Model reduction of mechanism-based pharmacodynamic models and its link to classical drug effect models

Dissertation

zur Erlangung des akademischen Grades
doctor rerum naturalium (Dr. rer. nat)
an der Mathematisch-Naturwissenschaftlichen Fakultät
Institut für Mathematik

1. Reviewer: Prof. Dr. Wilhelm Huisinga
2. Reviewer: Prof. Dr. Carsten Hartmann
3. Reviewer: Prof. Dr. Thorsten Lehr

Supervisors: Prof. Dr. Wilhelm Huisinga and
Prof. Dr. Charlotte Kloft

eingereicht von Jane Knöchel

Berlin, 6 Dezember 2019

Published online at the
Institutional Repository of the University of Potsdam:
https://doi.org/10.25932/publishup-44059
https://nbn-resolving.org/urn:nbn:de:kobv:517-opus4-440598

Jane Knöchel

Model reduction of mechanism-based pharmacodynamic models and its link to classical drug effect models

Dissertation zur Erlangung des akademischen Grades
doctor rerum naturalium (Dr. rer. nat), Berlin, 6 Dezember 2019
Reviewers: Prof. Dr. Wilhelm Huisinga, Prof. Dr. Thorsten Lehr and Prof. Dr. Carsten Hartmann
Supervisors: Prof. Dr. Wilhelm Huisinga and Prof. Dr. Charlotte Kloft

Universität Potsdam
Mathematische Modellierung und Systembiologie
Institut für Mathematik
Mathematisch-Naturwissenschaftliche Fakultät
Karl-Liebknecht-Str. 24/25
14476 Potsdam

Abstract

Continuous insight into biological processes has led to the development of large-scale, mechanistic systems biology models of pharmacologically relevant networks. While these models are typically designed to study the impact of diverse stimuli or perturbations on multiple system variables, the focus in pharmacological research is often on a specific input, e.g., the dose of a drug, and a specific output related to the drug effect or response in terms of some surrogate marker. To study a chosen input-output pair, the complexity of the interactions as well as the size of the models hinders easy access and understanding of the details of the input-output relationship.

The objective of this thesis is the development of a mathematical approach, in specific a model reduction technique, that allows (i) to quantify the importance of the different state variables for a given input-output relationship, and (ii) to reduce the dynamics to its essential features – allowing for a physiological interpretation of state variables as well as parameter estimation in the statistical analysis of clinical data. We develop a model reduction technique using a control theoretic setting by first defining a novel type of time-limited controllability and observability gramians for nonlinear systems. We then show the superiority of the time-limited generalised gramians for nonlinear systems in the context of balanced truncation for a benchmark system from control theory. The concept of time-limited controllability and observability gramians is subsequently used to introduce a state and time-dependent quantity called the input-response (ir) index that quantifies the importance of state variables for a given input-response relationship at a particular time. We subsequently link our approach to sensitivity analysis, thus, enabling for the first time the use of sensitivity coefficients for state space reduction. The sensitivity based ir-indices are given as a product of two sensitivity coefficients. This allows not only for a computational more efficient calculation but also for a clear distinction of the extent to which the input impacts a state variable and the extent to which a state variable impacts the output.

The ir-indices give insight into the coordinated action of specific state variables for a chosen input-response relationship. Our developed model reduction technique results in reduced models that still allow for a mechanistic interpretation in terms of the quantities/state variables of the original system, which is a key requirement in the field of systems pharmacology and systems biology and distinguished the reduced models from so-called empirical drug effect models. The ir-indices are explicitly defined with respect to a reference trajectory and thereby dependent on the initial state (this is an important feature of the measure). This is demonstrated for an example from the field of systems pharmacology, showing that the reduced models are very informative in their ability to detect (genetic) deficiencies in certain physiological entities. Comparing our novel model reduction technique to the already existing techniques shows its superiority.

The novel input-response index as a measure of the importance of state variables provides a powerful tool for understanding the complex dynamics of large-scale systems in the context of a specific drug-response relationship. Furthermore, the indices provide a means for a very efficient model order reduction and, thus, an important step towards translating insight from biological processes incorporated in detailed systems pharmacology models into the population analysis of clinical data.

Zusammenfassung

Die kontinuierliche Erforschung von biologischen Prozessen hat zur Entwicklung umfangreicher, mechanistischer systembiologischer Modelle von pharmakologisch relevanten Netzwerken beigetragen. Während diese Modelle in der Regel darauf ausgelegt sind, die Auswirkung von Stimuli oder Störungen auf die Systemdynamik zu untersuchen, liegt der Fokus in der pharmakologischen Forschung häufig auf einer bestimmten Kontrolle, z.B. der Dosis eines Wirkstoffes, und einer bestimmten Ausgangsgröße, welche in Bezug steht zu dem Wirkstoff-Effekt oder das Ansprechen auf einen Wirkstoff über einen Surrogatmarker. Die Untersuchung und ein einfaches Verständnis einer spezifischen Eingabe-Ausgabe-Beziehung wird durch die Komplexität der Interaktionen sowie der Größe des Modells erschwert.
Das Ziel dieser vorliegenden Arbeit ist die Entwicklung eines mathematischen Ansatzes, insbesondere eines Modellreduktionsverfahrens, der es ermöglicht, (i) die Bedeutung der verschiedenen Zustandsvariablen für eine gegebene Eingabe-Ausgabe-Beziehung zu quantifizieren, und (ii) die Dynamik des Systems auf seine wesentlichen Merkmale zu reduzieren, während gleichzeitig die physiologische Interpretierbarkeit von Zustandsvariablen sowie eine Parameterschätzung im Rahmen von einer statistischen Analyse klinischer Daten ermöglicht wird. Unter Verwendung eines kontrolltheoretischen Settings entwickeln wir eine Modellreduktionstechnik, indem wir vorerst einen neuartigen Typ von zeitlich begrenzten Kontrollliербarkeits- und Beobachtbarkeitsgramian für nichtlineare Systeme definieren. Anschließend zeigen wir die Überlegenheit der zeitlich begrenzten verallgemeinerten Gramian für nichtlineare Systeme im Kontext von Balanced Truncation am Beispiel eines Benchmark-Systems aus der Kontrolltheorie. Wir nutzten das Konzept der zeitlich begrenzten Kontrollliербarkeits- und Beobachtbarkeitsgramian, um eine neue Zustands- und zeitabhängige Größe, die als Input-Response (IR-) Index bezeichnet wird, einzuführen. Dieser Index quantifiziert die Bedeutung von Zustandsvariablen zu einem bestimmten Zeitpunkt für eine bestimmte Eingabe-Ausgabe-Beziehung. Schließlich verknüpfen wir unseren Ansatz mit der Sensitivitätsanalyse und ermöglichen so erstmals die Verwendung von Sensitivitätskoeffizienten im Rahmen der Reduktion des Zustandsraumes. Wir erhalten die sensitivitätsbasierten IR-Indizes als Produkt zweier Sensitivitätskoeffizienten. Dies ermöglicht nicht nur eine effizientere Berechnung, sondern auch eine klare Unterscheidung, inwieweit die Eingabe eine Zustandsvariable beeinflusst und inwieweit eine Zustandsvariable die Ausgabe beeinflusst.
Mit Hilfe der IR-Indizes erhalten wir einen Einblick in den koordinierten Ablauf der Aktivierung von spezifischen Zustandsvariablen für eine ausgewählte Eingabe-Ausgabe-Beziehung. Unser entwickeltes Modellreduktionsverfahren resultiert in reduzierten Modelle, welche eine mechanistische Interpretation hinsichtlich der Originalgrößen und Zustandsvariablen des Ursprungssystems zulassen. Dies war eine wichtige Anforderung an das Verfahren von Seiten der Systempharmakologie und -biologie. Die reduzierten Modelle unterscheiden sich damit wesentlich von den so genannten empirischen Wirkstoff-Effekt-Modellen. Die IR-Indizes sind explizit in Bezug auf eine Referenzlösung definiert und damit vom Anfangszustand abhängig (dies ist ein wichtiges Merkmal der Indizes). Wir zeigen anhand eines Beispiels aus dem Bereich der Systempharmakologie, dass die reduzierten Modelle sehr aussagekräftig sind, um (genetische) Mängel in bestimmten physiologischen Einheiten festzustellen. Der Vergleich unseres neuartigen Modellreduktionsverfahrens mit den bereits vorhandenen Verfahren zeigt

dessen Überlegenheit.
Der neuartige IR-Index als Maß für die Wichtigkeit von Zustandsvariablen bietet ein leistungsfähiges mathematisches Werkzeug zum Verständnis und der Analyse der komplexen Dynamik von großen Systemen im Kontext einer bestimmten Wirkstoff-Effekt-Beziehung. Darüber hinaus sind die Indizes eine wichtige Grundlage für das eingeführte und sehr effiziente Modellreduktionsverfahren. Insgesamt stellt dies einen wichtigen Schritt zur Nutzung von Erkenntnissen über biologische Prozesse in Form von detaillierten systempharmakologischen Modellen in der Populationsanalyse klinischer Daten dar.

Acknowledgements

I would like to acknowledge all the people without whom this thesis would not have been possible. Looking back on the four years at Potsdam University a number of people stand out and I would like to express my particular gratitude to the following individuals:

Foremost, I would like to thank my supervisor, Prof. Dr. Wilhelm Huisinga for enabling and supporting my doctoral studies. The fact that his door was always open and his enthusiasm for this field of research have markedly shaped and enriched this thesis.

My sincere thanks also goes to Prof. Dr. Charlotte Kloft for her insightful discussions during project presentations which sharpened my communication skills in the interdisciplinary research field.

I thank the Pharmacometrics group at AbbVie Ludwigshafen, in particular Dr. Sven Mensing, for many interesting visits to their group as well as their perspectives on my work.

For the unique and inspiring working atmosphere, I would like to thank my colleagues at the University of Potsdam and at the Freie University Berlin. In particular, I am very happy and grateful that Saskia Fuhrmann started her PhD thesis together with me. Over the years Saskia has been indispensable for me, thanks Saskia for sharing the whole experience with me and always answering any pharmaceutical or non-pharmaceutical question I posed to you. I would also like to thank Lisa Ehmann for introducing me in her interesting research topic and for her many valuable suggestions on this thesis. Furthermore I would like to acknowledge Dr. Niklas Hartung for the many mathematical discussion about the gramians, sensitivity analysis or other topics. As well as Corinna Mayer for new impulses in our working group and great inputs on parts of this thesis. I would also like to thank Undine Falkenhagen and Lena Klopp-Schulze for the many helpful comments on the manuscript of this thesis.

Last but not least I warmly thank my family for their continuous encouragement. Especially I want to thank my Sister, who not only always brightened my mood but also has found various new ways to motivate me.

This work was financially supported by the Graduate Research Training Program 'PharMetrX: Pharmacometrics and Computational Disease Modelling'

Publications

Parts of this thesis have been published during the PhD.

Parts of Chapter 4 (Section 4.3) and Chapter 5 (Section 5.1.1-5.1.2) have been published in *Journal of Pharmacokinetics and Pharmacodynamics* in collaboration with Charlotte Kloft and Wilhelm Huisinga, with the title 'Understanding and reducing complex systems pharmacology models based on a novel input-response index'.

Contents

List of Figures

List of Tables

Acronyms

BMD bone mineral density

CI confidence interval

CMT compartment

DAE differential algebraic equation

EGFR epidermal growth factor receptor

ERK extracellular signal-regulated kinase

F fibrin

Fg fibrinogen

ILDM Intrinsic Low-Dimensional Manifold

INR international normalised ratio

LTI linear time invariant

LTV linear time-varying

MAPK mitogen-activated protein kinase

MCA metabolic control analysis

MOR model order reduction

NPT normal prothrombin concentration

ODE ordinary differential equation

PK pharmacokinetics

PD pharmacodynamics

$PGF_{2\alpha}$ prostagladin $F2\alpha$

PT prothrombin time

QSSA quasi-steady state approximation

RC resistor-capacitor

TF tissue factor

1. Introduction

"Everything should be made as simple as possible, but not simpler."
- Albert Einstein -

In clinical pharmacology, it is of major importance to understand the dose-response relationship for a therapeutic agent [54]. During the drug development process an enormous amount of experimental and clinical data is generated, and using this information efficiently to guide decision-making across the various phases is a challenging task [64]. Mathematical models are used in clinical development, (i) to understand and guide the selection of optimal drug candidates in drug discovery and preclinical development; (ii) to asses safety, efficacy and characterise variability in exposure and response of drugs [71]. In drug discovery and preclinical development, continuous insight into biological processes has led to the development of large-scale, mechanistic systems biology models of pharmacologically relevant networks. In clinical development, the characterisation of the pharmacokinetics (PK) and pharmacodynamic (PD) of the drug are of interest [86]. To study the PK and PD, the drug is administered and the resulting drug concentration in the body, as well as a surrogate marker for the drug effect, are measured in several patients repeatedly over time. These types of data, collected in clinical trials, are usually referred to as population data. An important aspect of the analysis of population data is the identification of patient characteristics (covariates), that partly explain the observed variability in the concentration-time and/or drug response profiles [105], allowing, for example, to identify patient populations at risk e.g. therapy failure or personalise treatment. For the analysis of population data of a clinical trial, so-called empirical models are prevalently employed, because the large-scale mechanistic systems biology models can not be used for the estimation of individual parameters. For empirical modelling approaches, the model structure and the model parameters are solely derived from data and, thus, highly-dependent on the data quality and quantity. The resulting parsimonious models – i.e. a model with the fewest possible number of parameters – [2] are used for decision making (e.g. dose selection, clinical trial design), however, completely neglecting the insight including mechanistic covariates gained in the earlier drug development stages. The integration of covariates in the empirical models is mostly based on statistical significance often neglecting the correlation between covariates [105]. Thus, an approach that allows to translate the knowledge present in large-scale mechanistic models including mechanistic covariates across the different drug development phases is highly desirable, yet still a challenging task.

While the large-scale mechanistic models in the early drug development stages are typically designed to study the impact of diverse inputs or stimuli, the focus in later stages of drug development is often on a specific input, e.g. dose of a drug, and a specific output related to the drug effect or response in terms of some surrogate marker. Understanding the specific input-response relationship of interest in pharmacology is a challenging task due to the complexity of interactions between the state variables (nonlinearity) and the high-dimensionality of the model. In general, model reduction is a crucial step towards solving this issue.

In the literature, various model reduction approaches have been proposed in the different fields of application [110, 5, 92]. Due to the specific requirements on the model structure in systems pharmacology, mainly three model reduction techniques are used [92]: sensitivity

analysis [103, 95, 139], time scale analysis [73, 138] and lumping [1, 11, 77, 99].
Sensitivity analysis quantifies the influence of the parameters and initial states on a specific output variable [139]. Typically a reduced model is obtained by eliminating those reactions, whose parameters have little or no effect on the model predictions [79]. However, model reduction techniques solely based on the importance of parameters obtained by sensitivity analysis have resulted in models with low accuracy [123, 126]. Various attempts to use additional features or techniques have not resulted in a unique criterion for the elimination of parameters [25, 79]. Therefore, an established model reduction technique based on the sensitivities of the system is still missing. Time scale analysis partitions the systems into slow and fast state variables. The reduced model is then obtained by a quasi-steady-state approximation (QSSA) of the fast state variables. The mathematical justification of the QSSA is singular perturbation theory [113]. Often, however, a substantial reduction of the system is possible without time scale separation. In general, the technique is, thus, not efficient enough as a single tool. Lumping is the most widely used technique [92]. It is based on a coarse-graining of the state space by merging several state variables into pseudo-state variables. There exist two different variants: proper and improper lumping. Proper lumping can be understood as a partitioning of the state variables, each partition is then associated with a single dynamical state variable in the reduced system [118]. Finding an optimal proper lumping scheme is a challenging task, especially for large-scale systems, since it is practically impossible to test all possible lumped models due to the curse of dimensionality. As a consequence, heuristics are used to find a nearly optimal lumping scheme. In [1], a recursive greedy search strategy is proposed: among all possible lumps of two state variables the pair of state variables is lumped that results in the lowest approximation error. It remains unclear, however, (i) to what extent the sequence, in which state variables are lumped, and (ii) to what extent the state variables, which are not important for a specific input-response relationship, impact the final reduced model. This might have implications on both, the approximation quality as well as the interpretation of the reduced model.

Very recently, a model reduction technique based on balanced truncation was developed and applied to a large systems biology model of the epidermal growth factor receptor (EGFR) and nerve growth factor receptor pathways [117]. The approach, however, is based on improper lumping (with the problems mentioned above) and requires substantial preprocessing (non-dimensionalisation, conservation analysis, preconditioning).

In summary, a model reduction approach that allows (i) to quantify the importance of the different constituents for a given input-output relationship, and (ii) to reduce the dynamics to its essential features is still lacking, yet highly desirable in the context of clinical development.

In this thesis, we address this gap by exploiting a control-theoretic setting for our model reduction. First, we define a novel type of time-limited empirical gramians with respect to a reference trajectory of the nonlinear system. Subsequently, the concept of empirical gramian is used to develop a novel state and time-dependent quantity called the input-response (ir) index that allows one to understand the essential features of the system dynamics for a given input-response relationship over time. We introduce an automated and efficient model reduction technique based on the ir-indices. The state variables and parameters in the reduced model still allow for a mechanistic interpretation in terms of quantities of the

original system and, thus, physical entities. By showing that the ir-indices can be linked to sensitivity analysis, we enable for the first time the use of sensitivity coefficients for state-space reduction.

Using an example from the field of systems pharmacology, we additionally show that the definition of the ir-indices with respect to the reference solution is an important feature of the measure as we are able to reproduce experimental findings. Our introduced model reduction technique is dependent on the considered time interval and the ranking according to the maximal value of the ir-indices. Therefore, we investigate how the ranking and time interval influences our technique showing that the ir-indices provide potential reduced model structures for different time intervals and insight into the reduced models based on a random ranking. Comparing our novel model reduction technique to the already existing techniques shows its superiority.

Outline: In Chapter 2, we introduce the general concepts of pharmacodynamic modelling as well as the relevant large-scale mechanistic pharmacodynamic models, later used to illustrate the strength of the novel approach. In Chapter 3, we outline the various existing techniques for model reduction, to enable a thorough discussion of the different techniques in the application context. Subsequently, in Chapter 4, we develop the novel model reduction technique based on the input-response characterisation. In Chapter 5, we apply our novel model reduction approach to various input-response relationships of two large-scale mechanistic models, the blood coagulation network [130] and the EGFR signalling cascade [55]. Chapter 6 gives a discussion and outlook of this thesis. The conclusion is then presented in Chapter 7.

2. Modelling the drug effect: empirical versus systems pharmacology models

It is of major importance in clinical pharmacology to understand the dose-effect/response relationship to demonstrate the safety and efficacy of a therapeutic agent [54]. This is studied amongst other aspects by (i) describing the drug concentration-time profile in body fluids resulting from the administration of a drug (pharmacokinetics) and (ii) characterising the observed effect resulting from the drug concentration (pharmacodynamics). This thesis will focus on the pharmacodynamics (PD) of a drug assuming a given drug concentration-time profile or pharmacokinetic (PK) model. Mathematical models are used to analyse the repeatedly measured clinical drug response over time in several patients (population analysis of clinical data). The two different types of modelling approaches for the pharmacodynamic analysis are (i) data-driven and (ii) mechanism-based models [27]. In the first section, these two modelling approaches are briefly presented. We then introduce two exemplary mechanism-based models from two different fields of pharmacology: the epidermal growth factor receptor (EGFR) system and the blood coagulation network. To guide understanding, we provide a short biological background for each system. In the last section, the need for model reduction is illustrated by discussing potential modelling approaches for two example drugs affecting the blood coagulation network: warfarin and rivaroxaban.

2.1. Data-driven versus mechanism-based pharmacodynamic (PD) models

Often pharmacodynamics is characterised in terms of 'what the drug does to the body' [17]. The measured PD data in a clinical trial can be either continuous ($\in \mathbb{R}$), e.g. concentration-time profile of biomarkers, or discrete ($\in \mathbb{Z}$), e.g. survival analysis of patients on a specific treatment [125]. In this section, we will introduce two different modelling approaches for continuous PD data.

2.1.1. Empirical/data-driven PD models

For so-called empirical PD or data-driven models, the model structure and model parameters are derived from clinical data and often lack a physiological explanation in terms of physical quantities. The resulting parsimonious model is the model with the fewest possible number of parameters [2].

Examples for empirical PD models are:

- linear model:

$$E(t) = E_0 \pm E_{\text{slope}} \cdot C_{\text{plasma}}(t) \tag{1}$$

- $\mathrm{E}_{\max}$ model :

$$E(t) = E_0 \pm \frac{E_{\max} \cdot C_{\text{plasma}}(t)^{\gamma}}{\mathrm{EC}_{50}^{\gamma} + C_{\text{plasma}}(t)^{\gamma}} \tag{2}$$

- effect-compartment model:

$$\frac{dC_{\text{effect}}(t)}{dt} = k_{\text{effect}} \cdot (C_{\text{plasma}}(t) - C_{\text{effect}}(t)) \tag{3}$$

$$E(t) = E_0 \pm \frac{E_{\max} \cdot C_{\text{effect}}(t)^{\gamma}}{\mathrm{EC}_{50}^{\gamma} + C_{\text{effect}}(t)^{\gamma}} \tag{4}$$

- indirect response model [114]:

$$\frac{dE(t)}{dt} = k_{\text{in}} \cdot \frac{I_{\max} \cdot C_{\text{plasma}}(t)}{\mathrm{IC}_{50} + C_{\text{plasma}}(t)} - k_{\text{out}} \cdot E(t) \tag{5}$$

where $E_0 \in \mathbb{R}_+$ is the baseline value of the drug effect E (when no drug concentration is present), $E_{\max} \in \mathbb{R}_+$ is the maximal change in drug effect/response that the drug can induce, $\mathrm{EC}_{50} \in \mathbb{R}_+$ is the drug concentration that produces the half maximal drug effect/response and $\gamma \in \mathbb{R}_+$ is the so-called Hill exponent which is primarily used to provide a better fit, $I_{\max} \in \mathbb{R}_+$ is the maximal inhibiting effect that the drug can induce and $\mathrm{IC}_{50} \in \mathbb{R}_+$ is the drug concentration that produces the half maximal inhibitory effect.

The algebraic equations (1) and (2) are known in pharmacology as the so-called direct link, direct effect or direct response models [125, 54, 28]. These types of models are used if there is no time delay observed between the drug concentration in plasma and the drug effect/response. For these models, the assumption is made that the drug concentration in plasma is proportional to the concentration at the effect site (e.g. receptor site) [28].

If the maximal drug effect is delayed compared to the maximal drug concentration in plasma, various modelling approaches are used to account for this time delay. In general, this type of models is referred to as so-called indirect effect models in pharmacology [125]. Mathematically the time delay could be modelled by delay differential equations, this is, however, usually not done in the field of pharmacology. Instead, the delay differential equation is often approximated by the use of transit compartment chains of different length [81]. The simplest model is given by a transit compartment chain of length 1, also known as the effect-compartment model (cf. equ. (3)). The drug effect is then calculated based on an algebraic equation e.g. equ. (4). Another approach for modelling the time delay between the drug concentration and the drug effect is by an additional ordinary differential equation for the drug effect, where the drug concentration in plasma influences either production or loss of the drug effect directly by stimulation or inhibition of the process (example given in equ. (5)).

2.1.2. System pharmacology/mechanism-based PD models

The model structure of a mechanistic pharmacodynamic model explicitly and conceptually represents the constituent parts and mechanism of a general physiological process to the de-

gree of currently available knowledge [24, 17]. These models are often built by integrating diverse types of data to provide a comprehensive understanding of the underlying biological processes [17]. The model structure and the model parameters are, thus, derived from the knowledge of the physiology and biochemistry of the organism and the models do not necessarily require clinical data to allow for predictive simulations of clinical trials not yet completed [9]. In many examples of relevance, mechanistic models are given by a system of nonlinear ordinary differential equations based on mass-action principles [8].

Examples for mechanistic pharmacodynamic models are the calcium homeostasis and the bone model [97, 106] (cf. Figure 1(A)), the bovine estrous cycle [120] (cf. Figure 1(B)), epidermal growth factor receptor (EGFR) system [111, 55](cf. section 2.2, Figure 2) and blood coagulation network [130, 40](cf. section 2.3, Figure 6).

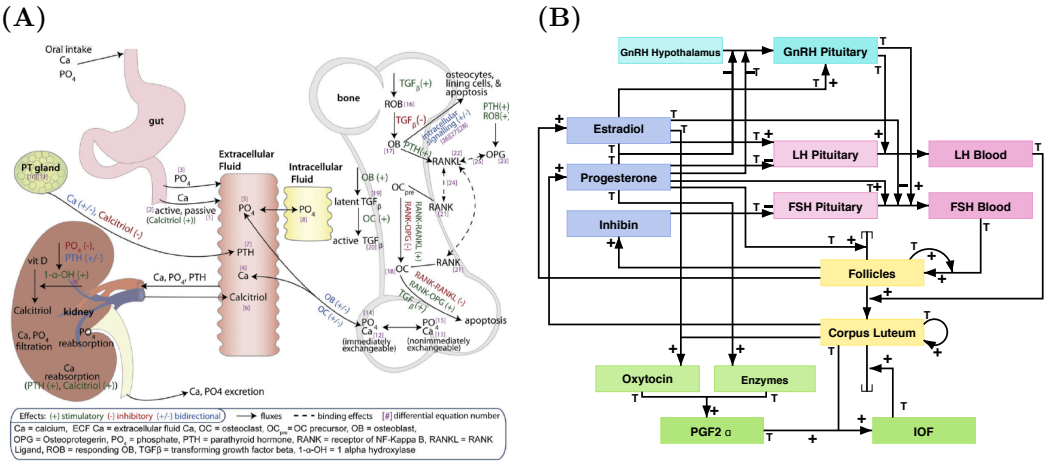


Figure 1: **Mechanistic pharmacodynamic model examples: (A)** calcium homeostasis and bone model [97, 106] **(B)** bovine estrous cycle [120]

The calcium homeostasis and bone model [97, 106] (cf. Figure 1(A)) was used to quantify the effect of estrogen suppression therapy on the long-term loss of bone mineral density (BMD). In [106] it was shown that using the comprehensive model for calcium homeostasis and bone remodelling one could identify reliable predictor biomarkers of the long-term BMD. Thus, answering questions such as which biomarker offers an early sensitivity measure of the long-term BMD outcome.

In contrast, the bovine estrous cycle [120] (cf. Figure 1(B)) was employed to study the estrus synchronisation protocol where Prostagladin $F2\alpha$ (PGFα) is administered to facilitate the timing of artificial insemination. Due to the short half-life of PGFα it is highly important to identify the optimal time points of administration of PGFα.

In general, mechanistic pharmacodynamic models can answer questions, which can not necessarily be addressed by traditional empirical models: e.g. (i) effects of treatment (changes), (ii) treatment discontinuation, (iii) effect of prior treatments or (iv) drug safety questions [96].

In the next section, two large-scale mechanistic models are presented more in detail. They were chosen based on their relevance for the field of oncology and haematology. Important

modelling aspects are highlighted and their application in the pharmacological context is shortly described. Both large-scale mechanistic models are used in section 5 as a challenging example for our novel model reduction technique.

2.2. Oncology: epidermal growth factor receptor (EGFR) system

In this section, a short biological background on the epidermal growth factor receptor (EGFR) system and associated therapeutics are provided. Subsequently, we introduce a comprehensive mathematical model of the epidermal growth factor (EGF) signalling pathway, which is used in section 5 as a challenging example for our novel model reduction technique.

2.2.1. Biological background of the EGFR signalling pathway

The EGFR signalling cascade is an important pathway in cell division, death, motility and adhesion [112, 137, 63]. In addition, it is of key interest in the development of anti-cancer therapies, as the pathway is often dysfunctional in tumour cells. A deeper understanding of the pathway can, therefore, yield direct clinical benefit.

The pathway is activated by the binding of EGF to its receptor (EGFR). Dimerised EGFR autophosphorylates, creating binding sites for the adapter proteins GAP, Shc, Grb2 and Sos. The signalling cascade has two major pathways: the Shc-dependent and the Shc-independent pathway. The two pathways, however, do not act independently from each other since there are many molecules involved in both pathways. Both pathways have the ability to activate Ras:GTP, a well-known oncogene. Ras:GTP [13] is the branching point of the two signalling pathways and activates the mitogen-activated protein (MAP) kinase cascade through the activation of Raf. The signalling output is the transient phosphorylation of extracellular signal-regulated kinase (ERK). A simplified reaction network of the EGF signalling pathway is given in Figure 2. While the principal flow of the downstream signalling and activation sequence is generally known, the kinetic network of this cascade is not completely understood [111].

In cancer cells two major mutations affecting the EGF signalling pathway have been reported: KRas, a mutation of the Ras protein, and Braf, a mutation of Raf [91]. These mutations do not occur concomitantly because their combined impact would be incompatible with proliferation, as an excess ERK signalling could lead to cell cycle arrest, differentiation, senescence or even cell death [91]. This shows the importance of the two oncogenes Ras and Raf. The similarities in the sequence and structure of Ras proteins [46] make targeting the Ras protein directly very challenging [62]. Therefore, the targeting of the EGFR signalling cascade is done by either targeting Raf, MEK or ERK [62]. The standard of care is the combination of a Raf inhibitor with a MEK inhibitor [62].

2.2.2. Comprehensive mathematical model of the EGFR system

The mitogen-activated protein kinase (MAPK) pathway is one of the best-studied systems because of its relevance to the development of cancer therapies [124]. The EGFR reaction

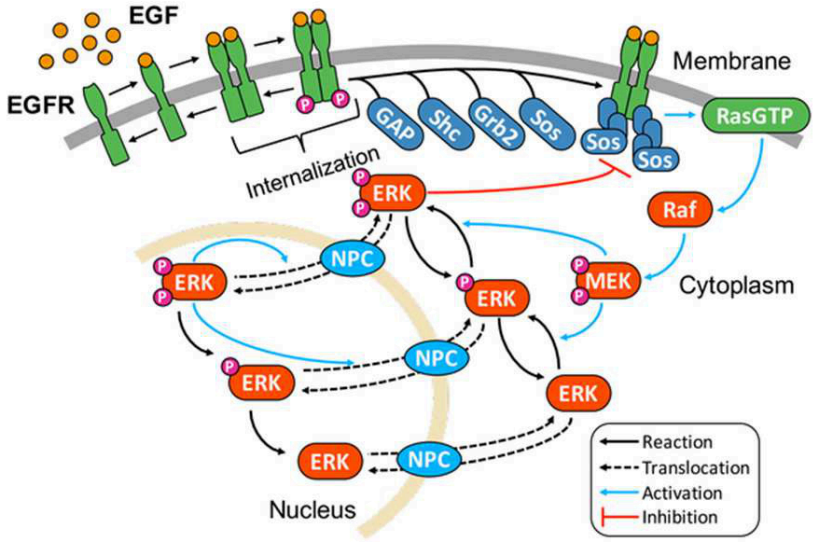


Figure 2: **Simplified reaction network of the EGF signalling pathway including nuclear ERK** (taken from [58]).

network including the MAPK pathway is initiated by the binding of the epidermal growth factor to its receptor and results in transient phosphorylation of the extracellular signal-regulated kinase. We used a detailed model of the EGFR reaction network [112, 55] consisting of 106 state variables, 148 (mostly reversible) reactions and 95 parameters. See Figure 4 for a common illustration that includes only the membrane-bound receptor part of the pathway, while there is in addition an internalised receptor-based part. The original model includes some 'degraded molecular species' that serve as a substitute for various degradation products. In some cases, this hindered the exploitation of conservation laws so that we modelled in these cases the degradation products as separate molecular species (thereby increasing the number of molecular species by six to 112). This extension did not change the remaining system dynamics.

In the absence of EGF, the stimulus of the system, all state variables are assumed to be in steady-state. All initial conditions and parameter values for the model can be found in [55, Suppl. Table 2]. The transient ERK-PP profile (cf. Figure 3) upon stimulation by EGF is an important determinant for the subsequent cellular response [55] and has been validated experimentally [56].

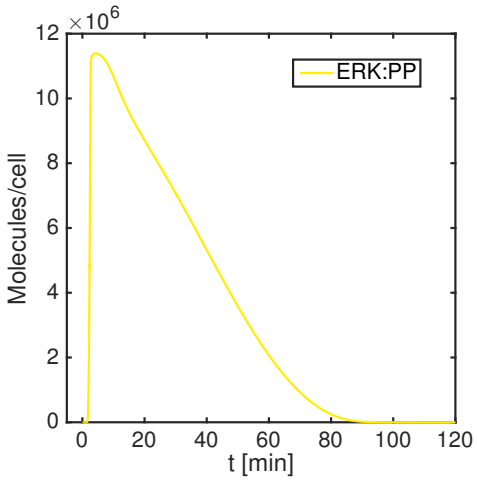


Figure 3: **ERK-PP transient signal**.

The EGF receptor signalling system is one of the best-studied pathways in systems biology [135]. Due to the complexity, the EGF receptor signalling system has been used many times to illustrate diverse model reduction approaches [79, 4, 55, 21, 132]. In [4] the assumption is made to ignore receptor internalisation and, thus, started from a nonlinear model with 35 state variables. The remaining system is then linearised around the equilibrium and a model reduction based on control theory is applied to further reduce the model to 8 state variables. In [79] sensitivity, principal component and flux analysis are used to reduce the full epidermal growth factor-mediated signalling model presented in [112]. They reduced the system by the elimination of reactions and, thus, found that the EGF mediated signalling was mainly mediated by the Shc-dependent pathway. In [55] an extensive sensitivity analysis for the three different characteristics of the output of the EGF-ERK-PP system (amplitude, duration of signal and integrated response) is performed. They showed that all characteristics were highly controlled by the processes involved in MEK phosphorylation by Raf and Raf dephosphorylation. However, this result was not used for model reduction. In [21] simulation studies were used to analyse the input-output behaviour of submodules of the system. This heuristic approach was then used to reduce the model. A similar approach was employed by [132]. Their analysis of the system revealed redundancies of the system that help maintain functional robustness. The various reduced models obtained by the different model reduction techniques presented in this paragraph are very heterogenous and there exists little consensus between the resulting reduced models.

Although the EGFR system has been intensively studied, it is still not fully understood how the signal is translated through the system and which parts are most important for the dynamics of the system. The identification of the most important state variables of the dynamics of the system might reveal new insights for the design or optimisation of anti-EGFR drug therapies.

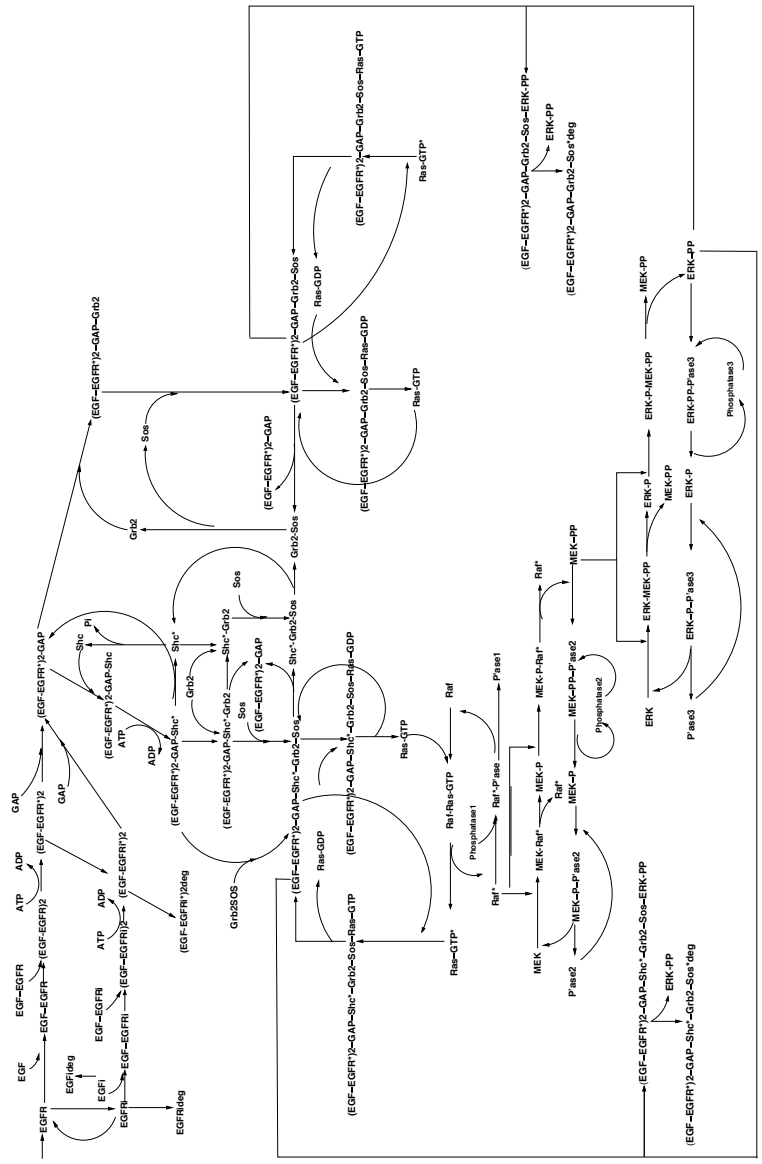


Figure 4: **EGF receptor - induced MAP kinase cascade model** (based on [55, Fig. 1]). The model comprises 106 molecular species (state variables) involved in 148 reactions. The input signal is the epidermal growth factor (EGF) that eventually activates the extracellular signal-regulated kinase (ERK). In the reaction graph, only parts of the state variables are shown; in particular, most of the so-called internalised forms are not shown.

2.3. Haematology: blood coagulation network

In this section, some background information on the blood coagulation network and associated anticoagulant therapy are presented. Subsequently, a comprehensive mathematical model is introduced, which is used in section 5.1 for the illustration of our novel model reduction technique in application to different aspects of the human blood coagulation network.

2.3.1. Biological background of the blood coagulation network

The process of clotting including the blood coagulation network is of vital importance for preventing leakage in the vascular system in cases of an injury [3]. The blood coagulation is a complex interaction network of many proteins and enzymes (cf. Figure 5). The blood coagulation functions reliably even though the activity and amount of proteins and enzymes vary between individuals [122].

The different coagulation factors are present in the blood in an inactive form [93]. The cascade can be initiated by the release of the tissue factor (TF) leading to the activation of the so-called extrinsic pathway (given in Figure 5 by red arrows), and/or the contact factor (CA), activating the so-called intrinsic pathway (yellow arrows in Figure 5), both pathways end in a common pathway (denoted by orange arrows in Figure 5) ultimately leading to the activation of fibrinogen to fibrin, which then forms the fibrin strands that strengthen the blood clot [69]. The various feedbacks in the network make it difficult to understand the PK-PD relationship for anticoagulants, like warfarin.

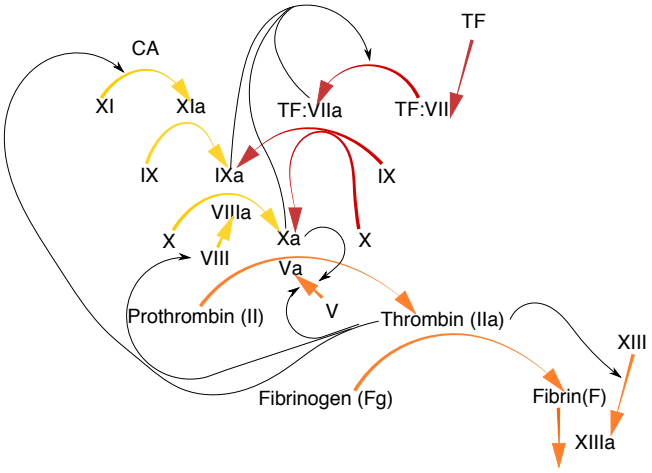


Figure 5: **Simplified interaction network of the blood coagulation cascade** (based on [88]). The cascade can be initiated by the release of the tissue factor (TF) leading to the activation of the extrinsic pathway (red arrows), and/or the contact factor (CA), activating the intrinsic pathway (yellow arrows), both pathways end in a common pathway (orange arrow) ultimately leading to the activation of fibrinogen to fibrin, which then forms the fibrin strands. The back arrows represent feedbacks of the proteins on the network.

The activated factors have a short half-life (on the time scale of seconds to minutes [59]) compared to the inactivated factors (on the time scale of hours to days) so that the clotting ceases after a short time period. There exist, however, other potent activators of specific factors in the blood coagulation system such as the brown snake venom which induces the activation of prothrombin [40].

To measure the clotting ability of the blood many clinical assays, usually, *in vitro* procedures, have been developed and are used in clinical practice [68]. In the following two of these assays are discussed in detail, because of their clinical relevance for the two model drugs analysed in this thesis (warfarin and rivaroxaban, compare section 2.4 and 5.1.3). One is the prothrombin time (PT) test, an *in vitro* blood coagulation test quantifying the activity of the so-called extrinsic pathway activated by the tissue factor TF. The reference range of the prothrombin time for a healthy donor is around 11.6-13.8 seconds [14]. Due to the usage of different types and batches of manufacturer's TF, the result of PT test is usually reported as the international normalised ratio (INR) for better comparability between laboratories. The INR is defined as

$$\text{INR} = \frac{t_{\text{PT,warfarin therapy}}}{t_{\text{PT,reference}}} \in [0, \infty) \tag{6}$$

where $t_{\text{PT,reference}}$ is the calculated geometric mean prothrombin time of a minimum of 20 healthy donors for the specific TF used in the laboratory [102] and $t_{\text{PT,warfarin therapy}}$ is the measured prothrombin time for the patient receiving warfarin therapy. The other is the activated partial thromboplastin time (aPTT) test, which quantifies the function of the so-called intrinsic pathway.

Anticoagulants are routinely prescribed as prevention after hip or knee arthroplasty, treatment of deep vein thrombosis or pulmonary embolism and stroke prevention [16]. Warfarin is the most prescribed anticoagulant in practice [37, 61, 66] since its approval for use in humans in 1954, even though it has a very complex dose-response relationship [45]. Warfarin is administered as a racemic mixture of two enantiomers [65] (50% R-warfarin and 50% S-warfarin [23]) with S-warfarin being 3-5 times more potent than R-warfarin [44]. Warfarin acts on the vitamin K cycle which in turn reduces the synthesis of various factors of the blood coagulation namely factor II, VII, IX, X, protein S and protein C [101]. The antidote for warfarin is vitamin K. The effect of warfarin therapy on the blood coagulation network can be assessed via the PT test, where the measured prothrombin time is prolonged. The target range for the INR with warfarin therapy is 2 to 3 [107]. Due to warfarin's narrow therapeutic window (INR $\in [2, 3]$) under- or over-anticoagulation during initiation of warfarin therapy as well as maintenance needs to be avoided [53]. Over-anticoagulation results in the most common side effect of warfarin bleeding that occurs in up to 41% of patients treated with warfarin [65]. The risk of bleeding (over-anticoagulation) is highest during the dose-titration period of warfarin use [134]. It has been shown that several patient characteristics (e.g. age, gender, body weight and body surface area), pharmacologic factors (vitamin K intake, drug-drug interactions) and genetic factors (CYP2C9, VKORC1 and CYP4F2 polymorphisms) impact warfarin maintenance dose [33]. An average maintenance dose of 4-5 mg daily has been reported, however, the individual maintenance dose can vary up to 15-fold between patients [66]. Additionally shortening the time needed to reach the correct maintenance dose is

of high value not only for the patient but also to reduce the cost of treatment [53]. Currently, the initiation of warfarin is a very iterative process taking anything between a few weeks to months [61]. Thus, to maintain effective drug concentrations of warfarin to avoid adverse effects, enhance safety and efficacy of the drug and reduce the associated costs to warfarin therapy, therapeutic drug monitoring is required [108]. However, the frequency of monitoring varies greatly in clinical practice due to direct costs associated with the service [100].

Modern modelling techniques, in the form of systems pharmacology model (i.e. the blood coagulation network model [130] described in the next section) can serve as a valuable tool to help with individualised dose titration and maintenance dose for warfarin therapy. However, due to their complexity, these models can not yet be employed in this context.

Due to warfarin's difficult use in clinical practice [65], new anticoagulant drugs have been developed (e.g. rivaroxaban, apixaban and dabigatran) [47, 78, 34, 89, 22]. Each of these anticoagulant targets only one factor of the coagulation network. It has been shown that they have a comparable or even superior effect as warfarin [94, 38, 20]. Furthermore, for these drugs in contrast to warfarin, it is not necessary to monitor the drug effect [87].

Rivaroxaban is a direct factor Xa inhibitor [87] and has an almost linear relationship between rivaroxaban concentration and prothrombin time. The prothrombin time is considered as the primary response variable in existing population PD models for rivaroxaban response [136, 36]. In contrast to warfarin, rivaroxaban has been shown to have predictable pharmacokinetics and pharmacodynamics, which makes it highly attractive to use in clinical practice [87].

2.3.2. Comprehensive mathematical model of the blood coagulation network

The detailed mathematical model of the interplay between the various coagulation factors was developed in [130] to study *in vivo* warfarin and vitamin K therapy as well as *in vitro* blood coagulation tests (prothrombin time (PT) test, activated partial thromboplastin time (aPTT) test); see Figure 6 for illustration. The original model consists of 51 states. The model describes the coordinate activation of different proteins (so-called coagulation factors) upon stimulation that eventually result in the activation of fibrinogen. To study the effect of the Australian elapid venoms on the blood coagulation, in [39, 40] the model was extended to include the effect of the brown snake, tiger snake, rough-scaled snake and snake of the genus Hoplocephalus spp venom [39, 40, 121], resulting in a model of 62 state variables and 178 parameters. As was already stated in [130] the complexity of the model makes any inferences about the behaviour of specific model parts difficult. This highlights the necessity for the development of new mathematical tools suitable for the analysis of complex systems pharmacology/biology models.

We will illustrate our novel model reduction technique in application to various aspects of the human blood coagulation network. Therefore, important modelling aspects that were included in the model, developed in [130, 40], are discussed in the next sections to enable understanding and discussion of the resulting reduced models in the sections 5.1.

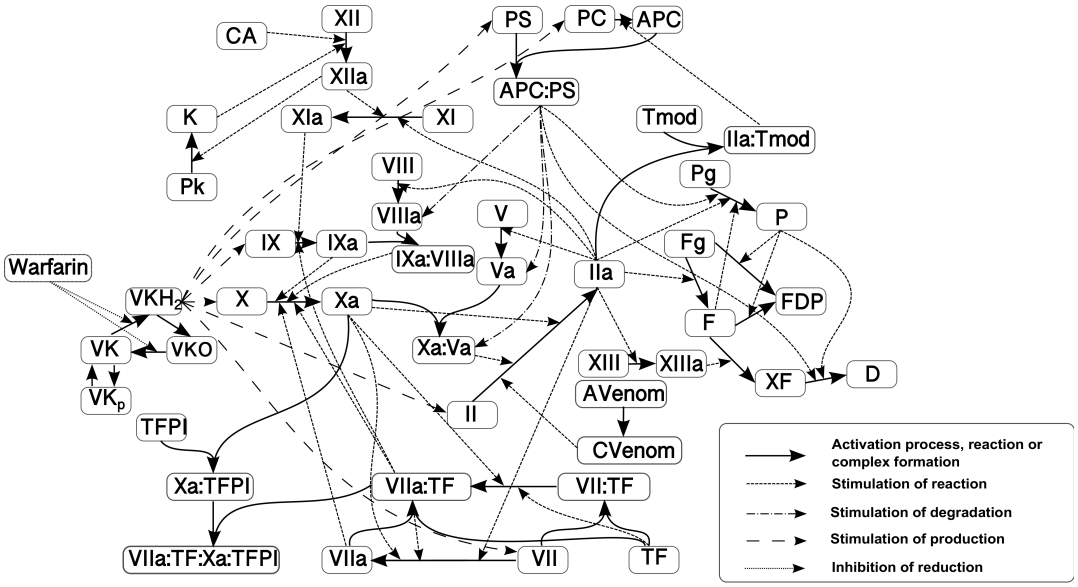


Figure 6: **Blood coagulation network including the effect of warfarin therapy and the brown snake venom on the activation of factor II**. Abbreviations used in the main text: activated protein C (APC); activator for the contact system (CA); fibrin (F); degradation product (FDP); fibrinogen (Fg); prothrombin (II); thrombin (IIa); plasmin (P); protein C (PC); plasminogen (Pg); protein S (PS); tissue factor (TF); vitamin K (VK). The graphic is based on [130, Fig. 1] and not all 62-state variables are shown.

Important modelling aspects of the blood coagulation model [130, 40]

The model consists of 62 state variables ($x \in \mathbb{R}^{62}$) and 178 parameters. In the following, we refer to a specific concentration of protein P by x_{P}, where P denotes the name of the factor/ protein. All activated factors are identified by the name of the factor followed by an additional small letter 'a' e.g. inactivated factor II has activated form IIa.

Inactive system and *in vivo*

Prior to the activation of the blood coagulation (the *in vivo* and inactivated setting) the model is a linear ordinary differential equation (ODE) system and all concentrations of the inactivated factors are constant (the system is in steady state). The activated forms of the factors are absent. The ODEs for the inactivated factor except for the vitamin K dependent factors (factor II, VII, IX, X, protein C and protein S) are given by

$$\frac{dx_i(t)}{dt} = p_i - d_i \cdot x_i(t) \tag{7}$$

with $p_i, d_i \in \mathbb{R}$ where x_i denotes the protein concentration and p_i, d_i the production and degradation rate constants of the protein/factor i [130]. The production rates were chosen by Wajima et al. [130] such that

$$p_i = d_i \cdot x_i(t_0) \tag{8}$$

with $x_i(t_0)$ denoting the physiological concentration of the inactivated factor/protein i. For the vitamin K dependent factors the production rate is dependent on vitamin K hydroquinone

(VKH2) and the ODEs for these factors become

$$\frac{dx_i(t)}{dt} = p_i(x_{\mathrm{VKH2}}(t)) - d_i \cdot x_i(t)$$

with $d_i \in \mathbb{R}$ for $i \in \{\mathrm{II}, \mathrm{VII}, \mathrm{IX}, \mathrm{X}, \mathrm{PC}, \mathrm{PS}\}$ denoting the indices of vitamin K dependent factors.

Active system / *in vitro*
In vitro e.g. clinical blood coagulation assays (PT test, aPTT test) the production of each inactivated factor and protein is assumed to be zero ($p_i = 0$) [130], so that the steady state *in vitro* is zero. The system only exhibits nonlinear behaviour if the blood coagulation is activated via the extrinsic (tissue factor) and/or the intrinsic (contact factor, CA in Figure 6) pathway. The intrinsic pathway is not of interest for the assessment of the effect of warfarin therapy (the PD model of warfarin) or rivaroxaban therapy due to the fact that both drugs don't significantly affect the factors in the intrinsic pathway. The model of the PT test is a nonlinear system.

In the active system setting the transformation rates $v(t)$ of inactivated factors to activated factors are given by Michaelis-Menten kinetics:

$$v(t) = \sum_{j \in J_{x_i}} \frac{V_{max,j} x_j(t)}{K_j + x_j(t)} \cdot x_i(t)$$

where J_{x_i} denotes the indices of the factors which activate the respective factor or induce its degradation and x_i denotes the concentration of the inactivated factor.

Brown snake venom - *in vivo* and activated system
The brown snake venom induces the degradation of fibrinogen by the activation of prothrombin (also called factor II), see Figure 2. After an initial steep decline, the fibrinogen concentration recovers to its initial steady-state concentration. In this case, the initial parts of both, the intrinsic and the extrinsic pathway do not play a role due to the direct activation of prothrombin by the brown snake venom.

2.3.3. Warfarin therapy - *in vivo* and *in vitro*

The understanding of the dose-response relationship to find the appropriate warfarin dosing scheme for each patient remains up to this day an important goal [107, 65]. The warfarin concentration influences the vitamin K cycle by inhibiting the chemical reduction reactions [129]. This, in turn, leads to a decrease in the synthesis of the vitamin K dependent factors (factor II, VII, IX, X, protein C and protein S).

There exist various empirical dose-effect models for warfarin. The most cited empirical dose-response model is by Hamberg et al. [45], consisting of two transit compartment chains with different length and mean transition time. This will be explained in more detail in the subsection 2.4.

As previously described, the drug effect of warfarin therapy is assessed by the PT test and

results in a prolongation of the prothrombin time. This is an *in vitro* procedure. Blood is drawn from the patient then the PT test is initialised experimentally by adding a well-defined amount of tissue factor (TF in Figure 6) to the diluted sample and the time is measured until the blood has clotted. By this procedure, the activity of the extrinsic pathway is measured.

This experimental procedure to assess the effect of the warfarin therapy via the PT test (as described in the previous paragraph) is mirrored in the model simulation (cf. for graphical depiction Figure 7). In [130], the PT test is initialised by 100 nM of initial TF concentration and the initial factor concentrations are scaled by 1/3 to account for the dilution process of the sample[1]. For $\mathrm{T_{Blood}} \in [t_0, t_1], t_0, t_1 \in \mathbb{R}$ the time of blood drawing (switching time ($\mathrm{T_{Blood}}$) from *in vivo* to *in vitro*) with t_0 time point of first warfarin dose, usually $t_0 = 0$, and t_1 is time point of e.g. end of warfarin therapy, the following equations are solved in the model

$$\dot{x} = \begin{cases} f_1(x(t), p_{in\ vivo}\ ), & t \in [t_0, \mathrm{T_{Blood}}], \quad \text{with} \quad x(0) = x_0 \\ f_2(x(t), p_{in\ vitro}\ ), & t \in (\mathrm{T_{Blood}}, T_1], \quad \text{with} \quad x(\mathrm{T_{Blood}}) = \frac{1}{3}x(\mathrm{T_{Blood}}) + x_{\mathrm{TF}}(0) \end{cases} \tag{9}$$

where $f_1(x(t), p_{in\ vivo}\ )$ represents the model in the *in vivo* setting, $f_2(x(t), p_{in\ vitro}\ )$ represents the *in vitro* situation and $x_{\mathrm{TF}}(0)$ is the initial TF concentration of 100 nM added to simulate the PT test. In this context the end time T_1 is given by $T_1 = \mathrm{T_{Blood}} + t_{\mathrm{PT}}$ where t_{PT} is of interest. The prothrombin time t_{PT} is dependent on $\mathrm{T_{Blood}}$ and thereby indirectly dependent on warfarin concentration until $\mathrm{T_{Blood}}$.

