Dissertation
submitted to the
Combined Faculties for the Natural Sciences and for Mathematics
of the Ruperto-Carola University of Heidelberg, Germany
for the degree of
Doctor of Natural Sciences

presented by

Diplom-Physiker:	Martin Bentele
born in:	Bietigheim, Germany

Oral examination: November 24$^{\text{th}}$, 2004

Stochastic Simulation and System Identification of large Signal Transduction Networks in Cells

Referees: **Prof. Dr. Dieter W. Heermann**
Prof. Dr. Roland Eils

Zusammenfassung

Komplexe und hochgradig unterbestimmte Netzwerke mit unklarer Netzwerktopologie, teilweise unzureichend verstandenen Mechanismen, vielen unbekannten Parametern sowie signifikanten stochastischen Effekten sind in der Wissenschaft, insbesondere in der molekularen Zellbiologie, weit verbreitet. Neue Ansätze zur mathematischen Modellierung und Systemidentifizierung sind hierbei erforderlich. Zellen übertragen und verarbeiten Signale durch Signaltransduktionsnetzwerke, die auf der biochemischen Wechselwirkung zwischen den beteiligten Molekülen basieren. Die hohe Komplexität ergibt sich aus der hohen Anzahl beteiligter Molekülsorten sowie aus der Vielzahl und Vielschichtigkeit interagierender Subprozesse. Die Modellierung solcher Netze beschränkt sich bislang zumeist auf kleine Subsysteme oder basiert auf rein qualitativer Information. Ein Ziel dieser Arbeit ist es, die Komplexität großer Systeme so zu reduzieren, dass eine Systemidentifizierung auf der Grundlage experimenteller Daten möglich wird. Der hierfür entwickelte Ansatz der "Sensitivität von Sensitivitäten", basierend auf der Auswertung stochastisch generierter Ensembles von Parametersätzen, legt zwei entscheidende, inhärente Systemeigenschaften offen: hohe Robustheit und eine modulare Struktur der Abhängigkeiten zwischen Zustandsvariablen und Parametern. Dies ist von entscheidender Bedeutung, um die Dimensionalität des Parameterschätzproblems deutlich zu reduzieren. Diese Methodik wird auf CD95-induzierte Apoptose, auch programmierter Zelltod genannt, angewendet. Fehler in der Regulierung von Apoptose haben eine Reihe schwerwiegender Krankheiten wie Krebs zur Folge. Obwohl die molekularen Mechanismen in zunehmendem Maße untersucht werden, fehlt bislang ein systemisches Verständnis des komplexen Signalwegs. Mit den hier geschätzten Parametern können beobachtete Prozesse reproduziert und wichtige Systemeigenschaften vorhergesagt werden. Diese wurden experimentell bestätigt und erlauben eine gezielte Planung weiterer Experimente. Dadurch konnte ein neuer regulatorischer Mechanismus entschlüsselt und ein Schwellenwertverhalten zwischen Überleben bzw. Tod einer Zelle identifiziert werden. Aufgrund hoher Fluktuationen und extrem geringer Teilchenzahlen zentraler Molekülsorten sind stochastisch exakte Simulationsmethoden unverzichtbar. Da diese für Systeme, deren Reaktionen auf deutlich unterschiedlichen Zeitskalen stattfinden, jedoch nicht praktikabel sind, wird ein effizienter Hybrid-Algorithmus entwickelt, der den exakten Gillespie-Algorithmus mit einem System stochastischer Differentialgleichungen kombiniert und stochastisch präzise wie auch effiziente Simulationen für jede Art von Markov-Prozessen erlaubt. Zusammenfassend ist die hier vorgestellte Methodik für hochgradig unterbestimmte Netzwerke besonders geeignet und für das neu entstehende Gebiet der Systembiologie von hoher Relevanz, zumal sie Anwendungen ermöglicht, die über Apoptose weit hinausgehen.

Abstract

New approaches are required for the mathematical modelling and system identification of complex networks, which are characterized by a large number of unknown parameters, uncertain network topologies, partially poorly understood mechanisms and significant stochastic effects. Networks with such properties are ubiquitous in many fields of science, especially in molecular cell biology, where, for example, large signal transduction networks are formed, by which cells transfer and process information, based on the biochemical interactions between signal transduction molecules. Complexity arises from the high number of different molecule species involved and the diversity of sub-processes interacting with each other. Previous attempts to model signal transduction were often limited to small systems or based on qualitative data only. One goal of this thesis is to reduce the complexity to enable system identification on the basis of experimental data. The concept of 'Sensitivity of Sensitivities', which is presented here for the first time and which is based on the evaluation of stochastically generated parameter set ensembles, reveals two important inherent system properties: high robustness and modular structures of the dependency between state variables and parameters. This is the key to drastically reduce the dimensionality of the parameter identification problem. The approach is applied to the signalling pathway of CD95-induced apoptosis, also called programmed cell death. Defects in the regulation of apoptosis result in a number of serious diseases such as cancer. Despite the ever-increasing number of studies of the molecular mechanisms of apoptotic signalling, a systemic understanding of this complex pathway is still missing. With the model and the estimated parameters of this thesis, it becomes possible to reproduce the observed system behaviour and to predict important system properties. The predictions have been experimentally confirmed and are used for the planning of further experiments. Thereby, a novel regulatory mechanism was revealed, i.e. a threshold between cell death and cell survival. High fluctuations and extremely low particle numbers of crucial molecule species require exact stochastic simulations. Computational problems arise from the huge differences among the timescales on which the reactions occur. Therefore, a stochastic hybrid algorithm is developed by combining the exact Gillespie algorithm with a system of stochastic differential equations. This enables stochastically accurate and highly efficient simulations for large reaction systems and for any other kind of Markov processes. In summary, this thesis provides a methodology specifically suited for highly underdetermined networks. This is of high relevance for the newly emerging field of systems biology going far beyond the present application of programmed cell death.

Contents

List of Figures

Chapter 1

Introduction

Computer simulation and mathematical modelling have become indispensable for most disciplines of natural sciences. In biology, quantitative models are established in many areas ranging from macromolecular structures to reaction-diffusion processes and other nonlinear dynamical systems [1]. However, up to now, quantitative modelling of most biological systems, like intracellular processes, is impaired by the high complexity arising from the diversity of different sub-processes interacting with each other and the lack of knowledge about the underlying mechanisms, especially on the quantitative level. Here, huge numbers of unknown parameters and large ranges of possible values for most quantities are ubiquitous. Whenever information is not directly accessible by observations, an investigation loop between hypothesis-driven modelling, simulation and experiments becomes essential for system identification [2]. In data-based approaches, unknown system parameters are identified based on the observed dynamical behaviour. Such approaches are, for example, well-established in physics or chemistry. In contrast, many biological questions raise qualitatively different problems due to less reliable experimental data, less information about the exact mechanisms and the high dimensionality of the space of unknown parameters.

One of the most challenging areas in cell biology is a better understanding of *signal transduction networks*. Cells show information processing by the biochemical interaction between molecules. Signals of external stimuli are, for example, passed into the nucleus to regulate gene expression, resulting in proliferation, mitosis (nuclear division), changes in metabolism or cell death [3]. Interactions like phosphorylation, exchange of smaller molecules, binding or cleavage, are the fundamental mechanisms, which form the signal transduction networks. Complexity arises from the huge number of different molecules and interactions between them. In eucaryotic cells, the steadily growing number of known signalling molecule species is currently in the order of magnitude of $10^4 - 10^5$ [4].

Different methods of translating signal transduction networks into the language of mathematics exist in order to analyse their complex and highly dynamical behaviour. Dynamic pathway models are constructed using a diversity of mathematical and computational methods. Petri Nets [5, 6, 7] are, for example, well suited to describe the state transi-

tion process of distributed systems. In agent-based approaches and cellular automata [8], macroscopic system properties emerge from the individual properties of the single entities, interacting with each other. Other methods originate from the analysis of biochemical systems ranging from the examination of steady states and flux modes to a large variety of control theories [9, 10, 11]. More recently, theoretical models for describing the signalling behaviour on systems level have been developed, using modular approaches [2, 12, 13]. Thus, simulation of signal transduction networks is either based on discrete models describing signalling as information processing, or on continuous models, where the information flux is modelled by a biochemical control system. In the latter approach, which goes back to the pioneering work of Garfinkel and Hess in the mid 60s [14, 15], the reaction network is translated into a system of ordinary differential equations [16, 17]. Today, there is a variety of sophisticated simulation methods to analyze complex biochemical reaction systems (e.g. [18, 19, 20, 21]). However, for all methods mentioned above, a fixed set of reaction partners and parameters is assumed in advance. Accordingly, the models are mostly limited to well investigated signalling pathways, where the biochemical mechanisms are even quantitatively well understood [22, 23]. As an alternative, methods for system identification like parameter estimation [24, 25] based on reliably measured time series of molecule concentrations, as successfully applied for chemical reaction systems [26], have been suggested recently [27].

In contrast to well-defined chemical reaction networks, signal transduction pathways usually cannot be regarded as isolated processes, which limits the applicability of *in vitro* data or data from experiments, referring to different experimental settings, cell types or states of cells. Moreover, the available information about signalling processes is spread over different levels of information quality, ranging from mechanistically well understood interactions to purely qualitative relations, like activation or inhibition. Quantitative information on e.g. reaction rate constants is mostly missing and the range of typical parameter values often covers several orders of magnitude. As a consequence, system identification of signal transduction networks - and many other biological systems - is severely impaired by the high number of unknown parameters and the *curse of dimensionality*, which refers to the problem that the space of possible parameter value sets grows exponentially with the number of unknown parameters impairing the search for the most probable set of parameters.

1.1 This Thesis

The goal of this thesis is to introduce quantitative modelling and system identification approaches to complex, high-dimensional and partially poorly understood cellular processes. The main objectives can be summarized as follows:

- Modelling approach for the quantitative description of mechanistically well understood subsystems embedded in larger networks, given by their connectivity.

- Introduction of a data-based approach to make the network model capable of reproducing the observed system behaviour and of predicting new system properties.
- Investigation of stochastic effects in signal transduction networks by development and application of stochastic simulation methods.
- Identification of an adequate signal transduction system as prototype application and formation of a collaboration with an experimental group in order to establish a loop between modelling, simulation and experiments. This is a crucial step for a better understanding of signalling networks on systems level.

So far, only a few models of signal transduction networks exist, which have been approached by data-based system identification methods [27]. Instead, simulations are performed using ad-hoc set parameters (e.g. [28]) or parameters, which refer to incomparable cell types, cell states or experimental settings. The high amount of data required for reliable system identification is of course limiting the size of assessable systems. To approach this problem, I make use of two important and ubiquitous system properties: modularity and robustness. It is a well-accepted fact that biological systems keep their system properties stable although they are subject to high parameter variations (e.g. [29, 30, 31]). Furthermore, cellular systems often exhibit a modular and hierarchical structure [32]. In addition, the effects of parameter variations on the system properties are often correlated. As a consequence, the original parameter sets can be replaced by a much lower number of 'effective' parameters. Based on these assumptions, a new systematic approach is developed to reduce the dimensionality of the system identification problem by determination of parameters, which are either irrelevant for certain network parts or correlated with other parameters.

Sensitivitiy of Sensitivities

Sensitivity analysis [33] plays an essential role in model reduction. However, in general, knowledge about the parameter values is required for identification of insensitive and correlated parameters. On the other hand, model reduction has to be applied prior to the determination of the unknown parameters. In this thesis, I therefore introduce the concept of 'Sensitivity of Sensitivities', a stochastic approach towards global sensitivity analysis, which shows that sensitivity values of typical signal transduction networks are often highly insensitive with respect to large variations of all parameters. On this basis, evidence about parameter dependencies and parameter correlation can be gained even without knowledge of the exact parameter values.

Sensitivity-based Parameter Estimation

Exploiting the latter information, a cluster-based and sensitivity-controlled parameter estimation framework is introduced. Typically, the sensitivities of specific molecules with respect to specific parameters are extremely low for many parameter-molecule combinations. If this property is fulfilled within the complete space of possible parameter values, at

least for a subset of these sensitivities, the model separates into clusters of molecules, which are dependent on a subset of parameters only. As a consequence, a hierarchical parameter structure can be introduced. Parameters are called global if more than one cluster depends on them, whereas local parameters are only relevant for the state variables of a single cluster, which drastically reduces the dimensionality of the parameter estimation problem. To my knowledge, this concept is demonstrated here for the first time. Furthermore, the approach is complemented by a sensitivity control within the parameter estimation algorithm in order to account for local sensitivities and parametric sensitivity correlations.

With the previous approaches, parameter estimation becomes practicable for much larger systems. Nevertheless, due to the current data situation, submanifolds of parameter sets providing almost the same fit with experimental data are expected rather than a unique solution. As a consequence, a large number of parameter estimations are performed with randomly chosen start values. Instead of relying on a single parameter set, providing the best fit with experimental data, I introduce the concept of ensembles of parameter fits. For statistical evaluation of system properties, the weight of each model sample based on a certain parameter set is given by a Boltzmann factor, which can be considered the probability that a system with the respective parameters reproduces the experimentally observed behaviour.

Hybrid Method for Stochastic Simulation

Another challenge in quantitative biochemical network modelling is to account for the stochastic behaviour and the discrete nature of the molecules, since high relative fluctuations and extremely low particle numbers of some species are ubiquitous [35]. As a consequence, the system behaviour is often governed by stochastic effects, which poses new problems for parameter estimation. Moreover, computational problems arise from the huge differences among the timescales on which the reactions occur, causing high cost for stochastically exact simulations [36].

In this thesis, I develop the 'General Stochastic Hybrid Method' combining the exact Gillespie algorithm with a system of stochastic differential equations. For this purpose, the reaction system is dynamically divided into two subsets: stochastically critical interactions, e.g. reactions between species of low particle numbers, are simulated as discrete events using the exact Gillespie method, whereas the remaining reactions are approximated by Langevin equations [37]. The interactions between both subsets are correctly considered by propagation of a 'generalized system state' consisting of both the state variables and the probability densities for the discrete events. Thus, a method is provided, which is exact in the sensitive low-particle number range, but still efficient for the fast reaction subset, where the Gillespie approach would no longer be feasible.

Model of Programmed Cell Death

The signal transduction network of programmed cell death, also called *apoptosis* [3], is a highly relevant field of research in cell biology. In this thesis, an important subsystem of this network is chosen as prototype application. Programmed cell death is one of the most complex signalling pathways and an essential property of all higher organisms. Defects in apoptosis result in a number of serious diseases such as cancer, autoimmunity and neurodegeneration (e.g. Altzheimer and Parkinson) [38, 39]. To develop efficient therapies, fundamental questions about molecular mechanisms and regulation of apoptosis remain to be answered.

Despite the ever-increasing number of studies on apoptosis, a systemic understanding of this complex signalling pathway is still missing. There is no experimental approach available at present matching the complexity and allowing monitoring of immediate and long-term changes of all affected molecules in the course of apoptosis. Instead, most experiments focus on the qualitative relation between single molecules and are often measured under conditions that are not comparable. For this purpose, I established an interdisciplinary collaboration with an experimental group focussing on the measurement of quantitatively reliable time series of molecule concentrations. A mathematical model of programmed cell death, integrating the presently distributed and heterogeneous knowledge, is of great benefit, since it allows the identification of most sensitive signalling molecules and predictions on the systemic behaviour of apoptotic signalling, e.g., upon stimulation by different molecules or through interaction of chemotherapeutics. In this study, CD95-induced apoptosis [38], which is one of the most important apoptotic signalling pathways, is investigated. To establish an integrated model of this pathway, I introduce 'Structured Information Models', an approach, which integrates information on various different levels in a unified form. Mechanistically well understood network parts are combined with less specified subsystems, mainly represented as black boxes and defined by the observed input-output behaviour. Besides the formulation of biological hypotheses, a mathematical model is also very beneficial for experimental design by suggesting the most promising next experiments to address a specific biological question.

Like in most signal transduction networks, quantitative information like reaction rates are missing and for most reactions, the respective parameters are not directly accessible *in vivo*. In a first attempt to theoretically describe apoptotic signalling a mathematical model including more than 20 reactions was proposed in the study of Fussenegger et al. [28]. However, this model was based on ad hoc fixed parameters and thus its potential for understanding the regulation of apoptosis remains very limited.

It is shown that the methods developed in this thesis are suited to overcome present obstacles in modelling and system identification of signal transduction networks, or, more generally speaking, stochastic reaction systems with many unknown parameters. The numerical simulations allow the prediction of the systemic behaviour of CD95-induced

apoptosis. By verifying the model hypotheses experimentally, it is demonstrated how iterations of theoretical modelling and experiments reveal new insights into the behaviour of complex biological systems. As a result, a mechanism for the control of CD95-induced apoptosis is identified, which answers highly debated questions about the regulation of programmed cell death.

1.2 Outline of this Thesis

The thesis is divided into three parts. In the first part, an introduction into deterministic and stochastic simulation methods for chemical reaction networks is given (Chapter 2) and the problem of processes running on multiple timescales is addressed (Chapter 3). Here, I derive the General Stochastic Hybrid Method, which combines the stochastically exact with the approximate simulation method. The algorithm is then applied to the signal transduction network of programmed cell death to demonstrate its efficiency and accuracy in comparison with alternative methods.

The second part of the thesis deals with modelling and system identification of signal transduction networks, especially with the problem of many unknown parameters. Chapter 4 presents the signal transduction pathway of CD95-induced apoptosis. In Chapter 5, I introduce the concept of structured information models in order to construct a simplified network model of the pathway. For reduction of the effective number of parameters, a new approach towards global sensitivity analysis is given in Chapter 6. Here, I introduce the new concept of Sensitivities of Sensitivities. After an introduction into parameter estimation in Chapter 7, I derive a cluster-based and sensitivity-controlled parameter estimation algorithm on the basis of Chapter 6. Chapter 8 presents the software package, which integrates the previous methods.

Finally, the results are given in Chapter 9. First, important characteristics of the distribution of sensitivities are shown. Then, the results of the new parameter estimation approach are demonstrated and the model-based predictions of the biological system behaviour are presented. Then, it is shown that these predictions could be experimentally verified and that detailed hypotheses about the underlying mechanisms became possible. At the end, the stochastic system behaviour is addressed and it is shown that simulations using the General Stochastic Hybrid Method of Chapter 3 were capable of reproducing the observed death rates. Chapter 10 puts the approaches and results of this thesis into a broader context and gives perspectives for future research.

Chapter 2

Simulation of Chemical Reaction Models

The quantitative simulation of signal transduction networks originates from (bio-)chemical reaction models. In this chapter, a deterministic simulation method for chemical reaction networks, based on a system of ordinary differential equations, is demonstrated. In addition, two stochastic simulation methods are derived: the exact Gillespie algorithm and an approximate method using stochastic differential equations.

2.1 Deterministic Kinetic Models

2.1.1 Definition of Reaction Systems

In the following, we assume a spatially homogeneous chemical reaction system [41] and consider a *well-stirred mixture* of m molecule species M_i and n different reactions R_j between them. The state of the system is defined by the molecule concentrations of each species

$$\mathbf{c}(t) = (c_1(t), \ldots, c_m(t)). \tag{2.1}$$

These concentrations are typically given in *mole per liter*. Alternatively, the state can be defined by the *amount of substance*, also called *material quantity*

$$\mathbf{n}(t) = (n_1(t), \ldots, n_m(t)), \tag{2.2}$$

which corresponds to the number of particles of each species (typically given in *mole*) within a given volume. Consequently, $(n_1, \ldots, n_m)$ are directly related to the concentrations $(c_1, \ldots, c_m)$.

The chemical reactions $\{R_1, \ldots, R_n\}$ between the molecules within the system are defined by the stoichiometric coefficients and the respective reaction velocities. Assuming that reaction R_j consumes K different educt species and generates L different product

species, the so-called *reaction equation* has the general form

$$(-\nu_{i_1 j})M_{i_1} + \ldots + (-\nu_{i_K j})M_{i_K} \rightarrow (\nu_{i_{K+1} j})M_{i_{K+1}} + \ldots + (\nu_{i_{K+L} j})M_{i_{K+L}}, \tag{2.3}$$

where $i_1, \ldots, i_K$ denote the indices of all educts, $i_{K+1}, \ldots, i_{K+L}$ the indices of all products, and ν_{ij} the stoichiometric coefficients with

$$\begin{aligned} \nu_{ij} &< 0 \text{ for } i \in \{i_1, \ldots, i_K\}, \\ \nu_{ij} &> 0 \text{ for } i \in \{i_{K+1}, \ldots, i_L\}. \end{aligned}$$

2.1.2 Reaction Rates

The *extent of reaction* ξ describes the progress of a chemical reaction and corresponds to the number of single, molecular reaction events occurred since time t_0. Considering reaction R_j only, the respective $\xi_j(t)$ is given by

$$\xi_j(t) := \frac{\Delta n_i}{\nu_{ij}} \qquad \forall \;\; i \in \{1, \ldots m\} \mid \nu_{ij} \neq 0, \tag{2.4}$$
$$\Delta n_i = n_i(t) - n_i(t_0).$$

This leads directly to the reaction velocity v_j of reaction R_j, defined as the extent of reaction per time and per volume. For simulation and analysis of chemical reaction systems, knowledge of the reaction velocities $v_j(\mathbf{c}(t), \mathbf{\Phi}(t)),\ (j = 1, \ldots, n)$ is crucial. In general, they depend on the molecule concentration of the species involved in the respective reaction (Section 2.1.3) and on the parameter set $(\Phi_1, \ldots, \Phi_q)$ containing reaction-specific parameters, like rate constants, or thermodynamic variables such as temperature.

2.1.3 Time Evolution of Reaction Systems

The time evolution of the state of a reaction system can be described by a system of ordinary differential equations (ODEs), generated as linear combination of the reaction velocities v_j. These equations are called *reaction rate equations* and defined by

$$\frac{dc_i}{dt} = \sum_{j=1}^{n} \nu_{ij} v_j(\mathbf{c}(t), \mathbf{\Phi}(t)) \tag{2.5}$$

with the *stoichiometric matrix*

$$\nu = \begin{pmatrix} \nu_{11} & \ldots & \nu_{1n} \\ \vdots & & \vdots \\ \nu_{m1} & \ldots & \nu_{mn} \end{pmatrix}.$$

This matrix contains the stoichiometric coefficients and links the reaction rates with the molecules affected. The determination of reaction velocities depends on the reaction type and is presented for elementary and enzymatic reactions in the following.

Elementary Reactions

In general, reaction velocities cannot be derived from the chemical reaction equations. However, in case of elementary reactions, they are determined by the concentration of the educts [48]:

First order: For unimolecular reactions (e.g. decay processes), the velocity is proportional to the molecule concentration of the consumed molecule

$$v(t) = kc_i(t). \tag{2.6}$$

Second order: For bimolecular reactions, $M_1 + M_2 \rightarrow \ldots$, the velocity is proportional to the concentration of both reaction partners

$$v(t) = kc_1(t)c_2(t). \tag{2.7}$$

In general, the velocity of an elementary reaction

$$(-\nu_{1j})M_1 + \ldots + (-\nu_{kj})M_k \rightarrow \ldots \tag{2.8}$$

has the form

$$v_j(t) = k \sum_{i=1}^{m} c_i^{-\nu_{ij}}(t). \tag{2.9}$$

Enzymatic Reactions

In biochemical systems, enzymatic reactions are of high relevance. A simple model was proposed by Michaelis and Menten:

$$E + S \underset{k_2}{\overset{k_1}{\rightleftharpoons}} ES \xrightarrow{k_3} E + P \ .$$

Enzyme E and substrate S form the complex ES, which can dissociate, or which can be processed to form product P. The production rate of P is derived in the following. Obviously, the time evolution of P and ES is given by

$$d[P]/dt = k_3[ES], \tag{2.10}$$
$$d[ES]/dt = k_1[E][S] - (k_2 + k_3)[ES]. \tag{2.11}$$

The Michaelis-Menten kinetic is based on the assumption that the concentration of ES remains constant (steady state)

$$d[ES]/dt = 0. \tag{2.12}$$

As a consequence, the relation

$$[ES] = \frac{1}{K_M}[E][S], \qquad K_M = \frac{k_2 + k_3}{k_1} \tag{2.13}$$

is fulfilled, where K_M is called *Michaelis constant*. The total enzyme concentration $[E_T]$ is defined as sum of the concentrations of uncombined enzyme E and complex ES

$$[E_T] = [E] + [ES].$$

Together with Eq. (2.10) and Eq. (2.13), where $[E]$ can be substituted by $[E_T] - [ES]$, the rate $V \equiv d[P]/dt$ for the production of P is given by

$$\begin{aligned} V &= k_3[ES] = k_3[E_T]\frac{[S]}{[S] + K_M}, \quad \text{or} \\ V &= V_{max}\frac{[S]}{[S] + K_M}, \qquad V_{max} = k_3[E_T]. \end{aligned} \tag{2.14}$$

The latter equation is the called *Michaelis-Menten equation*.

Numerical Propagation of ODE Systems

Once the initial molecule concentrations are given, the time evolution of chemical reaction systems is treated as an *initial value problem* [42] and the ODE system is numerically integrated. For this purpose, there is a huge variety of ODE solvers, which propagate the system state using finite time steps [43, 44]. Difficulties arise from the high non-linearity of the reaction velocities, which are typical for chemical and especially biochemical systems. Usually, there are periods, in which the time derivatives of state variables are subject to extremely fast changes, requiring extremely small step sizes to keep the numerical exactness at a constant level. For this reason, the required step sizes often differ by orders of magnitude. As a consequence, adaptive step size control or implicit methods [45, 46, 47] are required.

2.2 Stochastic Simulation of Markov Processes

In the previous approach, molecule populations were described by the continuous state variables $(c_1, \ldots, c_m)$, although they refer to molecules, which exist in discrete numbers on the microscopic level. Further, a predictable system was assumed for the determination of reaction velocities and the time evolution. These two assumptions are not appropriate for systems with extremely low particle numbers of at least some species and high relative fluctuations due to stochastic effects. In many cases, such effects lead to a qualitatively different system behaviour as, for example, demonstrated in Section 3.4.2. Moreover, data from biochemical experiments often refer to a population of many cells and the observed quantities are also influenced by variability among the individual cells. In the following, a stochastic system with fixed (bio-)chemical parameters and fixed initial conditions is considered.

Let us again assume a system of m molecule species and n reactions. The state of the system is defined by the discrete particle numbers $\mathbf{X} \equiv (X_1, \ldots, X_m)$[1] of each species, also called *molecular populations*, which are random variables. Thus, all possible states can be represented on a m-dimensional discrete lattice. Each state transition caused by a single molecular reaction event corresponds to a jump to another lattice site according to the stoichiometric coefficients ν_{ij}. In order to describe this process stochastically, *Markov processes* are introduced and the *master equation* [37, 50] is derived in the following.

2.2.1 The Markov Property and Markov Processes

Given a process and a set of s successive times $(t_1 < t_2 < \ldots < t_s)$, the *conditional probability*

$$P_{r|s-r}(\mathbf{X}_{r+1}, t_{r+1}; \ldots; \mathbf{X}_s, t_s \mid \mathbf{X}_1, t_1; \ldots; \mathbf{X}_r, t_r), \quad r < s, \tag{2.15}$$

is defined as the probability density for the states $\mathbf{X}_{r+1}, \ldots, \mathbf{X}_s$ at time $t_{r+1}, \ldots, t_s$ on condition that the states $\mathbf{X}_1, \ldots, \mathbf{X}_r$ have been passed at time $t_1, \ldots, t_r$.

A Markov process is defined as a stochastic process fulfilling the *Markov property*

$$P_{1|s-1}(\mathbf{X}_s, t_s \mid \mathbf{X}_1, t_1; \ldots; \mathbf{X}_{s-1}, t_{s-1}) = P_{1|1}(\mathbf{X}_s, t_s \mid \mathbf{X}_{s-1}, t_{s-1}) \tag{2.16}$$

for any time sequence $(t_1 < t_2 < \ldots < t_s)$. $P_{1|1}$ is called *transition probability*. Thus, in Markov processes the transition probability from state $\mathbf{X}_{s-1}$ at time t_{s-1} to state $\mathbf{X}_s$ at time t_s is uniquely determined by knowledge of the state at time t_{s-1}. As a consequence, a Markov process is fully determined by the probability density $P_1(\mathbf{X}_1, t_1)$ for the states at time t_1 and the conditional probability density $P_{1|1}(\mathbf{X}_2, t_2 \mid \mathbf{X}_1, t_1)$. On the microscopic level, chemical reaction processes are considered memoryless and therefore described as Markov processes [37].

Chapman-Kolmogorov Equation

At first, a continuous system, whose state is defined by the non-discrete random variable y, and the *conditional probability density* $\mathcal{P}_{r|s-r}$ is considered. From the Markov property, the *Chapman-Kolmogorov equation* can be directly derived. It has the general form

$$\mathcal{P}_{1|1}(y_3, t_3 \mid y_1, t_1) = \int \mathcal{P}_{1|1}(y_3, t_3 \mid y_2, t_2)\mathcal{P}_{1|1}(y_2, t_2 \mid y_1, t_1)dy_2. \tag{2.17}$$

For the m-dimensional chemical reaction system with the discrete states $\mathbf{X}_i = (X_{1,i}, \ldots, X_{m,i})$, the equation reads

$$P_{1|1}(\mathbf{X}_3, t_3 \mid \mathbf{X}_1, t_1) = \sum_{X_{1,2}=0}^{\infty} \cdots \sum_{X_{m,2}=0}^{\infty} P_{1|1}(\mathbf{X}_3, t_3 \mid \mathbf{X}_2, t_2)P_{1|1}(\mathbf{X}_2, t_2 \mid \mathbf{X}_1, t_1). \tag{2.18}$$

[1] For a system with a given volume, the concentrations c_i can be directly derived from X_i.

Since chemical reaction processes are assumed to be homogeneous in time, the transition probability $P_{1|1}(\mathbf{X}_2, t_2 \mid \mathbf{X}_1, t_1)$ is not dependent on t_1 and t_2, but only on the time difference $\tau = t_2 - t_1$. For this reason, the transition probability T_τ is introduced as

$$T_\tau(\mathbf{X}_2 \mid \mathbf{X}_1) \equiv P_{1|1}(\mathbf{X}_2, t_2 \mid \mathbf{X}_1, t_1),\ \tau = t_2 - t_1. \tag{2.19}$$

A continuous system with the non-discrete state variable y, in which the conditional probability and the transition probability are denoted by $\mathcal{P}$ and $\mathcal{T}$, can be simplified accordingly. With the definition

$$\mathcal{T}_\tau(y_2 \mid y_1) \equiv \mathcal{P}_{1|1}(y_2, t_2 \mid y_1, t_1),\ \tau = t_2 - t_1, \tag{2.20}$$

the Chapman-Kolmogorov equation (Eq. (2.17)) can be transformed into the form

$$\mathcal{T}_{\tau_0+\tau}(y_3 \mid y_1) = \int \mathcal{T}_\tau(y_3 \mid y_2)\mathcal{T}_{\tau_0}(y_2 \mid y_1)dy_2 \quad (\tau_0 > 0, \tau > 0). \tag{2.21}$$

2.2.2 The Master Equation

The *master equation* is a differential equation directly derived from Eq. (2.20) (or Eq. (2.19), respectively) for the limit $\tau \to 0$. If the second and higher orders of τ are omitted, Eq. (2.20) can be converted into the form

$$\mathcal{T}_\tau(y_2 \mid y_1) = (1 - a_0\tau)\delta(y_2 - y1) + \tau\mathcal{W}(y_2 \mid y_1) + \mathcal{O}(\tau^2). \tag{2.22}$$

Here, $\mathcal{W}(y_2 \mid y_1)$ is the non-negative *transition probability per time* for a state transition from y_1 to y_2 and $(1 - a_0\tau)$ the probability that the system remains in state y_1, depending on

$$a_0(y_1) = \int \mathcal{W}(y_2 \mid y_1)dy_2. \tag{2.23}$$

Inserting Eq. (2.22) into the modified Chapman-Kolmogorov equation (Eq. (2.21)) results in

$$\mathcal{T}_{\tau+\tau'}(y_3 \mid y_1) = (1 - a_0(y_3)\tau')\mathcal{T}_\tau(y_3 \mid y_1) + \tau' \int \mathcal{W}(y_3 \mid y_2)\mathcal{T}_\tau(y_2 \mid y_1)dy_2. \tag{2.24}$$

This equation can be converted into its differential form for $\tau \to 0$, which directly leads to the *master equation*

$$\frac{\partial}{\partial\tau}\mathcal{T}_\tau(y_3 \mid y_1) = \int [\mathcal{W}(y_3 \mid y_2)\mathcal{T}_\tau(y_2 \mid y_1) - \mathcal{W}(y_2 \mid y_3)\mathcal{T}_\tau(y_3 \mid y_1)]dy_2. \tag{2.25}$$

The equation can be simplified by replacing the expression $\mathcal{T}_\tau(y_2 \mid y_1)$ with the probability distribution $\mathcal{P}_1(y_2)$ at time τ. In the following, this probability distribution is called $\mathcal{P}(y, t)$, the probability for state y at time t. Consequently, Eq. (2.25) reads

$$\frac{\partial}{\partial t}\mathcal{P}(y, t) = \int [\mathcal{W}(y \mid y_2)\mathcal{P}(y_2, t) - \mathcal{W}(y_2 \mid y)\mathcal{P}(y, t)]dy_2. \tag{2.26}$$

This central equation can be reformulated for the discrete molecular populations of the chemical reaction system. If r denotes the index, which labels all possible states $\{\mathbf{X}_r\}$, it obtains the form

$$\frac{\partial}{\partial t}P(\mathbf{X}_r,t)=\sum_{r'}[W(\mathbf{X}_r\mid\mathbf{X}_{r'})P(\mathbf{X}_{r'},t)-W(\mathbf{X}_{r'}\mid\mathbf{X}_r)P(\mathbf{X}_r,t)], \tag{2.27}$$

where r' denotes the index of all states $\{\mathbf{X}_{r'}\}$, which are one molecular reaction event away from a given $\mathbf{X}_r$. In other words, the probability for state $\mathbf{X}_r$ increases with the probability that a transition from another state to $\mathbf{X}_r$ takes place (1^{st} term) and decreases with the probability that the system is already in state $\mathbf{X}_r$ and a transition occurs (2^{nd} term). The states $\{\mathbf{X}_{r'}\}$, to which a direct transition is possible once the system is in state $\mathbf{X}_r$, are given by the stoichiometric coefficients $\{\nu_{ij}\}$ of the reactions $\{R_j\}$ according to

$$(X_{1,r'},\ldots,X_{m,r'})=(X_{1,r},\ldots,X_{m,r})+(\nu_{1j},\ldots,\nu_{mj}),\quad j\in\{1,\ldots,n\}. \tag{2.28}$$

Now, the question arises how Eq. (2.27) can be applied in order to simulate the time evolution of the stochastic system once an initial probability distribution is given. In the next sections, two different approaches will be demonstrated:

- the *exact stochastic simulation of coupled chemical reactions*, developed by D.T. Gillespie [51, 36], and
- an approximate solution using stochastic differential equations (SDEs) [55].

