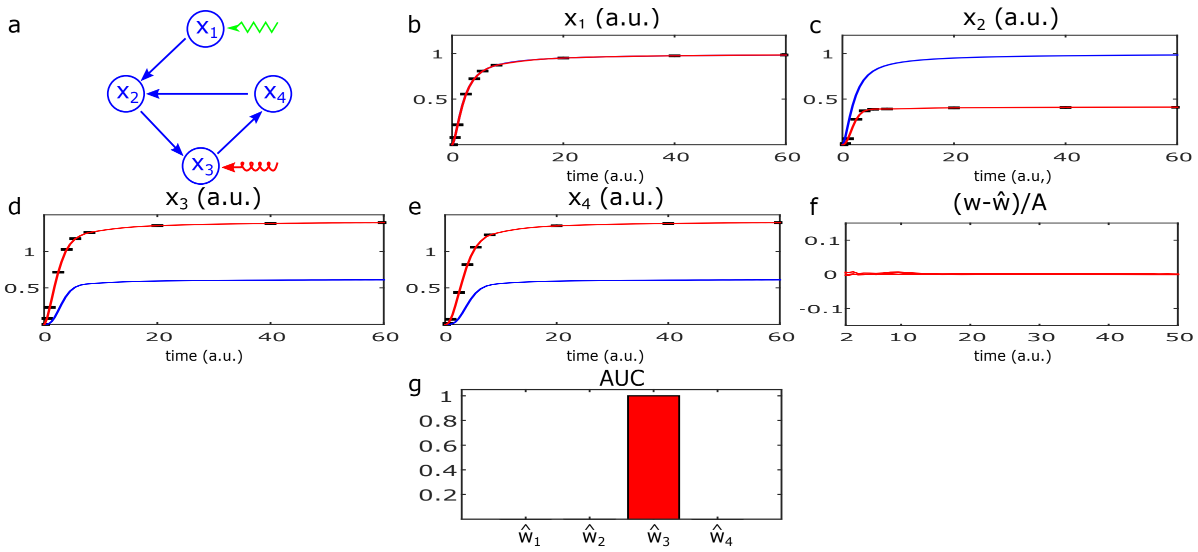


Figure C.14: The feedback loop with exogenous error. (a) Graph of the nominal model with known inputs (green) and additional hidden influences (red) (b-e) Simulated data for variables x_i (black) compared to the nominal model (blue) and the fit of the DEN (red). (f) Discrepancy between the true error signal and the DEN estimate $\hat{\boldsymbol{w}}$, scaled by the maximum value of the true hidden input. (g) AUC for the DEN estimates.

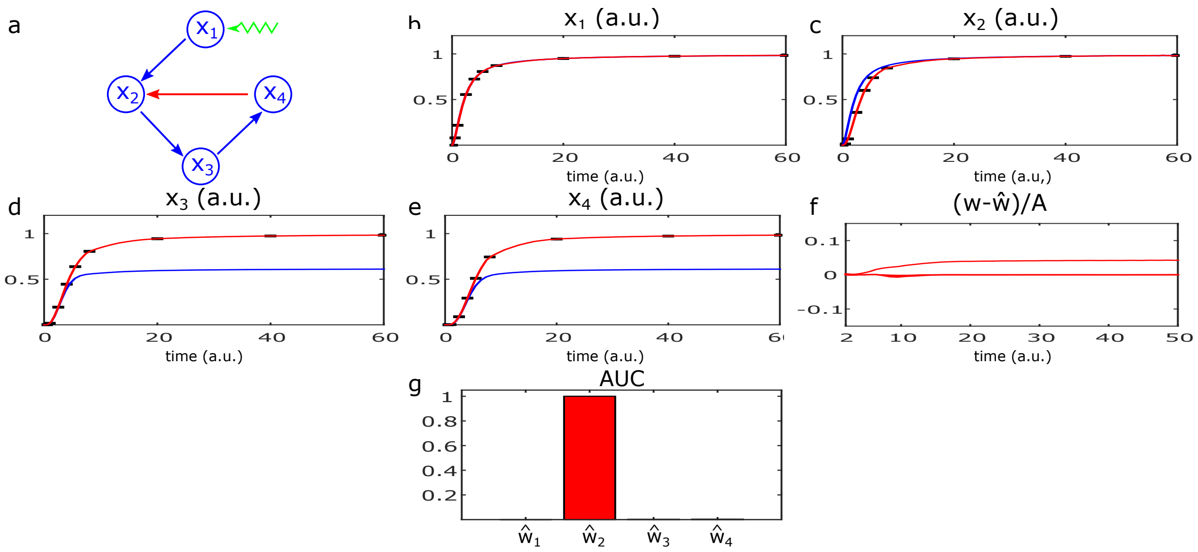


Figure C.15: The feedback loop with endogenous error. (a) Graph of the nominal model with known inputs (green) and missing feedback (red) (b-e) Simulated data for the variables x_i (black) compared to the nominal model (blue) and the fit of the DEN (red). (f) Discrepancy between the true error signal and the DEN estimate $\hat{\boldsymbol{w}}$, scaled by the maximum value of the true hidden input. (g) AUC for the DEN estimates.

Further Model Equations

Feed-forward Loop

True System

$$\begin{aligned}
\dot{x}_1(t) &= u(t) - x_1(t) + \left(1 - \frac{1}{1+t}\right) \\
\dot{x}_2(t) &= x_1(t) - x_2(t) \\
\dot{x}_3(t) &= x_2(t) + x_1(t) - x_3(t) + \left(1 - \frac{1}{1+t}\right)
\end{aligned}$$

Nominal System

$$\begin{aligned}
\dot{\tilde{x}}_1(t) &= u(t) - \tilde{x}_1(t) \\
\dot{\tilde{x}}_2(t) &= \tilde{x}_1(t) - \tilde{x}_2(t) \\
\dot{\tilde{x}}_3(t) &= \tilde{x}_2(t) + \tilde{x}_1(t) - \tilde{x}_3(t)
\end{aligned}$$

Diamond

True System

$$\begin{aligned}
\dot{x}_1(t) &= u(t) - x_1(t) \\
\dot{x}_2(t) &= x_1(t) - x_2(t) + \left(1 - \frac{1}{1+t}\right) \\
\dot{x}_3(t) &= x_1(t) - x_3(t) + \left(1 - \frac{1}{1+t}\right) \\
\dot{x}_4(t) &= x_2(t) + x_3(t) - x_4(t)
\end{aligned}$$

Nominal System

$$\begin{aligned}
\dot{\tilde{x}}_1(t) &= u(t) - \tilde{x}_1(t) \\
\dot{\tilde{x}}_2(t) &= \tilde{x}_1(t) - \tilde{x}_2(t) \\
\dot{\tilde{x}}_3(t) &= \tilde{x}_1(t) - \tilde{x}_3(t) \\
\dot{\tilde{x}}_4(t) &= \tilde{x}_2(t) + \tilde{x}_3(t) - \tilde{x}_4(t)
\end{aligned}$$

Bi-fan

True System

$$\begin{aligned}
\dot{x}_1(t) &= u(t) - x_1(t) \\
\dot{x}_2(t) &= u(t) - x_2(t) + \left(1 - \frac{1}{1+t}\right) \\
\dot{x}_3(t) &= \tilde{x}_2(t) + x_1(t) - x_3(t) - \left(1 - \frac{1}{1+t}\right) \\
\dot{x}_4(t) &= \tilde{x}_2(t) + x_1(t) - x_4(t)
\end{aligned}$$

Nominal System

$$\begin{aligned}
\dot{\tilde{x}}_1(t) &= u(t) - \tilde{x}_1(t) \\
\dot{\tilde{x}}_2(t) &= u(t) - \tilde{x}_2(t) \\
\dot{\tilde{x}}_3(t) &= \tilde{x}_2(t) + \tilde{x}_1(t) - \tilde{x}_3(t) \\
\dot{\tilde{x}}_4(t) &= \tilde{x}_2(t) + \tilde{x}_1(t) - \tilde{x}_4(t)
\end{aligned}$$

Feedback Loop 1

True System

$$\begin{aligned}
\dot{\tilde{x}}_1(t) &= u(t) - x_1(t) \\
\dot{\tilde{x}}_2(t) &= \frac{x_1(t)}{1 + x_4} - x_2(t) \\
\dot{\tilde{x}}_3(t) &= x_2(t) - x_3(t) + \left(1 - \frac{1}{1 + t}\right) \\
\dot{\tilde{x}}_4(t) &= x_3(t) - x_4(t)
\end{aligned}$$

Nominal System

$$\begin{aligned}
\dot{\tilde{x}}_1(t) &= u(t) - \tilde{x}_1(t) \\
\dot{\tilde{x}}_2(t) &= \frac{\tilde{x}_1(t)}{1 + \tilde{x}_4} - \tilde{x}_2(t) \\
\dot{\tilde{x}}_3(t) &= \tilde{x}_2(t) - \tilde{x}_3(t) \\
\dot{\tilde{x}}_4(t) &= \tilde{x}_3(t) - \tilde{x}_4(t)
\end{aligned}$$

Feedback Loop 2

True System

$$\begin{aligned}
\dot{x}_1(t) &= u(t) - x_1(t) \\
\dot{x}_2(t) &= x_1 - x_2(t) \\
\dot{x}_3(t) &= x_2(t) - x_3(t) \\
\dot{\tilde{x}}_4(t) &= x_3(t) - x_4(t)
\end{aligned}$$

Nominal System

$$\begin{aligned}
\dot{\tilde{x}}_1(t) &= u(t) - \tilde{x}_1(t) \\
\dot{\tilde{x}}_2(t) &= \frac{\tilde{x}_1(t)}{1 + \tilde{x}_4} - \tilde{x}_2(t) \\
\dot{\tilde{x}}_3(t) &= \tilde{x}_2(t) - \tilde{x}_3(t) \\
\dot{\tilde{x}}_4(t) &= \tilde{x}_3(t) - \tilde{x}_4(t)
\end{aligned}$$