As previously mentioned the effect of the warfarin therapy is reported by the INR. In the blood coagulation model [130] both $t_{\mathrm{PT,reference}} = t_{\mathrm{PT}}(t_0)$ and $t_{\mathrm{PT,warfarin\ therapy}}(\mathrm{T_{Blood}})$ are defined by

$$\begin{aligned} t_{\mathrm{PT}}(\mathrm{T_{Blood}}) &= \inf\left\{ t > 0 : \int_{\mathrm{T_{Blood}}}^{t+\mathrm{T_{Blood}}} x_{\mathrm{F}}(\tau)\mathrm{d}\tau > 1.500^{\mathrm{nmol}/\mathrm{l\cdot\ s}} \right\} \\ &= \inf\left\{ t > 0 : \int_0^t x_{\mathrm{F}}(\mathrm{T_{Blood}} + s)\mathrm{d}s > 1.500^{\mathrm{nmol}/\mathrm{l\cdot\ s}} \right\} \end{aligned} \tag{10}$$

where x_F is the fibrin concentration. The threshold of $1.500^{\mathrm{nmol}/\mathrm{l\cdot\ s}}$ of the integral of fibrin is equivalent to an $\sim 30\%$ fibrinogen reduction in the standard plasma [130] and results in a feasible $t_{\mathrm{PT,reference}}$ value of 11.8 seconds. The INR is then calculated as given in equation (6). The typical target INR range which should be achieved with warfarin therapy is 2-3 [107], dependent on the disease. See Figure 8 for a typical predicted concentration time course of warfarin therapy with 4 mg daily and the corresponding INR time course.

[1] In [130] the concentration for the high TF scenario is reported to be 300 nM, but subsequently scaled to 1/3 of its value to reflect a dilution process.

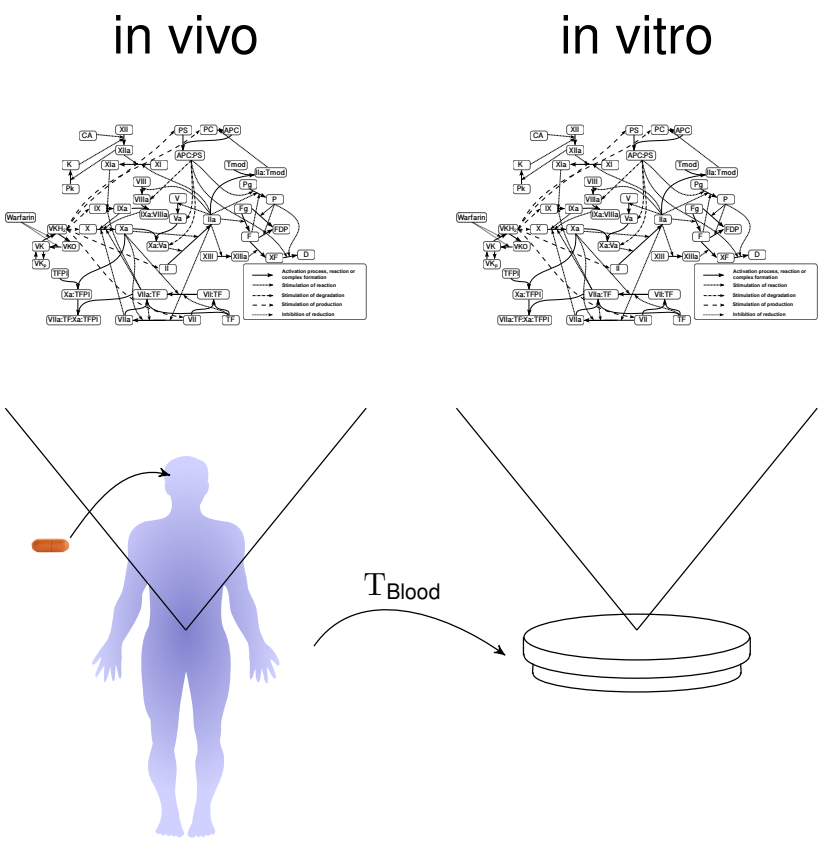


Figure 7: **Graphical representation of how warfarin therapy is simulated by the blood coagulation model [130].**

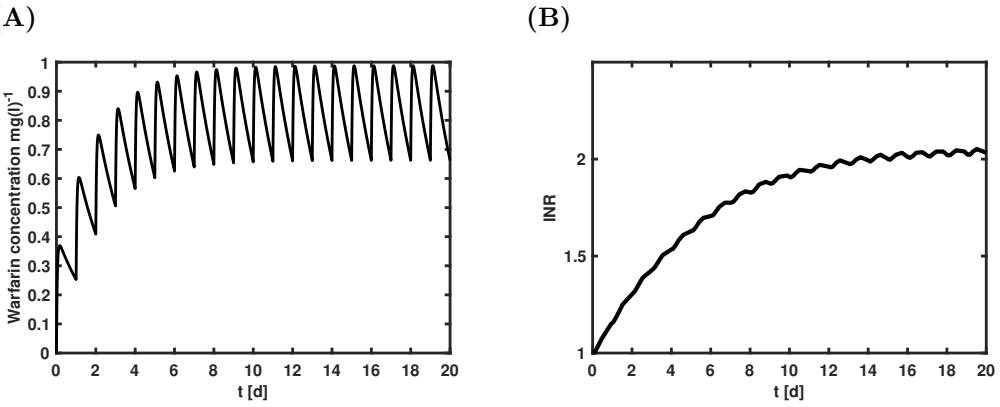


Figure 8: **Warfarin pharmacokinetics and pharmacodynamics for daily dosing of 4 mg (A)** Simulated time course of warfarin plasma concentration (daily dose 4mg). **(B)** Simulated time course of the INR (normalised prothrombin time) for warfarin dosing of 4 mg daily.

2.3.4. Rivaroxaban therapy - *in vitro*

We extended the comprehensive blood coagulation model (given in Figure 6) to include the effect of rivaroxaban on the blood coagulation. The pharmacokinetics of rivaroxaban was modelled as a one-compartment model (taken from Girgis et al. [36]). The interaction of rivaroxaban plasma concentration C_{Riva} with factors Xa and Xa:Va was implemented as in Burghaus et al. [12] choosing the same parameter values. The additional reactions are given by

$$\mathrm{Xa} + \mathrm{C_{Riva}} \underset{k_{\mathrm{Xa:C_{Riva}}}}{\overset{k_{\mathrm{Xa,C_{Riva}}}}{\rightleftharpoons}} \mathrm{Xa{:}C_{Riva}}\,,$$

$$\mathrm{Xa{:}Va} + \mathrm{C_{Riva}} \underset{k_{\mathrm{Xa:Va:C_{Riva}}}}{\overset{k_{\mathrm{Xa:Va,C_{Riva}}}}{\rightleftharpoons}} \mathrm{Xa{:}Va{:}C_{Riva}}\,.$$

In contrast to warfarin, rivaroxaban does not influence the blood coagulation system in its inactivated form, but only exhibits an effect when the system becomes activated by the intrinsic and/or the extrinsic pathway. Therefore, it is sufficient to consider the activation of the system including the extension with the pharmacokinetics of rivaroxaban (more specifically the *in vitro* assays - e.g. PT test) to assess the effect of rivaroxaban on the blood coagulation network.

2.4. Need for model order reduction of systems pharmacology models

The analysis of clinical trial data is of key importance for drug development. In drug discovery and preclinical development the drug properties, exposure, efficacy and drug-drug interactions are assessed and detailed knowledge about the underlying processes is available. By using and integrating many different data sources large-scale, mechanistic pharmacological models e.g. the EGFR system [55] (cf. section 2.2) and the blood coagulation network [130] (cf. section 2.3) can be developed, providing a comprehensive understanding of the underlying biological processes. The large-scale mechanistic models have successfully been employed to identify possible drug targets [111], investigate the action of drugs (or venoms) on the system [12, 40] or to study the effect of drug combinations [67]. However, due to the complexity of the interactions and the model size in terms of parameters and state variables these mechanistic models are not yet used for the population analysis of clinical data in late development of drug discovery and development. Translating the knowledge represented in the form of these detailed mechanistic models into the population analysis of clinical data is highly desirable, yet a still challenging task. For the analysis of population data of a clinical trial, empirical PD models are prevalently employed. Examples for the most frequently employed empirical models are given in section 2.1. For empirical modelling approaches, the model structure and the model parameters are derived from data and, thus, highly-dependent on the data quality (as described in section 2.1).

In the following section, the need for model reduction of systems pharmacology models is made apparent by illustrating the difficulties with the empirical modelling approach for two example drugs (e.g warfarin and rivaroxaban). The following empirical models will be used in section 5.1.3 as benchmarks to understand the resulting models obtained by our novel model reduction technique.

Warfarin was chosen due to the continuing efforts to individualise its dosing schedule. Warfarin is the most widely used anticoagulant [37, 61]. However, finding an optimal dosing strategy is challenging due to warfarin's narrow therapeutic window. Its large inter-individual variability often results in insufficient anticoagulation or increased bleeding risk [26]. Additionally, the dose-response relationship for warfarin has not been fully understood and, thus, remains to this day an important area of research [107, 65]. Knowing the factors that influence individual responses to warfarin therapy would help in tailoring the doses needed to maintain appropriate anticoagulation with fewer serious complications [60].

Rivaroxaban was chosen due to its novelty in clinics (approved in 2008 [87]) and predictable pharmacokinetics and pharmacodynamics. As discussed earlier, rivaroxaban inhibits one factor of the blood coagulation network directly, namely factor Xa. It is believed that rivaroxaban may ensure more consistent and predictable anticoagulation than warfarin [94]

Example 1: warfarin effect
In [53] a review of the different empirical PD models used to describe the response of warfarin is given. In [53] it is stated that the simplest system accounting for the changes in prothrombin time induced by warfarin is achieved by four independent models: (i) PK model: for the absorption, distribution and elimination of warfarin; (ii) PD model: for the effect of warfarin

on the synthesis of clotting factors collectively referred to as the prothrombin complex; (iii) physiological model: for the synthesis and degradation of the prothrombin complex; and (iv) a link model: for the relationship between the physiological model and the actually measured drug effect (prothrombin time) [53]. This highlights the complexity of an empirical PD model for warfarin. Of these different empirical PD models for warfarin, two of the most recent ones which are most frequently used in literature shall be discussed in detail in this section: the empirical pharmacodynamic models developed in [45, 44] (cf. Figure 9) and [90](cf. Figure 10).

The empirical pharmacodynamic model for warfarin effect in [45, 44] consists of two transit chains of different length and with different mean transition time. The INR is then calculated via a combination of baseline and a maximal effect depending exponentially on the last compartments of the two transit chains. The transit compartment chain approach was chosen to account for the time delay between drug concentration and an increase in INR. In [44] it is stated, that up to 3 parallel transit compartment chains were tested corresponding to the inhibition of the coagulation factors II, VII and X.

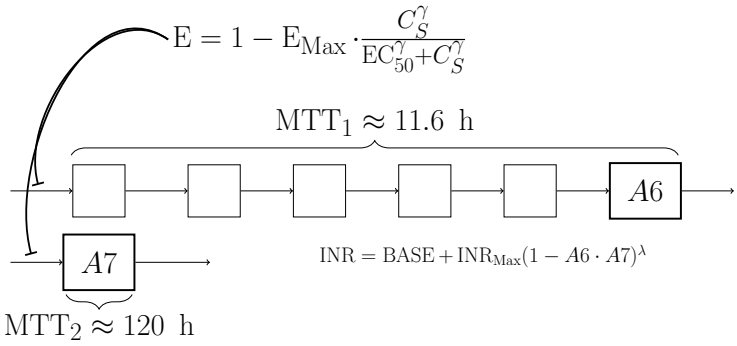


Figure 9: **Empirical pharmacodynamic model for warfarin by Hamberg et al.** (based on [45, 44]). Here C_S represents the S-warfarin concentration. The parameters: mean transition time (MTT), drug concentration that produces the half maximal inhibitory effect (EC_{50}), maximal inhibiting effect ($\mathrm{E_{Max}}$), baseline (BASE) of the international normalised ratio (INR) and exponent λ were taken from [45, 44].

The structure of the empirical pharmacodynamic model for warfarin in [90] differs from the model in [45, 44] and was developed based on the data of an Asian population (cf. Figure 10). In the model [90] the normal prothrombin concentration (NPT) was chosen as a biomarker for the coagulation activity and the synthesis of this biomarker was modelled to be inhibited by the warfarin concentration (similar to equ. (5)). The effect on the INR was again calculated as a change from baseline by a maximal effect $\mathrm{INR_{Max}}$ depending exponentially on the relative change in NPT.

The empirical pharmacodynamic models in [45, 44] (cf. Figure 9) and [90](cf. Figure 10) can actually be seen as a composition of these 4 models described in [53]. The Emax/Imax model in both models corresponds to the PD model, the transition chain model/the PD-1 model would then be the physiological model and the equation for the international normalised ratio (INR) equals the link model.

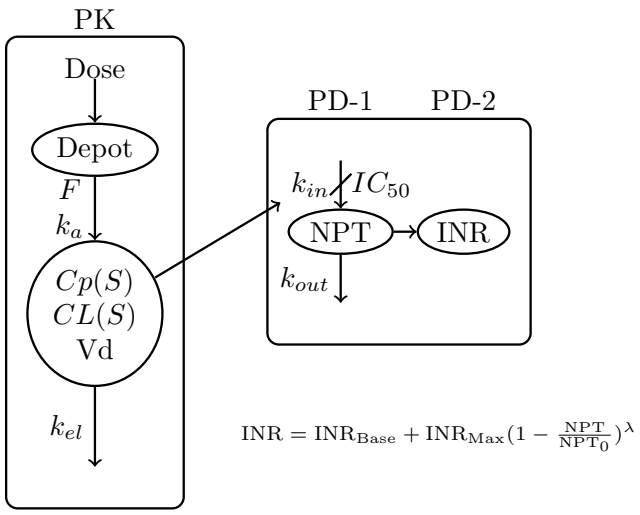


Figure 10: **Pharmacokinetic and empirical pharmacodynamic models of warfarin** (based on [90]). The pharmacokinetic model describes the time course of the plasma concentration of S-warfarin $Cp(S)$. The normal prothrombin (NPT) and international normalised ratio (INR) time course are described by the empirical pharmacodynamic model.

The empirical pharmacodynamic models for warfarin presented in this subsection can be categorised as indirect empirical models accounting for the time delay in the drug response. Both models were successfully employed for the analysis of clinical data. Due to the complex concentration-response relationship for warfarin, both models attempted a semi-mechanistic empirical PD model for warfarin. The chosen parameterisation in both models does not allow for a physiological interpretation.

Example 2: rivaroxaban effect

Rivaroxaban in contrast to warfarin only targets one factor within the coagulation cascade, Factor Xa [12]. The drug effect of rivaroxaban on the blood coagulation can be also be assessed by the PT test [36]. The PD relationship of rivaroxaban concentration C_{Riva} on prothrombin time prolongation was among others described by a so-called direct nonlinear empirical pharmacodynamic model [36]:

$$\mathrm{PT} = \mathrm{PT_{BASE}} + \mathrm{Slope} \cdot C_{\mathrm{Riva}}^{(1-\mathrm{Exponent}\cdot C_{\mathrm{Riva}})}. \tag{11}$$

In contrast, in [87] a linear relationship between rivaroxaban concentration and prothrombin time is reported.

Need for model order reduction of systems pharmacology models

All of the empirical models (given in Figure 9, 10 and in equ. (11)) were successfully employed for the population analysis of clinical data (including the statistical identification of covariates). There exist, however, various empirical models for the description of the same concentration-response relationship. This raises the questions 'How to identify an appropriate pharmacodynamic model?' and 'What model should be employed?'. As it was the case for warfarin, a simple concentration-response relationship is often not readily apparent, which makes the development of appropriate fit-for-purpose empirical PD models challeng-

ing. Furthermore, in general, empirical models can only be used to a very limited extent for extrapolation to e.g. outside the boundaries of the experimental data used for their development [83, 27].

In contrast, large-scale mechanistic models (e.g. the EGFR system [55](cf. section 2.2) and the blood coagulation network [130] (cf. section 2.3)) are typically designed to study the impact of diverse inputs or stimuli, often already incorporate important covariates or markers mechanistically and are more likely to be successfully used for extrapolation [83]. Additionally, as shown in section 2.3, the drug effect of various drugs acting on one particular system can be incorporated and described by the same mechanistic model.

In general, the two different modelling approaches presented for warfarin and rivaroxaban based on (i) the mechanistic blood coagulation model (section 2.3) and (ii) the empirical PD models (previous paragraphs) exist in parallel without cross-fertilisation between them. As described in this section, both approaches have their clinical relevance in the drug development process, however, could markedly benefit from a cross-link between them.

It is, therefore, highly desirable to develop a rational method to obtain suitable pharmacodynamic models from large-scale mechanistic pharmacodynamic models retaining the mechanistically incorporated covariates. In general, model reduction is a step towards this and can provide a link between mechanistic pharmacodynamic models and the classical drug effect models.

3. Relevant methods of model reduction for systems pharmacology/biology models from two different fields of application

The systems pharmacology/biology models of interest in this thesis are generally of the form:

$$\frac{\mathrm{d}x(t)}{\mathrm{d}t} = f\big(x(t), p\big), \quad x(0) = x_0 \tag{12}$$

where $x(t) \in \mathbb{R}^n$ is the vector of state variables at time $t \in [0, t_{\text{end}}]$, $p \in \mathbb{R}^d$ is the vector of parameters. The function $f : \mathbb{R}^n \times \mathbb{R}^d \to \mathbb{R}^n$ represents the systems pharmacology/biology model of interest, and is typically nonlinear. In this thesis the systems pharmacology/biology models are considered for a fixed vector of parameter values (cf. section 6 for discussion) and, thus, we drop p in the notation if not otherwise needed. Typically, systems pharmacology/biology models provide a detailed description of the underlying biological process reflecting the current state of knowledge. However, the complexity of interactions as well as the size of the models in terms of parameters and state variables render these models unsuitable for the estimation of individual parameters in a statistical setting due to identifiability issues. In general, model reduction is a step towards solving this identifiability issue.

The goal of model reduction techniques is to reduce large-scale complex models to 'suitable' reduced models. The requirements on the model reduction technique are set by the field of application as well as the desired application or task. Although model reduction and model order reduction are often used interchangeable [70], in this thesis, we will make a clear distinction between model reduction and model order reduction. In this thesis, model order reduction methods denote those methods that result in a reduction of the state space of a model, while model reduction will also refer to methods reducing the parameter space of a model. Furthermore, in this thesis, we focus on model order reduction techniques that apply a linear transformation to the system to obtain a reduced system.

All model order reduction techniques can be considered as a projection of the state space ($\subset \mathbb{R}^n$) to a state space of lower dimension ($\subset \mathbb{R}^{n_{\text{red}}}$ with $n_{\text{red}} < n$). Let us consider $T : \mathbb{R}^n \to \mathbb{R}^n$ a nonsingular transformation such that

$$Tx = x_{\text{trans}} = \begin{pmatrix} x_{\text{red}} \\ \hat{x} \end{pmatrix} \tag{13}$$

with $x_{\text{red}} \in \mathbb{R}^{n_{\text{red}}}$. Let us assume that the transformation T can be written as

$$T = \begin{pmatrix} W^\top \\ T_2^\top \end{pmatrix}$$

and its inverse is given by

$$T^{-1} = \begin{pmatrix} V & T_1 \end{pmatrix},$$

where $V, W \in \mathbb{R}^{n \times n_{\text{red}}}$ and $T_1, T_2 \in \mathbb{R}^{n \times (n - n_{\text{red}})}$. The transformed system is then given by

$$\dot{x}_{\text{trans}}(t) = T f(T^{-1} x_{\text{trans}}(t))$$

and, thus, the ordinary differential equation for x_{red} can be written as

$$\dot{x}_{\text{red}}(t) = W^\top f(V x_{\text{red}}(t) + T_1 \hat{x}) \,. \tag{14}$$

Note that this equation is exact. An approximation of the reduced system given in eq. (14) can then be obtained by neglecting $\hat{x}$, replacing $\dot{\hat{x}}$ with algebraic equations or combining $\hat{x}$ with other state variables. In the case of neglecting $\hat{x}$ the projection to the reduced state space is then obtained by $P = VW^\top$, since $W^\top V = I_{n_{\text{red}}}$. The projection is called Petrov-Galerkin projection if $V \neq W$ and Galerkin projection otherwise. The approximated reduced system $\tilde{x}_{\text{red}}$ is then given by

$$\begin{aligned} \dot{\tilde{x}}_{\text{red}}(t) &= W^\top f(V \tilde{x}_{\text{red}}(t)) \\ \Leftrightarrow \dot{\tilde{x}}_{\text{red}}(t) &= W^\top f(V (W^\top V)^{-1} \tilde{x}_{\text{red}}(t)) \,, \end{aligned} \tag{15}$$

In order to quantify the accuracy of the reduced model in eq. (15) in approximating the original model in eq. (14) an approximation error ε with an appropriate norm needs to be defined. The chosen approximation error ε and chosen norm have implications for the projection and, thus, the construction of the projection varies between the model order reduction techniques.

A large variety of model order reduction methods is available in the different fields of application [110, 5, 92]. However, not all of the methods fulfil the specific requirements inherent in the field of systems pharmacology/biology. In particular, the most important yet most challenging aspect to take into account is that the model parameters and state variables correspond to physical processes, e.g. production rate of a protein (cf. eq. (8)), and entities, e.g concentrations of proteins (cf. eq. (7)). In order to analyse patient data and understand variability in specific patient parameters (so-called inter-individual variability/ between patient variability), the parameters and state variables of the reduced model must allow for a mechanistic interpretation in terms of quantities of the original model. This leads to a restriction of the set of allowed transformation matrices in the context of model reduction. The three established methods of model reduction in the field of systems pharmacology/biology are lumping, time scale analysis and sensitivity analysis.

In the following sections, we will focus on the most common techniques in the context of systems pharmacology/biology models. The topic of model order reduction is an extensive field and there exist various variants of each method. Therefore, this section focuses on the general concepts and does not attempt to give a full review of all adapted model reduction techniques to systems pharmacology/biology models. For a survey of the current methods and trends in model reduction for large-scale biological systems, we refer to [118]. In this section, the applicability, advantages and drawbacks of the different concepts are discussed after all the different methods have been introduced.

3.1. Systems pharmacology/biology: model reduction techniques

In systems pharmacology/biology, mainly three model reduction approaches are used [92]: lumping [1, 11, 77, 99], time scale analysis [73, 138] and sensitivity analysis [103, 95, 139]. These three approaches shall be discussed in detail in this section.

3.1.1. Lumping

In the field of systems pharmacology/biology, the reduction of a system by a linear transformation is referred to as lumping. To allow for consistency with the literature on lumping, the lumping matrix is denoted by L in this subsection and the relationship between the projection matrix $V, W^\top$ and L is given after lumping has been introduced.

The first analysis of lumpability of monomolecular reaction systems was performed in 1969 by [133]. Since then lumping has become the most widely used technique in this field [92]. Two different variants of lumping are distinguished: proper and improper lumping.

Definition 3.1 (lumping matrix)
The matrix $L \in \mathbb{R}^{n_{red} \times n}$ is called a lumping matrix if $\sum_{i=1}^{n_{red}} L_{ij} = 1$ for all $1 \leq j \leq n$ and $L_{ij} \in [0, 1]$ for all $1 \leq j \leq n$ and $1 \leq i \leq n_{red}$.

Definition 3.2 (proper lumping matrix)
The lumping matrix $L \in \mathbb{R}^{n_{red} \times n}$ is called proper if $L_{ij} \in \{0, 1\}$ for all $1 \leq j \leq n$ and $1 \leq i \leq n_{red}$.

A lumping matrix that does not fulfil the definition 3.2 is called improper. In proper lumping the original state vector is partitioned and each partition is then reduced to a single dynamical state variable in the reduced system [118]. Note that the lumping matrix which is given in definition 3.1 or 3.2 is singular.

The reduced state variables are given by

$$x_{\mathrm{red}}(t) = Lx(t)\,.$$

By applying the Galerkin projection, the dynamics of the reduced system can be written as

$$\frac{\mathrm{d}x_{\mathrm{red}}(t)}{\mathrm{d}t} = Lf\big(\hat{L}x_{\mathrm{red}}(t)\big), \quad x_{\mathrm{red}}(0) = Lx_0 \tag{16}$$

with $\hat{L}$ a generalised inverse such that $L\hat{L} = I_{n_{\mathrm{red}}}$. There are infinitely many generalised inverses fulfilling this requirement [76, 31]. The choice of inverse $\hat{L}$ has an influence on the approximation error in the case of approximate lumping [117], as will be described in the subsequent paragraphs. The approximation of the original state variables is given by

$$x \approx \hat{L}x_{\mathrm{red}}\,.$$

In the general framework as introduced in eq. (13), (14) and (15), L is equal to $W^\top$ and $\hat{L} = V(W^\top V)^{-1}$.

There exist two ways to compute the differential equations for the reduced system: (i) $\dot{x}_{\text{red}}(t) = f_{\text{red}}(Lx(t))$ and (ii) $\dot{x}_{\text{red}}(t) = Lf(x(t))$ assuming that a n_{red} dimensional vector-function f_{red} exists. A system is called exactly lumpable by a matrix L if $f_{\text{red}}(Lx(t)) = Lf(x(t))$ [133]. In the case of an exactly lumpable system the approximation error $\varepsilon(x(t)) = f_{\text{red}}(Lx(t)) - Lf(x(t))$ is zero and, thus, $\varepsilon(x(t))$ is independent of the choice of the inverse of L. However, even if a system is exactly lumpable, the requirement, that a specific easily measurable state variable should be kept unlumped, will result in approximate lumping of the system [75]. If a system is not exactly lumpable, the lumping matrices are chosen such that the approximation error $\varepsilon(x(t))$ is minimised. The main objective is, therefore, to find an appropriate lumping scheme (L and $\hat{L}$) that will minimise the approximation error ε [76].

In the simple case of a linear system it was proven that for a chosen L a unique solution for $\hat{L}$ can be found by projection of the approximation error ε onto the subspace spanned by L such that $\varepsilon L^\top = 0$ or onto the subspace spanned by LX_{ss} such that $\varepsilon(X_{\text{ss}})L^\top = 0$ with $X_{\text{ss}} = \text{diag}\{x_{\text{ss}}\}$, where x_{ss} denotes the steady state of the system assuming that the system has a unique steady state [76]. The first projection will lead to the Moore-Penrose inverse $L^+ = L^\top(LL^\top)^{-1}$, which was proven to minimise the squared Frobenius norm of the approximation error ε (i.e. $||\varepsilon(x(t))||_F^2$) along the solution trajectory [76]. The second approach gives a generalised inverse $\hat{L} = X_{\text{ss}}L^\top(LX_{\text{ss}}L^\top)^{-1}$ ensuring that the steady state x_{ss} of the system is retained [117]. For a linear system $\dot{x} = Ax$ the approximation error

$$\varepsilon(X_{\text{ss}}) = \underbrace{LAX_{\text{ss}}}_{=0} - \underbrace{LAX_{\text{ss}}}_{=0} L^\top(LX_{\text{ss}}L^\top)^{-1}X_{\text{ss}} = 0$$

vanishes in the steady state. Clearly this choice of inverse only makes sense if the steady state of the system is distinct from zero.

Finding an optimal lumping scheme is a challenging task, especially for large-scale nonlinear systems, since it is practically impossible to test all possible lumped models due to the curse of dimensionality. As a consequence, heuristics are used to find a nearly optimal lumping scheme L and $\hat{L}$ with an optimality criterion e.g. minimising the squared difference of a subset of state variables of interest of the reduced and original system. In [1] a recursive greedy search strategy is proposed: at each step all possible lumps of two state variables are computed, among all these possible lumps the pair of state variables is lumped that results in the lowest approximation error ε between the state variables of interest of the reduced and original system. This process is continued until no further reduction is possible within the desired accuracy of the approximation. This algorithm uses the Moore-Penrose inverse, but can easily be adapted to use the generalised inverse retaining the steady state. The proper lumping method is one of the model reduction techniques of choice in the field of systems pharmacology/biology since this special case of transformation matrix allows for some conservation of physiological interpretability in the reduced system's structure [1].

3.1.2. Time scale separation

The motivation for time scale separation is that the dynamic behaviour of pharmacology/biology reaction systems often contains multiple time scales, i.e. after an initial transient period, some fast reactions can be considered almost instantaneously relative to the remaining reactions [92]. This provides the basis for the quasi-steady-state approximation (QSSA), which is mathematically justified by singular perturbation theory [113].

The first step of time scale analysis is to find a partition of the system into slow and fast state variables. The second step is to approximate the dynamics of the fast state variables by their steady state as a function of the slow state variables. The equation of the fast state variables as a function of the slow state variables can then be substituted into the differential equations for the slow state variables to obtain a reduced system of ODEs in terms of only the slow state variables. In the derivation of the method, we closely follow [98]. In the following paragraph, it is assumed that such a partition of the system into slow and fast state variables was found. Let x_{slow} denote the state variables which are evolving slowly compared to x_{fast} and x_{fast} denote the state variables evolving on a fast time scale, then the ODE system in eq. (12) can be partitioned and is given by

$$\begin{aligned}\frac{\mathrm{d}x_{\text{slow}}(t)}{\mathrm{d}t} &= f_1(x_{\text{slow}}(t), x_{\text{fast}}(t))\,,\\ \frac{\mathrm{d}x_{\text{fast}}(t)}{\mathrm{d}t} &= f_2(x_{\text{slow}}(t), x_{\text{fast}}(t))\end{aligned}$$

with $x(t_0) = (x_{\text{slow}}(0), x_{\text{fast}}(0))$. The dynamics of the fast state variables is then approximated by their steady state (quasi-steady state assumption)

$$f_2(x_{\text{slow}}(t), x_{\text{fast}}(t)) = 0.$$

A system reduced based on the quasi-steady state assumption results in the following system of differential algebraic equations (DAE)

$$\begin{aligned}\frac{\mathrm{d}x_{\text{slow}}(t)}{\mathrm{d}t} &= f_1(x_{\text{slow}}(t), x_{\text{fast}}(t))\,,\\ 0 &= f_2(x_{\text{slow}}(t), x_{\text{fast}}(t)).\end{aligned}$$

In general one hopes to find a function $x_{\text{fast}}(t) = g(x_{\text{slow}}(t))$ with $(x_{\text{slow}}(t), g(x_{\text{slow}}(t)))$ a root of $f_2(x_{\text{slow}}(t), x_{\text{fast}}(t)) = 0$.

The main drawback of this technique is that often it is not known *a priori* if a slow/fast partitioning of the system exists and how to obtain it. In order to obtain a good approximation of the original system by the reduced system with the QSSA a clear separation in the timescales is required. The difficulty of this technique is, thus, to find the partition of the system into slow and fast state variables. While the partition might not be apparent in the original state space, a transformation of the state space can lead to a clear time scale separation. Existing automated techniques such as Intrinsic Low-Dimensional Manifold (ILDM) [127] and computational singular perturbation (CSP) [72, 73, 138] transform either state or tangent space

of the state space (e.g. f) to find an appropriate partition into slow and fast dynamics. However, after the coordinate transformation has been applied, the physiological interpretation of the original state variables might be lost.

The ILDM technique provides a numerically stable means for obtaining an eigenvalue decomposition of the system. Therefore, we have chosen the ILDM technique to be used in section 5.3 in the comparison of model order reduction techniques on an example of pharmacological interest. We shortly describe the ILDM method closely following the derivation in [127].

The ILDM method automatically partitions the system into fast and slow dynamics by first applying a Schur decomposition to the Jacobian of the right-hand side at a particular state. The Schur decomposition applied to a square matrix gives an upper triangular matrix and some orthonormal basis. If it is applied to the Jacobian of the dynamical system in eq. (12) at a particular state e.g. steady state x_{ss} of the system, one obtains

$$Q^T J(x_{\mathrm{ss}}) Q = \begin{pmatrix} S_{\mathrm{slow}} & S_{\mathrm{coup}} \\ 0 & S_{\mathrm{fast}} \end{pmatrix} .$$

The upper triangular matrices S_{slow} and S_{fast} correspond to the slow and fast eigenvalues of the Jacobian respectively. The matrix S_{coup} gives the coupling between the fast and the slow dynamics of the system. The ILDM method then aims to find a transformation matrix $Z \in \mathbb{R}^{n_{\mathrm{slow}} \times n_{\mathrm{fast}}}$ of the system that decouples the fast and slow dynamics. The matrix Z can be obtained by solving the Sylvester equation

$$S_{\mathrm{slow}} Z - Z S_{\mathrm{fast}} + S_{\mathrm{coup}} = 0 .$$

The full transformation matrix T decoupling the system into slow and fast dynamics is given by

$$T = Q \begin{pmatrix} I & Z \\ 0 & I \end{pmatrix} .$$

After the transformation has been applied, QSSA is used to transform the system into a differential algebraic equation (DAE) system.

A classical example for this technique is the use of QSSA to reduce a simple enzym-substrate reaction system that leads to the Michaelis-Menten kinetic [84, 10]. For low-dimensional model systems, time scale separation techniques are often successfully employed, therefore, time scale separation methods are one of the most commonly employed model reduction techniques in the field of pharmacology/biology.

3.1.3. Sensitivity analysis

Sensitivity analysis plays an important role in the analysis of the dynamics of systems pharmacology/biology models, furthermore, it has proven to be useful in the context of model simplification [139]. In sensitivity analysis, the impact of parameter changes (incl. initial

conditions) on state variables as a function of time is assessed [139]. Typically a reduced model is obtained by eliminating those reactions, whose parameters have little or no effect on the model predictions [79]. State variables found to be least sensitive with respect to their initial condition are usually fixed to a constant value (e.g. steady state) [116].

The parameter sensitivity matrix $\mathcal{S}^p(t)$ is defined as

$$\mathcal{S}^p(t) := \begin{pmatrix} \frac{\partial x(t)}{\partial p_1} & \dots & \frac{\partial x(t)}{\partial p_d} \end{pmatrix} = \begin{pmatrix} \frac{\partial x_1(t)}{\partial p_1} & \dots & \frac{\partial x_1(t)}{\partial p_d} \\ \vdots & \vdots & \vdots \\ \frac{\partial x_n(t)}{\partial p_1} & \dots & \frac{\partial x_n(t)}{\partial p_d} \end{pmatrix} \in \mathbb{R}^{n \times d}$$

where x_k denotes the kth state variable and p_i denotes the ith parameter.

The parameter sensitivity matrix $\mathcal{S}^p(t)$ is calculated by the linear variational equation:

$$\frac{d\mathcal{S}^p(t)}{dt} = f_x(x(t), p)\mathcal{S}^p(t) + f_p(x(t), p)\,, \tag{17}$$

where f_x and f_p are the Jacobian matrices of the system with respect to states and parameters. The initial condition for the sensitivity matrix $\mathcal{S}^p$ is given by

$$\mathcal{S}^p(t_0) = \begin{pmatrix} \frac{\partial x_0}{\partial p_1} & \dots & \frac{\partial x_0}{\partial p_d} \end{pmatrix}$$

and equals zero, if the initial conditions do not depend on the parameters. Sensitivity coefficients are usually normalised to compensate for different scales and to eliminate units [95]. The normalised sensitivity coefficient $\hat{\mathcal{S}}^p_{i,j}$ is defined as

$$\hat{\mathcal{S}}^p_{i,j} = \frac{\partial x_i(t)}{\partial p_j} \cdot \frac{p_j}{x_i(t)} = \frac{\partial \log(x_i)}{\partial \log(p_j)}. \tag{18}$$

The sensitivity matrix of the state variables with respect to the initial condition is defined by

$$\mathcal{S}^{x_0}(t) = \frac{\partial x(t)}{\partial x_0}.$$

The initial condition sensitivity matrix $\mathcal{S}^{x_0}(t)$ is obtained by solving the differential equation

$$\frac{d\mathcal{S}^{x_0}(t)}{dt} = f_x(x(t), p)\mathcal{S}^{x_0}(t)$$

with initial condition $\mathcal{S}^{x_0}(t_0) = I_n$. The matrix $\mathcal{S}^{x_0}(t) = \frac{\partial x(t)}{\partial x_0}$ is called the Wronksi matrix.

The above given sensitivity matrices are considered the local sensitivities. The importance of reactions is then assessed by generating a ranking based on an appropriate summary statistics of the parameters sensitivities over time, such as the infinity norm, the Fisher information matrix, the time integral or the parameter sensitivity of the system at a particular state, e.g. steady state [95]. However, sensitivity analysis showing the effect of small parameter changes can not solely be used to classify the importance of reactions to the system of interest [123, 126]. In particular, a reduction based on the importance obtained only by

sensitivity analysis will result in reduced models with low approximation quality [123, 126]. Various attempts of model reduction techniques based on sensitivity analysis have tried to overcome this issue, usually by using additional features of the model such as flux analysis [79] and additional techniques such as principal component analysis of $\left(\hat{\mathcal{S}}^p\right)^\top \hat{\mathcal{S}}^p$ [25]. However, in [25] it is noted that no unique criterion for the elimination of parameters could be identified since even parameters with low importance caused significant model prediction discrepancies. In [25] the normalised parameter sensitivities in eq. (18) were used, the importance was then defined by the sum of squares of the normalised sensitivities for predefined discrete time point of a time interval of interest. The model reduction in [25] was, however, based on the eigenvalue decomposition of $\left(\hat{\mathcal{S}}^p\right)^\top \hat{\mathcal{S}}^p$. Therefore, there still does not exist an established model reduction technique based on the sensitivities of the system. In order to use sensitivity analysis as one of the methods that shall be compared for an example of pharmacological interest in section 5.3 we have not deleted parameters solely based on their sensitivity measure but additionally required the relative approximation error ε between the original and reduced system to be below a user-defined threshold.

We would like to note that for signalling cascades in biological systems (e.g. the EGFR signalling pathway in section 2.2), signals are often propagated through a series of activation steps. Activated forms temporarily rise from zero or very low concentrations to their maximum, while inactive forms almost vanish. In such a situation, the normalisation in eq. (18) is highly problematic, since $x_k(t_0) = 0$ for active forms and $x_k(t) \approx 0$ at later times for inactive forms, resulting in ill-defined coefficients and numerical problems, respectively. It is, therefore, highly desirable to extend the consideration of sensitivity coefficients of the state variables beyond the initial time point.

3.2. Control theory: gramian based model order reduction

In control theory, one aims to predict the effect that a particular action will have on a physical system [74, p. 5-2]. More precisely, how will the response, $y(t)$, of a specified system evolve over time $t \in [t_0, t_1]$ after an arbitrary input $u(t)$ has been applied over the same time interval [74, p. 5-2]. In order to analyse the system with respect to a particular input-output relationship, the system equation in eq. (12) is changed such that the input and output are explicitly stated. Following [109] the system of interest is, thus, written as

$$\begin{aligned} \dot{x}(t) &= f(x(t), u(t)), \quad x(t_0) = x_0\,, \\ y(t) &= h(x(t), u(t)) \end{aligned} \tag{19}$$

where

$x : [t_0, t_1] \mapsto \mathbb{R}^n$	is the vector function of state variables of the system,
$u : [t_0, t_1] \mapsto \mathbb{R}^s$	is the vector function of control signals which affect the system to achieve a desired behaviour,
$y : [t_0, t_1] \mapsto \mathbb{R}^q$	is the vector function of output signals which serve to assess whether the control achieved the desired goal,

$x_0 \in \mathbb{R}^n$ initial value,

f, h set of differential and algebraic equations, respectively, which describe the relationship between input, state and output variables.

In control theory, the system in eq. (19) is used to understand whether an optimal input that achieves a specific response exists, whether one can reconstruct the solution trajectory from partially observed measurements and to understand the input-output behaviour of the system. To characterise the input-output behaviour of a system the concepts of controllability and observability are introduced.

Model reduction techniques in the field of control theory focus on approximating the output of the system in eq. (19) as good as possible, e.g. the difference between the output of the original system and reduced system is small in an appropriate norm. Under the assumption that an appropriate state space transformation T was found, the reduced system is then given by

$$\begin{aligned} \dot{x}_{\text{red}}(t) &= f_{\text{red}}(x_{\text{red}}(t), u(t)), \quad x_{\text{red}}(t_0) = x_{\text{red},0}\,, \\ y_{\text{red}} &= h_{\text{red}}(x_{\text{red}}(t), u(t))\,, \end{aligned} \tag{20}$$

with $f_{\text{red}} : \mathbb{R}^{n_{\text{red}}} \times \mathbb{R}^s \to \mathbb{R}^{n_{\text{red}}}, h_{\text{red}} : \mathbb{R}^{n_{\text{red}}} \times \mathbb{R}^s \to \mathbb{R}^q$ and $y_{\text{red}}(t) \approx y(t) \in \mathbb{R}^q, n_{\text{red}} \ll n$.

In order to find an appropriate state space transformation T, in the gramian-based model reduction techniques it is investigated if the system in eq. (19) is controllable and observable.

In this section we make the dependence of the solution trajectory of the differential equation system (defined by (19)) on the initial value x_0 and on the input function u explicit, by writing $x(t)$ as $x(t; x_0, u(t))$ and $y(t) = h(x(t; x_0, u(t)), u(t)) = y(t; x_0, u(t))$. This notation is chosen to make the differences in the definitions in section 3.2.3 clear and to make the dependence of the solution on the input readily apparent.

Definition 3.3 (controllable)
The system in eq. (19) *is controllable on the time interval* $[t_0, t_1]$ *if there exists for all states* $x_0, x_1 \in \mathbb{R}^n$ *a control function* u*, such that the solution satisfies:* $x(t_0; x_0, u(t_0)) = x_0$ *and* $x(t_1; x_0, u(t_1)) = x_1$*.*

Definition 3.4 (observable)
The system in eq. (19) *is observable on the time interval* $[t_0, t_1]$ *if it is possible to uniquely determine the initial state* $x_0 \in \mathbb{R}^n$ *by measuring the output function* $y(t; x_0, u(t))$ *on* $[t_0, t_1]$ *given the input function* u *on* $[t_0, t_1]$*.*

These two properties of a system given in definition 3.3 and 3.4 play a central role for the gramian-based model reduction in control theory. These properties are closely related to the concept of gramians, which will be introduced in the next subsection. More specifically, whether a system is controllable/observable can be determined by the gramians. If both gramians are positive definite the linear system in eq. (21) or in eq. (34) is fully controllable and observable. In order to understand the concepts needed for nonlinear systems (e.g. sys-

tems pharmacology/biology model), we first introduce the analytical gramians for linear systems and subsequently for nonlinear systems. This allows us to show under which conditions on the linear systems the gramians for nonlinear systems are identical to the gramians for linear systems.

3.2.1. Overview of analytical gramians for linear systems

In this section, we will first introduce the analytical controllability and observability gramians for linear systems and in section 3.2.3 we will elucidate how this concept of controllability and observability gramians can be extended for nonlinear systems. In general, a gramian or gram matrix is a special square matrix, that is given by the pairwise scalar product of a set of vectors. All gram matrices are non-negative definite and symmetric. The controllability and observability gramians are fundamental for the most popular model order reduction technique in control theory called balanced truncation, which will be introduced at the end of this section.

A linear time invariant (LTI) system is given by:

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t), \quad x(t_0) = x_0\,, \\ y(t) &= Cx(t) \end{aligned} \tag{21}$$

with u the input and y the output function, $A \in \mathbb{R}^{n\times n}, B \in \mathbb{R}^{n\times s}$ and $C \in \mathbb{R}^{q\times n}$. Thus, we have that $f(x(t), u(t)) = Ax(t) + Bu(t)$ and $h(x(t), u(t)) = Cx(t)$ in eq. (19).

The solution of the system (21) is given by

$$\begin{aligned} x(t; x_0, u(t)) &= \Phi^{t,t_0} x_0 + \int_{t_0}^{t} \Phi^{t,s} B u(s) \mathrm{d}s \\ &= e^{A(t-t_0)} x_0 + \int_{t_0}^{t} e^{A(t-s)} B u(s) \mathrm{d}s\,, \\ y(t; x_0, u(t)) &= C\left(\Phi^{t,t_0} x_0 + \int_{t_0}^{t} \Phi^{t,s} B u(s) \mathrm{d}s\right) \end{aligned}$$

with Φ^{t,t_0} the state transition matrix, that maps the initial state at time t_0 to the solution $x(t; x_0, u(t))$.

For the LTI system given in eq. (21) the input-to-state map $\Psi_{u\to x}(t)$ is defined as $\Psi_{u\to x}(t) = e^{At}B$ and the state-to-output map $\Psi_{x\to y}(t)$ is defined as $\Psi_{x\to y}(t) = Ce^{At}$ [5, p. 79].

Definition 3.5 (unit/delta impulse [74, p.5-8ff])
Let $f(t)$ be any function that is continuous on the interval $-\varepsilon < t < \varepsilon$ for every $\varepsilon > 0$. Then the unit/delta impulse $\delta(t)$ satisfies

$$f(0) = \int_{-\infty}^{\infty} f(\tau)\delta(\tau)\mathrm{d}\tau\,.$$

If the input of the system is a unit/delta impulse, then the resulting state is $\Psi_{u\to x}(t)$ [5, p. 79].

When applying a model reduction technique in the field of control theory, it is important to understand how the state space transformation T in eq. (13) changes the input-output behaviour of an LTI system in eq. (21). In particular, one is interested in functions/ quantities with respect to the input-output behaviour that do not change under a state space transformation T. In the following paragraph, such a function is introduced for an LTI system in eq. (21).

The Laplace transform (also known as the frequency domain method [74, p.1-2]) of a function is defined by $\mathcal{L} : f(t) \mapsto F(s) := \mathcal{L}[f(t)] = \int_0^\infty e^{-st} f(t)\mathrm{d}t,\ s \in \mathbb{C}$. The application of the Laplace transformation to the state and output equations is often beneficial since it allows to use the frequency domain in system and control theory [74, p.9-26ff]. By applying the Laplace transformation to the LTI system in eq. (21) under the assumption that $x_0 = 0$, the transfer function G of the system is obtained. The transfer function G defines the relationship between the Laplace transform of the input and the Laplace transform of the output $Y(s) = G(s)U(s)$, and is given by

$$G(s) = C^\top (sI - A)^{-1} B\,. \tag{22}$$

In the frequency domain analysis $G(s)$ is evaluated for $s = i\omega$, where $\omega \in [0, \infty]$ has the physical interpretation of a frequency and the input is considered as a signal with frequency ω [74, p.9-26ff]. Additionally, a state space transformation T as given in eq. (13) of the system does not change the input-output behaviour of the system, such that the transfer function of the transformed system is equal to the transfer function of the original system. Due to these properties, the goal of many model reduction techniques is to achieve a good approximation between the transfer functions of the original and reduced system.

Let us now introduce the analytical finite time gramians that will subsequently be linked to the Definition 3.3 and 3.4. These gramians will then be used in section 3.2.2 to find an appropriate state space transformation T that achieves a good approximation between the transfer function of the original and reduced system.

Definition 3.6 (finite time controllability gramian [5, p.68])
For an LTI system eq. (21) *the $n \times n$ finite time controllability gramian $\mathcal{C}(0,t)$ is defined by*

$$\mathcal{C}(0,t) := \int_0^t e^{A(t-\tau)} B B^\top e^{A^\top (t-\tau)} \mathrm{d}\tau. \tag{23}$$

Definition 3.7 (finite time observability gramian [5, p.76])
For an LTI system eq. (21) *the $n \times n$ finite time observability gramian $\mathcal{O}(0,t)$ is defined by*

$$\mathcal{O}(0,t) := \int_0^t e^{A^\top \tau} C^\top C e^{A\tau} \mathrm{d}\tau. \tag{24}$$

Both gramians are positive semidefinite by definition [85, 5]. The link to the controllability/ observability of the system is given by the following corollaries.

Corollary 3.8 ([5, p.69])
The LTI system eq. (21) *is controllable if and only if* $\mathcal{C}(0,t)$ *is positive definite for some* $t > 0$.

Corollary 3.9 ([5, p.78])
The LTI system eq. (21) *is observable if and only if* $\mathcal{O}(0,t)$ *is positive definite for some* $t > 0$.

For the proof of these corollaries we refer to [5, p.69]. The unobservable states are exactly in the kernel of $\mathcal{O}(0,t)$ and the system contains uncontrollable states if $\mathcal{C}(0,t)$ does not have full rank.

For the limit case of $t \to \infty$ for asymptotically stable LTI systems, $\mathcal{C}(0,\infty)$ and $\mathcal{O}(0,\infty)$ are well-defined and subsequently always denoted by $\mathcal{C}(\infty), \mathcal{O}(\infty)$. For unstable LTI systems the infinite time gramians $\mathcal{C}(\infty)$ and $\mathcal{O}(\infty)$ are not defined.

Another important gramian for model reduction is the infinite time cross gramian [50].

Definition 3.10 (infinite time cross gramian [32])
For an asymptotically stable LTI system in eq. (21) *with* $q = s$ *the* $n \times n$ *infinite time cross gramian* $\mathcal{X}(\infty)$ *is defined by*

$$\mathcal{X}(\infty) := \mathcal{X}(0,\infty) = \int_0^\infty e^{At} B C e^{At} \mathrm{d}t\,.$$

The infinite time cross gramian carries information pertaining to both controllability and observability of the LTI system in eq. (21) [32]. For a symmetric and stable LTI system in eq. (21) with $q = s$ the cross gramian is the product of infinite time observability and controllability gramian $\mathcal{X}(\infty) = \mathcal{C}(\infty)\mathcal{O}(\infty)$ [50]. Thus, the infinite time cross gramian is a quantity that allows to investigate both controllability and observability of a system at the same time [32]. Additionally, it will be shown in section 4.3 how the novel input-response indices are related to the cross gramian.

In [7] a simpler way of calculating the gramians than computing the analytical expression given in eq. (23) and (24) is provided. For an asymptotically stable LTI system (fully controllable and observable) the controllability and observability gramian given in eq. (23) and (24) for $t \to \infty$ are the unique symmetric positive definite solution of the following algebraic matrix equations, also known as Lyapunov equations [7]

$$A\mathcal{C}(\infty) + \mathcal{C}(\infty)A^\top + BB^\top = 0, \tag{25}$$

$$A^\top\mathcal{O}(\infty) + \mathcal{O}(\infty)A + C^\top C = 0. \tag{26}$$

That eqs. (25) and (26) hold for $\mathcal{C}(\infty), \mathcal{O}(\infty)$, can be easily verified by inserting the expression

e.g. eq. (24) in eq. (26):

$$\begin{aligned} A^\top \mathcal{O}(\infty) + \mathcal{O}(\infty)A + C^\top C &= A^\top \int_0^\infty e^{A^\top \tau} C^\top C e^{A\tau} \mathrm{d}\tau + \int_0^\infty e^{A^\top \tau} C^\top C e^{A\tau} \mathrm{d}\tau A + C^\top C, \\ &= \int_0^\infty \underbrace{A^\top e^{A^\top \tau} C^\top C e^{A\tau} + e^{A^\top \tau} C^\top C e^{A\tau} A}_{\frac{\mathrm{d}}{\mathrm{d}\tau} e^{A^\top \tau} C^\top C e^{A\tau}} \mathrm{d}\tau + C^\top C, \\ &= \lim_{\tau \to \infty} e^{A^\top \tau} C^\top C e^{A\tau} - \underbrace{e^{A^\top 0}}_{I_n} C^\top C \underbrace{e^{A0}}_{I_n} + C^\top C. \end{aligned}$$

For asymptotically stable LTI systems, it holds that $\lim_{\tau \to \infty} e^{A^\top \tau} = 0$, such that one obtains

$$A^\top \mathcal{O}(\infty) + \mathcal{O}(\infty)A + C^\top C = -C^\top C + C^\top C = 0.$$

In the following paragraph, it is discussed under which conditions the infinite time gramians provide a good approximation for the finite time gramians [128]. By considering the time interval $[0, t]$ we get the finite time gramians, which satisfy the following functional equation (due to linearity of the integral and properties of the state transition matrix)

$$\begin{aligned} \mathcal{C}(\infty) &= \mathcal{C}(0,t) + e^{At} \mathcal{C}(\infty) e^{A^\top t}, \\ \mathcal{O}(\infty) &= \mathcal{O}(0,t) + e^{At} \mathcal{O}(\infty) e^{A^\top t}. \end{aligned} \tag{27}$$

The infinite time gramians can be used to approximate the finite time gramians. The approximation quality can be obtained by rearranging eq. (27) and taking a matrix norm. Then it is clear that the following error estimate between the infinite time gramian and finite time gramian holds [128]

$$\begin{aligned} ||\mathcal{C}(\infty) - \mathcal{C}(0,t)|| &\leq ||e^{At}||^2 ||\mathcal{C}(\infty)||, \\ ||\mathcal{O}(\infty) - \mathcal{O}(0,t)|| &\leq ||e^{At}||^2 ||\mathcal{O}(\infty)||. \end{aligned} \tag{28}$$

The finite time gramian is, thus, a good approximation for the infinite time gramian if the system given in eq. (21) is asymptotically stable and t is chosen large enough (twice the largest characteristic time in the system [128]) .