2.2.3 Exact stochastic simulation of chemical reaction systems

This section deals with the numerical simulation of Markov-processes. A Monte-Carlo procedure to simulate the chemical reaction system will be demonstrated as proposed by D.T. Gillespie [36]. In the following chapters, 'stochastically exact' is used in the sense of simulations driven by discrete reaction events on the molecular level. Although the master equation of the previous section provides an exact and elegant theoretical basis for the time evolution of the random variables, its mathematical solution is often untractable. Instead, the time evolution is approached by a random walk starting from an initial state and accounting for the transition probabilities of the master equation.

Reaction probabilities and propensity function

First, it is assumed that the chemical reaction system, defined by $(X_1,\ldots,X_m)$, is a well-stirred mixture (spatially homogeneous) and the probability for a single molecular reaction R_j $(j=1,\ldots,m)$ is well-defined in each state. On a microphysical basis, reaction probabilities can for example be derived for a gas phase, where collisions between molecules are considered based on their velocities and cross-sections [52].

As a fundamental hypothesis about the reaction probabilities in general, the existence of a constant c_j is claimed for each reaction R_j, which is defined as follows:

$c_j dt$: probability for the reaction R_j of a particular combination of single molecules within the infinitesimal time interval dt, averaged over all possible combinations of reactants for R_j on molecular level.

Next, this probability is multiplied by the number of all possible molecule combinations for reaction R_j, referred to as h_j. In case of a second order elementary reaction $R_j : M_\alpha + M_\beta \rightarrow \ldots$, for example, this number is given by $h_j = X_\alpha X_\beta$. For a reaction with identical educts (e.g. $M_\alpha + M_\alpha \rightarrow \ldots$), it reads $h_j = \frac{1}{2} X_\alpha (X_\alpha - 1)$. As a consequence, the probability that a reaction R_j will take place somewhere in the system within the time interval dt is described by

$$a_j dt \equiv h_j c_j dt, \tag{2.29}$$

where a_j denotes the probability density for reaction R_j, also called *propensity function*, which is defined as

$a_j(X_1, \ldots, X_m)dt$: probability that, given the state $(X_1, \ldots, X_m)$ at time t, reaction R_j will occur within the time interval $[t, t+dt)$.

Note that there is a proportional relation between the probability density a_j, referring to single molecules and the macroscopic reaction velocity v_j of the deterministic model, which refers to concentrations (e.g. Eq. (2.5)). In the Chapter 3, this property will be applied.

In the following, the master equation (Eq. (2.27)) is reformulated by using the propensities a_j, resulting in

$$\frac{d}{dt} P(X_1, \ldots, X_m, t) = \sum_{j=1}^{n} B_j - \sum_{j=1}^{n} a_j(X_1, \ldots, X_m) P(X_1, \ldots, X_m, t). \tag{2.30}$$

Here, B_j denotes the temporal probability density that state $(X_1, \ldots, X_m)$ is reached by reaction R_j starting from a state, which is exactly one transition R_j away. Thus,

$$B_j = b_j(X_1 - \nu_{1j}, \ldots, X_m - \nu_{mj}) P(X_1 - \nu_{1j}, \ldots, X_m - \nu_{mj}, t) \tag{2.31}$$

with the propensity function b_j for the respective reaction and state. For reasons of computational practicability, Eq. (2.30) is also given in the form

$$P(X_1, \ldots, X_m, t + dt) = P(X_1, \ldots, X_m, t)[1 - \sum_{j=1}^{n} a_j dt] + \sum_{j=1}^{n} B_j dt. \tag{2.32}$$

The stochastic simulation approach

For the stochastic simulation algorithm, the *reaction probability density function* $p(\tau, \mu)$ plays a central role. It is defined as

$p(\tau, \mu)d\tau$: probability that, given the state $(X_1, \ldots, X_m)$ at time t, the next reaction will take place in the time interval $[t + \tau, t + \tau + d\tau)$ and the reaction will be of reaction type R_μ, $\mu = (1, \ldots, n), 0 \leq \tau < \infty$.

The main idea of the Gillespie approach is to define a random walk, which makes use of this function. The time τ and the reaction type μ of the next reaction is chosen randomly according to the probability density function $p(\tau, \mu)$. Then, the reaction is executed, the new state is generated for time $t + \tau$, and the probability density $p(\tau, \mu)$ is recalculated. Therefore, the function $p(\tau, \mu)$ is derived as

$$p(\tau, \mu)d\tau = p_0(\tau)a_\mu d\tau, \tag{2.33}$$

where $p_0(\tau)$ is the probability that no reaction has taken place until the time point τ. The time evolution of $p_0(\tau)$ is consequently given by

$$p_0(\tau' + d\tau') = p_0(\tau')[1 - \sum_{j=1}^{n} a_j d\tau'], \tag{2.34}$$

leading to the simple exponential form

$$p_0(\tau) = e^{-\sum_{j=1}^{n} a_j \tau} \quad \text{or} \quad p_0(\tau) = e^{-a_0 \tau}. \tag{2.35}$$

Here, a_0 denotes the probability densitiy for any of the n possible reactions, consequently

$$a_0 = \sum_{j=1}^{n} a_j. \tag{2.36}$$

Substituting Eq. (2.33), we obtain the final result

$$p(\tau, \mu) = a_\mu e^{-a_0 \tau} \ , \tau \geq 0, \ \mu = (1, \ldots, n). \tag{2.37}$$

Given the state $(X_1, \ldots, X_m)$ at time t, a pair of two random variables (r_1, r_2) of the interval [0,1) can be generated and mapped to the time τ of the next reaction event (at time $t + \tau$) and the type μ of the reaction according to the reaction probability density function

$$\begin{aligned} &\tau = \frac{1}{a_0} \ \ln\left(\frac{1}{r_1}\right), \\ &\sum_{j=1}^{\mu-1} a_j \leq r_2 \ a_0 < \sum_{j=1}^{\mu} a_j. \end{aligned} \tag{2.38}$$

A simulation algorithm driven by this principle is stochastically exact since it is in accordance with the reaction probability density function derived from the master equation (under the assumption that Eq. (2.29) correctly describes the reaction probabilities). In contrast to the deterministic approach, it fully accounts for fluctuations. The stochastically chosen time steps are exact, whereas ODE solvers use finite time steps, always providing approximations with a numerical error bound. However, one stochastic simulation run only corresponds to a single realization of the total ensemble of possible time evolutions of the system starting from a given initial state. Therefore, a large number of runs with different sequences of random numbers is a requirement to obtain reliable values for the cumulants of the distribution functions. Typically, the mean values $\langle X_i \rangle_t$, the second moment $\langle X_i^2 \rangle_t$ and the cross-correlations $\langle X_i X_j \rangle_t - \langle X_i \rangle_t \langle X_j \rangle_t$ for specific times t are of interest. Note that an adequate generation of random numbers has to be applied (Section 8.2). The algorithm is given in the following.

The Gillespie algorithm

1. Initialize the initial particle numbers $(X_1, \ldots, X_m)$ for time t_0 and set $t = t_0$.
2. Calculate the propensities $a_j(X_1, \ldots, X_m)$ for each reaction R_j.
3. Generate 2 random numbers (r_1, r_2) ranging from 0 to 1.
4. Calculate time τ and type μ of the next reaction by mapping (r_1, r_2) to (τ, μ) according to Eq. (2.38).
5. Set the time t to $t + \tau$. Update $(X_1, \ldots, X_m)$ by executing reaction R_μ according to Eq. (2.28).
6. Go to step 2 (or stop if t_{max} is reached).

2.2.4 Approximate solution using Stochastic Differential Equations

In the previous section, a simulation algorithm, which constitutes an exact solution of the master equation, has been introduced. However, even a single simulation run is extremely time-consuming for macroscopic system sizes since the algorithm acts on the molecular level. Moreover, a high number of simulation runs is required to obtain statistically reliable information about mean values and higher moments of the concentrations. Therefore, approximate solutions for the master equation and their numerical treatment, adequate for macroscopic systems with sufficiently high molecule populations, are presented in the following.

Two approaches are demonstrated: the *Fokker-Planck equation*, by which the time evolution of the probability density is described, and the *Langevin equations* providing the time evolution of the state variables themselves. More details about the equations and their derivation can be found in [37].

The Fokker-Planck Equation

The Fokker-Planck equation describes the time evolution of the probability densities and can be directly derived from the master equation. First, the transition probability $\mathcal{W}(y \mid y_0)$ of Eq. (2.25) is substituted by

$$\mathcal{W}(y_0, r) \equiv \mathcal{W}(y \mid y_0), \quad r = y - y_0. \tag{2.39}$$

The introduced variable r quantifies the size of the jump at starting point y_0. Next, the resulting master equation

$$\frac{\partial}{\partial t}\mathcal{P}(y,t) = \int \mathcal{W}(y-r,r)\mathcal{P}(y-r,t)dr - \mathcal{P}(y,t)\int \mathcal{W}(y,-r)dr \tag{2.40}$$

can be simplified under the important assumption that both the transition probability $\mathcal{W}(y-r,r)$ and $\mathcal{P}(y,t)$ itself slowly vary with the first argument. In addition, it is assumed that the jumps are small. Then, a Taylor expansion can be applied and the master equation reads

$$\begin{aligned}\frac{\partial}{\partial t}\mathcal{P}(y,t) = &\int \mathcal{W}(y,r)\mathcal{P}(y,t)dr - \mathcal{P}(y,t)\int \mathcal{W}(y,-r)dr \\ &- \int r\frac{\partial}{\partial y}[\mathcal{W}(y,r)\mathcal{P}(y,t)]dr \\ &+ \frac{1}{2}\int r^2\frac{\partial^2}{\partial y^2}[\mathcal{W}(y,r)\mathcal{P}(y,t)]dr \mp \ \ldots .\end{aligned} \tag{2.41}$$

Note that the first two terms are identical. The remaining terms result in the so-called *Kramers-Moyal expansion*, given by

$$\frac{\partial}{\partial t}\mathcal{P}(y,t) = \sum_{q=1}^{\infty}\frac{(-1)^q}{q!}\frac{\partial^q}{\partial y^q}[a_q(y)\mathcal{P}(y,t)] \tag{2.42}$$

with

$$a_q(y) = \int_{-\infty}^{\infty} r^q \mathcal{W}(y,r)dr. \tag{2.43}$$

Stopping the Taylor expansion after second order, this leads directly to the *forward Fokker-Planck equation*

$$\frac{\partial}{\partial t}\mathcal{P}(y,t) = -\frac{\partial}{\partial y}[a_1(y)\mathcal{P}(y,t)] + \frac{1}{2}\frac{\partial^2}{\partial y^2}[a_2(y)\mathcal{P}(y,t)]. \tag{2.44}$$

The coefficient $a_1(y)$ is called *drift coefficient*, $a_2(y)$ is called *diffusion coefficient*.

Note that this approximation is based on the assumption of small jumps r, which is valid for macroscopic chemical reaction systems with high particle numbers of all species, indicating that the transition probability $\mathcal{W}(y-r,r)$ is a sharply peaked function of r. In addition, it is appropriate to assume that $\mathcal{W}(y-r,r)$ slowly varies with y as long as the reaction probabilities 'smoothly' depend on the state variables (e.g. linear or quadratic relations for first or second order reactions).

The Langevin Equation

The Langevin equation is a stochastic differential equation (SDE), by which the state trajectories are specified rather than the probability distribution of the stochastic process. The Langevin equations were originally introduced to formulate the dynamics of Brownian motion following Einstein's explanation. They can be written in the form

$$\dot{y} = v(y,t) + b(y,t)\eta(t). \tag{2.45}$$

The first term, also called the *drift term*, is deterministic, the second term consists of the deterministic function $b(y,t)$ and a stochastic term $\eta(t)$, for which in chemical reaction systems *Gaussian white noise* is typically assumed. Thus, $\eta(t)$ is a rapidly and irregularly fluctuating random function in time with the properties

$$\begin{aligned} \langle \eta(t) \rangle &= 0 \quad \forall\, t, \\ \langle \eta(t)\eta(t') \rangle &= 0 \quad \forall\, t \neq t', \\ \langle \eta^n(t) \rangle &= 0 \quad \forall\, n \geq 3. \end{aligned}$$

A formal definition of η is given in Section 2.2.4. It can be shown that under certain assumptions, the Langevin equation is equivalent to the Fokker-Planck equation [56].

The multivariate Langevin Equation

The Langevin equation can be derived in different ways [37, 53, 54]. Here, a derivation of the *multivariate Langevin equations*, required e.g. for chemical reaction systems, is outlined as suggested in [57, 58]. Let us assume a system whose state is defined by m continuous random variables $(x_i(t), \ldots, x_m(t))$, driven by a stochastic process, which fulfils the following conditions:

1. The m-variate stochastic process is a Markov process.
2. The random variable $\Xi_i(dt, \mathbf{x}, t) = x_i(t+dt) - x_i(t)$, specifying the increments of each component, given the state $\mathbf{x}$ at time t, is 'smoothly' depending on dt, $\mathbf{x}$ and t.
3. $\Xi_i(dt, \mathbf{x}, t) \to 0$ for $dt \to 0 \quad \forall\, \mathbf{x}$, t.

4. The mean and variance of each random variable Ξ_i is well-defined.

Then, it can be shown that the increments Ξ_i for such a process behave according to

$$\Xi_i(dt, \mathbf{x}, t) = a_i(\mathbf{x}, t)dt + \sum_{j=1}^{m} b_{ij} N_j(t)(dt)^{\frac{1}{2}}. \tag{2.46}$$

The statistically independent *unit normal* random variables $N_j(t) = \mathcal{N}(0, 1)$, where $\mathcal{N}(\mu, \sigma^2)$ denotes the normal random variable with mean μ and variance σ^2, has the properties

$$\begin{aligned} \langle N_j(t) \rangle &= 0 \quad \forall\ t, \\ \langle N_j^2(t) \rangle &= 1 \quad \forall\ t, \\ \langle N_j(t) N_k(t') \rangle &= \delta(t - t')\delta(j - k). \end{aligned}$$

In order to derive Eq. (2.46), the time interval dt is divided into q subintervals of length dt/q starting at the points

$$t_k = t + k\frac{dt}{q},\ (k = 0, \ldots, q). \tag{2.47}$$

For reasons of self-consistency of Markov processes, it is claimed that

$$x_i(t + dt) = \sum_{k=1}^{q} (x_i(t_k) - x_i(t_{k-1})), \tag{2.48}$$

and consequently

$$\Xi_i(dt, \mathbf{x}, t) = \sum_{k=1}^{q} \Xi_i \left(\frac{dt}{q}, \mathbf{x}(t_{k-1}), t_{k-1} \right). \tag{2.49}$$

Together with the upper condition 2 and 3, it can be concluded that $\mathbf{x}(t_{k-1}) \rightarrow \mathbf{x}(t)$ for $t_{k-1} \rightarrow t$, and that in this case, Eq. (2.49) results in

$$\Xi_i(dt, \mathbf{x}, t) = \sum_{k=1}^{q} \Xi_{ik} \left(\frac{dt}{q}, \mathbf{x}(t), t \right) + \mathcal{O}(dt^2), \quad (i = 1, \ldots, m), \tag{2.50}$$

where $\Xi_{i1}, \ldots, \Xi_{iq}$ are statistically independent of each other. According to the central limit theorem, the sum of the q statistically independent random variables with identical distributions is normally distributed for $q \rightarrow \infty$. Moreover, the mean and the variance of the sum of the random variables Ξ_{ik} is equal to the sum of the respective mean and variance, resulting in

$$\begin{aligned} \langle \Xi_i(dt, \mathbf{x}(t), t) \rangle &= q \cdot \langle \Xi_i(dt/q, \mathbf{x}(t), t) \rangle, & (2.51) \\ \mathrm{var}\{\Xi_i(dt, \mathbf{x}(t), t)\} &= q \cdot \mathrm{var}\{\Xi_i(dt/q, \mathbf{x}(t), t)\}. & (2.52) \end{aligned}$$

This can only be fulfilled if

$$\begin{aligned} \langle \Xi_i(dt, \mathbf{x}(t), t) \rangle &= A_i(\mathbf{x}(t), t)dt \ \text{ and} & (2.53) \\ \mathrm{var}\{\Xi_i(dt, \mathbf{x}(t), t)\} &= D_i(\mathbf{x}(t), t)dt. & (2.54) \end{aligned}$$

For the covariance, it can also be shown that

$$\mathrm{cov}\{\Xi_i(dt,\mathbf{x},t),\Xi_j(dt,\mathbf{x},t)\} = C_{ij}(\mathbf{x}(t),t)dt. \tag{2.55}$$

Note that A_i, C_{ij} and D_i are independent of dt. In summary, the four previously given conditions imply that the increments Ξ_i of the stochastic process are normal random variables with the above mean, variance and covariance. For the following, it is taken into account that a set of random variables $y_1,\ldots,y_m$, defined by

$$y_i = \alpha_i + \sum_{j=1}^{m} \beta_{ij} N_j, \tag{2.56}$$

where $N_1,\ldots,N_m$ are statistically independent unit normal random variables $\mathcal{N}(0,1)$, are normal random variables as well, which fulfil

$$\langle y_i\rangle = \alpha_i, \qquad \langle y_i^2\rangle - \langle y_i\rangle^2 = \sum_{j=1}^{m}\beta_{ij}^2 \qquad \text{and} \qquad \langle y_i y_j\rangle - \langle y_i\rangle\langle y_j\rangle = \sum_{k=1}^{m}\beta_{ik}\beta_{jk}.$$

As a consequence, the increments Ξ_i can be expressed in the form of Eq. (2.46) with the coefficients

$$A_i = a_i, \quad C_{ij} = \sum_{k=1}^{m} b_{ik}b_{jk}, \quad D_i = \sum_{k=1}^{m} b_{ik}^2. \tag{2.57}$$

Finally, Eq. (2.46) can be directly transformed into the *standard form multivariate Langevin equation*

$$x_i(t+dt) = x_i(t) + A_i(\mathbf{x}(t),t)dt + \sum_{j=1}^{m} b_{ij}(\mathbf{x}(t),t)N_j(t)(dt)^{\frac{1}{2}}. \tag{2.58}$$

The chemical Langevin Equation

The multivariate Langevin equations are now applied to a chemical reaction system, which is defined by non-discrete (macroscopic) state variables $(x_i(t),\ldots,x_m(t))$ for each of the m molecule species, corresponding to the particle numbers $\{X_i(t)\}$ introduced at the beginning of the chapter. In addition, the system is assumed to be driven by n chemical reactions R_j.

In order to find the correct coefficients A and b for Eq. (2.58), the propensity functions a_j, which have been introduced in Section 2.2.3, are applied [58]. These functions have been defined as the temporal probability densities for single molecular reaction events on the microscopic level. To derive an approximate solution for macroscopic systems, this concept is extended by introduction of an equivalent random variable $K_j(\mathbf{X},\tau)$, defined as the integer number of reaction events for reaction type R_j within a time interval of length τ given the state $(X_1,\ldots,X_m)$. Obviously, by using the stoichiometric coefficients from Section 2.1.1, the time evolution reads

$$X_i(t+\tau) = X_i(t) + \sum_{j=1}^{n} \nu_{ij} K_j(\mathbf{X}(t),\tau)\ , \quad (i=1,\ldots,m). \tag{2.59}$$

In macroscopic systems with 'well-behaved'[1] propensity functions a_j, it is possible to choose τ such that it is short enough to fulfil the requirement $a_j(\mathbf{X}(t')) \cong a_j(\mathbf{X}(t))$ for $t \leq t' \leq t + \tau$ on one hand, and long enough to expect many reactions within this time interval. Then, $\{K_j(\mathbf{X}(t), \tau)\}$ can be approximated by statistically independent *Poisson* random variables $\mathcal{P}_j(a_j(\mathbf{X}), \tau)$ with

$$\langle \mathcal{P}_j(a_j(\mathbf{X}(t)), \tau) \rangle = \mathrm{var}\{\mathcal{P}_j(a_j(\mathbf{X}(t)), \tau)\} = a_j(\mathbf{X}(t))\tau. \tag{2.60}$$

Moreover, $\mathcal{P}_j$ is approximated by normal random variables for sufficiently high values of $a_j(\mathbf{X}(t))\tau$ [58], leading to the relation

$$\mathcal{P}_j(a_j(\mathbf{X}(t)), \tau) \approx \mathcal{N}(a_j(\mathbf{X}(t))\tau, a_j(\mathbf{X}(t))\tau). \tag{2.61}$$

Replacing the discrete variables $(X_1, \ldots, X_m)$ by $(x_1, \ldots, x_m)$, specifying the non-discrete molecule concentrations[2], Eq. (2.59) results in

$$x_i(t + \tau) = x_i(t) + \sum_{j=1}^{n} \nu_{ij} \mathcal{N}_j(a_j(\mathbf{x}(t))\tau, a_j(\mathbf{x}(t))\tau), \quad (i = 1, \ldots, m). \tag{2.62}$$

With the linear combination theorem

$$\mathcal{N}(\mu, \sigma^2) = \mu + \sigma \mathcal{N}(0, 1), \tag{2.63}$$

it is possible to convert Eq. (2.62) into

$$x_i(t + \tau) = x_i(t) + \sum_{j=1}^{n} \nu_{ij} a_j(\mathbf{x}(t))\tau + \sum_{j=1}^{n} \nu_{ij} [a_j(\mathbf{x}(t))\tau]^{\frac{1}{2}} \mathcal{N}_j(0, 1), \quad (i = 1, \ldots, m). \tag{2.64}$$

On this basis, a special form of the multivariate Langevin equations can be directly derived. It is called the *chemical Langevin equation* and reads

$$x_i(t + dt) = x_i(t) + \sum_{j=1}^{n} \nu_{ij} a_j(\mathbf{x}(t))dt + \sum_{j=1}^{n} \nu_{ij} a_j^{\frac{1}{2}}(\mathbf{x}(t)) N_j(t)(dt)^{\frac{1}{2}}, \quad (i = 1, \ldots, m), \tag{2.65}$$

where $N_j(t)$ are n statistically independent and temporally uncorrelated unit random variables as introduced in the context of Eq. (2.46).

For numerical treatment, Eq. (2.65) can be transformed into a form using the *Wiener process* W [55]. It is then given by

$$dx_i(t) = \sum_{j=1}^{n} \nu_{ij} a_j(\mathbf{x}(t))dt + \sum_{j=1}^{n} \nu_{ij} a_j^{\frac{1}{2}}(\mathbf{x}(t)) dW_j, \quad (i = 1, \ldots, m), \tag{2.66}$$

[1] In typical reaction systems, a_j is dependent on non-negative powers of the molecule populations only.

[2] The discrete and non-discrete description of the system are considered equivalent for sufficiently high molecule populations. However, the propensity functions $a_j(\mathbf{X})$ have to be replaced by $a_j(\mathbf{x})$.

with the normal random variables $dW_j(t) = \mathcal{N}(0, dt)$. The numerical solution of this system of stochastic differential equations is addressed in the next chapter.

In conclusion, a deterministic and two stochastic methods for the simulation of chemical reaction processes were presented. The approximate stochastic method using Langevin equations was derived for macroscopic system sizes, whereas the stochastically exact Gillespie algorithm is appropriate for microscopic systems only. In the next chapter, a hybrid method will be developed, which combines the latter approaches.

Chapter 3

Hybrid Simulation Method for Multi-Timescale Markov Processes

Due to the enormous computation time required for exact stochastic simulations, I developed the 'General Stochastic Hybrid Method' (GSHM) [59], which combines the exact Gillespie algorithm with the approximate numerical solution of a system of stochastic differential equations (SDEs). The underlying algorithm, which is presented here for the first time, propagates a 'generalized system state' containing the deterministic terms of the chemical Langevin equation (see Section 2.2.4) and the probability densities required for the Gillespie method, thereby correctly considering all interactions. This method is particularly suitable for Markov processes running on multiple timescales. In this way, stochastically correct and reliable simulations of the large and complex signal transduction network investigated in this thesis are enabled.

3.1 The Timescale Problem

3.1.1 Deterministic versus Stochastic Methods

Three different numerical simulation approaches for chemical reaction systems were demonstrated in the previous chapter:

1. Deterministic Simulation Methods using Ordinary Differential Equations

The continuous molecule populations $(x_1, \ldots, x_m)$ in a system of m molecule species[1] and n reactions are supposed to evolve deterministically without fluctuations. The time evolution is specified by a system of ordinary differential equations (ODEs), defined by

$$\frac{dx_i}{dt} = \sum_{j=1}^{n} \nu_{ij} v_j(\mathbf{x}(t)), \tag{3.1}$$

[1]For simplicity and without loss of generality, the populations $(x_1, \ldots, x_m)$ are dimensionless and correspond to particle numbers per cell.

$$\begin{aligned} v_j &: \text{reaction rate of reaction } j,\ j = (1, \ldots, n), \\ \nu_{ij} &: \text{stoichiometric coefficients},\ i = (1, \ldots, m), \end{aligned}$$

which can be propagated starting from a set of initial concentrations.

2. Exact Stochastic Simulation Methods using Propensity Functions

A stochastically exact simulation method was proposed by Gillespie [36]. Here, the system is defined by discrete particle numbers $(X_1, \ldots, X_m)$. In a Monte-Carlo simulation [60], random numbers are mapped to the time point and the type of the next molecular reaction event according to the propensity functions $a_j(\mathbf{X})$, the temporal probability density for reaction R_j. The propensities $a_j(\mathbf{X})$ directly correspond to the reaction velocities $v_j(\mathbf{x})$ in the deterministic case.

3. Approximate Stochastic Simulation Methods using Stochastic Differential Equations

Approximate solutions of the time evolution of stochastic processes are based on the Langevin equations. The continuous system is driven by a set of stochastic differential equations, which read

$$dx_i(t) = \sum_{j=1}^{n} \nu_{ij} a_j(\mathbf{x}(t))dt + \sum_{j=1}^{n} \nu_{ij} a_j^{\frac{1}{2}}(\mathbf{x}(t))dW_j, \tag{3.2}$$

with a deterministic and a stochastic term (W_j: Wiener process). For high reaction rates, the second term is negligible and this method becomes equivalent with the deterministic method.

Simulation Methods and Timescales

The deterministic simulation methods are very efficient and return an exact and (numerically) differentiable solution after one simulation run. However, the deterministic assumption of a continuous and predictable system is not appropriate for systems containing molecule species of extremely low particle numbers. In this case, stochastic effects can influence the system's behaviour even qualitatively (see Section 3.4.2). On the other hand, stochastic methods require a high number of independent simulation runs in order to receive statistically reliable information. Moreover, even one simulation run using the exact Gillespie approach [36] requires enormous computation time since each single molecular reaction event is simulated. The third approach, which makes use of the Langevin equations, is a good approximation for large particle numbers and high reaction probabilities (see Section 2.2.4). In comparison with the deterministic approach, however, it is already applicable for a much lower range, since fluctuations are not neglected.

3.1.2 Subgroups of Reactions

Obviously, each of the three methods is adequate for a certain range of particle numbers and/or reaction rates, but none of them is correct and feasible at the same time if the reactions occur on different timescales, resulting in huge differences among the reaction probabilities. This is the typical situation in biochemical reaction networks, particularly in signal transduction systems, where reaction probability densities are separated by several orders of magnitude and where reactions of low probabilities are often crucial.

In this thesis, the timescale problem is approached by splitting the reaction system into three reaction subgroups, each of them corresponding to one of the simulation methods summarized in Section 3.1.1. Since in general, the reaction subgroups are not independent of each other, the interactions between them have to be carefully considered. In addition, the assignment of the reactions to the subgroups has to be organized dynamically, based on the constantly changing reaction probabilities. Furthermore, the system is partially described by discrete particle numbers. Consequently, a dynamic transition between continuous and discrete state variables has to be realized.

3.2 Stochastic-deterministic Hybrid Method

For simplicity, I first developed a stochastic-deterministic hybrid method, which is demonstrated in this section. Here, the Gillespie Algorithm is combined with a system of ODEs. I extended this concept by use of SDEs, leading to the 'General Stochastic Hybrid Algorithm', which is described in Section 3.3. The stochastic-deterministic approach does not account for stochastic fluctuations of the fast reaction subset and is therefore valid for certain reaction systems only. In contrast, the General Stochastic Hybrid Algorithm does not neglect these fluctuations. Instead, they are approximated by means of the Langevin equation.

3.2.1 Threshold Criteria

In this approach, stochastically relevant reactions are separated from those ones, which are assumed to be well-described by the deterministic method. For each system, rigorous quantitative criteria can be defined depending on the specific problem. Let $R = \{R_1, \ldots, R_n\}$ be the set of all reactions and $S_k \subseteq R$ the subset of all reactions to be treated stochastically exact according to criterion k. Two examples are given in the following. The first criterion, based on relative fluctuations, is defined as

$$S_1 = \left\{ R_j \;\middle|\; \exists i \in (1,\ldots,m) \;\middle|\; \frac{1}{\sqrt{dt}} \cdot \frac{\check{\sigma}_{ij}(dt)}{x_i} > \epsilon_1 \right\}, \quad \epsilon_1 > 0 \tag{3.3}$$

$$\text{with} \quad \check{\sigma}_{ij}(dt) = \sqrt{\text{var}(\check{x}_{ij}(t+dt))}.$$

Here, $\check{x}_{ij}(t+dt)$ denotes the particle number (or concentration, resp.) of molecule i at time $t+dt$, given state $\mathbf{x}$ at time t, under the assumption that from time t on, it is influenced by the reaction R_j only. Thus, a reaction is treated 'stochastically exact' as soon as it causes a *critical relative fluctuation* of the population of at least one species. For the threshold criterion, the relative standard deviation $\check{\sigma}_{ij}(dt)/x_i$ of this single-reaction process at time dt is chosen and divided by $\sqrt{dt}$, since the variance scales with dt according to the Langevin equation.

Obviously, (sub-)systems with extremely low particle numbers suggesting a discrete particle system, are not adequately addressed by this criterion, since the discrete nature causes increments, which might be much higher than the respective standard deviation. Therefore, a second criterion is introduced by

$$S_2 \;=\; \left\{ R_j \;\middle|\; \exists i \in (1,\ldots,m) \;\middle|\; \frac{\mid \nu_{ij} \mid}{X_i} > \epsilon_2 \right\}, \quad \epsilon_2 > 0. \tag{3.4}$$

Here, ν_{ij} denotes the stoichiometric coefficient of reaction R_j and X_i the discrete particle numbers. Their ratio describes the relative state change caused by a single reaction event, which is high for low particle numbers. Assuming no prior knowledge about the ranges of the particle numbers (concentrations), both criteria can be applied. It is not necessary to distinguish between subsets, for which the first or the second criterion has to be applied, since the second one becomes irrelevant for higher particle numbers by itself. Thus the stochastically exact subset of reactions can be defined as

$$S = S_1 \cup S_2.$$

Note that fast reactions do not necessarily influence molecule species with high populations only. However, high reaction rates usually indicate high particle numbers of educts and products.

3.2.2 Generalized System State

In a first step, I combined the *Next Reaction Method* [61], an extension of Gillespie's First Reaction Method demonstrated in Section 2.2.3, with the deterministic method. The discrete Next Reaction Method is applied to all reactions of the stochastic subset whereas the remaining reactions are described by a system of ODEs, which is propagated in parallel. Therefore, I introduced a 'generalized' state vector describing the system of m molecule species and n reactions

$$\begin{aligned} &(x_1,\ldots,x_m,p_1,\ldots,p_n),\\ &\qquad (s_1,\ldots,s_n), \quad s_j \in \{0,1\},\ s_j = 1 \Leftrightarrow R_j \in S,\\ &\qquad (r_1,\ldots,r_n), \quad 0 \le r_j < 1. \end{aligned}$$

Here, $\{x_i\}$ represent the molecule populations and p_j the probability that a discrete reaction event of type R_j, $R_j \in S$, has not occurred since its previous occurrence. The stochastic

flags $\{s_j\}$ are set for 'stochastic' reactions and $\{r_j\}$ contain random numbers, which are generated after each molecular reaction event of R_j.

3.2.3 Stochastic Reaction Subset

The state $\mathbf{x}$ is given at time t_0. The probability $p_j(t_0 + \Delta t)$ that reaction R_j does not occur within the time interval $[t_0, t_0 + \Delta t)$ can be derived for each stochastic reaction using the respective propensity functions a_j introduced in Section 2.2.3. Accordingly, it can be described by

$$p_j(t + dt) = p_j(t) - p_j(t) \cdot a_j(\mathbf{x}(t))dt, \tag{3.5}$$

which leads to the differential form

$$\frac{dp_j(t)}{dt} = -p_j(t)a_j(\mathbf{x}(t)). \tag{3.6}$$

In order to stochastically simulate the reaction events, the random numbers r_j are mapped to the putative time point $t_0 + \tau_j$ of the next reaction of type R_j in accordance with the 'probability of no reaction' $p_j(t)$. This is equivalent to Gillespie's Next Reaction Method [61] and the putative time points $t_0 + \tau_j$ are given by

$$p_j(t_0 + \tau_j) = r_j. \tag{3.7}$$

Purely Stochastic Systems

Let us first regard purely stochastic systems ($S \equiv R$). In this case, state $\mathbf{x}$ and the propensities $a_j(\mathbf{x}(t))$ remain constant until the next stochastic reaction is triggered. Therefore the time point of the next reaction is computed according to $t_{\text{next}} = t_0 + \min\{\tau_j\}$, based on the propensities $a_j(\mathbf{x}(t_0))$ at time t_0

$$a_j(\mathbf{x}(t)) \equiv a_j(\mathbf{x}(t_0)) \quad \text{for} \quad t_0 \leq t < t_{\text{next}}$$

$$\Rightarrow \quad p_j(t) = p_j(t_0) \cdot \mathrm{e}^{-a_j(\mathbf{x}(t_0))t} \quad \text{with} \quad p_j(t_0) = 1 \tag{3.8}$$

$$\Rightarrow \quad t_{\text{next}} = \min\left\{ -\frac{1}{a_j(\mathbf{x}(t_0))} \ln r_j \right\}. \tag{3.9}$$

Thus, for $S \equiv R$, the simulation algorithm is already given by Gillespie's Next Reaction Method.

Mixed Systems

Typically, only a subset of reactions is treated stochastically, leading to two important limitations:

1. In general, the propensities $a_j(\mathbf{x}(t))$ are also dependent on molecule concentrations influenced by reactions of the deterministic subset. Therefore, they cannot be considered constant between two stochastic events. In fact, they might even vary on the timescale of the 'fast' reaction subset and it might become necessary to steadily update them between the stochastic reaction events.

2. As a consequence, it is no longer possible to determine the real time point of the next stochastic reaction by knowledge of state $\mathbf{x}(t_0)$ only. Even the putative type of the next stochastic reaction R_j, given by the order of $\{\tau_j\}$, might change.