G protein Signaling Model

The model for the heterotrimeric G protein cycle in yeast was downloaded from the BioModels Database (BIOMD0000000072) (Yi et al., 2003; Li et al., 2010). This model was used as a nominal model in Appendix C, but the tilde is suppressed for simplicity of notation.

$$
\begin{aligned}
\frac{\mathrm{d}([R])}{\mathrm{d}t} &= -\left(3.32e^{-18}\right)[L]_{const.}\,[\mathrm{R}] + 0.01\,[LR] - 4 - 0.0004\,[R] \\
\frac{\mathrm{d}([LR])}{\mathrm{d}t} &= \left(3.32e^{-18}\right)[L]_{const.}\,[R] - 0.01\,[LR] - 0.004\,[LR] \\
\frac{\mathrm{d}\left(\left[\overline{\mathrm{GP}}\right]\right)}{\mathrm{d}t} &= \left[\overline{\mathrm{GP}}_{\alpha}\right]\left[GP_{\beta\backslash\gamma}\right] - \left(1e^{-05}\right)[LR]\left[\overline{\mathrm{GP}}\right] \\
\frac{\mathrm{d}\left(\left[GP_{\beta\backslash\gamma}\right]\right)}{\mathrm{d}t} &= -\left[\overline{\mathrm{GP}}_{\alpha}\right]\left[GP_{\beta\backslash\gamma}\right] + \left(1e^{-05}\right)[LR]\left[\overline{\mathrm{GP}}\right] \\
\frac{\mathrm{d}([\mathrm{GP}_{\alpha}])}{\mathrm{d}t} &= \left(1e^{-05}\right)[LR]\left[\overline{\mathrm{GP}}\right] - 0.11\,[\mathrm{GP}_{\alpha}] \\
\frac{\mathrm{d}\left(\left[\overline{\mathrm{GP}}_{\alpha}\right]\right)}{\mathrm{d}t} &= -\left[\overline{\mathrm{GP}}_{\alpha}\right]\left[GP_{\beta\backslash\gamma}\right] + 0.11\,[\mathrm{GP}_{\alpha}]
\end{aligned}
$$

These interactions are depicted in Figure C.10a. The initial conditions and the names of the state variables are given by

$$
\begin{aligned}
[R] &= x_1 & \qquad \left[GP_{\beta\backslash\gamma}\right] &= x_4 \\
[LR] &= x_2 & \qquad [\mathrm{GP}_{\alpha}] &= x_5 \\
\left[\overline{\mathrm{GP}}\right] &= x_3 & \qquad \left[\overline{\mathrm{GP}}_{\alpha}\right] &= x_6
\end{aligned}
$$

$$
\begin{aligned}
[R]_0 &= 100\ \mathrm{item}\backslash\mathrm{l} & \qquad \left[\overline{\mathrm{GP}}_{\alpha}\right]_0 &= 30\ \mathrm{item}\backslash\mathrm{l} \\
\left[\overline{\mathrm{GP}}\right]_0 &= 70\ \mathrm{item}\backslash\mathrm{l} & \qquad [\mathrm{GP}_{\alpha}]_0 &= 0\ \mathrm{item}\backslash\mathrm{l} \\
\left[GP_{\beta\backslash\gamma}\right]_0 &= 30\ \mathrm{item}\backslash\mathrm{l} & \qquad [LR]_0 &= 0\ \mathrm{item}\backslash\mathrm{l} \\
[L]_{const.} &= 0.02\ \mathrm{item}\backslash\mathrm{l} & &
\end{aligned}
$$

UV-B Network Model

The model equations for the UV-B signaling network were obtained from the BioModels Database (Li et al., 2010; Ouyang et al., 2014), see BIOMD0000000545. Although this model was used as a nominal model in the main text, the tilde is suppressed for simplicity of notation.

$$
\begin{aligned}
\frac{d\,[\mathrm{CS}]}{dt} =& -2 \cdot ka_1 \cdot [\mathrm{CS}]^2 \cdot [\mathrm{UVR8M}]^2 + 2kd_1 \cdot [\mathrm{UCS}] \\
&+ ks_1 \cdot (1 + UV \cdot n_3 \cdot ([\mathrm{HY5}] + FHY3)) \\
&- kdr_1 \cdot (1 + (n_1 \cdot UV)) \cdot [\mathrm{CS}] - kd_2 \cdot [\mathrm{CDCS}] \\
&- 2 \cdot ka_2 \cdot [\mathrm{CS}]^2 \cdot [\mathrm{CD}] \\
\frac{d\,[\mathrm{CD}]}{dt} =& - ka_2 \cdot [\mathrm{CS}]^2 \cdot [\mathrm{CD}] + kd_2 \cdot [\mathrm{CDCS}] \\
&+ ka_4 \cdot [\mathrm{CD}] \cdot [\mathrm{DWD}] + kd_4 \cdot [\mathrm{CDW}] \\
\frac{d\,[\mathrm{CDCS}]}{dt} =& - kd_2 \cdot [\mathrm{CDCS}] + ka_2 \cdot [\mathrm{CS}]^2 \cdot [\mathrm{CD}] \\
\frac{d\,[\mathrm{UVR8M}]}{dt} =& -2 \cdot k_1 \cdot [\mathrm{UVR8M}]^2 + 2 \cdot k_2 \cdot [\mathrm{UVR8D}] \\
&- 2 \cdot ka_1 \cdot [\mathrm{CS}]^2 \cdot [\mathrm{UVR8M}]^2 + 2 \cdot kd_1 \cdot [\mathrm{UCS}] \\
&- ka_3 \cdot [\mathrm{UVR8M}] \cdot [\mathrm{RUP}] \\
\frac{d\,[\mathrm{UCS}]}{dt} =& - kd_1 \cdot [\mathrm{UCS}] + ka_1 \cdot [\mathrm{CS}]^2 \cdot [\mathrm{UVR8M}]^2 \\
\frac{d\,[\mathrm{UVR8D}]}{dt} =& - k_2 \cdot [\mathrm{UVR8D}] + k_1 \cdot [\mathrm{UVR8M}]^2 + kd_3 \cdot [\mathrm{UR}]^2 \\
\frac{d\,[\mathrm{RUP}]}{dt} =& - ka_3 \cdot [\mathrm{UVR8M}] \cdot [\mathrm{RUP}] + ks_2 \cdot (1 + UV \cdot [\mathrm{UCS}]) \\
&- kdr2 \cdot [\mathrm{RUP}] + (2) \cdot kd_3 \cdot [\mathrm{UR}]^2 \\
\frac{d\,[\mathrm{UR}]}{dt} =& -2 \cdot kd_3 \cdot [\mathrm{UR}]^2 + ka_3 \cdot [\mathrm{UVR8M}] \cdot [\mathrm{RUP}] \\
\frac{d\,[\mathrm{HY5}]}{dt} =& - kdr3 \cdot \left(\frac{[\mathrm{CDCS}]}{kdr_{3a} + [\mathrm{CDCS}]} + \frac{[\mathrm{CDW}]}{kdr_{3b} + [\mathrm{CDW}]} \right) \cdot [\mathrm{HY5}] \\
&+ ks3_p \cdot (1 + n_2 \cdot UV) - kdr3 \cdot \left(\frac{[\mathrm{UCS}]}{ksr + [\mathrm{UCS}]} \right) \cdot [\mathrm{HY5}] \\
\frac{d\,[\mathrm{DWD}]}{dt} =& - ka_4 \cdot [\mathrm{CD}] \cdot [\mathrm{DWD}] + kd_4 \cdot [\mathrm{CDW}] \\
\frac{d\,[\mathrm{CDW}]}{dt} =& - kd_4 \cdot [\mathrm{CDW}] + ka_4 \cdot [\mathrm{CD}] \cdot [\mathrm{DWD}]
\end{aligned}
$$

$$
\begin{aligned}
UM_{Total} &= 2 \cdot [\mathrm{UCS}] + [\mathrm{UVR8M}] + [\mathrm{UR}] \\
COP1_{Total} &= 2 \cdot [\mathrm{UCS}] + 2 \cdot [\mathrm{CDCS}] + [\mathrm{CS}] \\
UVR8D_{obs.} &= [\mathrm{UVR8D}] \\
HY5_{obs.} &= [\mathrm{HY5}] \\
UVR8M_{obs} &= [\mathrm{UVR8M}]
\end{aligned}
$$

$$
\begin{aligned}
[\mathrm{CS}] &= x_1 & [\mathrm{UVR8M}] &= x_4 & [\mathrm{RUP}] &= x_7 & [\mathrm{DWD}] &= x_{10} \\
[\mathrm{CD}] &= x_2 & [\mathrm{UCS}] &= x_5 & [\mathrm{UR}] &= x_8 & [\mathrm{CDW}] &= x_{11} \\
[\mathrm{CDCS}] &= x_3 & [\mathrm{UVR8D}] &= x_6 & [\mathrm{HY5}] &= x_9 & &
\end{aligned}
$$

$$
\begin{aligned}
ks_1 &= 0.23 & ks_2 &= 4.0526 & k_1 &= 0.0043 \\
kdr_1 &= 0.1 & kdr_2 &= 0.2118 & k_2 &= 161.62 \\
ka_1 &= 0.0372 & ka_2 &= 0.0611 & ka_3 &= 4.7207 \\
kd_2 &= 50.6973 & kd_3 &= 0.5508 & kd_1 &= 94.3524 \\
ks_3 &= 0.4397 & kdr_3 &= 1.246 & UV &= 1 \\
kd_4 &= 1.1999 & n_1 &= 3 & ka_4 &= 10.1285 \\
n_2 &= 2 & n_3 &= 3.5 & kdr_{3a} &= 0.9735 \\
ksr &= 0.7537 & FHY3 &= 5 & kdr_{3b} &= 0.406
\end{aligned}
$$

$$
\begin{aligned}
[\mathrm{CS}]_0 &= 0.2\ \mathrm{mol} & [\mathrm{RUP}]_0 &= 0\ \mathrm{mol} \\
[\mathrm{CD}]_0 &= 10\ \mathrm{mol} & [\mathrm{UR}]_0 &= 0\ \mathrm{mol} \\
[\mathrm{CDCS}]_0 &= 2\ \mathrm{mol} & [\mathrm{HY5}]_0 &= 0.25\ \mathrm{mol} \\
[\mathrm{UVR8M}]_0 &= 0\ \mathrm{mol} & [\mathrm{DWD}]_0 &= 20\ \mathrm{mol} \\
[\mathrm{UCS}]_0 &= 0\ \mathrm{mol} & [\mathrm{CDW}]_0 &= 0\ \mathrm{mol} \\
[\mathrm{UVR8D}]_0 &= 20\ \mathrm{mol} & &
\end{aligned}
$$