No realistic system can be considered on an infinite time horizon, however, under asymptotic stability of the system given in eq. (21) the infinite time gramians provide a good approximation of the finite time (e.g $[0, t]$) or time-limited (e.g. $[t_0, t_1]$) gramians that actually describe any physical system [35]. In the case of unstable systems the infinite time gramians are not defined and, thus, cannot be used to obtain a reduced model.

As an alternative to the infinite time gramians, in [35] the time-limited controllability and observability gramians are introduced, these are defined by

$$\mathcal{C}(t_0, t_1) := \int_{t_0}^{t_1} e^{A(t_1 - \tau)} B B^\top e^{A^\top (t_1 - \tau)} \mathrm{d}\tau \,, \tag{29}$$

$$\mathcal{O}(t_0, t_1) := \int_{t_0}^{t_1} e^{A^\top (\tau - t_0)} C^\top C e^{A(\tau - t_0)} \mathrm{d}\tau \tag{30}$$

with $t_0 = 0$ and $t_1 = t$ these are exactly the gramians given in Definition 3.6 and 3.7. The time-limited gramians can be used for model reduction (described in detail in section 3.2.2).

Let us now provide a simpler way of calculating the time-limited gramians given in eq. (29) and (30), the derivation closely follows [35]. The time-limited gramians can be calculated by adapted Lyapunov equations for asymptotically stable LTI systems. This will be exemplified for the observability gramian. We start by linking the time-limited observability gramian (eq. (30)) to infinity gramian by simply exploiting the linearity of the integral

$$\begin{aligned}\mathcal{O}(t_0, t_1) &= \int_{t_0}^{t_1} e^{A^\top(\tau - t_0)} C^\top C e^{A(\tau - t_0)} \mathrm{d}\tau \\ \Leftrightarrow \mathcal{O}(t_0, t_1) &= e^{-A^\top t_0} \int_{t_0}^{t_1} e^{A^\top \tau} C^\top C e^{A\tau} \mathrm{d}\tau \; e^{-At_0}\end{aligned}$$

and

$$\begin{aligned}\int_{t_0}^{t_1} e^{A^\top \tau} C^\top C e^{A\tau} \mathrm{d}\tau &= \int_0^{\infty} e^{A^\top \tau} C^\top C e^{A\tau} \mathrm{d}\tau - \int_0^{t_0} e^{A^\top \tau} C^\top C e^{A\tau} \mathrm{d}\tau - \int_{t_1}^{\infty} e^{A^\top \tau} C^\top C e^{A\tau} \mathrm{d}\tau \\ &= \mathcal{O}(\infty) - \mathcal{O}(0, t_0) - \mathcal{O}(t_1, \infty).\end{aligned}$$

Now computing the first two terms of the Lyapunov equations (26) for $\int_{t_0}^{t_1} e^{A^\top \tau} C^\top C e^{A\tau} \mathrm{d}\tau$:

$$\begin{aligned}A^\top \int_{t_0}^{t_1} e^{A^\top \tau} C^\top C e^{A\tau} \mathrm{d}\tau + \int_{t_0}^{t_1} e^{A^\top \tau} C^\top C e^{A\tau} \mathrm{d}\tau A &= \int_{t_0}^{t_1} \frac{\mathrm{d}}{\mathrm{d}\tau} e^{A^\top \tau} C^\top C e^{A\tau} \mathrm{d}\tau \\ &= e^{A^\top t_1} C^\top C e^{At_1} - e^{A^\top t_0} C^\top C e^{At_0}.\end{aligned} \tag{31}$$

Multiplying equation (31) with $e^{-A^\top t_0}$ from the left, e^{-At_0} from the right and using $A^\top e^{A^\top t_1} = e^{A^\top t_1} A^\top$, equation (31) results in

$$A^\top \mathcal{O}(t_0, t_1) + \mathcal{O}(t_0, t_1) A = e^{A^\top (t_1 - t_0)} C^\top C e^{A(t_1 - t_0)} - C^\top C.$$

It was, therefore, shown as in [35] that the time-limited observability gramian $\mathcal{O}(t_0, t_1)$ for stable LTI systems can be calculated by solving the following adapted Lyapunov equation

$$A^\top \mathcal{O}(t_0, t_1) + \mathcal{O}(t_0, t_1) A + C^\top C - e^{A^\top (t_1 - t_0)} C^\top C e^{A(t_1 - t_0)} = 0. \tag{32}$$

Similarly, for the time limited controllability gramian, we have

$$\mathcal{C}(t_0, t_1) = \int_{t_0}^{t_1} e^{A(t_1 - \tau)} B B^\top e^{A^\top (t_1 - \tau)} \mathrm{d}\tau$$

and by using $s = t_1 - \tau$ we obtain

$$\begin{aligned}\mathcal{C}(t_0, t_1) &= -\int_{t_1 - t_0}^{0} e^{As} B B^\top e^{A^\top s} \mathrm{d}s, \\ &= \int_0^{t_1 - t_0} e^{As} B B^\top e^{A^\top s} \mathrm{d}s.\end{aligned}$$

Following the same steps as for the observability gramian, the time limited controllability

gramian $\mathcal{C}(t_0, t_1)$ can be calculated by

$$A\mathcal{C}(t_0,t_1) + \mathcal{C}(t_0,t_1)A^\top + BB^\top - e^{A(t_1-t_0)}BB^\top e^{A^\top(t_1-t_0)} = 0. \tag{33}$$

The time-limited Lyapunov eqs. (32) and (33) can be employed to calculate the finite time gramian for stable LTI systems. The computation of the analytical expression in eq. (23) and (24) is challenging from a numerical viewpoint due to the matrix exponential [5, p. 70], such that for the efficient computation of the gramians other approaches are used e.g. solving the Lyapunov equations in case of a stable LTI system.

For the LTI systems the questions of controllability/observability on a time interval $[t_0, t_1]$ can always be rephrased as the controllability/observability on the time interval $[0, t_1 - t_0] = [0, t_{\text{end}}]$ with $t_{\text{end}} = t_1 - t_0$. This is not the case for systems where the matrices A, B and C are time dependent.

Linear time-varying (LTV) systems naturally arise when one linearises a nonlinear system about a specific trajectory [131] and are, thus, of interest in section 3.2.3, where gramians are introduced for nonlinear systems.

An LTV system is given by

$$\begin{aligned} \dot{x}(t) &= A(t)x(t) + B(t)u(t), \quad x(t_0) = x_0\,, \\ y(t) &= C(t)x(t), \end{aligned} \tag{34}$$

where B and C are matrix functions. For LTV systems equation (29) and (30) have to be adapted.

Definition 3.11 (time-limited controllability gramian [6, p. 230])
For an LTV system eq. (34) *the $n \times n$ time-limited controllability gramian $\mathcal{C}(t_0, t_1)$ is defined by*

$$\mathcal{C}(t_0,t_1) = \int_{t_0}^{t_1} \Phi^{t_1,\tau} B(\tau)B(\tau)^\top \left(\Phi^{t_1,\tau}\right)^\top \mathrm{d}\tau\,, \tag{35}$$

with Φ^{t,t_0} the state transition matrix.

Definition 3.12 (time-limited observability gramian [6, p.249])
For an LTV system eq. (34) *the $n \times n$ time-limited observability gramian $\mathcal{O}(t_0, t_1)$ is defined by*

$$\mathcal{O}(t_0,t_1) = \int_{t_0}^{t} \left(\Phi^{\tau,t_0}\right)^\top C^\top(\tau)C(\tau)\Phi^{\tau,t_0}\mathrm{d}\tau\,, \tag{36}$$

with Φ^{t,t_0} the state transition matrix.

For LTV systems, there does not exist a well-defined transfer function G. Thus, model reduction methods based on providing a good approximation of the transfer function cannot necessarily be employed for LTV systems (no analytical error bound available). The charac-

terisation of controllability/observability (in analogy to LTI systems given by corollary 3.8 and 3.9) of the LTV system on the time interval $[t_0, t_1]$ is given if the controllability/observability gramian is positive definite [30].

The analytical gramians (e.g. eq. (23) and eq. (24)) introduced in this section provide a basis for a popular model order reduction technique in control theory, which shall be briefly introduced in the subsequent subsection.

3.2.2. Balanced truncation

Balanced truncation is one of the most commonly used methods of model reduction in the field of control theory and was first introduced by Moore [85]. In this section important aspects of balanced truncation are discussed in detail, closely following the derivations in [5, ch. 7]. The idea of balanced truncation is to remove state variables that have the least contribution to the overall input-output behaviour of the system. Important properties of balanced truncation for LTI system with zero initial condition ($x(t_0) = 0$) are that by the elimination of some of the transformed state variables stability of the system is preserved and there exists an a-priori computable error bound for the reduced system [5, p. 207]. This a-priori error bound explains the popularity of balanced truncation for reduction of large-scale systems [7].

Balanced truncation can only be applied to systems that are controllable and observable, i.e. the controllability and observability gramian are positive definite. The coordinate transformation is then constructed by simultaneously diagonalising the controllability and observability gramian.

In order to investigate the input-output behaviour of a system, the Hankel operator needs to be employed.

Definition 3.13 (Hankel operator [5, p. 135])
Given an LTI system eq. (21) *the Hankel operator* $\mathcal{H} : L_2^s(-\infty, 0) \to L_2^q(0, \infty), u_- \mapsto y_+$ *is defined by*

$$y_+(t) = \mathcal{H}(u_-)(t) = \int_{-\infty}^{0} Ce^{A(t-s)}Bu_-(s)\mathrm{d}s = Ce^{At}\int_{-\infty}^{0} e^{-As}Bu_-(s)\mathrm{d}s\,,$$

with $L_2^s(-\infty, 0), L_2^q(0, \infty)$ *the Lebesque function spaces with 2-norm over* $\mathbb{R}^s$,$\mathbb{R}^q$ *respectively.*

The Hankel operator, which is closely linked to the controllability and observability gramian, thus, maps past inputs u_- into future outputs y_+. A singular value decomposition of this linear operator leads to n singular values, called the Hankel singular values.

In balanced truncation, the goal is to find a basis in the state space in which states that are difficult to control are simultaneously difficult to observe. The problem is, thus, to find a transformation matrix T, such that $\mathcal{C}$ and $\mathcal{O}$ are equal and diagonal.

Definition 3.14 (balanced system [5, p. 210])
A fully controllable, fully observable and stable system as given in eq. (21) *is called balanced if* $\mathcal{C} = \mathcal{O} = \Sigma$.

In a balanced system the degree of controllability and the degree of observability of each state are the same [29].

A state space transformation T that balances an LTI system exists if the controllability and observability gramian are both positive definite [5, p. 209ff]. The transformation T is then constructed by first computing the Cholesky factorisation of $\mathcal{C} = RR^\top$. Then the Hankel singular values $\{\sigma_1, \ldots, \sigma_n\}$ are derived by applying the eigenvalue value decomposition to $R^\top \mathcal{O} R$

$$R^\top \mathcal{O} R = U \Sigma^2 U^\top \text{ with } \Sigma = \begin{pmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_n \end{pmatrix}.$$

The Hankel singular values can be used to assess the importance of state variables for the input-output behaviour of the system. The state variables of the transformed system corresponding to large Hankel singular values are more important for the system.

The transformation, balancing the system, is defined as

$$T = \Sigma^{\frac{1}{2}} U^\top R^{-1},$$

with inverse

$$T^{-1} = R U \Sigma^{-\frac{1}{2}}.$$

In the transformed system $x_{\text{trans}} = Tx$, state variables are orthogonal in terms of their contribution to the input-output behaviour [116]. By truncation of the $\hat{x}$ (cf. eq. (13)) corresponding to small Hankel singular values $\sigma_{n_{\text{red}}} \geq \sigma_i$ with n_{red} the dimension of the reduced system, a good approximation of the transfer function G (cf. eq. (22)) can be obtained. The truncation is obtained by applying the following Petrov-Galerkin projection

$$P = \begin{pmatrix} I_{n_{\text{red}}} \\ 0 \end{pmatrix}.$$

Then

$$\begin{aligned} V &= TP, \\ W &= (T^{-1})^\top P, \end{aligned}$$

which results in the following reduced system

$$\begin{aligned} \dot{x}_{\text{red}}(t) &= W^\top f(V x_{\text{red}}(t)) + W^\top g(u(t)), \\ y_{\text{red}}(t) &= h(V x_{\text{red}}(t)). \end{aligned}$$

For LTI system, the dimension n_{red} of the reduced system can be determined a-priori based on the error bound

$$\sigma_{n_{\text{red}}+1} \leq ||G - G_{\text{red}}||_\infty = \sup_{u \in L_2^s} \frac{||y - y_{\text{red}}||}{||u||} \leq 2 \sum_{i=n_{\text{red}}+1}^{n} \sigma_i \,,$$

with G, G_{red} the transfer functions of the full system (given in eq. (22)) and reduced system, respectively. Although no error-bound is available for nonlinear systems, balanced truncation has gained popularity for nonlinear systems, too [116]. We will use balanced truncation in section 4.2 to obtain a reduced model for a challenging benchmark system from the field of electrical circuits.

3.2.3. Overview of empirical gramians for nonlinear systems

So far we have introduced the characterisation of both controllability and observability by the analytical gramians of linear systems. To obtain analytical gramians for nonlinear system one can apply linearisation around (i) a specific state of interest, e.g. steady state, which results in an LTI system or (ii) a specific solution trajectory, which gives an LTV system. The linearisation will only provide a good approximation of the gramian for the nonlinear system around the (i) specific state or (ii) the specific solution trajectory. The so-called empirical gramians are the extension of gramians to nonlinear systems without linearisation and, thus, provide a means for a more accurate approximation over a larger operating region Ω, which is defined by the predefined perturbations in the input and initial state. They are computed by averaging over trajectory samples with perturbations in the input and initial state. The empirical gramians are a 'local' quantity in the sense that they are defined with respect to a pre-specified operating region Ω. In the literature there exist different variants of empirical gramians, which are discussed in this section.

The first attempt to characterise the controllability and observability of nonlinear systems was done by so-called covariance matrices [80], based on the following calculation

$$\mathcal{M} = \int_\Omega \int_0^\infty Q(x(t) - x^*)(x(t) - x^*)^\top Q^\top \mathrm{d}t \mathrm{d}\mu_\Omega$$

with a positive definite and quadratic weighting matrix Q, an arbitrary stationary reference state x^* and an appropriate measure μ_Ω on Ω.

Empirical gramians

The so-called empirical controllability and observability matrices were first introduced in [70]. The empirical gramians can be seen as an extension of the analytical LTI gramians to nonlinear time invariant systems. The empirical gramians were introduced by [70] for a control-affine (the control enters the differential equation linearly) nonlinear system of the form

$$\begin{aligned} \dot{x}(t) &= f(x(t)) + g(x(t))u(t), \quad x(0) = 0 \\ y(t) &= h(x(t)) \end{aligned} \tag{37}$$

with $u(t) \in \mathbb{R}^s, y \in \mathbb{R}^q$.

The construction of the empirical gramians in [70] is based on specific correlation matrices defined as

$$\mathcal{N} := \int_0^\infty (x(t) - \bar{x})(x(t) - \bar{x})^\top \mathrm{d}t \tag{38}$$

where the so called time averaged state is defined by

$$\bar{x} := \lim_{T\to\infty} \frac{1}{T} \int_0^\top x(t)\mathrm{d}t \tag{39}$$

and $x(t)$ the vector of state variables of the system. In [70] it is assumed that the nonlinear system given in eq. (37) is exponentially stable with zero being the steady state and the time averaged state as given in eq. (39) is, thus, zero.

The empirical gramians are then numerically computed using finite time trajectory samples with perturbations in input and initial states for a time interval $[0, t_{\mathrm{end}}]$, which give a good approximation of the infinite time gramian if t_{end} is chosen large enough. Note that this was shown for LTI systems in eq. (28), but was not proven for nonlinear systems.

To construct the different variants of empirical gramians, let us define the following two sets for the input and state perturbations:

$$\begin{aligned}\mathcal{U} = \{&u_{\mathrm{per}} \in \mathbb{R}^s \mid u_{\mathrm{per}} = c_m T_l e_i \text{ for } i \in \{1,\dots,s\}, l \in \{1,\dots,L\}, m \in \{1,\dots,M\},\\ &c_m \in \mathbb{R}_+, T_l \in \mathbb{R}^{s\times s}, T_l^\top T_l = I_s \text{ and } e_i \text{ the } i\text{th standard unit vector in } \mathbb{R}^s\}\end{aligned} \tag{40}$$

$$\begin{aligned}\mathcal{X} = \{&x_{\mathrm{per}}(0) \in \mathbb{R}^{n\times n} \mid x_{\mathrm{per}}(0) = c_m T_l \text{ for } l \in \{1,\dots,L\}, m \in \{1,\dots,M\},\\ &c_m \in \mathbb{R}_+, T_l \in \mathbb{R}^{n\times n}, T_l^\top T_l = I_n\}\end{aligned} \tag{41}$$

with L orthogonal matrices T_l and M different perturbation magnitudes c_m. The orthogonal matrices T_l represent the direction of the perturbation, the positive real constants c_m give the magnitude of the perturbation and the standard unit vector e_i represents the input or state variable that is affected by the perturbation. In this case the operating region Ω is, thus, given by $\mathcal{U} \times \mathcal{X}$.

For each perturbation, the correlation matrix as given in eq. (38) is calculated in [70] and then the empirical gramians are given by the weighted average over all the calculated correlation matrices. In order to understand the differences in the development of the empirical gramians and, in particular, our developed extension of the empirical gramians, we provide in this section a unified framework for the empirical gramians given in literature [70, 43, 19]. The definitions from [70] for an infinite time empirical controllability gramian and an infinite time empirical observability gramian in our unified framework are then given by:

Definition 3.15 (Infinite time empirical controllability gramian)
Let $\mathcal{U}$ be the set of input perturbations as defined in eq. (40)*. For the nonlinear system eq.* (37)*, an infinite time empirical controllability gramian Lall–$\mathcal{C}_{\mathcal{U}}(\infty)$ is defined by*

$$\textit{Lall–}\mathcal{C}_{\mathcal{U}}(\infty) = \int_0^\infty \overline{\Psi}(t,0)\overline{\Psi}(t,0)^\top \mathrm{d}t \tag{42}$$

where $\overline{\Psi}(t,0) \in \mathbb{R}^{n\times s}$ *is given by*

$$\overline{\Psi}(t,0) := \frac{s}{|\mathcal{U}|} \sum_{u_{per}\in\mathcal{U}} \frac{1}{||u_{per}||_2^2} \Psi_{u_{per}}(t,0) \tag{43}$$

and $\Psi_{u_{per}}(t,0)$ *is defined by*

$$\Psi_{u_{per}}(t,0) := (x(t;0,u_{per}(t)) - \bar{x}(u_{per})) \cdot u_{per}^{\top} \tag{44}$$

with $x(t;0,u_{per}(t))$ *the state of the system corresponding to the impulse input* $u_{per}(t) = u_{per}\delta(t)$.

Note that the impulse input could as well have been written as $x_{u_{\mathrm{per}}}(0) = g(0)u_{\mathrm{per}}$. It was shown in [70] that for stable LTI systems an infinite time empirical controllability gramian is equal to the infinite time analytical controllability gramian, as shown in the following lemmas given with proof. The proofs of both lemmas are provided for completeness and to facilitate understanding of subsequent proofs.

Lemma 3.16 ([70])
Given a stable LTI system of the form $\dot{x}(t) = Ax(t) + Bu(t)$ *and* $y = Cx(t)$ *on the time interval* $[0,\infty]$. *Then, for any non-empty perturbation set* $\mathcal{U}$ *as given in eq.* (40), *an infinite time empirical controllability gramian Lall–*$\mathcal{C}_{\mathcal{U}}(\infty)$ *is identical to the infinite time analytical controllability gramian* $C(\infty) = \int_0^\infty e^{At}BB^{\top}e^{A^{\top}t}\mathrm{d}t$.

Proof: The input-perturbed solution for the LTI system is $x(t;0,u_{\mathrm{per}}) = \int_0^t e^{A(t-s)}Bu_{\mathrm{per}}(s)\mathrm{d}s$ with $u_{\mathrm{per}}(t) = u_{\mathrm{per}}\delta(t)$ and, thus, $x(t;0,u_{\mathrm{per}}) = e^{At}Bu_{\mathrm{per}}$.
Using the input-perturbed solution for the LTI system, $\Psi_{u_{\mathrm{per}}}(t,0)$ becomes

$$\Psi_{u_{\mathrm{per}}}(t,0) = \left(e^{At}Bu_{\mathrm{per}} - \lim_{T\to\infty}\frac{1}{T}\int_0^T e^{At}Bu_{\mathrm{per}}\mathrm{d}t\right)u_{\mathrm{per}}^{\top}\,.$$

Next we use the fact that the system is stable and that the time averaged state is zero. Thus, $\Psi_{u_{\mathrm{per}}}(t,0)$ simplifies to

$$\Psi_{u_{\mathrm{per}}}(t,0) = e^{At}Bu_{\mathrm{per}}u_{\mathrm{per}}^{\top}\,,$$

with $u_{\mathrm{per}} \in \mathcal{U}$. Inserting $\Psi_{u_{\mathrm{per}}}(t,0)$ into eq. (43), $\overline{\Psi}(t,0)$ becomes

$$\overline{\Psi}(t,0) = \frac{s}{|\mathcal{U}|}\sum_{u_{\mathrm{per}}\in\mathcal{U}}\frac{1}{||u_{\mathrm{per}}||_2^2}e^{At}Bu_{\mathrm{per}}u_{\mathrm{per}}^{\top}\,,$$

and using $u_{\mathrm{per}} = c_m T_l e_i$, we get

$$\begin{aligned}\overline{\Psi}(t,0) &= \frac{s}{s\cdot L\cdot M}\sum_{i,l,m}\frac{1}{c_m^2}e^{At}Bc_mT_le_i(c_mT_le_i)^{\top}\\ &= \frac{1}{L\cdot M}\sum_{i,l,m}e^{At}BT_le_ie_i^{\top}T_l^{\top}\,,\end{aligned} \tag{45}$$

where

$$\sum_i e^{At} B T_l e_i e_i^\top T_l^\top = e^{At} B T_l T_l^\top .$$

Thus, eq. (45) further simplifies to

$$\begin{aligned} \overline{\Psi}(t,0) &= \frac{1}{L \cdot M} \sum_{l,m} e^{At} B \underbrace{T_l T_l^\top}_{=I_s} \\ &= \frac{1}{L \cdot M} \sum_{l,m} e^{At} B \\ &= e^{At} B \,. \end{aligned} \tag{46}$$

Then finally by inserting eq. (46) into eq. (42) an infinite time empirical gramian Lall–$\mathcal{C}_{\mathcal{U}}(\infty)$ is given by

$$\begin{aligned} \text{Lall–}\mathcal{C}_{\mathcal{U}}(\infty) &= \int_0^\infty e^{At} B \left(e^{At} B\right)^\top \mathrm{d}t \\ &= \int_0^\infty e^{At} B B^\top e^{A^\top t} \mathrm{d}t = \mathcal{C}(\infty) \,, \end{aligned}$$

which is the desired result. □

Definition 3.17 (Infinite time empirical observability gramian)
Let $\mathcal{X}$ be the set of initial state perturbations as defined in eq. (41). For the nonlinear system eq. (37), an infinite time empirical observability gramian Lall–$\mathcal{O}_{\mathcal{X}}(\infty)$ is defined by

$$\textit{Lall–}\mathcal{O}_{\mathcal{X}}(\infty) = \int_0^\infty \overline{\Theta}(t,0)^\top \overline{\Theta}(t,0) \mathrm{d}t \tag{47}$$

where $\overline{\Theta}(t,0) \in \mathbb{R}^{q \times n}$ is given by

$$\overline{\Theta}(t,0) = \frac{1}{|\mathcal{X}|} \sum_{x_{per} \in \mathcal{X}} \sum_{i=1}^{n} \frac{1}{||x_{per}||_2^2} \Theta_{x_{per,i}}(t,0) \tag{48}$$

and $\Theta_{x_{per,i}}(t,0)$ is defined by

$$\Theta_{x_{per,i}}(t,0) := (y(t; x_{per,i}(0), 0) - \bar{y}(x_{per,i})) \cdot x_{per,i}(0)^\top \tag{49}$$

with $y(t; x_{per,i}(0), 0)$ the output of the system corresponding to the initial condition $x_{per,i}(0) = x_{per} e_i$, e_i the ith standard unit vector in $\mathbb{R}^n$ and $u(t) = 0$.

Lemma 3.18 ([70])
Given a stable LTI system of the form $\dot{x}(t) = Ax(t) + Bu(t)$ and $y = Cx(t)$ on the time interval $[0, \infty]$. Then, for any non-empty perturbation set $\mathcal{X}$ as given in eq. (41), an infinite time empirical observability gramian Lall–$\mathcal{O}_{\mathcal{X}}(\infty)$ is identical to the infinite time analytical observability gramian $\mathcal{O}(\infty) = \int_0^\infty e^{A^\top \tau} C^\top C e^{A\tau} \mathrm{d}\tau$.

Proof: The initial state-perturbed solution for the LTI system is $x(t; x_{\mathrm{per},i}(0), 0) = e^{At} x_{\mathrm{per},i}(0)$ and $y(t; x_{\mathrm{per},i}(0), 0) = Ce^{At} x_{\mathrm{per},i}(0)$.

Inserting the initial state-perturbed solution into eq. (49) gives

$$\Theta_{x_{\mathrm{per},i}}(t,0) = \left(Ce^{At} x_{\mathrm{per},i}(0) - \lim_{T\to\infty} \frac{1}{T} \int_0^T Ce^{At} x_{\mathrm{per},i}(0) \mathrm{d}t \right) \cdot x_{\mathrm{per},i}(0)^\top .$$

Using the fact that the system is stable and the time averaged state is zero, $\Theta_{x_{\mathrm{per},i}}(t,0)$ simplifies to

$$\Theta_{x_{\mathrm{per},i}}(t,0) = Ce^{At} x_{\mathrm{per},i}(0) \cdot x_{\mathrm{per},i}(0)^\top$$

with $x_{\mathrm{per},i}(0) = x_{\mathrm{per}} e_i$. Inserting $\Theta_{x_{\mathrm{per},i}}(t,0)$ into eq (48), $\overline{\Theta}(t,0)$ becomes

$$\overline{\Theta}(t,0) = \frac{1}{|\mathcal{X}|} \sum_{x_{\mathrm{per}} \in \mathcal{X}} \sum_{i=1}^{n} \frac{1}{||x_{\mathrm{per}}||_2^2} Ce^{At} x_{\mathrm{per},i}(0) \cdot x_{\mathrm{per},i}(0)^\top .$$

and using $x_{\mathrm{per},i}(0) = x_{\mathrm{per}} e_i = c_m T_l e_i$, we get

$$\begin{aligned} \overline{\Theta}(t,0) &= \frac{1}{L \cdot M} \sum_{l,m} \sum_{i=1}^{n} \frac{1}{c_m^2} Ce^{At} c_m T_l e_i \cdot (c_m T_l e_i)^\top \\ &= \frac{1}{L \cdot M} \sum_{l,m} \sum_{i=1}^{n} Ce^{At} T_l e_i e_i^\top T_l^\top , \end{aligned} \tag{50}$$

where

$$\sum_{i=1}^{n} Ce^{At} T_l e_i e_i^\top T_l^\top = Ce^{At} T_l T_l^\top .$$

Thus, eq. (50) further simplifies to

$$\begin{aligned} \overline{\Theta}(t,0) &= \frac{1}{L \cdot M} \sum_{l,m} Ce^{At} \underbrace{T_l T_l^\top}_{I_n} \\ &= \frac{1}{L \cdot M} \sum_{l,m} Ce^{At} = Ce^{At} . \end{aligned} \tag{51}$$

Then by inserting eq. 51 into eq. 47, an infinite time empirical observability gramian Lall–$\mathcal{O}_{\mathcal{X}}(\infty)$ becomes

$$\begin{aligned} \text{Lall–}\mathcal{O}_{\mathcal{X}}(\infty) &= \int_0^\infty \left(Ce^{At} \right)^\top Ce^{At} \mathrm{d}t \\ &= \int_0^\infty e^{A^\top t} C^\top Ce^{At} \mathrm{d}t = \mathcal{O}(\infty) , \end{aligned}$$

which is the desired result. □

The infinite time empirical gramians with impulse inputs given in definitions 3.15 and 3.17, thus, are identical to the infinite time analytical gramians for stable time-invariant linear

systems independent of applied input and initial state perturbation. In order to provide a comprehensive overview over the properties of each empirical gramian definition in this section, we provide a summary of the conditions (e.g. stable LTI system etc.) in which the defined empirical gramians are identical to the analytical gramians (cf. Table 1).

Table 1: **Conditions on an LTI system under which the infinite time empirical gramians in eqs. (42) and (47) are identical to the infinite time analytical gramians in eqs. (23) and (24) with $t \to \infty$.**

Properties of the LTI system given as in eq. (21)			Lall–$\mathcal{C}_{\mathcal{U}}(\infty) = \mathcal{C}(\infty)$, Lall–$\mathcal{O}_{\mathcal{X}}(\infty) = \mathcal{O}(\infty)$
stable with $x_0 = x_{\text{ss}}$	finite time	$x_0 = 0$	sometimes
		$x_0 \neq 0$	no
	infinite time	$x_0 = 0$	yes
		$x_0 \neq 0$	no
stable with $x_0 \neq x_{\text{ss}}$	finite time	$x_0 = 0$	no
		$x_0 \neq 0$	no
	infinite time	$x_0 = 0$	no
		$x_0 \neq 0$	no
unstable	finite time	$x_0 = 0$	no
		$x_0 \neq 0$	no

Infinite time empirical covariance gramians

The infinite time empirical gramians given in definitions 3.15 and 3.17 are explicitly based on impulse inputs. In order to allow for different kinds of inputs the concept was extended in [41, 42, 43] by the introduction of infinite time empirical covariance matrices, which allow for step inputs as well as series of step inputs. The empirical observability covariance matrix is given by the definition 3.17 with the adaptation, that the steady state of the system must not be the zero state and, thus, $x_{\text{per},i}(0)$ becomes $x_{\text{per},i}(0) = x_{\text{per}}(0)e_i + x_{\text{ss}}$ with x_{ss} denoting the steady state of the system. The infinite time empirical controllability covariance matrix is given by the following definition.

Definition 3.19 (Infinite time empirical controllability covariance gramian)
Let $\mathcal{U}$ be the set of input perturbations as defined in eq. (40). For the nonlinear system eq. (37), an infinite time empirical controllability covariance gramian Hahn–$\mathcal{C}_{\mathcal{U}}(\infty)$ is defined by

$$\textit{Hahn–}\mathcal{C}_{\mathcal{U}}(\infty) = \int_0^\infty \overline{\Psi}(t,0)\overline{\Psi}(t,0)^\top \mathrm{d}t \tag{52}$$

where $\overline{\Psi}(t,0) \in \mathbb{R}^{n\times s}$ is given by

$$\overline{\Psi}(t,0) := \frac{s}{|\mathcal{U}|} \sum_{u_{per}\in\mathcal{U}} \frac{1}{||u_{per}||_2^2} \Psi_{u_{per}}(t,0) \tag{53}$$

and $\Psi_{u_{per}}(t,0)$ is defined by

$$\Psi_{u_{per}}(t,0) := (x(t; x_{ss}, u_{per}(t)) - x_{per,ss}) \cdot u_{per}^\top \tag{54}$$

with $x(t; x_{ss}, u_{per}(t))$ *the state of the system corresponding to the perturbed input* $u_{per}(t) = u_{per}u(t) + u_{ss}$*, where* u_{ss} *denotes the input at the original steady state,* x_{ss} *and* $x_{per,ss}$ *denote the steady state of the original and perturbed solution trajectory corresponding to the input* $u_{per}(t)$ *respectively.*

In [43] it was shown that both infinite time empirical covariance gramians are identical to the infinite time analytical gramians for stable time invariant linear systems for inputs defined by a delta impulse or a series of step functions given by

$$u(t) = \sum_{k=1}^{z} a_k S(t - t_k^{\text{step}}), \quad -1 \leq a_k \leq 1$$

with $S(t - t_k^{\text{step}}) = 1$ for $t < t_k^{\text{step}}$ and zero otherwise, t_k^{step} time points when a step change occurs, a_k is the size and direction of the step change. Furthermore, in [43] it is noted that it is possible to compute the infinite time empirical covariance gramians even for systems with multiple steady states, however, special attention should be given to the interpretation of results for cases in which perturbed solution trajectories leave the region of attraction of the original considered steady state. In Table 2 the cases are listed in which the infinite time empirical covariance gramians, defined in 3.19, are identical to the infinite time analytical gramians.

A major drawback of the so far discussed empirical gramians (cf. Definition 3.15, 3.17 and 3.19), however, is that only the relationship to stable LTI systems was established. Thus, completely neglecting the relationship between a nonlinear system and the linearised system along the trajectory resulting in an LTV system. Note that the definitions above do not produce the correct gramians for stable LTV systems since they are computed using delta impulses to characterise the input-to-state behaviour of the system. The characterisation of the input-state behaviour by delta impulses is sufficient for LTI systems, but not for LTV systems. In [19] an extension of the empirical gramians was presented and shown that the extended empirical gramians are identical to the analytical infinite time gramians for stable LTV systems. The extended empirical gramians provide a better approximation for the nonlinear system.

Extension of empirical gramians to be linked to stable LTV systems

In the following we consider the nonlinear system given in eq. (37) with $g(x(t)) = B(t)$. In [19] the steady state as well as the input into the system is assumed to be zero. Then the extended infinite time empirical controllability and observability gramian based on averaging of the state transition matrix are defined by

Definition 3.20 (Infinite time empirical controllability gramian)

Let $\mathcal{X}$ *be the set of initial condition perturbations as defined in eq.* (41)*. For the nonlinear system eq.* (37) *with* $g(x(t)) = B(t)$*, an infinite time empirical controllability gramian Condon–*$\mathcal{C}_{\mathcal{X}}(\infty)$ *is defined by*

$$\textit{Condon–}\mathcal{C}_{\mathcal{X}}(\infty) = \int_{-\infty}^{0} \overline{\Psi}(t,0)\overline{\Psi}(t,0)^{\top} \mathrm{d}t\,, \tag{55}$$

Table 2: **Conditions on an LTI system under which the infinite time empirical covariance gramians in eq.**(52) **are identical to the infinite time analytical gramian in eqs.** (23) **and** (24) **with** $t \to \infty$**.**

Properties of the LTI system given as in eq. (21)			Hahn–$\mathcal{C}_{\mathcal{U}}(\infty) = \mathcal{C}(\infty)$, Hahn–$\mathcal{O}_{\mathcal{X}}(\infty) = \mathcal{O}(\infty)$
stable with $x_0 = x_{\text{ss}}$	finite time	$x_0 = 0$	sometimes
		$x_0 \neq 0$	sometimes
	infinite time	$x_0 = 0$	yes
		$x_0 \neq 0$	yes
stable with $x_0 \neq x_{\text{ss}}$	finite time	$x_0 = 0$	no
		$x_0 \neq 0$	no
	infinite time	$x_0 = 0$	no
		$x_0 \neq 0$	no
unstable	finite time	$x_0 = 0$	no
		$x_0 \neq 0$	no

where $\overline{\Psi}(t,0) \in \mathbb{R}^{n \times s}$ *is given by*

$$\overline{\Psi}(t,0) = \left(\frac{1}{|\mathcal{X}|} \sum_{x_{per} \in \mathcal{X}} \sum_{i=1}^{n} \frac{1}{||x_{per}||_2^2} \Psi_{x_{per}}(t,0) \right)^{-1} \cdot B(t) \tag{56}$$

and $\Psi_{x_{per}}(t,0)$ *is defined by*

$$\Psi_{x_{per}}(t,0) = x(t; x_{per,i}(0); 0) \cdot x_{per,i}(0)^{\top} \tag{57}$$

with $x(t; x_{per,i}(0); 0)$ *the state of the system for the initial condition* $x_{per,i}(0) = x_{per} e_i$ *and* $e_i \in \mathbb{R}^n$ *the standard unit vector.*

Definition 3.21 (Infinite time empirical observability gramian)
Let $\mathcal{X}$ *be the set of initial condition perturbations as defined in eq.* (41)*. For the nonlinear system eq.* (37) *with* $g(x(t)) = B(t)$*, an infinite time empirical observability gramian Condon–*$\mathcal{O}_{\mathcal{X}}(\infty)$ *is defined by*

$$\textit{Condon–}\mathcal{O}_{\mathcal{X}}(\infty) = \int_0^{\infty} \overline{\Theta}(t,0)^{\top} \overline{\Theta}(t,0) \mathrm{d}t \tag{58}$$

where $\overline{\Theta}(t,0) \in \mathbb{R}^{q \times n}$ *is given by*

$$\overline{\Theta}(t,0) = \frac{1}{|\mathcal{X}|} \sum_{x_{per} \in \mathcal{X}} \sum_{i=1}^{n} \frac{1}{||x_{per}||_2^2} \Theta_{x_{per,i}}(t,0) \tag{59}$$

and $\Theta_{x_{per,i}}(t,0)$ *is defined by*

$$\Theta_{x_{per,i}}(t,0) := y(t; x_{per,i}(0), 0) \cdot x_{per,i}(0)^{\top} \tag{60}$$

with $y(t; x_{per,i}(0), 0)$ *the state of the system for the initial condition* $x_{per,i}(0) = x_{per} e_i$ *and* $e_i \in \mathbb{R}^n$ *the standard unit vector.*

Additionally in [19], it is assumed that the perturbation $x_{\mathrm{per},i}(0)$ does not result in a solution trajectory that leaves the region of attraction of the equilibrium point $x_{\mathrm{ss}} = 0$. In the case of multiple steady states of the system given in eq. (37) it is suggested to construct empirical gramians for each steady state separately [19].

In Table 3 the condition on an LTI systems are listed in which the infinite time empirical gramians defined in 3.20 and 3.21 are identical to the infinite time analytical gramians for linear systems.

Table 3: **Conditions on an LTI system or an LTV system under which the infinite time empirical gramians in 3.20 and 3.21 are identical to the infinite time analytical gramians in eqs.** (23) **and** (24) **with** $t \to \infty$**.**

Properties of the LTI system given as in eq. (21)			Condon–$\mathcal{C}_{\mathcal{U}}(\infty) = \mathcal{C}(\infty)$, Condon–$\mathcal{O}_{\mathcal{X}}(\infty) = \mathcal{O}(\infty)$
stable with $x_0 = x_{\mathrm{ss}}$	finite time	$x_0 = 0$	sometimes
		$x_0 \neq 0$	sometimes
	infinite time	$x_0 = 0$	yes
		$x_0 \neq 0$	sometimes
stable with $x_0 \neq x_{\mathrm{ss}}$	finite time	$x_0 = 0$	no
		$x_0 \neq 0$	no
	infinite time	$x_0 = 0$	no
		$x_0 \neq 0$	no
unstable	finite time	$x_0 = 0$	no
		$x_0 \neq 0$	no

Properties of the LTV system given as in eq. (34)			Condon–$\mathcal{C}_{\mathcal{U}}(\infty) = \mathcal{C}(\infty)$, Condon–$\mathcal{O}_{\mathcal{X}}(\infty) = \mathcal{O}(\infty)$
stable with $x_{\mathrm{ss}} = x_0$	finite time	$x_0 = 0$	sometimes
		$x_0 \neq 0$	sometimes
	infinite time	$x_0 = 0$	yes
		$x_0 \neq 0$	sometimes
stable with $x_{\mathrm{ss}} \neq x_0$	finite time	$x_0 = 0$	no
		$x_0 \neq 0$	no
	infinite time	$x_0 = 0$	no
		$x_0 \neq 0$	no
unstable	finite time	$x_0 = 0$	no
		$x_0 \neq 0$	no

Although the empirical gramians can easily be obtained by simple matrix computations, they have not gained much attention in the literature, until recently with the introduction of the empirical cross gramian in [49, 50, 51], which also can be used for model order reduction of nonlinear systems. Furthermore, in [49] a unified software framework for the empirical gramians is provided. In Figure 11 the development of the empirical gramians given in this section is summarised.

So far, so-called infinite time empirical gramians (e.g. Lall–$\mathcal{C}_{\mathcal{U}}(\infty)$, Hahn–$\mathcal{C}_{\mathcal{U}}(\infty)$, Condon–$\mathcal{C}_{\mathcal{U}}(\infty)$) were considered, with the underlying assumption that the nonlinear system is stable. No realistic system can be considered on an infinite time horizon. The infinite time gramians constitute only an approximation of the finite time/time-limited gramians that describe any physical system [35]. A formulation of the empirical gramians for finite time

Figure 11: **Summary of the development of the different empirical gramians introduced in section 3.2.3.**

(e.g. time interval $[0, t]$) or time-bounded (e.g. time interval $[t_0, t_1]$ with $t_0 \neq 0$) gramians is, thus, needed to provide not only a more accurate approximation of the nonlinear gramians but also to allow for the consideration of the transient behaviour of the nonlinear system.

In section 4.1 we, therefore, will generalise the empirical gramian formulation for nonlinear control affine system (not necessarily stable) which links to the time-limited analytical gramians for both LTI and LTV systems (not necessarily stable). In the next section, the presented relevant model reduction techniques are discussed in the context of system pharmacology/biology models.

3.3. Discussion of applicability of methods in the context of systems pharmacology/biology models

In the previous sections, we have introduced four different general concepts which are used for model order reduction in two different fields of application. The systems of interest in this thesis are from the field of systems pharmacology/biology. The models are usually nonlinear, have a meaningful interpretation of state variables and the number of parameters exceeds the number of state variables. In drug development and discovery the focus, under which systems pharmacology/biology models are investigated, is often on a specific drug influencing the system and one wants to characterise the drug effect or response in terms of some surrogate marker. Thus, we are interested in the characterisation of a specific input-output relationship for these systems. The system of interest is then given by

$$\begin{aligned} \dot{x}(t) &= f(x(t)) + u(t), \quad x(0) = x_0\,, \\ y(t) &= h(x(t)). \end{aligned} \tag{61}$$

The function $f : \mathbb{R}^n \to \mathbb{R}^n$ represents the systems pharmacology/biology model of interest, and the function $h : \mathbb{R}^n \to \mathbb{R}^q$ maps the state vector x to the output y of interest (e.g. an experimentally observed quantity). In a typical setting, $u(t) \in \mathbb{R}^n$ represent some drug administration (including intravenous bolus (i.v. bolus), intravenous infusion (i.v. infusion) and per os administration (p.o. administration)) or other stimulus of the system.

In the following sections, we discuss the introduced methods under aspects which are crucial for the application in system pharmacology/biology, namely the approximation accuracy, interpretability of state variables/parameters and consideration of the input-output relationship. All the introduced methods are in general suitable for nonlinear systems, however, the local sensitivities might not reflect the true importance of parameters and, in particular, a reduction based on the importance of reactions according to their sensitivity could lead to models with low approximation quality [123, 126]. Furthermore, the larger the nonlinear system the more difficult it is to find a suitable slow/fast partition of the system for time scale analysis and it must not necessarily exist. Finding an optimal lumping scheme is also a challenging task, especially for large-scale systems, since it is practically impossible to test all possible lumped models due to the combinatorial explosion in the number of possibilities. This problem can be overcome by using greedy search algorithms (e.g. in [1]), where then, however, it has not been considered to what extent the sequence in which state variables

are lumped impacts the final reduced model. This might have possible implications on the approximation quality of the reduced model.

The interpretability of the reduced model system is an important requirement in the field of systems pharmacology/biology, which is met by the proper lumping approach, time scale separation (if such a partition is known/ exists) and sensitivity analysis. As soon as a coordinate transformation is applied, this requirement might be violated, e.g. for balanced truncation and ILDM. Nonetheless, model reduction techniques for system pharmacology/biology models have used balanced truncation [117]. Due to the balanced truncation applied in this approach, the final reduced model lost its interpretability (as mentioned above).

We are interested in systems pharmacology/ biology models under a specific input-output relationship and want to reduce the system to its essential features for this relationship. Balanced truncation is very well suited for this purpose since it is specifically designed to maintain a particular input-output setting of a system. Time scale separation, on the other hand, does not explicitly take into account a specific input-output behaviour. For example, based on the chosen input-output setting, only certain pathways might be active. For the sensitivity analysis, an extension to take the input-output relationship into account exists. This is done by introducing an additional function for the output and considering how sensitive the output trajectory is to small changes in the parameters and initial values. It is, however, not clear how to extend the impact of state variables beyond the initial time point. We remark, that especially for signalling cascades in system pharmacology/biology, it is feasible to assume that a state variable might have no/ only a small impact initially on the output but a significant impact at later times. It is, therefore, highly desirable to extend the consideration of sensitivity coefficients of the model system state variables.

4. Novel input-response characterisation for model analysis & reduction

Often the transient output of systems pharmacology/biology models (cf. section 2.2 and 2.3) is of particular interest. For linear systems time-limited analytical gramians were introduced in order to capture the transient system behaviour [35]. For nonlinear systems, empirical gramians have been introduced to characterise the controllability and observability in the steady state of the system. In order to characterise the transient system behaviour for nonlinear systems, we will introduce time-limited empirical gramians in this section. To show the applicability of the novel empirical gramians beyond system pharmacology/biology models, we have chosen a nonlinear benchmark example from the field of control theory, which has been previously employed to show the superiority of a novel empirical gramian formulation [18]. Additionally, based on the introduced time-limited empirical gramians we will derive the novel quantity called the input-response index. The input-response index quantify the importance of state variables for a given input-output relationship as a function of time. We will show that a local variant of the input-response indices can be defined as the product of two sensitivity coefficients over time, the indices thus obtained will be termed sensitivity based input-response indices. These indices, as well as the empirical based indices, can be used to understand the dynamics of the system upon perturbation or stimulation by the input with respect to the specified output. Based on the input-response indices we introduce a very efficient model order reduction based on a four-step procedure: (i) elimination of state variables whose dynamics have only negligible impact on the input-response relationship, (ii) application of the quasi-steady state approximation to eliminate fast changing molecular species, (iii) exploiting the conservation laws, and a potential last step where (iv) proper lumping is applied to the remaining state variables (cf. section 4.5). The reduced model still allows for a mechanistic interpretation in terms of the quantities of the original system, which is a key requirement in the pharmacology domain of application.

4.1. New generalised empirical gramians

The novel generalised empirical gramians are defined with respect to a reference trajectory and given in a time-limited formulation. It can be shown that for LTI and LTV systems, the generalised empirical gramians are identical to the analytical time-limited gramians.

To construct the empirical gramians, we use the perturbation sets for the input and initial state perturbations as defined in section 3.2.1 in eq. (40) and (41).

Definition 4.1 (Time-limited empirical controllability gramian)
Let $\mathcal{U}$ be the set of input perturbations as defined in eq. (40)*. For the nonlinear system eq.* (37) *with reference solution $x_{ref}(t) = x(t; x_{ref}(t_0), u_{ref}(t))$, input u_{ref} and time interval $[t_0, t_1]$, a time-limited empirical controllability gramian REF-$\mathcal{C}_{\mathcal{U}}(t_0, t_1)$ with respect to the reference*

solution x_{ref} is defined by

$$REF\text{–}\mathcal{C}_{\mathcal{U}}(t_0,t_1) = \int_{t_0}^{t_1} \overline{\Psi}(t_1,\tau)\overline{\Psi}(t_1,\tau)^\top \mathrm{d}\tau \tag{62}$$

where $\overline{\Psi}(t_1,\tau) \in \mathbb{R}^{n\times s}$ is given by

$$\overline{\Psi}(t_1,\tau) := \frac{s}{|\mathcal{U}|} \sum_{u_{per}\in\mathcal{U}} \frac{1}{||u_{per}||_2^2} \Psi_{u_{per}}(t_1,\tau) \tag{63}$$

and $\Psi_{u_{per}}(t_1,\tau)$ is defined by

$$\Psi_{u_{per}}(t_1,\tau) = \big(x(t;x_{u_{per}}(\tau),u_{ref}(t)) - x_{ref}(t;x_{ref}(\tau),u_{ref}(t))\big)\cdot u_{per}^\top \tag{64}$$

and $x(t;x_{u_{per}}(\tau),u_{ref})$ is the state corresponding to the initial state value $x_{u_{per}}(\tau) = x_{ref}(\tau) + g(x_{ref}(\tau))u_{per}$.

Lemma 4.2
Given an LTV system of the form $\dot{x}(t) = A(t)x(t) + B(t)u(t)$ and $y(t) = C(t)x(t)$ on the time interval $[t_0,t_1]$. Then, for any non-empty perturbation set $\mathcal{U}$ as given in eq. (40), a time-limited empirical controllability gramian $REF\text{–}\mathcal{C}_{\mathcal{U}}(t_0,t_1)$ is identical to the analytical time-limited controllability gramian $\mathcal{C}(t_0,t_1) = \int_{t_0}^{t_1} \Phi^{t_1,\tau}B(\tau)B(\tau)^\top \left(\Phi^{t_1,\tau}\right)^\top \mathrm{d}\tau$.

Proof: The input-perturbed solution for the LTV system is $x(t;x_{u_{\mathrm{per}}}(\tau),u_{\mathrm{ref}}(t)) = \Phi^{t,\tau}x_{u_{\mathrm{per}}}(\tau) + \int_\tau^t \Phi^{t,s}B(s)u_{\mathrm{ref}}(s)\mathrm{d}s$ with $x_{u_{\mathrm{per}}}(\tau) = x_{\mathrm{ref}}(\tau) + B(\tau)u_{\mathrm{per}}$.

Using the input-perturbed solution of the LTV system, $\Psi_{u_{\mathrm{per}}}(t_1,\tau)$ becomes

$$\begin{aligned}\Psi_{u_{\mathrm{per}}}(t_1,\tau) &= \big(\Phi^{t_1,\tau}x_{u_{\mathrm{per}}}(\tau) - \Phi^{t_1,\tau}x_{\mathrm{ref}}(\tau)\big)\cdot u_{\mathrm{per}}^\top \\ &= \big(\Phi^{t_1,\tau}x_{\mathrm{ref}}(\tau) + \Phi^{t_1,\tau}B(\tau)u_{\mathrm{per}} - \Phi^{t_1,\tau}x_{\mathrm{ref}}(\tau)\big)\cdot u_{\mathrm{per}}^\top \\ &= \big(\Phi^{t_1,\tau}B(\tau)u_{\mathrm{per}}\big)\cdot u_{\mathrm{per}}^\top\end{aligned}$$

with $u_{\mathrm{per}} \in \mathcal{U}$. Inserting $\Psi_{u_{\mathrm{per}}}(t_1,\tau)$ into eq. (63), $\overline{\Psi}(t_1,\tau)$ becomes

$$\overline{\Psi}(t_1,\tau) = \frac{s}{|\mathcal{U}|}\sum_{u_{\mathrm{per}}\in\mathcal{U}} \frac{1}{||u_{\mathrm{per}}||_2^2}\Phi^{t_1,\tau}B(\tau)u_{\mathrm{per}}\cdot u_{\mathrm{per}}^\top$$

and using $u_{\mathrm{per}} = c_m T_l e_i$, we get

$$\begin{aligned}\overline{\Psi}(t_1,\tau) &= \frac{s}{s\cdot L\cdot M}\sum_{i,l,m}\frac{1}{c_m^2}\Phi^{t_1,\tau}B(\tau)c_m T_l e_i\cdot(c_m T_l e_i)^\top \\ &= \frac{1}{L\cdot M}\sum_{i,l,m}\frac{c_m^2}{c_m^2}\Phi^{t_1,\tau}B(\tau)T_l e_i e_i^\top T_l^\top\,,\end{aligned} \tag{65}$$

where

$$\sum_i \Phi^{t_1,\tau} B(\tau) T_l e_i e_i^\top T_l^\top = \Phi^{t_1,\tau} B(\tau) T_l T_l^\top .$$

Thus, eq. (65) further simplifies to

$$\begin{aligned} \overline{\Psi}(t_1,\tau) &= \frac{1}{L \cdot M} \sum_{l,m} \Phi^{t_1,\tau} B(\tau) \underbrace{T_l T_l^\top}_{=I_s} \\ &= \frac{1}{L \cdot M} \sum_{l,m} \Phi^{t_1,\tau} B(\tau) \\ &= \Phi^{t_1,\tau} B(\tau) , \end{aligned} \tag{66}$$

Finally, by inserting eq. (66) into (62), a time-limited empirical controllability gramian REF–$\mathcal{C}_\mathcal{U}(t_0,t_1)$ is given by

$$\text{REF–}\mathcal{C}_\mathcal{U}(t_0,t_1) = \int_{t_0}^{t_1} \Phi^{t_1,\tau} B(\tau) B(\tau)^\top \left(\Phi^{t_1,\tau}\right)^\top \mathrm{d}\tau = \mathcal{C}(t_0,t_1),$$

which is the desired result. □

Definition 4.3 (Time-limited empirical observability gramian)
Let $\mathcal{X}$ be the set of input perturbations as defined in eq. (41). For the nonlinear system given in eq. (37) with reference solution $x_{ref}(t) = x(t; x_{ref}(t_0), u_{ref}(t))$, input u_{ref} and time interval $[t_0,t_1]$, a time-limited empirical observability gramian REF–$\mathcal{O}_\mathcal{X}(t_0,t_1)$ with respect to the reference solution x_{ref} is defined by

$$\text{REF–}\mathcal{O}_\mathcal{X}(t_0,t_1) := \int_{t_0}^{t_1} \overline{\Theta}(\tau,t_0)^\top \overline{\Theta}(\tau,t_0) \mathrm{d}\tau \tag{67}$$

where $\overline{\Theta}(\tau,t_0) \in \mathbb{R}^{q \times n}$ is given by

$$\overline{\Theta}(\tau,t_0) = \frac{1}{|\mathcal{X}|} \sum_{x_{per} \in \mathcal{X}} \sum_{i=1}^{n} \frac{1}{||x_{per}||_2^2} \Theta_{x_{per,i}}(\tau,t_0) \tag{68}$$

and $\Theta_{x_{per},i}(\tau,t_0)$ is defined by

$$\Theta_{x_{per},i}(\tau,t_0) = \left(y(\tau; x_{per,j}(t_0), u_{ref}(\tau)) - y(\tau; x_{ref}(t_0), u_{ref}(\tau))\right) \cdot (x_{per} \cdot e_i)^\top \tag{69}$$

with $y(\tau; x_{per,i}(t_0), u_{ref}(\tau))$ the state of the system corresponding to the initial value $x_{per,i}(t_0) = x_{ref}(t_0) + x_{per} \cdot e_i$ and e_i the ith standard unit vector in $\mathbb{R}^n$.