Thus, Eqs. (3.8) - (3.9) are not valid in general. Instead, Eqs. (3.6) - (3.7) have to be solved numerically, taking into account the time evolution of the deterministic part.

3.2.4 Deterministic Reaction Subset

Between the stochastic reaction events, the changes caused by the deterministic subset are determined by propagation of the ODE system of Eq. (3.1), reduced by the stochastic reactions subset S, resulting in

$$\frac{dx_i}{dt} = \sum_{j \notin S} \nu_{ij} v_j(\mathbf{x}(t)). \tag{3.10}$$

For the numerical propagation (see Section 2.1.3), the fourth-order Runge-Kutta method with an adaptive stepsize control was used here [42, 45].

3.2.5 Interaction between the Reaction Subsets

The time evolution of state $\mathbf{x}$ is governed by both reaction subsets and in general, the reaction rates of the deterministic subset are dependent on state variables also influenced by the stochastic subset and vice versa. A correct handling of the interactions between both subsets is critical and approached as follows:

1. Both parts of the generalized system state, namely the populations x_i and the probability functions p_j of the stochastic reactions, are propagated according to Eq. (3.6) and (3.10) using one ODE solver for the complete ODE system

$$\frac{dx_i(t)}{dt} = \sum_{j=1}^{n} (1 - s_j) \cdot \nu_{ij} v_j(\mathbf{x}(t)), \quad i = (1, \ldots, m), \tag{3.11}$$

$$\frac{dp_j(t)}{dt} = -s_j \cdot p_j(t) a_j(\mathbf{x}(t)), \quad j = (1, \ldots, n). \tag{3.12}$$

   As a result, the influence of the deterministic reaction subset on the reaction probabilities is entirely considered. The step size of the ODE solver is also controlled by the probabilities p_j, thus, the integration is numerically exact for both reaction subsets.

2. Since the propensities a_j are no longer constant, the time points for the next stochastic reaction events have to be determined *implicitly* according to

$$p_j(t_0 + \tau_j) = r_j, \tag{3.13}$$

which is efficiently realized by estimating the putative reaction time points $t_0 + \tau_j$ within the ODE solver algorithm, according to the logarithmic equation (3.9), which is now an approximation only. The solver is stopped as soon as the time point of the next stochastic event is hit within a given tolerance. This is achieved by iteration. Using the above approximation given by Eq. (3.13), it is usually not necessary to insert more than one iteration step.

3. Once the time point of the next stochastic reaction, called R_{j^*}, is reached, the reaction is executed by updating the populations $\{x_i\}$ according to the stoichiometry

$$x_i = x_i + \nu_{ij^*}, \quad i = (1, \ldots, m). \tag{3.14}$$

4. The continuous molecule populations $\{x_i\}$ are directly transformed into discrete particle numbers $\{X_i\}$ whenever required for the stochastic treatment.

3.2.6 Stochastic-Deterministic Hybrid Algorithm

The algorithm for the hybrid solver, which is given in the following, can be realized with most common ODE solvers, since only little modifications are required for the implicit determination of the time points at which the solver has to be interrupted for stochastic events.

1. Set the initial concentrations $(x_1, \ldots, x_m)$.
Set the probabilities of no reaction $(p_1, \ldots, p_n) = (1, \ldots, 1)$.
Generate the random numbers $(r_1, \ldots, r_n)$.
Set $t_0 = t_{\text{start}}$.

2. Determine the stochastic subset $S \subseteq R$ according to the upper (or another user-defined) criteria and set the stochastic flags $(s_1, \ldots, s_n)$.

3. Start ODE solver with the following task: Propagate the generalized system

$$(x_1, \ldots, x_m, p_1, \ldots, p_n),$$

according to Eqs. (3.11)-(3.12); start at t_0 and stop at t_{stop}.

Determine t_{stop} implicitly by steadily[2] updating τ_j for $R_j \in S$, according to Eq. (3.9)

$$\begin{aligned} t_{\text{stop}} &= t_0 + \tau_{j^*}, \\ \tau_{j^*} &= \min_{j \in S}\{\tau_j\}, \ R_{j^*} : \text{type of the next putative reaction event.} \end{aligned}$$

[2]The update is usually performed after one of many time steps only, depending on to the time difference to the next putative stochastic reaction, the step size and the timescale, on which the propensity functions change.

Stop the ODE solver as soon as $|t - t_{\text{stop}}| \leq tol$ or repeat the last time step if $t_{\text{stop}} - t > tol$ with an appropriate step size (overshoot mode).

4. Execute reaction R_{j^*} by updating $\{x_i\}$ according to Eq. (3.14).
5. Generate new random number r_{j^*} and set $p_{j^*} = 1$.
6. Set $t_0 = t$ and go to step 3, or to step 2, if an update of $\{s_j\}$ is necessary[3].

This algorithm is highly efficient whenever there are subsets with extremely high and extremely low reaction rates and/or particle numbers within one system. However, for systems whose reactions are also distributed over the timescales in-between, there are molecule species with high particle numbers and high relative fluctuations. Thus, either a stochastic simulation has to be performed for those reactions, again resulting in extremely high computation times, or the respective fluctuations are neglected.

However, it is not correct to treat the stochastic behaviour in an exact way below a certain threshold and to completely and abruptly ignore stochastic fluctuations above this threshold, although the absolute fluctuations become even higher for reactions with increased reaction rates or propensity functions. A solution to this problem is given in the next section.

3.3 General Stochastic Hybrid Method

In this approach, stochastic differential equations are introduced to approximate relevant stochastic effects of reactions, which have been described by ODEs so far. Therefore, an additional reaction subset $L \subseteq R$ is introduced for those reactions, which influence molecule species with particle numbers in a higher range and which cause fluctuations that are not negligible. It is assumed that the stochastic behaviour of this subset is well approximated by Langevin equations and that an exact simulation on molecular level is therefore not required (see Section 2.2.4). The deterministic approach can be considered a special case of the Langevin approach for extremely fast reactions whose relative fluctuations (corresponding to the stochastic term of the Langevin equations) become negligible. For simplicity, the respective reactions are therefore not treated separately from the reactions of the Langevin subset and consequently, $L = R \setminus S$.

3.3.1 Numerical Solution of the Langevin Equation

In order to numerically solve the Langevin equation

$$dx_i(t) = \sum_{j=1}^{n} \nu_{ij} a_j(\mathbf{x}(t))dt + \sum_{j=1}^{n} \nu_{ij} \sqrt{a_j(\mathbf{x}(t))}\, dW_j, \tag{3.15}$$

[3]Usually, the assignment of reactions to the stochastic or deterministic subset changes on a much slower timescale and has to be rechecked after many executions of step 3 only.

$$dW_j(t) \;=\; \mathcal{N}(0, dt), \qquad \mathcal{N}(\mu,\, \sigma^2)\text{: normal random variable,}$$

a Wiener process is generated with finite increments $\Delta W = \mathcal{N}(0, \Delta t)$ as suggested in the Euler-Maruyama approximation [55]. Different from the Euler-Maruyama method, the time evolution is split into a deterministic and a stochastic part: the deterministic term $d\hat{x}_i(t)$ and the deterministic part of the stochastic term, denoted by $d\tilde{x}_i(t)$, are propagated like ODEs. In parallel, Wiener increments are generated, leading to

$$dx_i(t) = d\hat{x}_i(t) + d\tilde{x}_i(t) \cdot n_i, \quad n_i = \mathcal{N}(0{,}1). \tag{3.16}$$

This is realized in a way that enables the use of most ODE solvers and that is therefore not restricted to the explicit Euler method:

- The step size of the ODE solver is limited to the step size Δt introduced for generation of the Wiener process. However, smaller sizes, proposed by an adaptive step size control, are allowed, depending on both terms of Eq. (3.15). Thereby, I introduced a 'relaxed Wiener process' and departed from the exact definition of a Wiener process as a trade-off with computational feasibility.

- The solver propagates the terms

$$d\hat{x}_i(t) \;=\; \sum_{j=1}^{n} \nu_{ij} a_j(\mathbf{x}(t)) dt, \tag{3.17}$$

$$d\tilde{x}_i(t) \;=\; \sum_{j=1}^{n} \nu_{ij} \sqrt{a_j(\mathbf{x}(t)) \cdot dt}\,. \tag{3.18}$$

  The right hand side of Eq. (3.18) multiplied by a unit normal random variable $n_i = \mathcal{N}(0, 1)$ corresponds to an increment of the Wiener process for each i. In this way, prior knowledge about the step size is not required for determination of the Wiener-increments. The step size is therefore adaptable on the basis of the deterministic terms and the deterministic part of the stochastic terms.

- After each completed time step, both terms are added according to Eq. (3.16) and new random numbers $\{n_i\}$ are generated.

3.3.2 General Stochastic Hybrid Algorithm

The General Stochastic Hybrid Algorithm efficiently combining Gillespie's stochastic simulation method with the approximate stochastic Langevin approach is demonstrated here. The set of reactions $R = \{R_j\}$ is divided into the respective subsets and the flags $(s_1, \ldots, s_n)$ are set for all reactions $R_j \in S$ (S denotes again the Gillespie subset).
The relative fluctuations defined by Eq. (3.3) and Eq. (3.4) in the previous section can again be chosen as threshold criterion between the exact and the approximate stochastic subset using an appropriate ϵ_1 and ϵ_2.

The algorithm for the GSHM, directly derived from the stochastic-deterministic hybrid solver is given here:

1. Set the initial concentrations $(x_1, \ldots, x_m)$.
   Set the probabilities of no reaction $(p_1, \ldots, p_n) = (1, \ldots, 1)$.
   Generate the random numbers $(r_1, \ldots, r_n)$.
   Generate the unit normal random variables $(n_1, \ldots, n_m)$.
   Set $t_0 = t_{\text{start}}$.

2. Determine the stochastically exact subset $S \subseteq R$, according to the upper (or another user-defined) criteria and set the flags $(s_1, \ldots, s_n)$.

3. Start ODE solver with the following task: Propagate the generalized system

$$(x_1, \ldots, x_m, p_1, \ldots, p_n),$$

   according to the equations

$$\begin{aligned}
\frac{dx_i(t)}{dt} &= \sum_{j=1}^{n}(1 - s_j) \cdot \nu_{ij} v_j(\mathbf{x}(t)), \quad i = (1, \ldots, m) \\
\frac{dp_j(t)}{dt} &= -s_j \cdot p_j(t) a_j(\mathbf{x}(t)), \quad j = (1, \ldots, n), \\
d\tilde{x}_i(t) &= \sum_{j=1}^{n}(1 - s_j) \cdot \nu_{ij} \sqrt{a_j(\mathbf{x}(t)) \cdot dt}\ , \quad i = (1, \ldots, m), \quad \text{see Eq. (3.18).}
\end{aligned}$$

   Start at t_0 and stop at t_{stop} with a maximum time step size Δt.

   Add the stochastic Langevin term after each solver step:

   - $x_i = x_i + \Delta\tilde{x}_i \cdot n_i,\ i = (1, \ldots, m)$,

     $\Delta\tilde{x}_i \cdot n_i$ denotes the finite increment for solver step size $\Delta\tilde{t}$
     $(dt \to \Delta\tilde{t} :\ d\tilde{x}_i \to \Delta\tilde{x}_i)$.
   - Generate new set of unit normal random variables $(n_1, \ldots, n_m)$.

   Determine t_{stop} implicitly by steadily[1] updating τ_j for $R_j \in S$, according to Eq. (3.9)

$$\begin{aligned}
t_{\text{stop}} &= t_0 + \tau_{j^*}, \\
\tau_{j^*} &= \min_{j \in S}\{\tau_j\},\ R_{j^*} : \text{type of the next putative reaction event.}
\end{aligned}$$

   Stop the ODE solver as soon as $|t - t_{\text{stop}}| \leq tol$ or repeat the last time step if $t_{\text{stop}} - t > tol$ with an appropriate step size (overshoot mode).

[1] The update is usually performed after one of many time steps only, depending on to the time difference to the next putative stochastic reaction, the step size and the timescale, on which the propensity functions change.

4. Execute Reaction R_{j^*} by updating $\{x_i\}$ according to Eq. (3.14).
5. Generate new random number r_{j^*} and set $p_{j^*} = 1$.
6. Set $t_0 = t$ and go to step 3 or to step 2, if an update of $\{s_j\}$ is necessary[2].

Generation of Standard Gaussian Random Variables

For the generation of standard Gaussian random variables, the *Box-Muller Method* is applied [55], which maps two independent, uniformly distributed random variables r_1, r_2 between 0 and 1 to the two independent random variables

$$n_1 = \sqrt{-2\ln(r_1)}\cos(2\pi r_2), \tag{3.19}$$

$$n_2 = \sqrt{-2\ln(r_1)}\sin(2\pi r_2), \tag{3.20}$$

which are then normally distributed. The generation of random numbers is addressed in Section 8.2.

3.4 Results

3.4.1 Comparison between Hybrid and Gillespie Algorithm

As a proof of principle, the accuracy and the computational efficiency of the hybrid method I developed was tested by comparing the hybrid and the exact algorithm using the small reaction system applied by Gillespie in [36] (see Reaction System 29) to demonstrate his algorithm. The system is given by

$$\begin{aligned} \bar{X} + Y &\xrightarrow{k_1} 2Y, \\ 2Y &\xrightarrow{k_2} Z. \end{aligned} \tag{3.21}$$

The population of X is assumed to remain constant. Single simulation runs as shown in Fig. 3.1 (left) were repeated 5000 times. The average time evolution of the molecule populations and its standard deviation were computed for the complete time interval (Fig. 3.1, right). Obviously, the hybrid method provides an almost perfect approximation of the exact result. The computation time required for this method was significantly lower, as shown in the following table.

Method	computation time	relative computation time
Gillespie Algorithm	54350 sec	100 %
Hybrid Algorithm	1477 sec	2.7 %

The times refer to 5000 stochastic simulation runs on an Intel Pentium III, 1 GHz.

[2]Usually, the assignment of reactions to the stochastic or deterministic subset changes on a much slower timescale and has to be rechecked after many executions of step 3 only.

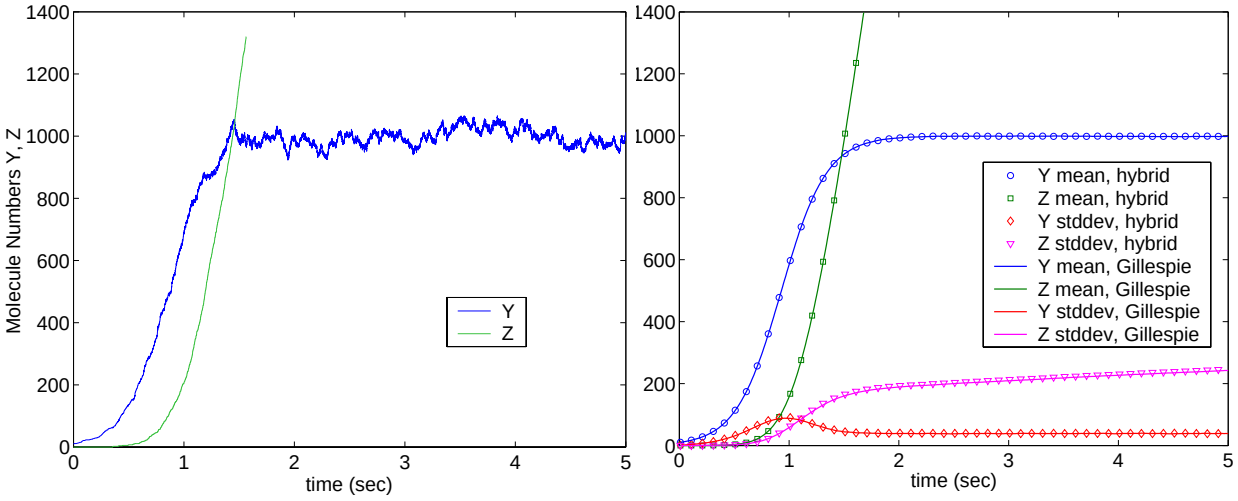

Figure 3.1: Simulation for Reaction System (3.21). Left: Single stochastic simulation run. Right: Mean and standard deviation based on 5000 simulation runs using the Gillespie algorithm (lines) and the hybrid method (dots). Start values: $Y = 10$, $Z = 0$, $k_1 X = 5$, $k_2 = 0.005$, threshold: 200 particles (fixed).

3.4.2 Exemplary Application: Hybrid Algorithm applied to Programmed Cell Death

As demonstrated in the previous section, the Gillespie algorithm is computationally limited even in case of small reaction systems. To simulate the reaction network of the signal transduction application of this thesis (Section 5.2), the GSHM was applied. The results were compared with the Gillespie, the pure Langevin, the pure deterministic and the deterministic-stochastic hybrid method. As expected, the computation time for a sufficiently high number of independent Gillespie simulation runs was enormous, whereas the other alternatives led to significantly different (and wrong) results.

Stochastic Simulation of Programmed Cell Death

As an exemplary application, the reaction system of CD95-induced apoptosis (programmed cell death), which is explained in Chapter 4, was chosen. This system exhibits high sensitivity with respect to variations of some crucial molecule species, called executioner caspases, since a low particle number of these molecules is sufficient to trigger a positive feedback loop, which then activates the execution of apoptosis. In this thesis, activation scenarios with different ligand concentrations corresponding to different intensities of activation signals of the pathway are of main interest. As demonstrated in Chapter 9, high ligand concentrations always result in apoptosis, whereas extremely low concentrations do not cause any cell death. However, the area in-between is likely to be dominated by stochastic effects since the particle numbers of some critical molecule species are extremely low (often

between 0 and 100 per cell). These species are interacting with an environment of much higher molecule concentrations, suggesting the use of the hybrid method.

The simulation results for a weak activation signal (low CD95-ligand concentration) are given in Fig. 3.2, where the concentrations of some key molecules of the pathway, which trigger and indicate execution of the cell death process, are shown. The exact stochastic simulations show a relatively broad period of caspase activity and an incomplete PARP cleavage (Fig. 3.2 A). From single simulation runs, however, it is known that in the model, either the caspase activation loop is triggered, resulting in a fast and complete cleavage of PARP and active caspases (e.g. Fig. 3.2 C), or the execution of cell death does not take place at all. Thus, the incomplete average PARP cleavage and the great variability of the reactant concentrations among different simulation runs is governed by stochastic effects, triggering the execution of cell death with a certain probability only. These effects also cause the high variability of the time points at which the activation is triggered. The set of independent stochastic simulation runs corresponds to an ensemble of independent cells[3]. The stochastically simulated behaviour could also be observed in experiments that refer to a population of many cells (see Section 9.4). The complete results of the simulation of cell death rates are shown in Section 9.4.

3.4.3 Comparison with other Methods

The results of the GSHM were compared with the pure Gillespie, the Langevin, the deterministic and the stochastic-deterministic hybrid algorithm. Apart from the deterministic and the Gillespie case, 1000 simulation runs were performed with independent sets of random numbers. On this basis, the mean time evolution and the respective standard deviation were computed for the molecule concentrations (see Fig. 3.2). The results of the stochastic hybrid method and the exact Gillespie algorithm converge for a sufficient number of simulation runs, however, the computation time for the GSHM is significantly lower:

Method	computation time	relative computation time
Gillespie Algorithm	1904 sec	100 %
Hybrid Algorithm	27.6 sec	1.45 %

Average time for a single stochastic simulation run. Intel Pentium III, 1 GHz.

The deterministic simulations do not incorporate the variability caused by stochastic effects and do not even provide the correct averaged behaviour. In case of the pure Langevin simulation, the increase of active caspase concentration starts much earlier than in the deterministic case. This is caused by fluctuations of the critical molecules, which trigger the positive feedback loop. However, due to the fact that the Langevin method does not account for the discrete nature of molecules, this is not an adequate approach, since the

[3]It is assumed that interactions between the cells are not affecting the considered pathway once the scenario is started at $t = 0$.

regarded process is driven by reactions between molecules of extremely low particle numbers at the beginning. As a consequence, the Langevin method results in a bias towards earlier activation. According to both the exact and the stochastic hybrid method, cell death can be executed in earlier or later stages with a certain probability or does not take place at all, resulting in incomplete average PARP cleavage (Fig. 3.2 A). The results of the stochastic-deterministc algorithm are much closer to the exact result, however, the PARP cleavage and the death rate show differences as well (Fig. 3.2 D).

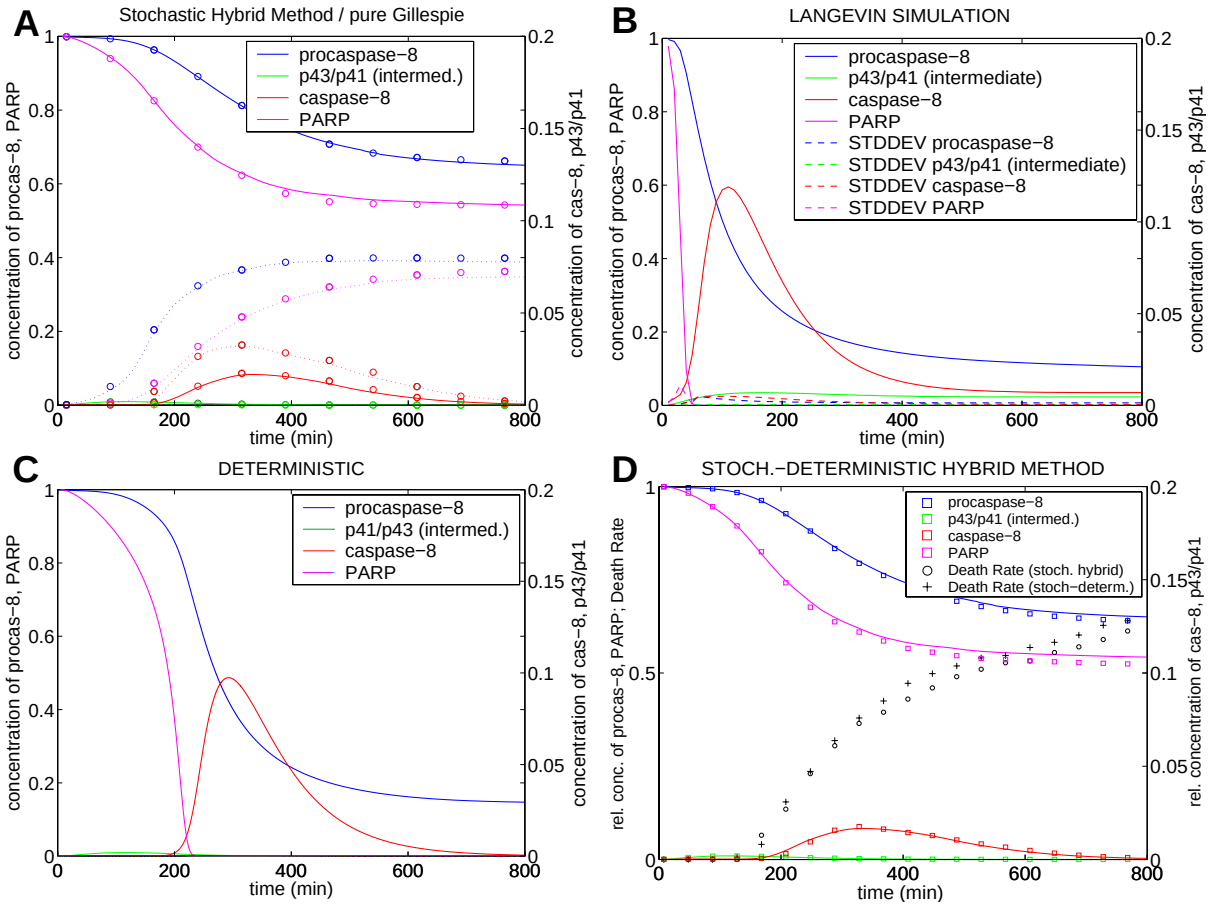

Figure 3.2: The simulated execution of programmed cell death (initial CD95-ligand concentration: 7 ng), indicated by the decrease of PARP upon conversion of procaspase-8 into active caspase-8 (see Chapter 5), shows that none of the alternative methods was capable of correctly simulating the diversity of stochastic effects. Even the Langevin approach predicts a different activation course by neglecting the discreteness of the molecular events, whereas the stochastic hybrid method led to the same results as the exact Gillespie simulation (circles in A, dotted lines: standard deviation), thereby saving $> 98\%$ of computation time. The concentrations are given in relative units referring to the initial concentrations. All simulation runs are started with the same set of parameters.

3.5 Conclusion

In summary, the General Stochastic Hybrid Method provides an algorithm, which is stochastically exact in the sensitive low-particle number range, but still efficient for the fast reaction subset, where the Gillespie approach would no longer be feasible. The stochastic effects are well approximated for fast reactions and/or high particle numbers by the use of SDEs. In general, it does not make sense to treat the stochastic behaviour in an exact way below a certain threshold and to abruptly ignore the even higher stochastic fluctuations above, as realized in the stochastic-deterministic method or proposed in the study of Kiehl et al. [67], which was performed at the same time. Furthermore, a solver with adaptive step size control was chosen, influenced by the deterministic term and the deterministic part of the stochastic term of the Langevin equation, as well as by the 'probabilities of no reaction' of the stochastically exact part. This was realized by propagating one equation system, which integrates all three sets of equations, for which I introduced the 'generalized' state vector. As a consequence, all interactions between the reaction subsets are automatically considered. Therefore, the exactness of the algorithm is only limited by the exactness of the numerical solution of the ODEs and by the Wiener process. Rigorous criteria were defined to assign the reactions to the subsets in a standardized and dynamic way. As a result, transitions of reactions from one subset to another are possible during the simulation run.

The CD95-induced cell death signalling system was simulated for activation scenarios close to the threshold between life and death. Here, the experimentally observed death rate of an ensemble of many cells could be explained by stochastic effects (see also Section 9.4). The hybrid method was more than 50 times faster than the exact algorithm, whereas the other simulation methods did not provide correct results.

To my knowledge, the combination of a stochastically exact and approximate algorithm is presented here for the first time. Since the stochastic methods used here, are directly derived from the master equation, the hybrid approach can be applied to any other type of Markov processes as well.

Chapter 4

The Signal Transduction Network of Programmed Cell Death

In this chapter, an introduction into the signal transduction network of *Programmed Cell Death* is given. Hereby, the pathway of *CD95-induced apoptosis*, the signalling system investigated in this thesis, will be explained in detail. In addition, I will present the network topology, derived from databases and literature.

Programmed Cell Death, called *apoptosis*, is the natural and controlled death of cells, in which the cell and its nucleus shrinks, condenses and fragments [3, 68], as opposed to *necrosis*, where the cells die due to acute injury. Apoptosis is one of the most complex signalling pathways and an essential property of all higher organisms. Defects in apoptosis result in a number of serious diseases such as cancer, autoimmunity and neurodegeneration [38, 39]. To develop efficient therapies, fundamental questions about molecular mechanisms and regulation of apoptosis remain to be answered. Apoptosis is triggered by a number of factors, including UV-light, γ-radiation, chemotherapeutic drugs, growth factor withdrawal ('death by neglect') and signalling from the death receptors [69, 70]. Apoptosis pathways can generally be divided into signalling via the death receptors at the membrane (extrinsic pathway) or the mitochondria [3] (intrinsic pathway). Both pathways imply *caspases* as effector molecules [71]. Caspases belong to the family of proteases, which are enzymes that degrade proteins by hydrolyzing some of their peptide bonds [72]. Caspases mostly exist in their inactive proforms (procaspases) and become active after getting cleaved. Various caspases are involved in both the initiation of the apoptotic process and the execution of the final apoptotic program.

4.1 CD95-induced Apoptosis

CD95-induced apoptosis is one of the best-studied apoptosis pathways. A detailed overview on this mechanism is for example given in [38] and [73]. CD95 is a member of the death receptor family, a subfamily of the TNF-receptor superfamily (e.g. [74]). Crosslinking of

the CD95-receptor either with its natural ligand CD95L or with agonistic[1] antibodies such as anti-APO-1 induces apoptosis in sensitive cells.

DISC, Caspases and Execution of Cell Death

Upon CD95 stimulation with CD95L or anti-APO-1, the Death-Inducing Signalling Complex (DISC) is formed (see Fig. 4.1). The DISC consists of oligomerized CD95, the death domain-containing adaptor molecule FADD, procaspase-8, procaspase-10 and c-FLIP. The interactions between the molecules in the DISC are based on homophilic contacts. The death domain (DD) of CD95 interacts with the DD of FADD, while the death effector domain (DED) of FADD interacts with the DED of procaspase-8. Once the DISC is formed, procaspase-8 can be autocatalytically cleaved: two procaspase-8 molecules bound at the DISC form the intermediate product p43/p41, followed by generation of an active caspase-8 complex p18/p10 (see Fig. 5.3) [75]. This process can be inhibited by c-FLIP, which binds to the DISC in various ways and blocks the latter mechanism [76].

After formation of active caspase-8, the apoptotic signalling cascade starts. A simplified network topology of the CD95-induced apoptosis pathway is given in Fig. 4.2. Caspase-8 cleaves and activates caspase-3 and -7; caspase-3 itself activates caspase-6, which again activates caspase-8, thereby establishing a self-amplifying activation loop. Caspase-3, -6, and -7 are involved in the execution of the death process, e.g. the chromosomal degradation of DNA [77] and, therefore, called *executioner caspases*, whereas the others, responsible for transferring the death signal, are referred to as *initiator caspases*. The DNA degradation plays an important role in the cell death process. It is started after ICAD gets cut off the CAD-ICAD complex by caspase-3 and -7, thereby terminating the inhibition of CAD, which directly fragments the DNA [78]. In parallel, PARP, a molecule, which repairs broken DNA strands is cleaved by executioner caspases as well. Once the DNA fragmentation process is triggered, a complete degradation of the cell starts irreversibly.

Type I versus Type II Cells and the Mitochondria

Two different CD95-signalling pathways are established in different cell types [79, 80]. Type I cells are characterized by intensive DISC formation and mitochondria independent caspase-3 activation. In Type II cells, the formation of the DISC complex is reduced and the activation of caspase-3 occurs downstream of the mitochondria: the active form of caspase-8 cleaves Bid, followed by translocation of the cleavage product tBid to mitochondria, which results in the release of cytochrome-C [81]. Subsequently, apoptosome[2] formation [82] takes place, leading to the activation of caspase-9, which then activates caspase-3, triggering the subsequent apoptotic events (see Fig. 4.1). Here, a feedback loop is established by caspase-2: it is activated by caspase-3 downstream of mitochondria and it cleaves Bid, which in response leads to mitochondrial cytochrome-C release (Fig. 4.2).

[1] An agonist is a substance that binds to a receptor and triggers a response by the cell.

[2] The apoptosome is a complex consisting of Apaf-1, cytochrome-C and caspase-9.

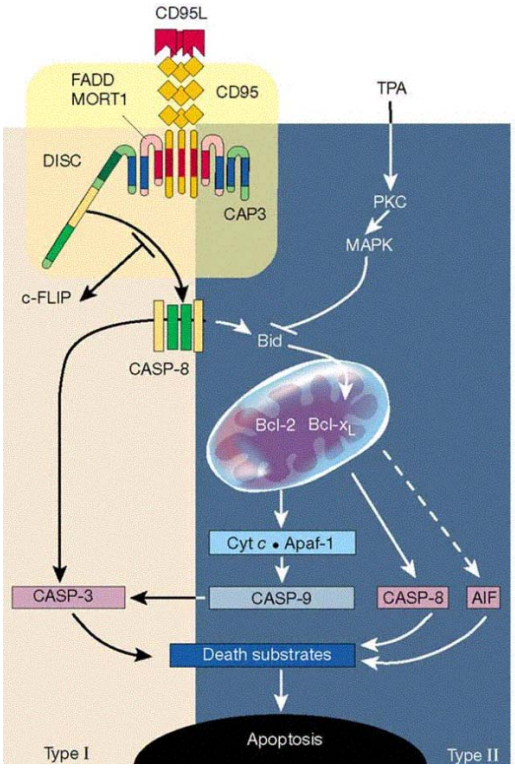

Figure 4.1: Simplified signalling pathway of CD95-induced apoptosis. The CD95-receptors (CD95) are activated at the membrane by CD95-ligands (CD95L), followed by DISC formation (yellow box) and processing of procaspase-8, resulting in the generation of active caspase-8 complexes. These complexes trigger the death process by activation of executioner caspases like caspase-3 directly (Type I cells), or via the mitochondria and caspase-9 (Type II cells). (Figure taken from P. Krammer [38].)

Regulation of CD95-induced Apoptosis

Furthermore, CD95-induced signalling is influenced and regulated by many other molecules, which mostly inhibit or amplify the apoptotic process, like XIAP and IAP1/2 (inhibitors of caspase-3,-7,-9)[83] or the BCL-2 family [84] consisting of pro-apoptotic (e.g. Bak, Bax) and anti-apoptotic (e.g. BCL-XL) members, regulating the critical cytochrome-C release. Even though gene regulation usually occurs on a slower time scale than the activation of caspases upon CD95 stimulation, it might also play an important role by influencing some of the most important inhibitors (e.g. expression of c-FLIP and IAP1/2 by NF-κ-B) [85, 86]. An overview about the molecule families, which play an important role in this pathway, is given in [87].

4.2 Network Topology

Despite the ever-increasing number of studies on CD95-induced apoptosis, a systemic understanding of this complex signalling pathway is still missing. Therefore, the network topology of CD95-induced apoptosis was established by critically searching databases [88, 89] and the literature. Molecules and reactions directly or indirectly interacting with the main components of this pathway were incorporated, leading to a network topology with more than 60 molecules and about 100 interactions (Fig. 4.2). This complexity cannot be matched by experimental data at present. Moreover, a high percentage of these interactions are known on the semantic level only (e.g. 'A activates B') and even for those interactions, whose biochemical mechanisms are well understood, quantitative information is mostly missing. In the next chapter, a modelling approach will be demonstrated as a basis for quantitative system analysis.