APPENDIX D

The Bayesian Dynamic Elastic-Net

Further results

Model	Error Level	Method	AUROC	BS
JAK-STAT	0%	BDEN	0.90 (0.15)	0.11 (0.11)
		DEN	0.60 (0.40)	0.16 (0.06)
	2.5%	BDEN	0.67 (0.15)	0.16 (0.10)
		DEN	0.51 (0.32)	0.32 (0.12)
	7.5%	BDEN	0.61 (0.18)	0.19(0.15)
		DEN	0.50 (0.33)	0.33 (0.13)
	12.5%	BDEN	0.59 (0.08)	0.26(0.16)
		DEN	0.43 (0.33)	0.33 (0.13)
G protein	0%	BDEN	1.00 (0.00)	0.04 (0.03)
		DEN	1.00 (0.00)	0.09 (0.02)
	2.5%	BDEN	1.00 (0.00)	0.05 (0.03)
		DEN	1.00 (0.00)	0.01 (0.01)
	7.5%	BDEN	0.96 (0.06)	0.12 (0.06)
		DEN	0.76 (0.20)	0.42 (0.05)
	12.5%	BDEN	0.95 (0.07)	0.15 (0.09)
		DEN	0.63 (0.23)	0.46 (0.05)
UV-B	0%	BDEN	0.91 (0.11)	0.19 (0.06)
		DEN	0.80 (0.19)	0.22 (0.08)
	2.5%	BDEN	0.78 (0.11)	0.14 (0.06)
		DEN	0.75 (0.14)	0.20 (0.06)
	7.5%	BDEN	0.76 (0.10)	0.15 (0.04)
		DEN	0.72 (0.15)	0.20 (0.06)
	12.5%	BDEN	0.73 (0.11)	0.16 (0.05)
		DEN	0.70 (0.19)	0.22 (0.07)
Motifs	0%	BDEN	1.00 (0.00)	0.01 (0.00)
		DEN	0.90 (0.14)	0.11 (0.09)
	2.5%	BDEN	1.00 (< 0.01)	0.01 (0.01)
		DEN	0.93 (0.11)	0.06 (0.08)
	7.5%	BDEN	1.00 (< 0.01)	(< 0.01) (< 0.01)
		DEN	0.93 (0.11)	0.06 (0.08)
	12.5%	BDEN	0.99 (0.03)	0.01 (0.02)
		DEN	0.94 (0.10)	0.06 (0.08)

Table D.1: Performance of the BDEN and the DEN in dependence on the error of kinetic parameter estimates (median) for fixed measurement noise of 2.5%. The median absolute deviation for the AUROC and Brier Scores are given in brackets.

Percentage Hidden influences	AUROC	BS
17%	1.00 (< 0.01)	0.02 (0.02)
33%	0.96 (0.06)	0.11 (0.09)
50%	0.86 (0.12)	0.22 (0.17)

Table D.2: Performance of the BDEN in dependence on an increasing number of hidden influences for the G protein cycle in yeast relative to the number of nodes in the nominal model (median) for a fixed measurement noise of 2.5%. The median absolute deviations of the AUROC and Brier Scores are given in brackets.

	Model	Noise	Class.	Overall
Missing Interaction	JAK-STAT	2.5%	1.00 (0.01)	0.80 (0.24)
		7.5%	1.00 (0.03)	0.70 (0.30)
		12.5%	0.91 (0.10)	0.70 (0.31)
	G protein	2.5%	1.00 (0.05)	0.71 (0.27)
		7.5%	0.93 (0.09)	0.58 (0.25)
		12.5%	0.85 (0.16)	0.54 (0.33)
	UV-B	2.5%	0.93 (0.09)	0.53 (0.15)
		7.5%	0.87 (0.15)	0.53 (0.09)
		12.5%	0.80 (0.23)	0.50 (0.14)

	Model	Noise	Class.	Overall
Wrong Interaction	JAK-STAT	2.5%	1.00 (0.02)	0.87 (0.19)
		7.5%	0.91 (0.06)	0.85 (0.19)
		12.5%	0.89 (0.12)	0.85 (0.29)
	G protein	2.5%	1.00 (0.01)	0.67 (0.19)
		7.5%	0.92 (0.14)	0.64 (0.10)
		12.5%	0.83 (0.25)	0.64 (0.18)
	UV-B	2.5%	0.91 (0.08)	0.68 (0.20)
		7.5%	0.87 (0.07)	0.61 (0.28)
		12.5%	0.83 (0.13)	0.50 (0.25)

Table D.3: Performance of of the BDEN correctly detect and classify interactions in dependence on the level of relative measurement noise (median AUROC). The column "Class." reflects the accuracy for calling a correctly detected hidden influence as "missing/wrong stimulation" and "missing/wrong inhibition", respectively. In contrast, the column "Overall" reflects the accuracy for correctly detecting a hidden influence AND correctly classifying it as missing/wrong interaction AND calling it correctly "simulation" and "inhibition", respectively.

	Model	Error Level	AUROC
Missing Interaction	JAK-STAT	0%	1.00 (0.00)
		2.5%	1.00 (0.00)
		7.5%	0.83 (0.28)
		12.5%	0.67 (0.44)
	G protein	0%	0.81 (0.19)
		2.5%	0.85 (0.21)
		7.5%	0.72 (0.40)
		12.5%	0.68 (0.43)
	UV-B	0%	1.00 (0.00)
		2.5%	1.00 (0.00)
		7.5%	0.75 (0.38)
		12.5%	0.65 (0.42)

	Model	Error Level	AUROC
Wrong Interaction	JAK-STAT	0%	1.00 (0.00)
		2.5%	0.93 (0.13)
		7.5%	0.92 (0.14)
		12.5%	0.83 (0.28)
	G protein	0%	1.00 (0.00)
		2.5%	0.94 (0.10)
		7.5%	0.89 (0.20)
		12.5%	0.76 (0.34)
	UV-B	0%	0.81 (0.20)
		2.5%	0.84 (0.22)
		7.5%	0.83 (0.18)
		12.5%	0.74 (0.32)

Table D.4: Performance of the BDEN to detect wrong and missing interactions in dependence on the error of the kinetic parameter estimates (median) for measurement noise of 2.5%.

	Model	Error	Class.	Overall
Missing Interaction	JAK-STAT	0%	1.00 (0.01)	0.80 (0.24)
		2.5%	1.00 (0.08)	0.63 (0.19)
		7.5%	0.94 (0.10)	0.58 (0.23)
		12.5%	0.88 (0.12)	0.57 (0.26)
	G protein	0%	1.00 (0.05)	0.71 (0.27)
		2.5%	1.00 (0.03)	0.68 (0.32)
		7.5%	0.95 (0.11)	0.67 (0.27)
		12.5%	0.84 (0.13)	0.68 (0.25)
	UV-B	0%	0.93 (0.08)	0.53 (0.15)
		2.5%	0.93 (0.08)	0.52 (0.08)
		7.5%	0.85 (0.11)	0.52 (0.11)
		12.5%	0.78 (0.09)	0.52 (0.12)

	Model	Error	Class.	Overall
Wrong Interaction	JAK-STAT	0%	1.00 (0.02)	0.87 (0.19)
		7.5%	0.99 (0.06)	0.82 (0.27)
		12.5%	0.91 (0.11)	0.77 (0.31)
	G protein	0%	1.00 (0.01)	0.67 (0.19)
		2.5%	1.00 (0.05)	0.62 (0.27)
		7.5%	0.93 (0.10)	0.52 (0.34)
		12.5%	0.81 (0.12)	0.48 (0.32)
	UV-B	0%	0.91 (0.08)	0.68 (0.20)
		2.5%	0.89 (0.09)	0.64 (0.18)
		7.5%	0.83 (0.12)	0.61 (0.22)
		12.5%	0.76 (0.16)	0.53 (0.32)

Table D.5: Performance of the BDEN to correctly detect and classify interactions in dependence on the error of the kinetic parameter estimates (median AUROC) for a fixed measurement noise of 2.5%. The column "Class." reflects the accuracy for calling a correctly detected hidden influence as "missing/wrong simulation" and "missing/wrong inhibition", respectively. In contrast, the column "Overall" reflects the accuracy for correctly detecting a hidden influence AND correctly classifying it as wrong interaction AND calling it correctly "simulation" and "inhibition", respectively.

Details on the Simulation Studies

Simulation of synthetic data was done as follows: First, a published ODE-based state-observation model with respective initial conditions was employed to generate noise-free observations at 6 time points, depending on the respective model (JAK-STAT, UV-B signaling, G protein cycling, network motifs). Subsequently Gaussian noise with fixed variance was added. Notably an (unrealistic) noise variance of zero is resulting into numerical problems for the BDEN and was thus not considered.