Lemma 4.4
Given an LTV system of the form $\dot{x}(t) = A(t)x(t) + B(t)u(t)$ and $y(t) = C(t)x(t)$ on the time interval $[t_0,t_1]$. Then, for any non-empty initial value perturbation set $\mathcal{X}$ eq. (41), a time-limited empirical observability gramian REF–$\mathcal{O}_\mathcal{X}(t_0,t_1)$ is identical to the time-limited analytical observability gramian $\mathcal{O}(t_0,t_1) = \int_{t_0}^{t_1} \left(\Phi^{\tau,t_0}\right)^\top C(\tau)^\top C(\tau) \Phi^{\tau,t_0} \mathrm{d}\tau$.

Proof: The initial state-perturbed solution for the LTV system is $x(t; x_{\mathrm{per},i}(t_0), u_{\mathrm{ref}}(t)) = \Phi^{t,t_0} x_{\mathrm{per},i}(t_0) + \int_{t_0}^{t} \Phi^{t,\tau} B(\tau) u_{\mathrm{ref}}(\tau) \mathrm{d}\tau$ with $x_{\mathrm{per},i}(t_0) = x_{\mathrm{ref}}(t_0) + x_{\mathrm{per}} \cdot e_i$.

Inserting the initial state-perturbed solution into eq. (69) gives

$$\begin{aligned}
\Theta_{x_{\mathrm{per}},i}(\tau, t_0) = &\left[C(\tau) \left(\Phi^{\tau,t_0} x_{\mathrm{per},i}(t_0) + \int_{t_0}^{\tau} \Phi^{\tau,s} B(s) u_{\mathrm{ref}}(s) \mathrm{d}s \right) \right. \\
&\left. - C(\tau) \left(\Phi^{\tau,t_0} x_{\mathrm{ref}}(t_0) + \int_{t_0}^{\tau} \Phi^{\tau,s} B(s) u_{\mathrm{ref}}(s) \mathrm{d}s \right) \right] \cdot (x_{\mathrm{per}} \cdot e_i)^\top \\
= &\left(C(\tau) \left(\Phi^{\tau,t_0} x_{\mathrm{per},i}(t_0) - \Phi^{\tau,t_0} x_{\mathrm{ref}}(t_0) \right) \right) \cdot (x_{\mathrm{per}} \cdot e_i)^\top \\
= &\left(C(\tau) \Phi^{\tau,t_0} x_{\mathrm{per}} \cdot e_i \right) \cdot (x_{\mathrm{per}} \cdot e_i)^\top
\end{aligned}$$

Inserting $\Theta_{x_{\mathrm{per}},i}(\tau, t_0)$ into eq. (68), $\overline{\Theta}(\tau, t_0)$ becomes

$$\overline{\Theta}(\tau, t_0) = \frac{1}{|\mathcal{X}|} \sum_{x_{\mathrm{per}} \in \mathcal{X}} \sum_{i=1}^{n} \frac{1}{||x_{\mathrm{per}}||_2^2} \left(C(\tau) \Phi^{\tau,t_0} x_{\mathrm{per}} \cdot e_i \right) \cdot (x_{\mathrm{per}} \cdot e_i)^\top .$$

Using $x_{\mathrm{per}} = c_m T_l$, we get

$$\begin{aligned}
\overline{\Theta}(\tau, t_0) &= \frac{1}{L \cdot M} \sum_{l,m} \sum_{i=1}^{n} \frac{1}{c_m^2} C(\tau) \Phi^{\tau,t_0} c_m T_l e_i e_i^\top T_l^\top c_m^\top \\
&= \frac{1}{L \cdot M} \sum_{l,m} \sum_{i=1}^{n} C(\tau) \Phi^{\tau,t_0} T_l e_i e_i^\top T_l^\top ,
\end{aligned} \tag{70}$$

where

$$\sum_{i=1}^{n} C(\tau) \Phi^{\tau,t_0} T_l e_i e_i^\top T_l^\top = C(\tau) \Phi^{\tau,t_0} T_l T_l^\top .$$

Thus, eq. (70) further simplifies to

$$\begin{aligned}
\overline{\Theta}(\tau, t_0) &= \frac{1}{L \cdot M} \sum_{l,m} C(\tau) \Phi^{\tau,t_0} \underbrace{T_l T_l^\top}_{=I_n} \\
&= \frac{1}{L \cdot M} \sum_{l,m} C(\tau) \Phi^{\tau,t_0} = C(\tau) \Phi^{\tau,t_0} .
\end{aligned} \tag{71}$$

Then, by inserting eq. (71) into eq. (67), a time-limited empirical observability gramian REF–$\mathcal{O}_{\mathcal{X}}(t_0, t_1)$ becomes

$$\mathrm{REF}\text{–}\mathcal{O}_{\mathcal{X}}(t_0, t_1) = \int_{t_0}^{t_1} \left(\Phi^{\tau,t_0} \right)^\top C(\tau)^\top C(\tau) \Phi^{\tau,t_0} \mathrm{d}\tau = \mathcal{O}(t_0, t_1) ,$$

which is the desired result. □

In Table 4 the conditions on the LTI or LTV system are summarised in which the novel time-limited empirical gramians are identical to the time-limited analytical gramians.

As described in section 3.2.3 the infinite time empirical gramians [70, 19, 42] defined in 3.15,

3.17, 3.19, 3.20 and 3.21 are equal to the infinite time analytical gramians for stable LTI or LTV system with initial state equal to zero or a particular steady state of the system, respectively. In the numerical approximation of these infinite time empirical gramians, the calculation time interval needs to be chosen 'large' enough to ensure that the perturbed system trajectory has reached a steady state. However, there does not exist an explicit definition or mathematical justification as to when the chosen time interval is large enough. As a consequence the numerical approximation of the infinite time empirical gramians is dependent on both the pre-defined perturbations as well as the chosen time interval $[0, t]$. Additionally, if the system has multiple steady states, the perturbation sets eq. (40) and (41) need to be chosen such that the perturbed system trajectory does not leave the region of attraction of the specified equilibrium point [19].

In contrast to the infinite time empirical gramians, our time-limited extension of the empirical gramians allows for a characterisation of the transient system behaviour. This is explicit of interest in the context of systems pharmacology/biology models which will be considered in chapter 5. However, the characterisation of transient system behaviour is also of relevance in the context of control theory, as previously described in section 3.2.1. In the subsequent section, we will, thus, consider a benchmark example for model order reduction techniques to demonstrate the usefulness of the novel formulation over the previously defined empirical gramians in section 3.2.1.

Table 4: **Conditions on an LTI system or an LTV system under which the time-limited empirical gramians in eqs.** (62) **and** (67) **are identical to the time-limited analytical gramians in eqs.** (35) **and** (36)**.**

Properties of the LTI system given as in eq. (21)			REF–$\mathcal{C}_{\mathcal{U}}(t_0,t_1) = \mathcal{C}(t_0,t_1)$, REF–$\mathcal{O}_{\mathcal{X}}(t_0,t_1) = \mathcal{O}(t_0,t_1)$
stable with $x_0 = x_{\text{ss}}$	finite time	$x_0 = 0$	yes
		$x_0 \neq 0$	yes
	infinite time	$x_0 = 0$	yes
		$x_0 \neq 0$	yes
stable with $x_0 \neq x_{\text{ss}}$	finite time	$x_0 = 0$	yes
		$x_0 \neq 0$	yes
	infinite time	$x_0 = 0$	yes
		$x_0 \neq 0$	yes
unstable	finite time	$x_0 = 0$	yes
		$x_0 \neq 0$	yes

Properties of the LTV system given as in eq. (34)			REF–$\mathcal{C}_{\mathcal{U}}(t_0,t_1) = \mathcal{C}(t_0,t_1)$, REF–$\mathcal{O}_{\mathcal{X}}(t_0,t_1) = \mathcal{O}(t_0,t_1)$
stable with $x_{\text{ss}} = x_0$	finite time	$x_0 = 0$	yes
		$x_0 \neq 0$	yes
	infinite time	$x_0 = 0$	yes
		$x_0 \neq 0$	yes
stable with $x_{\text{ss}} \neq x_0$	finite time	$x_0 = 0$	yes
		$x_0 \neq 0$	yes
	infinite time	$x_0 = 0$	yes
		$x_0 \neq 0$	yes
unstable	finite time	$x_0 = 0$	yes
		$x_0 \neq 0$	yes

4.2. Numerical example: resistor-capacitor (RC) ladder

The nonlinear resistor-capacitor (RC) ladder system is often employed to test model order reduction techniques for nonlinear systems [15, 18, 104]. It is an electronic circuit consisting of a parallel connected resistor with a diode. For a large number of nodes, the simulation of the behaviour of this system becomes computationally demanding [104]. Furthermore, the series of nonlinear resistors produces a strong global nonlinearity, which makes the system a challenging test system for model order reduction techniques [18].

The systems equations of the RC ladder are given by

$$\begin{aligned} \dot{x}(t) &= f(x(t)) + g(x(t))u(t), \quad x(t_0) = 0 \\ y(t) &= h(x(t)) \,, \end{aligned}$$

where

$$f(x(t)) = \begin{pmatrix} -2 & 1 & & & \\ 1 & -2 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & 1 & -2 & 1 \\ & & & 1 & -1 \end{pmatrix} x(t) + \begin{pmatrix} 2 - \exp(40x_1(t)) - \exp(40(x_1(t) - x_2(t))) \\ \exp(40(x_1(t) - x_2(t))) - \exp(40(x_2(t) - x_3(t))) \\ \vdots \\ \exp(40(x_{N-2}(t) - x_{N-1}(t))) - \exp(40(x_{N-1}(t) - x_N(t))) \\ \exp(40(x_{N-1}(t) - x_N(t))) - 1 \end{pmatrix},$$

and $g(x(t)) = B$, $h(x(t)) = Cx(t)$ and $C = B^T = (1, 0, \ldots, 0) \in \mathbb{R}^N$ with N denoting the number of nodes and x_i denoting the voltage at node i [19]. This nonlinear system is asymptotically stable around zero if no input is applied, i.e. $x_{\text{ss}} = 0$.

In order to compare our novel empirical gramians against the infinite time empirical gramians in definition 3.15 and 3.17, we have chosen the number of nodes to be $N = 64$ and the time interval of interest to be $t \in [0, 1]$ (as in [48]). Based on the empirical controllability and observability gramian a nonlinear system can be reduced by balanced truncation (as outlined in section 3.2.2). Therefore, we will use balanced truncation to show the benefits of our novel empirical gramians formulation. In contrast to linear systems, there does not exist an a-priori error bound for balanced truncation for nonlinear systems. We have chosen the following approximation error ε_{rel}

$$\varepsilon_{\text{rel}} = \frac{\left(\int_0^{t_{\text{end}}} (y_{\text{ref}}(t) - y_{\text{red}}(t))^2 \mathrm{d}t \right)^{\frac{1}{2}}}{\left(\int_0^{t_{\text{end}}} y_{\text{ref}}(t)^2 \mathrm{d}t \right)^{\frac{1}{2}}}. \tag{72}$$

We compare our new approach, the introduced generalised empirical gramians (cf. definition 4.1 and 4.3) with subsequent model reduction by balanced truncation, with two others: (i) analytical gramians based on linearisation of the nonlinear system in the steady state, (ii) infinite time empirical gramians and subsequent balanced truncation. The resulting orders of the reduced models for all three approaches for varying error tolerance ε_{rel} are given in Table 5.

The comparison of the obtained order of the reduced system based on the different user-defined relative error tolerance (cf. Table 5) it becomes evident that for every error tolerance the order of reduced model obtained by balanced truncation based on the time-limited empirical is lower than for the two other approaches, e.g. for $\varepsilon_{\text{rel}} = 0.1$ order 6 versus 12 and 36. This demonstrates the superiority of the novel empirical gramians for this numerical example.

In this section we have shown that the novel generalised empirical gramian formulation can be used in the context of balanced truncation for nonlinear systems exemplified for a benchmark

Table 5: **Order of reduced models obtained by balanced truncation based on different controllability and observability gramians for varying relative error tolerance.** The resulting orders of the reduced model obtained by balanced truncation based on (i) the analytical gramians calculated for the steady state of the nonlinear system (linearisation in steady state), (ii) the infinite time empirical gramian formulation in definition 3.15 and 3.17 and (iii) the time-limited empirical gramians.

	Dimension of the reduced model obtained by BT based on		
error tolerance	**analytical gramians calculated for the steady state of the nonlinear system**	**Lall**–$\mathcal{C}_{\mathcal{U}}(\infty)$ **Lall**–$\mathcal{O}_{\mathcal{X}}(\infty)$	**REF**–$\mathcal{C}_{\mathcal{U}}(t_0, t_1)$ **REF**–$\mathcal{O}_{\mathcal{X}}(t_0, t_1)$
0.1	12	36	6
0.01	61	51	19
0.001	64	64	31

model from the field of control theory, furthermore they can be used to calculate approximations of the analytical gramians in the case where one cannot solve the Lyapunov equations (e.g. unstable system, uncontrollable/ unobservable system parts).

The drawback of balanced truncation in the context of systems pharmacology/biology models has already been discussed in section 3.3. Balanced truncation might be applied to systems pharmacology/biology models, but then the interpretability of the system components can be lost. We will use the novel empirical gramians and balanced truncation for the PT test setting in section 5.3 to compare the method with other model order reduction methods. If the requirement of interpretability of the reduced model components can be neglected, balanced truncation can be used to obtain a reduced system for systems pharmacology/biology models.

The goal of this thesis is to develop a model reduction technique which retains the physiological interpretability of the reduced system. To that end, we will introduce in the next section so-called input-response indices, which are based on the novel time-limited empirical gramians introduced in this section.

4.3. Derivation of input-response indices

In a typical pharmacology/biology setting, $u(t)$ represents some drug administration (including i.v. bolus, i.v. infusion and p.o. administration) or some other stimulus of the system (e.g. the release of the tissue factor as in section 2.3). In the case of single dosing most inputs to the system are, thus, at $t = 0$. Therefore, we will focus in the following section on the case that $u(t) = \delta(t) = u_0$, such that the system in equ. (61) simplifies to

$$\frac{\mathrm{d}x}{\mathrm{d}t}(t) = f\big(x(t)\big), \quad x(0) = x_0 + u_0\,, \tag{73}$$

$$y(t) = h\big(x(t)\big). \tag{74}$$

The output of interest y is then either the concentration-time profile of a specific state variable, its associated nadir concentration or the time point at which the concentration of a specific state variable reaches a threshold. The initial condition $x(0)$ comprises two parts, an input-independent part x_0 and the input u_0. The input-independent part x_0 characterises the state of the system prior to the input, which typically corresponds to some steady state. The input-dependent part u_0 is chosen such that $u_{0,i} \neq 0$ only for every ith state variable that is associated with the input. However, how the concept can be extended to account also for time-dependent inputs, like multiple dosing, will be discussed in section 6.

In the following sections, we will drop the explicit dependence of the solution trajectory of the differential equation with respect to the initial value and input function in the notation for simplicity.

As in section 3.2.1, the operator $\Phi^{t,0}$ (i.e. the state transition matrix) maps the initial condition $x(0)$ to the solution $x(t)$ of the differential equation (73) at time $t > 0$, thus

$$x(t) = \Phi^{t,0}(x_0 + u_0). \tag{75}$$

We are ultimately interested in the dynamics of the output y as a response to the input u_0 on the time interval $[0, t_{\mathrm{end}}]$ with $t_{\mathrm{end}} > 0$. Using the evolution operator Φ, we obtain

$$y(t) = h\big(\Phi^{t,0}(x_0 + u_0)\big),$$

which makes the dependence of the output y on the input u_0 explicit.

The input-response indices are based on two constituents that characterise

1. to what extent the input at $t = 0$ has an impact on the state variable x_i at some time $t^* \in [0, t_{\mathrm{end}}]$, and
2. to what extent variations in x_i at time t^* impact the output y on the remaining time interval $[t^*, t_{\mathrm{end}}]$,

In control theory, the first part is associated with the time-bounded controllability of x_i by the input u_0, while the second part is associated with the time-bounded observability of x_i through the output y; cf. Figure 12 for illustration. In contrast to existing approaches in control theory, however, we combined time-bounded controllability and observability in a

novel way to quantify the overall impact of a state variable on the input-output relationship. For an efficient numerical realisation, we borrowed ideas from the concept of the time-limited empirical gramians (introduced in section 4.1).

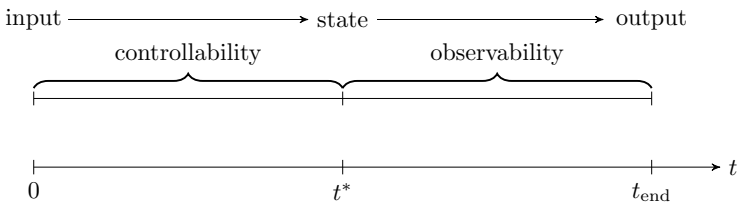


Figure 12: **Schematic representing the consideration of observability and controllability of the system**. Controllability characterises how the input affects the state variables x_i and is considered from 0 to t^*. Observability then quantifies how the state variables impact the output and this is considered on the interval $[t^*, t_{\text{end}}]$.

We define the importance of state variables relative to some reference trajectory x_{ref}, defined as the solution of eq. (73) with initial value $x_{\text{ref}}(0) = x_0 + u_{\text{ref}}$, where u_{ref} denotes a reference input. Using the evolution operator defined in eq. (75), it is

$$x_{\text{ref}}(t) = \Phi^{t,0}(x_0 + u_{\text{ref}}), \quad t \in [0, t_{\text{end}}]. \tag{76}$$

To quantify the extent to which the input has an impact on a given state variable x_i at time $t^* \in [0, t_{\text{end}}]$, we considered perturbations Δu of the reference input u_{ref} at time $t = 0$ and determined the resulting perturbation of x_i at time t^*. In other words, it is determined how the perturbation of the input propagates through the system to affect the different state variables x_i. To this end, we denoted by $x_{\Delta u}$ the solution of the system based on the input perturbation Δu, i.e.

$$x_{\Delta u}(t^*) = \Phi^{t^*,0}(x_0 + u_{\text{ref}} + \Delta u)\,.$$

Note that $x_{\Delta u}(t^*)$ is a vector of state variables, with the ith component $(x_{\Delta u}(t^*))_i$ referring to state variable x_i at time t^*. Thus, a perturbation Δu of the input u_{ref} results in a perturbation

$$\Delta x_i(t^*) = \left(x_{\Delta u}(t^*) - x_{\text{ref}}(t^*)\right)_i \tag{77}$$

of x_i at time t^* relative to the reference solution x_{ref}. The larger $\Delta x_i(t^*)$, the larger the impact of the input on the state variable x_i at time t^*. If the input has no impact on the state variable x_i, then $\Delta x_i(t^*) = 0$ for all $t^* > 0$, since the ith component of the perturbed trajectory x_u will be identical to the ith component of the reference trajectory x_{ref}. The size of the perturbation Δx_i, however, does not allow to infer anything about the relevance of a state variable for the input-response relationship. For example, if x_i does not affect the output y at all, it will be considered as unimportant for the input-response relationship, independent of the size of Δx_i. Therefore, we next quantified the impact of the perturbation $\Delta x_i(t^*)$ on the output y during the remaining time interval $[t^*, t_{\text{end}}]$. To this end, the reference output

is determined

$$y_{\text{ref}}(t) = h(x_{\text{ref}}(t)), \quad t \in [t^*, t_{\text{end}}]\,,$$

and the specific perturbation of the ith component is defined as

$$\Delta u_i(t^*) = (0, \ldots, 0, \underbrace{\Delta x_i(t^*)}_{i\text{th entry}}, 0, \ldots, 0)^\top \in \mathbb{R}^{i-1} \times \mathbb{R} \times \mathbb{R}^{n-(i+1)}\,.$$

This allowed us to define the output $y_{\Delta u_i(t^*)}$ based on the perturbation of the ith component of the reference trajectory x_{ref} at time t^*

$$y_{\Delta u_i(t^*)}(t) = h\left(\Phi^{t,t^*}(x_{\text{ref}}(t^*) + \Delta u_i(t^*))\right)$$

for $t \in [t^*, t_{\text{end}}]$. In other words, $y_{\Delta u_i(t^*)}$ is the output that results from a perturbation $\Delta x_i(t^*)$ of the ith state variable of the reference trajectory x_{ref} at time t^* (cf. Figure 13 for graphical illustration). This perturbation is specific to the state variable x_i and is the result of a perturbation of the input u_{ref} by Δu at $t = 0$.

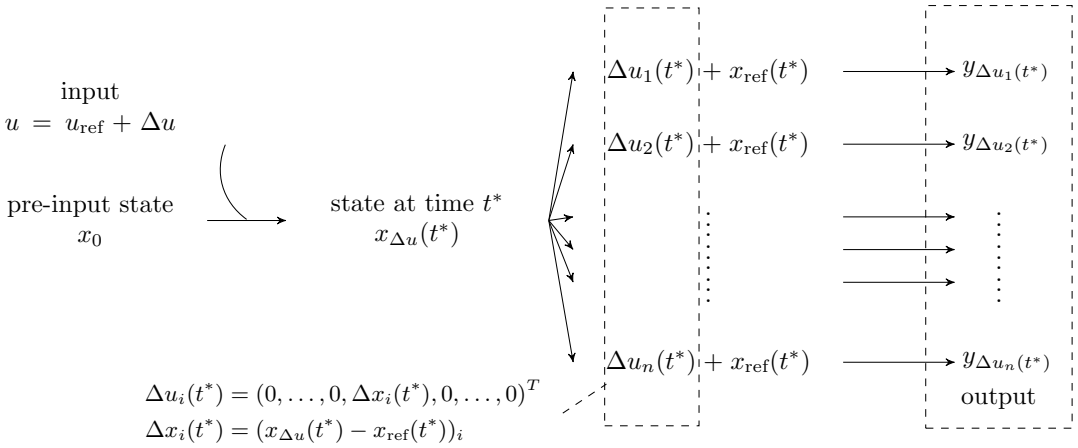


Figure 13: **Schematic representing of the derivation of the input-response indices for one input perturbation Δu**

To quantify the impact of the input $u_{\text{ref}} + \Delta u$ on the response y via x_i at time t^*, we determined the squared absolut difference between the reference response y_{ref} and the resulting perturbed response $y_{u_i(t^*)}$, integrated over the remaining time interval $[t^*, t_{\text{end}}]$

$$\int_{t^*}^{t_{\text{end}}} \left| y_{\Delta u_i(t^*)}(t) - y_{\text{ref}}(t) \right|^2 \mathrm{d}t\,. \tag{78}$$

To remove the dependence on a particular perturbation, the average over a number of initial perturbations

$$\mathcal{U} = \{\Delta u_1, \ldots, \Delta u_N\}$$

of the reference input is determined. The input perturbations Δu are chosen to reflect the expected variability of the input and defined as in eq. (40) with $s = n$, $i \in I_u$ and I_u the

indices of the state variables associated with non-zero input. This resulted in the following definition:

Definition 4.5 (input-response (ir-) index)
For the nonlinear system in eqs. (73) *and* (74) *with reference solution* $x_{ref}(t)$ *and time interval* $[0, t_{\text{end}}]$ *of interest, the input-response index* ir_i *of the* i*th state component at time* t^* *are defined by*

$$\text{ir}_i(t^*) = \left(\frac{1}{N} \sum_{\Delta u \in \mathcal{U}} \frac{1}{Z_u} \int_{t^*}^{t_{\text{end}}} \left| y_{\Delta u_i(t^*)}(t) - y_{\text{ref}}(t) \right|^2 \mathrm{d}t \right)^{\frac{1}{2}}, \tag{79}$$

where Z_u *denotes a normalisation constant.*

We will use $Z_u = t_{\text{end}}^2 \cdot \|(\Delta u)\|_2^2$, but alternative constants could be used. The input-response index ir_i is a measure of the importance of the ith state variable to the input-output behaviour of the system as a function of time. A large value of $\text{ir}_i(t^*)$ indicates that the impact of the input u on x_i at time t^* impacts the output y on the time interval $[t^*, t_{\text{end}}]$ in a substantial way. It is a distinct feature of the input-response index that it combines both, the impact of the input on the state variables as well as the impact of the state variables on the output. Neither one nor the other impact on its own is informative for the relevance of a state variable x_i for the input-response relationship. Since the input-response index is a function of time, it not only characterises whether a state variable is important but also when. For example, a state variable might have a large impact on the output at some early time point, but a minor impact at much later times. This will be illustrated for the blood coagulation system in section 5.1.

Note that as for the time-limited empirical gramians in section 4.1 the ir-indices for nonlinear systems are dependent on the set of chosen input perturbations $\mathcal{U}$. It can, however, be shown that the ir-indices for LTI systems with a single input ($B \in \mathbb{R}^{n \times 1}$) and a single output $y(t) \in \mathbb{R}$, are independent of the magnitude of the chosen input perturbations (for any non-empty perturbation set $\mathcal{U}$).

As before, the ir-indices are defined with respect to the reference trajectory x_{ref}

$$x_{\text{ref}}(t) = e^{A(t-t_0)}x_0 + \int_{t_0}^{t} e^{A(t-s)}u_{\text{ref}}(s)ds.$$

However, since $u_{\text{ref}}(t) = B\delta(t)$, the reference solution $x_{\text{ref}}(t)$ becomes

$$x_{\text{ref}}(t) = e^{A(t-t_0)}(x_0 + B)\,.$$

To quantify the extend to which an input impacts a state, as in section 4.3, an input perturbation Δu is applied. In this case, $\Delta u = Bu_{\text{per}}$ with $u_{\text{per}} \in \mathcal{U}$ as defined in eq. (40) to enable the clear link to previously introduced control-theoretic concepts. The resulting perturbed solution is denoted by $x_{\Delta u}(t^*)$

$$x_{\Delta u}(t^*) = e^{A(t^*-t_0)}(x_0 + B + \Delta u).$$

Thus, the perturbation $\Delta x_i(t^*)$ on each state variable induced by the input perturbation becomes

$$\Delta x_i(t^*) = \left(x_{\Delta u}(t^*) - x_{\text{ref}}(t^*)\right)_i = \left(e^{A(t^*-t_0)}\Delta u\right)_i ,$$

which characterises the controllability of the system. The input into the observability consideration of the system at time point $x(t^*)$ is then

$$\begin{aligned}\Delta u_i(t^*) &= (0,\dots,0,\underbrace{\Delta x_i(t^*)}_{i\text{th entry}},0,\dots,0)^\top \\ &= (0,\dots,0,\left(e^{A(t^*-t_0)}\Delta u\right)_i,0,\dots,0)^\top .\end{aligned}$$

We define for the linear system a projection Π_i selecting the ith component of the state vector

$$\begin{aligned}\Pi_i : \mathbb{R}^n &\to \mathbb{R}^n \\ x &\mapsto (0,\dots,0,x_i,0,\dots,0)^\top ,\end{aligned}$$

such that we can write $u_i(t^*)$ as

$$\Delta u_i(t^*) = \Pi_i e^{A(t^*-t_0)}\Delta u .$$

The next step is to quantify the impact of the perturbation $\Delta x_i(t^*)$ on the output y. To that end we first determine the reference output y_{ref} by

$$y_{\text{ref}}(t) = Ce^{A(t-t^*)}\left(e^{A(t^*-t_0)}(x_0+B)\right).$$

The output $y_{\Delta u_i(t^*)}$ based on a perturbation of the ith component of the state vector at time t^* is given by

$$y_{\Delta u_i(t^*)}(t) = Ce^{A(t-t^*)}\left(e^{A(t^*-t_0)}(x_0+B) + \Pi_i e^{A(t^*-t_0)}\Delta u\right).$$

Finally inserting y_{ref} and $y_{\Delta u_i(t^*)}$ into eq. (79) the input-response indices for the LTI system are given by

$$\begin{aligned}\text{ir}_i(t^*) &= \left(\frac{1}{N}\sum_{u_{\text{per}}\in\mathcal{U}}\frac{1}{||u_{\text{per}}||_2^2}\int_{t^*}^{t_{\text{end}}}\left|Ce^{A(t-t^*)}\Pi_i e^{A(t^*-t_0)}\Delta u\right|^2 \text{d}t\right)^{\frac{1}{2}} , \\ &= \left(\frac{1}{N}\sum_{u_{\text{per}}\in\mathcal{U}}\int_{t^*}^{t_{\text{end}}}\left|Ce^{A(t-t^*)}\Pi_i e^{A(t^*-t_0)}B\frac{u_{\text{per}}}{||u_{\text{per}}||_2^2}\right|^2 \text{d}t\right)^{\frac{1}{2}} .\end{aligned}$$

Using the initial assumption, that we are considering a single input system, we obtain

$$\begin{aligned}\mathrm{ir}_i(t^*) &= \left(\frac{1}{N} \sum_{u_{\mathrm{per}} \in \mathcal{U}} \int_{t^*}^{t_{\mathrm{end}}} \left| Ce^{A(t-t^*)} \Pi_i e^{A(t^*-t_0)} B \right|^2 \underbrace{\frac{||u_{\mathrm{per}}||_2^2}{||u_{\mathrm{per}}||_2^2}}_{=1} \mathrm{d}t \right)^{\frac{1}{2}} \\ &= \left(\int_{t^*}^{t_{\mathrm{end}}} \left| Ce^{A(t-t^*)} \Pi_i e^{A(t^*-t_0)} B \right|^2 \mathrm{d}t \right)^{\frac{1}{2}}\end{aligned} \tag{80}$$

Now, we can, furthermore, show how the input-response indices can be expressed in terms of the input-to-state and state-to-output map introduced in section 3.2.1. To that end, we define the operator inside the integral by

$$\begin{aligned}&\tilde{\mathcal{H}} : \mathbb{R} \to L_2(t^*, t_{\mathrm{end}}, \mathbb{R}), \\ &y(t) = \tilde{\mathcal{H}}(u_0)(t) = Ce^{A(t-t^*)} \Pi_i e^{A(t^*-t_0)} B u_0.\end{aligned} \tag{81}$$

Using the input-to-state $\Psi_{u\to x}(t)$ and state-to-output maps $\Psi_{x\to y}(t)$ defined as follows

$$\begin{aligned}&\Psi_{u\to x}(t) : \mathbb{R}^1 \to \mathbb{R}^n \\ &x(t) = \Psi_{u\to x}(t) u = e^{At} B u \,,\end{aligned}$$

and

$$\begin{aligned}&\Psi_{x\to y}(t) : \mathbb{R}^n \to \mathbb{R}^1 \\ &y(t) = \Psi_{x\to y}(t) x = Ce^{At} x \,,\end{aligned}$$

the operator $\tilde{\mathcal{H}}$ becomes

$$\tilde{\mathcal{H}}(u_0)(t) = \Psi_{x\to y}(t - t^*) \Pi_i \Psi_{u\to x}(t^* - t_0) u_0 \,.$$

This operator maps, similar to the Hankel operator, past inputs into future outputs. Such that we can write the eq. (80) as

$$\begin{aligned}\mathrm{ir}_i(t^*) &= \left(\int_{t^*}^{t_{\mathrm{end}}} \left| \Psi_{x\to y}(t - t^*) \Pi_i \Psi_{u\to x}(t^* - t_0) \right|^2 \mathrm{d}t \right)^{\frac{1}{2}} = \left(\int_{t^*}^{t_{\mathrm{end}}} \tilde{\mathcal{H}} \tilde{\mathcal{H}}^\top \mathrm{d}t \right)^{\frac{1}{2}} \\ &= \left(\int_{t^*}^{t_{\mathrm{end}}} \Psi_{x\to y}(t - t^*) \Pi_i \Psi_{u\to x}(t^* - t_0) \left(\Psi_{x\to y}(t - t^*) \Pi_i \Psi_{u\to x}(t^* - t_0) \right)^\top \mathrm{d}t \right)^{\frac{1}{2}} \\ &= \left(\int_{t^*}^{t_{\mathrm{end}}} \Psi_{x\to y}(t - t^*) \underbrace{\Pi_i \Psi_{u\to x}(t^* - t_0) \Psi_{u\to x}(t^* - t_0)^\top \Pi_i}_{ii\text{th entry of squared impulse response at time } t^*} \Psi_{x\to y}(t - t^*)^\top \mathrm{d}t \right)^{\frac{1}{2}}\end{aligned}$$

We have, thus, shown that for the LTI system (single input-single output), the input-response indices are independent of the magnitude of the perturbation. The indices were linked to the impulse response of the system. In definition 3.10 the infinite time analytical cross

gramian was given. The infinite time cross gramian gives a measure for controllability and observability of a system in a single matrix. However, in contrast to the infinite time cross gramian the input-response indices quantify the impact of each state variable on the input-output relationship separately. Thus, the indices provide for linear systems a novel insight into the importance of a state variable for the impulse-response of the system.

4.4. Derivation of the sensitivity based input-response indices

We will show in this section that the sensitivity based ir-indices can be given as a product of two sensitivity coefficients over time. This allows for a clear distinction between the extent to which the input impacts a state variable, a state variable impacts the output. In addition, it allows for a computationally more efficient calculation. We will subsequently show how the concept of sensitivity coefficients in the context of our newly introduced sensitivity based ir-indices can be leveraged for model order reduction.

To quantify the extent to which the input has an impact on the state variables, we again consider a perturbation Δu of the input u_{ref} and determine the resulting perturbation of the state variables x_i at time t^*. We, thus, determine how the perturbation in the input is propagated through the system. We denote by $x_{\Delta u}$ the solution of the system with input perturbation Δu

$$x_{\Delta u}(t^*) = \Phi^{t^*,0}(x_0 + u_{\text{ref}} + \Delta u).$$

Thus, a perturbation Δu results in a perturbation of the ith state variable at time t^* of size

$$\Delta x_i(t^*) = \left(x_{\Delta u}(t^*) - x_{\text{ref}}(t^*)\right)_i$$

with x_{ref} the reference trajectory as defined in eq. (76). Subsequently, as in section 4.3, we derive a quantity that measures how large the impact of the input perturbation is on each state variable at time point t^*. In contrast to the previous calculation via multiple perturbations, we consider the first order perturbation of the state variables at time point t^* with respect to the input. This impact of the input on the state variables is given by the ith row of the Jacobian

$$J_u(t^*, t_0) = \frac{\partial}{\partial u}\Phi^{t^*,t_0}(x_0 + u)\Big|_{u=u_{\text{ref}}} = \frac{\partial x_{\text{ref}}(t^*)}{\partial u}.$$

Whenever $[J_u(t^*, t_0)]_{i,j}$ is large, the jth input u_j has a large impact on the ith state variable x_i at time t^*. If $[J_u(t^*, t_0)]_{i.j} = 0$, then the jth input has no impact on the ith state variable at time t^*. As in section 4.3, the quantification of the impact on the state variable alone, however, is not sufficient to infer the importance of state variables for the input-response relationship.

As a second step we quantified the extent, to which a perturbation of the ith state variable at t^* impacts the output y on the remaining interval. Rather than using equal or normalised perturbations like in eq. (18), we used the input-induced perturbation $\Delta x_i(t^*)$. To this end,

we reinterpreted this perturbation as some input

$$\Delta u_i(t^*) = (0, \ldots, 0, \Delta x_i(t^*), 0, \ldots, 0)^\top \in \mathbb{R}^{i-1} \times \mathbb{R} \times \mathbb{R}^{n-(i+1)} ,$$

of the ODE (73) at time $t_0 = t^*$. The input-independent part x_0 was chosen as the reference trajectory x_{ref} at t^*. Thus, we considered the ODE (73) on $[t^*, t_{\text{end}}]$ with initial condition given by $x(t^*) = x_{\text{ref}}(t^*) + \Delta u_i(t^*)$. This results for $t \geq t^*$ in the output

$$y_{\Delta u_i(t^*)}(t) = h\left(\Phi^{t,t^*}\left(x_{\text{ref}}(t^*) + \Delta u_i(t^*)\right)\right) ,$$

while the reference output is

$$y_{\text{ref}}(t) = h\left(\Phi^{t,t^*}\left(x_{\text{ref}}(t^*)\right)\right) = h\left(\Phi^{t,t_0}\left(x_0 + u_{\text{ref}}\right)\right) .$$

Using first order perturbations, this impact was quantified by the Jacobian

$$J_y(t, t^*) = \left.\frac{\partial h\left(\Phi^{t,t^*} x\right)}{\partial x}\right|_{x = x_{\text{ref}}(t^*)} .$$

We finally define the input-response index $\text{ir}_i(t^*)$ of the ith state variable at time t^* as the (component-wise squared entries of the) product of the two Jacobian's $J_u(t^*, t_0)$ and $J_y(t, t^*)$, integrated over the remaining time span $[t^*, t_{\text{end}}]$

Definition 4.6 (sensitivity based input-response (ir-) index)
For the nonlinear system in eqs. (73) *and* (74) *with reference solution $x_{ref}(t)$ and time interval $[0, t_{\text{end}}]$ of interest, the sensitivity based input-response indices of the ith state variable for the jth input and kth output are defined by*

$$[\text{ir}_i(t^*)]^2_{kj} = \frac{1}{t_{\text{end}}} \int_{t^*}^{t_{\text{end}}} \left([J_y(t, t^*)]_{k,i} \; [J_u(t^*, t_0)]_{i,j} \right)^2 \mathrm{d}t \, . \tag{82}$$

In the special case of a single-state input (u_0 has only a single non-zero entry, e.g. $u_j = x_j$) and the output being identical to a (different) state variable x_o (i.e. $y(t) = h(x(t)) = x_o(t) \in \mathbb{R}$), the input-response index ir_i is real-valued and given by

$$\text{ir}_i(t^*) = \left(\frac{1}{t_{\text{end}}} \int_{t^*}^{t_{\text{end}}} \mathcal{S}_{o,i}(t, t^*)^2 \mathrm{d}t \right)^{\frac{1}{2}} \cdot |\mathcal{S}_{i,j}(t^*, t_0)| \tag{83}$$

with $|z|$ doting the modulus of $z \in \mathbb{R}^n$ and

$$\mathcal{S}_{m,n}(t, s) = \left.\frac{\partial \left[\Phi^{t,s} x\right]_m}{\partial x_n}\right|_{x = x_{\text{ref}}(s)} .$$

As before, there is also a link to control theory, where the first factor quantifies the observability and the second factor the controllability of the state x_i. The eq. (83) is a local variant of eq. (79).

Since the sensitivity based input-response indices are stated in terms of the product of two

sensitivity coefficients, they allow for a distinction of effects. The second factor of the product can be interpreted in a control theoretical context as some controllability index

$$c_i(t^*) = |\mathcal{S}_{i,j}(t^*, t_0)|\,,$$

while the first factor can be interpreted as some observability index

$$o_i(t^*) = \frac{1}{t_{\text{end}}} \left(\int_{t^*}^{t_{\text{end}}} \mathcal{S}_{o,i}(t, t^*)^2 \mathrm{d}t \right)^{\frac{1}{2}}.$$

These additional indices provide further insight in the case, when an input-response index is small. For example, a state variable can have a large controllability index, in which case perturbations in the input have a large impact on this state variables. If, however, at the same time the observability index is very small or even zero, the input-response index might be small or zero. See section 5.2.2 for an application to the EGFR system. Note that all three indices ir_i, c_i and o_i are unitless.

In [117], alternative control-theoretical indices are presented based on so-called empirical observability and controllability gramians. As the authors state themselves, however, 'they are somewhat limited in their usefulness' [117, p.15]. Another approach is metabolic control analysis (MCA) [52] characterising how state variables influence the system *in steady state*. Here, flux control coefficients, quantifying how changes in the state variable effect the flux, and concentration control coefficients, quantifying the sensitivity of the response of the system with respect to the state variables, are defined. MCA has proven useful in elucidating rate-controlling steps in enzymatic reaction networks [119, Chap. 11]. Both MCA coefficients have little predictive value for the transient behaviour of the system since they are defined only in the steady state of a system [119, Chap. 11].

4.5. Model order reduction method based on input-response indices

In the following section, we will introduce a model order reduction technique that allows for a mechanistic interpretation in terms of the quantities of the original system. To that end we combine various aspects of the model reduction techniques introduced in section 3: namely the elimination of the systems state variables based on non-important part as in balanced truncation, the consideration of fast and slow dynamics from time scale separation that leads to the quasi-steady-state approximation, as well as the aspect that state variables are not dynamically important and can be considered as constants (as it is done in model reduction techniques based on sensitivity analysis). Additionally, we will exploit conservation laws if they are present in the model system.

For a given input-response relationship, reference trajectory and time interval of interest, we used the following classification of state variables to subsequently reduce the complexity of the system:

(a) environmental state variables x_{env}: considered to be constant on the time interval of interest. We assumed environmental state variables to be identical to their initial value.

(b) negligible state variables x_{neg}: considered unimportant for the given setting and set to zero, thereby neglected in the model for the specific input-output behaviour of interest.

(c) state variables in quasi-steady state x_{qss}: assumed to evolve on a fast time scale than the time scale of the response such that they can be determined by the well-known quasi-steady state approximation.

(d) conserved quantities x_{con}: eliminated by exploiting a conservation law.

(e) dynamical state variables x_{dyn}: all remaining state variables

In the reduced model, the dynamic state variables are characterised by differential equations (as in the original model). The environmental state variables enter differential equations as covariates, while the negligible state variables completely disappear from the system. The state variables in quasi-steady state are represented by algebraic equations, as are the conserved state variables.

The key problem is to decide for a given state variable x_i, whether it is an environmental, negligible, quasi-steady state, conserved or dynamic state variable. Since for large-scale systems, it is prohibitive to test all possible combinations, iterative approaches are the methods of choice. To this end, the input-response indices (defined in eq. (79) and (83)) will play a crucial role.

For illustration, consider the following reaction scheme:

$$A + E \underset{k_{\text{d}}}{\overset{k_{\text{a}}}{\rightleftharpoons}} E^* \xrightarrow{k_{\text{in}}} E_{\text{i}}; \qquad M + E^* \xrightarrow{k_{\text{on}}} M^* + E^*; \qquad M^* \xrightarrow{k_{\text{off}}} M$$

of a simple enzyme-mediated activation pathway. An activator A that binds to an enzyme E with rate constant k_{a} in [1/min/nM]. The resulting complex E*, the active enzyme, can either dissociate or be inactivated to E_{i} with rate constants k_{d} and k_{in} in [1/min], respectively. The activated enzyme transforms a protein M to become M* (an active form of M) with rate constant k_{on} in [1/min/nM], while M* turns back to its inactive form M with rate constant k_{off} in [1/min].

The resulting system of six ODEs for the rate of change of the corresponding concentrations is then

$$\begin{aligned} \frac{\mathrm{d}A}{\mathrm{d}t} = \frac{\mathrm{d}E}{\mathrm{d}t} &= -k_{\text{a}}AE + k_{\text{d}}E^* \\ \frac{\mathrm{d}E^*}{\mathrm{d}t} &= k_{\text{a}}AE - (k_{\text{d}} + k_{\text{in}})E^* \\ \frac{\mathrm{d}E_{\text{i}}}{\mathrm{d}t} &= k_{\text{in}}E^* \\ -\frac{\mathrm{d}M}{\mathrm{d}t} = \frac{\mathrm{d}M^*}{\mathrm{d}t} &= k_{\text{on}}E^*M - k_{\text{off}}M^*\,, \end{aligned}$$

with initial conditions $A_0 > 0$, $E_0 > 0$, $E_0^* = 0$, $M_0 > 0$ and $M_0^* = 0$. Let us consider A as input and M* as output. For illustration, we assume that $A_0 \gg E_0$ and that E^* is in quasi-steady state because the complex formation usually happens on a much faster time scale then the output dynamic in this case of M*. In the chosen setting $A(t) \approx A_0$ because the concentration of A_0 is assumed to be much larger then the enzyme concentration E_0. Then,

the activator is classified as an environmental state variable with $A_{\mathrm{env}} \equiv A_0$; the complex E^* as a quasi-steady state variable with

$$E^*_{\mathrm{qss}} = \frac{k_{\mathrm{a}} A_0 E}{k_{\mathrm{d}} + k_{\mathrm{in}}}\,. \tag{84}$$

The enzyme is kept as a dynamic state variable with

$$\frac{\mathrm{d}E}{\mathrm{d}t} = -k_{\mathrm{a}} A_0 E + k_{\mathrm{d}} E^*_{\mathrm{qss}} = -k_{\mathrm{a}} \frac{k_{\mathrm{in}}}{k_{\mathrm{d}} + k_{\mathrm{in}}} A_0 E\,. \tag{85}$$

The inactivated enzyme E_{i} has no impact on the output and is classified as a neglected state variable, i.e., $E_{\mathrm{i}} \equiv 0$ and thereby removed from the reduced model. Finally, we would eliminate M by the conservation law $M(t) + M^*(t) = M_0$ and keep its active variant as dynamical state variable with

$$\frac{\mathrm{d}M^*}{\mathrm{d}t} = k_{\mathrm{on}} E^*_{\mathrm{qss}} M - k_{\mathrm{off}} M^* = k_{\mathrm{on}} \frac{k_{\mathrm{a}} A_0 E}{k_{\mathrm{d}} + k_{\mathrm{in}}} (M_0 - M^*) - k_{\mathrm{off}} M^*\,.$$

This would result in a reduction from six to two ODEs.

A reduced model is uniquely determined by the sets of indices (or state variables) that correspond to the dynamic, environmental, neglected, quasi-steady state approximated or by conservation law eliminated state variables, stated as five-tuple

$$\mathcal{M} = (\mathcal{M}_{\mathrm{dyn}}, \mathcal{M}_{\mathrm{env}}, \mathcal{M}_{\mathrm{neg}}, \mathcal{M}_{\mathrm{qss}}, \mathcal{M}_{\mathrm{con}}).$$

In the illustrative example, the reduced model is $(\{E, M^*\}, \{A\}, \{E_i\}, \{E^*\}, \{M\})$. For a given reduced model with output y_{red}, we defined the relative approximation error $\varepsilon_{\mathrm{rel}}$ as

$$\varepsilon_{\mathrm{rel}} = \frac{\left(\int_0^{t_{\mathrm{end}}} (y_{\mathrm{ref}}(t) - y_{\mathrm{red}}(t))^2 \mathrm{d}t\right)^{\frac{1}{2}}}{\left(\int_0^{t_{\mathrm{end}}} y_{\mathrm{ref}}(t)^2 \mathrm{d}t\right)^{\frac{1}{2}}} \tag{86}$$

with y_{ref} denoting the output of the original (reference) model. For a user-defined relative error threshold $\delta > 0$ (e.g., 10%), the challenge is to find a reduced model that still satisfies $\varepsilon_{\mathrm{rel}} \leq \delta$. Since this optimisation problem suffers from the 'curse of dimensionality', we applied the following iterative procedure (for a visual illustration, see the flowchart in Fig. 14):

1. Order state variables according to maximal value of their ir-index

Compute the input-response indices ir_i for all n state variables. Then, order all state variables in order of increasing maximal value of their input-response indices (short: maximal *ir*-index) such that ir_1 corresponds to the state variable with lowest maximal value, and ir_n to the state variable with largest. Set $\mathcal{M}_0 = (\{1, \ldots, n\}, \{\}, \{\}, \{\}, \{\})$ representing the original model.

2. Check for classification as neglected or environmental state variable

Iterate in order of increasing maximal ir-index starting with ir_1 and $j = 1$. At the jth iteration and based on the reduced model $\mathcal{M}_{j-1}$ of the previous iteration, compute the

reduced-order models $\mathcal{M}_{j,\text{env}}$ and $\mathcal{M}_{j,\text{neg}}$ by considering the jth ordered state variable as environmental or neglected, respectively. Based on the resulting relative errors ε_{env} and ε_{neg}, set

$$\mathcal{M}_j = \begin{cases} \mathcal{M}_{j,\text{env}}; & \varepsilon_{\text{env}} \leq \delta \text{ and } \varepsilon_{\text{env}} < \varepsilon_{\text{neg}} \\ \mathcal{M}_{j,\text{neg}}; & \varepsilon_{\text{neg}} \leq \delta \text{ and } \varepsilon_{\text{neg}} \leq \varepsilon_{\text{env}} \\ \mathcal{M}_{j-1}; & \varepsilon_{\text{env}} > \delta \end{cases}$$

Iterate until $j = n$.

3. Check for classification as quasi-steady state variable

Amongst all (say m) remaining dynamic state variables of $\mathcal{M}_n$, iterate again in order of increasing maximal ir-index. At the jth iteration and based on $\mathcal{M}_{n+j-1}$, compute the reduced-order model $\mathcal{M}_{n+j,\text{qss}}$ by considering the jth dynamic state variable of $\mathcal{M}_n$ in quasi-steady state. Based on the resulting relative error ε_{qss}, set

$$\mathcal{M}_{n+j} = \begin{cases} \mathcal{M}_{n+j,\text{qss}}; & \varepsilon_{\text{qss}} \leq \delta \\ \mathcal{M}_{n+j-1}; & \text{otherwise} \end{cases}$$

Iterate until $j = m$.

4. Exploit remaining conservation laws

Amongst all (say l) remaining dynamic state variables of $\mathcal{M}_{n+m}$, iterate again in order of increasing maximal ir-index. At the jth iteration and based on $\mathcal{M}_{n+m+j-1}$, check whether the jth dynamic state variable of $\mathcal{M}_{n+m}$ is part of a conservation law of the original model M_0. If so, compute the reduced-order model $\mathcal{M}_{n+m+j,\text{con}}$ by eliminating the jth dynamic state variable by exploiting the conservation law. Based on the resulting relative error ε_{con}, set

$$\mathcal{M}_{n+m+j} = \begin{cases} \mathcal{M}_{n+m+j,\text{con}}; & \varepsilon_{\text{con}} \leq \delta \\ \mathcal{M}_{n+m+j-1}; & \text{otherwise} \end{cases}$$

Iterate until $j = l$.

The final reduced-order model is then $\mathcal{M}_{\text{final}} = \mathcal{M}_{n+m+k}$. As a potential 5th step, one could further reduce the model $\mathcal{M}_{\text{final}}$ based on proper lumping (e.g., based on the approach described in [1]).

Exploiting conservation laws in the original model $\mathcal{M}_0$ comes without any approximation error. It is therefore tempting to start with step 4. The problem is that we do not know initially, which state variable to eliminate via a conservation law. We could, however, use a first (pre-)run to determine in step 4, which state variables to eliminate from the original model $\mathcal{M}_0$ via conservation laws (in a preliminary pre-run). In the second run, we eliminate all state variables determined by the pre-run in step 4 using conservation laws, and then proceed with steps 1-3. While this increases the computational cost, we found that at the same time it also increases numerical stability.