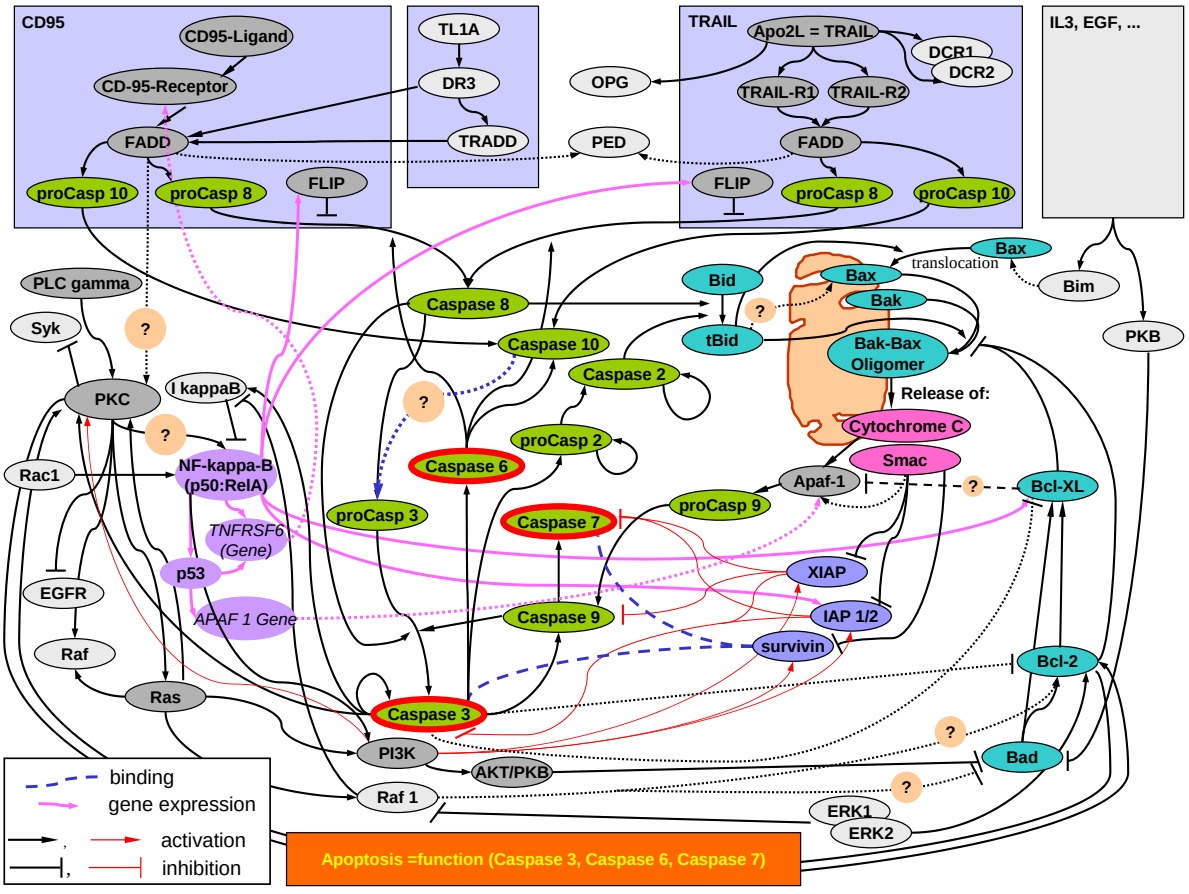

Figure 4.2: Simplified network topology of the CD95-induced apoptosis signalling pathway. The death process is executed in dependence of the activity of the executioner caspases. Note that some of the interactions are not proven yet (question marks). The model also contains the TRAIL- and EGF-receptor [90, 91].

Chapter 5

Modelling Signal Transduction

In the following, a quantitative modelling approach for signal transduction networks will be demonstrated. To overcome the problem of heterogeneous information about the cellular processes, I will introduce *Structured Information Models*. On this basis, a model of CD95-induced apoptosis will be derived. To my knowledge, this is the first data-based model of this pathway covering the complete process from activation to cell degradation. The model will also provide the basis for system identification based on experimental data.

5.1 The Information Problem

Signal transduction networks are described on different levels of information quality (Fig. 5.1). In most cases, interactions between two molecules are known on the semantic level only (e.g. A inhibits B or A activates B), thereby providing a network topology. For some well-investigated molecules and interactions, the biochemical mechanism is also known (e.g. enzymatic process, formation of complexes, ...), which is required for quantitative modelling. However, information about the underlying biochemical parameters like reaction rate constants, Michaelis-Menten constants or dissociation rates, is mostly completely missing. Even if quantitative information is available, its usability is limited if it refers to different experimental settings, cell types or states of cells, since these quantities are mostly influenced by many factors.

5.1.1 Structured Information Models

To reduce the complexity of the model without sacrificing essential components of the network, subsystems of different information qualities were identified and incorporated: subsystems mainly consisting of interactions with well-understood biochemical mechanisms are modelled as chemical reaction systems, whereas all others are modelled as 'black boxes', defined by their experimentally observed input-output behaviour. The subsystems are identified according to the following criteria:

- The input/output behaviour should be measurable.

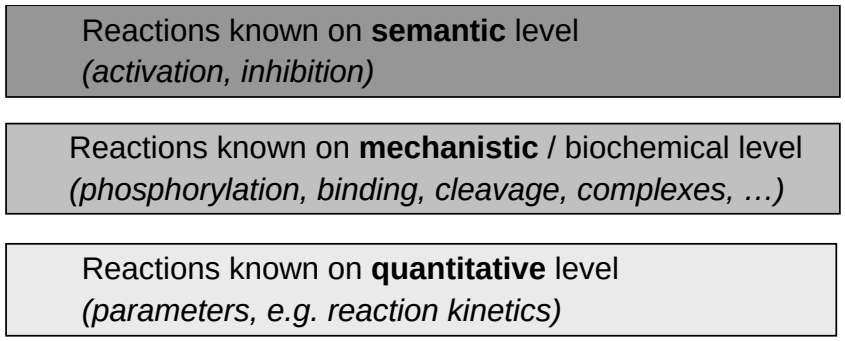

Figure 5.1: Information about biochemical networks is typically spread over three levels of quality.

- The number of input/output variables should be low.
- Subsystems should represent real functional systems (e.g. mitochondria).
- The information within one subsystem should be on the same level.

As an example, the decomposition of the signalling system of Fig. 4.2 is shown in Fig. 5.2. Notably, the black boxes do not assume knowledge of the exact underlying mechanisms. Instead, they reproduce the behaviour of the respective subsystems in a simplified way, e.g. by introduction of B-Splines [92] interpolating between the observed input-output behaviour of black boxes. Moreover, minimum sets of state variables and 'effective' parameters are introduced that do not necessarily correspond to molecule concentrations and biochemical parameters. As a consequence, the number of unknown parameters can be drastically reduced a priori. Note that this concept is motivated by a typical situation in system identification of biochemical networks: due to missing information and limited experimental data, the models should be small and restricted to those network parts, which are well understood. On the other hand, parts of biochemical networks can in general not be regarded independently of their environment. Thus, black boxes are introduced to reproduce the relevant effects of the surrounding network parts on a mechanistically well-understood subsystem, rather than for system identification of the surrounding network itself, for which the data basis would be missing anyway.

The degradation process of CD95-induced apoptosis (see Section 4.1) is, for example, not in the main focus of the signalling pathway investigated in this thesis, however, experimental data cannot be matched if it is ignored. On the other hand, the mechanism of the degradation process is not completely known and would involve many additional molecules, interactions and unknown parameters. Therefore, the degradation is modelled as a decay function depending on a virtual state variable describing the 'apoptotic activity', which is influenced by executioner caspases. Thus, this function approximates the experimental observations, thereby requiring a few parameters only.

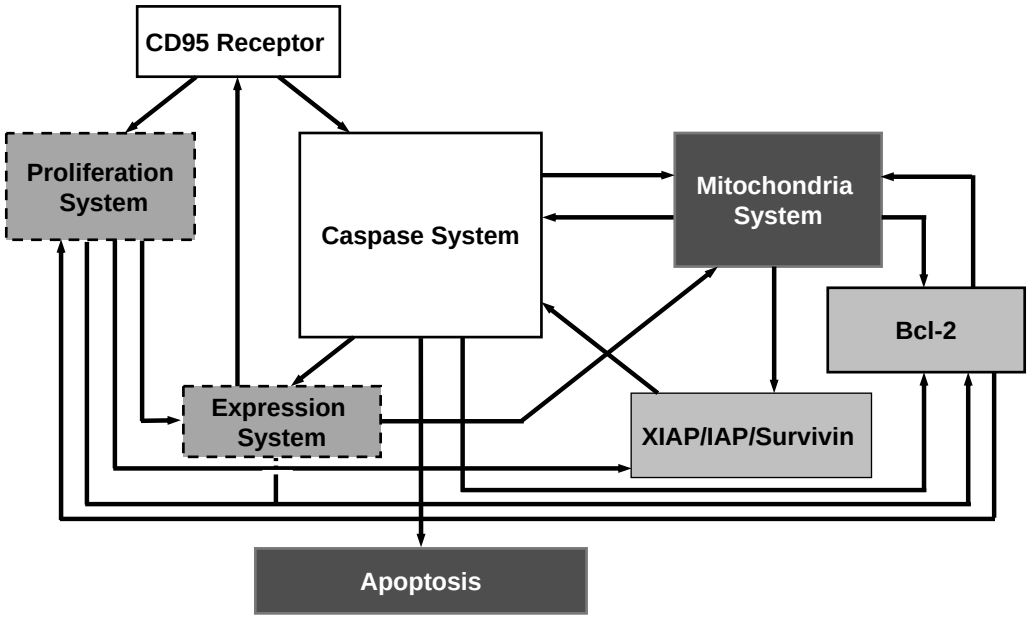

Figure 5.2: Principle of Structured Information Models based on the network topology of Fig. 4.2. The grey scale level of the boxes corresponds to the information quality of the subsystems. Information on a mechanistic level is available for the well-investigated CD95-receptor and for the interactions between the caspases (white boxes). It is partially available for the IAP/XIAP/survivin-family and the Bcl-2-family [87], which will be modelled in a simplified way (Fig. 5.3). The mitochondria and the final death process (upon activation by executioner caspases) are modelled as black boxes. The gene expression and the proliferation [3] subsystems involved in this signalling system could also be modelled as black boxes, however, they were taken out scope since their impact during the activation scenarios investigated in this thesis seems to be low. Interactions between the subsystems are indicated by arrows.

The decomposition of the complete system into subsystems is an iterative and adaptive process. Based on new information, a subsystem might be split into further subsystems. A great advantage of the so-obtained 'Structured Information Models' is that it combines heterogeneous information in one model instead of dealing with isolated models.

5.1.2 Combined Model Definition

For the mathematical description of the mechanistic part of structured information models, interactions are modelled based on reaction rate equations (see Chapter 2). The state of this system part is described by the concentration of l relevant signal transduction molecules $(x_1, \ldots, x_l)$. The reaction rates depend on these concentrations and also on biochemical parameters $(\Phi_1, \ldots, \Phi_r)$ like binding constants. To describe the temporal behavior, a system of ODEs is automatically generated as linear combinations of the reaction rates

according to Eq. 2.5. Note that the initial concentrations $x_i(t=0)$ are often unknown. In this case, they are considered unknown parameters.

Description of Black Boxes

Black boxes are defined by their experimentally observed input-output behavior. Additional state variables $(x_{l+1}, \ldots, x_m)$ that do not necessarily correspond to molecule concentrations and additional parameters $(\Phi_{r+1}, \ldots, \Phi_s)$ can be introduced. The q-th black box is represented by the function $\mathbf{f}^q(x_1, \ldots, x_l, x_{l+1}, \ldots, x_m, \Phi_{r+1}, \ldots, \Phi_s, t)$ that describes the changes of molecule concentrations and other state variables it affects. Boundary conditions like conservation laws have to be taken into account. Thus, the combination of subsystems modelled as chemical reaction systems and black boxes leads an ODE system, which reads

$$\begin{aligned}
\frac{dx_i}{dt} &= \sum_{j=1}^{n} \nu_{ij} v_j(x_1, \ldots, x_l, \Phi_1, \ldots, \Phi_r) + \sum_{q=1}^{N} f_i^q(\mathbf{x}, \Phi_{r+1}, \ldots, \Phi_s, t), \quad i = (1, \ldots, l), \\
\frac{dx_i}{dt} &= \sum_{q=1}^{N} f_i^q(\mathbf{x}, \Phi_{r+1}, \ldots, \Phi_s, t), \quad i = (l+1, \ldots, m),
\end{aligned}$$

where n denotes the number of reactions, N the number of black boxes, ν_{ij} the stoichiometric matrix and v_j the reaction velocities.

5.2 The Model of CD95-induced Apoptosis

In this section, a quantitative model of CD95-induced apoptosis is derived from the network topology of Fig. 4.2. Thereby, existing information about the respective reaction mechanisms and the experimental accessibility are taken into account.

5.2.1 Mechanistic Subsystems

A detailed reaction mechanism could be established for the DISC- and the caspase-system. The mechanisms at the DISC are largely described by elementary reactions (e.g. [28]), whereas the caspase cleavage process is considered an enzymatic process (e.g. [104]). In principle, these interactions could have been modelled in a more simplified way. The influence of caspase-3 on Bid, for example, could have been modelled directly thereby using 'effective' parameters without accounting for the intermediate caspase-2 cleavage. However, since time series about the concentration of caspase-2 are measurable, this molecule was kept in the system in order to gain more information for system identification. In contrast, many molecules and interactions with equivalent properties were replaced by 'effective' molecules and interactions based on the analysis of parameter sensitivity correlations (see Section 6.3). The molecules XIAP, IAP1/2, survivin and their interactions with caspase-3,-7, and -9 are, for example, reduced to one 'effective' molecule called IAP

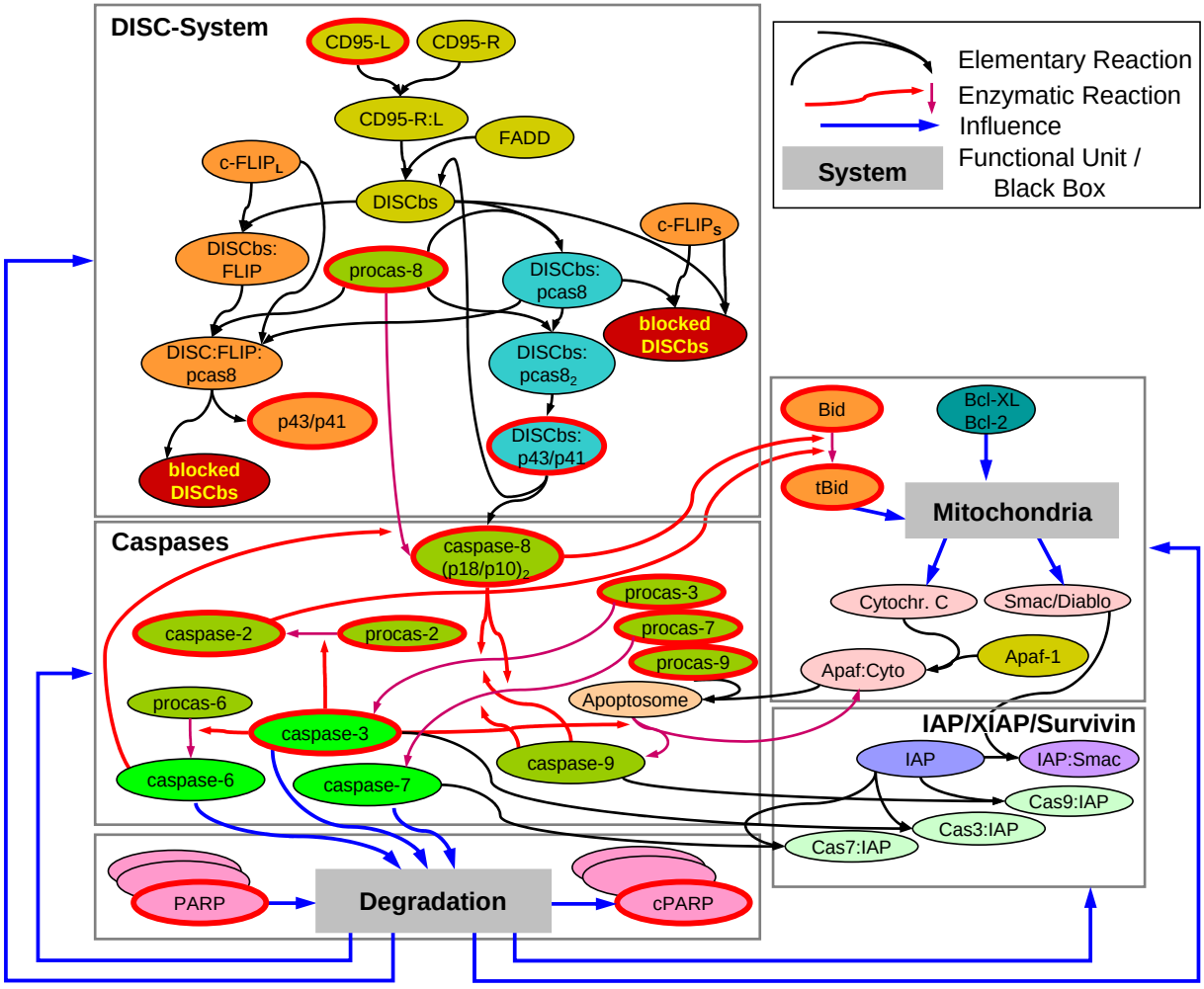

Figure 5.3: Structured information model of CD95-induced apoptosis. In the mechanistic part (DISC, caspases, IAP), interactions are modelled as elementary reactions including competitive inhibitions [49] and enzymatic reactions. Receptors are activated by ligands initiating the DISC formation. After binding to the DISC binding site (DISCbs), procaspase-8 is cleaved (initiator caspase), followed by the activation of executioner caspases (3, 6, 7). PARP cleavage was chosen as experimental end-point of the pathway. The mitochondria and the degradation process, which influences all molecules, are modelled as black boxes defined by their input-output behavior (see supp. online material). Each reaction contains one or more unknown parameters. Since the black boxes are based on observations, the number of unknown parameters is low here. Experimental time series were measured for all molecules framed in red. For details on reactions and parameters see Appendix B.1. Note that due to model simplifications some molecule species are replaced by virtual substitutes (e.g. IAP).

and the interactions are described by effective binding parameters. According to the experimental situation, information about the parameters belonging to the real biochemical reactions, cannot be gained anyway.

The complete system is shown in Fig. 5.3. Details about the chemical reactions are given in Appendix B.1. In the following, the definition of the black boxes is given.

5.2.2 Black Box Subsystems

Two subsystems with a significant influence on the signalling system are identified, whose mechanisms are not fully understood: the degradation process and the mitochondrial cytochrome-C release (Section 4.1). These are modelled in the form of black boxes as described in the following.

Degradation System

The execution of cell death after activation of executioner caspases is a complex and only partially understood process, which results in the degradation of all molecules considered in the model of CD-95 induced apoptosis. Although the model mainly addresses the signalling and the regulation of apoptosis, which mainly corresponds to the processes prior to degradation, the degradation itself cannot be ignored since it also affects the time evolution of the modelled signalling molecules. In order to substitute the complex degradation process of the real system, a virtual state variable called x_{apop} is introduced, which directly describes the velocity of the cell's degradation. Thus, this state variable can be considered to quantify the 'apoptotic activity'. In accordance with the network topology, x_{apop} increases upon activity of the executioner caspases, namely caspase-3, -6, and -7. It is assumed that the increase of the 'apoptotic activity' behaves similar to the experimentally observable PARP cleavage, since this molecule is one of the main players in the execution of cell death. Corresponding to the PARP cleavage, the increase of x_{apop} is modelled as 'virtual' enzymatic process, in which the executioner caspases are considered enzymes. Comparability with experimental data is provided by the observable molecule degradation. In addition, the activity itself is subject to a slow decay, corresponding to the half-live times of the real molecules causing the degradation.

Let $(x_1, \ldots, x_l)$ denote the concentration of all molecules and $x_{l+1} \equiv x_{apop}$. Thus, according to Section 5.1.2, the time evolution of x_{apop} is defined by

$$\begin{aligned} \frac{dx_{\text{apop}}}{dt} &= f_{l+1}^{\text{degrad}}(x_1, \ldots, x_l, x_{\text{apop}}) \\ &= \sum_{\xi=1}^{3} \frac{k_{\text{ecasp}_\xi} \cdot (1 - x_{\text{apop}}) \cdot x_{i_{\text{ecasp}\xi}}}{K_{\text{m, ecasp}_\xi} + (1 - x_{\text{apop}})} - K_{\text{apop\_decay}} \cdot x_{\text{apop}}, \end{aligned}$$

where $x_{i_{\text{ecasp}\xi}}$ denotes the concentrations of the three executioner caspases and k_{ecasp_ξ}, $K_{\text{m, ecasp}_\xi}$ the rate constants and Michaelis-Menten constants (Section 2.1.3) with respect

to the apoptotic activation. $K_{\text{apop\_decay}}$ denotes the constant for the decay of x_{apop}. Note that x_{apop} is normalized to $[0,1]$ and $x_{\text{apop}} = 0$ for $t = 0$. The introduced parameters are subject to parameter estimation.

In order to describe the degradation process affecting all molecules of the model, a decay function is introduced depending on x_{apop}. Based on experimental data, a quadratic relation between x_{apop} and the degradation rate is suggested in order to account for a superproportional influence of the apoptotic activity. In addition, a steady degradation of all active caspases like the caspase-8 complex $(\text{p18/p10})_2$ is assumed with the rate constant $k_{\text{std\_degrad}}$, regardless of the apoptotic activity. The steady degradation of the other pathway molecules like procaspases is not taken into account explicitly, since this 'non-apoptotic' degradation is supposed to be in equilibrium with the cell's steady expression of proteins. Note that the active caspase complexes are not directly expressed, but formed by cleavage of procaspases within the regarded pathway. Consequently, the black box behaviour describing the degradation process for all molecules of the model is given by the function

$$f_i^{\text{degrad}}(x_1,\ldots,x_l,x_{\text{apop}}) = -x_i \cdot k_{\text{degrad}} \cdot x_{\text{apop}}^2 - \begin{cases} k_{\text{std\_degrad}} \cdot x_i & : \text{ active caspase complexes,} \\ 0 & : \text{ otherwise,} \end{cases}$$

where $i = (1,\ldots,l)$. For simplicity, the degradation constant k_{degrad} is the same for all molecules[1]. This constant will also be subject to parameter estimation. The results in Section 9.2 show that this black box is well-suited to reproduce the observed behaviour. More details about the black box definition are given in Appendix B.2.

Mitochondrial Cytochrome-C Release

The cytochrome-C release of the mitochondria (Section 4.1) is modelled based on experimental observations [93], which describe a complete release within approximately 5 minutes as soon as pro-apoptotic substrates like tBid reach a certain level in comparison to anti-apoptotic substrates like Bcl-2/Bcl-XL. According to the network topology, the main link from the caspase system to the mitochondria is established by Bid. The exact mechanism of triggering the cytochrome-C release is only partially understood and consequently modelled as black box: once tBid reaches a threshold level in comparison with the anti-apoptotic substrate concentration of Bcl-Xl and Bcl-2, cytochrome-C release is triggered ($t = t_{\text{trigger}}$). The release rate is defined by the function $f_{\text{release}}(t)$ describing a smooth and complete release within a specified time interval [93, 94]. Thus, the time evolution of cytochrome-C is given by

$$\begin{aligned} x_{\text{cyt\_strd}}(t) &= x_{\text{cyt\_strd}}(t_{\text{trigger}}) \cdot f_{\text{release}}(t - t_{\text{trigger}}), \\ \frac{dx_{\text{cyt\_rel}}}{dt} &= -\frac{dx_{\text{cyt\_strd}}}{dt}, \end{aligned}$$

[1] The degradation of PARP is modelled separately by introduction of $k_{\text{degrad\_PARP}}$, since PARP is an active part of the degradation process itself.

where $x_{\text{cyt\_strd}}$ and $x_{\text{cyt\_rel}}$ denote the state variables of the respective cytochrome-C concentrations. Details are given in Appendix B.2.

5.2.3 Experimental Data and Observation Functions

Based on sensitivity analysis (Chapter 6) and on experimental feasibility, a set of experiments to measure time series of concentrations of M different molecules[2] after activation of CD95-receptors was designed. The measured molecules and molecule complexes are shown in Fig. 5.3 (red frames). Since the concentrations could be directly measured, the observation functions $y_\nu^{\text{model}}, \nu = (1, \ldots, M)$ (see Section 7.1) are equivalent to the measurable state variables $\{x_{i_1}, \ldots, x_{i_M}\}$, where i_ν denotes the index of the state variable assessed by the observation function y_ν^{model}. Due to the fact that some concentrations cannot be measured in absolute quantities, a scaling factor ζ_ν is introduced and

$$y_\nu^{\text{model}}(t) = \zeta_\nu \cdot x_{i_\nu}(t).$$

Cells were stimulated with different concentrations of agonistic anti-APO-1 antibody (Section 4.1), also referred to as ligands in the following, for various periods of time (from 5 minutes to 4 days). Each sample was evaluated by three independent approaches. See Appendix C for experimental details. In a first set of experiments, time series were measured for a 'fast' activation scenario with an oversaturated ligand concentration corresponding to more than one ligand per CD95-receptor. To gain additional information about the system's dynamic, several experiments with much lower ligand concentrations were performed resulting in a slower activation of apoptosis.

In summary, the resulting model consists of 41 molecules and molecule complexes, 32 reactions, and 2 black boxes. It contains more than 50 missing parameters. Therefore, it is still too complex for reliable system identification and requires further reduction of complexity considering the limited number of data points. Therefore, sensitivity analysis will be applied prior to system identification, as described in the next chapter.

[2]Note that complexes of molecules are also considered molecules.

Chapter 6

Sensitivity Analysis of Complex Systems

For analysing complex systems, parametric sensitivities are of high interest [33], however, their computation usually requires knowledge of all system parameters. In the following, I will present a new approach towards global sensitivity analysis, which is developed by introducing the concept of 'Sensitivity of Sensitivities'. I will show that sensitivities in biological systems can be highly insensitive to parameter variations, which in general provides a good basis for reduction of the system's dimensionality even without knowledge of the 'true' parameter values. The results are given in Section 9.1.

6.1 Parametric Sensitivity Analysis

The dynamics of complex biological systems are governed by many parameters. In general, each of these parameters affects the system to different extents. Parametric sensitivity analysis determines the changes of the system behaviour as a result of parameter variations. This methodology is adopted by many fields of sciences like process optimization, systems control, or cell biology. Reviews on mathematical theories about sensitivity analysis can be found in [95, 96].

In case of a system with m state variables and n parameters, the changes of the state variables $x_1, \ldots, x_m$ as a result of changes of the parameters $\Phi_1, \ldots, \Phi_n$ are determined. In the signal transduction system, which is investigated here, the state variables mostly correspond to molecule concentrations and the parameter set consists of biochemical constants related to the systems dynamic (e.g. reaction rate constants) and initial conditions like unknown initial concentrations of some species. The absolute sensitivities are defined by

$$s_{ij}^{\mathrm{abs}}(\mathbf{\Phi}) \quad = \quad \frac{\partial x_i}{\partial \Phi_j}. \tag{6.1}$$

In most cases however, the relative sensitivities s_{ij}, defined as

$$s_{ij}(\mathbf{\Phi}) \quad = \quad \frac{\Phi_j}{x_i}\,\frac{\partial x_i}{\partial \Phi_j} \quad = \quad \frac{\partial \ln(x_i)}{\partial \ln(\Phi_j)}, \tag{6.2}$$

are more meaningful. Consequently, the relative sensitivity matrix $\mathbf{S}$ will be used in the following. The matrix is defined by

$$\mathbf{S} = \begin{pmatrix} s_{11} & \dots & s_{1n} \\ \vdots & & \vdots \\ s_{m1} & \dots & s_{mn} \end{pmatrix}.$$

There are several motives for applying sensitivity analysis. On one hand, it allows conclusions about intrinsic system properties like stability and robustness of the system behaviour with respect to parameter fluctuations in a systematic way. On the other hand, information about insensitive and critical system parameters and about correlations between the parameter sensitivities provides a basis for system identification. Moreover, sensitivity analysis is required to identify those parameters which can be approached by parameter estimation based on experimental data (Chapter 7) and to quantify the reliability of parameter estimations. Furthermore, sensitivity analysis can be applied for optimal experimental design [97].

6.1.1 Calculating Sensitivities

An analytical derivation of sensitivities is usually not possible for large models. In the following, a model, whose time evolution is given by a system of ordinary differential equations (ODEs), defined by

$$\begin{aligned} dx_i/dt &= f_i(\mathbf{x}, \mathbf{\Phi}, t), \qquad i = (1, \dots, m), \\ x_i(0) &= x_i^I, \end{aligned} \tag{6.3}$$

is considered with the initial values $\{x_i^I\}$. In general, the state variables are numerically integrated using an ODE solver as described in Section 2.1.3. The sensitivities $s_{ij}(\mathbf{\Phi}, t)$ can be numerically determined in different ways, as demonstrated in the following.

Finite difference method

The absolute sensitivities are approximated by calculating the differences between two time evolutions with small variations $\Delta\Phi_j$ of parameter Φ_j[1]

$$s_{ij}^{\text{abs}}(\mathbf{\Phi}, t) \approx \frac{\Delta x_i}{\Delta \Phi_j} = \frac{x_i(\mathbf{\Phi} + \Delta\Phi_j \mathbf{e_j}, t) - x_i(\mathbf{\Phi}, t)}{\Delta \Phi_j}, \tag{6.4}$$

[1] A numerical refinement can be achieved by running the first time evolution with $+\frac{1}{2}\Delta\Phi_j$ and the second one with $-\frac{1}{2}\Delta\Phi_j$.

where $\mathbf{e_j}$ is the unit vector in j-direction. Problems arise from the fact that differences between two propagations of the complete system are evaluated. Since the numerical exactness of ODE solvers is always limited, a lower boundary for the parameter variation $\Delta\Phi_j$, namely $\Delta\Phi_{j,\min}$, is introduced. This value depends on the relative errors ϵ_{x_i} of each state variable x_i caused by the numerical integration. According to [33], a criteria for $\Delta\Phi_{j,\min}$ is given by

$$\epsilon_{s_{ij}} \cdot s_{ij}^{\text{abs}}(\mathbf{\Phi}, t) = \frac{\epsilon_{x_i} \cdot x_i}{\Delta\Phi_{j,\min}}, \tag{6.5}$$

where $\epsilon_{s_{ij}}$ is the allowed fractional error of sensitivity s_{ij}^{abs}. Note that different magnitudes of the variables x_i require different $\Delta\Phi_j$ in order to achieve the same accuracy. Obviously, the minimum difference $\Delta\Phi_{j,\min}$ is higher for low sensitivities. Taking into account that, on the other hand, the finite difference approximation requires small parameter variations, $\Delta\Phi_j$ should be adapted close to the minimum.

Another problem is related to the finite time steps that are used for the integration of ODEs. Due to adaptive step size control (see Section 2.1.3), parameter variations result in different integration time steps causing considerable inaccuracies. In this thesis, the problem is approached by applying the same time step sequence for both integrations. This is realized by recording the stepping sequence chosen by the adaptive step size control for the first integration (with parameter Φ_j). For the second integration (with parameter $\Phi_j + \Delta\Phi_j$), the solver is forced to step through the same sequence. Additional steps are inserted where required. In order to make the insertion of additional steps unlikely, a stronger step size requirement is applied for the first integration, resulting in smaller step sizes.

Direct differential method

In this method, sensitivities are determined by their time evolution, described by a system of ordinary differential equations, called *variational equations*, given by

$$\begin{aligned} \frac{ds_{ij}^{\text{abs}}(\mathbf{x}, \mathbf{\Phi})}{dt} &= \frac{d(\partial x_i / \partial \Phi_j)}{dt} \\ &= \sum_{k=1}^{m} \frac{\partial f_i}{\partial x_k} \frac{\partial x_k}{\partial \Phi_j} + \frac{\partial f_i(t)}{\partial \Phi_j} \\ &= \sum_{k=1}^{m} J_{ik}(t) \cdot s_{kj}^{\text{abs}}(\mathbf{x}, \mathbf{\Phi}) + \frac{\partial f_i(t)}{\partial \Phi_j} \end{aligned} \tag{6.6}$$

with the Jacobian matrix $\mathbf{J} = \{J_{ik}\},\ J_{ik} = \frac{\partial f_i}{\partial x_k}$. Thus,

$$\frac{d\mathbf{s}_j^{\text{abs}}(\mathbf{x}, \mathbf{\Phi})}{dt} = \mathbf{J}(t) \cdot \mathbf{s}_j^{\text{abs}}(\mathbf{x}, \mathbf{\Phi}) + \frac{\partial \mathbf{f}(t)}{\partial \Phi_j} \tag{6.7}$$

with $\mathbf{s}_j^{\text{abs}} = (s_{1j}^{\text{abs}}, \ldots, s_{mj}^{\text{abs}})^T$. The initial sensitivities s_{ij}^{abs} at time $t = 0$ are set to zero unless they refer to an initial value, consequently

$$s_{ij}^{\text{abs}}|_{t=0} = \begin{cases} 1 & : \quad x_i^I \equiv \Phi_j \\ 0 & : \quad x_i^I \not\equiv \Phi_j \end{cases} .$$

Together with the model system given by Eq. (6.3) itself, a system of $(m+1) \times n$ equations has to be integrated for computation of the sensitivities s_{ij}^{abs} of all state variables with respect to all parameters. Apart from the high computational effort, step size problems may occur if the sensitivities are propagated together with the state variables. The method can be improved by decoupling the model equations (6.3) from the sensitivity equations and by solving the state trajectory $\mathbf{x}(t)$ separately in a first step, followed by computation of the sensitivities using interpolation of the solution $\mathbf{x}(t)$ for computation of the Jacobian matrix $\mathbf{J}$ and $\partial f_i / \partial \Phi_j$ [98].

The choice of the appropriate method depends on the specific problem[2]. The finite difference method is, for example, applied for the sensitivity-controlled parameter estimation algorithm (see Chapter 7), where only the sensitivities of some state variables with respect to some parameters are required and the computational effort of the finite difference method is much lower. However, a correct choice of the parameter variation $\Delta\Phi_j$ is essential. Note that the usability of the direct differential method is limited if the model contains 'black boxes' (see Chapter 5), since in this case, an analytical calculation of the Jacobian matrix and of $\partial f_i / \partial \Phi_j$ is not always possible.

6.1.2 Time Dependency of Sensitivities

The time points $\{t_k\}$, for which the sensitivities $s_{ij}(\mathbf{\Phi}, t_k)$ are computed, have to be chosen carefully. In *Metabolic Control Analysis* [9, 99] for example, steady states depending on parameter variations are investigated, whereas in signal transduction systems, the transient behaviour is of high interest to analyse the system's regulation. As a consequence, the complete time period, during which e.g. a signalling pathway is active and exhibits a dynamical behaviour, is considered relevant rather than a distinct time point. In this thesis, two strategies are applied:

- For sensitivity analysis performed in the context of parameter estimation (Section 7.2.2), the time points are chosen according to the measuring points of the experiments. Here, the *objective sensitivity* is computed, defined as the relative sensitivity of the objective function, also called χ^2-function (Section 7.1), with respect to parameter variations. The χ^2-function, which has to be minimized, depends on the differences between simulation and experimental data. As a consequence, only those points, for which experimental data exist, are relevant.

[2] A third method - the Greens function method - is not explained here, since it is not efficient if there are more state variables than parameters for which sensitivities are required [33].