Hidden influences were simulated by adding one (Tables 7.2, D.1) or up to three randomly picked (Table D.2) state variables on the right hand side of the ODE system, yielding a wrong nominal ODE system. Likewise, wrong interactions were simulated by randomly adding one interaction on the right hand side of the ODE system (Tables 7.3, D.3, D.4, D.5). Missing interactions were simulated by randomly removing one of the existing interactions (Tables 7.3, D.3, D.4, D.5).

Due to the randomness of the process described above we repeated all simulations a number of times (Table D.6). The exact number of repeats varies slightly to take into account the different number of combination possibilities. ODE systems with more state variables or allow for a larger number of possible combinations of hidden influence signals.

Table	Repeats
7.2	450
7.3	400
D.1	450
D.2	400
D.3	350
D.4	400
D.5	350

Table D.6: Number of independent simulation repeats for each of the presented results.

Additive Hidden Inputs and Model Error

The assumption of a hidden input signal incorporates many different types of model errors including parameter errors, hidden or misspecified interactions as well as inputs from exosystems. This can be seen as follows. Assume that the true system generating the data has the form

$$\begin{pmatrix} \dot{\boldsymbol{x}}(\boldsymbol{t}) \\ \dot{\boldsymbol{z}}(\boldsymbol{t}) \end{pmatrix} = \begin{pmatrix} \boldsymbol{\phi}(\boldsymbol{x}(t), \boldsymbol{z}(t), \boldsymbol{u}(t)) \\ \boldsymbol{\psi}(\boldsymbol{x}(t), \boldsymbol{z}(t), \boldsymbol{u}(t)) \end{pmatrix} \tag{D.1a}$$

$$\boldsymbol{y}(t) = \boldsymbol{h}(\boldsymbol{x}(t)) \; . \tag{D.1b}$$

Here, $\boldsymbol{x}(t) = (x_1(t), \ldots, x_N(t))^T \in \mathbb{R}^N$ is the state vector also included in the nominal model

$$\dot{\boldsymbol{x}}(t) = \boldsymbol{f}(\boldsymbol{x}(t), \boldsymbol{u}(t)) \tag{D.2a}$$

$$\boldsymbol{y}(t) = \boldsymbol{h}(\boldsymbol{x}(t)) \; , \tag{D.2b}$$

which represents our current knowledge or assumptions about the true system. The output map $\boldsymbol{h}$ is assumed to be exactly known and depends only on $\boldsymbol{x}$, so (D.2b) and (D.1b) have the same form. The exostate $\boldsymbol{z}(t) = (z_1(t), \ldots, z_P(t))^T$ in the true system (D.1) represents dynamic variables ignored in the nominal model. Let us denote the solution of the true system (D.1a) as

$$\boldsymbol{v}(t) = \begin{pmatrix} \boldsymbol{x}(t) \\ z(t) \end{pmatrix} \qquad \text{with initial condition} \quad \boldsymbol{v}(t_0) = \begin{pmatrix} \boldsymbol{x}(t_0) \\ z(t_0) \end{pmatrix}. \tag{D.3}$$

Now the functions $\boldsymbol{\phi}$ and $\boldsymbol{f}$ are compared along the true state trajectory (D.3) and define

$$\boldsymbol{w}(t = \dot{\boldsymbol{x}}(t) - \boldsymbol{f}\left(\boldsymbol{x}(t), \boldsymbol{u}(t)\right) = \boldsymbol{\phi}\left(\boldsymbol{x}(t), z(t), \boldsymbol{u}(t)\right) - \boldsymbol{f}\left(\boldsymbol{x}(t), \boldsymbol{u}(t)\right). \tag{D.4}$$

Thus,

$$\dot{\boldsymbol{x}}(t) = \boldsymbol{f}\left(\boldsymbol{x}(t), \boldsymbol{u}(t)\right) + \boldsymbol{w}(t). \tag{D.5}$$

represents the true dynamics $\boldsymbol{x}(t)$ through the nominal system $\boldsymbol{f}$ and the hidden input $\boldsymbol{w}(t)$. Note that the dependence on parameters is suppressed in the notation. However, this can be seen as part of the systems specification, i.e., as defining the properties of $\boldsymbol{f}$, $\boldsymbol{\phi}$ and $\boldsymbol{\psi}$, respectively. Thus, model errors can be represented by hidden (unknown) inputs.

Full Derivation of Equation (7.6)

According to (Zacher et al., 2012), Equation (7.6) can be derived as follows:

Given $y_{i,l}\left|\boldsymbol{x}_l, \xi^2_{i,l}\right. \sim N\left(y_{i,l}\left|\boldsymbol{x}_l, 1/\tau_{i,l}\right.\right)$ with unknown $\tau = \frac{1}{\xi^2_{i,l}}$ and $\tau \sim G\left(\alpha,\beta\right)$, we obtain

$$\begin{aligned}
p\left(y_{i,l}\left|\boldsymbol{x}_l, \alpha, \beta\right.\right) &\propto \int p\left(y_{i,l}\left|\boldsymbol{x}_l, \xi^2_{i,l}\right.\right) \times p\left(\xi^2_{i,l}\right|\alpha,\beta\Big)d\xi^2_{i,l} \\
&= \int N\left(y_{i,l}\left|\boldsymbol{x}_l, \tau_{i,l}\right.\right) \times G\left(\tau_{i,l}\right|\alpha,\beta\Big)d\tau_{i,l} \\
&= \int \frac{\beta^\alpha}{\Gamma\left(\alpha\right)}\tau^{\alpha-1}\exp\left(-\tau\beta\right)\left(\frac{\tau}{2\pi}\right)^{\frac{1}{2}} exp\left(-\frac{\tau}{2}\left(y_{i,l} - h_i\left(\boldsymbol{x}_l\right)\right)^2\right)d\tau_{i,l} \\
&= \frac{\beta^\alpha}{\Gamma\left(\alpha\right)}\frac{1}{\sqrt{2\pi}}\int \tau^{\alpha+\frac{1}{2}-1}\exp\left(-\frac{\tau\left(2\beta + \left(y_{i,l} - h_i\left(\boldsymbol{x}_l\right)\right)^2\right)}{2}\right)d\tau_{i,l} \\
&= \frac{\beta^\alpha}{\Gamma\left(\alpha\right)}\frac{1}{\sqrt{2\pi}}\frac{\Gamma\left(\alpha + \frac{1}{2}\right)}{\left(\beta + \frac{1}{2}\left(y_{i,l} - h_i\left(\boldsymbol{x}_l\right)\right)^2\right)^{\alpha+\frac{1}{2}}} \\
&= \frac{\Gamma\left(\alpha + \frac{1}{2}\right)}{\Gamma\left(\alpha\right)}\frac{1}{\left(2\pi\beta\right)^{\frac{1}{2}}}\frac{1}{\left(1 + \frac{1}{2\beta}\left(y_{i,l} - h_i\left(\boldsymbol{x}_l\right)\right)^2\right)^{\alpha+\frac{1}{2}}}.
\end{aligned}$$

Methodology

The full hierarchical model of the Bayesian elastic-net is given by (Zou and Hastie, 2005; Kyung et al., 2010)

$$\begin{aligned}
w_{n,l} \left| \sigma^2, \boldsymbol{\tau}^2, \lambda_2, w_{n,l-1} \right. &\sim N\left(w_{n,l-1}, \frac{\sigma^2 \tau_n^2}{\lambda_2 \tau_n^2 + 1}\right) \\
\boldsymbol{\tau}^2 \left| \boldsymbol{\lambda}_1^2 \right. &\sim \prod_{n=1}^{N} \left[\frac{\lambda_{1,n}^2}{2} \exp\left(-\frac{\lambda_{1,n}^2 \tau_n^2}{2}\right)\right], \qquad \tau_1^2, \ldots, \tau_N^2 > 0 \\
\boldsymbol{\lambda}_1^2 &\sim \prod_{n=1}^{N} \left[\frac{\delta_{1,n}^{r_{1,n}}}{\Gamma\left(r_{1,n}\right)} \left(\lambda_{1,n}^2\right)^{r_{1,n}-1} \exp\left(-\delta_{1,n} \lambda_{1,n}^2\right)\right] \\
\lambda_2 &\sim \frac{\delta_2^{r_2}}{\Gamma\left(r_2\right)} \lambda_2^{r_2-1} \exp\left(-\delta_2 \lambda_2\right) \\
\sigma^2 &\sim p\left(\sigma^2\right), \qquad \sigma^2 > 0.
\end{aligned}$$

For σ^2 we chose a standard non-informative, improper and scale-invariant prior $p\left(\sigma^2\right) \propto \sigma^{-2}$ (Park and Casella, 2008; Kyung et al., 2010). In contrast to the variance of the measurement noise $\boldsymbol{\xi}_l$, σ^2 represents the variance of the hidden influences. The parameters $\boldsymbol{\lambda}_1$ and λ_2 control the sparsity and smoothness of the resulting hidden influence dynamics, respectively. Please note that the parameter $\boldsymbol{\lambda}_1$ in contrast to λ_2 is controlled by an additional hyperparameter $\boldsymbol{\tau}^2$. For l = 1 the full Bayesian elastic net prior corresponds to a product of a Gaussian and Laplace distribution (Zou and Hastie, 2005; Kyung et al., 2010).