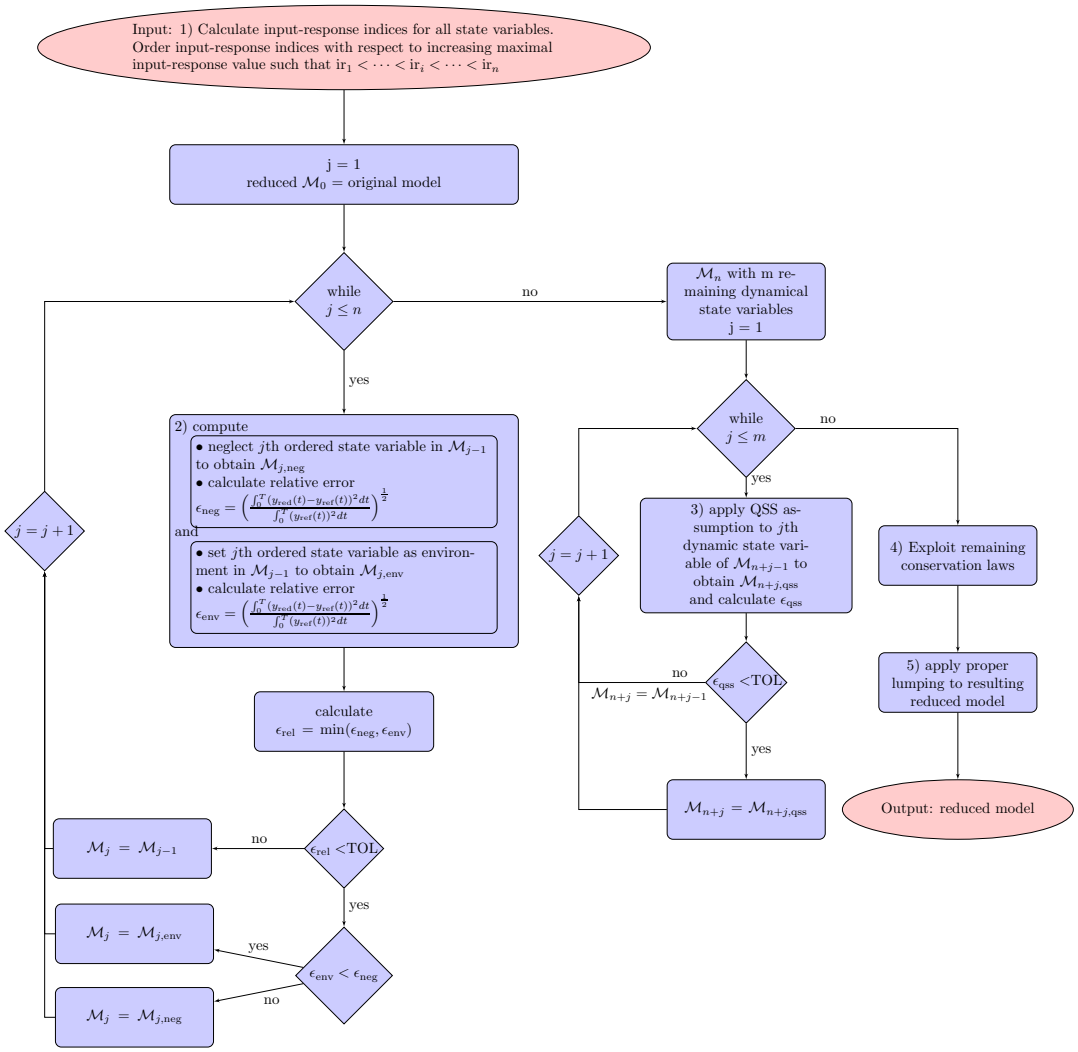


Figure 14: **Flowchart for the model reduction technique presented in this thesis.**

5. Application: analysis & reduction of two large-scale systems pharmacology models

To illustrate the potential of the novel input-response indices to obtain insight into the dynamic behaviour and to reduce complex systems pharmacology models, the blood coagulation network and the EGFR signalling cascade were chosen as challenging model systems. We will address different aspects (i.e. the influence of time interval, the sequence of elimination steps and input domain) of our introduced model order reduction technique. First in section 5.1 we will benchmark our model order reduction technique against a previously employed model reduction technique in the context of the brown snake venom-fibrinogen system. In the brown-snake venom-fibrinogen setting we will focus on the first step of our proposed model order reduction method (cf. section 4.5) in order to analyse the impact of the time interval and the sequence of elimination steps on the reduced model, hereinafter, referred to as elimination-reduced model. Subsequently, it will be demonstrated how the choice of reference trajectory impacts the resulting elimination-reduced models in the context of the PT test. This will allow us to understand the lack of impact of certain genetic modifications on the outcome of the standard blood coagulation test and their impact on the results of a modified test. Next, we will apply our model order reduction technique to extract the essential features of two different anticoagulant drug effects, namely of warfarin and rivaroxaban. This will allow us to answer the previously stated questions of 'How to identify an appropriate pharmacodynamic model?' and 'What model should be employed?'. To show the general applicability of our model reduction technique, we have further chosen the EGFR signalling cascade. While for the EGFR signalling cascade the principal downstream signalling and activation sequence has been intensively studied, the relative importance of the different pathways and molecular constituents is still controversially discussed. In section 5.2 we will demonstrate how the sensitivity based ir-indices can be applied to provide a better understanding of these features. Finally in section 5.3, we will compare our introduced model reduction technique for the PT test setting to the most commonly employed MOR techniques in the field of systems pharmacology/biology.

5.1. Blood coagulation network

In this section, we illustrate our novel model reduction technique in application to different aspects of the human blood coagulation network [130, 39, 40] described in detail in section 2.3.

5.1.1. The brown snake venom - fibrinogen system: Impact of the time interval of interest and sequence of elimination steps on the reduced model

The brown snake venom-fibrinogen system was studied in detail in [39, 40]. The authors simplified the system to obtain a reduced model of 5 state variables for the effect of the brown snake venom on fibrinogen (Fg) [40]. Their simplification was based on proper lumping [1] in addition to heuristically manipulating parameter values. Their results were considered as a benchmark for testing our novel model reduction technique. As in [40], the brown snake

venom, triggering the activation of the blood coagulation through activation of factor II (as described in section 2.3.2), was considered as the input and fibrinogen as the response (output).

Input-response indices during the first hour after envenomation

The reference dynamics was chosen to be the dynamics of the blood coagulation network with an initial amount of brown snake venom of 0.0015 mg (as in [40]). The empirical ir-indices were determined based on ten perturbations of the snake venom input ranging from 50% to 150% of the reference input according to eq. (79). Since activation of the coagulation factors occurs on a fast time scale compared to the recovery phase, we first studied the system on a time period up to 1 h (and subsequently up to 40 h, see section "Input-response indices of the brown snake venom-fibrinogen system including recovery").

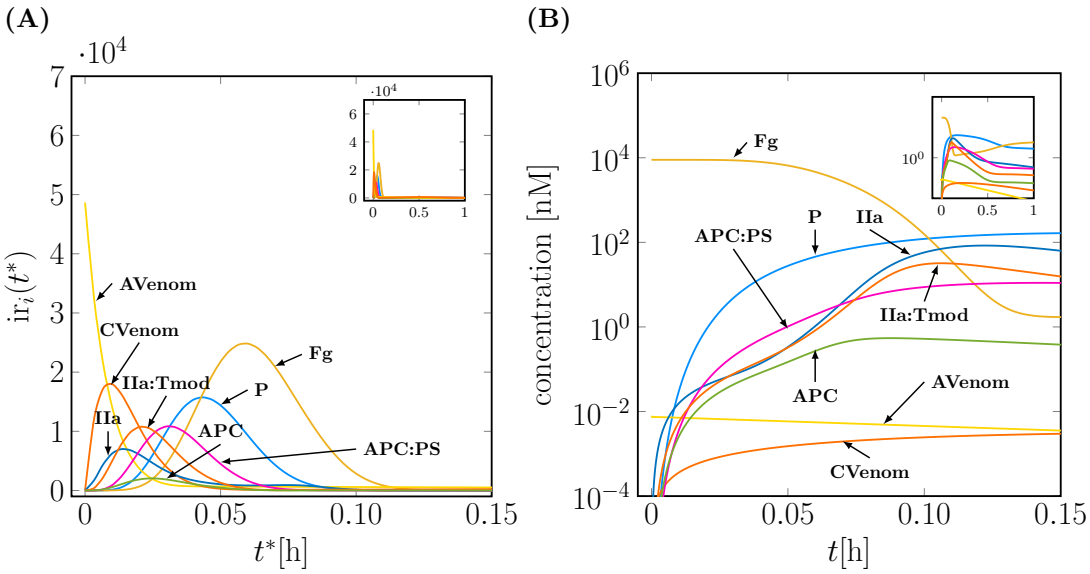


Figure 15: **Empirical ir-indices for the brown snake venom-fibrinogen system up to 1h and time profiles of corresponding factors**. **(A)** Shown are the empirical ir-indices of the most important state variables for the time interval $[0, 0.15]$ h. The insets shows the ir-indices on the entire interval $[0, 1]$ h. Ir-indices not shown are below $2 \cdot 10^3$ in magnitude. **(B)** Concentration-time profiles of the most important model molecular species (according to the indices) predicted by the original model. The inset shows the concentration-time profiles on the entire time interval. The concentration-time profiles of the remaining states are given in Figure 43 in the appendix.

Figure 15(A) shows the ir-indices for the time interval $[0, 1]$ h. The ir-indices nicely reflect the coordinated activation of the different coagulation factors. These include the factors: IIa, complex of thrombin and thrombmodulin (IIa:Tmod), complex of activated protein C and protein S (APC:PS), and P. The remaining factors do not substantially impact the output Fg up to 1h after envenomation, since their ir-indices are to small to be visible in Figure 15(A). Figure 15(B) shows the concentration-time courses of the corresponding coagulation factors.

The time courses of the remaining factors are shown in Figure 43 in the appendix. In contrast to the ir-indices, a clear insight into the relevance of factors for the system dynamics is difficult to obtain. The ir-indices highlight the most relevant path through the coagulation system from the activation by the snake venom (input - AVenom) to the resulting blood coagulation response (output - Fg). Furthermore, the ir-indices indicate that coagulation factors are only of relevance during certain time windows. For example, while the impact of the factors IIa and IIa:Tmod on Fg is largely confined to the initial time window up to 0.05h (cf. panel A), the concentrations of these factors change during the whole time interval (cf. panel B).

Elimination-reduced model of the brown snake venom-fibrinogen system

We next reduced the 62-state brown snake venom-fibrinogen system based on the ir-indices (cf. Figure 15). To this end, we recursively eliminated state variables in the order of the maximum of the input-response indices by determining if they can be classified as neglected or environmental state variables, starting with the lowest value. The relative approximation error in eq. (86) was required to satisfy $\varepsilon_{\text{rel}} \leq 0.1$. The elimination process using only step one of the model order reduction technique (cf. section 4.5) resulted in a so-termed elimination-reduced model with 8 dynamical state variables (cf. Figure 16), in addition to 5 environmental state variables (indicated with $*$ in Figure 16). While environmental state variables are important for the input-response relationship, they are considered as constant for the time scale of interest, in contrast to the dynamical state variables. The remaining $49 = 62 - 8 - 5$ state variables were neglected for the input-output relationship since they were not important for the effect of the brown snake venom on Fg during the first hour after envenomation.

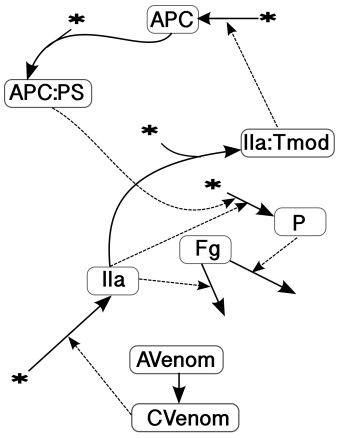


Figure 16: **Elimination-reduced model of the brown snake venom-fibrinogen system.** Model reduction was based on maximal value of the input-response indices with $\varepsilon_{\text{rel}} \leq 0.1$. Shown are the 8 dynamical state variables. The environmental state variables (indicated by '*') are II, Pg, PC, Tmod and PS. The corresponding differential equations and further details are given in the appendix. Cf. Figure 6 for a legend of the different arrow types.

A distinction of state variables into dynamical, environmental and negligible is in line with biological and pharmacological insight and expectations: The blood coagulation cascade

consists of two main pathways: the extrinsic pathway activated by TF and the intrinsic pathway activated by the contact factor (CA). In case of an injury, both pathways act together to form fibrin strands. The brown snake venom directly activates the blood coagulation through factor II. In this case, the initial parts of both, the intrinsic and extrinsic pathway do not play a role. Therefore, both initial parts (TF, VII, VII:TF, etc. as well as CA, XII, etc) can be eliminated. The activation of factor II is key for the brown snake venom-fibrinogen system. While factor II concentration is important, it does not change significantly during the time of interest. Thus it was considered as an environmental state.

As expected, the elimination-reduced model (8-state variable model) approximates the Fg concentration-time profile of the original model (62-state variable model) with high accuracy for the time interval up to 1h (cf. Figure 17). As can be inferred from the inset in Figure 17, the approximation quality also extends to the time interval up to 40 h (cf. Table 14 in the appendix for a quantification of the relative error pre and post nadir of the fibrinogen concentration).

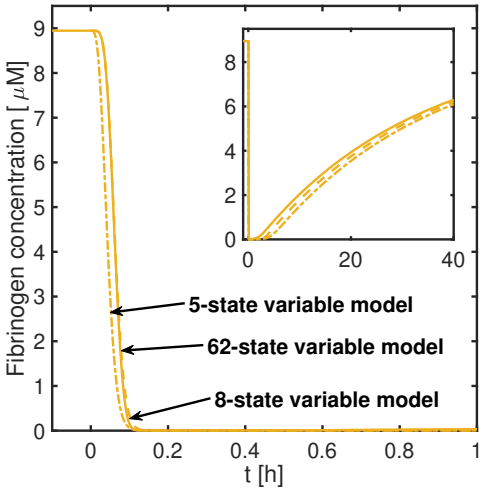


Figure 17: **Comparison of fibrinogen (output) for the brown snake venom-fibrinogen system** based on the original 62-state variable model (solid), the elimination-reduced 8-state variable model (cf. Figure 16) (dashed) and the lumped elimination-reduced 5-state variable model (cf. Figure 20) (dot-dashed). The inset shows the comparison of fibrinogen for the time interval of 40 hours. The comparison of the concentration time profile for the remaining 7 state variables is shown in Figure 44 in the appendix.

The elimination-reduced model (8-state variable model) in Figure 16 gives further insight into the full coagulation dynamics (cf. Figure 6). The activated factor IIa impacts the decline of Fg in three different ways: (i) directly by increasing the transformation of Fg to F; (ii) indirectly via the factor P that mediates the transformation of Fg to the fibrin degradation product (FDP); and (iii) indirectly via the activated protein C (APC) that in turn increases the impact of P on Fg. In order to quantify the relative contributions between the direct and the indirect impact, we determined the transformation rates of Fg to F (direct impact) and

Fg to FDP (indirect impact) from the original model. As can be inferred from Figure 18, the indirect impact largely outperforms the direct impact. This is in agreement with the concentration-time courses shown in Figure 15(B), in which factor P concentrations were considerably higher than factor IIa levels.

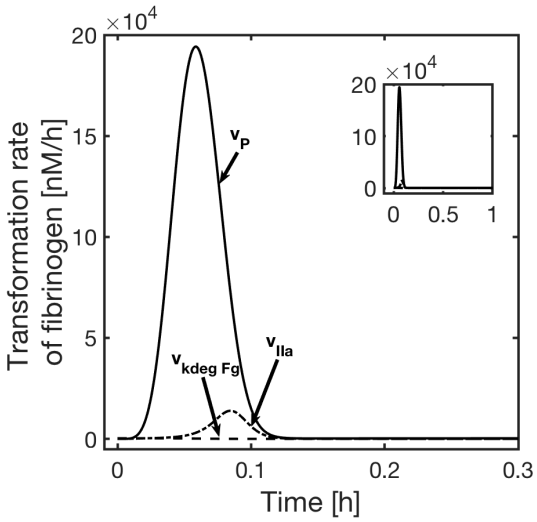


Figure 18: **Transformation rates of Fg for the brown snake venom-fibrinogen system.** Shown are the transformation rates of Fg mediated by P and IIa in addition to the degradation rate based on the original 62-state variable model. Transformation of Fg is dominated by v_{P}. The inset shows the rates for an extended time interval.

The reduced model obtained in this section depends on the choice of user-defined error tolerance. If a higher approximation accuracy of the reduced model is required, the error tolerance should be decreased. As a result, the order of the reduced model is expected to increase. This is demonstrated and summarised in Table 15 in the appendix, showing the resulting order of the reduced model for decreasing values of the user-defined error tolerance.

Input-response indices for fibrinogen recovery up to 40 h

We next studied the brown snake venom-fibrinogen system on the time period up to 40 h after envenomation thereby including the recovery phase of Fg. The corresponding ir-indices are shown in Figure 19. In contrast to the ir-indices obtained for the time period up to 1h, we observed that the pathway including the factors Xa and Xa:Va becomes relevant on the longer time period up to 40h. This highlights the fact that also the time interval is important, on which one aims to approximate the full dynamics.

To understand the importance of factors V, Va, Xa and Xa:Va as indicated by the indices, we compared the dynamics of Fg predicted by the 62-state variable model and a model, where (i) the factors V and X, or (ii) only factor V were artificially knocked out. This was realised by setting the production rates of the corresponding factors and their initial concentration to zero. Consequently, also the activated factors Xa, Va and Xa:Va were absent in the knock-out model.

As can be inferred from Figure 42 in the appendix, the pathway via activation of X and V influences the recovery phase of Fg. Performing an analogous analysis by knocking-out factor Pg, the inactive form of factor P, a key factor in the 8-state variable elimination-reduced model in Figure 16, we inferred that factor P is important for the initial transient decay of Fg (cf. Figure 42 in the appendix).

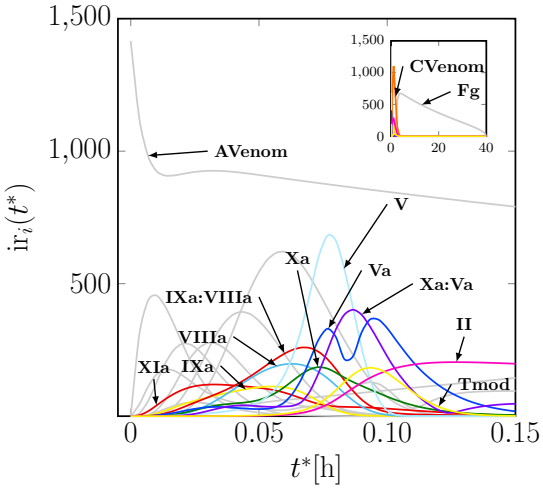


Figure 19: **Empirical ir-indices for the brown snake venom-fibrinogen system up to 40h.** Shown are the indices of the most important state variables; input-response indices below 10^2 in magnitude are not shown. To reduce the complexity of the figure, the ir-indices of the states already shown in Figure 15 are coloured in grey, since their initial transient peak did not change (up to a normalisation constant). This was done for illustrative purposes and did not influence the subsequent model reduction. The inset shows the ir-indices on the entire time interval [0,40h].

For the elimination of state variables according to step one of our model order reduction technique (cf. section 4.5), we again proceeded recursively in the order of the maximum of their input-response indices, starting with the lowest value. The relative approximation error in eq. (86) was required to satisfy $\varepsilon_{\mathrm{rel}} \leq 0.1$. Interestingly, the elimination-reduced model based on the indices up to 40h is identical to the corresponding elimination-reduced model based on the indices up to 1h. An explanation of this unexpected result is provided in the next subsection.

Influence of sequence of elimination steps on the elimination-reduced model

In our approach, the elimination of states is based on the maximum of the ir-indices, proceeding from low to high values. Thereby, we first eliminate state variables that have a low impact on the output. Recursively, it is tested whether a state variable can be eliminated by fixing it to its initial value (environmental state variable) or by fixing it to zero (negligible state variable). The question arises whether one could simply have chosen a random order of the state variables instead and subsequently proceeded with the elimination steps as above. In other words, the question is: How does the sequence of elimination steps affect the resulting reduced model?

To address this question, we randomly chose a sequence of state variables and performed the elimination steps as outlined above. Note that all state variables were tested for elimination, proceeding in the (randomly) chosen order. The time interval was set to be $[0, 40]$ h, thus including both, the initial drop and the subsequent recovery phase of Fg. The relative approximation error was required to satisfy $\varepsilon_{\text{rel}} \leq 0.1$. The resulting elimination-reduced models for ten random sequences are stated in Table 6. We observed that the order of the models differs, ranging from 8 to 13, and that the models fall into one of two groups: (i) reduced models comprising the factors P, APC and APC:PS (including the model shown in Figure 16); and (ii) reduced models including the factors Xa, Va and Xa:Va. A model of the latter type based on run no. 2 is shown in the appendix (cf. Figure 41).

Table 6: **Influence of the sequence of elimination steps on the reduced model** Listed are the dynamical state variables of the elimination-reduced model based on ten randomly chosen sequences in which state variables were tested for elimination.

Run	**Order**	**Dynamical state variables of the elimination-reduced model**	**Environmental state variables**
1	8	Fg, IIa, IIa:Tmod, AVenom, CVenom, P, APC, APC:PS (identical to model depicted in Figure 16)	II, Pg, PC, PS, Tmod
4	8	identical to run no. 1	identical to run no. 1
5	8	identical to run no. 1	identical to run no. 1
8	8	identical to run no. 1	identical to run no. 1
3	9	Fg, IIa, IIa:Tmod, AVenom, CVenom, Pg, P, APC, APC:PS	II, PC, PS, Tmod
6	9	Fg, IIa, IIa:Tmod, AVenom, CVenom, P, APC, APC:PS, XIIIa	II, Pg, PC, PS, Tmod, XIII
9	11	Fg, II, IIa, CVenom, Xa, Va, Xa:Va, XIa, IXa ,VIIIa, IXa:VIIIa	AVenom, VIII, XI, IX, V
2	12	Fg, IIa, AVenom, CVenom, Xa, Va, Xa:Va, XIa, IX, IXa, VIIIa, IXa:VIIIa	II, V, VIII, X, XI
7	13	Fg, II, IIa, AVenom, CVenom, Xa, Va, Xa:Va, XIa, IXa, VIII, VIIIa, IXa:VIIIa	XI, X, V, IX
10	13	identical to run no. 7	identical to run no. 7

As stated in the previous section, the elimination-reduced model based on the indices up to 40h does not include the factors V, Xa, Va and Xa:Va, despite their importance indicated by the input-response indices shown in Figure 19. They only appeared in reduced models based on the elimination of state variables in a randomly chosen order that is substantially different from the order defined by the ir-indices.

As outlined in the previous section, also the order, in which state variables are tested for elimination, has an impact on an iterative model reduction process. In this regard it is important to notice that the ir-indices of the activated factors Va and Xa have a lower

maximum than factor P. Consequently, they are earlier tested for elimination than factor P. If Va and Xa are eliminated, then factor V does not play any role and can be eliminated as well. Thus, the relevance of factor V is conditioned on the presence of Va, Xa and/or Xa:Va; only in the presence of Va, Xa and/or Xa:Va, the factor V does have a large impact. If, however, factor P is eliminated before the factors Xa, Va, Xa:Va and V in the iterative model reduction process[2], then Xa, Va, Xa:Va and V are important for the recovery phase and are not neglected in subsequent iterations. In this case, the reduced model approximates the recovery phase well at the expense of the initial transient decay (cf. Table 14 in the appendix for quantification of the relative error pre and post nadir of the fibrinogen concentration).

The ir-indices quantify the importance of state variables in the original model. It is important to emphasise that the ir-indices do not depend on any error tolerance that a user has to specify to control the approximation error in the model reduction process. Thus, while the ir-indices do not change with different error tolerances, one would, however, expect to obtain different reduced models when changing the error tolerance. This is demonstrated in Table 15 (appendix): If the error tolerance is decreased, the reduced model contains more state variables. In particular, we observed that the reduced models obtained with a lower error tolerance contain the pathways involving both, factor P and factor V. For the chosen error tolerance of 0.01, both the 8- and 12-state variable model (cf. Figure 16 and 41) satisfy the error threshold (see Table 14 in the appendix).

From a pure theoretical point of view, both the 8-state variable and 12-state variable elimination-reduced models in Figures 16 and 41 satisfy the approximation error as set by the user-defined tolerance. From a clinical point of view, their utility and the interpretation of the two models could be considerably different, depending on the question of interest. Importantly, the ir-indices allow to understand the origin of these differences between the two reduced models as well as potential differences between the reduced and the original model, as outlined above.

Elimination-reduced and lumped model of the brown snake venom-fibrinogen system up to 40h

Finally, we further simplified the 8-state variable model given in Figure 16 via proper lumping (described in detail in section 3.1.1). For the proper lumping approach, the same threshold on the approximation error, i.e, $\varepsilon_{\text{rel}} < 0.1$, was used and the time interval of $[0, 40]$ h was considered. This resulted in an elimination-reduced and lumped (short: reduced) model of 5 state variables in addition to 3 environmental state variables. See Figure 20 for a graphical representation of the reduced model. The differential equations of the reduced 5-state variable model are:

$$
\begin{aligned}
\frac{\mathrm{d}x_{\text{L1}}}{\mathrm{d}t} &= -\,d_{\text{CVenom}} \cdot \frac{1}{2} x_{\text{L1}} \\
\frac{\mathrm{d}x_{\text{L2}}}{\mathrm{d}t} &= V_{\text{Xa:Va,IIa}} \frac{\frac{1}{2}x_{\text{L1}}}{\frac{1}{2}x_{\text{L1}} + K_{\text{Xa:Va,IIa}}} \cdot x_{\text{env,1}} - (d_{\text{IIa:Tmod}} + d_{\text{IIa}}) \cdot \frac{1}{2} x_{\text{L2}} \\
\frac{\mathrm{d}x_{\text{L3}}}{\mathrm{d}t} &= V_{\text{IIa:Tmod,APC}} \frac{\frac{1}{2}x_{\text{L2}}}{\frac{1}{2}x_{\text{L2}} + K_{\text{IIa:Tmod,APC}}} \cdot x_{\text{env,2}} - (d_{\text{APC:PS}} + d_{\text{APC}}) \cdot \frac{1}{2} x_{\text{L3}}
\end{aligned}
$$

[2] e.g., due to a different order of state variables in a randomly chosen order, cf. Table 6, e.g., run no. 2

$$\frac{\mathrm{d}x_{\mathrm{L4}}}{\mathrm{d}t} = \left(V_{\mathrm{IIa,P}} \frac{\frac{1}{2}x_{\mathrm{L2}}}{\frac{1}{2}x_{\mathrm{L2}} + K_{\mathrm{IIa,P}}} + V_{\mathrm{APC:PS,P}} \frac{\frac{1}{2}x_{\mathrm{L3}}}{\frac{1}{2}x_{\mathrm{L3}} + K_{\mathrm{APC:PS,P}}} \right) \cdot x_{\mathrm{env},3} - d_{\mathrm{P}} \cdot x_{\mathrm{L4}}$$

$$\frac{\mathrm{d}x_{\mathrm{L5}}}{\mathrm{d}t} = p_{\mathrm{Fg}} - \left(V_{\mathrm{IIa,Fg}} \frac{\frac{1}{2}x_{\mathrm{L2}}}{\frac{1}{2}x_{\mathrm{L2}} + K_{\mathrm{IIa,Fg}}} + V_{\mathrm{P,Fg}} \frac{x_{\mathrm{L4}}}{x_{\mathrm{L4}} + K_{\mathrm{P,Fg}}} + d_{\mathrm{Fg}} \right) \cdot x_{\mathrm{L5}}.$$

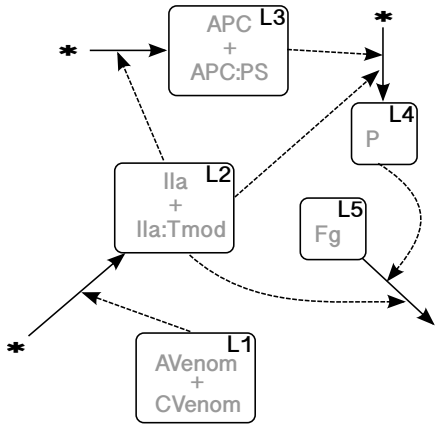


Figure 20: **Reduced 5-state variable model of the brown snake venom-fibrinogen system with $\varepsilon_{\mathrm{rel}} \leq 0.1$.** Shown are the lumped state variables with their corresponding state variables of the 8 state variable model. The environmental state variables are indicated by '*' (cf. Table 7 for further details).

Table 7 lists the reduced and environmental state variables and their corresponding (initial) values. The parameter values are given in Table 12 in the appendix. The snake venom input corresponds to the state variable x_{L1}, while the Fg output corresponds to the state variable x_{L5}. The reduced 5-state variable model still reproduced the Fg response sufficiently well as a result of the snake venom input (cf. Figure 17 and inset).

Table 7: **Reduced 5-state variable model of the brown snake venom-fibrinogen system.** Definition of the reduced and environmental state variables.

Lumped state variable	**Original state variables**	**Initial condition**
x_{L1}	AVenom + CVenom	0.0075 [nM]
x_{L2}	IIa + IIa:Tmod	0 [nM]
x_{L3}	APC + APC:PC	0 [nM]
x_{L4}	P	0 [nM]
x_{L5}	Fg	8945.5 [nM]

Environmental state variable	**Original state variable**	**Constant value**
$x_{\mathrm{env},1}$	II	1394.4 [nM]
$x_{\mathrm{env},2}$	PC	60 [nM]
$x_{\mathrm{env},3}$	Pg	2154.3 [nM]

5.1.2. The PT test: Impact of the magnitude of the input on the resulting reduced model

It is well known that model reduction techniques for nonlinear systems typically only apply locally, i.e., the reduced model is a good approximation to the original model for a given domain of the input, the initial state and the parameter values. While this is often considered as a disadvantage of model reduction techniques, we believe that it should be rather considered as a feature allowing for valuable insight into the dynamics. We illustrated this in the following section by studying the blood coagulation system *in vitro* in the absence of any pharmacological agent.

The prothrombin time (PT) test is an *in vitro* blood coagulation test quantifying the activity of the so-called extrinsic pathway activated by the tissue factor TF. It has been studied in detail by Wajima et al. [130]; the model is a variant of the coagulation network shown in Figure 6. The PT test is initiated experimentally by adding a well-defined amount of TF to a sample of blood and measuring the time until the blood has clotted. In [130], the PT test was studied with the TF concentration as input and fibrin (F) as the response.

It has been reported in the literature that the common PT test—characterised by a high TF concentration—is insensitive to genetic deficiencies of factors VIII and IX, while a modified PT test—characterised by a much lower TF concentration (roughly by a factor 1000)—is sensitive to these deficiencies [14]. This suggests that depending on the magnitude of the input the same *in vitro* PT test model should result in different reduced models (in particular regarding the inclusion of factors VIII and IX).

Input-response indices for two PT test scenarios

The reference dynamics were chosen to be the dynamics of the PT test with an initial TF concentration of 5pM (low TF scenario) and 100nM (high TF scenario), in line with [14, 130] using the 62-state variable model (cf. Figure 6).[3] The ir-indices were determined based on ten perturbations of the TF input ranging from 50% to 150 % of the reference input according to eq. (79). The time interval of interest was determined based on the corresponding reference dynamics to include the transformation of Fg to F (which is required for the PT test). For the low TF scenario, we chose $[0, 240]$ seconds, while for the high TF scenario, we chose $[0, 30]$ seconds. The latter include the prothrombin time of about 11.6-13.8 seconds for a healthy individual [14]. Figure 21 shows the corresponding input-response indices for the low and high TF scenario.

Both scenarios have seven factors of high importance according to the ir-indices in common: TF, IIa, VII:TF, VIIa:TF, Xa, Fg, and F. In the low TF scenario, in line with expectations, also the factors V, VIII and IX (and their activated forms) impact the output via the factors Xa and Xa:Va.

[3] In [130], the concentration for the high TF scenario is reported to be 300 nM, but subsequently scaled to 1/3 of its value to reflect a dilution process.

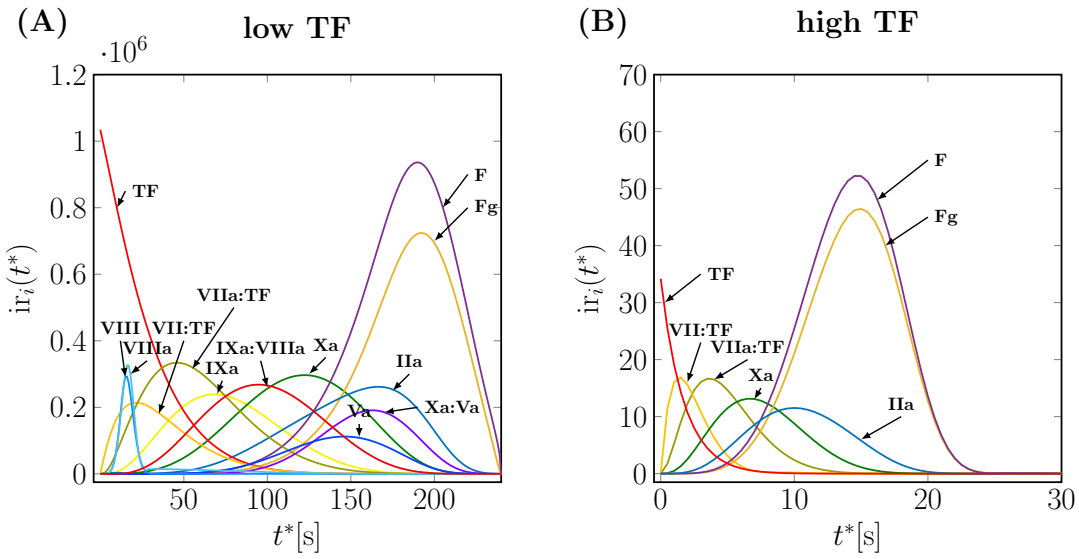


Figure 21: **Empirical ir-indices for two PT test scenarios.** Shown are the input-response indices of the most important state variables as a function of time for **(A)** the low TF and **(B)** high TF scenario. Input-response indices not shown are below $3 \cdot 10^4$ in magnitude in the low TF and 3 in magnitude in the high TF scenario.

Reduced models for two PT test scenarios

We next reduced the 62-state variable model (cf. section 2.3.2) for the low and high TF PT test scenarios by applying step one of our model order reduction procedure (as described in detail in section 4.5). The relative approximation error was required to satisfy $\varepsilon_{\mathrm{rel}} \leq 0.1$. The resulting elimination-reduced models comprised 13 state variables for the low TF scenario and 7 state variables for the high TF scenario (cf. Figure 22).

The difference between the two reduced models is the pathway of factor V, VIII and IX, which is included in the low TF scenario but not in the reduced model for the high TF scenario. This is due to the fact that the activation of Fg is much faster in the high TF scenario such that delayed positive feedback via the factors V, VIII and IX has no substantial impact on the output. In the low TF scenario, this delayed feedback amplifies the weak TF-induced activation of factor II to a substantial extent.

The two TF scenarios clearly demonstrate that the magnitude of the input can have a profound impact on the resulting elimination-reduced model, and this impact may be important for our understanding of the dynamics.

The elimination-reduced models for both TF scenarios were further reduced via proper lumping. The same threshold on the approximation error and time interval was used. The resulting elimination-reduced and lumped models consist of 12 state variables and 6 environmental state variables in the low TF case and 5 state variables and 4 environmental state variables in the high TF scenario (cf. Table 8).

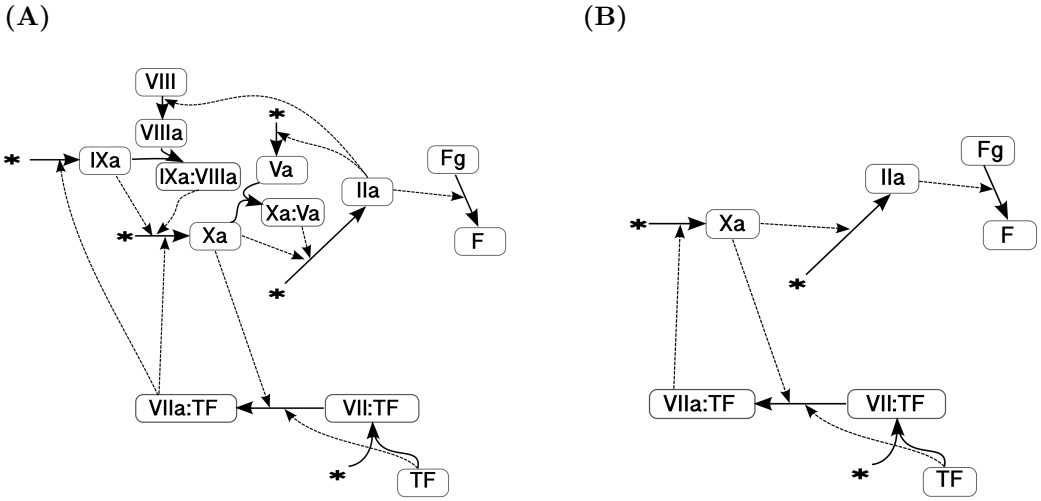


Figure 22: **Elimination-reduced models for two PT test scenarios** the **(A)** low TF and **(B)** high TF scenario with $\varepsilon_{\text{rel}} \leq 0.1$. Shown are the dynamical state variables. The environmental state variables (indicated by '*') are II, V, VII, IX and X for low TF and II,VII and X for high TF. The corresponding differential equations and further details are given in the appendix. Compare Figure 6 for a legend of the different arrow types.

Table 8: **Reduced models for two PT test scenarios** for low TF and high TF with relative approximation error $\varepsilon_{\text{rel}} \leq 0.1$. Definition of the reduced and environmental state variables.

low TF

Lumped state variables	**Original state variables**
x_{L1}	F
x_{L2}	Fg
x_{L3}	TF
x_{L4}	IXa
x_{L5}	IIa
x_{L6}	IXa:VIIIa
x_{L7}	VIII
x_{L8}	VIIIa
x_{L9}	VII:TF
x_{L10}	VIIa:TF
x_{L11}	Va + Xa
x_{L12}	Xa:Va

Environmental state variable	**Original state variables**
$x_{\text{env},1}$	II
$x_{\text{env},2}$	VII
$x_{\text{env},3}$	X
$x_{\text{env},4}$	TFPI
$x_{\text{env},5}$	IX
$x_{\text{env},6}$	V

high TF

Lumped state variables	**Original state variables**
x_{L1}	F
x_{L2}	Fg
x_{L3}	TF
x_{L4}	Xa
x_{L5}	IIa + VII:TF + VIIa:TF

Environmental state variable	**Original state variables**
$x_{\text{env},1}$	II
$x_{\text{env},2}$	VII
$x_{\text{env},3}$	X
$x_{\text{env},4}$	TFPI

5.1.3. Extracting the essential features of two different drug effects: warfarin and rivaroxaban

In this section, the usefulness of model order reduction in the context of a systems pharmacology model is highlighted by using the blood coagulation network model (cf. section 2.3.2) to extract two drug effect models - warfarin and rivaroxaban effect. As described in section 2.4 warfarin was chosen due to its clinical relevance and pharmacological challenges to date (narrow therapeutic window, large inter-individual variability, genetic polymorphisms in e.g. CYP2C9 and VKORC1), while rivaroxaban was chosen due to its novelty in clinics and its completely different mode of action. In order to include the effect of rivaroxaban on the blood coagulation model we extended the original 62-state variable model [130, 40] resulting in 66 state variables (detailed description in section 2.3.4). This extension had no impact on the original systems dynamics.

Warfarin effect

The general procedure of warfarin effect assessment by drawing blood from the patient and initialising the PT test by adding a well-defined amount of tissue factor was described in section 2.3.3. It is possible to describe this with the blood coagulation network model (cf. Figure 7). Therefore, we have chosen the dynamics of the blood coagulation system under warfarin therapy (4 mg daily) with a range of sample times $T_{Blood} \in [0, 20]$ days of blood withdrawal and an addition of TF concentration of 100 [nM] to the diluted factor concentrations (dilution to one third) at time T_{Blood} as the reference dynamics. Instead of computing multiple dosing for warfarin, we considered a constant infusion. This was chosen for numerical reasons and closely matches the dynamics under multiple dosing (as it is the limit case).

The sensitivity based ir-indices for the warfarin-fibrin system during the initiation of warfarin therapy (cf. Figure 23) clearly shows that during the *in vivo* setting, the warfarin concentration (the input) only influences the vitamin K-dependent factors as well as vitamin K hydroquinone (VKH2). All the other factors/proteins of the blood coagulation system did not change from their initial value (excluding the vitamin K cycle). Furthermore, the ir-indices reflect that during therapy initiation factor VII is the most important state variable for the *in vivo* response. This is in line with expectations since it is known that the early changes in prothrombin time reflect the changes in VII levels [45]. In the *in vitro* setting after the addition of TF the blood coagulation network becomes activated and thus the ir-indices reflect the sequential activation of the blood coagulation by the most important factors. However, in contrast to the high TF PT test scenario (cf. Figure 21(B)), factor VII and not the tissue factor (TF) is considered as important for the dynamics of the system.

The sensitivity based ir-indices for later time points of blood withdrawal are shown in the appendix (cf. Figure 45 and 46). Factor VII is the most important factor during the initiation of warfarin therapy, however, closer to the so-called clinical 'steady state' under warfarin therapy, factor II and X become more important than factor VII (cf. Figure 45 and 46 in the appendix).

To illustrate how the input-response indices change for different time points of blood with-

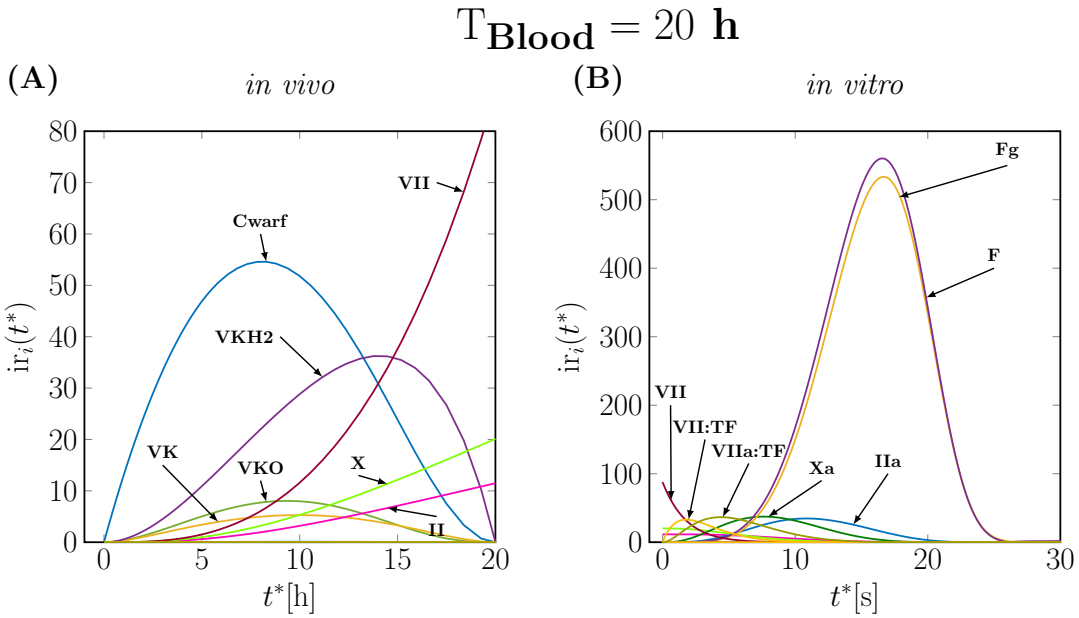


Figure 23: **Sensitivity based ir-indices for the warfarin-fibrin system during the initiation of warfarin therapy.** Time of blood withdrawal was chosen to be $T_{Blood} = 20$ h. Shown are the sensitivity based ir-indices for the most important state variables of the **(A)** *in vivo* and **(B)** *in vitro* parts under warfarin therapy (4 mg daily). For better readability, the ir-indices were shown for *in vivo* and *in vitro* part in two different panels. Sensitivity based ir-indices not shown are below 5 in magnitude.

drawal during warfarin therapy, we summarise the indices by their maximal value for $T_{Blood} \in [0, 20]$ days (cf. Figure 24). As previously mentioned, the summarised indices indicate that factor VII is the most important factor during the first 3 days of warfarin therapy, while in the maintenance phase of warfarin therapy factor X and factor II are of more/similar importance. Additionally, one can infer from the Figure 24, that although the importance of each factor changes over the duration of warfarin therapy the order of importance based on their maximal value at each time point T_{Blood} is largely retained, especially in the *in vitro* setting (exceptions are factor X and II).

Next, we reduced the 66-state blood coagulation model by successively eliminating states according to their importance to the system dynamics of interest, starting with the lowest value (as described in detail in section 4.5) based on the sensitivity based input-response indices for various T_{Blood}. In all cases, the resulting relative approximation error was required to satisfy $\varepsilon_{rel} < 0.05$ (cf. eq. 86). The resulting reduced models for the maintenance phase of warfarin therapy had 11 dynamical state variables of the original blood coagulation network model and 2 additional state variables for the pharmacokinetics of warfarin (cf. Figure 25).

In section 2.4 two empirical pharmacodynamic models were introduced for the warfarin effect. In order to assess the plausibility of the resulting INR prediction from both the original mechanistic 66-state variable and reduced mechanistic 13-state variable model (cf. Figure 25) we have chosen the most widely used empirical PK-PD model for S-warfarin [45] in our comparison. As was previously described in section 2.4 the empirical PD model consists of two

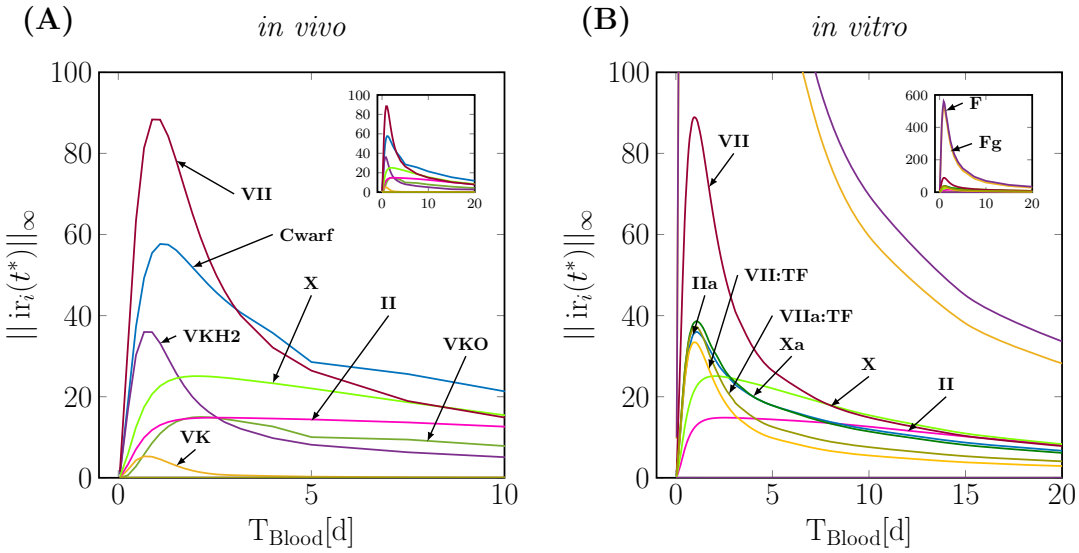


Figure 24: **Maximal values for sensitivity based ir-indices for the warfarin concentration-fibrin system over time of blood withdrawal ($\mathrm{T}_{\mathbf{Blood}}$).** Shown are maximal values of the sensitivity based ir-indices for the most important state variables over $\mathrm{T}_{\mathrm{Blood}}$ of the **(A)** *in vivo* and **(B)** *in vitro* parts under warfarin therapy (4 mg daily). For better readability, the ir-indices were shown for *in vivo* and *in vitro* part in two different panels. Sensitivity based ir-indices not shown are below 5 in magnitude. The inset shows the maximal values of the sensitivity based ir-indices on the entire time interval $[0, 20]$ days.

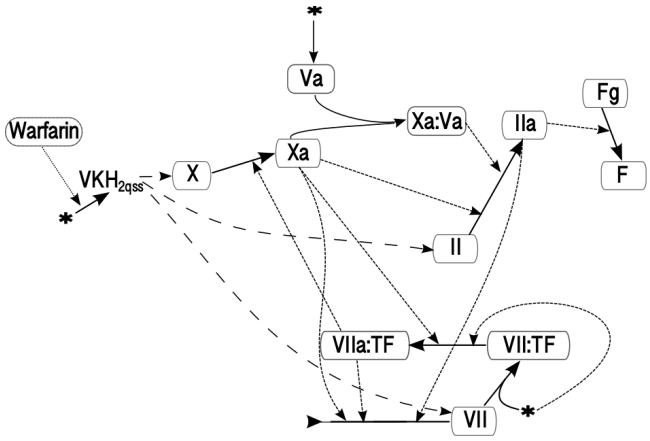


Figure 25: **Reduced model for warfarin therapy**. Shown are the dynamical state variables. The environmental state variables (indicates by '*') are TF and VK. The state variable approximated by its quasi-steady state is indicated by the subscript 'qss'. In the model order reduction process the relative approximation error was required to satisfy $\varepsilon_{\mathrm{rel}} \leq 0.05$.

transit compartment chains of different length and different mean transition time (cf. Figure 9). In order to compare the predictions, it was necessary to account for the differences in pharmacokinetics. In the mechanistic model, a PK model for the racemic warfarin concentration was used, while in the empirical PK-PD model S-warfarin concentration was chosen. Therefore, we assumed a dose of 2 mg daily in the empirical PK-PD model to be comparable

to 4 mg daily dosing in the mechanistic model. This assumption was based on the finding in the literature that the racemic mixture consists of 50% R-warfarin and 50% S-warfarin [23, Ch. 20] and [82, 33]. In Figure 26 we compare the typical INR predictions of the original (blue line) and reduced (red dashed line) mechanistic model with the 90% CI of the empirical PK-PD model based on 1000 simulations of the typical patient accounting for parameter uncertainty. Thus, the 90% CI of the empirical PD model reflects the effect of the uncertainty in the PD parameter estimates as reported in [45]. We assumed the parameters to be uncorrelated since no uncertainty covariance matrix was published in [45].

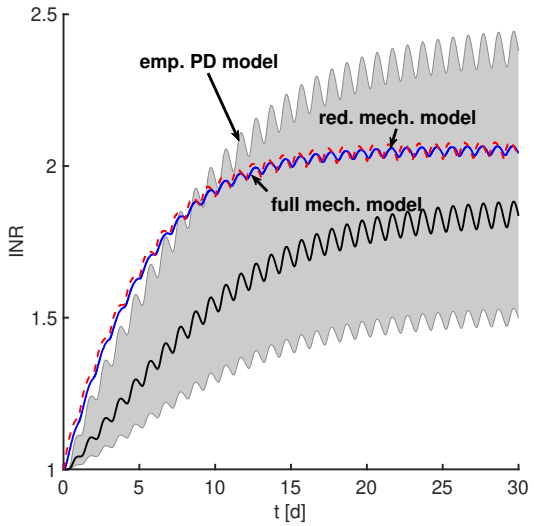


Figure 26: **Comparison of international normalised ratio (INR) for warfarin therapy of 4mg daily** based on the original 66-state variable model (solid blue), the reduced 13-state variable model (dashed red) and the empirical PD model in [45] with median (black) and 90% CI (grey area).

As can be inferred from Figure 26 during the induction of warfarin therapy the two different modelling approaches produce divergent predictions for the INR both in actual levels as well as time to maintenance phase ('steady state' - stable limit cycle). However, the mechanistic based 'steady state' INR predictions lie within the 90% CI of the empirical PK-PD model. The difference in the predictions of the induction phase of warfarin therapy can in parts be explained by the fact that only data for a single dose and during stable maintenance dose were used to develop the empirical PK-PD model with slightly more than half of the data coming from the single dose administration. As was remarked in [45, p. 535], 'the model needs to be further informed by data from the induction phase'. This informative difference in the model INR prediction might thus indicate model misspecification and under-prediction of warfarin initiation therapy by the empirical PK-PD model. The slight differences in the intensity of anticoagulation during one dosing interval between the two modelling approaches is due to the differences in warfarin pharmacokinetics. In the mechanistic blood coagulation model, a 1 compartment (CMT) model for the racemic warfarin concentration is used while in the PK-PD model a 2 CMT model for S-warfarin concentration was chosen. The concentration-

time profiles of the respective pharmacokinetics are shown in Figure 47 in the appendix. In a recent study, the clinical significant contribution of R-warfarin to the biological effect was shown [82]. This indicates that the R-warfarin is of clinical relevance and should be accounted for in the PK model.

In summary, we believe that by incorporating the known covariates for warfarin mechanistically in the reduced model obtained from the mechanistic blood coagulation model, the reduced model can help improve predictions of a more individualised loading and maintenance dose. This, in turn, will help to improve the efficacy and safety of warfarin therapy by achieving the target INR range faster as well as avoid over- or under-anticoagulation.

Rivaroxaban effect

In contrast to warfarin, rivaroxaban does not influence the blood coagulation system prior to activation. Therefore, to assess the effect of rivaroxaban it is sufficient to consider the activation of the system - more specifically - the *in vitro* assays: PT test.

The reference dynamic was chosen to be the dynamics of the blood coagulation network with an initial TF concentration of 100 [nM] and initial rivaroxaban concentration of 114.5 [nM] (average rivaroxaban concentration at so-called clinical 'steady state' of rivaroxaban therapy with 20 mg twice daily [36]) in the time interval [0, 90] seconds. The sensitivity based indices were determined according to eq. (83). Figure 27(A) shows the most important sensitivity based ir-indices during the time interval of interest.

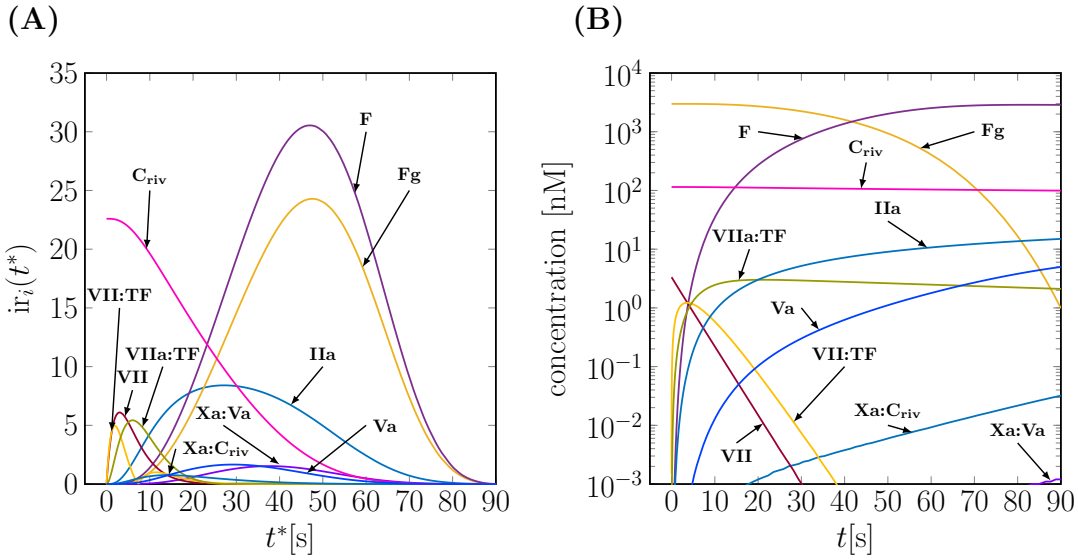


Figure 27: **Sensitivity based ir-indices for the rivaroxaban-fibrin system and time profiles of the corresponding factors**. **(A)** Shown are the sensitivity based ir-indices of the most important state variables for the time interval [0, 90] seconds. Sensitivity-based ir-indices not shown are below 0.5 in magnitude. **(B)** Concentration-time profiles of the most important molecular species (according to the ir-indices) predicted by the original 66-state variable model.