- Sensitivity analysis is also applied for model reduction prior to parameter estimation and independently of experimental data at first. In order to avoid artefacts related to the choice of the time points $\{t_k\}$, the absolute values of the relative sensitivities[3] are averaged over the complete time period of interest instead of choosing distinct time points. These 'averaged' sensitivities are defined by

$$\hat{s}_{ij}(\mathbf{\Phi}) = \frac{1}{\Delta t} \int_{t_0}^{t_0+\Delta t} \Big| s_{ij}(t, \mathbf{\Phi}) \Big| dt. \tag{6.8}$$

  The time period of interest starts at t_0, the beginning of an investigated scenario, e.g. the time of activation of a signal transduction pathway. The time interval Δt corresponds to the period, in which the system shows reactions to parameter variations, e.g. the period until equilibrium is reached. In case of the apoptosis application investigated in this thesis, the interval ends when the cell is completely degraded.

In the following, the sensitivity matrix $\mathbf{S}$ refers to the averaged relative sensitivities $\hat{s}_{ij}$.

6.1.3 Local versus Global Sensitivity Analysis

In most cases, the time evolution and the sensitivities of large models can be computed numerically only. As a consequence, a general relation between sensitivities $\hat{s}_{ij}(\mathbf{\Phi})$ and the parameter values $\mathbf{\Phi}$, at which the sensitivities are determined, cannot be deduced. Instead, sensitivities are determined for specific points in parameter space only. These are called *local sensitivities*. In contrast, *global sensitivity analysis* is required to provide information about the sensitivities $\hat{s}_{ij}(\mathbf{\Phi})$ as a function of $\mathbf{\Phi}$. In complex biological systems, the parameters Φ_j are often unknown and only a range of possible parameter values $\Phi_{j,\min} \leq \Phi_j \leq \Phi_{j,\max}$ can be given, resulting in a high-dimensional space of possible parameter value sets. Strictly speaking, global sensitivity analysis has to provide the local sensitivities for the complete space of possible parameter values. A general solution to this problem does not exist. Section 6.2 describes how this problem is approached in this thesis.

6.1.4 Sensitivity and Robustness of Real Systems

Considering the lack of information about system parameters, sensitivity analysis applied to complex biological systems, like biochemical reaction networks, appears to be questionable. On the other hand, it can be taken advantage of the fact that these systems are often characterized by high robustness. Many biological systems keep their system properties constant, although they are subject to high parameter fluctuations (e.g. [29, 30, 100, 101]). The reaction kinetics and initial concentrations of real systems are, for example, not fine-tuned.

[3] The main goal of sensitivity analysis prior to parameter estimation is the identification of dependencies between parameters and state variables. Thus, the absolute values are integrated in order to detect any dependency.

Nevertheless, complex molecular networks must operate reliably under ever-changing environmental conditions that can cause large changes in these internal parameters [102]. Robustness and low sensitivities are even proposed as a measure of plausibility in models of biochemical networks [103]. Since the adaptation property is often a consequence of the network's connectivity [30], appropriate models of biochemical networks are likely to provide information about the system behaviour even without exact knowledge of the parameters. This is an important assumption for the next section.

6.2 Stochastic Approach to Global Sensitivity Analysis

High dimensionality of the space of unknown parameters and large ranges of possible parameter values are ubiquitous in complex biological systems. New methods for system identification like parameter estimation are required and can be considered one of the main challenges in this area.

In this thesis, sensitivity analysis is used as an essential tool for system identification and model reduction. It is applied to classify the parameters into critical, correlated and insensitive ones and to identify 'independent' subsystems, providing a basis for efficient parameter estimation. Consequently, these steps have to be performed prior to the determination of parameters. Therefore, the usefulness of local sensitivity analysis is limited if there is little information about the parameters at first. In addition, global sensitivity analysis is usually impaired by the high dimensionality of the parameter space leading to an enormous computational effort. Moreover, general evidence cannot be provided if the sensitivities strongly depend on unknown parameters.

6.2.1 Sensitivity of Sensitivities: Random Walk through Parameter Space

In a virtual experiment, I performed sensitivity analysis for a large number of randomly chosen points in parameter space within specified ranges, covering up to three orders of magnitude. For the application investigated in this thesis, the ranges are defined for each parameter type (e.g. bimolecular reaction rate constants, initial concentrations, Michaelis-Menten constants, etc.), unless more precise information is available. Thereby, the robustness of sensitivities with respect to large parameter variations can be tested, motivated by two important facts: Typically, the behaviour of biological systems exhibits high robustness to parameter fluctuations (see Section 6.1.4), indicating that the area of impact of certain parameters is limited. This suggests that the value of certain sensitivities behaves insensitive with respect to certain parameter variations as well. In the following, this assumption is referred to as 'low Sensitivity of Sensitivities'. Second, structure and connectivity of biochemical networks suggests that the influence of some parameters on

some state variables of distant network regions is limited and that the associated sensitivities are extremely low - independent of the parameter values. Different strategies for the generation of random parameter sets are given:

- The logarithms of the randomly generated parameter values are equally distributed within the respective ranges $[\ln(\Phi_{j,\min}); \ln(\Phi_{j,\max})]$. Typically, this is applied if the orders of magnitude are the only information about the parameter values.

- If, for example, literature values are available (e.g. [104]), it is more appropriate to generate the random parameter values according to a normal distribution around the given value Φ_j^* with the standard deviation σ_j. Often, biochemical parameter values refer to different experiments and different conditions and σ_j has to be chosen sufficiently high.

The distributions of the computed sensitivities $\hat{s}_{ij}$ for the randomly generated points in parameter space are plotted in the form of histograms for the sensitivity value of each (i, j)-combination. In Fig. 9.3, histograms are shown for the signal transduction application of this thesis. The x-axis shows the sensitivity value and the y-axis corresponds to the number of occurrences. Thus, a sharp peak can be considered an indication that the respective sensitivity is likely to keep about the same value, independent of the parameter set.

6.2.2 Statistical Evidence of Stochastic Sensitivity Analysis

A systematic scan of a multi-dimensional parameter space is computationally limited due to the Curse of Dimensionality (Chapter 1). For the same reason, the significance of stochastically chosen sensitivities is also questionable. However, I will demonstrate that under the assumption that the parameter space can be projected onto an 'effective' parameter space of a lower dimension, it is a reasonable approach. This assumption is typically fulfilled in case of biological systems exhibiting high robustness.

Let the sensitivity matrix $\mathbf{S}(\mathbf{\Phi})$ be computed for a set of randomly chosen parameter sets $\{\mathbf{\Phi}^q\},\ q = (1, \ldots, N)$. The number N is of course limited by computational feasibility. The space of possible parameter values is defined according to

$$P = P_1 \times P_2 \times \ldots \times P_n\ , \quad P_k = [\Phi_{k,\min}; \Phi_{k,\max}].$$

Now, a minimum density of parameter points per volume is introduced, for which sensitivities have to be computed to provide representative information about possible sensitivity values within space P. If $\rho_k^{ij}(\mathbf{\Phi})$ denotes the required density for $\hat{s}_{ij}$ in k-direction, the total number N_{ij}^* of required samples is given by

$$N_{ij}^* \ \geq \int_{\Phi_{1,\min}}^{\Phi_{1,\max}} d\Phi_1 \cdots \int_{\Phi_{n,\min}}^{\Phi_{n,\max}} d\Phi_n \prod_{k=1}^{n} \rho_k^{ij}(\Phi_1, \ldots, \Phi_n) \tag{6.9}$$

$$\text{with} \quad \int_{\Phi_{k,\min}}^{\Phi_{k,\max}} d\Phi_k \rho_k^{ij}(\mathbf{\Phi}) \to 1 \quad \text{for} \quad \frac{\partial \hat{s}_{ij}}{\partial \Phi_k} \to 0. \tag{6.10}$$

Note that ρ_k^{ij} is influenced by the variation of the sensitivity matrix with respect to parameter k, thus ρ_k^{ij} is dependent on $\mathbf{\Phi}$. Examining large systems, one is often interested in the behaviour of a subset of state variables only[4], or in observation functions (Section 7.1) depending on a subsets of state variables. Therefore, the sub-matrix S^*, defined as

$$S^* = \{\hat{s}_{ij}(\mathbf{\Phi})\}, \quad i \in M^* \subseteq M = \{1, \ldots, m\}, \quad j \in \{1, \ldots, n\}, \tag{6.11}$$

is considered in the following. M^* denotes the set of indices of the state variables of interest, for example those state variables, which can be experimentally observed.

Now, it is claimed as a crucial system property that the behaviour and the sensitivities of each single state variable x_i, $i \in M^*$, are insensitive to a considerable number of parameters, described by the index set $\mathcal{P}_{ij}^0 \subseteq \{1, \ldots, n\}$, which is defined by

$$\exists\, \mathcal{P}_{ij}^0 \;\Big|\; \frac{\partial \hat{s}_{ij}}{\partial \Phi_k} \approx 0 \;\; \forall\, k \in \mathcal{P}_{ij}^0.$$

This assumption is motivated by the fact that biological systems typically show extremely high robustness to most parameter fluctuations (see Section 6.1.4). Moreover, hierarchical structures are suggested for large biochemical systems (e.g. [105, 32]), resulting in stable interdependencies and low sensitivities of the high-level state variables with respect to parameters on the lower level. It is well-accepted that complex biological systems contain subsystems acting as control-modules [12], providing a stable input-output behaviour on the upper level. In addition, subprocesses often run on timescales, which are completely different from the main process. Therefore, the behaviour of interest is mostly insensitive towards many parameters of subprocesses within certain ranges. Aside from the general context, there is strong evidence that the assumed system property is fulfilled in the signal transduction application of this thesis.

Since Eq. (6.12) results in $\rho_k^{ij} \approx 0$ for $k \in \mathcal{P}_{ij}^0$, the number N_{ij}^* reads

$$N_{ij}^* \approx \int_{\Phi_{k_1,\min}}^{\Phi_{k_1,\max}} d\Phi_{k_1} \cdots \int_{\Phi_{k_l,\min}}^{\Phi_{k_l,\max}} d\Phi_{k_l} \prod_{r=1}^{l} \rho_{k_r}(\Phi_{k_1}, \ldots, \Phi_{k_l}), \tag{6.12}$$
$$\text{where } \{k_1, \ldots, k_l\} \text{ is the index set } \mathcal{P}_{ij}^* := \{1, \ldots, n\} \setminus \mathcal{P}_{ij}^0.$$

N_{ij}^* corresponds to the sample number required for an l-dimensional parameter space. Note that l is specific to each sensitivity $\hat{s}_{ij}$ and that in most cases $l \ll n$. For each sensitivity,

[4] e.g. the key molecules in biological systems, providing the core functionality, or the subset of experimentally accessible state variables in case of parameter estimation problems.

the parameter space can thus be written as

$$P = P^*_{ij} \times P^0_{ij} \qquad \text{with} \quad P^*_{ij} = P_{k_1} \times \ldots \times P_{k_l}$$
$$\text{and} \quad P^0_{ij} = P_{k^0_1} \times \ldots \times P_{k^0_{n-l}}, \quad k^0_\kappa \in \mathcal{P}^0_{ij},$$

and consequently, each random parameter set $\mathbf{\Phi}^q \in P$ can be projected onto P^*_{ij}.

Thus, even if the relevant dimensionality l of the parameter space with respect to the sensitivities $\hat{s}_{ij}(\mathbf{\Phi})$ is not known in advance, the set of N generated random parameter values, for which the sensitivities of interest are computed, automatically contains N parameter sets of space P^*_{ij} for all i, j of interest. Depending on dimension l, the number N might yield a density, which reaches an adequate level to gain global information about the sensitivity distributions. For an isotropic parameter space for example, the actual density would result in $\rho^{ij}_k = \sqrt[l]{N}$.

Another important characteristic of biochemical systems is the occurrence of coupled parameters or parameters showing multiple correlation [106]. This means that the effects of parameter variations on the system behaviour are correlated. In case of a reaction chain for example, the behaviour of the final product concentration is influenced by the rate constants of the intermediate reactions in a similar way. Such processes can be reproduced by simplified models using a lower number of effective parameters [27]. Thus the parameter space of correlated parameters can be mapped onto a space of 'effective' parameters, P^E, described by

$$P_{k_1} \times \ldots \times P_{k_\mu} \;\rightarrow P^E_{k^E_1} \times \ldots \times P^E_{k^E_\nu}, \quad \nu < \mu,$$

which has a lower dimensionality and by which the original system can be approximated after redefining the model. Accordingly, the respective transformation of the random parameter points $\{\mathbf{\Phi}^q\}$ results again in a higher density of samples with respect to the efficient parameter space. A systematic approach can for example be derived from the *High Dimensional Model Representation* method [107].

In practice, the computation of sensitivities and sensitivity correlations has shown that the dimensionality of the parameter space related to the signal transduction application of this thesis can be drastically reduced. This is demonstrated in Section 9.1. Thus, even without detailed prior knowledge of the dependencies of the sensitivities on the parameters and merely under the assumption that there is a large number of parameters, which are either correlated, or which do not influence certain sensitivities, the stochastic sensitivity approach provides general information about $S^*(\mathbf{\Phi})$.

6.2.3 Data-based Weighting of Sensitivities

The distribution of the sensitivity values $\hat{s}_{ij}$ for all parameter sets $\{\mathbf{\Phi^q}\}$ are visualized as histograms (Fig. 9.3). In this first approach, all random parameter sets are equally

weighted, independent on how much the resulting system dynamics deviate from the real system. As an extension, I therefore incorporated information from experimental data by introducing a weighting factor:

Based on experimental time series of molecule concentrations, an objective function E_q (equivalent to the χ^2-function in Section 7.1) is computed for each parameter set $\mathbf{\Phi}^q$. This objective function is based on the differences between the experimental and simulation data according to

$$E_q = E(\mathbf{\Phi^q}) = \sum_{i,k} \frac{1}{2} \frac{[x_i(t_k, \mathbf{\Phi^q}) - x_{ik}^{\exp}]^2}{\sigma_{ik}^2}, \tag{6.13}$$

$x_{ik}^{\exp}$: experimental value for molecule i at time t_k,

σ_{ik} : standard deviation for molecule i at time t_k.

Then, it is assumed that p_q, the probability that a system with parameter set $\mathbf{\Phi^q}$ produces the experimental data $\{x_{ik}^{\exp}\}$, follows a Boltzmann distribution with the objective function as energy term:

$$p_q \propto \exp(-E_q/kT). \tag{6.14}$$

This assumption is motivated as follows: considering Gaussian random measurement errors, each characterized by σ_{ik}, the probability p_q can be written as product [108]

$$p_q \propto \prod_{i,k} \exp\Big(-\frac{1}{2} \frac{[x_i(t_k, \mathbf{\Phi^q}) - x_{ik}^{\exp}]^2}{\sigma_{ik}^2} \Big). \tag{6.15}$$

Using Eq. (6.13), this relation is equivalent to Eq. (6.14) with $kT = 1$.

For generation of a second type of sensitivity histograms, the Boltzmann factor is applied. It is used as weighting factor characterizing the weight of the individual contributions. Thus, the histograms are computed as follows: instead of counting the number of parameter sets with sensitivities $\hat{s}_{ij}$ within a certain sensitivity interval, the contribution of each parameter set $\mathbf{\Phi}^q$ is given by $\exp(-E_q/kT)$. Note that for $kT \to \infty$, the weighted sensitivity histograms become identical with the non-weighted histograms.

Additional information can be gained by varying kT. Starting from the equally weighted histograms ($kT = \infty$), the system can be 'cooled down', resulting in histograms based on the Boltzmann factor. If a sensitivity behaves insensitive with respect to parameter variations, particularly within the subspace of probable solutions (areas in parameter space with a low objective function), decreasing kT results in histograms with a much sharper peak. Sensitivity histograms showing more than one distinct peak when the system is cooled down, indicate that the respective sensitivity strongly depends on the exact parameter set within the parameter subspace of probable solutions. Details are given in Section 9.1.2, where the method is applied to the apoptotic signal transduction system.

6.3 Analysis of Parameter Sensitivity Correlations

Apart from detecting parameters with low sensitivities, information about the parameter correlations is also crucial to reduce the system's dimensionality. As a consequence, I computed the parameter sensitivity correlations, which were then applied for the sensitivity-based parameter estimation algorithm (see Chapter 7.2.2).

Let us consider the state variables $x_i(t_k), i \in M^*$ (Eq. (6.11)), for which we assume that experimental data are given at the time points $\{t_k\}$. In complex networks, determination of parameters based on such data typically leads to sub-manifolds, where different combinations of parameters result in the same values for the respective state variables [106]. Thus, only relations between the parameters of the respective subsets can be determined since their influences on the resulting state variables are correlated.

6.3.1 Normalized Fisher Information Matrix

Let us consider a system with the unknown parameter set $\mathbf{\Phi}$. Information about the parameters has to be extracted from observations $\mathbf{y} = (y_1, \ldots, y_N)$. These observations can be described as realizations of random variables, for which a probability density function $\rho(\mathbf{\Phi}, z_1, \ldots, z_N)$ is given [109]. Let $\hat{\mathbf{\Phi}}(\mathbf{y})$ be the result of an estimator, such that $\langle \hat{\mathbf{\Phi}}(\mathbf{y}) \rangle = \mathbf{\Phi}_0$, where $\mathbf{\Phi}_0$ denotes the 'true' parameter set. Then, the precision of the estimated parameters is limited depending on the observations. A lower bound for the deviation of any estimator (even for the estimator of the highest quality) is given by

$$\left\langle [\hat{\mathbf{\Phi}}(\mathbf{y}) - \mathbf{\Phi}_0][\hat{\mathbf{\Phi}}(\mathbf{y}) - \mathbf{\Phi}_0]^T \right\rangle \geq \mathbf{J}^{-1}, \tag{6.16}$$

called the Cramér-Rao inequality, for which a proof is given in [109]. The matrix $\mathbf{J}$ is known as the Fisher Information matrix and consists of the elements

$$J_{pq} = \left\langle \frac{\partial \ln \rho(\mathbf{\Phi}, \mathbf{y})}{\partial \Phi_p} \frac{\partial \ln \rho(\mathbf{\Phi}, \mathbf{y})}{\partial \Phi_q} \right\rangle. \tag{6.17}$$

Under the assumption that the deviation of the observations is governed by Gaussian noise [106], the information matrix can be written as

$$J_{pq} = \sum_{\nu=1}^{N} \omega_\nu^2 \frac{\partial \hat{y}_\nu}{\partial \Phi_p} \frac{\partial \hat{y}_\nu}{\partial \Phi_q}. \tag{6.18}$$

The weighting factors ω_ν correspond to the reciprocal standard deviation and $\hat{y}_\nu$ denote the model values for y_ν based on an estimated parameter set $\hat{\mathbf{\Phi}}$. Applied to the previous reaction system, where the observations $\{x_i(t_k)\}$ of molecule populations are obtained, the information matrix can be expressed by the vectors $\mathbf{S}_j$ containing the absolute sensitivities

and reads

$$J_{pq} = \mathbf{S}_p \mathbf{S}_q \quad \text{with} \quad \mathbf{S}_j = \begin{pmatrix} \omega_1 \, \partial\hat{x}_1/\partial\Phi_j \\ \omega_2 \, \partial\hat{x}_2/\partial\Phi_j \\ \vdots \\ \omega_N \, \partial\hat{x}_N/\partial\Phi_j \end{pmatrix}. \tag{6.19}$$

The index $\nu = (1, \ldots, N)$ replaces the indices i and k in order to enumerate the combinations of observable state variables x_i at the times t_k: $x_i(t_k) \rightarrow \hat{x}_\nu$. To identify correlated parameters, the *normalized Fisher Information matrix* $\mathbf{J}^n$ [106] can be computed as

$$J^n_{pq} = \mathbf{S}^{\mathrm{n}}_p \mathbf{S}^{\mathrm{n}}_q \qquad \text{with} \quad \mathbf{S}^{\mathrm{n}}_j = \frac{\mathbf{S}_j}{|\mathbf{S}_j|}. \tag{6.20}$$

Obviously, each element of $\mathbf{J}^{\mathrm{n}}$ corresponds to the cosine of the angle between $\mathbf{S}_p$ and $\mathbf{S}_q$, and $\mathbf{J}^{\mathrm{n}}$ can be applied as a measure for parameter sensitivity correlations. Parallel sensitivity vectors ($\mathbf{S}^{\mathrm{n}}_p \mathbf{S}^{\mathrm{n}}_q = 1$), for example, indicate that both parameters can be varied in a way, which results in exactly the same effect with respect to all observed state variables x_i for the respective time points t_k.

Remark: The elements of $\mathbf{S}^{\mathrm{n}}_j$ are weighted with $\omega_\nu = 1/\sigma_\nu$. If $\mathbf{S}^{\mathrm{n}}_j$ contained the absolute sensitivities only, the matrix elements of $\mathbf{J}^{\mathrm{n}}$ would in general be dominated by sensitivities of state variables with the highest absolute values, thereby neglecting correlations between all other sensitivities. However, since the orders of magnitude of absolute standard deviations are mostly proportional to the absolute value ranges of the state variables themselves, relative standard deviations $\sigma^{\mathrm{rel}}_\nu$ are more appropriate, resulting in a new form for $\mathbf{S}^{\mathrm{n}}_j$, which reads

$$\begin{aligned} S_{\nu j} &= \frac{1}{\sigma^{\mathrm{rel}}_\nu \hat{x}_\nu} \frac{\partial \hat{x}_\nu}{\partial \Phi_j} \\ \Rightarrow S^n_{\nu j} &= \frac{1}{|\mathbf{S_j}|} \left(\frac{1}{\sigma^{\mathrm{rel}}_\nu} \frac{\Phi_j}{\hat{x}_\nu} \frac{\partial \hat{x}_\nu}{\partial \Phi_j} \right) \quad = \quad \frac{1}{|\mathbf{S_j}|} \frac{1}{\sigma^{\mathrm{rel}}_\nu} \cdot s_{\nu j} \end{aligned} \tag{6.21}$$

with the relative sensitivities $s_{\nu j}$. The factor Φ_j could be inserted since this does not change the normalized vector. Thus, $\mathbf{J}^n$ can be considered a matrix of scalar products of the dimensionless relative sensitivities, weighted with the relative standard deviation, which lowers the impact of those data points, for which the relative standard deviation is high.

6.3.2 Application of Parameter Correlations

In many cases, parameter sensitivities were found to be highly correlated as an intrinsic system property, regardless of the parameter values. If data, which is required for the identification of each single parameter, is not accessible, the respective model part is replaced

by an 'effective' subsystem containing less parameters ('effective parameters'), thereby breaking down the sub-manifold of equivalent solutions. For example, reaction chains can often be replaced by one reaction using an 'effective' reaction rate (and time delay, where appropriate). Different molecules with similar properties and reactions can be replaced by one 'effective' molecule whose reactions are described by 'effective' reaction rates. An example is given in Section 5.2.1, where the molecules XIAP, IAP1/2 and survivin are replaced by the 'effective' IAP molecule (compare Fig. 4.2 and Fig. 5.3).

Whenever parameter sensitivity correlations are detected for certain areas in parameter space only, the model is not modified. Instead, for parameter estimation a threshold Θ_{corr} is introduced and correlated parameter subsets P_κ^{corr}, defined as

$$P_\kappa^{\text{corr}} = \{\Phi_{i_1}, \ldots, \Phi_{i_r}\} \;\Big|\; |\mathbf{S}_p^n \mathbf{S}_q^n| \geq 1 - \Theta_{\text{corr}} \;\; \forall\, \Phi_p, \Phi_q \in P_\kappa^{\text{corr}}, \tag{6.22}$$

are identified. Only the most sensitive parameter within such a subset is varied while the rest is kept constant as long as the correlation remains above the threshold. In summary, either the model is simplified a priori, or all parameters are kept in the system, but only a subset of them is estimated depending on the local correlations.

The methods introduced in this chapter were applied to the signal transduction system of CD95-induced apoptosis. The results are described in Section 9.1.

Chapter 7

Parameter Estimation

In this chapter, a quick introduction into parameter estimation methods is given and local estimation algorithms are demonstrated. Then, a general problem of large nonlinear systems with a high number of unknown parameters and large ranges of possible parameter values is addressed. These features are ubiquitous in biochemical networks and impair the search for the globally best parameter set. For this reason, I developed a sensitivity-controlled and cluster-based parameter estimation approach. Based on the 'global' sensitivity analysis of Section 6.2, the system's complexity is reduced a priori. This reduction is complemented by a sensitivity-control within the parameter estimation algorithm based on local sensitivities. Furthermore, I will demonstrate the new concept of 'ensembles' of estimated parameter sets.

7.1 Maximum Likelihood Estimation

In typical parameter estimation problems, a model M with m state variables $x_1, \ldots, x_m$ is given, which contains a set of n unknown parameters $\Phi_1, \ldots, \Phi_n$. The goal is to find the parameter values, for which the model predictions are in the best agreement with the given experimental data D. In the *maximum likelihood estimation* [108], this is achieved by searching for a parameter set $\mathbf{\Phi}$, which maximizes $p(D \mid M, \mathbf{\Phi})$, the probability that model M with the parameter values $\mathbf{\Phi}$ produces the given experimental data, thereby requiring an error model for the measurements.

Typically, it is assumed that the errors of the experimental data follow a Gaussian distribution [108]. Let $y_i^{\text{model}}(x_1, \ldots, x_m, t), i = (1, \ldots, r)$, denote observation functions and $\{t_k\}, k = (1, \ldots, s)$, the time points for which experimental data are measured. In the simplest case, y_i^{model} corresponds to the state variables themselves, provided that they are directly measurable ($y_i^{\text{model}} \equiv x_i, r = m$). Due to the Gaussian distribution of the experimental errors, the probability that the measured value for the i-th observation at

time t_k, given a parameter set $\mathbf{\Phi}$, lies in the interval $[y_{ik}, y_{ik} + dy_{ik}]$ reads

$$p(y_{ik} \mid \mathbf{\Phi}) = \frac{1}{\sqrt{2\pi\sigma_{ik}^2}} \exp\Big(- \frac{[y_{ik} - y_i^{\text{model}}(t_k, \mathbf{\Phi})]^2}{2\sigma_{ik}^2}\Big) dy_{ik} \tag{7.1}$$

with the variance σ_{ik}^2 originating from the distribution of the experimental errors, typically equivalent to the variance of multiple experimental measurements for each data point. Let us assume that experimental values $\{y_{ik}^{\text{exp}}\}$ are given for all time points of the time series $\{t_k\}$ and all observation functions[1]. As a consequence of Eq. (7.1) and on condition that the measurement errors are independent of each other, the probability for a data set $\{y_{ik}^{\text{exp}}\}$ fulfils the relation

$$p(\{y_{ik}\} \mid \mathbf{\Phi}) \propto \prod_{i,k} \exp\Big(- \frac{[y_{ik} - y_i^{\text{model}}(t_k, \mathbf{\Phi})]^2}{2\sigma_{ik}^2}\Big). \tag{7.2}$$

According to the maximum likelihood estimation, the goal is to find a parameter set, which maximizes this probability. Consequently, the term

$$\chi^2(\mathbf{\Phi}) = \sum_{i,k} \Big(\frac{[y_{ik}^{\text{exp}} - y_i^{\text{model}}(t_k, \mathbf{\Phi})]^2}{\sigma_{ik}^2}\Big), \tag{7.3}$$

also referred to as *objective function*, has to be minimized.

Thus, the sum of squares of differences between experimentally measured and simulated data, divided by the standard deviation, which lowers the impact of experimentally less reliable data, has to be minimized in order to optimally fit the model with data from experiments. In the signal transduction application of this thesis, the main problem arises from the high number of unknown parameters and their highly nonlinear influence on the system behaviour.

7.1.1 Nonlinear Least Squares

If the χ^2-function depends on the parameters in a nonlinear way, a general method, which guarantees to find the global minimum of χ^2 does not exist. Instead, there is a variety of iterative methods starting from an initial parameter guess. A review on the methods commonly used for biochemical reaction networks is given in [25]. Most methods are based on derivative-driven algorithms. This means that $\chi^2(\mathbf{\Phi})$ and its derivatives are evaluated locally and the parameters are moved accordingly in order to reach the next minimum. For this reason, computation of the gradient, whose components are given by

$$(\nabla\chi^2)_\nu = \frac{\partial\chi^2}{\partial\Phi_\nu} = -2\sum_{i,k} \frac{y_{ik}^{\text{exp}} - y_i^{\text{model}}(t_k, \mathbf{\Phi})}{\sigma_{ik}^2} \cdot \frac{\partial y_i^{\text{model}}(t_k, \mathbf{\Phi})}{\partial\Phi_\nu}, \tag{7.4}$$

[1] The assumption that experimental data are given for all observation functions at all time points $\{t_k\}$ is not always fulfilled; the generalization is, however, trivial.

is required. Furthermore, the Hessian Matrix, which is defined as the matrix of second derivatives,

$$\begin{aligned} H_{\mu\nu} &= \frac{\partial^2 \chi^2}{\partial \Phi_\mu \partial \Phi_\nu} \\ &= 2 \sum_{i,k} \frac{1}{\sigma_{ik}^2} \left[\frac{\partial y_i^{\text{model}}(t_k, \mathbf{\Phi})}{\partial \Phi_\mu} \frac{\partial y_i^{\text{model}}(t_k, \mathbf{\Phi})}{\partial \Phi_\nu} - [y_{ik}^{\text{exp}} - y_i^{\text{model}}(t_k, \mathbf{\Phi})] \frac{\partial^2 y_i^{\text{model}}(t_k, \mathbf{\Phi})}{\partial \Phi_\mu \partial \Phi_\nu} \right] \end{aligned} \tag{7.5}$$

has to be computed.

The *steepest descent* or *gradient descent* method seeks for the minimum by stepping into the direction of the gradient

$$\mathbf{\Phi}^{(l+1)} = \mathbf{\Phi}^{(l)} - \alpha \nabla \chi^2(\mathbf{\Phi}^{(l)}), \tag{7.6}$$

where l denotes the index of the iteration step and $\mathbf{\Phi}^{(0)}$ the initial parameter guess. The step width is controlled by α.

In *Newton's method*, it is assumed that χ^2 can be expanded around $\mathbf{\Phi}_0$ as

$$\chi^2(\mathbf{\Phi}) \approx \chi^2(\mathbf{\Phi}_0) + (\nabla \chi^2(\mathbf{\Phi}_0))(\mathbf{\Phi} - \mathbf{\Phi}_0) + \frac{1}{2}(\mathbf{\Phi} - \mathbf{\Phi}_0)\mathbf{H}(\mathbf{\Phi} - \mathbf{\Phi}_0). \tag{7.7}$$

Since $\nabla \chi^2(\mathbf{\Phi}) = 0$ is fulfilled for $\mathbf{\Phi} = \mathbf{\Phi}_0$, the iteration of Newton's method is given by

$$\mathbf{\Phi}^{(l+1)} = \mathbf{\Phi}^{(l)} \mathbf{H}^{-1} \nabla \chi^2(\mathbf{\Phi}^{(l)}). \tag{7.8}$$

This algorithm converges quickly if the initial parameters are chosen close to the minimum, whereas it is not appropriate far away from it. In the latter case, the steepest descent method is more adequate. A robust method combining both approaches is the *Levenberg-Marquardt method* [110, 111], which has also been applied and modified in the scope of this thesis and which is demonstrated in the following.

7.1.2 Levenberg-Marquardt Method

The Levenberg-Marquardt method can be considered an interpolation between the steepest descent and Newton's method. For this reason, the dimensionless factor λ is introduced as a weighting factor for the contribution of both methods. For computation of the step size of the gradient descent, the diagonal elements of the Hessian are used since they can be regarded as a measure for the curvature of the χ^2-hyperplane. If $\delta\mathbf{\Phi}$ denotes the iteration step, the steepest descent contribution to $\delta\mathbf{\Phi}$ is therefore given by the components

$$(\delta \Phi_\nu)_{\text{gradient}} = -\frac{1}{\lambda H_{\nu\nu}} \frac{\partial \chi^2}{\partial \Phi_\nu}. \tag{7.9}$$

Thus, high curvatures result in a low step width and vice versa. In order to add the contribution of Newton's method, the matrix $\mathbf{M}$ with elements

$$M_{\mu\nu} = \begin{cases} \frac{1}{2}\frac{\partial^2\chi^2}{\partial\Phi_\nu^2}\cdot(1+\lambda) & \text{for } \mu = \nu \\ \\ \frac{1}{2}\frac{\partial^2\chi^2}{\partial\Phi_\nu\partial\Phi_\mu} & \text{for } \mu \neq \nu \end{cases}$$

is introduced and the iteration steps of the Levenberg-Marquardt method are given by

$$\delta\mathbf{\Phi} = -\mathbf{M}^{-1}\nabla\chi^2. \tag{7.10}$$

Obviously, for $\lambda = 0$ this is equivalent to Newton's method, whereas for high values of λ, the steepest descent contribution becomes dominant. The value of λ is controlled dynamically: the algorithm starts with a high value (steepest descent) and decreases λ with decreasing values of χ^2. Details about this method are given in [112].

7.2 Parameter Estimation Approach for high-dimensional Systems

The χ^2-function of nonlinear systems with many unknown parameters typically contains many local minima. Consequently, local parameter estimation methods like the Levenberg-Marquardt method converge to the next local minimum and cannot guarantee to find the global minimum, unless the initial parameter guess is chosen close enough to it. Even global estimation methods like *Simulated Annealing* [113], which was also applied for the investigated signal transduction system (see study of Vacheva et al. [117]), are in general not capable of finding the global minimum in case the number of local minima is too large. Since this number usually grows with the space of possible parameter values and, as a consequence, exponentially with the number of unknown parameters, it becomes crucial to keep the number of unknown parameters as low as possible. In Chapter 6, an approach to reduce the dimensionality of the parameter space based on sensitivities and sensitivity correlations is given and applied for parameter estimation in the following.

The reduction of the system's dimensionality is spread over two stages. Prior to the start of the parameter estimation algorithm, an approach to cluster the system based on 'global' properties of the parametric sensitivities is introduced. This approach is complemented by a sensitivity-control based on local sensitivities and local sensitivity correlations within the parameter estimation algorithm.

7.2.1 Clusters-based Parameter Estimation

This approach takes advantage of the fact that clusters of state variables and parameters can be identified in such a way as to have subsets of state variables whose temporal

behaviour depends on a subset of parameters only (see Section 9.1.1). For simplicity and corresponding to the experimental situation of the signal transduction system of this thesis (Section 5.2.3), it is assumed that the observation functions $y_i^{\text{model}}(\mathbf{x}, t)$ are equivalent to state variables $x_i(t)$, which mostly correspond to molecule concentrations.