Sampling Algorithm

Algorithm 1 : Pseudo Code

```
for $l \leftarrow 1$ to $T$ do
    for $s \leftarrow 1$ to $S$ do
        1. $i \leftarrow$ random species
           $w^*_{n,t_l} \leftarrow w_{n,l-1} + N(0, J)$
           Accept $w^*_{n,l}$ with probability $\boldsymbol{r}\left(\boldsymbol{w}^{(s)}{}_l \middle| \boldsymbol{w}^*_l\right)$, otherwise set $w^{(s+1)}_{n,l} \leftarrow w^{(s)}_{n,l}$
        2. Draw $\sigma^2 \sim inverseGamma\left(\frac{n}{2}, \frac{\zeta + \sum_{n=1}^{N}\left(w^{(s+1)}_{n,l} - w_{n,l-1}\right)^2 \frac{\lambda_2 \tau_n^2 + 1}{\tau_n^2}}{2}\right)$
        3. Draw $\tau_n^{-2} \sim inverseGaussian\left(\sqrt{\frac{\lambda^2_{1,n}\sigma^2}{w^{(s+1)}_{n,l} - w_{n,l-1}}}, \lambda^2\right)$ for $n = 1, \ldots, N$
        4. Draw $\lambda^2_{1,n} \sim Gamma\left(r_{1,n} + 1, \frac{\tau_n^2}{2} + \delta_{1,n}\right)$ for $n = 1, \ldots, N$
        5. Draw $\lambda_2 \sim Gamma\left(r_2 + \frac{n}{2}, \frac{1}{2\sigma^2}\left\|\boldsymbol{w}^{(s+1)}_l - \boldsymbol{w}_{l-1}\right\|_2^2 + \delta_2\right)$
    $\boldsymbol{w}_l = \frac{1}{S}\sum_{s=1}^{S} \boldsymbol{w}^s_l;$
    $x_l = \int_{t_{l-1}}^{t_l} \boldsymbol{f}\left(\boldsymbol{x}\left(t'_l\right)\boldsymbol{u}\left(t'_l\right)\right) + \boldsymbol{w}\left(t'_l\right) dt' \Big|_{x_{l-1}}, \quad \boldsymbol{x}_0 = \eta$
```

$$\boldsymbol{r}\left(\boldsymbol{w}^{(s)}{}_l \middle| \boldsymbol{w}^*_l\right) = \min\left\{1, \frac{p\left(\boldsymbol{y}_l \middle| \boldsymbol{x}_l, \alpha, \beta\right)}{p\left(\boldsymbol{y}_l \middle| \boldsymbol{x}_l, \alpha, \beta\right)} \times \frac{p\left(\boldsymbol{x}_l \middle| \boldsymbol{x}_{l-1}, \boldsymbol{w}^*_l, \boldsymbol{w}_{l-1}\right)}{p\left(\boldsymbol{x}_l \middle| \boldsymbol{x}_{l-1}, \boldsymbol{w}^s_l, \boldsymbol{w}_{l-1}\right)} \times \frac{p\left(\boldsymbol{w}^*_l \middle| \boldsymbol{\theta}\right)}{p\left(\boldsymbol{w}^s_l \middle| \boldsymbol{\theta}\right)}\right\}$$

$$= \min\left\{1, \prod_{i=1}^{I}\left(\frac{2\beta + (y_{i,l} - h_i(\boldsymbol{x}_l))^2}{2\beta + (y_{i,l} - h_i(\boldsymbol{x}_l))^2}\right)^{\alpha + \frac{1}{2}} \times \frac{p\left(\boldsymbol{w}^*_l \middle| \sigma^2, \boldsymbol{\tau}^2, \lambda_2, \boldsymbol{w}_{l-1}\right)}{p\left(\boldsymbol{w}^s_l \middle| \sigma^2, \boldsymbol{\tau}^2, \lambda_2, \boldsymbol{w}_{l-1}\right)}\right\}$$

Step 1 is given by a Metropolis-Hastings move (Algorithm 2) with respect to J (Brooks, 1998). Avoiding autocorrelation, the mean for $\boldsymbol{w}_l$ is adapted via thinning (Gelman et al., 2013).

Algorithm 2 : Independent MH

```
for $s' \leftarrow 1$ to $S'$ do
    Draw $\tilde{n} \sim U\{1, N\}$
    Draw $w^{*}_{i=\tilde{n},l} \sim N(w_{n=\tilde{n},l-1}, J)$
    $w^{*}_{n\neq\tilde{n},l} = w^{(s')}_{n\neq\tilde{n},l}$
    if $\boldsymbol{r}\left(\boldsymbol{w}^{(s)}{}_{l} \middle| \boldsymbol{w}^{*}_{l}\right) \geq \alpha$ then $w^{s'+1}_{n,l} = w^{*}_{n,l}$;
    else $w^{s'+1}_{n,l} = w^{(s')}_{n,l}$;
$\boldsymbol{w}_l = \frac{1}{S'} \sum_{s'=1}^{S'} w^{s'}_{l}$
```

Hyperparameter Settings

Following an empirical Bayesian approach, α and β in the variance prior are estimated by fitting the inverse sample variance distribution via maximum likelihood. An initial value for σ^2 can be approximated by the maximal distance between two subsequent $\boldsymbol{w}(t_l)$ in the time series

$$\max_{n,l>1}\left(|w_n(t_l)| - |w_n(t_{l-1})|\right) \propto \max_{i,l>1}\left(\left|\frac{d\left(y_i(t_l) - h_i(\boldsymbol{x}(t_l))\right)}{dt_l}\right| - \left|\frac{d\left(y_i(t_{l-1}) - h_i(\boldsymbol{x}(t_{l-1}))\right)}{dt_{l-1}}\right|\right). \tag{D.6}$$

We advise to include a parameter ζ to omit values next to zero for σ^2, which would lead to numerical issues. Here a good conservative choice is also given by Equation (D.6).

The parameter J of the candidate distribution $\pi \sim N(0, J)$ can be defined as the maximum of the mean largest value of the approximated hidden influence with respect to the least square error. To increase performance we suggest a burn-in phase of about 33% of the total number of iterations, which was set to $1500 \times N$ as a compromise between computational time and sampling quality. The number of MH moves was set to be ten times larger than the number of Gibbs steps.

Due to the low number of measurement points and their variance we introduce the additional parameter $\gamma \leq 1$ which means that the measurement noise is reduced by γ and $\xi_{i,l}$ is re-parametrized as

$$\xi^2{}_{i,l} \sim IG\left(\alpha, \frac{\beta}{\gamma}\right). \tag{D.7}$$

Since the re-parametrized variance estimator yields a sharper prior distribution, it is more conservative.

The parameters λ_1 and λ_2 themselves depend on hyperparameters, which can be set in a non-informative manner or with respect to prior knowledge about the degree of shrinkage and smoothness of the hidden influences (Kyung et al., 2010).

Model Details

Motifs

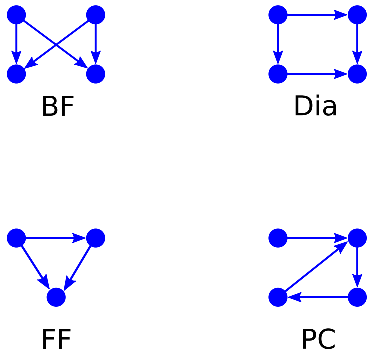


Figure D.1: Graphical representation of the investigated bi-fan motif (BF), diamond motif (Dia), feed-forward loop (FF) and the protein cascade (PC) (Milo, 2002).

For all motifs (see (Milo, 2002)) we used the monotonic function

$$u(t) = \left(1 - \frac{1}{1+t}\right)$$

as known input stimulus and the initial condition $\boldsymbol{x}(0) = \mathbf{0}$ for all state variables with $t = [0, 1]$. The system is assumed to be fully observable.

Bi-fan

$$\begin{aligned}\dot{x}_1(t) &= u(t) - x_1(t)\\ \dot{x}_2(t) &= u(t) - x_2(t)\\ \dot{x}_3(t) &= x_2(t) + x_1(t) - x_3(t)\\ \dot{x}_4(t) &= x_2(t) + x_1(t) - x_4(t)\end{aligned}$$

Diamond

$$\begin{aligned}\dot{x}_1(t) &= u(t) - x_1(t)\\ \dot{x}_2(t) &= x_1(t) - x_2(t)\\ \dot{x}_3(t) &= x_1(t) - x_3(t)\\ \dot{x}_4(t) &= x_2(t) + x_3(t) - x_4(t)\end{aligned}$$

Feed-forward Loop

$$\begin{aligned}
\dot{x}_1(t) &= u(t) - x_1(t) \\
\dot{x}_2(t) &= x_1(t) - x_2(t) \\
\dot{x}_3(t) &= x_2(t) + x_1(t) - x_3(t)
\end{aligned}$$

Protein Cascade

$$\begin{aligned}
\dot{x}_1(t) &= u(t) - x_1(t) \\
\dot{x}_2(t) &= \frac{x_1(t)}{1 + x_4} - x_2(t) \\
\dot{x}_3(t) &= x_2(t) - x_3(t) \\
\dot{x}_4(t) &= x_3(t) - x_4(t)
\end{aligned}$$

G protein Signaling Model

The model for the heterotrimeric G protein cycle in yeast was downloaded from the BioModels Database (Li et al., 2010; Yi et al., 2003) (BIOMD0000000072). The system is assumed to be fully observable.

$$\begin{aligned}
\frac{\mathrm{d}([R])}{\mathrm{d}t} &= -\left(3.32e^{-18}\right)[L]_{const.}\,[\mathrm{R}] + 0.01\,[RL] - 4 - 0.0004\,[R] \\
\frac{\mathrm{d}([\mathrm{GP_i}])}{\mathrm{d}t} &= [\mathrm{GP}_{\alpha\mathrm{i}}]\left[GP_\beta\right] - \left(1e^{-05}\right)[RL]\,[\mathrm{GP_i}] \\
\frac{\mathrm{d}\left(\left[GP_\beta\right]\right)}{\mathrm{d}t} &= -\,[\mathrm{GP}_{\alpha\mathrm{i}}]\left[GP_\beta\right] + \left(1e^{-05}\right)[RL]\,[\mathrm{GP_i}] \\
\frac{\mathrm{d}\left(\left[\mathrm{GP}_{\alpha_\mathrm{i}}\right]\right)}{\mathrm{d}t} &= -\,[\mathrm{GP}_{\alpha\mathrm{i}}]\left[GP_\beta\right] + 0.11\,[\mathrm{GP}_\alpha] \\
\frac{\mathrm{d}([\mathrm{GP}_\alpha])}{\mathrm{d}t} &= \left(1e^{-05}\right)[RL]\,[\mathrm{GP_i}] - 0.11\,[\mathrm{GP}_\alpha] \\
\frac{\mathrm{d}([RL])}{\mathrm{d}t} &= \left(3.32e^{-18}\right)[L]_{const.}\,[R] - 0.01\,[RL] - 0.004\,[RL]
\end{aligned}$$