The differences to the high TF PT test setting are more apparent in form of the resulting reduced model. To this end, we exploit one mass conservation law to eliminate the molecular species Xa:C_{Riva}. The state variable to be eliminated was determined in a pre-run (see the

last paragraph in section 4.5). Next, we recursively eliminate state variables in order of increasing magnitude of their sensitivity based ir-index (as described in detail in section 4.5). In all cases, the resulting relative approximation error was required to satisfy $\varepsilon_{\text{rel}} < 0.05$ (see Eq. 86). Surprisingly, factor Xa remains only implicitly present in the reduced dynamics. The resulting reduced model comprised 8 dynamic state variables and 2 state variables describing the pharmacokinetics of rivaroxaban.

The reduced model is depicted in Figure 28, in contrast to the reduced model for the high TF PT test setting (cf. Figure 22) factors Va, Xa:Va and VII are included as dynamical state variables. Furthermore, as is already evident from the sensitivity based ir-indices factor Xa loses its importance once a rivaroxaban concentration is present in the system dynamics of the PT test. This can be explained by the rapid formation of the Xa:C_{Riva} complex, such that factor Xa can actually be considered in quasi-steady state.

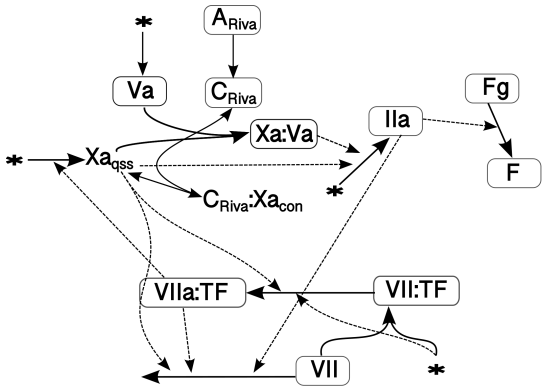


Figure 28: **Reduced model for rivaroxaban therapy**. Shown are the dynamical state variables after application of our model order reduction technique (as described in detail in section 4.5) with user defined error tolerance of 5%. The environmental state variables (indicates by '*') are II, V, X and TF. The state variable approximated by its quasi-steady state is indicated by the subscript 'qss'. The state variable eliminated by mass conservation is indicated by the subscript 'con'.

Similarly, as for the warfarin effect model obtained from the blood coagulation network model, we again aim to compare the mechanistic model prothrombin time predictions with an available pharmacodynamic model. In section 2.4 two different PD model for rivaroxaban were presented. We will employ the empirical PD model by [36] to investigate the plausibility of the resulting prothrombin time (PT) prediction of the reduced mechanistic model. For the rivaroxaban models, we compare the typical prothrombin time prediction of the original (blue line) and reduced (red dashed line) mechanistic model with the 90% CI of the empirical PD model (cf. Figure 29) based on 1000 simulations. The 90% CI of the empirical PD model shows the effect of the parameter uncertainty in the PD parameters that were reported in [36]. We assumed the parameters to be uncorrelated since no uncertainty covariance matrix was published in [36]. The PK model was chosen to be the same for both mechanistic and empirical PD model.

The comparison of the different model predictions shows a clear discrepancy between the

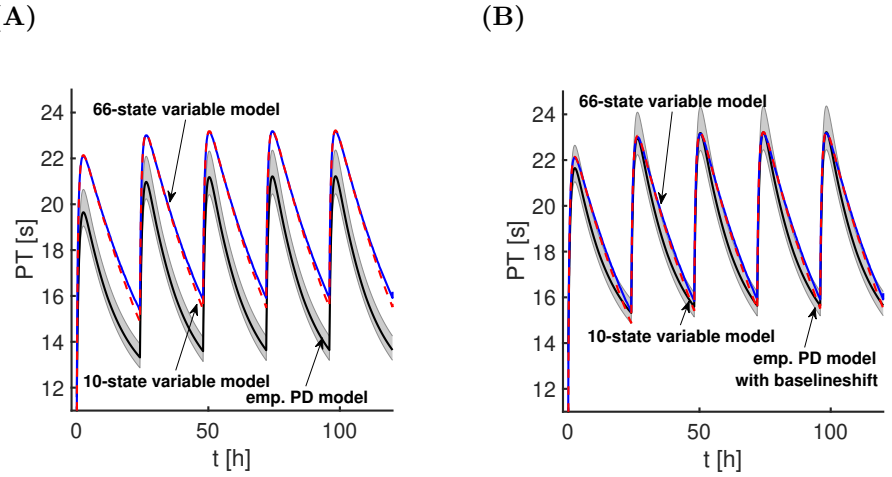


Figure 29: **Comparison of prothrombin time (PT) for rivaroxaban therapy of 20mg daily** based on the original 66-state variable model (solid blue), the elimination-reduced 10-state model (dashed red) and the empirical PD model in [36] with median (black) and 90% CI (grey area) in **(A)** and **(B)** with increased typical baseline value for empirical PD model by 2 seconds.

mechanism based and the empirical PT time predictions (cf. Figure 29(A)). The mechanism based PT time predictions are 2 seconds higher than the typical empirical PD model prediction. In order to show that the discrepancy is due to different assumed baseline in the different models, we increased the typical baseline level by 2 seconds in the PT time predictions of the empirical PD model (cf. Figure 29(B)). Consequently, in Figure 29(B) one can see that the mechanism based predictions lie within the 90% CI of the empirical PD model with the updated typical baseline value. Further, it is apparent that the predicted curves during one dosing interval decline slightly different between the two models. This might suggest model miss-specifications.

In future, the obtained reduced models for warfarin and rivaroxaban (cf. Figure 25 and 28) could be applied to perform the population analysis of clinical data. Although the reduced models obtained for warfarin and rivaroxaban from the large-scale mechanistic model depend on the chosen reference solution and time interval, the model order reduction technique allows for easy extraction of a new reduced system depending on the desired accuracy and requirements on the model. Model order reduction can thus answer the questions 'How to identify an appropriate pharmacodynamic model' as well as 'What model should be employed'. It should be noted that due to the (yet large) number of state variables and parameters of the reduced models for warfarin and rivaroxaban, both models might not be identifiable when used in a population analysis approach and a further reduction might be necessary to achieve full identifiability. However, our introduced model order reduction technique allows for a substantial reduction while at the same time retaining the physiological interpretability. In future, it should additionally be investigated how to implement covariates as well as random interindividual variability and how to incorporate their possible effects into the model reduction technique to obtain reduced models that automatically consider these components.

5.2. EGFR signalling cascade

The biological background and a comprehensive mathematical model for the EGFR signalling cascade were introduced in section 2.2. As previously described, the signalling cascade is activated by binding of EGF to EGFR and ultimately results in a transient ERK-PP output signal (cf. Figure 3). In view of the input (EGF) and the output (ERK-PP), the signalling cascade is in this section sometimes termed the EGF–ERK-PP system.

While the principal flow of the downstream signalling and activation sequence is well known, the relative importance of the different pathways (with/without Shc, membrane-bound vs. internalized forms) and its molecular constituents is still not well understood. We have, therefore, chosen this system to illustrate the potential of the sensitivity based input-response indices in gaining a detailed understanding of the transient behaviour of complex networks.

5.2.1. Model reduction of the EGF–ERK:PP signalling cascade using sensitivity based input-response indices

The reference dynamics was chosen to be the dynamics of the EGFR signalling cascade with a constant EGF concentration of $5 \cdot 10^{-8}$ molecules per cell in the time interval [0, 100] min as in [55]. The sensitivity based ir-indices were determined according to eq. (83). Figure 30(A) shows the most important ir-indices during the onset of signal propagation up to 0.3 min, while Figure 31(A) shows the time interval up to 3 min, when the output signal ERK-PP already peaks (compare also inset in Figure 36). The indices nicely illustrate how the signal propagates through the network. This signal propagation is not explicitly inferable from the concentration-time profiles (see corresponding panels (B). A detailed discussion is given in section 5.2.2. The complete list of input-response indices, also given on the full time interval, is shown in the appendix (cf. Figure 48-56).

In the attempt to understand the EGF–ERK-PP system, we first used the ir-indices to determine a reduced model of the signalling system, which subsequently guided our view and understanding of the activation cascade. This, in turn, allowed us to understand the transient dynamics and critical stoichiometry of key molecular species of the signalling pathway by analysing their ir-indices.

At start, all state variables were ordered according to the magnitude of their input-response index; see Figure 32. In a first reduction step, we exploited mass conservation laws to eliminate eight molecular species (see Table 16 in the appendix, column 'con'). The state variables to be eliminated were determined in a pre-run (see the last paragraph in section 4.5). Next, we recursively eliminated state variables in the order of increasing magnitude of their input-response index. This included complete removal from the reaction system (neglected state variables) or considering the state variables as being constant (environmental state variables). In all cases, the resulting relative approximation error was required to satisfy $\varepsilon_{\mathrm{rel}} < 0.1$ (see Eq. 86). This error tolerance was chosen to allow a substantial reduction of the system, while still maintaining a good approximation of the output signal. As a result, additional 42 state variables were eliminated (see Table 16 in the appendix, columns 'neg' and 'env'). Finally, we exploited time-scale separation, i.e., a quasi-steady state assumption,

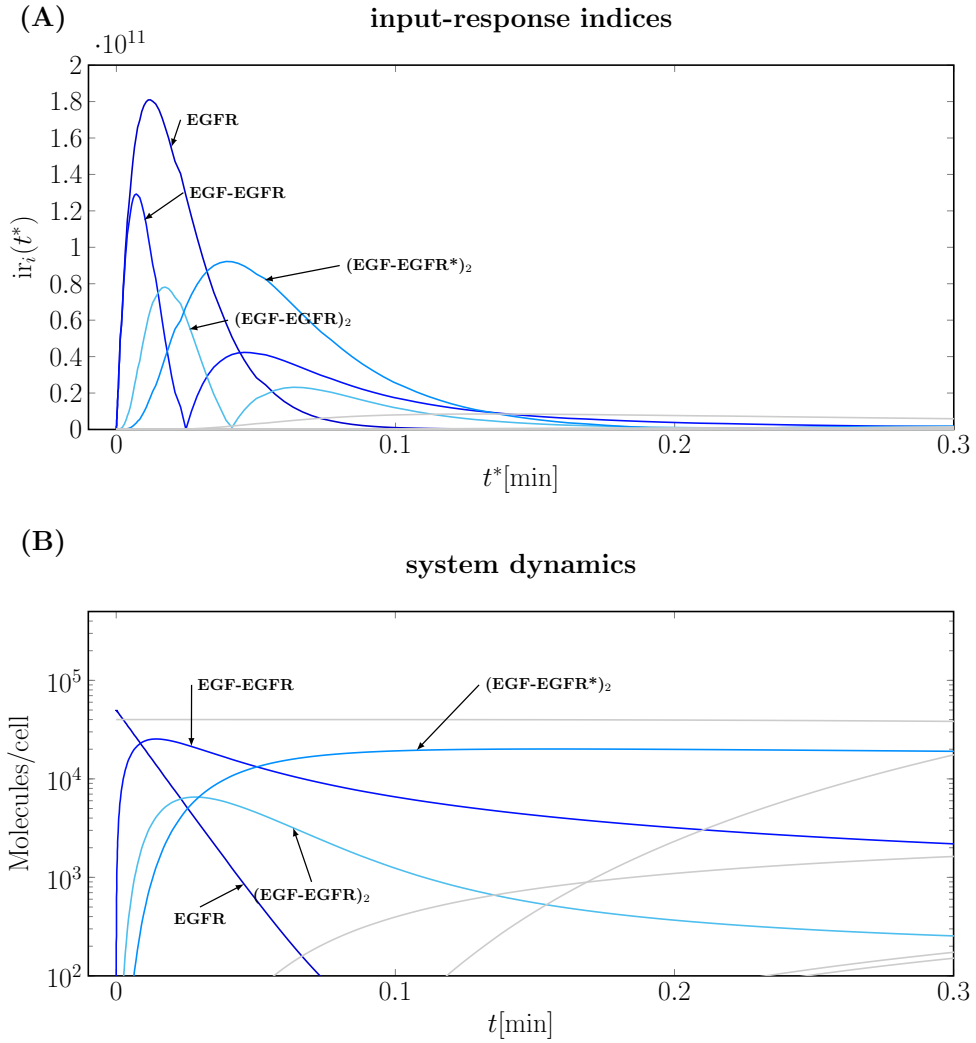


Figure 30: **Sensitivity based ir-indices and concentration-time profiles during onset of signal propagation.** State variables with largest input-response indices **(A)** and corresponding concentration-time profiles **(B)** during the time interval [0, 0.3] min.

to eliminate additional 34 state variables (see Table 16 in the appendix, column 'qss'). These molecular species are implicitly present in the reduced model, comparable to the substrate-enzyme complex in the Michaelis-Menten approximation. Surprisingly, also Raf* and some of its complexes are only present implicitly due to their fast dynamics. The resulting reduced model comprised only 28 dynamic state variables (modelled via differential equations). Thus, the number of state variables was reduced by approximately 75%.

The reduced model is depicted in Figure 33. It provides a clear view on the propagation of the input signal through the signalling cascade. While the first part of the signalling cascade is a linear sequence of activation steps (from EGF to (EGF-EGFR*)$_2$-GAP-Shc*-Grb2-Sos), the remaining part of the network consists of a cyclic module and a sequence of activation cascades. The complex (EGF-EGFR*)$_2$-GAP-Shc*-Grb2-Sos is key in the activation of Ras-GDP to Ras-GTP. Finally, the output ERK-PP feeds back to the degradation of (EGF-

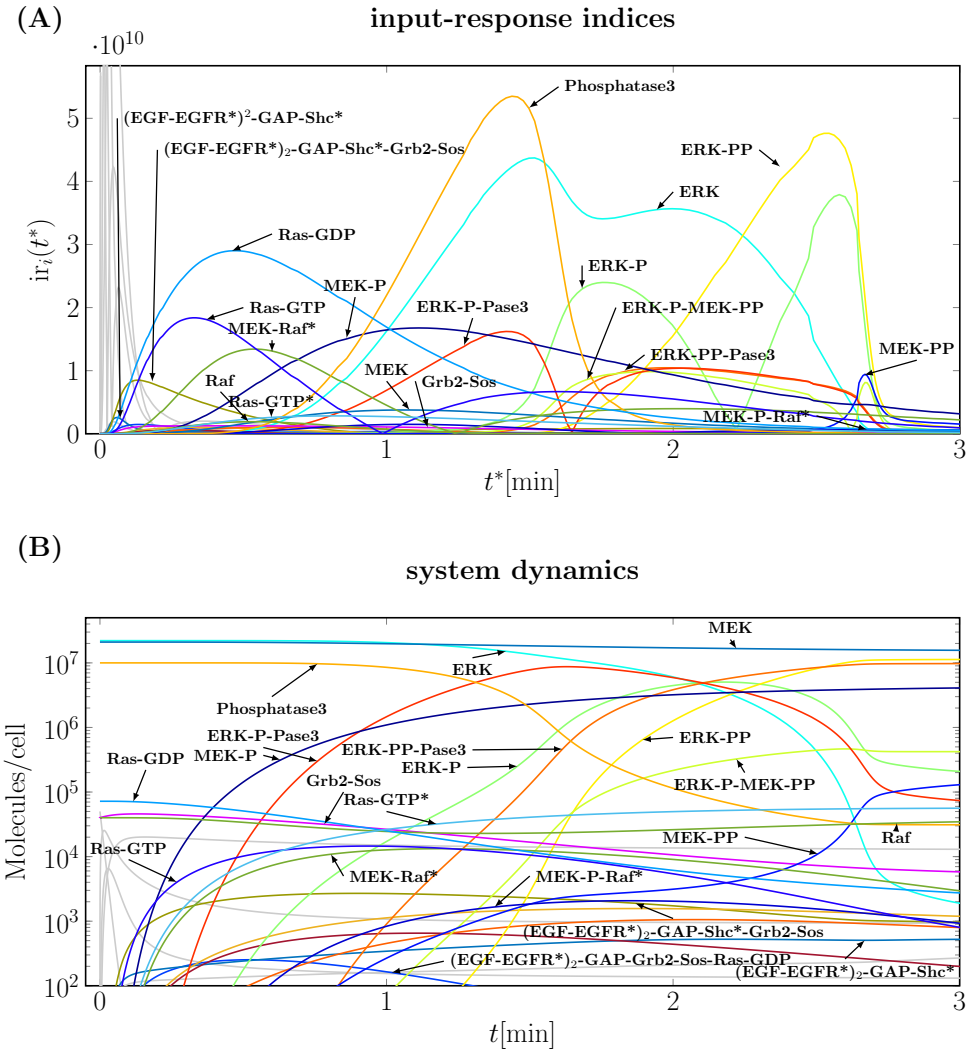


Figure 31: **Sensitivity based ir-indices and concentration-time profiles until peak of output signal.** State variables with largest input-response indices **(A)** and corresponding concentration-time profiles **(B)** during the time interval [0, 3] min.

EGFR*)$_2$-GAP-Shc*-Grb2-Sos, thereby attenuating the input signal.

The EGFR signalling network comprises membrane-bound and internalised forms of the receptor and its complexes. Previous reduced models of the EGFR signalling pathway simply neglected those molecular species as a way to reduce the order [4, 21]. In contrast, our approach highlights the relevance of these internalised forms. They are considered explicitly, like (EGF-EGFRi*)$_2$-GAP-Shc*-Grb2-Sos, or implicitly in form of quasi-steady state species (Table 16 in the appendix, column 'qss', and Figure 33) and constitute an alternative way to activate Ras-GDP.

Figure 32 shows the maximal value of the sensitivity based ir-indices for all state variables. The classification of a state variable into dynamic (dyn), environmental (env), neglected (neg), in quasi-steady state (qss) or conserved (con) is colour-coded. As can be inferred from the Figure, only states with a larger maximal value of the ir-index were considered as

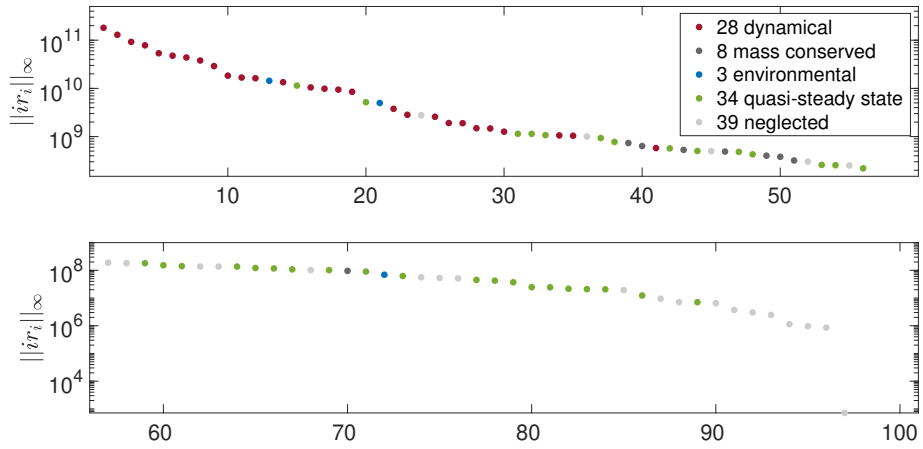


Figure 32: **Maximal value of the ir-indices in decreasing order and corresponding state classification.** Indices were determined based on a constant reference stimulus of EGF=$5 \cdot 10^{-8}$ molecules per cell and a time interval of [0, 100] min as in [55]. The relative approximation error was set to $\varepsilon_{\text{rel}} < 0.1$. The colour coding indicates the classification of a state as: dynamical (red), mass conserved (dark grey), environmental (blue), in quasi-steady state (green) and neglected (light grey). Input-response indices of value zero are not shown. The bottom panel is the continuation of the top panel.

dynamic. Amongst the states with the 30 largest values, only five states were not classified as dynamic. Beyond the 30 largest values, most state variables were either classified as neglected or in quasi-steady state.

Finding the lowest order reduced model is an optimisation problem suffering from the 'curse of dimensionality'. Therefore, model reduction approaches of large networks typically proceed sequentially. Figure 32 a-posteriori underpins the usefulness of the ir-indices to determine an informed order of state variables for the sequential reduction process. In contrast to many greedy approaches, this order is based (via the ir-indices) on the *original* model (determined prior to the reduction process), rather than recursively based on some *intermediate reduced* model.

5.2.2. Understanding the dynamics of EGFR signalling using the input-response indices

Figures 30+31 show the input-response indices and the temporal change of the 28 dynamic state variables of the reduced model. They allow to get further insight into the two questions: When is a state variable important? Why is it important?

The input-response index is a time dependent measure of importance. It gives insight about when a state variable is important for a given input-response relationship. Figure 30(A) shows the input-response indices on the time interval [0, 0.3] min. The indices show that the importance of EGFR is restricted to the first 10-20 sec and does not influence the system in the sequel (cf. Figure 48-56 given in the appendix). The indices for EGF-EGFR and (EGF-EGFR)$_2$ become small in comparison to the other indices ($< 10^9$). This is due to the fact

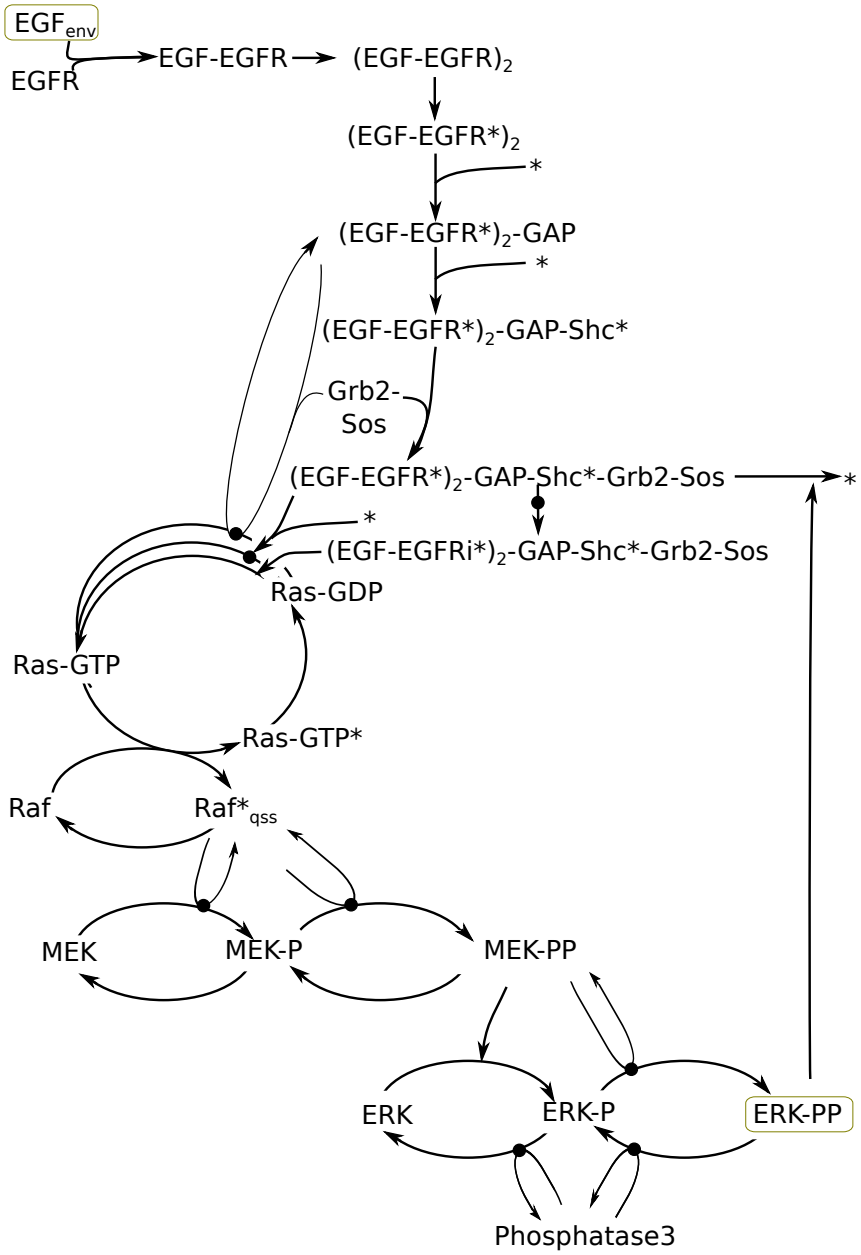


Figure 33: **Graphical representation of the reduced EGF–ERK:PP signalling cascade based on ir-indices.** The reduced model comprises 28 dynamical state variables (75% reduction compared to the original model) and was derived based on a relative error tolerance of 10%. The environmental state variables are indicated by '*'. For a clearer layout, dynamical state variables representing intermediate complexes (e.g., ERK-PP-Pase3) are not explicitly listed, but rather indicated by a dot on the corresponding reaction arrow. For better readability, we included the Raf* although it was eliminated via quasi-steady state approximation.

that at those time points the input perturbation does not lead to a substantial perturbation in the concentration-time profile of these molecular species. Most other state variables are not yet important during this time span and have indices that are too small to be visible

in the figure. Figure 30(B) shows the concentration-time course of the corresponding state variables. In contrast to the indices, it does not become evident by only considering the time profiles, when a state variable has a large impact on the system.

Figure 31(A) shows the input-response indices on a 10x larger time interval [0, 3] min. While some of the indices meet our expectations (e.g., temporal ordering of Ras-GTP, MEK-Raf*, MEK-P etc), others are more surprising, as the indices of phosphatase 3 (most prominent on [0.5, 2] min) or MEK-PP (only visible after 2.5 min). In combination with the time profiles, we infer that upon phosphorylation of MEK-P, double phosphorylated MEK immediately forms a complex with ERK. Once ERK has been phosphorylated, the dominant binding partner seems to be phosphatase 3 (see a strong increase of ERK-P-Pase3 starting roughly at 0.5 min). Analogously, any ERK-PP seems to immediately bind to phosphatase 3, resulting in a strong increase of ERK-PP-Pase3 up to 2 min. Only when the impact of phosphatase 3 is sufficiently diminished (in form of reduced free phosphatase 3), ERK-PP levels increase. Thus, one effect of phosphatase 3 is a delay in the onset of ERK activation. Due to the fast complex formation of MEK-PP with ERK and ERK-P during the first 2.5 min, its input-response index is low (almost quasi-steady state condition of MEK-PP). Only after ERK-PP decreases, the impact of MEK-PP becomes more dominant.

A large input-response index may be due to a large controllability coefficient or a large observability coefficient (cf. eq. (83)). Figure 35(A) depicts the controllability indices, i.e., how the EGF signal influences the state variables. The system is mostly controlled by EGF during the time window of [0, 3] min. Largest controllability is exerted on the ERK module, in particular, ERK and ERK-PP. Given that EGF is the stimulus for ERK phosphorylation, this result is in line with expectations. The observability indices, shown in Figure 35(B), provide a measure of how the state variable at time t^* impacts the output on the remaining time interval. Interestingly (EGF-EGFRi*)$_2$-GAP-Shca-Grb2-Sos has initially the largest observability index. To understand this phenomenon, we performed an in silico experiment with an additional stimulus of (EGF-EGFRi*)$_2$-GAP-Shca-Grb2-Sos (magnitude $5 * 10^3$ molecules/cell at time t=0), see Figure 34. The resulting output demonstrates that (EGF-EGFRi*)$_2$-GAP-Shca-Grb2-Sos does have a large impact on the time course of ERK-PP, rising to much lower values. If, however, the same stimulus of (EGF-EGFRi*)$_2$-GAP-Shca-Grb2-Sos is given at a t=3 min, the output signal of ERK-PP is only very slightly perturbed.

This distinction into controllability and observability indices allows to understand due to which processes the state variables might be important. If the controllability dominates, the state variable is largely influenced by changes in the input signal. If the observability index dominates, then small changes in the level of the state variable have a large impact on the output concentration.

5.2.3. Further reduction of the model complexity by applying proper lumping

As the final step of the model reduction process, we applied proper lumping to the 28-state variable model, using the same threshold on the approximation error as before. This resulted in a lumped 19-state variable model that implicitly includes 3 environmental and 34

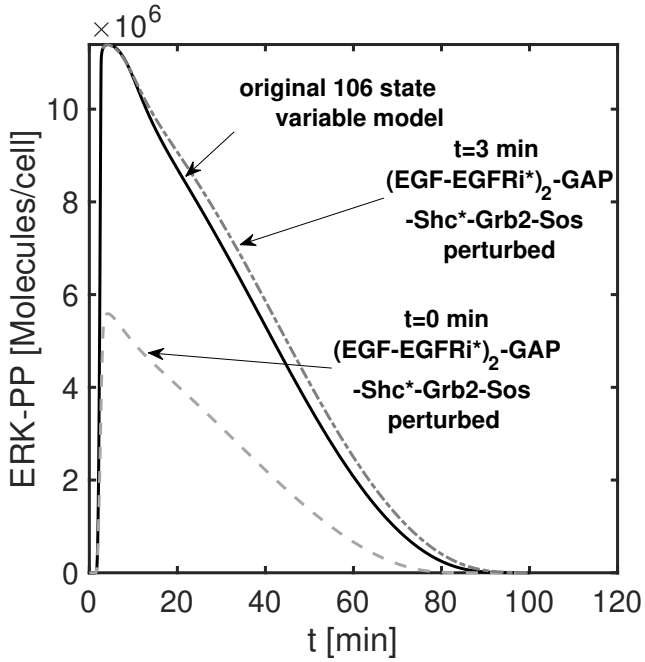


Figure 34: **Comparison of ERK-PP (output) of the original 112-state variable model** based on three different inputs: a) original input of EGF only (black line); b) additional input of (EGF-EGFRi*)$_2$-GAP-Shca-Grb2-Sos at time $t = 0$ min of magnitude $5 * 10^3$ molecules/cell (dashed light grey line); c) additional input of (EGF-EGFRi*)$_2$-GAP-Shca-Grb2-Sos at time $t = 3$ min of magnitude $5 * 10^3$ molecules/cell (dot-dashed dark grey line).

quasi-steady state variables. Both reduced models, the 28-state and lumped 19-state model reproduced the transient output signal ERK-PP with the desired accuracy (see Figure 36).

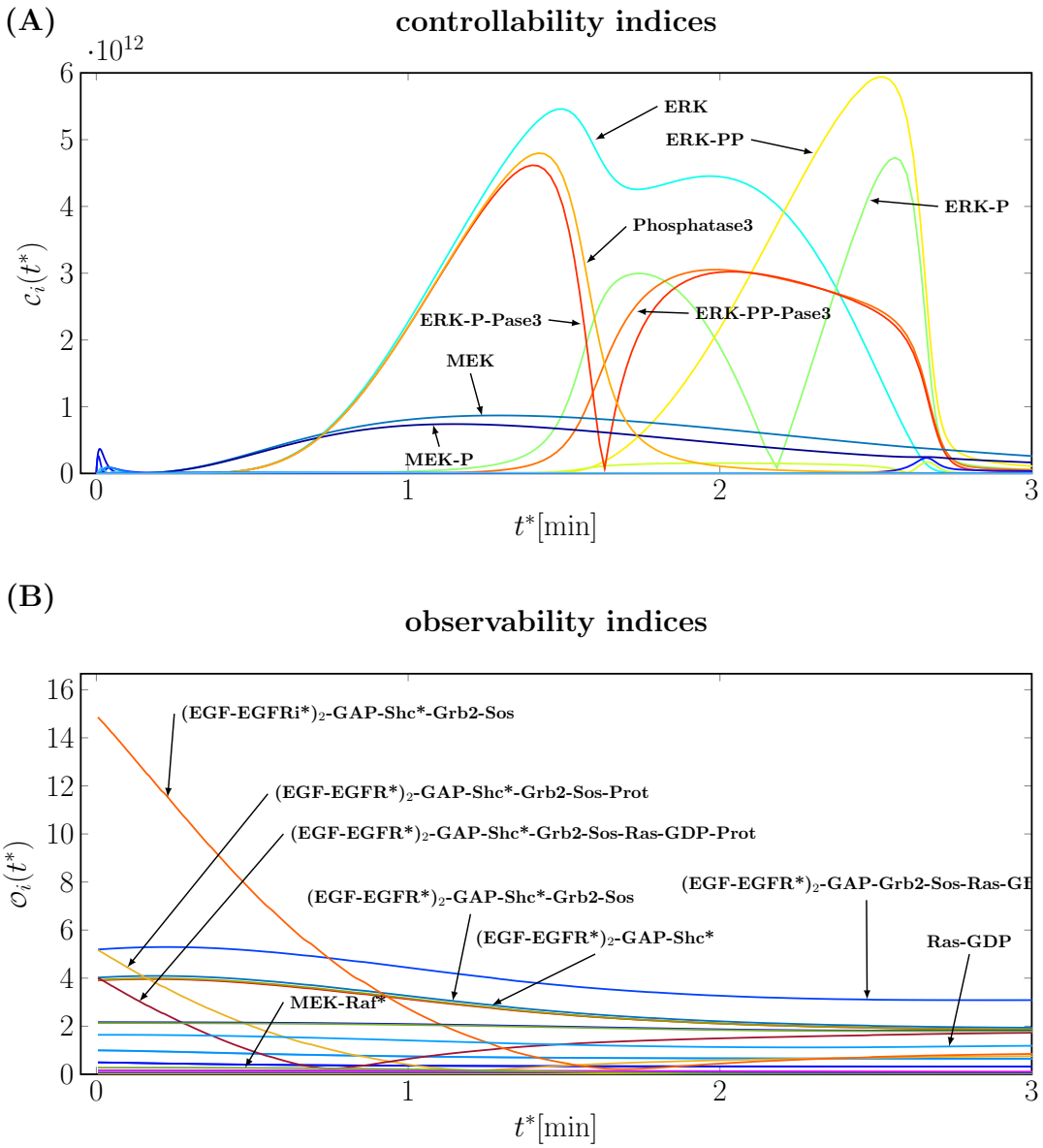
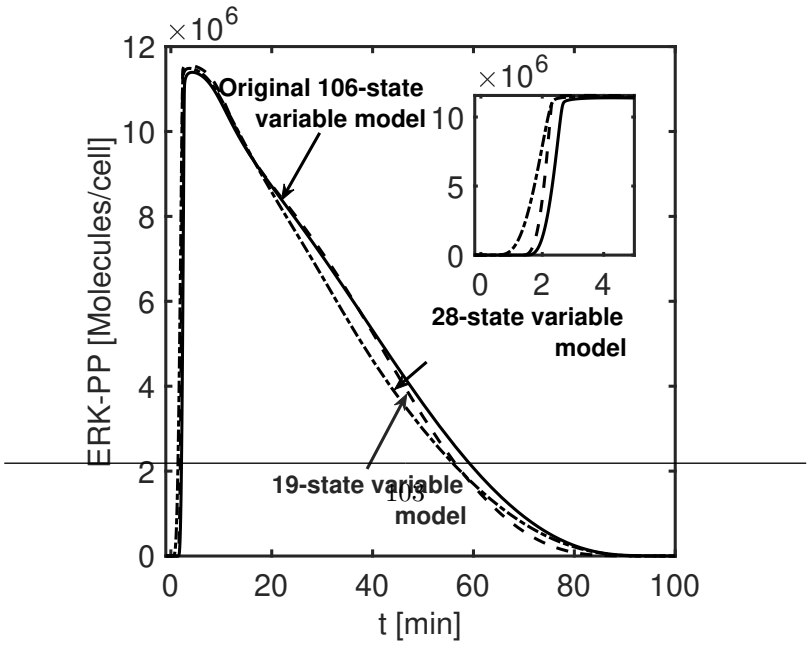


Figure 35: **Controllability and observability indices of the dynamical state variables of the reduced model (A)** Controllability indices and **(B)** observability indices of the dynamical state variables of the reduced model for $t^* \in [0, 3]$ min.



5.3. Comparison of different MORs

In this section, we provide a comparison of the following model reduction techniques: sensitivity analysis, proper lumping, ILDM and balanced truncation introduced in section 3 to reduce the complexity of a specific pharmacological example of interest: the PT test of the blood coagulation model. We will discuss the advantages and disadvantages of each method in the application context. The different magnitudes of the input (TF concentration) are chosen as 5pM (low TF scenario) and 100nM (high TF scenario) (cf. section 5.1.2). The fibrin concentration is chosen as a surrogate for the response. The user-defined error tolerance was chosen to be 20% for each model reduction method.

Employing scaled sensitivity coefficients for model order reduction

First we use the local sensitivities of the system, obtained by solving the variational equation in eq. (17). Usually for systems pharmacology/biology models the sensitivities are computed for the steady state of the system. However, in the PT test setting we are interested in the transient behaviour of the fibrin concentration $x_F(t)$ and, thus, compute the scaled sensitivity $\hat{S}^p_{F,j}(t) = \frac{\delta x_F(t)}{\delta p_j} \cdot \frac{p_j}{x_F(t)}$ of each $t \in [0, 30]$ seconds for the high TF case and $t \in [0, 240]$ seconds for the low TF case with respect to each parameter. The resulting scaled sensitivities with respect to the parameters are given in Figure 37 and the scaled sensitivities with respect to the initial condition are given in Figure 38.

(A) **(B)**

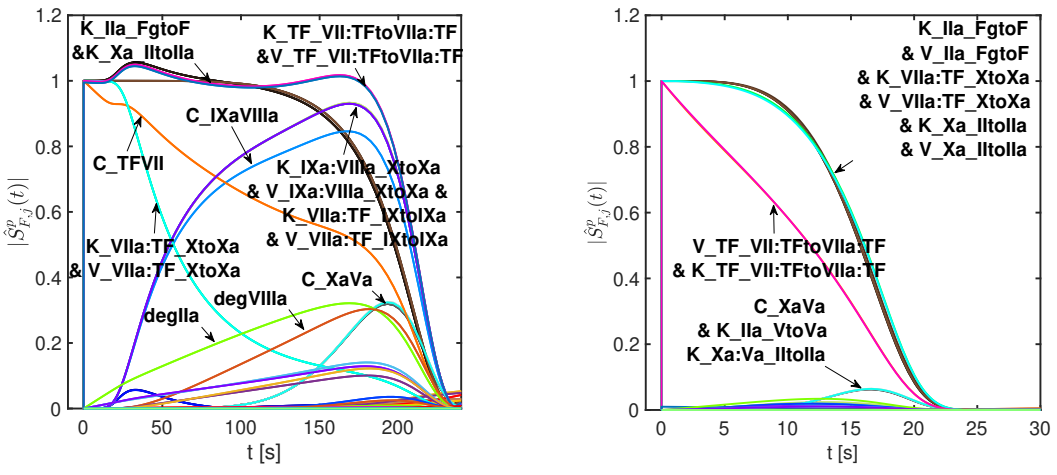


Figure 37: **Absolute scaled parameter sensitivities of the output x_F (fibrin concentration) over time for two PT test scenarios** for **(A)** low TF and **(B)** high TF. Shown are the sensitivities of the most important parameters as a function of time for the **(A)** low TF and **(B)** high TF scenario.

(A) **(B)**

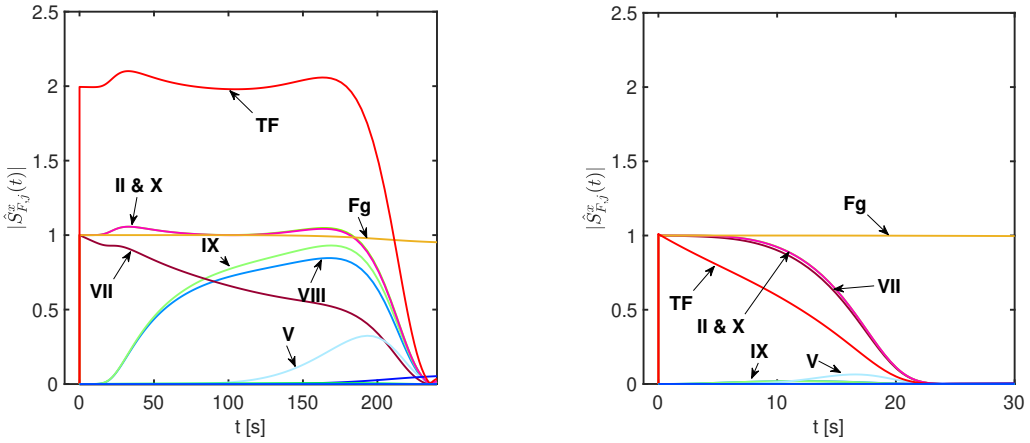


Figure 38: **Absolute scaled sensitivities of the output x_F (fibrin concentration) with respect to the initial condition for two PT test scenarios** for **(A)** low TF and **(B)** high TF. Shown are the sensitivities of the most important initial conditions as a function of time for **(A)** the low TF and **(B)** high TF scenario.

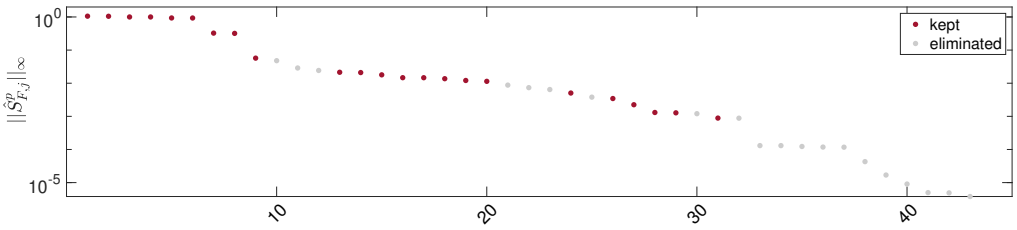
The model reduction based on the sensitivities with respect to the parameters was performed in a similar manner as the first step of our novel model reduction technique (cf. section 4.5): First computing the scaled parameter sensitivities for all parameters. Then, we ordered all of the parameters according to increasing maximal value of their scaled parameter sensitivities. In the iteration, each parameter according to the previously determined order is considered for elimination. The parameter is only set to zero if the relative l2-error is below the user-defined threshold[4]. The most important parameters according to their maximal value of the scaled parameter sensitivities for each setting are given in Figure 39. As was already stated in section 3.1.3 some parameters which were considered important based on the scaled sensitivities with respect to parameters were eliminated in the process and some parameters which were considered less important were kept in the reduced model. In Figure 39 the classification of parameters into eliminated and kept is color coded. For the high TF setting all parameters from the reduced model are shown, for the low TF case parameters with very low sensitivity, which are not included in the figure, are kept in the reduced system.

The resulting reduced models are given in Figure 40. All degradation reactions were eliminated from the model in the high TF setting. The similarities between these reduced models in Figure 40 and the elimination-reduced model in Figure 22 are obvious. This can be explained by the similarities in the model reduction technique and the fact that the input-response indices can be derived as the product of two sensitivity coefficients with respect to the states (cf. section 4.4). The fact that the state variables II, VII, X can be considered to be constant and thus environmental states do not become obvious with using the sensitivities with respect to the parameters or initial conditions. Rather based on Figure 38 the state variables II, VII, X and Fg would have been considered important for the system dynamics.

[4]To allow for parameter elimination the model was re-parameterised without changing the systems dynamics.

Important in which sense can not be distinguished since it is not clear how to extend the impact of state variables on the system dynamics beyond the initial time point.

(A)



(B)

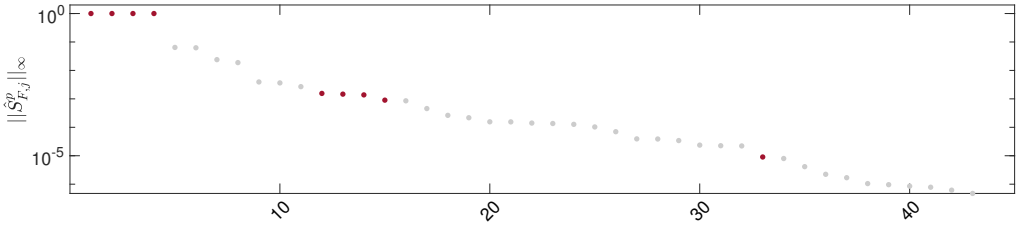


Figure 39: **Magnitude of the scaled parameter sensitivities of two PT test scenarios** for **(A)** low TF and **(B)** high TF. Shown are the sensitivities of the most important parameters (a forth of all parameters) for **(A)** the low TF and **(B)** high TF scenario. The color coding indicates how the parameter was classified in the reduction process: kept (red) and eliminated (grey).

(A) **(B)**

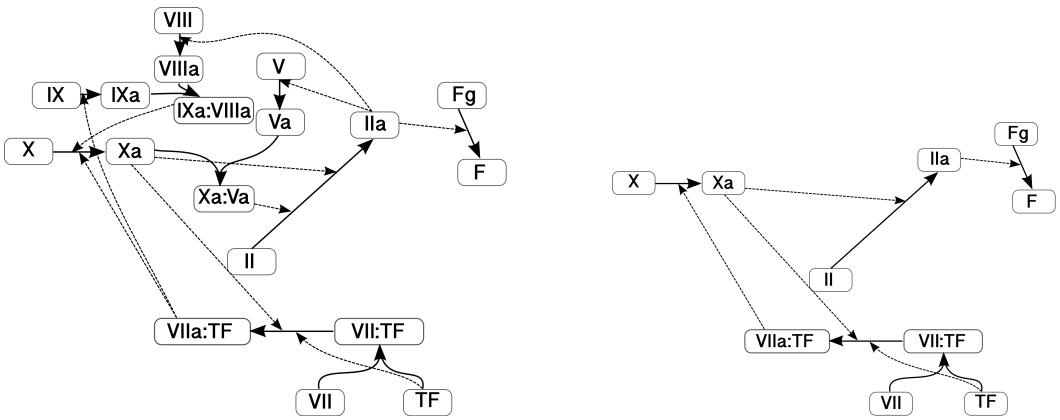


Figure 40: **Reduced models for two PT test scenarios based on parameter sensitivities** for **(A)** low TF and **(B)** high TF. Shown are the dynamical state variables. See Figure 6 for a legend of the different arrow types.

In the case of using sensitivity analysis for model reduction not actually the scaled sensitivities

(cf. Figure 37 and 38) but the reduced models (which are comparable to ones obtained by our model reduction technique) give insight into the important parts of the system. This stresses the point that our introduced measure of importance (the ir-index) uses the state sensitivities in an efficient way to immediately filter the important dynamics.

Proper lumping using a recursive greedy search strategy

In this section, we use proper lumping to reduce the PT test setting for the two settings (low and high TF). We have used the previously mentioned recursive greedy search strategy [1]. The resulting lumped models are given in Table 9 and Table 10. As with our novel model order reduction technique in Table 8, one obtains two different reduced models for the different PT test settings (low and high TF). However, the interpretability of the lumps e.g. x_{L1} and x_{L2} in Table 9 is obscured in contrast to the reduced models in Table 8. Physiologically it is intuitive that the snake venoms do not play a role for the PT test, they are, however, included in the lumped state variables. Furthermore, we can not clearly distinguish based on the automated greedy lumping search why we obtain two different reduced models for the PT test.

Table 9: **Resulting reduced model for the low TF PT test setting obtained by proper lumping** using a recursive greedy search strategy [1] with $\varepsilon_{\mathrm{rel}} \leq 0.2$.

Lumped state variables	**Original state variables**
x_{L1}	XII, XI, X, V, VIIa: TF: Xa: TFPI, TAT, FDP, XIII, Pg, PC, VK, VKH2, VKO, $\mathrm{VK_p}$, PS, Pk, AWarf, CWarf, AUC
x_{L2}	XIIa, Xa:Va, D, APC, Tmod, IIa: Tmod, K, CA, AVenom, CVenom, A_Enox, ENOp, Taipan Venom, AT: III, delay Taipan CMT 1, delay Taipan CMT 2, AVenom Tiger, CVenom Tiger, AT: III: UFH
x_{L3}	VIII, VIIIa, TFPI
x_{L4}	IX, II
x_{L5}	IXa, XIa, VIIa, Va, Xa: TFPI, XF, XIIIa, P, APC: PS, AT: III: Heparin
x_{L6}	VII
x_{L7}	Xa, VIIa: TF
x_{L8}	IIa
x_{L9}	Fg
x_{L10}	F
x_{L11}	IXa: VIIIa
x_{L12}	TF
x_{L13}	VII: TF

Table 10: **Resulting reduced model for the high TF PT test setting obtained by proper lumping** using a recursive greedy search strategy [1] with $\varepsilon_{\text{rel}} \leq 0.2$.

Lumped state variables	**Original state variables**
x_{L1}	XII, VIIIa, IX, XI, X, V, II, VIIa: TF: Xa: TFPI, TAT, FDP, XIII, Pg, PC, TF, VK, VKH2, VKO, VK_p, PS, Pk, AWarf, CWarf, AUC, AT: III, AT: III: UFH
x_{L2}	XIIa, VIII, IXa, XIa, VIIa, Xa: TFPI, XF, XIIIa, TFPI
x_{L3}	VII, Xa, Va, P, IXa: VIIIa, VII: TF, VIIa: TF
x_{L4}	Xa:Va, IIa
x_{L5}	Fg
x_{L6}	F
x_{L7}	D, APC, Tmod, IIa: Tmod, APC: PS, K, CA, AVenom, CVenom, AT: III: Heparin, A_Enox, ENOp, Taipan Venom, delay Taipan CMT 1, delay Taipan CMT 2, AVenom Tiger, CVenom Tiger

Time scale separation - ILDM method

In this section, we use the ILDM method to reduce the blood coagulation model for the PT test setting. The ILDM method (cf. section 3.1.2) is based on the Jacobian of the system in a specific state. Usually, the steady state is chosen for the ILDM method. For the PT test setting the steady state is zero, we have therefore chosen the initial state. We have eliminated the state variables that were known to be zero for this reduction. The ILDM based on the Jacobian in the initial state did lead to a resulting reduced model of dimension 45 for the low TF and no further reduction for the high TF setting.

Using the ILDM based on the Jacobian of the initial state as a subsequent step after our elimination model reduction step, we are actually able to reduce the system for both high and low TF case to a system with only one dynamical state variable for the given error tolerance.

We can conclude that ILDM can not be applied to the blood coagulation system, in particular, the PT test setting without preconditioning. Although highlighting the fact that time scale separation is a powerful model reduction technique, it can not be used as a single tool for large-scale models.

Gramian based model reduction - balanced truncation

In this section, we use balanced truncation based on the time-limited empirical gramians (presented in section 4.1). The time-limited empirical gramians were determined based on ten perturbations ranging from 50% to 150% of the reference input and initial conditions. An important prerequisite for the application of the balanced truncation method (cf. section 3.2.2) is that the system is fully controllable and observable. Unobservable and uncontrollable state variables can be deleted by identifying zero rows and columns in the empirical controllability and observability gramian respectively. The minimal system where each state variable is both controllable as well as observable of the PT test consists of 27 state variables for the

high TF case and 29 state variables for the low TF setting. Using the time-limited empirical gramians, we were able to reduce the system to 24 state variables. No further reduction was possible due to the stiffness of the transformed system leading to unfeasible small step sizes with an implicit ode solver, such that the integration with ode15s was aborted.

Table 11: **Summary of the results from the comparison of using IR based MOR (cf. section 4.5), sensitivity analysis, proper lumping, ILDM and balanced truncation for model reduction of the blood coagulation for the PT test setting with different magnitudes of the input.**

Model reduction technique	**Dimension of the reduced model**	
	low TF	high TF
IR based MOR (cf. section 4.5)	13	6
MOR based on parameter sensitivity coefficient	18	10
Proper lumping	13	7
ILDM	46	47
Balanced truncation	24	24

In Table 11 the resulting model dimensions obtained from the different model reduction techniques are summarised. As expected model reduction based on parameter sensitivity and proper lumping perform best in the context of the PT test setting. Both ILDM and balanced truncation required preconditioning in terms of removing state variables that are zero or that are uncontrollable/unobservable respectively. However, the dimension of the reduced model obtained by balanced truncation is comparable to the dimension obtained by model reduction based on parameter sensitivity and proper lumping in the low TF setting. Interestingly there exists no clear difference in terms of dimension of the obtained model from both ILDM and balanced truncation. This highlights that our introduced model reduction technique (described in section 4.5) using ir-indices based on the concept of empirical gramian and incorporating the consideration of environmental state variables as well as approximating some state variable by their quasi-steady state is a valuable addition in the field of model order reduction of systems pharmacology/biology models.

6. Discussion

In this section, we will set the main results of this thesis into a larger context, discuss features, limitations and give possible extensions for future work.

Local model reduction method

Model reduction techniques are often discussed in terms of 'local' and 'global' approaches. Typically, the connotation with 'local' is negative, meaning that one should aim for a 'global' approach (e.g., [31]). We believe, however, that the property of a reduced model being valid only locally can also be a positive feature, rather than a negative one. To be able to distinguish different local regimes, our ir-indices are defined with respect to a reference trajectory. It was illustrated for the PT test that different reference trajectories (resulting from different magnitudes of the inputs: low vs. high initial TF concentration) can give rise to different reduced models (cf. Figure 22). Rather than aiming for a single compromise reduced model, we believe that the difference between the two reduced models is very informative about the PT test, in particular about its ability to detect (genetic) deficiencies in the factors VIII and IX. Our results confirm the statement in [14] that only the low TF, but not the high TF prothrombin time test is sensitive to detecting genetic deficiencies of factors VIII and IX.