Now, a sensitivity matrix $\mathbf{S}$ consisting of the matrix elements $\{\hat{s}_{ij}\}$ (see Chapter 6) referring to m state variables and n parameters is considered. It is assumed that the property

$$|\hat{s}_{ij}| < \Theta, \quad \Theta = \Theta_s \cdot \langle s \rangle, \tag{7.11}$$

where $\langle s \rangle$ denotes the average sensitivity (averaged over all matrix elements) and Θ_s a low relative threshold, indicates that the respective state variable can be considered to be independent of the respective parameter. Whenever a high percentage of sensitivities fulfils this property - and this is the typical case in large signal transduction systems (e.g. Fig. 9.1) - clusters of the remaining above-threshold sensitivities can be established after reordering the rows of the sensitivity matrix as shown in Fig. 7.1. A cluster C_q is defined as a sub-matrix of the sensitivity matrix, whose elements cover only those parameters, whose influence on the state variables of the same submatrix, is not negligible. Thus, C_q consists of the matrix elements $\{\hat{s}_{ij}\},\ i \in \mathcal{I}_q,\ j \in \mathcal{J}_q$ according to

$$\begin{aligned} \mathcal{I}_q &= \{M_q, \ldots, M_{q+1} - 1\}, \\ &\quad M_q : \text{index of first state variable of cluster } q \text{ (Fig. 7.1)}, \\ \mathcal{J}_q &\subset \{1, \ldots, n\} \;\Big|\; j \in \mathcal{J}_q \Leftrightarrow \left(\exists\, i \in \mathcal{I}_q \;\Big|\; |\hat{s}_{ij}| > \Theta\right). \end{aligned}$$

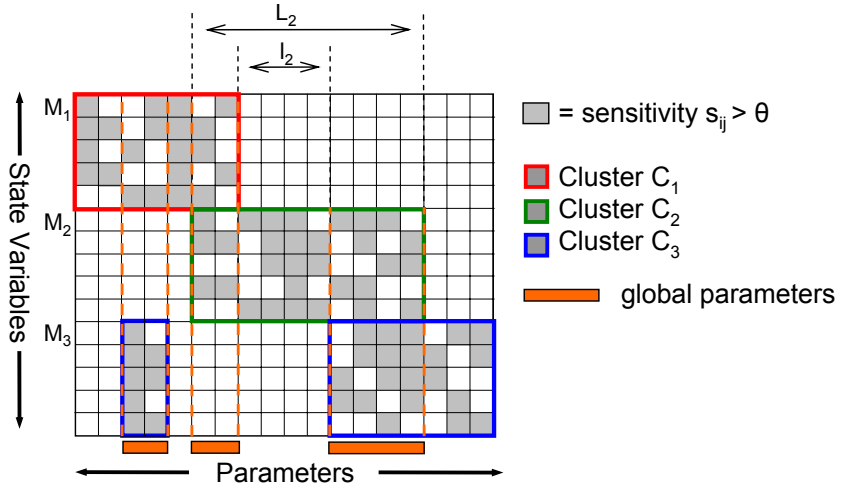

Figure 7.1: The sensitivity matrix exemplifies the clustered dependence of state variables (vertical axis) and parameters (horizontal axis). Red bars indicate the global parameters on which state variables of more than one cluster are dependent on. For cluster C_2, the parameter number L_2 and number of local parameters l_2 is shown.

Let C_q denote the q-th cluster and l_q the number of 'local parameters' of C_q. Local parameters are defined as those parameters whose sensitivities are below a threshold Θ for all state variables outside the cluster they belong to. If all parameters of one cluster influence state variables of the same cluster only, they can be estimated independently from all other clusters, leading to a parameter estimation of a much lower dimensionality. However, in general this is not the case and a cluster-wise estimation would determine the same parameters in the context of different clusters leading to inconsistencies. Therefore, parameters are split into local and global ones: a parameter, such that the state variables of more than one cluster depend on, is called a 'global parameter'. Consequently, the total number of global parameters, denoted by g, is given by

$$g = n - \sum_q l_q.$$

Hierarchical Parameter Estimation Approach

To address the estimation of local and global parameters, a hierarchical approach was designed, in which parameter estimation is performed on two levels. On the upper level, global parameters are estimated by optimising all clusters: for each cluster, parameter estimation is recursively called at the lower level. The χ^2-function of the upper level is based on the estimated parameters of the single clusters (lower level). On the lower level, the local parameters are estimated separately for each cluster, depending on the values of the global parameters proposed by the algorithm of the upper level, but independent of the parameters of all other clusters. If $W(d)$ denotes the cost for a parameter estimation of dimension d, the total cost W_{tot} of the algorithms can be compared as follows:

$$\begin{aligned} \text{unclustered algorithm:} \quad & W_{tot} \propto W(n), \\ \text{clustered algorithm:} \quad & W_{tot} \propto W(g) \cdot \sum_q W(l_q). \end{aligned}$$

Thus, the approach reduces the dimensionality from the number of all parameters n to the sum of the number of global parameters g and the maximum number of local parameters ($\max\{l_q\}$). Note that $W(d)$ strongly increases with dimension d, a fact also known as the curse of dimensionality. Some methods even show an exponential relationship [25]. As a consequence, this relation reflects a drastic reduction of computational costs whenever the relative number of global parameters g/n is low and the largest cluster is significantly smaller than the original matrix. These properties are typically fulfilled in signal transduction systems. To optimize the computation time for very large systems, it would be adequate to choose the clusters so that $g + \max\{l_q\}$ is minimized.

Since the clustering is applied as a basis for efficient parameter estimation and thus before parameter values are given, there is no basis to compute a sensitivity matrix, which is, however, required for the clustering. A solution to this problem is given by the 'global' sensitivity analysis approach demonstrated in Section 6.2. Since the clustering method only requires information about which sensitivities are below threshold Θ, the sensitivity

histograms (see also Section 9.1.2) are evaluated. The clustering is then performed by neglecting those sensitivities, whose histograms indicate that the property $|\hat{s}_{ij}(\mathbf{\Phi})| < \Theta$ is probable to be fulfilled within the complete space of possible parameters (Fig. 9.3).

In order to further reduce the dimensionality of the estimation problem, sensitivities, which are below a threshold only locally in parameter space, are taken into account within the parameter estimation algorithm. Moreover, local sensitivity correlations (see Section 6.3) are evaluated in order to detect highly correlated parameters.

7.2.2 Sensitivity-controlled Parameter Estimation

By integrating a sensitivity-control in the parameter estimation algorithm, a method is introduced, which adaptively reduces the dimensionality during the estimation. The local sensitivity matrix ($\hat{s}_{ij}$) and the derivations $\partial\chi^2/\partial\Phi_j$ are computed and evaluated after each iteration step. A subset of parameters, whose relative impact on the χ^2-function is significantly lower than the average impact of all parameters, is generated according to the criterion

$$\Big|\frac{\partial\chi^2}{\partial\Phi_j}\frac{\Phi_j}{\chi^2}\Big| < \Theta_\chi \cdot \Big\langle\Big|\frac{\partial\chi^2}{\partial\Phi_\nu}\cdot\frac{\Phi_\nu}{\chi^2}\Big|\Big\rangle, \tag{7.12}$$

where Θ_χ is a low relative threshold (e.g. $\Theta_\chi = 0.01$). The parameters belonging to this subset are kept constant for the next iteration step. This prevents the algorithm from being misguided by irrelevant parameters. Since the distinction between relevant and irrelevant parameters is generally valid for a specific region in the parameter space only, the sensitivities are automatically redetermined after each iteration step.

Parametric Sensitivity Correlations

In order to account for parameters with highly correlated sensitivities (see Section 6.3), the normalized Fisher Information Matrix $\mathbf{J}^{\mathrm{n}}$ defined in Eq. (6.20) is computed before each iteration step. On this basis, parameters are identified, which affect the system behaviour (almost) identically and only one of these parameters is varied for parameter estimation. Herefore, the parameter with the highest value of $|\frac{\partial\chi^2}{\partial\Phi_j}\frac{\Phi_j}{\chi^2}|$ is chosen. The criterion to select pairs of correlated parameters Φ_q, Φ_p is given by

$$|\mathbf{S}^{\mathrm{n}}_p \mathbf{S}^{\mathrm{n}}_q| \geq 1 - \Theta_{\mathrm{corr}},$$

where $\mathbf{S}^{\mathrm{n}}_j$ denotes the normalized sensitivity vector (Eq. (6.19)-(6.20)), and Θ_{corr} a low relative threshold (e.g. $\Theta_{\mathrm{corr}} = 0.01$). Note that only those parameters are checked for sensitivity correlations, which do not already fulfil the previous sensitivity criterion given by Eq. (7.12).

To apply the previous concept, the Levenberg-Marquardt algorithm was extended by the sensitivity- and sensitivity-correlation-control. A comparison with the unmodified algorithm shows that this combined method leads to significantly better parameter fits. The results are discussed in Section 9.2.3.

7.2.3 Ensembles of Estimated Parameter Sets

Due to the high nonlinearity of the χ^2-function and the large ranges of possible parameter values, it has to be taken into account that the χ^2-function may contain many local minima within the considered parameter space. Thus, parameter estimation algorithms are likely to find different local minima for different initial parameter sets, as indicated in Fig. 9.6 by the broad distribution of χ^2-values associated to the resulting parameter fits. In this thesis, it is suggested to combine parameter estimation methods with a random multi-start algorithm and to run a series of parameter estimations with different randomly chosen start parameter sets. The ranges of the random parameter values correspond to those used for the sensitivity analysis (see Section 6.2). As local parameter estimation method, the Levenberg-Marquardt algorithm extended by the previously discussed clustering and sensitivity control is chosen. Furthermore, the *Multiple Shooting* parameter estimation approach [114] using the MUSCOD software package [115] is applied (see study of Mesecke [116]). In order to exemplify global estimation algorithms, the simulated annealing method [113] is also examined in the study of Vacheva et al. [117].

The execution of many parameter estimation runs for the signal transduction application of this thesis confirms that parameter estimations with different initial parameter sets often lead to different local minima. This is also the case for the 'global' simulated annealing method. Moreover, the value of χ^2 associated to the best parameter fit is typically not significantly lower than the χ^2-values of the next best fits. Instead, a set of local minima with well-fitting parameters is found, whose χ^2-values are in the same range. A typical distribution of these values is shown in Fig. 9.6. As a consequence, it is suggested in this thesis, to consider ensembles of estimated parameter sets providing a good fit with experimental data rather than to focus onto one single parameter set, which provides the best fit. For this purpose, all estimated parameter sets, for which the χ^2-function is below a certain threshold, are selected. It is assumed that these parameter sets are elements of the submanifold of possible solutions in the parameter space, given the experimental data.

Next, the system behaviour simulated for the selected parameter fits is evaluated in order to gain information about the respective system properties and about their evidence. It is suggested to preselect all those system properties and predictions for further investigation, which are found to be (at least qualitatively) identical for all parameter sets of the ensemble. On this basis, model predictions are generated, which are subject to experimental validation. If these predictions fail, a wrong model choice is more likely than discrepancies in the parameters. On the other hand, all those system properties, which strongly vary between different parameter sets of the ensemble, are subject to further inves-

tigation and provide a basis for the design of new experiments [97, 118]. These experiments have to be planned in such a way as to reflect the investigated system properties in the χ^2-function. Typically, a sub-ensemble of the original ensemble of estimated parameters can then be identified, for which the respective model simulations reproduce the experimentally observed behaviour. If adequate parameter fits cannot be found anymore based on data including the new experiments, a wrong model choice is indicated.

Consider, for instance, the following example. It refers to the parameter estimation described in Section 9.2. For the fast activation scenario of the CD95-induced apoptosis (5 $\mu g/ml$ of ligand concentration), many parameter sets were found providing good fits with the respective experimental data (Fig. 9.4 A). Model predictions based on these parameter fits for a much slower activation scenario are, however, not consistent. For one part of the parameter fits, a slow increase of active caspase-8 is predicted, which starts immediately after stimulating the cell with ligands. For most other parameter fits, the model predicts no significant caspase-8 activity within a certain time period, followed by a second period, which is also characterized by an increase of active caspase-8. As a consequence, experiments for a slow activation scenario of 200 ng/ml were performed. The observed behaviour confirms a significant time delay before the activity of caspase-8 increases (Fig. 9.4 B). Thereby, a subset of the latter parameter set ensemble was invalidated. New parameter estimations were started based on experimental data referring to both scenarios. Good fits were found especially for those estimation runs using the delay-predicting parameter fits as initial values. Another, more systematic application of this concept is demonstrated in Section 9.4 in the context of the investigation of cell death rates.

Thus, a method is provided, which successively decreases the size of the solution submanifold in parameter space by iteratively incorporating data from new experiments, which are designed in a way to maximize the gain of information. As an extension to the concept of parameter set ensembles referring to the same model, alternative models can also be taken into account.

Chapter 8

Software Framework

In this chapter, the software package is described, which I developed to integrate the modelling, simulation and system identification methods presented in Chapter 3, 5, 6 and 7 in one common framework. An important requirement for this software was high flexibility and extensibility regarding the size and the function of user-defined network models. Furthermore, the framework had to be designed in a way, which is extensible to different numerical components like parameter estimation methods or ODE solvers. Existing software packages for the simulation of biochemical networks, like Gepasi, ECELL, STOCHSIM and many others [18, 19, 35], are stand-alone tools, which do not allow user-defined parameter identification methods, user-defined solvers, functional subunits, or integration of new sensitivity analysis approaches, as it was applied in this thesis.

As a result, a modular and object-oriented software package was implemented in C++, which is capable of running any user-defined network model of nearly unlimited sizes (only limited by memory). For this reason, the software contains a model-generator, which automatically translates a user-defined model file (see Appendix A), containing the reaction network and functional subunits (black-boxes), if required, into a system of differential equations. For stochastic simulation, the probability density functions and stochastic differential equations (Section 2.2) are automatically generated as well. As an example for the modular software design, the simulated annealing method [113] has been plugged in as a global parameter estimation method by Dr. Ivayla Vacheva [117].

8.1 Functionality

The software can be run in seven different modes. This run mode, but also the numerical settings, the model definition, the parameters and the experimental data (parameter estimation only) have to be provided by the user in form of an input file, which is described in Appendix A. In the following, the functionalities of the different run modes are described.

8.1.1 Software Modes

Single-Simulation Mode. A single simulation run is performed by either using the deterministic method, the Gillespie algorithm, the stochastic-deterministic algorithm or the General Stochastic Hybrid Method. The simulation result is provided in the form of a time series of all state variables.

Multiple Stochastic Simulations. A series of independent stochastic simulations either using the pure Gillespie, the stochastic-deterministic, or the general stochastic hybrid method are run. The output contains a time series of the mean values and standard deviation of each state variable.

Local Sensitivity Analysis. Local sensitivity analysis is performed for a given parameter set. The output provides the sensitivity matrix.

Sensitivity Histogram Generation. A series of random parameter sets is generated within given value ranges. For each parameter set, all sensitivities are computed and stored in a data structure, based on which sensitivity histograms, a matrix of average sensitivities and a matrix of the standard deviation of the sensitivities are generated. Optionally, a Boltzmann factor based on the difference between simulation and experimental data is applied for the generation of the sensitivity histograms. In this case, time series of experimental data have to be provided through the input file.

Parameter Estimation. The parameter estimation is run for a given initial parameter set according to Section 7.2. Time series of experimental data have to be provided through the input file. Within the algorithm, a sensitivity matrix and a sensitivity correlation matrix is computed. The output contains the results after each iteration step, including the parameter values, the χ^2-function (Section 7.1), the sensitivities, the final parameter fit, and, optionally, the parametric sensitivity correlation matrix (normalized Fisher information matrix, Section 6.3.1).

Random-Start Parameter Estimation. A series of random parameter sets is generated within given ranges of parameter values. For each parameter set, the parameter estimation algorithm is started as described in the previous mode. Time series of experimental data have to be provided. The output contains the resulting parameter values and the corresponding χ^2-functions. The parameter values are provided in a format, which can be directly reused in the input file, e.g. for further simulation.

Cluster-based Parameter Estimation. The cluster-based parameter estimation method as defined in Section 7.2.1 is run. Here, hierarchical clusters of parameters and state variables have to be defined as well as the global parameters. Time series of experimental data have to be provided. The output contains the results after each iteration step, including parameter values and χ^2-function.

Moreover, user-defined parameter estimation methods can be plugged in. All other components of the software can thereby be accessed, especially the simulation for a certain parameter set, the manipulation of parameter values and the generation of the χ^2-function.

Parallel computing is enabled for the generation of sensitivity histograms and the parameter estimation for randomly chosen initial parameter sets, since these two modes are extremely time-consuming.

8.1.2 Parameter Estimation based on multiple experimental Scenarios

An important feature of the parameter estimation framework is the parameter estimation based on multiple scenarios. I.e. experimental data referring to different initial conditions or different experimental settings can be incorporated and the parameters are estimated to optimally fit all scenarios. The weighting of the respective experimental data is defined by the standard deviation and by the number of data sets. Whenever the simulation is called within the parameter estimation algorithm, it is executed for each of the different scenarios. All simulated scenarios are based on the same model and the same set of parameters except for the parameter(s), by which the scenarios differ (Fig. 8.1). In case of CD95-induced apoptosis, experimental values were provided for different activation scenarios, corresponding to different initial ligand concentrations.

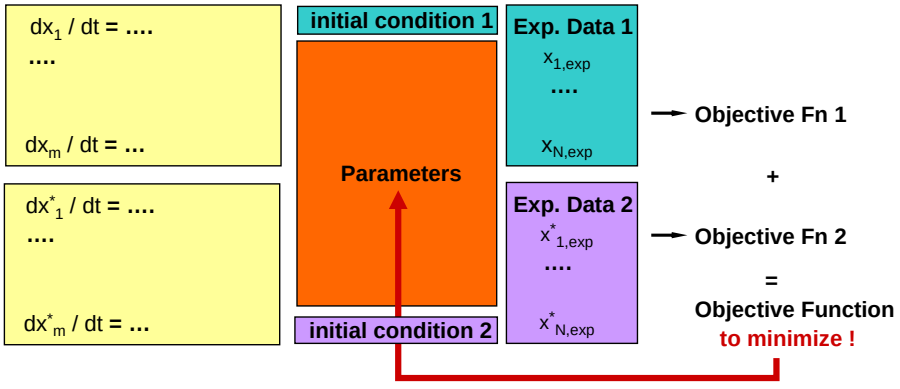

Figure 8.1: Example for parameter estimation based on two scenarios, which only differ in the initial conditions. Experimental data are provided for both scenarios and the objective function (χ^2-function) is computed accordingly.

8.2 Software Structure and Components

An overview about the software structure and the workflow is given in the flowcharts Fig. 8.2, Fig. 8.3 and Fig. 8.4. The first figure focuses on the core elements, in particular the model generation and the simulation. The other flowcharts describe the higher level processes for all possible modes, which were given in Section 8.1, like single or multiple (stochastic) simulation runs, sensitivity analysis or parameter estimation.

Automatic Model Generation

Based on the input data consisting of state variables, model definition (reaction network and functional subunits), parameters and the definition of different scenarios, where applicable (Section 8.1.2), the model is generated in the form of internal data structures for state variables and differential equations, called Differential Equation Definition Objects (DEDO). If parameter estimation is based on multiple scenarios, the model is multiplied, which means that parallel models are generated, which differ according to the different scenarios.

Model Simulation

The simulation either runs in the deterministic, stochastic, deterministic-stochastic, or in the general stochastic hybrid mode. For parameter estimation and sensitivity analysis, only the deterministic mode is used. The solver accesses the DEDO for propagation of the system states. The reaction rates, on which the differential equations depend, are also used for computation of probability density functions in the stochastic case. In addition, the black boxes are called in order to add their contribution to the time evolution. The solver contains an adaptive step size control.

Generation of Random Numbers

Random numbers is required for stochastic simulations as well as for the generation of random parameter sets, e.g. for parameter estimation. The generation of random numbers of high quality is thereby crucial (see e.g. [120]). Here, the *Mersenne Twister* random number generator [121], which is based on a *generalized feedback shift register generator* [122], is applied. In case of parallel computing, as it was performed for parameter estimation and computation of sensitivities using randomly chosen parameter sets, the jobs were started with different sequences of random numbers in order to prevent correlation.

Time Step Constraints for ODE Solver

The solver can be forced to stop at a sequence of specified time points for the following reasons. In case of sensitivity analysis, the integral of absolute values of the differences between simulation results for slightly varied parameters have to be computed (Section 6.1.2). For parameter estimation, the solver has to stop at the time points, for which

experimental data are given in order to evaluate the simulation data and in case of multiple stochastic runs, the mean and standard deviation of state variables is evaluated for a sequence of distinct time points.

Integration of spatial Information and Diffusion Processes

As a possible future extension, simulation of spatial distribution of molecules can be integrated. Therefore, the reaction system has to be replaced by a diffusion-reaction system. For the software, the following extensions would be required:

- The state variables $x_i(t)$ describing molecule concentrations under the assumption that the system is a 'well-stirred mixture' have to be replaced by molecule densities $x_i(\mathbf{r}, t)$ depending on the time t and position $\mathbf{r}$.
- The diffusion system (system of partial differential equations) has to be discretised and transformed into a system of ordinary differential equations. The molecule changes due to diffusion and due to reactions have to be added.
- Instead of one value for the initial molecule concentration x_i^I for each species, the initial spatial distribution $x_i^I(\mathbf{r})$ has to be provided, e.g. in the form of images or image stacks containing the spatial information about molecule concentrations.
- For parameter estimation, simulated data $x_i(\mathbf{r}, t)$ have to be compared with image files from experiments.

In summary, the software provides a general framework for the simulation and system identification of large, user-defined reaction networks. The software package consists of more than 10000 lines of code, organized in 25 classes and about 250 subroutines. A high degree of flexibility is provided by the definition of black boxes. A description of the input files required to run the software is given in Appendix A.

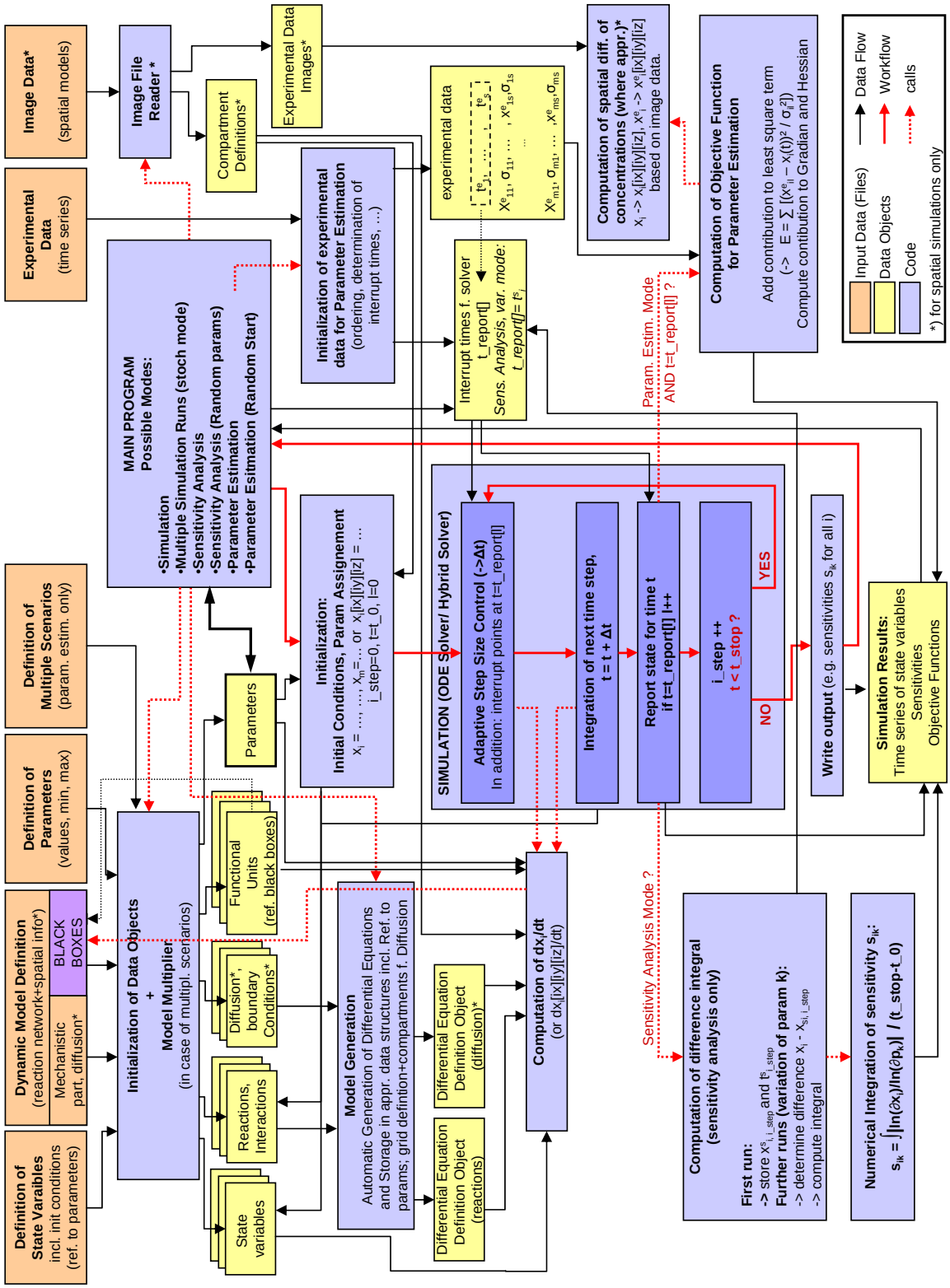

Figure 8.2: Software Structure. Pseudocode presented as flowcharts of the software framework. Detailed description of the simulation part, which is called by the main program running in one of seven possible modes.

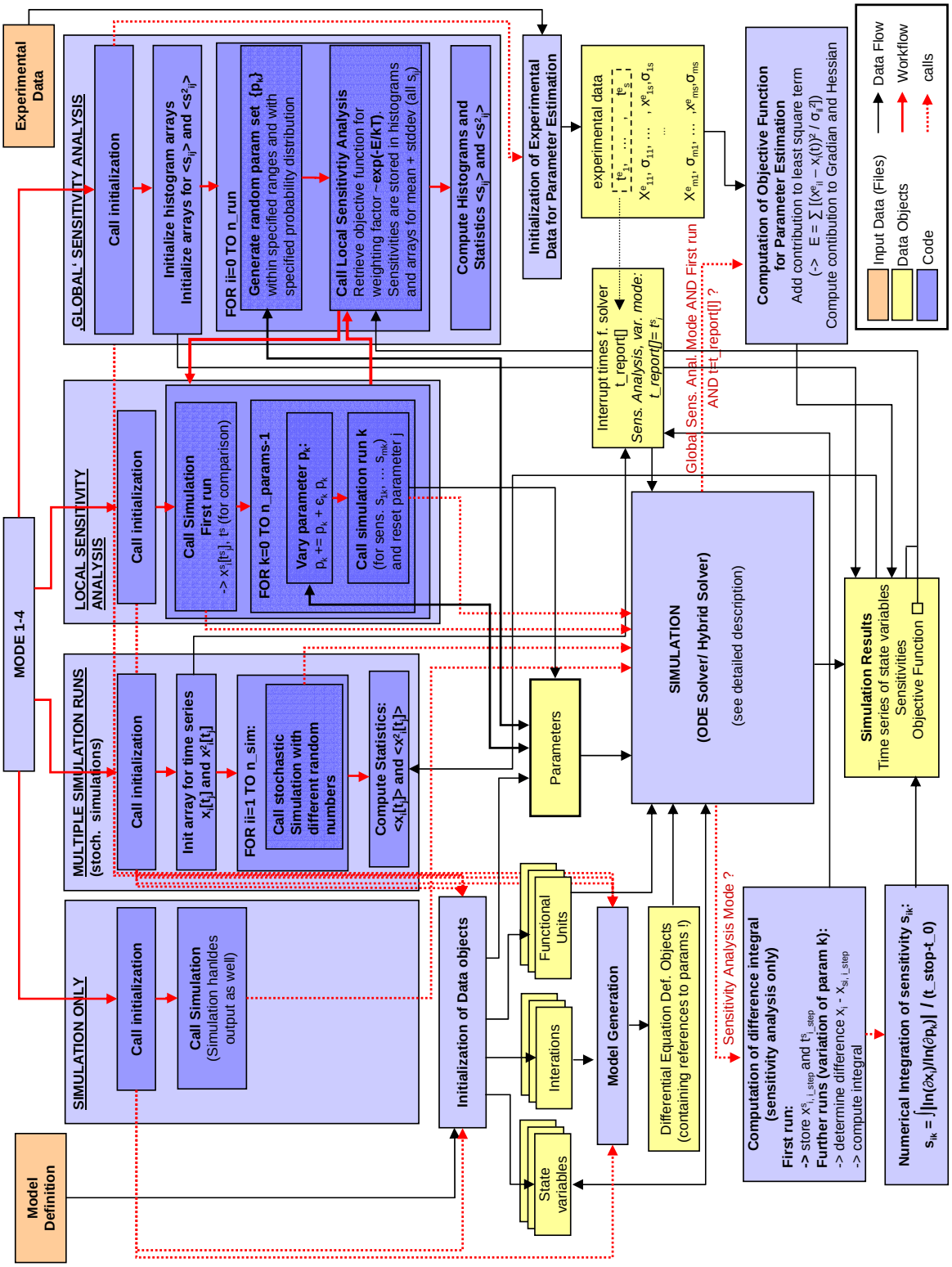

Figure 8.3: Flowchart for the main program running in one of the following modes: simulation only, multiple stochastic simulation runs, local and 'global' sensitivity analysis.

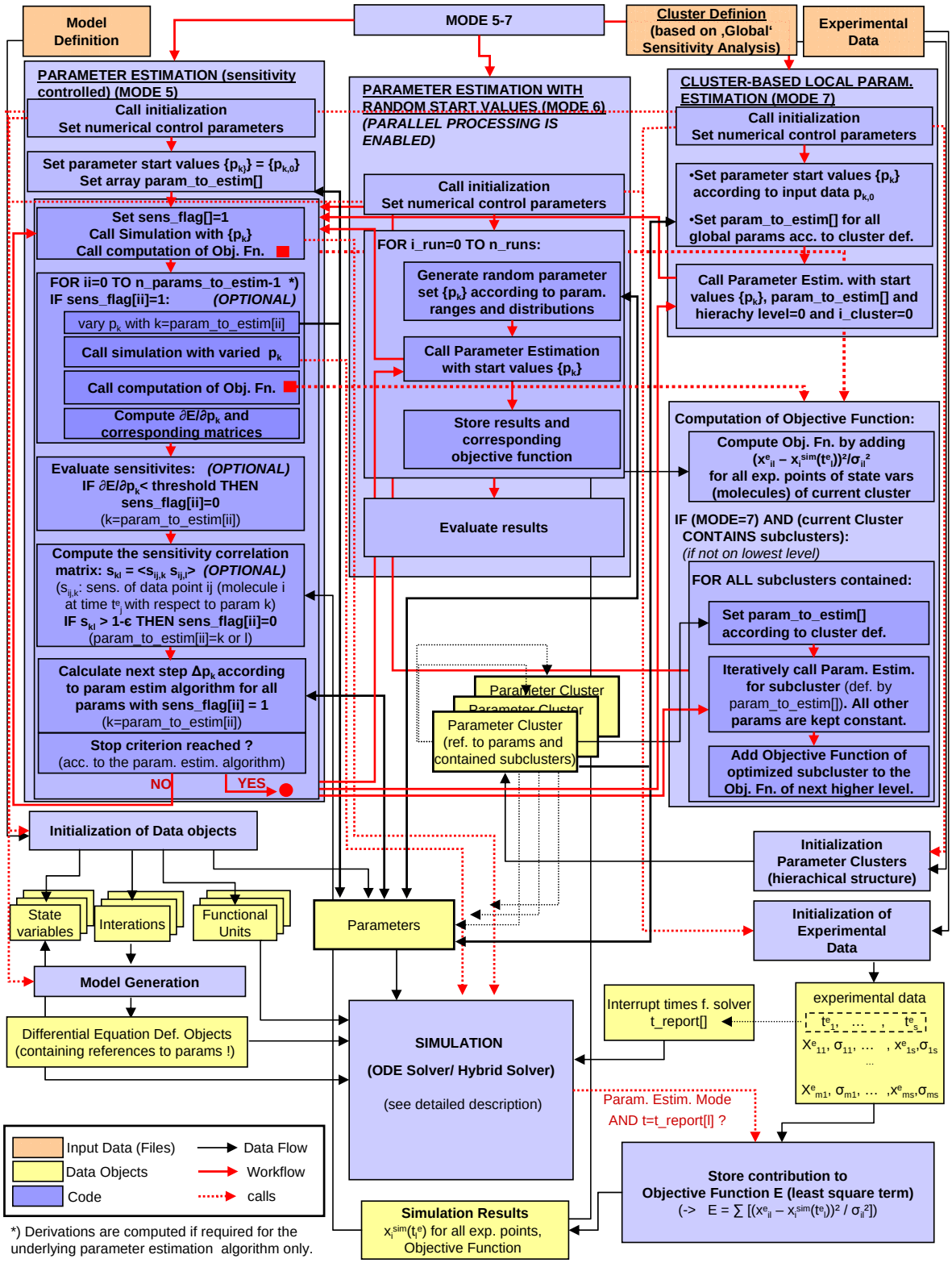

Figure 8.4: Flowchart with focus on parameter estimation running in one of three modes: fixed initial parameters, stochastic initial parameters and cluster-based approach.

Chapter 9

Results

The framework presented in Chapter 5, 6, 7 and 8, was applied to programmed cell death, in particular to CD95-induced apoptosis, which constitutes one of the most important signalling pathways in all higher organisms (Chapter 4). Controversially discussed questions like a possible threshold mechanism for life and death or the point of no return were thereby addressed.

The chapter is structured as follows. After model generation based on structured information models, the results of the stochastic sensitivity analysis approach I developed in Section 6.2 are demonstrated, especially the important fact that the 'Sensitivties of Sensitivities' are extremely low. On this basis, the parameter estimation method I derived in Section 7.2 is applied, thereby significantly reducing the dimensionality of the parameter identification problem. Furthermore, it is shown that the sensitivity-controlled estimation method finds significantly better parameter fits and converges much faster then the unmodified method. Then, it is demonstrated that the model with the estimated parameters is capable of reproducing experimental observations referring to different scenarios and of predicting biologically important system properties. The predictions were experimentally verified by our biological collaboration partners and detailed hypotheses about the exact mechanisms are generated. Finally, it is demonstrated that for certain scenarios, the system is mainly governed by stochastic effects due to low particle numbers of critical molecules. This suggests the use of a stochastically exact simulation method (Chapter 2.2.3). Because of the heterogeneity among the reaction time-scales, the General Stochastic Hybrid Method I developed in Chapter 3 was applied. Thereby, the observed stochastic features are reproduced.

Structured information model of CD95-induced apoptosis

The network topology of CD95-induced apoptosis was reconstructed by critically searching databases and the literature, resulting in a model with about 60 molecules and more than 120 unknown parameters (Fig. 4.2). Based on this topology, the concept of 'structured information models' was applied (Section 5.1.1). The core elements of the pathway

(e.g. DISC, caspases) were modelled mechanistically, whereas the remaining network was described by simplified functional subunits (see Section 5.3). As a result, a model of CD95-induced apoptosis consisting of 41 molecules (or molecule complexes), 32 interactions and more than 50 missing parameters was established. This model system is still too large to allow reliable parameter estimation given the limited number of experimental data points (Section 9.2).