The initial conditions are given by

$$\begin{aligned}
[R]_0 &= 100 \text{ item\textbackslash l} & \left[\overline{\mathrm{GP}}_\alpha\right]_0 &= 30 \text{ item\textbackslash l} \\
\left[\overline{\mathrm{GP}}\right]_0 &= 70 \text{ item\textbackslash l} & [\mathrm{GP}_\alpha]_0 &= 0 \text{ item\textbackslash l} \\
\left[GP_{\beta\backslash\gamma}\right]_0 &= 30 \text{ item\textbackslash l} & [LR]_0 &= 0 \text{ item\textbackslash l} \\
[L]_{const.} &= 0.02 \text{ item\textbackslash l} & &
\end{aligned}$$

and $t = [0, 25]$

UV-B Network Model

The model equations for the UV-B signaling network were obtained from the BioModels Database (Li et al., 2010; Ouyang et al., 2014), see BIOMD0000000545.

$$
\begin{aligned}
\frac{d\,[\mathrm{CS}]}{dt} =& -2 \cdot ka_1 \cdot [\mathrm{CS}]^2 \cdot [\mathrm{UVR8M}]^2 + 2kd_1 \cdot [\mathrm{UCS}] \\
& + ks_1 \cdot (1 + UV \cdot n_3 \cdot ([\mathrm{HY5}] + FHY3)) \\
& - kdr_1 \cdot (1 + (n_1 \cdot UV)) \cdot [\mathrm{CS}] - kd_2 \cdot [\mathrm{CDCS}] \\
& - 2 \cdot ka_2 \cdot [\mathrm{CS}]^2 \cdot [\mathrm{CD}] \\
\frac{d\,[\mathrm{CD}]}{dt} =& - ka_2 \cdot [\mathrm{CS}]^2 \cdot [\mathrm{CD}] + kd_2 \cdot [\mathrm{CDCS}] \\
& + ka_4 \cdot [\mathrm{CD}] \cdot [\mathrm{DWD}] + kd_4 \cdot [\mathrm{CDW}] \\
\frac{d\,[\mathrm{CDCS}]}{dt} =& - kd_2 \cdot [\mathrm{CDCS}] + ka_2 \cdot [\mathrm{CS}]^2 \cdot [\mathrm{CD}] \\
\frac{d\,[\mathrm{UVR8M}]}{dt} =& -2 \cdot k_1 \cdot [\mathrm{UVR8M}]^2 + 2 \cdot k_2 \cdot [\mathrm{UVR8D}] \\
& - 2 \cdot ka_1 \cdot [\mathrm{CS}]^2 \cdot [\mathrm{UVR8M}]^2 + 2 \cdot kd_1 \cdot [\mathrm{UCS}] \\
& - ka_3 \cdot [\mathrm{UVR8M}] \cdot [\mathrm{RUP}] \\
\frac{d\,[\mathrm{UCS}]}{dt} =& - kd_1 \cdot [\mathrm{UCS}] + ka_1 \cdot [\mathrm{CS}]^2 \cdot [\mathrm{UVR8M}]^2 \\
\frac{d\,[\mathrm{UVR8D}]}{dt} =& - k_2 \cdot [\mathrm{UVR8D}] + k_1 \cdot [\mathrm{UVR8M}]^2 + kd_3 \cdot [\mathrm{UR}]^2 \\
\frac{d\,[\mathrm{RUP}]}{dt} =& - ka_3 \cdot [\mathrm{UVR8M}] \cdot [\mathrm{RUP}] + ks_2 \cdot (1 + UV \cdot [\mathrm{UCS}]) \\
& - kdr2 \cdot [\mathrm{RUP}] + (2) \cdot kd_3 \cdot [\mathrm{UR}]^2 \\
\frac{d\,[\mathrm{UR}]}{dt} =& -2 \cdot kd_3 \cdot [\mathrm{UR}]^2 + ka_3 \cdot [\mathrm{UVR8M}] \cdot [\mathrm{RUP}] \\
\frac{d\,[\mathrm{HY5}]}{dt} =& - kdr3 \cdot \left(\frac{[\mathrm{CDCS}]}{kdr_{3a} + [\mathrm{CDCS}]} + \frac{[\mathrm{CDW}]}{kdr_{3b} + [\mathrm{CDW}]} \right) \cdot [\mathrm{HY5}] \\
& + ks3_p \cdot (1 + n_2 \cdot UV) - kdr3 \cdot \left(\frac{[\mathrm{UCS}]}{ksr + [\mathrm{UCS}]} \right) \cdot [\mathrm{HY5}] \\
\frac{d\,[\mathrm{DWD}]}{dt} =& - ka_4 \cdot [\mathrm{CD}] \cdot [\mathrm{DWD}] + kd_4 \cdot [\mathrm{CDW}] \\
\frac{d\,[\mathrm{CDW}]}{dt} =& - kd_4 \cdot [\mathrm{CDW}] + ka_4 \cdot [\mathrm{CD}] \cdot [\mathrm{DWD}]
\end{aligned}
$$

$$
\begin{aligned}
UM_{Total} &= 2 \cdot [\mathrm{UCS}] + [\mathrm{UVR8M}] + [\mathrm{UR}] \\
COP1_{Total} &= 2 \cdot [\mathrm{UCS}] + 2 \cdot [\mathrm{CDCS}] + [\mathrm{CS}] \\
UVR8D_{obs.} &= [\mathrm{UVR8D}] \\
HY5_{obs.} &= [\mathrm{HY5}] \\
UVR8M_{obs} &= [\mathrm{UVR8M}]
\end{aligned}
$$

$$
\begin{aligned}
[\mathrm{CS}] &= x_1; & [\mathrm{UVR8M}] &= x_4; & [\mathrm{RUP}] &= x_7; & [\mathrm{DWD}] &= x_{10};\\
[\mathrm{CD}] &= x_2; & [\mathrm{UCS}] &= x_5; & [\mathrm{UR}] &= x_8; & [\mathrm{CDW}] &= x_{11};\\
[\mathrm{CDCS}] &= x_3; & [\mathrm{UVR8D}] &= x_6; & [\mathrm{HY5}] &= x_9
\end{aligned}
$$

The initial conditions are given by

$$
\begin{aligned}
[\mathrm{CS}]_0 &= 0.2 \text{ mol} & [\mathrm{RUP}]_0 &= 0 \text{ mol}\\
[\mathrm{CD}]_0 &= 10 \text{ mol} & [\mathrm{UR}]_0 &= 0 \text{ mol}\\
[\mathrm{CDCS}]_0 &= 2 \text{ mol} & [\mathrm{HY5}]_0 &= 0.25 \text{ mol}\\
[\mathrm{UVR8M}]_0 &= 0 \text{ mol} & [\mathrm{DWD}]_0 &= 20 \text{ mol}\\
[\mathrm{UCS}]_0 &= 0 \text{ mol} & [\mathrm{CDW}]_0 &= 0 \text{ mol}\\
[\mathrm{UVR8D}]_0 &= 20 \text{ mol}
\end{aligned}
$$

and $t = [0, 6]$.

JAK-STAT Model

The measurements for the JAK-STAT system were obtained from http://webber.physik.uni-freiburg.de/ ~jeti/PNAS_Swameye_Data. Mass conservation was accounted for by the constraint $2x_4(t) + 2x_3(t) + x_1(t) + x_2(t) = \text{const.}$ for all t (Swameye et al., 2003; Raue et al., 2009). The nominal model is given by (Swameye et al., 2003):

$$
\begin{aligned}
\dot{x}_1 &= -\psi_1 x_1 u\\
\dot{x}_2 &= \psi_1 x_1 u - 2\psi_2 x_2^2\\
\dot{x}_3 &= \psi_2 x_2^2 - \psi_3 x_3\\
\dot{x}_4 &= \psi_3 x_3\\
y_2 &= \psi_5 (x_1 + x_2 + 2x_3)\\
y_1 &= \psi_4 (x_2 + 2x_3)\,.
\end{aligned}
$$

$$
\begin{aligned}
u &= \text{Erythropoietin Receptor}\\
x_1 &= \text{STAT5}\\
x_2 &= \text{STAT5}_\text{p}\\
x_3 &= \text{STAT5}_\text{di}\\
x_4 &= \text{STAT5}_\text{n}\\
y_1 &= \text{total STAT5}\\
y_2 &= \text{total STAT5}_\text{p}
\end{aligned}
$$

and $t = [0, 60]$.