Multiple reduced models can also be integrated in a statistical setting. Assume that the model reduction of a large-scale model of a pharmacologically targeted system results in two different reduced models, depending on the reference trajectory. In a statistical context like NLME, one would then aim to estimate parameters based on a mixture modelling approach. One could further try to identify covariates in the model reduction process that indicate the expected reduced model. An approach based on a compromise reduced model, however, would most likely run into parameter identifiability issues—unless additional prior information within a Bayesian approach is assumed (as in [31]).

Dependence of ir-indices on time interval

The focus of this thesis was on relatively short observation periods for the various application examples. For applications to chronic progressive diseases, longer time periods are of interest. Then, slow changes in endogenous or exogenous factors may determine the rate of change in the system, while such changes are likely to be irrelevant on short time intervals. The model reduction should be based on a time span that covers the entire disease progression period. The application of our model reduction technique can then result in two different outcomes: for the period of chronic progression, the input-response indices of slow-changing endogenous or exogenous factors become important; or remain unimportant. In the first case, these factors are considered in the reduced model, while in the second case, these factors are not included. In the latter situation, either the model does not yet fully account for the importance of this slow changing factors (indicating that the model should be revised), or the factors do indeed not play an important role (indicating that clinical expectations might need to be revised).

Ranking of state variables according to ir-indices

The proposed model order reduction technique is based on a ranking of the state variables according to their importance for the input-response setting. The maximal value of the ir-index was chosen as a summary metric for the ranking. This choice was motived by the fact that in signalling cascades some state variables are important for very short periods of time and others are relevant during longer time periods. Therefore if a state variable is important at some point in time (even, if only for a very brief period of time), we consider it to be potentially relevant. Alternative metrics to rank state variables, like the AUC (1-norm) of the ir-index, are possible. In addition to the summary metric, the model reduction process also depends on the user-defined error tolerance δ (as for any model reduction technique) and the time span on which the dynamics was considered.

Global search strategies for complex networks are challenging due to the combinatorial explosion in the number of possibilities, therefore many model reduction approaches are greedy and recursive: the reduced model at the current step is a reduction of the reduced model obtained at the previous step. As a consequence, the final result can potentially depend on the sequence, in which the steps are taken. We showed for the brown snake venom-fibrinogen system that randomly chosen orders can result in different reduced models (cmp. Figure 16 with Figure 41). Note that all models are 'valid' approximations obeying the same criterion on the relative approximation error. When interpreting different reduced models, however, this dependence could result in confusion, since different researchers (using different algorithms) on the same problem could end up with different models. Our ir-indices (cf. Figure 19) are indicative for the two different reduced model structures of the brown snake venom-fibrinogen system since both, the pathway involving the factors P and APC as well as the pathway involving the factors VIII, IX and X are expected to impact the output Fg. We further infer that the first pathway is more important during the first 1h after envenomation, since in Figure 15, where the indices have been determined up to 1h, factors of the second pathway do not show up. Only beyond the 1h period, the second pathway becomes relevant. A purely 'objective' model reduction criterion in terms of meeting a given threshold criterion 'randomly' choses one of the two potential reduced model structures, depending on the sequence of reduction steps. It does, however, in general not give further information about alternatives.

Our model order reduction technique versus other model reduction techniques

A central component of our model order reduction approach is the elimination of state variables—either by completely neglecting them or by considering them as environmental variables. In [31] it is argued that lumping is more powerful than elimination, with the latter being almost a special case of lumping. We believe, in contrast, that the elimination of state variables is a very valuable and efficient model order reduction approach that moreover reflects biological and experimental expectations. When a complex pathway with multiple potential inputs and multiple outputs is stimulated by a specific input and analysed in terms of a specific output, one would expect parts of the pathway to be negligible for the given input-output relation. These parts correspond to state variables that will be eliminated by neglecting them. If moreover, the time scale of interest is such that some constituents do not change in a relevant way, but impact the output, we would expect them to be eliminated by

considering them as environmental states. Lumping all these state variables into one or more pseudo-state variables might formally result in the same approximation quality—however the interpretation of the reduced model will be obscured as was shown in the context of the PT test in section 5.3. This can also be seen in [40], where after lumping 58 state variables into a single pseudo-state variable, it is given *a-posteriori* the interpretation of a specific coagulation factor (namely factor II) that meets the expectations of the authors, but that is not consistent with the pseudo-state variable originating from 58 state variables of the original model.

Various other model reduction approaches have been employed for the EGFR system. Existing reduced models, however, suffer from the loss of the transient output signal [4], high complexity [79] or heuristic model reduction techniques that are not easily translatable to other signalling cascades [21]. In addition, reduced models have been proposed by exploiting the redundancy within the EGFR signalling cascade, i.e., the existence of parallel pathways. These include the Shc-dependent and Shc-independent pathway as well as signal processing through membrane-bound or internalised receptors. While simply neglecting a parallel pathway will reduce the order of the model, it might lead to misleading conclusions when interpreting the full model based on the reduced one. In contrast, our approach maintains parallel pathways (cf. Figure 33) and is easily applied to other systems biology models. In addition, our model order reduction technique scales well with the size of the models due to the usage of the ir-indices to obtain a sequence for the sequential model order reduction procedure.

Link to sensitivity analysis

The theory of sensitivity analysis is well established; yet its use in model order reduction is unclear. It has been noted that a reduction based on the importance of states obtained solely by sensitivity analysis will result in reduced models with low approximation quality [123, 126]. The sensitivity based ir-indices show how the concept of sensitivity coefficients, suitably expanded, can be leveraged for model reduction.

Leveraging ir-indices for model analysis

Although the ir-indices were particularly developed for our model order reduction approach, they additionally allow to gain a better understanding of the model dynamics of large-scale mechanistic models. The ir-indices nicely illustrate how the signal propagates through the chosen network. This was examplified for the EGFR system, where several studies reported signalling via the internalised receptor species as being unimportant [79, 112, 55]. The sensitivity based ir-indices revealed, however, that not all of the internalised state variables can be neglected. For a given maximal relative error, some of the internalised species need to be considered dynamically or in quasi-steady state, as can be inferred from Table 33. As a consequence, they enter the differential equations implicitly via algebraic equations (like E^*_{qss} in eq. (84) entering the eq. for E in (85)). This way, although not explicitly present in the reduced model, the internalised species indirectly influence the dynamic behaviour.

A detailed study of the EGFR signalling pathway is presented in [55]. The authors analyse, which molecular species of the complex network do control the ERK-PP output profile. To this end, they *first* simplified the extensive reaction scheme given in Figure 4 into a simpler

textbook type cartoon [55, Fig. 4] using their expert knowledge on the studied system. The authors *then* quantified the relevance of reaction processes based on the concept of 'impact control coefficients'. Interestingly, our automated model reduction process resulted in a reduced model scheme (Figure 33) that has many features in common with their cartoon. In addition, it also shows informative differences. These include the mentioned relevance of the Shc-independent pathway.

Future work

There are some important open questions related to our novel approaches that will be subject to future research.

Approximation error for non-observed states

By construction, only the approximation error of the output variables is controlled during the model reduction process. We noticed, however, that the approximation error of the remaining state variables of the reduced model is also small during their time interval of importance (i.e., when their input-response index is large). This could potentially be theoretically studied to see if an a-priori analytical or numerical error bound can be derived based on the ir-indices. For now, if one needs to control the approximation error of additional state variables, then theses variables should be included as 'output state variables'.

Using time-limited gramians for model order reduction

The focus in this thesis was the development of a model order reduction technique for systems pharmacology/biology models. However, in section 4.1, we introduced time-limited empirical gramians and showed how these could be used for balanced truncation of the nonlinear RC ladder - a benchmark model from the field of control theory. For this numerical example we were able to demonstrate the superiority of the time-limited empirical gramians over (i) the analytical gramians based on linearisation in the steady state and (ii) the infinite time empirical gramians in the context of balanced truncation. To further demonstrate the usefulness of the time-limited empirical gramians, they should be applied to a wider range of control theoretic model systems.

Importance of state variables during specific time windows

The ir-indices indicate that state variables are only important during different time windows (cf. section 5.1.1). Thus a possible extension of the model order reduction technique is to obtain a set of reduced models for a specific discretisation of the time interval. In [104] such an approach was introduced based on trajectory piecewise-linearisation and Krylov methods for the model order reduction of nonlinear systems dynamics. In contrast the extension of our model order reduction technique would result in nonlinear models for each time window and, thus, more accurately approximate the nonlinear dynamics.

Generalising ir-indices to account for parameter uncertainty and time-dependent inputs

In this thesis, we introduced the ir-index for a fixed set of model parameter values of the large-scale model and for an input that is restricted to perturbations at the initial time only. Both assumptions can be relaxed. Uncertainty and variability in parameter values can be accounted for by considering different reference trajectories. Thus an important next step for the model order reduction technique based on the input-response indices is to show how robust the reduced model is with respect to parameter perturbations (region of the parameter space). The extension of the ir-indices for time-dependent inputs $u(\cdot)$ is also possible. In this case, the system equation in (73) changes to $\frac{\mathrm{d}x}{\mathrm{d}t}(t) = f(x(t), p) + g(x(t))u(t)$ with a function $g : \mathbb{R}^n \to \mathbb{R}^{n\times s}$, where s denotes the number of inputs into the system. In this notation the link to control theory becomes even more apparent.

Applicability of mechanistic PD model for population analysis of clinical data

We successfully extracted mechanistic PD models for both warfarin and rivaroxaban (cf. Figure 25 and 28). The emphasis was set on obtaining reduced models retaining the physiological interpretability. As a logical next step both the large-scale mechanistic model as well as the model order reduction technique need to be extended (by allowing for parameter perturbations). Into the large-scale mechanistic model covariates as well as interindividual variability should be incorporated. Then the extracted minimal mechanistic PD models can be employed for population analysis of clinical data. Additionally, the reduction process offers a systematic means to derive the covariate relationship for the mechanistic PD model based on the integration of the covariates in the large-scale mechanistic model. In contrast to the empirical PD models, where covariates are mainly included based on statistical significance, this will allow for a physiological interpretation of the covariates in the reduced model. In the context of physiologically based pharmacokinetics models an approach already exists that employs lumping to obtain mechanistically justified covariate models [57].

Dependency on the chosen user-defined error tolerance

We have shown in the context of the brown snake venom-fibrinogen system that the reduced model was unique for a given user-defined error tolerance when considered on a short time span (activation of the cascade), while there were two competing models on the longer time span (activation and recovery of the cascade). This allows to gain additional insight into the role of the different constituents of the signalling pathway (like feedback loops). However, a more in depth analysis of the entirety of the reduced models for large-scale systems pharmacology/biology models, including its dependency on the chosen user-defined error tolerance δ should be subject to future research. In this context the usage of artificial intelligence approaches might be advantageous.

Extension from deterministic to stochastic model dynamics

In this thesis, we introduced the ir-index for large-scale models based on ordinary differential equations. As possible future extension, the ir-indices could be developed for models based on stochastic differential equations.

7. Conclusion

There is a clear need for predictive rather than descriptive modelling approaches – mechanistic models as opposed to empirical models – in drug development [115]. A growing understanding of complex processes in biology has led to large-scale mechanistic models of pharmacologically relevant processes. While these models are designed to study the impact of various inputs or stimuli on the system behaviour, the focus in pharmacology is often on describing the relationship between a specific input (e.g administration of a drug) to specific output (e.g. drug effect, drug response, surrogate marker). These particular questions arising in drug development can be set into a control theoretical framework and allowed us to develop a novel simple and automated model order reduction algorithm based on the introduction of a novel time- and state-dependent quantity called the input-response index. The method has been particularly tailored for systems pharmacology models and thus maintains the interpretability of the parameters and state variables in the reduced model. For the first time, we show how sensitivity analysis can be systematically used for efficient model order reduction. Rather than relying on a single reduction technique, our proposed model reduction approach leverages conservation laws, time-scale separation (QSSA) and time-limited controllability/observability ideas. We can conclude that this work presents a step towards the usage of mechanistic pharmacodynamic models in the context of population analysis of clinical data and thus can provide model continuity from preclinical or early clinical development to the later stages of drug development. All in all, we believe that the proposed concept of input-response indices and the thereon based introduced model order reduction technique significantly broaden our means to analyse and understand complex systems pharmacology/biology models.

References

[1] Aarons, L., Dokoumetzidis, A.: Proper lumping in systems biology models. IET Systems Biology **3**(October 2007), 40–51 (2009)

[2] Agoram, B.: Evaluating systems pharmacology models is different from evaluating standard pharmacokinetic-pharmacodynamic models. CPT: Pharmacometrics and Systems Pharmacology **3**(2) (2014)

[3] Andersen, N.M., Sørensen, M.P., Efendiev, M.A., Olsen, O.H., Ingwersen, S.H.: Modelling of the blood coagulation cascade in an in vitro flow system. International Journal of Biomathematics and Biostatistics **1**(1), 1–7 (2010)

[4] Anderson, J., Chang, Y.C., Papachristodoulou, A.: Model decomposition and reduction tools for large-scale networks in systems biology. Automatica **47**(6), 1165–1174 (2011)

[5] Antoulas, A.: Approximation of Large-Scale Dynamical Systems. SIAM, Philadelphia (2005)

[6] Antsaklis, P.J., Michel, A.N.: Linear Systems, 2nd edn. Birkhäuser Basel (2006)

[7] Benner, P.: Solving Large-Scale Control Problems. IEEE Control Systems **24**(1), 44–59 (2004)

[8] Bonate, P.L.: Pharmacokinetic/pharmacodynamic modeling and simulation, 1st edn. Springer (2008)

[9] Bosley, J., Boren, C., Lee, S., Grøtli, M., Nielsen, J., Uhlen, M., Boren, J., Mardinoglu, A.: Improving the economics of NASH/NAFLD treatment through the use of systems biology. Drug Discovery Today **22**(10), 1532–1538 (2017)

[10] Briggs, G.E., Haldane, J.B.: A Note on the Kinetics of Enzyme Action. Biochemical Journal **19**(2), 338–339 (1925)

[11] Brochot, C., Tóth, J., Bois, F.Y.: Lumping in pharmacokinetics. Journal of pharmacokinetics and pharmacodynamics **32**(5-6), 719–36 (2005)

[12] Burghaus, R., Coboeken, K., Gaub, T., Kuepfer, L., Sensse, A., Siegmund, H.U., Weiss, W., Mueck, W., Lippert, J.: Evaluation of the efficacy and safety of rivaroxaban using a computer model for blood coagulation. PloS one **6**(4), e17,626 (2011)

[13] Campbell, S., Khosravi-Far, R., Rossman, K., Clark, G., Der, C.: Increasing complexity of Ras signaling. Oncogene. **17**, 1395–1413 (1998)

[14] Cawthern, K.M., van 't Veer, C., Lock, J.B., DiLorenzo, M.E., Branda, R.F., Mann, K.G.: Blood coagulation in hemophilia A and hemophilia C. Blood **91**(12), 4581–4592 (1998)

[15] Chen, Y., White, J., Others: A quadratic method for nonlinear model order reduction. In: Inter. Conf. on Modelling and Simulation of Microsystems Semiconductors, Sensors and Actuators, pp. 6–9 (2000)

[16] Ciurus, T., Sobczak, S., Cichocka-radwan, A., Lelonek, M.: New oral anticoagulants - a practical guide. Kardiochirurgia i Torakochirurgia Polska **12**(2), 111–118 (2015)

[17] Clegg, L.E., Gabhann, F.M.: Computational Models for Drug Development. Pharmacol. Res. **99**, 149–154 (2015)

[18] Condon, M., Ivanov, R.: Model reduction of nonlinear systems. COMPEL - The international journal for computation and mathematics in electrical and electronic engineering **23**(2), 547–557 (2003)

[19] Condon, M., Ivanov, R.: Empirical balanced truncation of nonlinear systems. Journal of Nonlinear Science **14**(5), 405–414 (2004)

[20] Connolly, S.J., Ezekowitz, M.D., Yusuf, S., Eikelboom, J., Oldgren, J., Parekh, A., Pogue, J., Reilly, P.A., Themeles, E., Varrone, J., Wang, S., Alings, M., Xavier, D., Zhu, J., Diaz, R., Lewis, B.S., Darius, H., Diener, H.C., Joyner, C.D., Wallentin, L.: Dabigatran versus Warfarin in Patients with Atrial Fibrillation. The New England journal of medicine **361**(12), 1139–1151 (2009)

[21] Conzelmann, H., Saez-Rodriquez, J., Sauter, T., Bullinger, E., Allgöwer, F., Gilles, E.D.: Reduction of mathematical models of signal transduction networks: simulation-based approach applied to egf receptor signalling. Systems biology **1**(1), 159–169 (2004)

[22] Dansirikul, C., Lehr, T., Liesenfeld, K.H., Haertter, S., Staab, A.: A combined pharmacometric analysis of dabigatran etexilate in healthy volunteers and patients with atrial fibrillation or undergoing orthopaedic surgery. Thrombosis and Haemostasis **107**(4), 775–785 (2012)

[23] Dasgupta, A., Wahed, A.: Clinical Chemistry, Immunology and Laboratory Quality Control, 1 edn. Elsevier (2014)

[24] Davies, M.R., Wang, K., Mirams, G.R., Caruso, A., Noble, D., Walz, A., Lave, T., Schuler, F., Singer, T., Polonchuk, L.: Recent developments in using mechanistic cardiac modelling for drug safety evaluation. Drug Discovery Today **21**(6), 924–938 (2016)

[25] Degenring, D., Froemel, C., Dikta, G., Takors, R.: Sensitivity analysis for the reduction of complex metabolism models. Journal of Process Control **14**(7), 729–745 (2004)

[26] Deng, J., Vozmediano, V., Rodriguez, M., Cavallari, L.H., Schmidt, S.: Genotype-guided dosing of warfarin through modeling and simulation. European Journal of Pharmaceutical Sciences **109**(May), S9–S14 (2017)

[27] Derendorf, H., Lesko, L.J.: Pharmacokinetic / Pharmacodynamic Modeling in Drug Research and Development. Journal of Clinical Pharmacology **40**, 1399–1418 (2000)

[28] Derendorf, H., Meibohm, B.: Modeling of Pharmacokinetic/Pharamcodynamic Relationship: Concepts and Perspectives. Pharmaceutical research **16**(2), 176–185 (1999)

[29] Desai, S., Prasad, R.: Balanced Truncation Methods of Model Order Reduction on TIDSP. Proceedings of SARC-IRF International Conference pp. 91–96 (2014)

[30] Dewilde, P., van der Veen, A.j.: Time-varying systems and computations. Computers & Mathematics with Applications **37**(8), 146 (1999)

[31] Dokoumetzidis, A., Aarons, L.: A method for robust model order reduction in pharmacokinetics. Journal of Pharmacokinetics and Pharmacodynamics **36**(6), 613–628 (2009)

[32] Fernando, K.V., Nicholson, H.: On the Structure of Balanced and other Principal Representations of SISO Systems. IEEE Transactions on Automatic Control **28**(2), 228–231 (1983)

[33] Ferrari, M., Pengo, V., Barolo, M., Bezzo, F., Padrini, R.: Assessing the relative potency of (S)- and (R)-warfarin with a new PK-PD model, in relation to VKORC1 genotypes. European Journal of Clinical Pharmacology **73**(6), 699–707 (2017)

[34] Frost, C., Nepal, S., Wang, J., Schuster, A., Byon, W., Boyd, R.A., Yu, Z., Shenker, A., Barrett, Y.C., Mosqueda-Garcia, R., Lacreta, F.: Safety, pharmacokinetics and pharmacodynamics of multiple oral doses of apixaban, a factor Xa inhibitor, in healthy subjects. British Journal of Clinical Pharmacology **76**(5), 776–786 (2013)

[35] Gawronski, W., Juang, J.N.: Model reduction in limited time and frequency intervals. International Journal of Systems Science **21**(2), 349–376 (1990)

[36] Girgis, I.G., Patel, M.R., Peters, G.R., Moore, K.T., Mahaffey, K.W., Nessel, C.C., Halperin, J.L., Califf, R.M., Fox, K.A.A., Becker, R.C.: Population Pharmacokinetics and Pharmacodynamics of Rivaroxaban in Patients with Non-valvular Atrial Fibrillation : Results from ROCKET AF. The Journal of Clinical Pharmacology **54**(8), 917–927 (2014)

[37] Gombarska, D., Benova, M.: Computer model of anticoagulation treatment. Przeglad Elektrotechniczny **89**(2), 262–263 (2013)

[38] Granger, C.B., Alexander, J.H., McMurray, J.J.V., Lopes, R.D., Hylek, E.M., Hanna, M., Al-Khalidi, H.R., Easton, J.D., Ezekowitz, J.A., Flaker, G., Garcia, D., Geraldes, M., Gersh, B.J., Golitsyn, S., Goto, S., Hermosillo, A.G., Hohnloser, S.H., Horowitz, J., Mohan, P., Jansky, P., Lewis, B.S., Lopez-Sendon, J.L., Pais, P., Parkhomenko, A., Veheugt, F.W.A., Zhu, J., Wallentin, L.: Apixaban versus Warfain in Patients with Atrial Fibrillation. The New England journal of medicine **365**(11), 981–992 (2011)

[39] Gulati, A., Isbister, G.K., Duffull, S.B.: Effect of Australian elapid venoms on blood coagulation: Australian Snakebite Project (ASP-17). Toxicon **61**(1), 94–104 (2013)

[40] Gulati, A., Isbister, G.K., Duffull, S.B.: Scale reduction of a systems coagulation model with an application to modeling pharmacokinetic-pharmacodynamic data. CPT: Pharmacometrics Systems Pharmacology **3**, e90 (2014)

[41] Hahn, J., Edgar, T.F.: A Gramian Based Approach to Nonlinearity Quantification and Model Classification. Industrial & Engineering Chemistry Research **40**(24), 5724–5731 (2001)

[42] Hahn, J., Edgar, T.F.: An improved method for nonlinear model reduction using balancing of empirical gramians. Computers & Chemical Engineering **26**(10), 1379–1397 (2002)

[43] Hahn, J., Edgar, T.F., Marquardt, W.: Controllability and observabiliy covariance matrices for the analysis and order reduction of stable nonlinear systems. Internal Report **13**, 1–33 (2003)

[44] Hamberg, A.K.: Pharmacometric Models for Individualisation of Warfarin in Adults and Children. Ph.D. thesis, Uppsala University (2013)

[45] Hamberg, A.K., Dahl, M.L., Barban, M., Scordo, M.G., Wadelius, M., Pengo, V., Padrini, R., Jonsson, E.N.: A PK-PD model for predicting the impact of age, CYP2C9, and VKORC1 genotype on individualization of warfarin therapy. Clinical Pharmacology and Therapeutics **81**(4), 529–538 (2007)

[46] Han, C.W., Jeong, M.S., Jang, S.B.: Structure, signaling and the drug discovery of the Ras oncogene protein. BMB Reports **50**(7), 355–360 (2017)

[47] Hartmann, S., Biliouris, K., Lesko, L.J., Nowak-Göttl, U., Trame, M.N.: Quantitative Systems Pharmacology Model to Predict the Effects of Commonly Used Anticoagulants on the Human Coagulation Network. CPT: Pharmacometrics and Systems Pharmacology **5**(10), 554–564 (2016)

[48] Himpe, C.: Empirical Gramian Framework. URL `http://gramian.de`

[49] Himpe, C., Ohlberger, M.: A Unified Software Framework for Empirical Gramians. Journal of Mathematics **2013**(2), 1–6 (2013)

[50] Himpe, C., Ohlberger, M.: Cross-gramian-based combined state and parameter reduction for large-scale control systems. Mathematical Problems in Engineering **2014** (2014)

[51] Himpe, C., Ohlberger, M.: Cross-Gramian-Based Model Reduction: A Comparison. In: P. Benner, M. Ohlberger, A. Patera, G. Rozza, K. Urban (eds.) Model reduction of Parametrized Systems, chap. 17, pp. 271–283. Springer, Cham (2016)

[52] Hofmeyr, J.H.S.: Metabolic control analysis in a nutshell. Proceedings of the 2nd International Conference on Systems Biology pp. 291–300 (2001)

[53] Holford, N.H.G.: Clinical pharmacokinetics and Pharmacodynamics of warfarin. Clin Pharmacokinet **11**(6), 483–504 (1986)

[54] Holford, N.H.G., Sheiner, L.B.: Understanding the Dose-Effect Relationship: Clinical Application of Pharmacokinetic-Pharmacodynamic Models. Clinical Pharmacokinetics **6**, 429–453 (1981)

[55] Hornberg, J.J., Binder, B., Bruggeman, F.J., Schoeberl, B., Heinrich, R., Westerhoff, H.V.: Control of MAPK signalling: from complexity to what really matters. Oncogene **24**(36), 5533–5542 (2005)

[56] Hornberg, J.J., Tijssen, M.R., Lankelma, J.: Synergistic activation of signalling to extracellular signal-regulated kinases 1 and 2 by epidermal growth factor and 4 beta-phorbol 12-myristate 13-acetate. European journal of biochemistry / FEBS **271**(19), 3905–13 (2004)

[57] Huisinga, W., Solms, A., Fronton, L., Pilari, S.: Modeling interindividual variability in physiologically based pharmacokinetics and its link to mechanistic covariate modeling. CPT: Pharmacometrics and Systems Pharmacology **1**(1), 1–10 (2012)

[58] Iwamoto, K., Shindo, Y., Takahashi, K.: Modeling Cellular Noise Underlying Heterogeneous Cell Responses in the Epidermal Growth Factor Signaling Pathway. PLoS Computational Biology **12**(11), 1–18 (2016)

[59] Jesty, J.: Analysis of the generation and inhibition of factor Xa: Area under generation curves is independent of enzyme generation rate. Journal of Biological Chemistry **265**(29), 17,539–17,544 (1990)

[60] Jezovnik, M.K., Poredos, P.: Clinical relevance of the pharmacogenetics of warfarin. ESC Council for Cardiology Practice **9**(21), 1–4 (2011)

[61] Johnson, J.A., Caudle, K.E., Gong, L., Whirl-Carrillo, M., Stein, C.M., Scott, S.A., Lee, M.T., Gage, B.F., Kimmel, S.E., Perera, M.A., Anderson, J.L., Pirmohamed, M., Klein, T.E., Limdi, N.A., Cavallari, L.H., Wadelius, M.: Clinical Pharmacogenetics Implementation Consortium (CPIC) Guideline for Pharmacogenetics-Guided Warfarin Dosing: 2017 Update. Clinical Pharmacology and Therapeutics **102**(3), 397–404 (2017)

[62] Karoulia, Z., Gavathiotis, E., Poulikakos, P.I.: New perspectives for targeting RAF kinase in human cancer. Nature reviews. Cancer **17**(11), 676–691 (2017)

[63] Kholodenko, B.N., Demin, O.V., Moehren, G., Hoek, J.B.: Quantification of short term signaling by the epidermal growth factor receptor. Journal of Biological Chemistry **274**(42), 30,169–30,181 (1999)

[64] Kimko, H., Pinheiro, J.: Model-based clinical drug development in the past, present and future: A commentary. British Journal of Clinical Pharmacology **79**(1), 108–116 (2014)

[65] Kimmel, S.: Warfarin therapy: in need of improvement after all these years. Expert Opin Pharmacother. **9**(5), 677–686 (2008)

[66] Klein, T.E., Altman, R., Eriksson, N., Gafe, B., Kimmel, S., Lee, M.T.M., Limdi, N.A., Page, D., Roden, D., Wagner, M., Caldwell, M., Johnson, J., Wadelius, M.: Estimation of the Warfarin Dose with Clinical and Pharmacogenetic Data. N engl J Med. **360**(8), 753–764 (2009)

[67] von Kleist, M., Menz, S., Stocker, H., Arasteh, K., Schütte, C., Huisinga, W.: HIV quasispecies dynamics during pro-active treatment switching: Impact on multi-drug resistance and resistance archiving in latent reservoirs. PLoS ONE **6**(3) (2011)

[68] Kogan, A.E., Kardakov, D.V., Khanin, M.A.: Analysis of the activated partial throm-

boplastin time test using mathematical modeling. Thrombosis Research **101**(4), 299–310 (2001)

[69] Lacroix, D.E.: A reduced equation mathematical model for blood coagulation and lysis in quiescent plasma. International journal of structural changes in solids - mechanics and applications **4**, 23–35 (2012)

[70] Lall, S., Marsden, J.E., Glavaški, S.: Empirical Model Reduction of Controlled Nonlinear Systems. World pp. 473–478 (1999)

[71] Lalonde, R.L., Kowalski, K.G., Hutmacher, M.M., Ewy, W., Nichols, D.J., Milligan, P.A., Corrigan, B.W., Lockwood, P.A., Marshall, S.A., Benincosa, L.J., Tensfeldt, T.G., Parivar, K., Amantea, M., Glue, P., Koide, H., Miller, R.: Model-based drug development. Clinical Pharmacology and Therapeutics **82**(1), 21–32 (2007)

[72] Lam, S.H.: Singular Perturbation for Stiff Equations Using Numerical Methods. In: C. Casci (ed.) Recent Advances in the Aerospace Sciences, chap. 1, pp. 2–20. Plenum Press, New York and London (1985)

[73] Lam, S.H.: Reduced Chemistry Modeling and Sensitivity Analysis. 1994-1995 Lecture Series Programme, the Von Karman Institute for Fluid Dynamics pp. 1–28 (1995)

[74] Levine, W.S.: Control System Fundamentals, 2nd edn. CRC Press, Inc., Boca Raton, FL, USA (2011)

[75] Li, G., Rabitz, H.: A General Analysis of Exact Lumping in Chemical Kinetics. Chemical Engineering Science **44**(6), 1413–1430 (1989)

[76] Li, G., Rabitz, H.: A general analysis of approximate lumping in chemical kinetics. Chemical Engineering Science **45**(4), 977–1002 (1990)

[77] Li, G., Tomlin, A.S., Rabitz, H., Tóth, J.: A general analysis of approximate nonlinear lumping in chemical kinetics. I. Unconstrained lumping. The Journal of Chemical Physics **101**(2), 1172 (1994)

[78] Liesenfeld, K.H., Schäfer, H.G., Trocóniz, I.F., Tillmann, C., Eriksson, B.I., Stangier, J.: Effects of the direct thrombin inhibitor dabigatran on ex vivo coagulation time in orthopaedic surgery patients: a population model analysis. British Journal of Clinical Pharmacology **62**(5), 527–537 (2006)

[79] Liu, G., Swihart, M.T., Neelamegham, S.: Sensitivity, principal component and flux analysis applied to signal transduction: The case of epidermal growth factor mediated signaling. Bioinformatics **21**(7), 1194–1202 (2005)

[80] Löffler, H.P., Marquardt, W.: Order reduction of non-linear differential-algebraic process models. Journal of Process Control **1**(1), 32–40 (1991)

[81] MacDonald, N.: Time delay in prey-predator models. Mathematical Biosciences **28**(3-4), 321–330 (1976)

[82] Maddison, J., Somogyi, A.A., Jensen, B.P., James, H.M., Gentgall, M., Rolan, P.E.: The pharmacokinetics and pharmacodynamics of single dose (R)- and (S)-warfarin

administered separately and together: Relationship to VKORC1 genotype. British Journal of Clinical Pharmacology **75**(1), 208–216 (2012)

[83] Marshall, S., Macintyre, F., James, I., Krams, M., Jonsson, N.E.: Role of mechanistically-based pharmacokinetic/pharmacodynamic models in drug development : a case study of a therapeutic protein. Clinical Pharmacokinetics **45**(2), 177–197 (2006)

[84] Michaelis, L., Menten, M.L.: Die Kinetik der Invertinwerkung (The kinetics of invertase activity). Biochemische Zeitschrift **49**(February), 333–369 (1913)

[85] Moore, B.: Principal component analysis in linear systems: Controllability, observability, and model reduction. IEEE Transactions on Automatic Control **26**(1), 17–32 (1981)

[86] Morgan, P., Van Der Graaf, P.H., Arrowsmith, J., Feltner, D.E., Drummond, K.S., Wegner, C.D., Street, S.D.: Can the flow of medicines be improved? Fundamental pharmacokinetic and pharmacological principles toward improving Phase II survival. Drug Discovery Today **17**(9-10), 419–424 (2012)

[87] Mueck, W., Stampfuss, J., Kubitza, D., Becka, M.: Clinical pharmacokinetic and pharmacodynamic profile of rivaroxaban. Clinical Pharmacokinetics **53**(1), 1–16 (2014)

[88] Mueller-Esterl, W.: Biochemie - Eine Einführung für Mediziner und Naturwissenschaftler. Spektrum akademischer Verlag (2011)

[89] Nayak, S., Lee, D., Patel-Hett, S., Pittman, D.D., Martin, S.W., Heatherington, A.C., Vicini, P., Hua, F.: Using a Systems Pharmacology Model of the Blood Coagulation Network to Predict the Effects of Various Therapies on Biomarkers. CPT: Pharmacometrics and Systems Pharmacology **4**(7), 396–405 (2015)

[90] Ohara, M., Takahashi, H., Lee, M.T.M., Wen, M.S., Lee, T.H., Chuang, H.P., Luo, C.H., Arima, A., Onozuka, A., Nagai, R., Shiomi, M., Mihara, K., Morita, T., Chen, Y.T.: Determinants of the over-anticoagulation response during warfarin initiation therapy in Asian patients based on population pharmacokinetic- pharmacodynamic analyses. PLoS ONE **9**(8), 1–11 (2014)

[91] Oikonomou, E., Koustas, E., Goulielmaki, M., Pintzas, A.: BRAF vs RAS oncogenes: Are mutations of the same pathway equal? Differential signalling and therapeutic implications. Oncotarget **5**(23), 11,752–11,777 (2014)

[92] Okino, M.S., Mavrovouniotis, M.L.: Simplification of Mathematical Models of Chemical Reaction Systems. Chemical Reviews **98**(2) (1998)

[93] Palta, S., Saroa, R., Palta, A.: Overview of the coagulation system. Indian Journal of Anaesthesia **58**(5), 515–523 (2014)

[94] Patel, M.R., Mahaffey, K.W., Garg, J., Pan, G., Singer, D.E., Hacke, W., Breithardt, G., Halperin, J.L., Hankey, G.J., Piccini, J.P., Becker, R.C., Nessel, C.C., Paolini, J.F., Berkowitz, S.D., Fox, K.A.A., Califf, R.M.: Rivaroxaban versus Warfarin in Nonvalvular Atrial Fibrillation. The New England journal of medicine **364**(10), 883– 891 (2011)

[95] Perumal, T.M., Gunawan, R.: Understanding dynamics using sensitivity analysis: caveat and solution. BMC systems biology **5**(41), 1–10 (2011)

[96] Peterson, M., Riggs, M.: FDA Advisory Meeting Clinical Pharmacology Review Utilizes a Quantitative Systems Pharmacology (QSP) Model: A Watershed Moment? CPT: Pharmacometrics & Systems Pharmacology **4**(3), 189–192 (2015)

[97] Peterson, M.C., Riggs, M.M.: A physiologically based mathematical model of integrated calcium homeostasis and bone remodeling. Bone **46**(1), 49–63 (2010)

[98] Petrov, V., Nikolova, E., Wolkenhauer, O.: Reduction of nonlinear dynamic systems with an application to signal transduction pathways. IET Systems Biology **1**(1), 2–9 (2007)

[99] Pilari, S., Huisinga, W.: Lumping of physiologically-based pharmacokinetic models and a mechanistic derivation of classical compartmental models. Journal of Pharmacokinetics and Pharmacodynamics **37**(4), 365–405 (2010)

[100] Pirmohamed, M.: Warfarin: Almost 60 years old and still causing problems. British Journal of Clinical Pharmacology **62**(5), 509–511 (2006)

[101] Pitsiu, M., Parker, E., Aarons, L., Rowland, M.: Population pharmacokinetics and pharmacodynamics of warfarin in healthy young adults. European Journal of Pharmaceutical Sciences **1**(3), 151–157 (1993)

[102] Poller, L., Hirsh, J.: A Simple System for the Derivation of International Normalized Ratios for the Reporting of Prothrombin Time Results with North American Thromboplastin Reagents. Am J Clin Pathol **92**, 124–126 (1989)

[103] Rabitz, H., Kramer, M., Dacol, D.: Sensitivity analysis in chemical kinetics. Annual Review Physics Chemistry **34**, 419–461 (1983)

[104] Rewieński, M.J.: A trajectory piecewise-linear approach to model order reduction of nonlinear dynamical systems. dissertation, Massachusetts institute of technology (2003)

[105] Ribbing, J., Niclas Jonsson, E.: Power, selection bias and predictive performance of the population pharmacokinetic covariate model. Journal of Pharmacokinetics and Pharmacodynamics **31**(2), 109–134 (2004)

[106] Riggs, M., Bennetts, M., van der Graaf, P., Martin, S.: Integrated Pharmacometrics and Systems Pharmacology Model-Based Analyses to Guide GnRH Receptor Modulator Development for Management of Endometriosis. CPT: Pharmacometrics & Systems Pharmacology **1**(October), e11 (2012)

[107] Saffian, S.M., Duffull, S.B., Wright, D.F.B.: Warfarin Dosing Algorithms Underpredict Dose Requirements in Patients Requiring 7 mg Daily : A Systematic Review and Meta-analysis. Clinical Pharmacology & Therapeutics **102**(2), 297–304 (2017)

[108] Sasaki, T., Tabuchi, H., Higuchi, S., Ieiri, I.: Warfarin-dosing algorithm based on a population pharmacokinetic / pharmacodynamic model combined with Bayesian forecasting. Pharmacogenomics **10**, 1257–1266 (2009)

[109] Savo, D., Sarić, A.T.: Dynamic Model Reduction : An Overview of Available Techniques with Application to Power Systems. Serbian Journal of Electrical Engineering **9**(2), 131–169 (2012)

[110] Schilders, W.H., van der Vorst, H.A., Rommes, J.: Model Order Reduction: Theory, Research Aspects and Applications. Springer Verlag (2008)

[111] Schoeberl, B.: Mathematical modeling of signal transduction pathways in mammalian cells at the example of the EGF induced MAP kinase cascade and TNF receptor crosstalk. Ph.D. thesis, Universität Stuttgart (2003)

[112] Schoeberl, B., Eichler-Jonsson, C., Gilles, E.D., Müller, G.: Computational modeling of the dynamics of the MAP kinase cascade activated by surface and internalized EGF receptors. Nature Biotechnology **20**(4), 370–375 (2002)

[113] Segel, L.A., Slemrod, M.: The quasi-steady state assumption: A case study in perturbation. Society of Industrial and Applied Mathematics **31**(3), 446–477 (1989)

[114] Sharma, A., Jusko, W.J.: Characteristics of indirect pharmacodynamic models and applications to clinical drug responses. Br J Clin Pharmacol **45**(3), 229–239 (1998)

[115] Sheiner, L.B.: Pharmacokinetic/Pharmacodynamic Modeling in Drug Development. Annual Review of Pharmacology and Toxicology **40**(1), 67–95 (2000)

[116] Snowden, T.: Methods of model reduction for quantitative systems pharmacology. dissertation, University of Reading (2015)

[117] Snowden, T.J., van der Graaf, P.H., Tindall, M.J.: A combined model reduction algorithm for controlled biochemical systems. BMC Systems Biology **11**(17), 1–18 (2017)

[118] Snowden, T.J., van der Graaf, P.H., Tindall, M.J.: Methods of Model Reduction for Large-Scale Biological Systems: A Survey of Current Methods and Trends. Bulletin of Mathematical Biology **79**(7), 1449–1486 (2017)

[119] Stephanopoulos, G.N., Aristidou, A.A., Nielsen, J.: Metabolic Engineering: Principles and Methodologies, 1 edn. Academic Press (1998)

[120] Stötzel, C., Plöntzke, J., Heuwieser, W., Röblitz, S.: Advances in modeling of the bovine estrous cycle: Synchronization with PGF2α. Theriogenology **78**(7), 1415–1428 (2012)

[121] Tanos, P.P., Isbister, G.K., Lalloo, D.G., Kirkpatrick, C.M.J., Duffull, S.B.: A model for venom-induced consumptive coagulopathy in snake bite. Toxicon **52**(7), 769–780 (2008)

[122] Tormoen, G.W., Khader, A., Gruber, A., McCarty, O.J.T.: Physiological levels of blood coagulation factors IX and X control coagulation kinetics in an in vitro model of circulating tissue factor. Phys Biol. **10**(3), 83–96 (2013)

[123] Turányi, T.: Sensitivity analysis of complex kinetic systems. Tools and applications. Journal of Mathematical Chemistry **5**, 203–248 (1990)

[124] Ünal, E.B., Uhlitz, F., Blüthgen, N.: A compendium of ERK targets. FEBS Letters **591**(17), 2607–2615 (2017)

[125] Upton, R.N., Mould, D.R.: Basic concepts in population modeling, simulation, and model-based drug development: part 3-introduction to pharmacodynamic modeling methods. CPT: pharmacometrics & systems pharmacology **3**(October 2013), e88 (2014)

[126] Vajda, S., Valko, P., Turanyi, T.: Principal component analysis of kinetic models. International Journal of Chemical Kinetics **17**, 55–81 (1985)

[127] Vallabhajosyula, R.R., Sauro, H.M.: Complexity Reduction of Biochemical Networks. Proceedings of the 2006 Winter Simulation Conference pp. 1690–1697 (2006)

[128] Verriest, E.I.: Finite Time System Operator and Balancing for Model Reduction and Decoupling. Proceedings of the 19th International Symposium on Mathematical Theory of Networks and Systems **2**, 1985–1988 (2010)

[129] Wadelius, M., Pirmohamed, M.: Pharmacogenetics of warfarin: current status and future challenges. The Pharmacogenomics Journal **7**(2), 99–111 (2007)

[130] Wajima, T., Isbister, G., Duffull, S.: A Comprehensive Model for the Humoral Coagulation Network in Humans. Journal of Clinical Pharmacology and Therapeutics **86**, 290–298 (2009)

[131] Wan, J.M.: Explicit solution and stability of linear time-varying differential state space systems. International Journal of Control, Automation and Systems **15**(4), 1553–1560 (2017)

[132] Wang, D.Y., Cardelli, L., Phillips, A., Piterman, N., Fisher, J.: Computational modeling of the EGFR network elucidates control mechanisms regulating signal dynamics. BMC Systems Biology **3**(118), 1–17 (2009)

[133] Wei, J., Kuo, J.C.W.: Lumping Analysis in Monomolecular Reaction Systems. Analysis of the Exactly Lumpable System. Industrial & Engineering Chemistry Fundamentals **8**(1), 114–123 (1969)

[134] White, R.H., Hong, R., Venook, A.P., Daschbach, M.M., Murray, W., Mungall, D.R., Coleman, R.W.: Initiation of Warfarin Therapy: Comparison of Physician Dosing With Computer-Assisted Dosing. Journal of General Internal Medicine **2**(May/Jun), 141–148 (1987)

[135] Wiley, H.S., Shvartsman, S.Y., Lauffenburger, D.A.: Computational modeling of the EGF-receptor system: A paradigm for systems biology. Trends in Cell Biology **13**(1), 43–50 (2003)

[136] Xu, X.S., Moore, K., Burton, P., Stuyckens, K., Mueck, W., Rossenu, S., Plotnikov, A., Gibson, M., Vermeulen, A.: Population pharmacokinetics and pharmacodynamics of rivaroxaban in patients with acute coronary syndromes. British Journal of Clinical Pharmacology **74**(1), 86–97 (2012)

[137] Yarden, Y., Sliwkowski: Untangling the ErbB signaling network. Nature Reviews Molecular Cell Biology **2**(February), 127–137 (2001)

[138] Zagaris, A., Kaper, H.G., Kaper, T.J.: Fast and slow dynamics for the computational singular perturbation method. Multiscale Modeling & Simulation **2**(4), 613–638 (2004)

[139] Zi, Z.: Sensitivity analysis approaches applied to systems biology models. IET Systems Biology **5**(6), 336–346 (2011)

I. Appendix

I.1. Elimination reduced model for the fibrinogen-brown snake venom setting

Differential equations for the elimination-reduced 8-state variable model of the snake venom system

The model equations for the elimination-reduced model are:

$$
\begin{aligned}
\frac{dx_{\mathrm{red,IIa}}}{dt} &= V_{\mathrm{Xa:Va,IIa}} \frac{x_{\mathrm{red,CVenom}}}{x_{\mathrm{red,CVenom}} + K_{\mathrm{Xa:Va,IIa}}} \cdot x_{\mathrm{env,II}} \\
&\quad - (c_{\mathrm{IIa,Tmod}} x_{\mathrm{env,Tmod}} + d_{\mathrm{IIa}}) \cdot x_{\mathrm{red,IIa}} \\
\frac{dx_{\mathrm{red,Fg}}}{dt} &= p_{\mathrm{Fg}} - \left(V_{\mathrm{IIa,Fg}} \frac{x_{\mathrm{red,IIa}}}{x_{\mathrm{red,IIa}} + K_{\mathrm{IIa,Fg}}} \right. \\
&\quad \left. + V_{\mathrm{P,Fg}} \frac{x_{\mathrm{red,P}}}{x_{\mathrm{red,P}} + K_{\mathrm{P,Fg}}} + d_{\mathrm{Fg}} \right) \cdot x_{\mathrm{red,Fg}} \\
\frac{dx_{\mathrm{red,P}}}{dt} &= \left(V_{\mathrm{IIa,P}} \frac{x_{\mathrm{red,IIa}}}{x_{\mathrm{red,IIa}} + K_{\mathrm{IIa,P}}} \right. \\
&\quad \left. + V_{\mathrm{APC:PS,P}} \frac{x_{\mathrm{red,APC:PS}}}{x_{\mathrm{red,APC:PS}} + K_{\mathrm{APC:PS,P}}} \right) \\
&\quad \cdot x_{\mathrm{env,Pg}} - d_{\mathrm{P}} \cdot x_{\mathrm{red,P}} \\
\frac{dx_{\mathrm{red,APC}}}{dt} &= x_{\mathrm{env,PC}} \cdot V_{\mathrm{IIa:Tmod,APC}} \\
&\quad \cdot \frac{x_{\mathrm{red,IIa:Tmod}}}{x_{\mathrm{red,IIa:Tmod}} + K_{\mathrm{IIa:Tmod,APC}}} \\
&\quad - (c_{\mathrm{APC,PS}} \cdot x_{\mathrm{env,PS}} + d_{\mathrm{APC}}) \cdot x_{\mathrm{red,APC}} \\
\frac{dx_{\mathrm{red,APC:PS}}}{dt} &= c_{\mathrm{APC,PS}} \cdot x_{\mathrm{env,PS}} \cdot x_{\mathrm{red,APC}} \\
&\quad - d_{\mathrm{APC:PS}} \cdot x_{\mathrm{red,APC:PS}} \\
\frac{dx_{\mathrm{red,IIa:Tmod}}}{dt} &= c_{\mathrm{IIa,Tmod}} \cdot x_{\mathrm{red,IIa}} \cdot x_{\mathrm{env,Tmod}} \\
&\quad - d_{\mathrm{IIa:Tmod}} \cdot x_{\mathrm{red,IIa:Tmod}} \\
\frac{dx_{\mathrm{red,AVenom}}}{dt} &= -k_{\mathrm{abs}} \cdot x_{\mathrm{red,AVenom}} \\
\frac{dx_{\mathrm{red,CVenom}}}{dt} &= k_{\mathrm{abs}} \cdot x_{\mathrm{red,AVenom}} \\
&\quad - d_{\mathrm{CVenom}} \cdot x_{\mathrm{red,CVenom}}
\end{aligned}
$$

All parameter values can be found in [130, Suppl. Fig. 2] and [39, Tab. 1&Tab. 2]. The variable $x_{\mathrm{env,II}}$ is equal to the initial value of II ($x_{\mathrm{II}}(0)$). This holds for all environmental state variables. All initial conditions can be found in [130, Suppl. Fig. 3].

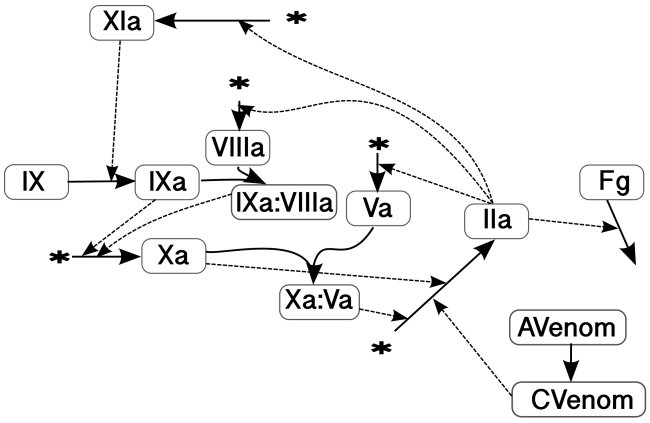


Figure 41: **Alternative elimination-reduced model of the brown snake venom-fibrinogen system.** Model reduction based on randomly chosen ranking of state variables (based on run no 2). Shown are the 12 dynamical state variables. The environmental state variables (indicated by '*') are II, V, VIII, X and XI.

Table 12: **Parameter values of the elimination-reduced and lumped 5-state variable model of the brown snake venom-fibrinogen system.** Explanation of notation (by examples): $V_{\mathrm{IIa,P}}$ represents maximal reaction velocity and $K_{\mathrm{IIa,P}}$ the concentration of IIa were half the maximal velocity is obtained (both are Michaelis-Menten constants). In both cases, P represents the factor which is being activated by this reaction. A degradation rate constants is denoted by d_{P} and production rate constants by p_{IIa}. A constant $c_{\mathrm{IIa,Tmod}}$ denotes a reaction rate constant for complex formation.

Parameter	Value	Unit	Reference
$V_{\mathrm{Xa:Va,IIa}}$	100	1/h	[130, Suppl. Fig. 2]
$K_{\mathrm{Xa:Va,IIa}}$	10	nM	[130, Suppl. Fig. 2]
$V_{\mathrm{IIa:Tmod,APC}}$	7	1/h	[130, Suppl. Fig. 2]
$K_{\mathrm{IIa:Tmod,APC}}$	1	nM	[130, Suppl. Fig. 2]
$V_{\mathrm{IIa,P}}$	7	1/h	[130, Suppl. Fig. 2]
$K_{\mathrm{IIa,P}}$	5000	nM	[130, Suppl. Fig. 2]
$V_{\mathrm{APC:PS,P}}$	2	1/h	[130, Suppl. Fig. 2]
$K_{\mathrm{APC:PS,P}}$	1	nM	[130, Suppl. Fig. 2]
$V_{\mathrm{IIa,Fg}}$	21000	1/h	[39, Tab. 2]
$K_{\mathrm{IIa,Fg}}$	30000	nM	[39, Tab. 2]
$V_{\mathrm{P,Fg}}$	500	1/h	[39, Tab. 2]
$K_{\mathrm{P,Fg}}$	500	nM	[39, Tab. 2]
d_{Fg}	0.032	1/h	[130, Suppl. Fig. 2]
d_{P}	20	1/h	[130, Suppl. Fig. 2]
$d_{\mathrm{APC:PS}}$	20	1/h	[130, Suppl. Fig. 2]
d_{APC}	20.4	1/h	[130, Suppl. Fig. 2]
$d_{\mathrm{IIa:Tmod}}$	20	1/h	[130, Suppl. Fig. 2]
d_{IIa}	67.4	1/h	[130, Suppl. Fig. 2]
d_{CVenom}	3.5	1/h	[39, Tab. 1]
p_{Fg}	286.256	nM/h	[130, Suppl. Fig. 2]
$c_{\mathrm{APC,PS}}$	2	$1/(\mathrm{nM}\cdot\mathrm{h})$	[130, Suppl. Fig. 2]
$c_{\mathrm{IIa,Tmod}}$	2	$1/(\mathrm{nM}\cdot\mathrm{h})$	[130, Suppl. Fig. 2]

Table 13: **Ordering of the state variables based on the input-response indices for the brown snake venom-fibrinogen system up to 1h**. State variables not included in the table had an input-response index of zero.

	State variable	$\|\|\mathrm{ir}_i\|\|_\infty$		**State variable**	$\|\|\mathrm{ir}_i\|\|_\infty$
1	AVenom	4.87e4	19	F	2.09e2
2	Fg	2.48e4	20	II	1.21e2
3	CVenom	1.80e4	21	PC	9.87e1
4	P	1.57e4	22	XI	1.39e1
5	APC:PS	1.08e4	23	IX	6.82
6	IIa:Tmod	1.07e4	24	X	6.44
7	IIa	6.99e3	25	PS	3.94
8	APC	2.03e3	26	XIIIa	2.22
9	XIa	1.65e3	27	TFPI	1.05
10	Xa:Va	1.56e3	28	XF	0.19
11	IXa:VIIIa	1.51e3	29	XIII	0.19
12	Xa	1.31e3	30	D	0.18
13	IXa	1.17e3	31	AUC	0.18
14	Va	1.17e3	32	FDP	0.18
15	V	6.73e2	33	TAT	0.18
16	VIIIa	3.30e2	34	VII	0.18
17	Pg	2.58e2	35	VIIa	0.18
18	Tmod	2.36e2	36	Xa:TFPI	0.18

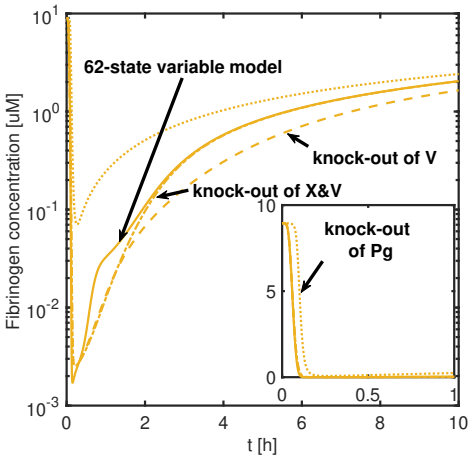


Figure 42: **Comparison of fibrinogen concentration-time profile** based on 62-state variable model (solid) and the model with knock-out of X and V (dot-dashed), knock-out of V (dashed) and knock-out of Pg, the inactive form of factor P (dotted). Note that the larger impact of a knock-out of V is due the fact that in this case, the impact of the activated form of factor X changes (due to the lacking complex formation of Xa and Va).