9.1 Reduction of System Complexity by Sensitivity Analysis

For further reduction of complexity, the sensitivity analysis approach presented in Chapter 6.2 was applied. Since in general, sensitivities can be determined for specific sets of parameter values only (local sensitivity analysis), the usefulness of sensitivity analysis is limited if most parameters are unknown at first. For this reason, sensitivities were determined by 'randomly walking' through the parameter space. In a virtual experiment, local sensitivities $\{\hat{s}_{ij}(\mathbf{\Phi}^q)\}$ (see Eq. 6.8), $q = (1, \ldots, N)$, were computed for $N \approx 300000$ randomly chosen parameter sets $\{\mathbf{\Phi}^q\}$ requiring huge computational effort. Hereby, the parameter value ranges, covering up to three orders of magnitude, were set up according to the typical values for the respective parameter types (e.g. Michaelis-Menten constants, bimolecular reaction rates or initial concentrations). Information about these values was taken from the literature (e.g. [49]). Thus, for each parameter set $\mathbf{\Phi}^q$, a sensitivity matrix $\mathbf{S}(\mathbf{\Phi}^q)$ was computed.

9.1.1 Sensitivity Matrix

Although, in general, the sensitivities depend on $\mathbf{\Phi}^q$, sensitivity matrices (see Fig. 9.1 for an exemplary parameter set) typically show two important properties:

- Sensitivities are low in general ($|\hat{s}_{ij}| \ll 1$ for most (i, j)) indicating high robustness and the possibility of lowering the 'effective' dimensionality of the parameter space (see Section 6.2.2).
- Apparently, modular structures can be identified, i.e. clusters that contain a subset of molecules, whose concentrations depend on a subset of parameters only. This inherent system property is an important feature for simplification of the parameter estimation problem.

Furthermore, the sensitivities of state variables, describing the utmost downstream processes that capture the main function of the signalling network, are extremely low, as shown in Fig. 9.2. The execution of programmed cell death is mainly governed by the executioner caspases, affecting the cleavage of PARP and the 'apoptotic activity' (see Section 5.2.2). Once the executioner caspases become activated, the process of cell death cannot be stopped anymore. Thus, low sensitivities of these state variables indicate a high

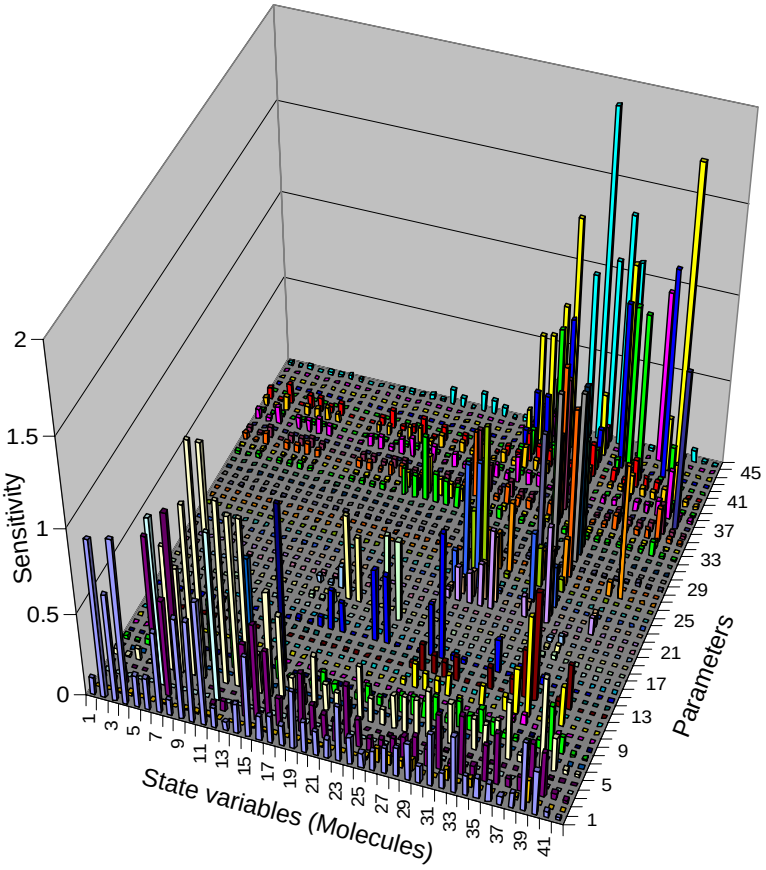

Figure 9.1: The sensitivity matrix of relative sensitivities $\{\hat{s}_{ij}\}$ shows the relative changes of each state variable i (left to right), mostly referring to molecule concentrations, with respect to changes of each parameter j (front to back). For simplicity, the absolute values $|\hat{s}_{ij}|$ of the relative sensitivities are displayed. The indices refer to the tables in Appendix B.1. The sensitivities visualized here were computed for the parameter set providing the best fit with experimental data as a result of parameter estimation. For different parameter sets, however, qualitatively similar structures were found.

robustness of the system's core functionality with respect to parameter fluctuations. This is a biologically interesting result. It indicates that in the case of CD95-induced apoptosis the assumption that complex molecular networks operate reliably under changing system parameters (see Section 6.1.4) is also fulfilled.

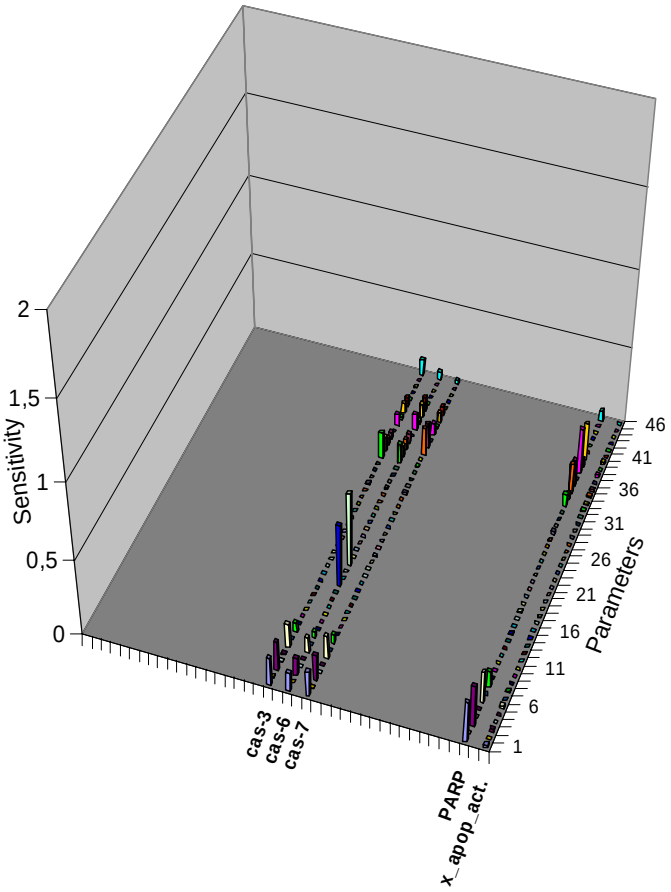

Figure 9.2: This diagram shows the sensitivities of Fig. 9.1 for the executioner caspases, namely caspase-3, -6, and -7, for PARP, and for the 'apoptotic activity' only. The comparatively low values indicate high robustness of the key functionality.

9.1.2 Sensitivity of Sensitivities

In order to gain information about sensitivities within the complete space of possible parameter values, the distribution of each sensitivity $\hat{s}_{ij}$ is plotted in the form of histograms for all random parameter sets $\{\mathbf{\Phi}^q\}$. Additionally, the Boltzmann factor $\exp(-E/kT)$, which was introduced in Section 6.2.3, was applied in order to amplify the statistical impact of sensitivities for those parameter sets that are more consistent with the experimental observations and consequently more probable.

The Sensitivity of Sensitivities is extremely low

The most crucial outcome of this 'global' sensitivity analysis approach is the fact that 'Sensitivities of Sensitivities' are extremely low in most cases. This is shown in Fig. 9.3 in the form of some exemplary sensitivity histograms. They were chosen in such a way as to

represent typical histograms of the sensitivities $\hat{s}_{ij}$ for all i, j. Obviously, most distributions show distinct and narrow peaks, indicating high robustness of the sensitivity values with respect to large variations of the parameter values.

Whenever the sensitivity of a sensitivity $\hat{s}_{ij}$ is extremely low and in addition, the distribution shows a sharp peak at zero (or extremely close to zero), an indication is given that the state variable i is likely not to be influenced by parameter j, regardless of the exact parameter value set. In order to gain additional information, the weighting factor $\exp(-E/kT)$ was applied and the system was 'cooled down' by decreasing the value of kT, starting from $kT = \infty$. Note that $kT = \infty$ corresponds to the case, in which all parameter samples are equally weighted, regardless of the differences between the respective simulation data and experimental data (Fig. 9.3, blue plots). The Y-axis of the histograms therefore corresponds to the number of times this sensitivity occurs within the N parameter sets. In the second case (red plots), the contribution of each parameter sample to the histogram is weighted according to the Boltzmann factor, which lowers the impact of parameter samples according to the inconsistency with experimental data. In this case, most peaks become sharper (e.g. Fig. 9.3 R, X, Z). Histograms, whose distributions are broad in the equally weighted case often turn into histograms showing peaks as well (e.g. C, D, E) as a consequence to the weighting factor. In most cases, only one peak is left over. This is a strong indication that, within the subset of samples, whose parameters result in system behaviour more consistent with experimental data, sensitivities are even less sensitive. In some cases however, the number of peaks increases as a consequence of the weighting factor (e.g. L, P, O), indicating that the system runs in different modes for certain areas of the parameters space. In this case, the parameter dependency of the respective sensitivity cannot be neglected, even if the sensitivity histogram shows one strong peak in the equally weighted case. Thus, it might be misleading to regard equally weighted distributions only, since small subsets of parameter samples are likely to be neglected although they might be highly relevant since they fit the experimental data very well.

In summary, it was shown that the novel sensitivity approach applied here for the first time is well-suited to gain important information about general parameter dependencies. As a consequence of the low sensitivity of sensitivities, subsets of parameters can be determined, which are unlikely to influence certain state variables. Considering the high number of sensitivities with $\hat{s}_{ij} \approx 0$, this step is crucial to reduce the system dimensionality if knowledge of the true parameter values is missing. Thus, the method provides a good basis for high-dimensional parameter estimation.

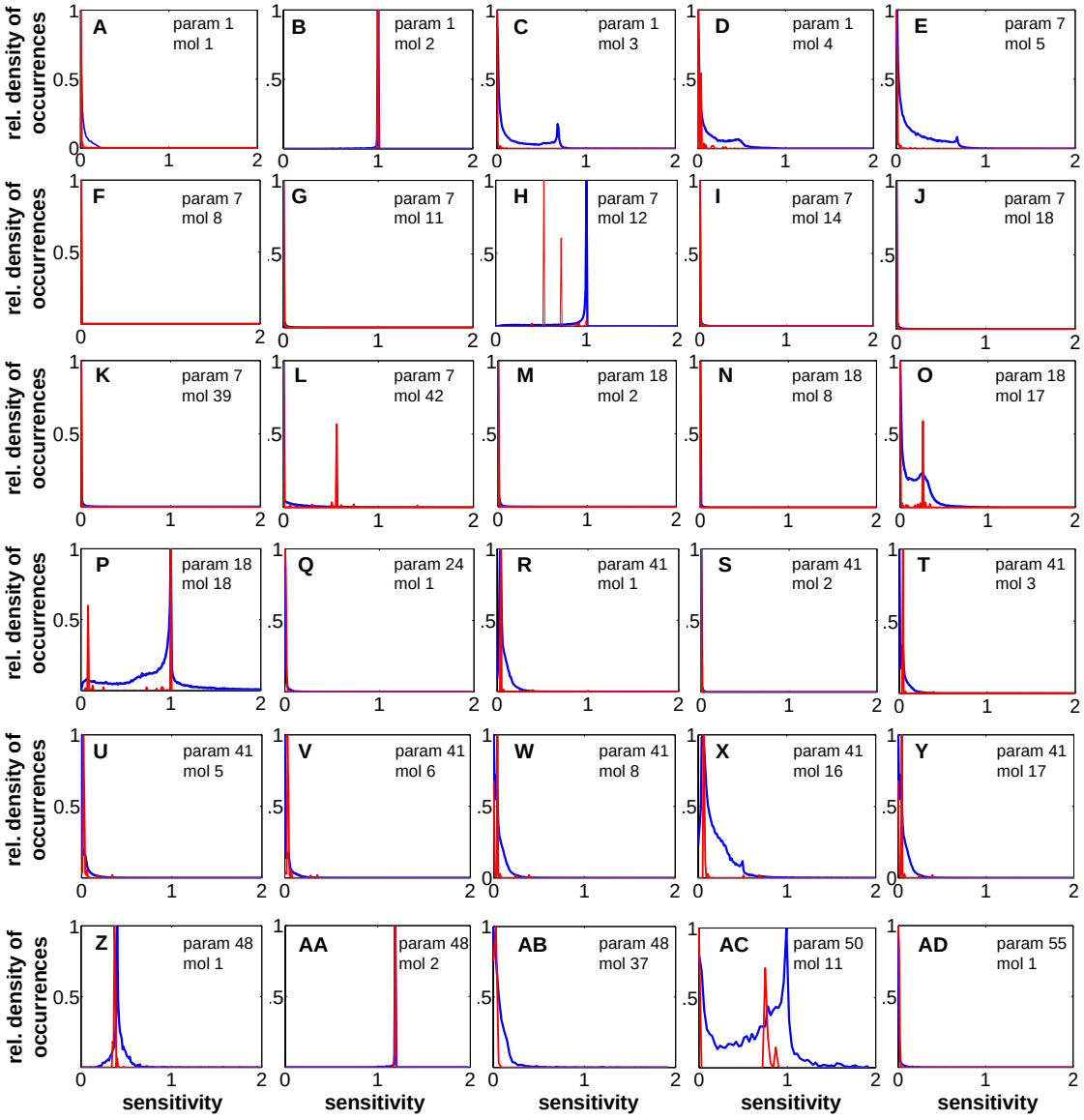

Figure 9.3: Each box shows a histogram for the absolute value of a specific relative sensitivity, $|\hat{s}_{ij}|$, computed for a large number of randomly chosen points in the parameter space. Parameter and molecule indices refer to the tables in Appendix B.1. The X-axis represents the relative sensitivity values from 0 to 2 and the Y-axis corresponds to the density of occurrences. The blue plot shows the equally weighted distribution of sensitivities, whereas for the red one, each contribution was weighted with the Boltzmann factor, resulting in sharper and sometimes slightly shifted peaks. The histograms are exemplary for the sensitivities of all i, j. Many histograms show clear peaks close to zero - an important property for further modularization. However, distributions like O, P, X, AB or AC indicate that the respective sensitivities are not robust with respect to large parameter variations and therefore not informative for modularization.

9.2 Parameter Estimation and Experiments

In the sensitivity matrix, clusters are identifiable that contain a subset of molecules whose concentrations depend on a subset of parameters only (Fig. 9.1). Consequently, the cluster-based parameter estimation approach presented in Section 7.2.1 was applied. Therefore, the matrix of sensitivities averaged over all parameter samples $\{\mathbf{\Phi^q}\}$ was evaluated and all (i, j)-combinations with negligible sensitivities were pre-selected. For these combinations, histograms were evaluated and only those combinations were chosen, whose histograms indicated that the respective sensitivity value is likely to be negligible in the complete space of possible parameter values. On this basis, the clustering was performed. As a result, the sensitivity matrix could be subdivided a priori into 4 clusters, which contain 16 global parameters. Thereafter, the effective system dimensionality was further reduced by introduction of the adaptive sensitivity control within the parameter estimation algorithm (see Section 7.2.2). Thus, the parameter identification problem reached a much lower dimension.

9.2.1 Experiments for probing Regulatory Mechanisms of CD95-induced Apoptosis

A set of experiments to measure time series of concentrations of 14 different molecules after activation of CD95 receptors was designed (see Fig. 5.3, red frames). For the experiments, the human B-lymphoblastoid cell line SKW 6.4 was chosen [123], previously classified as type I cells by their high amount of DISC formation (see Section 4.1). These cells are highly sensitive to CD95-mediated apoptosis. Cells were stimulated with different concentrations of agonistic anti-APO-1 antibody or LZ-CD95L, which can both be considered CD95 ligands (Section 5.2), for various periods of time (from 5 minutes to 4 days). Each sample was evaluated by three independent approaches. Cell death was determined by flow cytometry [124], caspase activity was measured by fluorometric activity [124], and the change of concentration of major apoptotic molecules was evaluated by western blot [124]. For all measurements, standardization of experiments was crucial. The experiments were performed by our collaboration partner Inna Lavrik [34]. Note that quantitatively reliable measurements in this area require a huge effort. See Appendix C for details.

In a first set of experiments, time series were measured for a 'fast' activation scenario with an oversaturated ligand concentration corresponding to more than one ligand per CD95 receptor. Oversaturation was achieved by 5 $\mu g/ml$ anti-APO-1 corresponding to a ligand-receptor ratio of about $5:1$. The ratio was determined under the assumption that there are approximately 40000 CD95 receptors per cell. This number was estimated from measurements by flow cytometry.

9.2.2 Parameter Estimation based on different scenarios

A good fit between model simulation and experimental data could be achieved reproducing the fast cleavage of procaspase-8 into its active form via the intermediate product p43/p41, followed by activation of the executioner caspases and cleavage of Bid and PARP (Fig. 9.4 A,C,D). Due to the oversaturation of receptors, the process is very fast and PARP gets almost completely cleaved after 10 minutes, followed by the degradation of all observed proteins. The mathematical model is well-suited to quantitatively describe the activation of CD95-induced apoptosis. Moreover, the black box representation of the complex degradation process, reproduced by a simplified decay-function, matches the experimental data very well.

However, the model is still underdetermined and many different parameter sets are able to match the same experimental data. Accordingly, generalization of the model for biological predictions is likely limited. In order to gain additional information about the system, a different activation scenario with a lower initial ligand concentration was measured and the parameter estimation was based on these multiple conditions as described in Section 8.1.2. Thus, an integrated model including different activation scenarios was automatically generated. The integrated model is based on a common set of biochemical parameters but different initial values of the ligand concentration. As a result of the combined parameter estimation approach, the model fits several activation scenarios for the different ligand concentrations (Fig. 9.4 A-E).

This combined parameter estimation problem required high computational effort. The Levenberg-Marquardt algorithm (Section 7.1.2), which was mainly applied here, was run for several thousand times with randomly chosen initial parameter values in order to find parameter sets fitting the activation scenarios. Without prior system reduction based on sensitivity analysis, the same algorithm did not provide any adequate parameter fit.

9.2.3 Evaluation of Sensitivity-controlled Parameter Estimation Algorithm

By integrating a local sensitivity control in the Levenberg-Marquardt algorithm (Section 7.2.2), the number of parameters, which are considered relevant, could further be reduced by $\approx 50\%$ due to extremely low sensitivities in the respective areas of the parameter space. Moreover, correlated parameters (Section 6.3) were detected. In order to show typical parameter sensitivity correlations, the normalized Fisher Information matrix $\mathbf{J}^n$ (Section 6.3) is visualized in Fig. 9.5 for an exemplary parameter set. These parameter sensitivity correlations were evaluated and taken into account after each iteration step.

The sensitivity-controlled and sensitivity-correlation-controlled algorithm was then compared with the original Levenberg-Marquardt algorithm. For this reason, both methods were run for a series of random start parameter sets. The distribution of the resulting

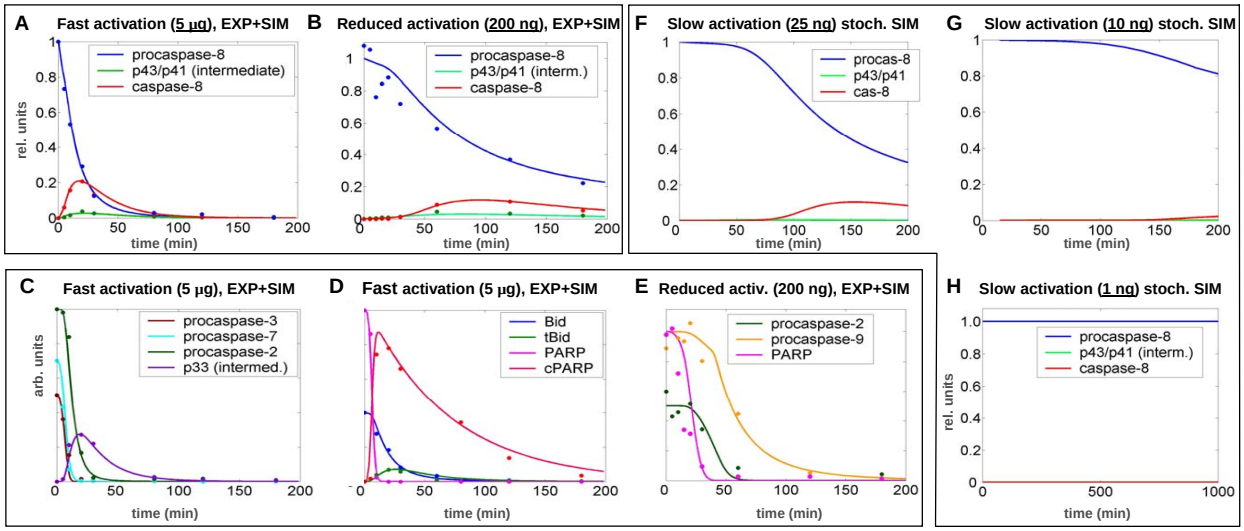

Figure 9.4: **A-E**: Parameter estimation on the basis of two activation scenarios with ligand concentrations of 5 $\mu g/ml$ and 200 ng/ml led to a good fit between model simulations (solid lines) and experimental data (dots) for both scenarios. A: The high ligand concentration leads to an early activation of receptors, followed by fast DISC formation, resulting in a high cleavage capacity of procaspase-8 via the intermediate product (p43/p41). C,D: Early generation of active caspase-8 is followed by the cleavage of caspase-3, -7, and -2 as well as by cleavage of Bid and PARP. After PARP cleavage, decomposition of cellular components starts. B,E: Simulation of slower activation (200 ng/ml) using the same set of biochemical parameters. Due to the smaller percentage of receptors activated by ligands, the capacity of caspase-8 cleavage is much lower. However, there is still a cleavage of 100% of the executioner caspases and PARP, resulting in apoptosis. **F-H**: Model prediction for a series of lower ligand concentrations (25 ng/ml, 10 ng/ml and 1 ng/ml) using the previously estimated parameter set. Note that for the latter activation scenarios stochastic instead of deterministic simulations (see Section 9.4) were applied. As expected, procaspase-8 cleavage slows down (F, G). However, for 1 ng/ml (H) the death process is completely stopped. Note that in F and G, the procaspase-8 cleavage starts after a significant delay. Y-axis: A, B, F, G, H: relative units since the (pro-) caspase-concentrations are directly comparable; C, D, E: arbitrary units. The standard deviation of experimental data was $\approx 20\%$ on average (see Appendix C.2).

χ^2-values (see Section 7.1) and the computation times is visualized in Fig. 9.6. Here, the χ^2-function, which corresponds to the sum of the squares of differences between experimental and simulated data, divided by the standard deviation, is used to quantify the quality of the respective parameter fits. Obviously, the sensitivity-controlled algorithm was capable of finding significantly more parameter sets with low χ^2-values. Furthermore, the modified algorithm required significantly less computation time (about 80% reduction).

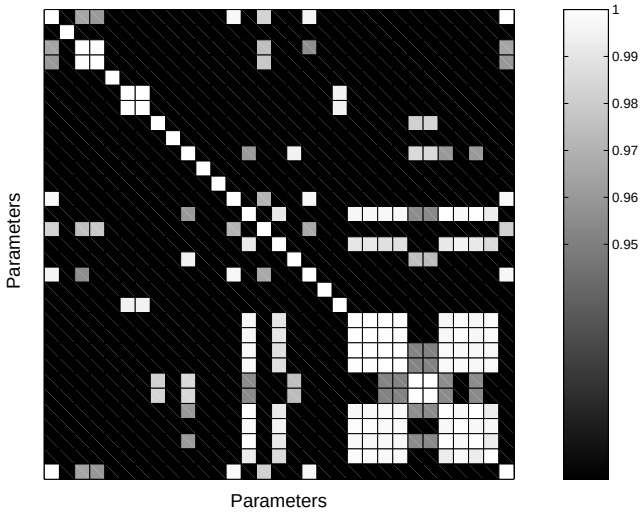

Figure 9.5: Normalized Fisher Information matrix $\mathbf{J}^n$ for an estimated parameter set. Here, the absolute values $|J_{ij}^n|$ are displayed for those matrix elements, which fulfil $|J_{ij}^n| > 1 - \Theta_{\text{corr}}$, $\Theta_{\text{corr}} = 0.05$. Parameters, whose relative sensitivities are extremely low with respect to all state variables, are already taken out of scope.

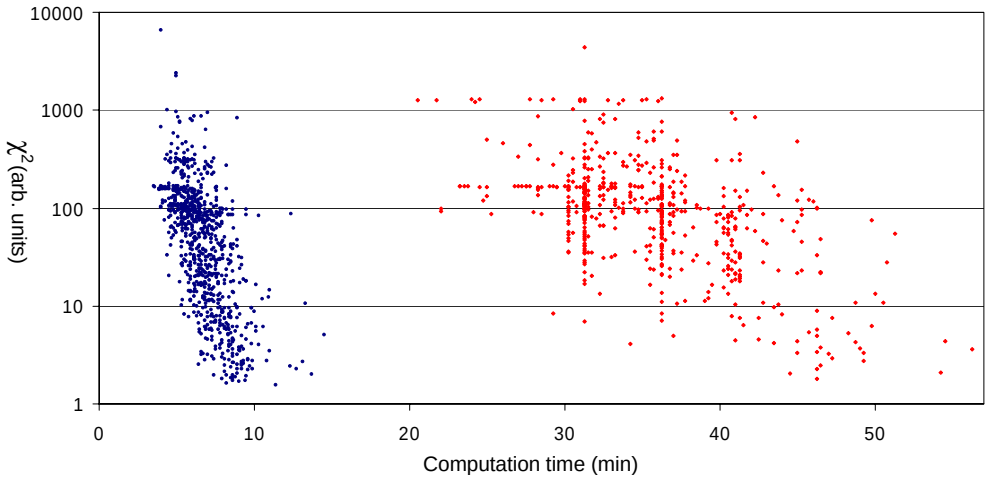

Figure 9.6: The performance of the sensitivity-controlled Levenberg-Marquardt algorithm (blue points, left) is compared with the original algorithm (red squares, right). Each point represents the result of one parameter estimation, started with a random initial parameter set. Y-axis: Value of the χ^2-function of the resulting parameter set. X-axis: Computation time for the respective parameter estimation run. The sensitivity control also checks for correlated parameters. In this application, 'good' parameter fits correspond to a χ^2-value of below 10.

9.3 Model Prediction and Experimental Validation

In the following, model predictions, generated on the basis of the estimated parameters, are presented. Since several parameter sets were found, which provide good fits with similar χ^2-values (see Fig. 9.6 and Section 7.2.3), only those predictions were taken into account, which were at least qualitatively consistent for all of these parameter sets[1].

Both the model and the experimental data show that with decreasing ligand concentration apoptosis is slowed down considerably (Fig. 9.4 A-E). However, cell death is still achieved. To address the question whether the apoptotic process slows down continuously with lower ligand concentrations or whether there is a threshold for induction of apoptosis at a distinct receptor-ligand ratio, induction of apoptosis was simulated for very low ligand concentrations (see Fig. 9.4 F,G,H). The model predicts that below a critical concentration corresponding to a ligand-receptor ratio of approx. $1 : 10^2$, apoptosis is completely stopped (see Fig. 9.4 H). This prediction was then validated by experiments (Fig. 9.7 A).

9.3.1 Threshold Mechanism for CD95-induced Apoptosis

In the next step, the model was used to elucidate the exact underlying threshold mechanism. It remains puzzling that even for the below-threshold scenario a sufficient number of receptors should be activated to cleave procaspase-8, thereby triggering all subsequent caspases. The caspase-8 cleavage capacity at the DISC is assumed to be proportional to the number of active CD95 receptors since the DISCs are supposed to remain active after cleaving procaspase-8 molecules. Consequently, the rate of caspase-8 cleavage continuously decreases with lower ligand concentration. Based on this fact only, it has to be assumed that even for ligand concentrations below the threshold, apoptosis cannot be stopped entirely, but would only be slowed down. The apparent contradiction between model prediction and the latter considerations were addressed by revealing the responsible molecules and molecular interactions in the model.

Binding of the short and the long variants of c-FLIP (c-FLIP$_S$, c-FLIP$_L$) to the DISC competes with activation of caspase-8 [125]. According to the parameter estimation, there are many more CD95 receptors and procaspase-8 molecules than c-FLIP molecules. Based on the sensitivity analysis, this estimate is considered very reliable since the investigated system behaviour and the χ^2-function are highly sensitive with respect to the parameters of this network part. In the meantime, the low c-FLIP concentration in comparison with procaspase-8 could also be confirmed by experiments (data not shown). The cleavage rate of procaspase-8 depends on the number of active receptors. Whenever c-FLIP binds to a DISC, the respective binding site is blocked. The simulation of a scenario with subthreshold concentrations of activating ligands shows a steady decrease of active DISCs until all

[1]The diagrams showing the simulated behaviour are based on the parameter set providing the best fit (lowest χ^2-function).

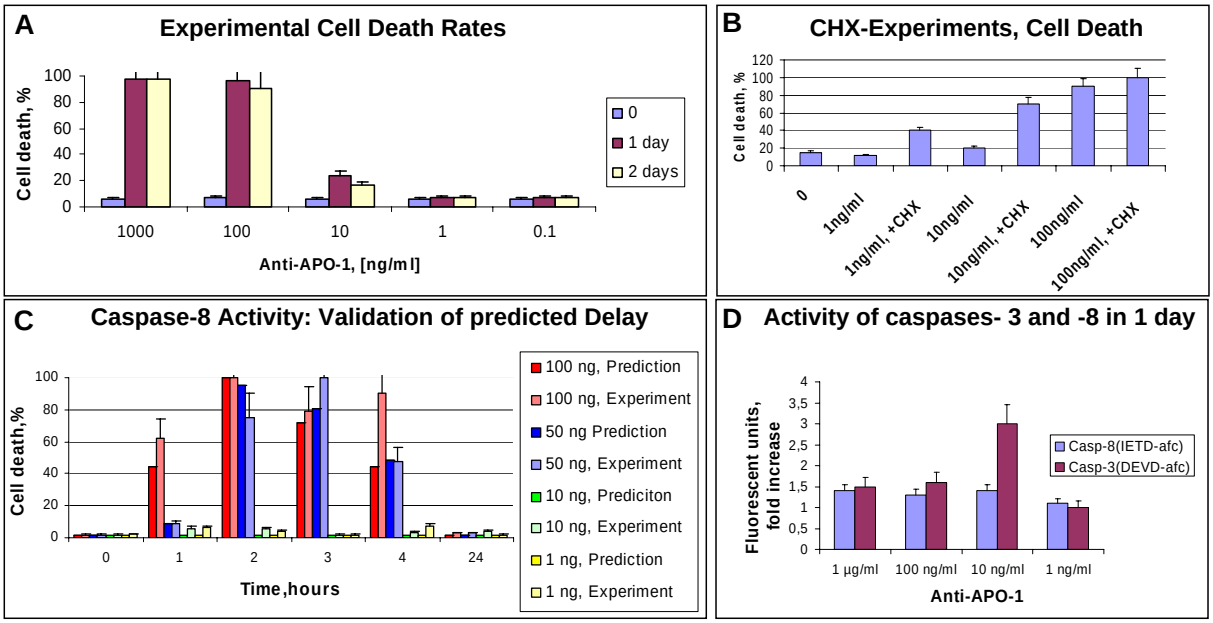

Figure 9.7: Experimental validation of predicted behaviour. **A**: The death rates are in a good agreement with the model, which predicts triggering of the death process for the scenarios of 10 ng/ml and above, whereas for 1 ng/ml, apoptosis is not executed at all. However, the measured death rate for 10 ng/ml was below 100%. Note that these rates were measured for a population of many cells. Variability of parameters and of numbers of ligands as well as intrinsic stochastic effects due to low particle numbers might account for significant fluctuations in scenarios close to the activation threshold. This assumption is confirmed by stochastic simulations in Section 9.4. **B**: SKW 6.4 cells were treated with CHX for 2 hours. Thereafter, the cells were stimulated with the indicated concentrations of anti-APO-1 antibodies (corresponding to CD95-ligands) for one day. Cell death was determined using FACS analysis. **C**: Caspase-8 activity: The model predictions for slow activation scenarios (ligand concentrations between 100 ng/ml and 1 ng/ml) were confirmed by experiments. In particular, the predicted delay of caspase activation was validated. An increase of active caspase-8 was already observed after 1 hour for 100 ng/ml. For 50 ng/ml, the activation starts significantly later. In case of 10 ng/ml, activity was observed after more than 4 hours (data not shown) as predicted in Fig. 9.4 G, whereas no increase occurred for 1 ng/ml. **D**: In the sub-threshold scenario, only a very low increase of active caspase-8 and no increase of active caspase-3 was measured (Y-axis shows the fold-increase). For higher ligand concentrations, the low caspase activity is due to prior degradation (data refer to one day after stimulation). The experiments were executed by Inna Lavrik [34].

of them are blocked by c-FLIP (Fig. 9.8).