Model of Information Processing at EpoR

The Model of the information processing at EpoR was obtained from (Becker et al., 2010). Data was obtained from dMod.

$$
\begin{aligned}
\dot{x}_1 =& k_t \cdot B_{max} - k_t \cdot x_1 - k_o n \cdot x_1 \cdot x_2 + k_{off} \cdot x_3 + k_{ex} \cdot x_4 \\
\dot{x}_2 =& - k_{on} \cdot x_1 \cdot x_2 + k_{off} \cdot x_3 + k_{ex} \cdot x_4 \\
\dot{x}_3 =& k_{on} \cdot x_1 \cdot x_2 - k_{off} \cdot x_3 - k_e \cdot x_3 \\
\dot{x}_4 =& k_e \cdot x_3 - k_{ex} \cdot x_4 - k_{di} \cdot x_4 - k_{de} \cdot x_4 \\
\dot{x}_5 =& k_{di} \cdot x_4 \\
\dot{x}_6 =& k_{de} \cdot x_4 \\
y_2 =& \psi_1 \left(x_2 + 2x_6\right) \\
y_1 =& \psi_2 \left(x_3\right) \\
y_3 =& \psi_3 \left(x_4 + x_5\right)
\end{aligned}
$$

$$
\begin{aligned}
x_1 &= \text{EpoR} \\
x_2 &= \text{Epo} \\
x_3 &= \text{Epo-EpoR} \\
x_4 &= \text{Epo-EpoR}_\text{i} \\
x_5 &= \text{dEpo}_\text{i} \\
x_6 &= \text{dEpo}_\text{e} \\
y_1 &= \text{Epo concentration in the medium} \\
y_2 &= \text{Epo concentration on the surface} \\
y_3 &= \text{Epo concentration in the cells,}
\end{aligned}
$$

and $t = [0, 300]$.

Model of α-Pinene Isomerization

The model of the thermal isomerization of αP in the liquid phase was obtained from (Fuguitt and Hawkins, 1947).

$$\begin{aligned}
\dot{x}_1 &= -(p_1 + p_2) \cdot x_1 \\
\dot{x}_2 &= p_1 \cdot x_1 \\
\dot{x}_3 &= p_2 \cdot x_1 - (p_3 + p_4) \cdot x_3 + p_5 \cdot x_4 \\
\dot{x}_4 &= p_4 \cdot x_3 - p_5 \cdot x_4
\end{aligned}$$

x_1 = α-Pinene
x_2 = Dipentene
x_3 = Dimer
x_4 = allo-ocimene,

and $t = [0, 25]$.

Gene Regulatory Network

The gene regulatory network was obtained from DREAM6 Challenge (Meyer et al., 2014). For further information about this challenge we refer the reader to the DREAM6 Challenge.

$$
\begin{aligned}
\dot{x}_1 =& pro1_{strength} - mrna1_{degradation-rate} \cdot x_1 \\
\dot{x}_2 =& rbs1_{strength} \cdot x_1 - p_{degradation-rate} \cdot x_2 \\
\dot{x}_3 =& pro2_{strength} \cdot \frac{\left(\frac{x_1}{v2_{Kd}}\right)^{v2_h}}{1 + \left(\frac{x_2}{v2_{Kd}}\right)^{v2_h}} \cdot \frac{1}{1 + \left(\frac{x_6}{v5_{Kd}}\right)^{v5_h}} - mrna2_{degradation-rate} \cdot x_3 \\
\dot{x}_4 =& rbs2_{strength} \cdot x_3 - p_{degradation-rate} \cdot x_4 \\
\dot{x}_5 =& pro3_{strength} \cdot \frac{\left(\frac{x_1}{v3_{Kd}}\right)^{v3_h}}{1 + \left(\frac{x_2}{v3_{Kd}}\right)^{v3_h}} \cdot \frac{1}{1 + \left(\frac{x_2}{v4_{Kd}}\right)^{v4_h}} - mrna3_{degradation-rate} \cdot x5 \\
\dot{x}_6 =& rbs3_{strength} \cdot x_5 - p_{degradation-rate} \cdot x_6 \\
\dot{x}_7 =& pro4_{strength} \cdot \frac{\left(\frac{x_1}{v1_{Kd}}\right)^{v1_h}}{1 + \left(\frac{x_2}{v1_{Kd}}\right)^{v1_h}} \cdot \frac{1}{1 + \left(\frac{x_5}{v8_{Kd}}\right)^{v8_h}} - mrna4_{degradation-rate} \cdot x_7 \\
\dot{x}_8 =& rbs4_{strength} \cdot x_7 - p_{degradation-rate} \cdot x_8 \\
\dot{x}_9 =& pro5_{strength} \cdot \frac{1}{1 + \left(\frac{x_4}{v6_{Kd}}\right)^{v6_h}} - mrna5_{degradation-rate} \cdot x_9 \\
\dot{x}_{10} =& rbs5_{strength} \cdot x_9 - p_{degradation-rate} \cdot x_1 0 \\
\dot{x}_{11} =& pro6_{strength} \cdot \frac{1}{1 + \left(\frac{x_4}{v7_{Kd}}\right)^{v7_h}} - mrna6_{degradation-rate} * x_1 1 \\
\dot{x}_{12} =& rbs6_{strength} \cdot x_1 1 - p_{degradation-rate} \cdot x_1 2
\end{aligned}
$$

$$
\begin{aligned}
y_1 &= \mathrm{pp1}_{\mathrm{mrna}} \\
y_2 &= \mathrm{p1} \\
y_3 &= \mathrm{p2} \\
y_4 &= \mathrm{p3} \\
y_5 &= \mathrm{p4} \\
y_6 &= \mathrm{p5} \\
y_7 &= \mathrm{p6}
\end{aligned}
$$

$$
\begin{aligned}
x_1 &= \mathrm{pp1}_{\mathrm{mrna}} \\
x_2 &= \mathrm{p1} \\
x_3 &= \mathrm{pp2}_{\mathrm{mrna}} \\
x_4 &= \mathrm{p2} \\
x_5 &= \mathrm{pp3}_{\mathrm{mrna}} \\
x_6 &= \mathrm{p3} \\
x_7 &= \mathrm{pp4}_{\mathrm{mrna}} \\
x_8 &= \mathrm{p4} \\
x_9 &= \mathrm{pp5}_{\mathrm{mrna}} \\
x_{10} &= \mathrm{p5} \\
x_{11} &= \mathrm{pp6}_{\mathrm{mrna}} \\
x_{12} &= \mathrm{p6}
\end{aligned}
$$

The initial conditions for the proteins $p1$ to $p6$ were set to 1 and, for all mRNA species, $\mathrm{pp1}_{\mathrm{mrna}}$ to $\mathrm{pp6}_{\mathrm{mrna}}$ were set to 0. The observation period is given by $t = [0, 20]$.

APPENDIX E

Refined Model of the M_2 Receptor-dependent Signaling

Material and Methods

Cell Culture

Flp-In™-CHO cells stably expressing the hM2 receptor (CHO-hM2 cells) were cultured as described previously in Ham's nutrient mixture F-12 (Ham's F-12) supplemented with 10% (v/v) fetal calf serum (FCS), 100 U mL^{-1} penicillin, 100 mg mL^{-1} streptomycin and 2 mM l-glutamine (Sigma Aldrich, UK) (Schrage et al., 2013). The cells were grown in a humidified incubator at 37 °C and 5% CO_2, and passaged by trypsinization at nearly confluence.

cAMP Assay

Quantification of the agonist-induced rise of the intracellular cAMP was performed using cAMP Complete ELISA kit (ADI-900-163, Enzo Life Sciences, Inc) according to the manufacturer's introductions. Cells were incubated for 30 minutes with iperoxo.

Western Blots

Cells were seeded into 6 well plates and stimulated with 0.1 μM iperoxo for 2, 5, 10, 15, 20, 30 minutes. After specified treatments, cellular lysates were prepared in lysis buffer [25 mM Hepes, 2.5 mM EDTA, 50 mM NaCl, 50 mM NaF, 30 mM sodium pyrophosphate, 10% (v/v) glycerol, 1% (v/v) Triton X-100, pH 7.5, containing Complete™ EDTA-free protease inhibitor cocktail tablets (Roche)].

Proteins were separated by SDS/PAGE (4–12% Bis-Tris gels) and transferred onto nitrocellulose membranes for western blotting using the NUPAGE system (Invitrogen). The membranes were blocked in 5% phosphoblocker. The blots were probed with the related antibodies (see

Antibody	Manufacturer	Code
PKA C-α	Cell Signaling Technology	4782
Epac (A-5)	Santa Cruz Biotechnology	sc-28366
Anti-AKT1 (phospho S473)	abcam	ab81283
Anti-GRK6	abcam	ab109302
Anti-M2	abcam	ab109226
Anti-GRK2 (phospho S29)	abcam	ab58520
Anti-RGS14	abcam	ab137893
Phospho-PKA Substrate	Cell Signaling Technology	9624
Akt1 (2H10)	Cell Signaling Technology	2967
phospho-UCR1	non-purchasable	-
pan-PDE4	Millipore	ABS181

Table E.1: List of used antibodies. The used Anti-GRK6 antibody detects GRK2 as well.

Table E.1 diluted in the range from 1/500 to 1/5000 in 1% phosphoblocker diluted in TBS-T (150 mM NaCl, 20 mM Tris–HCl, pH 7.6, 0.1% Tween-20). The bands were visualized using the appropriate secondary antibody and visualized using an Odyssey scanner (Licor Biosciences) and the band intensity was quantified using Image Studio (Licor Biosciences, v. 5.2).