Table 14: **Relative error for the 8-state variable and the 12-state variable model** on the interval $[0, 40]$h. Given is the squared relative error for the output fibrinogen based on the 62-state variable and the reduced models (Figure 16 and Figure 41) for different time intervals of interest (pre nadir, post nadir and for the whole interval).

Reduced model	**Squared relative error pre nadir**	**Squared relative error post nadir**	**Squared relative error for** $[0, 40\mathbf{h}]$
8-state variable model depicted in Figure 16	3.78e-6	0.0022	0.0022
12-state variable model depicted in Figure 41	0.0002	0.002	0.0022

Table 15: **Order of the elimination-reduced model obtained by varying the user-defined error tolerance for the brown snake venom-fibrinogen system for the first hour after envenomation.** Given are the dynamical state variables and the environmental state variables for each of the obtained elimination-reduced model for the different user-defined error tolerance.

Error tolerance	**Order**	**Dynamical state variables**	**Environmental state variables**
0.001	23	APC, APC:PS, AVenom, CVenom, F, Fg, II, IIa, IIa:Tmod, IXa, IXa:VIIIa, P, PC, Pg, Tmod, V, VIII, VIIIa, Va, XI, XIa, Xa, Xa:Va	VKH2, TFPI, PS, X, IX
0.01	16	APC, APC:PS, AVenom, CVenom, Fg, IIa, IIa:Tmod, IXa, IXa:VIIIa, P, VIII, VIIIa, Va, XIa, Xa, Xa:Va	TFPI, PS, X, IX, XI, PC, II, Tmod, Pg, V
0.1	8	APC, APC:PS, AVenom, CVenom, Fg, IIa, IIa:Tmod, P	II, Pg, PC, Tmod, PS

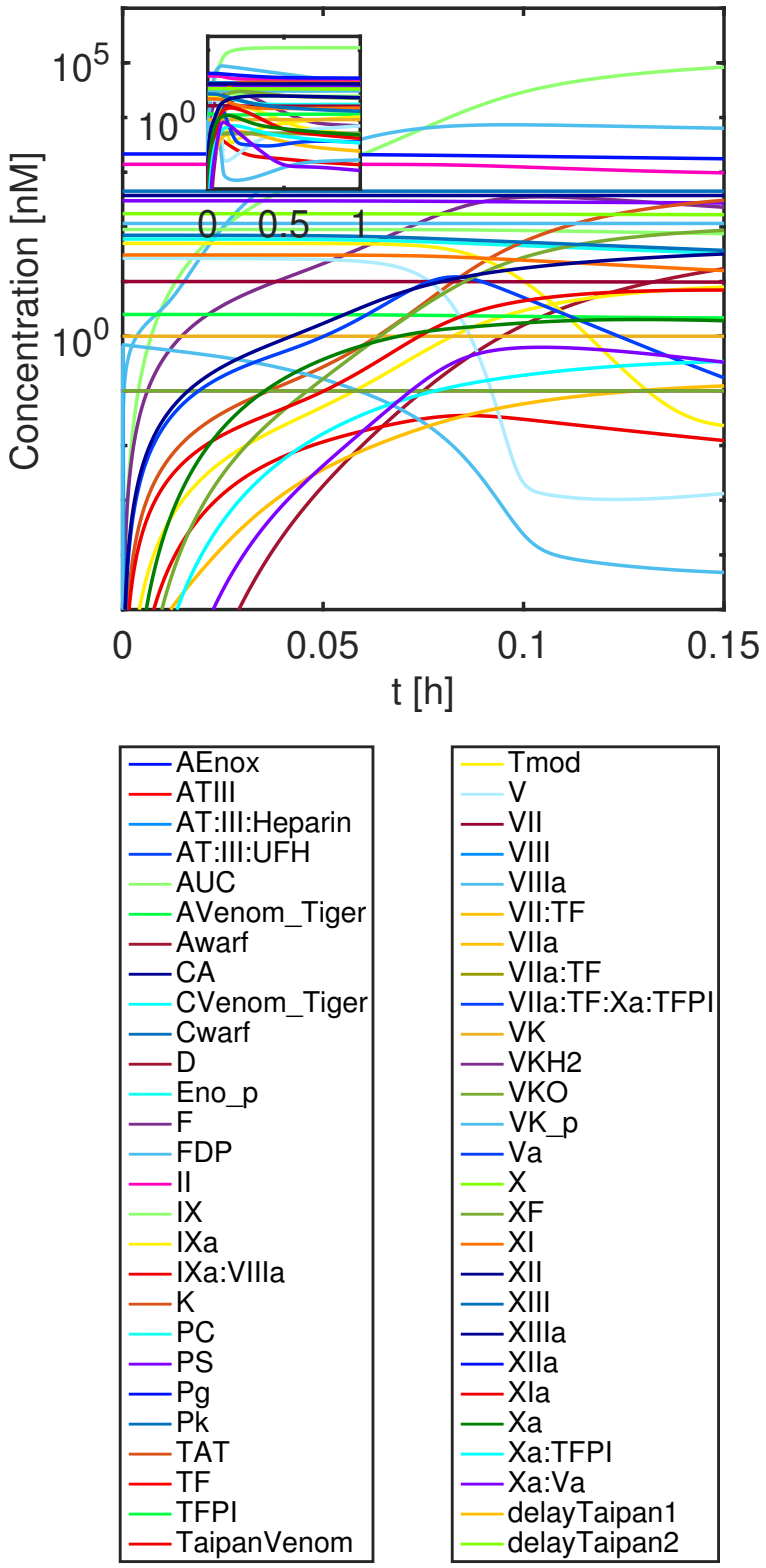


Figure 43: **Concentration-time profiles of all state variables of the blood coagulation** predicted by the original 62-state variable model excluding those concentration-time profiles already given in Figure 15B.

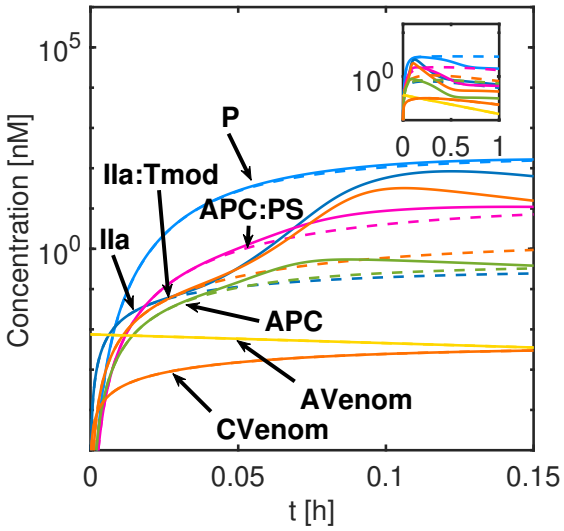


Figure 44: **Comparison of concentration-time profiles for dynamical state variables of the elimination-reduced model (given in Figure 16)** based on 62-state variable model (solid) and the elimination-reduced 8-state variable model (dashed). For AVenom and CVenom the dashed and solid line coincide.

I.2. Elimination reduced model for the PT test

The model equations for the 7-state variable elimination-reduced model for the high TF are:

$$\frac{dx_{\mathrm{red,IIa}}}{dt} = V_{\mathrm{Xa,IIa}} \frac{x_{\mathrm{red,Xa}}}{x_{\mathrm{red,Xa}} + K_{\mathrm{Xa,IIa}}} \cdot x_{\mathrm{env,II}} - d_{\mathrm{IIa}} \cdot x_{\mathrm{red,IIa}}$$

$$\frac{dx_{\mathrm{red,Fg}}}{dt} = -V_{\mathrm{IIa,Fg}} \frac{x_{\mathrm{red,IIa}}}{x_{\mathrm{red,IIa}} + K_{\mathrm{IIa,Fg}}} \cdot x_{\mathrm{red,Fg}} - d_{\mathrm{Fg}} \cdot x_{\mathrm{red,Fg}}$$

$$\frac{dx_{\mathrm{red,F}}}{dt} = V_{\mathrm{IIa,Fg}} \frac{x_{\mathrm{red,IIa}}}{x_{\mathrm{red,IIa}} + K_{\mathrm{IIa,Fg}}} \cdot x_{\mathrm{red,Fg}} - d_{\mathrm{F}} \cdot x_{\mathrm{red,F}}$$

$$\frac{dx_{\mathrm{red,Xa}}}{dt} = V_{\mathrm{VIIa:TF,X}} \frac{x_{\mathrm{red,VII:TF}}}{x_{\mathrm{red,VII:TF}} + K_{\mathrm{VII:TF,X}}} \cdot x_{\mathrm{env,X}} - c_{\mathrm{TFPI,Xa}} \cdot x_{\mathrm{env,TFPI}} \cdot x_{\mathrm{red,Xa}} - d_{\mathrm{Xa}} \cdot x_{\mathrm{red,Xa}}$$

$$\frac{dx_{\mathrm{red,VII:TF}}}{dt} = c_{\mathrm{VII,TF}} \cdot x_{\mathrm{red,TF}} \cdot x_{\mathrm{env,VII}} - \left(V_{\mathrm{Xa,VII:TF}} \frac{x_{\mathrm{red,Xa}}}{x_{\mathrm{red,Xa}} + K_{\mathrm{Xa,VII:TF}}} + V_{\mathrm{TF,VII:TF}} \frac{x_{\mathrm{red,TF}}}{x_{\mathrm{red,TF}} + K_{\mathrm{TF,VII:TF}}} \right) \cdot x_{\mathrm{red,VII:TF}} - d_{\mathrm{VII:TF}} \cdot x_{\mathrm{red,VII:TF}}$$

$$\frac{dx_{\mathrm{red,VIIa:TF}}}{dt} = \left(V_{\mathrm{Xa,VII:TF}} \frac{x_{\mathrm{red,Xa}}}{x_{\mathrm{red,Xa}} + K_{\mathrm{Xa,VII:TF}}} + V_{\mathrm{TF,VII:TF}} \frac{x_{\mathrm{red,TF}}}{x_{\mathrm{red,TF}} + K_{\mathrm{TF,VII:TF}}} \right) \cdot x_{\mathrm{red,VII:TF}} - d_{\mathrm{VIIa:TF}} \cdot x_{\mathrm{red,VIIa:TF}}$$

$$\frac{dx_{\mathrm{red,TF}}}{dt} = -c_{\mathrm{VII,TF}} \cdot x_{\mathrm{red,TF}} \cdot x_{\mathrm{env,VII}} - d_{\mathrm{TF}} \cdot x_{\mathrm{red,TF}}$$

All parameter values and initial conditions for the state variables can be found in [130, Suppl.Fig. 2&Fig. 3]. The state variable $x_{\mathrm{env,II}}$ is equal to the initial value of II ($x_{\mathrm{II}}(0)$).

This holds for all environmental state variables.
The model equations for the 13-state variable elimination-reduced model for the low TF are:

$$\begin{aligned}
\frac{dx_{\text{red,IIa}}}{dt} &= V_{\text{Xa,IIa}}\frac{x_{\text{red,Xa}}}{x_{\text{red,Xa}}+K_{\text{Xa,IIa}}}\cdot x_{\text{env,II}} - d_{\text{IIa}}\cdot x_{\text{red,IIa}}\\
\frac{dx_{\text{red,Fg}}}{dt} &= -V_{\text{IIa,Fg}}\frac{x_{\text{red,IIa}}}{x_{\text{red,IIa}}+K_{\text{IIa,Fg}}}\cdot x_{\text{red,Fg}} - d_{\text{Fg}}\cdot x_{\text{red,Fg}}\\
\frac{dx_{\text{red,F}}}{dt} &= V_{\text{IIa,Fg}}\frac{x_{\text{red,IIa}}}{x_{\text{red,IIa}}+K_{\text{IIa,Fg}}}\cdot x_{\text{env,Fg}} - d_{\text{F}}\cdot x_{\text{red,F}}\\
\frac{dx_{\text{red,Xa}}}{dt} &= \left(V_{\text{VIIa:TF,X}}\frac{x_{\text{red,VII:TF}}}{x_{\text{red,VII:TF}}+K_{\text{VII:TF,X}}} + V_{\text{IXa,X}}\frac{x_{\text{red,IXa}}}{x_{\text{red,IXa}}+K_{\text{IXa,X}}}\right.\\
&\quad \left. + V_{\text{IXa:VIIIa,X}}\frac{x_{\text{red,IXa:VIIIa}}}{x_{\text{red,IXa:VIIIa}}+K_{\text{IXa:VIIIa,X}}}\right)\cdot x_{\text{env,X}}\\
&\quad - c_{\text{TFPI,Xa}}\cdot x_{\text{env,TFPI}}\cdot x_{\text{red,Xa}} - d_{\text{Xa}}\cdot x_{\text{red,Xa}}\\
\frac{dx_{\text{red,VII:TF}}}{dt} &= c_{\text{VII,TF}}\cdot x_{\text{red,TF}}\cdot x_{\text{env,VII}} - \left(V_{\text{Xa,VII:TF}}\frac{x_{\text{red,Xa}}}{x_{\text{red,Xa}}+K_{\text{Xa,VII:TF}}}\right.\\
&\quad \left. + V_{\text{TF,VII:TF}}\frac{x_{\text{red,TF}}}{x_{\text{red,TF}}+K_{\text{TF,VII:TF}}}\right)\cdot x_{\text{red,VII:TF}} - d_{\text{VII:TF}}\cdot x_{\text{red,VII:TF}}\\
\frac{dx_{\text{red,VIIa:TF}}}{dt} &= \left(V_{\text{Xa,VII:TF}}\frac{x_{\text{red,Xa}}}{x_{\text{red,Xa}}+K_{\text{Xa,VII:TF}}} + V_{\text{TF,VII:TF}}\frac{x_{\text{red,TF}}}{x_{\text{red,TF}}+K_{\text{TF,VII:TF}}}\right)\cdot x_{\text{red,VII:TF}}\\
&\quad - d_{\text{VIIa:TF}}\cdot x_{\text{red,VIIa:TF}}\\
\frac{dx_{\text{red,TF}}}{dt} &= -c_{\text{VII,TF}}\cdot x_{\text{red,TF}}\cdot x_{\text{env,VII}} - d_{\text{TF}}\cdot x_{\text{red,TF}}\\
\frac{dx_{\text{red,VIII}}}{dt} &= -V_{\text{IIa,VIII}}\frac{x_{\text{red,IIa}}}{x_{\text{red,IIa}}+K_{\text{IIa,VIII}}}\cdot x_{\text{red,VIII}} - d_{\text{VIII}}\cdot x_{\text{red,VIII}}\\
\frac{dx_{\text{red,VIIIa}}}{dt} &= V_{\text{IIa,VIII}}\frac{x_{\text{red,IIa}}}{x_{\text{red,IIa}}+K_{\text{IIa,VIII}}}\cdot x_{\text{red,VIII}} - (c_{\text{IXa,VIIIa}}\cdot x_{\text{red,IXa}} + d_{\text{VIIIa}})\cdot x_{\text{red,VIIIa}}\\
\frac{dx_{\text{red,IXa}}}{dt} &= V_{\text{VIIa:TF,IX}}\frac{x_{\text{red,VIIa:TF}}}{x_{\text{red,VIIa:TF}}+K_{\text{VIIa:TF,IX}}}\cdot x_{\text{env,IX}} - (c_{\text{IXa,VIIIa}}\cdot x_{\text{red,VIIIa}} + d_{\text{IXa}})\cdot x_{\text{red,IXa}}\\
\frac{dx_{\text{red,IXa:VIIIa}}}{dt} &= c_{\text{IXa,VIIIa}}\cdot x_{\text{red,IXa}}\cdot x_{\text{red,VIIIa}} - d_{\text{IXa:VIIIa}}\cdot x_{\text{red,IXa:VIIIa}}\\
\frac{dx_{\text{red,Va}}}{dt} &= V_{\text{IIa,V}}\frac{x_{\text{red,IIa}}}{x_{\text{red,IIa}}+K_{\text{IIa,V}}}\cdot x_{\text{env,V}} - d_{\text{IXa:VIIIa}}\cdot x_{\text{red,Va}}\\
\frac{dx_{\text{red,Xa:Va}}}{dt} &= c_{\text{Xa,Va}}\cdot x_{\text{red,Xa}}\cdot x_{\text{red,Va}} - d_{\text{Xa:Va}}\cdot x_{\text{red,Xa:Va}}
\end{aligned}$$

All parameter values and initial conditions for the state variables can be found in [130, Suppl.Fig. 2&Fig. 3]. The state variable $x_{\text{env,II}}$ is equal to the initial value of II ($x_{\text{II}}(0)$). This holds for all environmental state variables.

I.3. Sensitivity input-response indices for the warfarin-fibrin system for varying time of blood withdrawal T_{Blood}

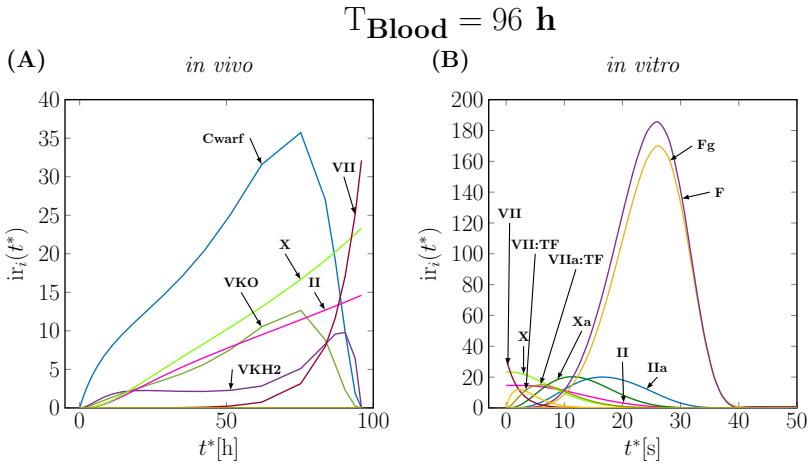


Figure 45: **Sensitivity based ir-indices for the warfarin-fibrin system during the initiation of warfarin therapy.** Time of blood withdrawal was chosen to be $T_{Blood} = 96$ h. Shown are the sensitivity based ir-indices for the most important state variables of the **(A)** *in vivo* and **(B)** *in vitro* parts under warfarin therapy. For better readability, the ir-indices were shown for *in vivo* and *in vitro* part in two different panels. Sensitivity based ir-indices not shown are below 2 in magnitude.

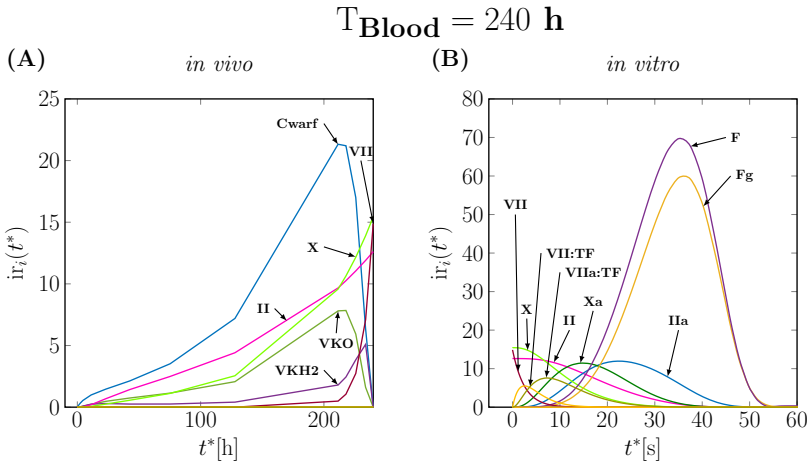


Figure 46: **Sensitivity based ir-indices for the warfarin-fibrin system during the maintance phase of warfarin therapy.** Time of blood withdrawal was chosen to be $T_{Blood} = 240$ h. Shown are the sensitivity based ir-indices for the most important state variables of the **(A)** *in vivo* and **(B)** *in vitro* parts under warfarin therapy. For better readability, the ir-indices were shown for *in vivo* and *in vitro* part in two different panels. Sensitivity based ir-indices not shown are below 1 in magnitude.

Comparison of warfarin plasma concentrations predicted by the empirical PKPD [45] and mechanistic model [130]

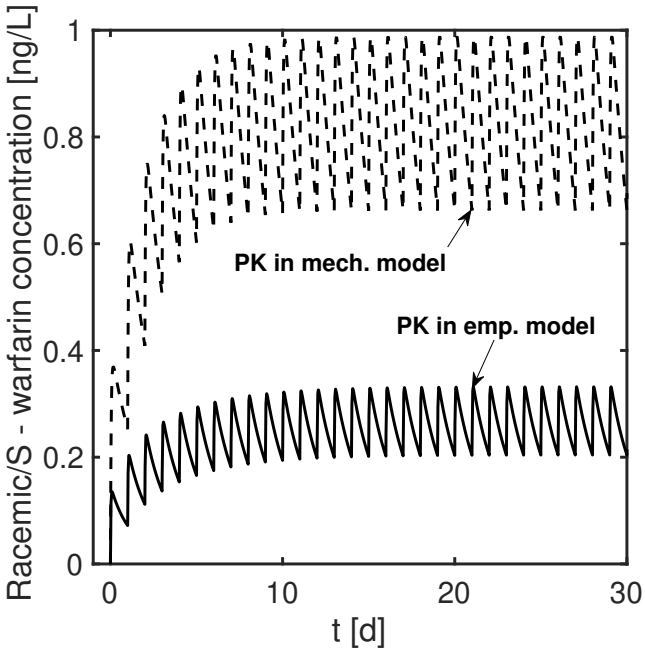


Figure 47: **Comparison of warfarin plasma concentrations predicted by the empirical PKPD [45] and mechanistic model [130]** Warfarin plasma concentration given for warfarin therapy of 4 mg racemic warfarin dose.

I.4. Sensitivity input-response indices of the EGF-ERK-PP signalling cascade

Table 16: **Classification of state variables of the EGFR system based on their ir-indices.** State variables were classified as dynamic (**dyn**); environmental (**env**); neglected (**neg**); in quasi-steady state (**qss**), or implicitly determined via conservation laws (**con**). For details, see section 5.2.1. For any complex with a phosphatase, we (as others) used the abbreviation P'ase for Phosphatase.

dyn	env	neg	qss	con	State variable
		x			AUC ERK-PP
	x				EGF
x					EGFR
		x			EGFRi
		x			$EGFRi_{deg}$
x					EGF-EGFR
x					$(EGF\text{-}EGFR)_2$
x					$(EGF\text{-}EGFR^*)_2$
x					$(EGF\text{-}EGFR^*)_2$-GAP
			x		$(EGF\text{-}EGFR^*)_2$-GAP-Grb2
			x		$(EGF\text{-}EGFR^*)_2$-GAP-Grb2-Prot
			x		$(EGF\text{-}EGFR^*)_2$-GAP-Grb2-Sos
			x		$(EGF\text{-}EGFR^*)_2$-GAP-Grb2-Sos-ERK-PP
			x		$(EGF\text{-}EGFR^*)_2$-GAP-Grb2-Sos-Prot
x					$(EGF\text{-}EGFR^*)_2$-GAP-Grb2-Sos-Ras-GDP
		x			$(EGF\text{-}EGFR^*)_2$-GAP-Grb2-Sos-Ras-GDP-Prot
			x		$(EGF\text{-}EGFR^*)_2$-GAP-Grb2-Sos-Ras-GTP
			x		$(EGF\text{-}EGFR^*)_2$-GAP-Grb2-Sos-Ras-GTP-Prot
		x			$(EGF\text{-}EGFR^*)_2$-GAP-Grb2-Sos_{deg}
			x		$(EGF\text{-}EGFR^*)_2$-GAP-Shc
x					$(EGF\text{-}EGFR^*)_2$-GAP-Shc*
			x		$(EGF\text{-}EGFR^*)_2$-GAP-Shc*-Grb2
			x		$(EGF\text{-}EGFR^*)_2$-GAP-Shc*-Grb2-Prot
x					$(EGF\text{-}EGFR^*)_2$-GAP-Shc*-Grb2-Sos
			x		$(EGF\text{-}EGFR^*)_2$-GAP-Shc*-Grb2-Sos-ERK-PP
x					$(EGF\text{-}EGFR^*)_2$-GAP-Shc*-Grb2-Sos-Prot
			x		$(EGF\text{-}EGFR^*)_2$-GAP-Shc*-Grb2-Sos-Ras-GDP
x					$(EGF\text{-}EGFR^*)_2$-GAP-Shc*-Grb2-Sos-Ras-GDP-Prot
			x		$(EGF\text{-}EGFR^*)_2$-GAP-Shc*-Grb2-Sos-Ras-GTP
			x		$(EGF\text{-}EGFR^*)_2$-GAP-Shc*-Grb2-Sos-Ras-GTP-Prot
		x			$(EGF\text{-}EGFR^*)_2$-GAP-Shc*-Grb2-Sos_{deg}
		x			EGF-EGFRi
		x			$(EGF\text{-}EGFRi)_2$

Table is continued on next page

Table 16 – cont.

dyn	env	neg	qss	con	State variable
			x		$(EGF\text{-}EGFRi^*)_2$
			x		$(EGF\text{-}EGFRi^*)_2$-GAP
			x		$(EGF\text{-}EGFRi^*)_2$-GAP-Grb2
			x		$(EGF\text{-}EGFRi^*)_2$-GAP-Grb2-Sos
		x			$(EGF\text{-}EGFRi^*)_2$-GAP-Grb2-Sos-ERKi-PP
			x		$(EGF\text{-}EGFRi^*)_2$-GAP-Grb2-Sos-Ras-GDP
			x		$(EGF\text{-}EGFRi^*)_2$-GAP-Grb2-Sos-Ras-GTP
		x			$(EGF\text{-}EGFRi^*)_2$-GAP-Grb2-Sos-Ras_{deg}
		x			$(EGF\text{-}EGFRi^*)_2$-GAP-Grb2-Sos_{deg}
		x			$(EGF\text{-}EGFRi^*)_2$-GAP-$Grb2_{deg}$
			x		$(EGF\text{-}EGFRi^*)_2$-GAP-Shc
		x			$(EGF\text{-}EGFRi^*)_2$-GAP-Shc_{deg}
			x		$(EGF\text{-}EGFRi^*)_2$-GAP-Shc*
			x		$(EGF\text{-}EGFRi^*)_2$-GAP-Shc*-Grb2
x					$(EGF\text{-}EGFRi^*)_2$-GAP-Shc*-Grb2-Sos
		x			$(EGF\text{-}EGFRi^*)_2$-GAP-Shc*-Grb2-Sos-ERKi-PP
			x		$(EGF\text{-}EGFRi^*)_2$-GAP-Shc*-Grb2-Sos-Ras-GDP
			x		$(EGF\text{-}EGFRi^*)_2$-GAP-Shc*-Grb2-Sos-Ras-GTP
		x			$(EGF\text{-}EGFRi^*)_2$-GAP-Shc*-Grb2-Sos-Ras_{deg}
		x			$(EGF\text{-}EGFRi^*)_2$-GAP-Shc*-Grb2-Sos_{deg}
		x			$(EGF\text{-}EGFRi^*)_2$-GAP-Shc*-$Grb2_{deg}$
		x			$(EGF\text{-}EGFRi^*)_2$-GAP_{deg}
		x			$(EGF\text{-}EGFRi^*)_{2,deg}$
		x			EGFi
		x			EGFideg
x					ERK
			x		ERK-MEK-PP
x					ERK-P
x					ERK-PP
x					ERK-PP-P'ase3
x					ERK-P-MEK-PP
x					ERK-P-P'ase3
		x			ERKi-MEKi-PP
		x			ERKi-P
		x			ERKi-PP
		x			ERKi-PP-P'ase3i
		x			ERKi-P-MEKi-PP
		x			ERKi-P-P'ase3i
	x				GAP
			x		Grb2

Table is continued on next page

Table 16 – cont.

dyn	env	neg	qss	con	State variable
x					Grb2-Sos
x					MEK
x					MEK-P
x					MEK-PP
			x		MEK-PP-P'ase2
				x	MEK-P-P'ase2
x					MEK-P-Raf*
x					MEK-Raf*
		x			MEKi-P
		x			MEKi-PP
		x			MEKi-PP-P'ase2i
		x			MEKi-P-P'ase2i
		x			MEKi-P-Rafi*
		x			MEKi-Rafi*
				x	Phosphatase1 (P'ase1)
				x	Phosphatase2 (P'ase2)
x					Phosphatase3 (P'ase3)
				x	Prot
		x			Proti
x					Raf
			x		Raf-Ras-GTP
			x		Raf*
				x	Raf*-P'ase
		x			Rafi-Rasi-GTP
		x			Rafi*
		x			Rafi*-P'ase
x					Ras-GDP
x					Ras-GTP
x					Ras-GTP*
			x		Rasi-GTP
				x	Rasi-GTP*
				x	Shc
			x		Shc*
			x		Shc*-Grb2
				x	Shc*-Grb2-Sos
	x				Sos
			x		Sos-ERK-PP
		x			Sos-ERKi-PP
		x			Sosi
28	**3**	**39**	**34**	**8**	**Sum**

Table 17: **Modularisation of the state variables of the EGFR system** adapted from existing modularisation in literature [132, 4, 79, 21] . Note that state variables can belong to more than one module. The modularisation is introduced to display the sensitivity based ir-indices into meaningful groups in Figs. 48-56.

Module name	**Module state variables**
1 - EGF and EGFR	EGF, EGFR, EGF-EGFR, $(EGF\text{-}EGFR)_2$, $(EGF\text{-}EGFR^*)_2$, EGFRi, $EGFRi_{deg}$, EGFi, EGF-EGFRi, $(EGF\text{-}EGFRi)_2$, $(EGF\text{-}EGFRi^*)_2$, $(EGF\text{-}EGFRi^*)_{2deg}$, $EGFi_{deg}$
2a - Pathway without SHC (membrane forms)	$(EGF\text{-}EGFR^*)_2$-GAP, $(EGF\text{-}EGFR^*)_2$-GAP-Grb2, $(EGF\text{-}EGFR^*)_2$-GAP-Grb2-Prot, $(EGF\text{-}EGFR^*)_2$-GAP-Grb2-Sos, $(EGF\text{-}EGFR^*)_2$-GAP-Grb2-Sos-Prot, $(EGF\text{-}EGFR^*)_2$-GAP-Grb2-Sos-Ras-GDP, $(EGF\text{-}EGFR^*)_2$-GAP-Grb2-Sos-Ras-GTP, $(EGF\text{-}EGFR^*)_2$-GAP-Grb2-Sos-Ras-GDP-Prot, $(EGF\text{-}EGFR^*)_2$-GAP-Grb2-Sos-Ras-GTP-Prot, $(EGF\text{-}EGFR^*)_2$-GAP-Grb2-Sos-ERK-PP, $(EGF\text{-}EGFR^*)_2$-GAP-Grb2-Sos_{deg}
2b - Pathway without SHC (internalised forms)	$(EGF\text{-}EGFRi^*)_2$-GAP, $(EGF\text{-}EGFRi^*)_2$-GAP-Grb2, $(EGF\text{-}EGFRi^*)_2$-GAP-Grb2-Sos, $(EGF\text{-}EGFRi^*)_2$-GAP-Grb2-Sos-Ras-GDP, $(EGF\text{-}EGFRi^*)_2$-GAP-Grb2-Sos-Ras-GTP, $(EGF\text{-}EGFRi^*)_2$-GAP-Grb2-Sos-ERKi-PP, $(EGF\text{-}EGFRi^*)_2$-GAP-Grb2-Sos_{deg}
3a - Pathway including SHC (membrane forms)	$(EGF\text{-}EGFR^*)_2$-GAP, Shc, Shca, Shca-Grb2, Shca-Grb2-Sos, $(EGF\text{-}EGFR^*)_2$-GAP-Shc, $(EGF\text{-}EGFR^*)_2$-GAP-Shca, $(EGF\text{-}EGFR^*)_2$-GAP-Shca-Grb2, $(EGF\text{-}EGFR^*)_2$-GAP-Shca-Grb2-Prot, $(EGF\text{-}EGFR^*)_2$-GAP-Shca-Grb2-Sos, $(EGF\text{-}EGFR^*)_2$-GAP-Shca-Grb2-Sos-Prot, $(EGF\text{-}EGFR^*)_2$-GAP-Shca-Grb2-Sos-Ras-GDP, $(EGF\text{-}EGFR^*)_2$-GAP-Shca-Grb2-Sos-Ras-GTP, $(EGF\text{-}EGFR^*)_2$-GAP-Shca-Grb2-Sos-Ras-GDP-Prot, $(EGF\text{-}EGFR^*)_2$-GAP-Shca-Grb2-Sos-Ras-GTP-Prot, $(EGF\text{-}EGFR^*)_2$-GAP-Shca-Grb2-Sos-ERK-PP, $(EGF\text{-}EGFR^*)_2$-GAP-Shca-Grb2-Sos_{deg}
3b - Pathway including SHC (internalised forms)	$(EGF\text{-}EGFRi^*)_2$-GAP-Shc, $(EGF\text{-}EGFRi^*)_2$-GAP-Shca, $(EGF\text{-}EGFRi^*)_2$-GAP-Shca-Grb2, $(EGF\text{-}EGFRi^*)_2$-GAP-Shca-Grb2-Sos, $(EGF\text{-}EGFRi^*)_2$-GAP-Shca-Grb2-Sos-Ras-GDP, $(EGF\text{-}EGFRi^*)_2$-GAP-Shca-Grb2-Sos-Ras-GTP, $(EGF\text{-}EGFRi^*)_2$-GAP-Shca-Grb2-Sos-ERKi-PP
4 - State variables common to both pathways	Prot, Proti, GAP, Grb2, Sos, Sosi, Grb2-Sos, Sos-ERK-PP, Sos-ERKi-PP

Table is continued on next page

Table 17 – cont.

Module name	Module state variables
5 - Ras and Raf	Ras-GTP, Ras-GDP, Raf, Raf-Ras-GTP, Ras-GTP*, Raf*, Raf*-Pase, Phosphatase1, Rasi-GTP, Rafi-Rasi-GTP, Rasi-GTPa, Rafia, Rafi*-Pase
6 - MEK	MEK, MEK-Raf*, MEK-P, MEK-P-Raf*, MEK-PP, MEK-PP-Pase2, Phosphatase2, MEK-P-Pase2, MEKi-Rafi*, MEKi-P, MEKi-P-Rafi*, MEKi-PP, MEKi-PP-Pase2i, MEKi-P-Pase2i
7 - ERK	ERK, ERK-MEK-PP, ERK-P, ERK-P-MEK-PP, ERK-PP, Phosphatase3, ERK-PP-Pase3, ERK-P-Pase3, ERKi-P, ERKi-MEKi-PP, ERKi-P-MEKi-PP, ERKi-PP, ERKi-PP-Pase3i, ERKi-P-Pase3i

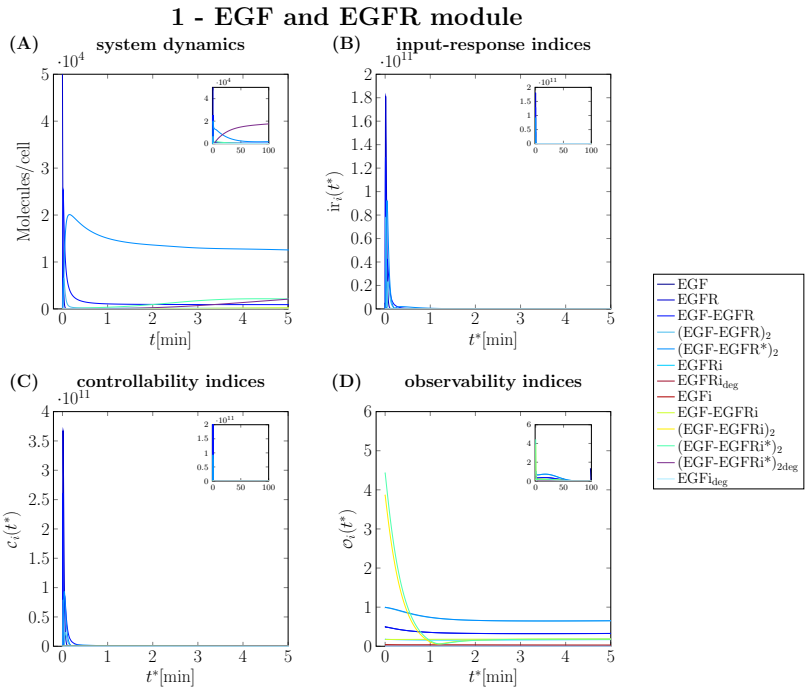


Figure 48: **Concentration-time profiles, input-response, controllability and observability indices of the 'EGF and EGFR' module** (see Table 17) on the time interval $[0, 5]$min (inset: $[0, 100]$min): concentration-time profiles **(A)**; input-response **(B)**, controllability **(C)** and observability indices **(D)**.

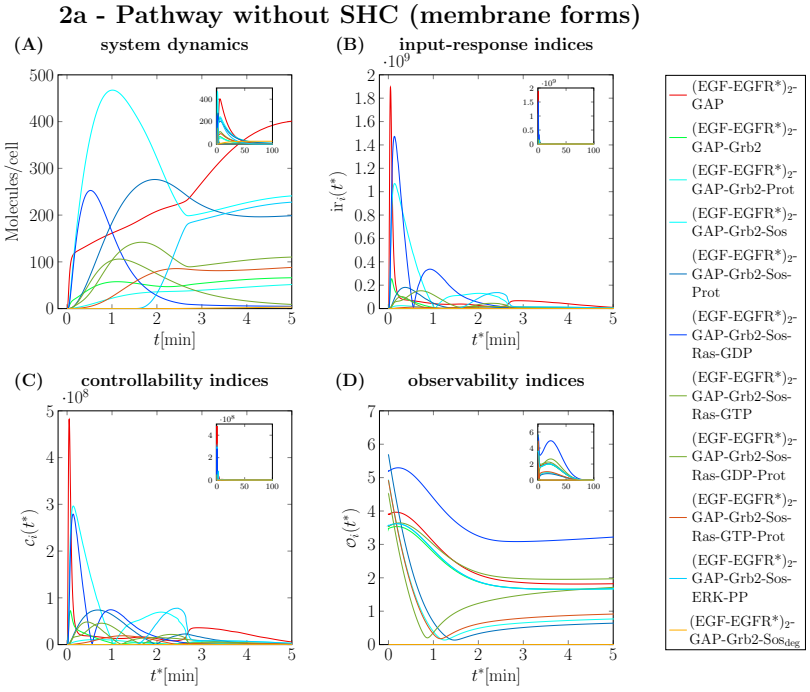


Figure 49: **Concentration-time profiles, input-response, controllability and observability indices of the 'Pathway without SHC (membrane forms)' module** (see Table 17) on the time interval $[0, 5]$min (inset: $[0, 100]$min): concentration-time profiles **(A)**; ir- **(B)**, controllability **(C)** and observability indices **(D)**.

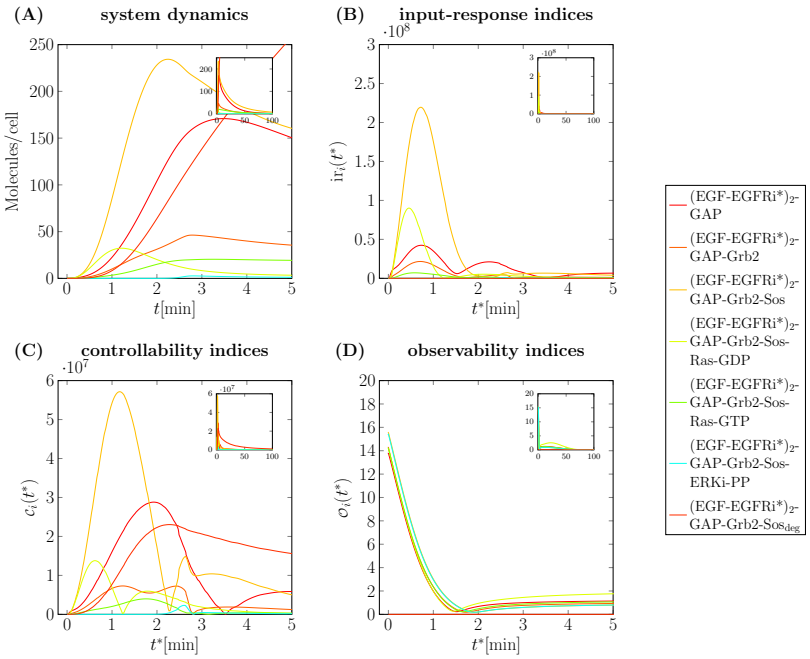


Figure 50: **Concentration-time profiles, input-response, controllability and observability indices of the 'Pathway without SHC (internalised forms)' module** (see Table 17) on the time interval $[0, 5]$min (inset: $[0, 100]$min): concentration-time profiles **(A)**; input-response **(B)**, controllability **(C)** and observability indices **(D)**.

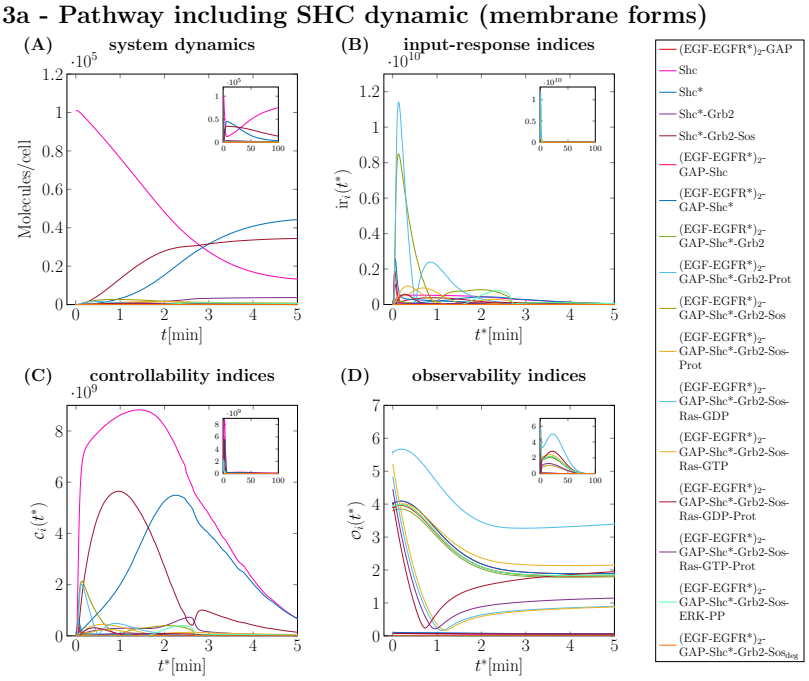


Figure 51: **Concentration-time profiles, input-response, controllability and observability indices of the 'Pathway with SHC (membrane forms)' module** (see Table 17) on the time interval $[0, 5]$min (inset: $[0, 100]$min): concentration-time profiles **(A)**; input-response **(B)**, controllability **(C)** and observability indices **(D)**.

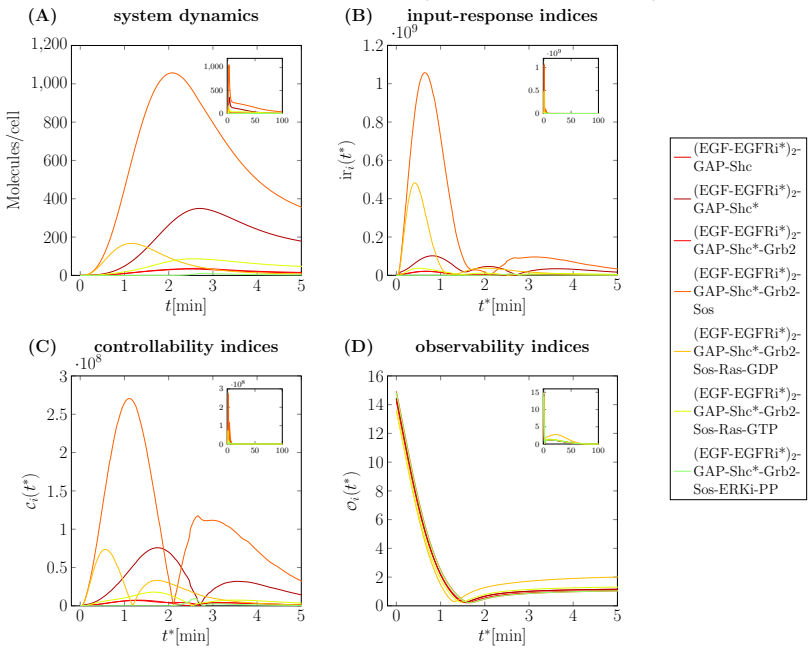


Figure 52: **Concentration-time profiles, input-response, controllability and observability indices of the 'Pathway with SHC (internalised forms)' module** (see Table 17) on the time interval $[0, 5]$min (inset: $[0, 100]$min): concentration-time profiles **(A)**; input-response **(B)**, controllability **(C)** and observability indices **(D)**.

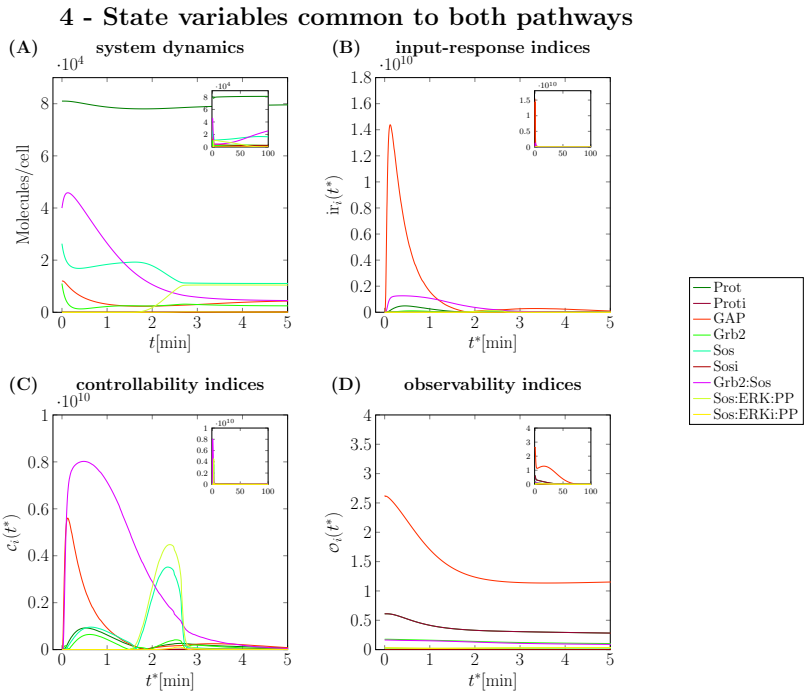


Figure 53: **Concentration-time profiles, input-response, controllability and observability indices of the 'State variables common to both pathways' module** (see Table 17) on the time interval $[0, 5]$min (inset: $[0, 100]$min): concentration-time profiles **(A)**; input-response **(B)**, controllability **(C)** and observability indices **(D)**.

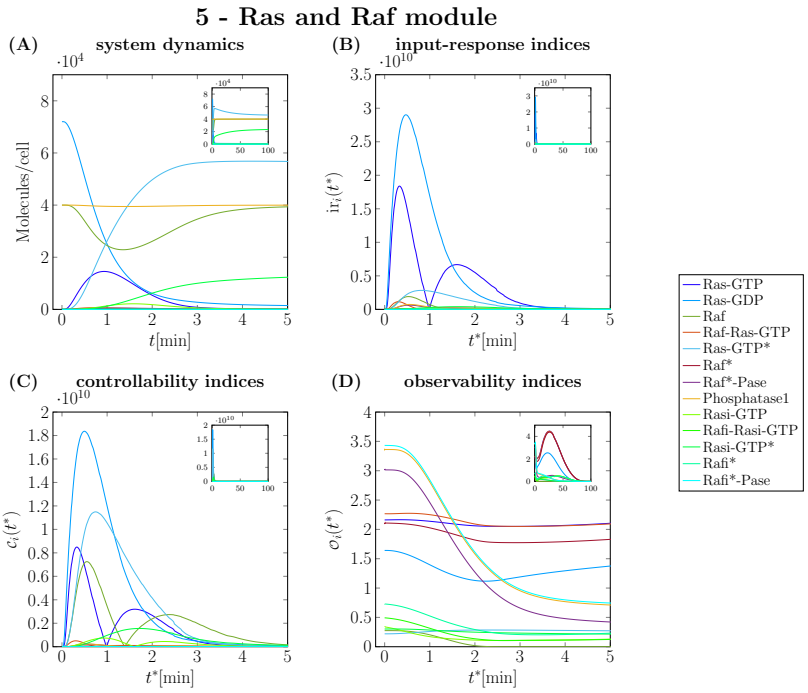


Figure 54: **Concentration-time profiles, input-response, controllability and observability indices of the 'Ras and Raf' module** (see Table 17) on the time interval $[0, 5]$min (inset: $[0, 100]$min): concentration-time profiles **(A)**; input-response **(B)**, controllability **(C)** and observability indices **(D)**.

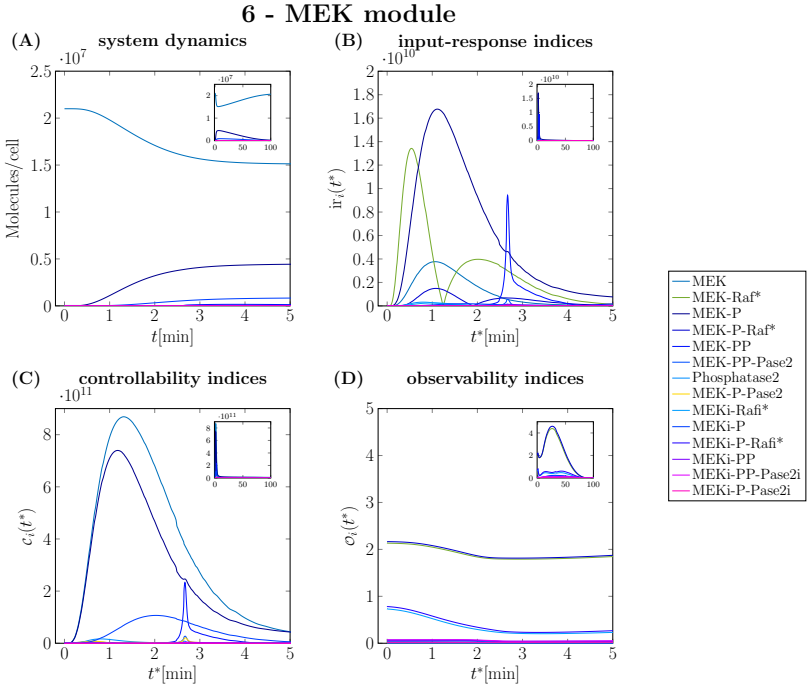


Figure 55: **Concentration-time profiles, input-response, controllability and observability indices of the 'MEK' module** (see Table 17) on the time interval $[0, 5]$min (inset: $[0, 100]$min): concentration-time profiles **(A)**; input-response **(B)**, controllability **(C)** and observability indices **(D)**.

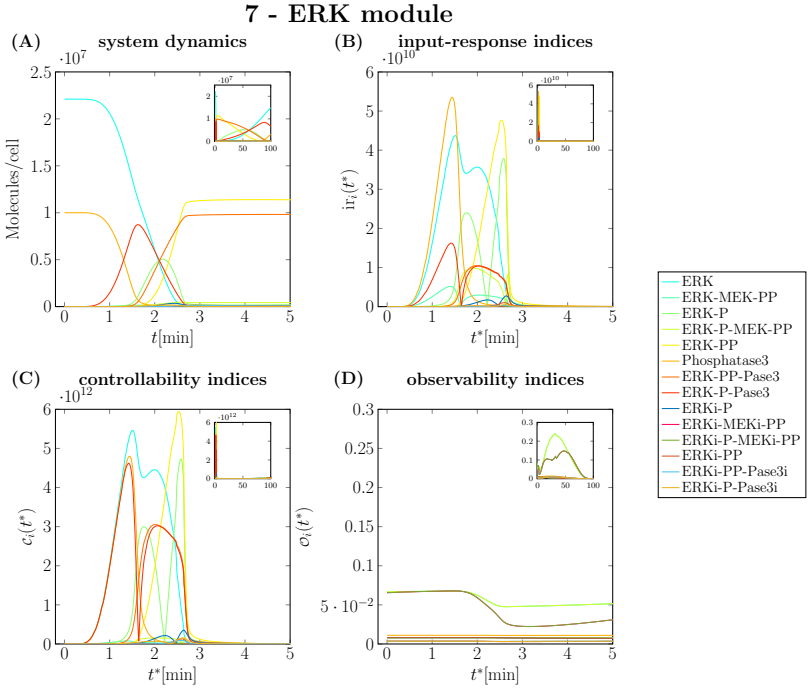


Figure 56: **Concentration-time profiles, input-response, controllability and observability indices of the 'ERK' module** (see Table 17) on the time interval $[0, 5]$min (inset: $[0, 100]$min): concentration-time profiles **(A)**; input-response **(B)**, controllability **(C)** and observability indices **(D)**.

II. Declaration

I hereby declare that I have completed the work solely and only with the help of the mentioned references.

Lübeck, December 9, 2019