Revised Model

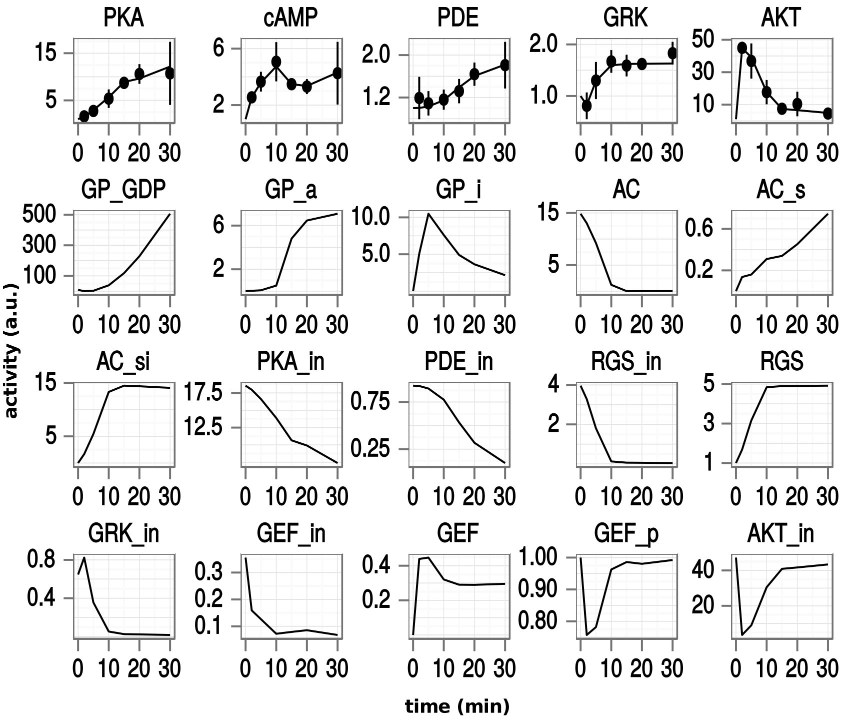


Figure E.1: Complete model fitting results for the revised model.

List of Reactions

The first number (red) counts the reactions, the second number is linked to Table B.1.

$$\overline{GP} \rightarrow GP \tag{1|1; 2; 3; 4; 11}$$

$$GP \rightarrow \alpha s\backslash o \tag{2|1; 2; 3; 4; 11}$$

$$GP \rightarrow \alpha i \tag{3|1; 2; 3; 4; 11}$$

$$\alpha s\backslash o \rightarrow \emptyset \tag{4|2; 11}$$

$$\alpha_i \rightarrow \emptyset \tag{5|7; 11}$$

$$AC + \alpha_s \rightleftharpoons AC_{\alpha_s} \tag{6|9; 11}$$

$$AC_{\alpha_s} + \alpha_i \rightleftharpoons AC_{\alpha_s \alpha_i} \tag{7|9; 10; 11}$$

$$ATP \rightarrow cAMP \tag{8|15; 18; 25}$$

$$cAMP \rightarrow AMP \tag{9|16}$$

$$2\overline{PKA} + 4cAMP \rightleftharpoons 2PKA \tag{10|24}$$

$$\overline{PDE} \rightleftharpoons PDE \tag{11|16}$$

$$\overline{RGS} \rightleftharpoons RGS \tag{12|6; 14}$$

$$\overline{GRK} \rightleftharpoons GRK \tag{13|13}$$

$$\overline{GEF} + cAMP \rightleftharpoons GEF \tag{14|17}$$

$$GEF \rightleftharpoons GEF_P \tag{15|12}$$

$$\overline{AKT} \rightleftharpoons AKT \tag{16|12}$$

$$X : Y = \text{Complex of X and Y}$$

$$\overline{X} = \text{Inactive form of protein X}$$

$$X_P = \text{Phosporylated form of protein X}$$

List of Rate Laws

Detailed list of rate laws including the parameter estimates. Numbers count the reactions linked to the List of Reactions.

	Rate-Law						
Nr.	Forward	Backward	Parameter Estimates				
(1)	$k_{eq}\,(1-\exp(-k_{obs}\times t))$		$k_{eq} = 50.5$	$k_{obs} = 0.076$			
(2)	$(1-Frac)\times\frac{V_1\times[GP]}{Km_1+[GP]}$		$Frac = 0.881$	$V_1 = 14.994$	$Km_1 = 0.247$		
(3)	$Frac\times\frac{V_1[GP]}{Km_1+[GP]}$		$Frac = 0.881$	$V_1 = 14.994$	$Km_1 = 0.247$		
(4)	$V_2\times[\alpha s\backslash o]$		$V_2 = 0.245$				
(5)	$\frac{V_3\times[RGS]\times[\alpha i]}{Km_3+[\alpha i]}$		$V_3 = 2.790$	$Km_3 = 0.066$			
(6)	$k_1\times[AC]\times[\alpha s\backslash o]$	$k_2\times[AC_s]$	$k_1 = 2.240$	$k_2 = 0.318$			
(7)	$k_3\times[AC_s]\times[\alpha i]$	$k_4\times[AC_{si}]$	$k_3 = 1.521$	$k_4 = 0.176$			
(8)	$\frac{V_4\times[AC_s]}{(Km_4+[AC_s])\left(1+\frac{[PKA]}{Ki_4}\right)}$		$V_4 = 47.223$	$Km_4 = 0.658$	$Ki_4 = 12.880$		
(9)	$\frac{V_5\times[cAMP]\times[PDE]}{Km_5+[cAMP]}$		$V_5 = 6.953$	$Km_5 = 0.167$			
(10)	$k_5\times\left[\overline{PKA}\right]\times[cAMP]^4$	$k_6\times[PKA]^2$	$k_5 = 8e-05$	$k_6 = 4e-04$			
(11)	$\frac{V_6\times[PKA]\times\left[\overline{PDE}\right]}{Km_6+\left[\overline{PDE}\right]}$	$V_7\times[PDE]$	$V_6 = 0.018$	$Km_6 = 0.516$	$V_7 = 0.010$		
(12)	$\frac{V_8\times[PKA]\times\left[\overline{RGS}\right]}{Km_8+\left[\overline{RGS}\right]}+\frac{V_{12}\times[GEF]\times\left[\overline{RGS}\right]}{Km_{12}+\left[\overline{RGS}\right]}$	$V_9\times[RGS]$	$V_8 = 0.382$	$Km_8 = 0.201$	$V_{12} = 0.312$	$Km_{12} = 0.354$	$V_9 = 0.166$
(13)	$\frac{V_{10}\times[PKA]\times\left[\overline{GRK}\right]}{Km_{10}+\left[\overline{GRK}\right]}$	$V_{11}\times[GRK]$	$V_{10} = 0.575$	$Km_{10} = 0.057$	$V_{11} = 0.961$		
(14)	$k_7\times\left[\overline{GEF}\right]\times[cAMP]$	$k_8\times[GEF]$	$k_7 = 0.369$	$k_8 = 0.363$			
(15)	$\frac{V_{13}\times[GEF]\times[GRK]}{Km_1 3+[GEF]}$	$k_9\times\left[GEF_p\right]$	$V_{13} = 0.423$	$Km_1 3 = 0.351$	$k_9 = 0.317$		
(16)	$\frac{V_{14}\times\left[\overline{AKT}\right]}{Km_{14}+\left[\overline{AKT}\right]}$	$\frac{V_{15}\times[PKA]\times[AKT]}{Km_{15}+[AKT])}$	$V_{14} = 55.452$	$Km_{14} = 5.054$	$V_{15} = 19.580$	$Km_{15} = 18.4603$	

APPENDIX F

List of Publications

Publications

- **Engelhardt B**, Fröhlich H, Kschischo M, Learning (from) the errors of a systems biology model, *Nature Scientific Reports*, 6, 20772, 2016.

- **Engelhardt B**, Kschischo M, Fröhlich H, A Bayesian Approach to Estimating Hidden Variables as well as Missing and Wrong Molecular Interactions in ODE Based Mathematical Models, *Journal of the Royal Society Interface*, under review.

- **Engelhardt B**, Holze J, Elliott C, Baillie GS, Kschischo M, Fröhlich H, Modeling and Mathematical Analysis of the M2-Receptor Dependent Joint Signaling and Secondary Messenger Network in CHO Cells, *Mathematical Medicine and Biology*, accepted.

- **Engelhardt B**, König J, Blettner M, Wild P, Münzel T, Lackner K, Blankenberg S, Pfeiffer N, Beutel M, Zwiener I, Combining Cross-sectional Data on Prevalence with Risk Estimates from a Prediction Model, *Methods of Information in Medicine*, 5, 371 - 379, 2014

Conference Presentations

- **Engelhardt B**, Fröhlich H, Kschischo M, Inferring model errors and unmeasured system states with incomplete models: The dynamic elastic net, presented at the Meeting of the International Study Group for Systems Biology 2016 in Jena (poster).

- **Engelhardt B**, Fröhlich H, Kschischo M, Learning (from) the errors of a systems biology model, presented at the International Symposium 2016 in Bonn (paper presentation and poster).

- **Engelhardt B**, Kschischo M, Fröhlich H, A New Computational Method to Make Dynamic Mathematical Models Complete and Consistent, presented at the European Conference on Mathematical and Theoretical Biology 2016 in Göttingen (poster).

- **Engelhardt B**, Kschischo M, Fröhlich H, A Bayesian Approach for estimating hidden variables in ODE based systems biology models, presented at the European Conference on Mathematical and Theoretical Biology 2016 in Notthingham (contributed talk and poster).
- **Engelhardt B**, Schrage R, Tränkle C, Mohr K, Kschischo M, Fröhlich H, G Protein-dependent Signaling in CHO Cells: Dynamic Systems and Beyond, presented at the Summer Academy Pharmacology 2015 in Bonn (contributed talk and poster).
- **Engelhardt B**, Kschischo M, Fröhlich H, Detection of Hidden Influences and Missed Reactions in Quantitative Models of Biological Systems, presented at the Basel Computational Biology Conference 2015 (contributed talk).
- **Engelhardt B**, Schrage R, Tränkle C, Mohr K, Kschischo M, Fröhlich H, Mathematical Analysis and Joint Modeling of the M2 Receptor - dependent Signaling And Secondary Messenger Network in CHO Cells, presented at the European Conference on Mathematical and Theoretical Biology 2014 in Gothenburg (contributed talk).

List of Figures

Elements in Figures 2.1, 2.3, 3.1 and 3.2 are adapted from Servier Medical Art; used under CC-BY.

List of Tables