As a consequence, the simulation shows a limited generation of the intermediate caspase-

8 cleavage product p43/p41, mainly due to the presence of c-FLIP$_L$ (Fig. 9.8 A,B), but no significant generation of active caspase-8 as a result of the early and complete DISC-blockage. In contrast, the simulation for a ligand-receptor ratio above the threshold shows an entirely different behavior: due to the higher number of active receptors, the amount of c-FLIP is not sufficient to block all DISCs before active caspase-8 can be generated in a quantity that is sufficient to trigger apoptosis. Thus, the c-FLIP mechanism identified in

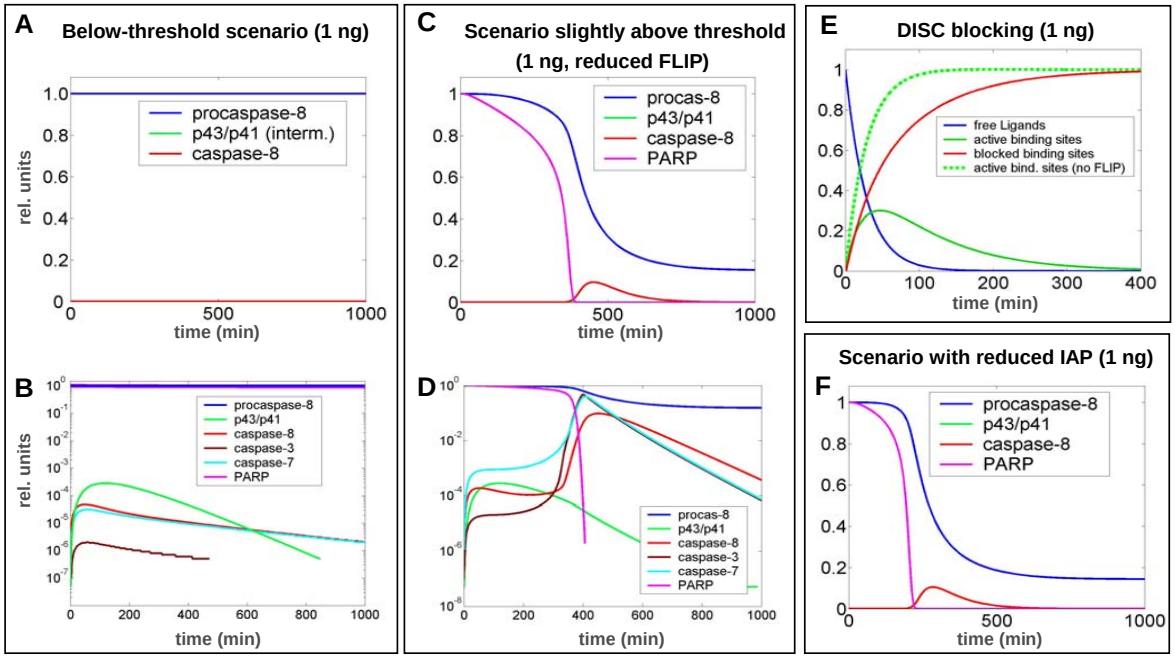

Figure 9.8: Simulation of threshold mechanism. The sub-threshold behaviour (**A, B**) is compared with an above-threshold scenario (**C, D**). Below the threshold, active caspase-8 and all the subsequent caspases cannot be generated in a number sufficiently high to trigger apoptosis. According to the model, c-FLIP is blocking the low number of active DISCs (see red and green curve in **E**) before caspase-8 can be generated in a sufficiently high amount. Without c-FLIP, the number of active DISCs and, therefore, their cleavage capacity would be significantly higher (dotted line in E). The simulation for c-FLIP reduced by 75% (C,D), corresponding to an above-threshold scenario, shows a slow and steady cleavage of procaspase-8 until caspase-3 is generated in a number sufficiently high to trigger the feedback loop via caspase-6, accelerating the activation of caspase-8 and resulting in apoptosis after a delay of many hours (compare B and D on a log-scale). **F**: A similar effect could be simulated for IAP reduced by 75% showing the importance of this inhibitor in case of slow activations (compare A and F).

the model can be considered a switch, which blocks the activation of caspase-8 for signals (ligand concentrations) below a critical quantity and passes on the activation signal above this level. Consequently, the threshold is highly sensitive to the concentration of c-FLIP (Fig. 9.8 C,D).

9.3.2 Experimental Validation of the Threshold Mechanism

To confirm the model predictions experimentally, the FLIP level was down-regulated using the translation inhibitor cyclohexamide (CHX)[3]. Since proteins are steadily degraded and regenerated, this technique mainly affects proteins with short half-life times like FLIP. The experiments confirmed that the addition of CHX decreased c-FLIP level up to 70% and did not change the amount of the caspases (Fig. 7 in [34]). Downregulation of c-FLIP under these conditions resulted in cell death already occurring upon a ligand concentration of only 1 ng/ml (see Fig. 9.7 B). This concentration was shown both experimentally and theoretically to be below the critical value required for apoptosis without CHX. These experiments show the important role of c-FLIP concentration in the regulation of CD95-induced apoptosis and clearly confirm the model predictions (see Fig. 9.8 C,D).

Model-based hypothesis checking of competing threshold mechanisms

The modelling framework was then used to address the currently discussed alternative inhibition mechanisms involving downstream inhibitors like IAP or XIAP [83, 126]. Especially in the case of a low caspase-8 activity, IAP concentration is highly relevant because it directly influences the critical 'threshold' value of the caspase-8 activity, above which the positive feedback loop caspase-8 → caspase-3 → caspase-6 → caspase-8 is triggered. The triggering of this loop is highly sensitive with respect to the concentration of active caspase-8. Once the loop is activated via caspase-3 cleaved by caspase-8, the death process cannot be stopped anymore. Thus, IAP is considered to induce a similar threshold mechanism by effectively blocking caspase-3, thereby disrupting the amplification loop. According to simulations, this is possible up to a critical quantity of caspase-8 only. Above this level, it is predicted that caspase-3 is activated in a quantity, which cannot be effectively blocked by IAP anymore, and thus, the irreversible death process starts. Consequently, for decreased IAP concentrations, this loop becomes already active for lower concentrations of active caspase-8 resulting in a complete cell death (Fig. 9.8 F), whereas high IAP concentrations either inhibit or delay this event for many hours. Thus, IAP also influences the threshold of ligand concentration. However, IAP alone is not sufficient to inhibit apoptosis in the absence of c-FLIP, since it can block signalling only in the case of low caspase-8 activities. Therefore, the influence of IAP is low for ligand concentrations significantly above the threshold. Consequently, the model suggests that the main threshold of CD95-induced apoptosis is determined upstream in the DISC by preventing a steady increase of active caspase-8, which would otherwise trigger the amplification loop even for sub-threshold ligand concentrations. Thereby, the ratio between active receptors and c-FLIP as well as the ratio between binding rates of c-FLIP to DISC and of procaspase-8 to DISC, respectively, are highly relevant parameters for this threshold.

Delay of Apoptosis and Point of no Return

Another important model prediction addresses the system behavior above the threshold, where the combination of the c-FLIP mechanism with the amplification loop does not lead to a steadily decreased caspase cleavage rate upon a decreased ligand concentration. Instead, the simulation shows that the caspase cleavage, the amplification loop and the subsequent death process are delayed but still complete. As shown in Fig. 9.4 F,G and Fig. 9.7 C for low ligand concentrations, there are no observable system changes for up to many hours after activation of the pathway. Then, the death process suddenly starts without any external stimulation of the system. This is due to an extremely slow increase of the active caspase concentration, which reaches a critical level upon which the amplification loop is triggered (see Fig. 9.8 C,D). Thus, for up to many hours, there is no observable difference to cells, which are not (or not sufficiently) stimulated by ligands, however, the death process is irreversibly started form the beginning of the stimulation on and cannot be stopped anymore (point of no return). These predicted delays were also confirmed by experiments (Fig. 9.7 C).

The proposed threshold mechanism was verified by testing the model predictions for several scenarios. The caspase-8 activation was measured for a series of lower ligand concentrations, quantitatively confirming the predicted delays, the complete cleavage of procaspase-8 above and the blockage of the active caspase-8 generation below the threshold (e.g. Fig. 9.7 C). In order to further prove the proposed mechanism, the activity of up- and downstream molecules below the threshold was systematically scanned. The experiments confirmed that a low amount of p43/41 and an extremely low amount of active caspase-8 were generated below the critical activation threshold as predicted by the model (Fig. 9.7 D and Fig. 9.8 B). No significant activity of caspase-3 was observed, which would otherwise have triggered the feedback loop (Fig. 9.7 D). Furthermore, neither PARP cleavage nor cell death was observed. This is a clear indication that the main signal is stopped upstream at the DISC by c-FLIP, and that IAP, the second important inhibitor, prevents the sensitive caspase-3 activity from reaching a significant level upon low amounts of caspase-8 as predicted by simulation.

9.4 Stochastic System Behaviour

As a consequence of the previously described threshold behaviour, deterministic simulation methods yield a death rate of either 0% for ligand concentrations below, or 100% for ligand concentration above one distinct threshold. The low particle numbers of critical molecules[2] for slow activation scenarios, however, suggest that the system behaviour is also governed by stochastic effects. Moreover, the experimentally observed death rates referring to a population of many cells are between 0% and 100% for certain ligand concentrations (Fig.

[2]The simulation shows that for slow activation scenarios, the particle numbers of active caspases can be below 100.

9.7 A). This also indicates that due to stochastic effects, apoptosis is triggered for a subset of all cells only. For this reason, stochastic simulations were performed for different activation scenarios with initial ligand concentrations between 0.1 ng/ml and 10 $\mu g/ml$. The simulated activation scenarios were then scanned for stochastic effects. Multiple stochastic simulation runs (up to 1000) were performed for each investigated activation scenario and the time evolution of the mean molecule concentrations and of the respective standard deviation were evaluated. Results for one specific scenario have already been discussed in Section 3.4.2.

For activation scenarios with ligand concentrations of more than ≈ 50 ng/ml, no significant stochastic effects could be found. The deviations between different stochastic simulation runs for the same scenario were negligible. Furthermore, the mean values are identical to the result of deterministic simulations. This fact confirms that the use of a deterministic simulation method for the parameter estimation (ligand concentrations ≥ 200 ng/ml) was appropriate. For lower ligand concentrations, however, stochastic simulations show a system behaviour, which is qualitatively different from the deterministic results. Thus, the deterministic simulation methods are not appropriate in this area. An example is given in Fig. 3.2 for a ligand concentration of 7 ng/ml.

In Section 9.3.1, it was demonstrated that the concentration of active caspase-8 and caspase-3 is highly critical since slight variations can be crucial for triggering the apoptotic pathway resulting in cell death. It is assumed that this effect is mainly due to the self-amplifying feedback loop (caspase-8 $\rightarrow$ caspase-3 $\rightarrow$ caspase-6 $\rightarrow$ caspase-8). The feedback loop is triggered by active caspase-8 molecules that are generated at the DISC depending on the number of active receptors and thus on the initial ligand concentration. Since active caspases are subject to steady degradation and inhibition (e.g. binding with IAP), there is a critical particle number of active caspases, above which the loop is triggered. If this number is not reached due to insufficient caspase-8 generation, apoptosis will not be executed and the cell will survive. This case occurs for ligand concentrations below a certain level (see Fig. 9.9). On the other hand, it can be predicted that the loop is always triggered for much higher ligand concentrations. However, for the area between both extremes, it cannot be predicted when the critical particle number for triggering the amplification loop is achieved. Furthermore, it cannot even be predicted whether the loop and the subsequent apoptotic process is triggered at all. This is a result of the fluctuations of the active caspase concentrations. Thus, only a probability for reaching the critical number of active caspases and for the execution of apoptosis can be determined depending on the ligand concentrations.

The results of the stochastic simulations performed for different activation scenarios are visualized in Fig. 9.9. It was shown that for ligand concentrations of ≈ 1.5 ng/ml and below, cell death is not achieved. For ≈ 20 ng/ml and above, the death rate is 100%, whereas the area between 1.5 ng/ml and 20 ng/ml is dominated by stochastic effects resulting in death rates between 0% and 100%. Thus, there is a value range of more than

one order of magnitude, in which only a part of the cells within a cell population dies due to fluctuations of the caspase activity.

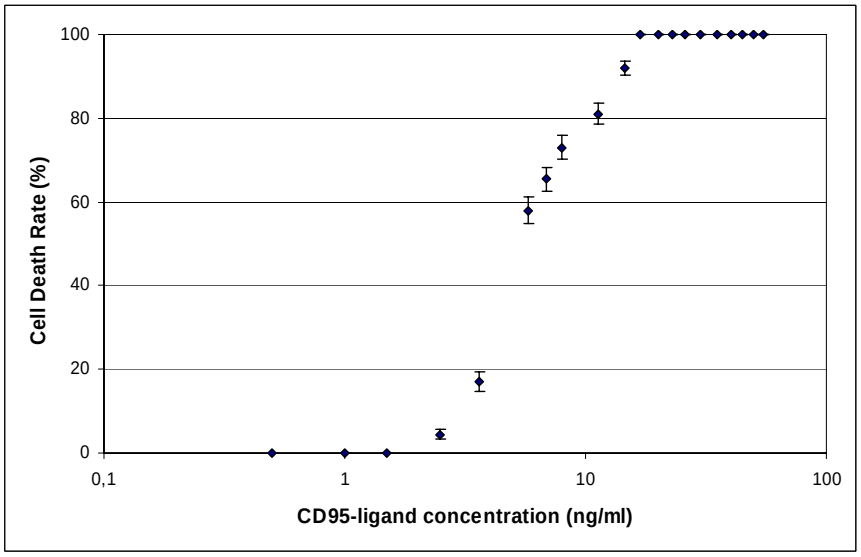

Figure 9.9: Cell Death Rates computed by stochastic simulation. This diagram shows the probability that cell death is triggered depending on the CD95-ligand concentration. A series of independent stochastic simulation runs was performed for each activation scenario (ligand concentration). According to these simulations, apoptosis is always executed above an upper threshold of ligand concentration, whereas the cells survive below a lower threshold of ligand concentration. Between both values, the cell behaviour shows high fluctuations either leading to cell death or cell survival. Note that a single simulation run corresponds to the behaviour of a single cell and that the probabilities reflect the death rates of a cell population.

As already demonstrated in Section 3.4.2, stochastically exact simulations using the original Gillespie algorithm are impaired by the high computation time since only some of the molecule species are of low particle numbers and reaction rates. On the other hand, it was shown that the approximate stochastic simulation method using Langevin equations could not provide correct results (Fig. 3.2). This is due to the fact that in this method, interactions between molecules are not simulated as discrete events, leading to qualitatively different results whenever particle numbers of critical molecules are low. For this reason, the General Stochastic Hybrid Method, which was developed in this thesis (Chapter 3.3), was used. This method applies the stochastically exact Gillespie algorithm only for the reaction subset of molecules with critically low particle numbers, whereas the remaining reactions are simulated using Langevin equations. As already presented in Section 3.4.2, the method was capable of saving $> 98\%$ of computation time in comparison with the pure Gillespie method and reliable stochastic simulations became feasible.

Critical System Behaviour revealed by Ensembles of Parameter Sets

A very interesting result is the fact that the critical threshold area between cell death and cell survival could also be detected by a completely different method. As an extension of the approach of evaluating sensitivities on the basis of ensembles of parameter sets, I suggested in Section 7.2.3 to investigate other system properties on this basis as well. This was proposed as an alternative to the examination of distinct parameter fits. Especially discrepancies in the predictions are thereby expected to reveal biologically interesting areas. In order to apply this concept, the concentration of important molecules like PARP, which is an indicator for the execution of cell death, was evaluated for a large number of parameter sets in the space of possible parameter values. On this basis, the variance of the prediction of cell death is plotted for different ligand concentrations to detect areas of high inconsistencies among the predictions. This plot is visualized in Fig. 9.10. It clearly indicates that only within a distinct area, which approximately corresponds to the area governed by stochastic effects (see Fig. 9.9), the predictions strongly vary. Thus, the approach of evaluating parameter set ensembles has proven to be capable of predicting an area of utmost biological relevance, which is also crucial for systematic planning of new experiments. This method can therefore be considered an enormously powerful tool for system identification of highly underdetermined biological networks.

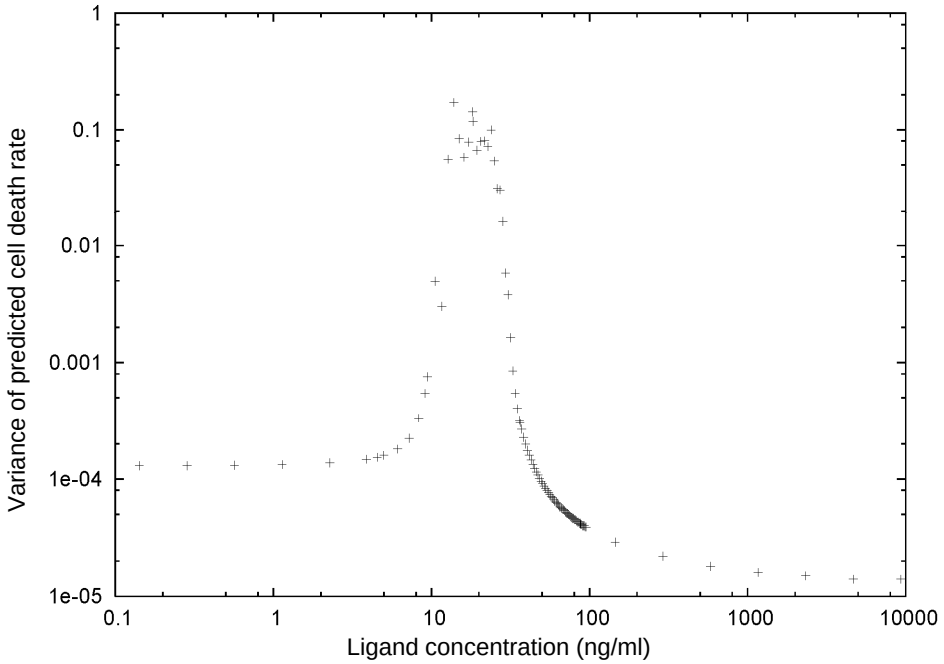

Figure 9.10: The variance of predicted cell death rates (Y-axis) is plotted for different ligand concentrations (X-axis), based on the evaluation of a parameter fit ensemble. Cell survival and cell death of a single system is quantified by 0 and 1. The plot reveals an area of high inconsistencies among the predictions for different parameter sets, corresponding to an area of high biological relevance. The plots were generated by Dr. Ivayla Vacheva.

Chapter 10

Conclusions and Perspectives

The motivation of this thesis was data-based modelling and system identification of complex networks in cell biology, which are characterized by a large number of unknown parameters, poorly understood mechanisms, uncertain network topologies and stochastic effects. Such systems are ubiquitous in molecular cell biology and in many other fields of science like econometrics. For this purpose, I implemented a modelling technique integrating information of heterogeneous quality. Furthermore, I developed a stochastic sensitivity analysis method and introduced the concept of 'Sensitivity of Sensitivities' based on the evaluation of sensitivity histograms. Applied to the signal transduction network of CD95-induced apoptosis, the method was able to reveal two important system properties even without having knowledge about the true parameters: modularity and robustness. On this basis, I created an efficient parameter estimation approach, leading to parameters, for which the model has been found to be capable of reproducing the observed system behaviour and of predicting biologically highly relevant system properties. These predictions could be confirmed by experiments. Moreover, I developed a stochastic hybrid algorithm, which enables stochastically accurate simulations for large reaction systems running on multiple timescales, thereby revealing the stochastic systems behaviour. In conclusion, I have demonstrated that the methods of this thesis were well-suited to identify and to elucidate key mechanisms in the complex regulation of programmed cell death.

Data-based Modelling of complex Signal Transduction Networks

It has been shown that the mathematical model of CD95-induced apoptosis, based on the system identification and modelling framework I developed in thesis, provides novel insights into important regulatory mechanisms for induction of apoptosis. To overcome the problem of heterogeneous information about cellular networks, 'Structured Information Models' have been introduced, which divide the network into subsystems according to information quality. Thereby, information of different levels can be incorporated in one model instead of dealing with isolated models. This approach provides high extensibility, e.g. by splitting the introduced black boxes into further subsystems once new information

becomes available. Therefore, the technique is well-adapted to uncertain network topologies and especially to the presently limited availability of reliable qunatitative data.

Parameters have been estimated based on experimental time-series and predictions could be generated that were thoroughly validated by experiments. In comparison to existing models of signal transduction systems, e.g. the study of Fussenegger et al. [28], which is based on ad hoc set parameter values only, or the data-based study of Swameye et al. [27], which addresses a rather small and isolated system, a data-based method has been developed here, which is capable of handling large models with a large number of unknown parameters. Two inherent system properties, i.e. modularity and high robustness of parameter sensitivities have been revealed by the new stochastic sensitivity analysis approach (Section 6.2). They were the key to drastically reduce the dimensionality of the parameter identification problem (Section 7.2). The developed framework provides a general basis for large-scale modelling and simulation of complex networks and can be readily applied to other applications such as modelling of metabolic networks, cell proliferation or differentiation. The intrinsic reduction of dimensionality proposed here is systematic and adaptive to both the original model and the experimental data.

Biological Relevance of Mathematical Modelling

A biologically highly relevant result of the combined theoretical and experimental approach was the finding of a 'threshold between life and death' in the regulation of CD95-mediated apoptosis. Model-based hypotheses about the underlying regulatory mechanism have been generated, also addressing the highly debated question of the exact influence and interaction of different inhibitors. Especially the fact that it became possible to predict scenarios, in which the influence of certain inhibitors becomes crucial for apoptotic regulation, is of utmost relevance for the planning of new experiments. Detailed predictions, e.g. concerning a shift of the threshold upon concentration changes of the inhibitors could be experimentally confirmed. Furthermore, the model has elucidated the mechanism responsible for the phenomenon that for up to many hours, no observable system change takes place in case of low activations, followed by a sudden start of the apoptotic program (point of no return). The threshold mechanism was predicted to be closely related to the upstream factor c-FLIP that efficiently blocks caspase-8 activation at the DISC at low ligand concentrations consequently stopping the apoptotic program. An important pre-condition for this mechanism was a significantly lower c-FLIP concentration in comparison with procaspase-8, which was revealed by parameter estimation and has been experimentally confirmed in the meantime. Moreover, this regulatory mechanism could be probed in silico by simulation of alternative mechanisms [34], which were found not to be consistent with experimental data.

Despite the ever-increasing number of studies on CD95-induced apoptosis, a systemic understanding of this complex signalling pathway is still missing. The question of a threshold for induction of apoptosis plays a central role in understanding the sensitivity and resistance of cells towards various chemotherapeutic agents. Abnormal c-FLIP expression has

been identified in various diseases connected with dysregulation in CD95 signalling such as multiple sclerosis, Alzheimer's disease, diabetes mellitus, rheumatoid arthritis, Hodgkin's disease and different cancers [127, 128]. The developed mathematical framework now enables simulations of the mechanism under different conditions, thereby predicting a higher or lower resistance to apoptosis. Thus, a powerful tool for describing potential effects of chemotherapeutics and for studying the mechanism behind the regulation of apoptosis by drugs in treatment of cancer and other diseases is provided. As an example, the highly debated question between IAP and c-FLIP as possible inhibitors, could be answered by identifying IAP to be crucial for extremely low activation scenarios only.

An important role of modelling signal transduction in this thesis was related to hypothesis generation and experimental design. The model predicted a delay of cell death with decreasing stimulation strength of CD95-induced apoptosis, which surprisingly resulted in a complete change of system behaviour, i.e. the complete stop of the cell death program. According to the limited numbers of experimental data, it cannot be expected that the parameters in the model are fine-tuned. However, by revealing the ranges, in which these qualitative changes of the system behaviour are expected, a basis for the definition of new experiments is given. Moreover, the model was capable of predicting that there is a range of ligand concentration covering more than one order of magnitude, in which the system is mainly governed by stochastic effects, which is a possible explanation for the observed death rates (Section 9.4). Thus, the model has revealed mechanisms for observed system behaviour, which have answered highly relevant questions about the regulation of programmed cell death for the first time. In this sense, mathematical modelling in the context of programmed cell death has proven to be an indispensable part of biological knowledge discovery.

Stochastic System Behaviour and General Stochastic Hybrid Algorithm

Stochastic simulations showed that the system behaviour is mainly governed by stochastic effects within a threshold-close range of ligand concentrations. Due to extremely low particle numbers of important molecule species (e.g. active caspase-8), where small variations can already be crucial for triggering the death process, exact stochastic simulations are required. It has been shown in Section 3.4.2 that deterministic or approximate stochastic simulations lead to qualitatively incorrect results. Exact stochastic simulations using the Gillespie algorithm are, however, severely impaired by the enormous computation time required in case the model also contains molecules and interactions of much higher particle numbers and reaction rates. In order to address this very general problem, I have developed the General Stochastic Hybrid Method (Chapter 3), which combines the Gillespie algorithm with a system of stochastic differential equations. Using this method for simulation of CD95-induced apoptosis, the computation time could be reduced by more than 98% and stochastically reliable simulations for the prediction of the cell death rates (Section 9.4) became possible. Moreover, the hybrid algorithm is a very general approach and can be applied to any other kind of Markov processes as well.

Distribution of Sensitivities in Parameter Space

The concept of evaluating the distribution of sensitivities in parameter space was introduced to reduce the high dimension of the parameter identification problem. The common approach of manual model simplification potentially introduces a user bias into the model, since simplification, in general, requires an understanding of the network functionality, which is typically not given without prior knowledge of the true parameters. Instead, the reduction of dimensionality was based on the evaluation of sensitivity histograms, which provide information about the distribution of sensitivities within the space of possible parameter values (Section 6.2). This can be considered as an approach towards global sensitivity analysis, for which a general solution does not exist. The main objective of this procedure was to gain information about general dependencies between the temporal behaviour of state variables and the parameters.

The fact that the sensitivity histograms provide crucial information (Section 9.1) is an important result. As an intrinsic system property, a large percentage of sensitivities have been found to be extremely low and extremely insensitive with respect to parameter variations. On this basis, subsets of state variables could be identified, which depend on a subset of parameters only, regardless of the true parameter values. This is crucial for the cluster-based and sensitivity-controlled parameter estimation method developed in Section 7.2. In this approach, reduction of dimensionality is already achieved by a modular structure of the dependencies between parameters and state-variables. Thereby, it is not a requirement to identify parameters, which are irrelevant for the complete system. As a result, the performance of the parameter estimation has been significantly increased (Section 9.2).

Although a systematic scan of the parameter space is impaired by its high dimension raising questions about the evidence of the latter approach, it has been shown in Section 6.2.2 that under certain assumptions, even a limited number of parameter samples provide meaningful information. These assumptions are not based on the specific sensitivity-parameter dependencies, they are merely based on properties, which are typically fulfilled in biochemical systems, like a high percentage of parameters, which are either correlated, or which do not influence certain system parts.

Furthermore, the approach of evaluating the distribution of sensitivities is always correct for predicting which dependencies between state variables and parameters must not be neglected for system identification and model simplification (even though they might be found to be irrelevant at the end). Thereby, it can also be predicted, which areas of a network have to be simulated quantitatively due to a strong influence of parameter values on the system behaviour, and for which parts qualitative simulations might be sufficient due to intrinsic and parameter-independent system properties.

Data-based Ensembles of Parameter Sets

The approach of evaluating sensitivities for an ensemble of parameter sets instead of considering local sensitivities can be put into a broader context. First, data from observations were incorporated by introduction of a Boltzmann factor, amplifying the impact of sensitivity information according to the difference between simulation data for the respective parameter set and experimental data (Section 6.2.3). Thereby, typical patterns were found in the histograms, which provide additional information.

Second, the procedure of evaluating sensitivities for a large number of different parameter sets has been extended to the investigation of other system properties on the basis of an ensemble of parameter fits. It has been shown in Section 9.2.3 that many different parameter fits were found by parameter estimation, which match the experimental data with almost the same quality. This indicates that system properties and predictions based on one single parameter fit are not very meaningful. As a consequence, it has been suggested in this thesis to consider a complete ensemble of parameter fits for the investigation of system properties and for the generation of predictions (Section 7.2.3). Especially inconsistencies in the predictions are of high relevance for the definition of new experiments. In this way, the critical activation range for CD95-induced apoptosis could be identified.

Future Perspectives

The investigation and analysis of complex biological networks and mechanisms in cells is probably one of the most challenging and fastest growing fields of science. The classic experimental approaches, which mostly focus on the investigation of molecule interactions in an isolated context and under specific experimental settings, cannot keep up with the steadily increasing number of potential interaction partners and the diversity and complexity of the real networks. The potential of pure experimental approaches will therefore be limited for revealing the network functionalities in living cells. Even the outcome of DNA sequencing remains below its potential as long as protein functions and interactions within the life systems are not sufficiently understood.

But also from the theoretical perspective, system identification of networks, characterized by an enormous complexity and a lack of information about the underlying mechanisms on both the qualitative and quantitative level, constitutes a new class of problems, which has not been sufficiently approached yet. Methods to describe complex networks in a qualitative way, e.g. as scale free networks [129, 130], are promising for revealing general principles like robustness and fragility [131]. Their ability to describe the real system behaviour is, however, limited, since single interactions are not quantitatively considered although the properties of real biological networks are often related to very concrete features of single mechanisms.

The emerging field of systems biology, which has been recently started with great en-

thusiasm, is an important step towards the investigation of biological processes on systems level. However, it has not proved yet that it also provides qualitatively new methods, which address the new dimension of complexity. Instead, approaches from the field of engineering and numerical mathematics are now applied to biological systems, especially to subsystems, where they are expected to be well-suited and promising. Moreover, the majority of studies address specific applications with the goal of answering specific biological questions. Although this is a first important step to establish systematic and theoretical procedures in the field of cell biology, it has to be followed by a second, much more challenging step concerning the development of new theoretical approaches for the description and particularly for the system identification of complex and highly underdetermined biological systems.

Numerical parameter identification methods, for example, which originate from disciplines, where experimental data are generated in huge amounts and with high precision and where the number of unknown parameters is low, cannot be expected to be appropriate for the completely different situation in cell biology. Here, uncertain models and a large number of unknown parameters are facing a low number of experimental data points with high measurement errors. Whereas the current numerical methods are based on the assumption that the best parameter fit corresponds to the problem's solution, it has to be taken into account that the data situation of biological systems typically results in a huge solution space and that a single parameter fit should be considered rather meaningless. One could argue that in this situation, better experimental data should be generated and more quantitative information concerning the single mechanisms should be obtained at first. Considering the fact that quantitatively reliable *in vivo* measurements like time series of concentrations are still enormously time-consuming, this cannot be expected in the near future.

Instead, I believe that it would be fatal - especially concerning the great potential of developing new theoretical methods - not to attempt to extract information from available quantitative experimental data in the context of existing qualitative knowledge about e.g. network topologies. Therefore, alternative ways disregarding the restrictions of current numerical methods and by-passing the requirement of finding one distinct parameter set have to be found. This is one of the main conclusions of this thesis and reflects the situation of real biological systems, as it is well-accepted that biological systems mostly keep their system properties constant although the real parameters are also subject to high variations. Thus, intrinsic biological properties like robustness and the fact that the function of biological systems does not require fine-tuned parameters, indicates that the extraction of information is possible even without exact knowledge of the true parameters. This is a new principle, which was approached in this thesis by evaluating randomly chosen parameter sets and by the generation of ensembles of estimated parameter fits based on randomly chosen initial values. It has been shown that histograms of parameter sensitivities, for example, provide enormous information, in particular by incorporation of a Boltzmann factor based on experimental data.

Although the huge space of unknown parameters raises questions about the evidence of system properties claimed on the basis of evaluating ensembles of parameter fits, the approach has proved that already based on inconsistencies in the predictions associated to the parameter sets, it has become possible to reveal crucial areas of impact of critical molecules and to define new experiments. So far, in most cases, this was only possible by good intuitive guesses of experimentalists. A systematic and theoretical approach was, however, missing.

I believe that the currently developing methods such as provided in this thesis or in the recent study of Brown and Sethna [132] dealing with model extraction also based on multiple parameter fits, mark the beginning of a new methodology to investigate highly underdetermined systems. The approach of this thesis has to be further refined, e.g. by the 'High Dimensional Model Representation' [107] for a systematic description of an 'effective' parameter space. Furthermore, it has to be extended by the concept of also considering alternative model choices instead of different parameter sets only, leading to model discrimination, which is currently examined in the study of Vacheva, Bentele and Eils [117]. For a better understanding of the regulation of programmed cell death, the established loop between modelling, theoretical predictions and experiments has already proved to be highly efficient and has raised a lot of new detailed questions. These have defined further areas of investigation of CD95-induced apoptosis, e.g. the influence of spatial aspects on the network function (study of Ulrich, Bentele, et al. [133]) or more detailed investigation of certain key regulatory mechanisms (study of Lavrik, Bentele, et al. [134]). The modular and hierarchical structure of the presented modelling framework provides a high degree of flexibility for future model extensions in various ways, either by adding additional pathways and systems like proliferation or gene expression, or by adding more detailed biochemical mechanisms with more information becoming available.

The prediction of molecules and mechanisms, relevant for the threshold between cell death and cell survival, is of utmost relevance for cancer research and for the identification of potential chemotherapeutical drugs. Thus, the methods of this thesis have proved to be well-suited for complex and highly underdetermined networks and are therefore of high relevance for many other research areas, going far beyond programmed cell death.

Appendix A

Sotware Package ISLANDS

In this chapter, a user manual for the software package ISLANDS (Integrated Signal Transduction Modelling and System Identification Tool), which I developed for integrating the modelling, stochastic simulation, sensitivity analysis and parameter estimation approaches of this thesis, is given. The software structure and functionality is outlined in Chapter 8.

The user interface is provided by an input file, in which the run mode and configuration parameters have to be specified. Furthermore, the input file contains a section for model definition and experimental data (parameter estimation). The output structure of the software depends on the run mode. For simulations, a time series of all state variables is given, for sensitivity analysis, the sensitivity matrix and sensitivity histograms are provided. In case of parameter estimations, the parameter fits and the least square term are given.

The input file is split into several data sections, which are demonstrated in the following. Each data section is introduced by the respective keyword. An exemplary input file containing an explanation of each element is given at the end of this chapter.

Run Mode and Settings

The settings section contains general setting like the run mode, options for the output (e.g. log level), configurations and numerical settings, e.g. for the ODE and hybrid solver. The data section is introduced with the keyword *SETTINGS*.

Model Definition

In this section, the model is defined by the molecules (state variables), reaction schema, reaction rate definition, black box definition and parameters. The molecule and parameter names can chosen in a user-defined way. Links between initial concentrations and parameters, between reactions and molecules, between reactions and parameters, etc., are given by these names. The data sections are introduced by the keywords *MOLECULES, REACTIONS, RATE_FACTORS, PARAMETERS* and *FN_UNITS*.

Experimental Data

For parameter estimation and data-based weighting of sensitivities, time series of experimental data have to be provided. Each time series consists of four lines. In the first line, the link to the state variable (molecule name) is given, to which the time series refer. The second, third and fourth line contain the time points, experimental value and standard deviation. The experimental data section is introduced by the keyword *EXPERIMENTAL_DATA*.

Multiple experimental scenarios can be included (see Section 8.1.2). For each additional scenario, an identifier has to be provided as well as the different initial conditions, which is realized by linking initial values of the respective state variables (molecules) to additional parameters. The experimental data have to refer to molecule names with the identifier for the respective scenario. The data section is introduced by the keyword *MODEL_MULTIPLIER*.

Example for Input File

```
////////////////////////////////////////////////////////////////////////////////////////////////////
//
//        I N P U T   F I L E   (E X A M P L E)
//
//        S E T T I N G S   A N D   M O D E L   D E F I N I T I O N
//
////////////////////////////////////////////////////////////////////////////////////////////////////

SETTINGS


// Run Mode:

RUN_MODE                          0           //  SIMULATION:                                0
                                              //  MULTIPLE_STOCH_RUNS:                       1
                                              //  SENSITIVITY_ANALYSIS:                      2
                                              //  STOCH. SENSITIVITY_ANALYSIS:               2
                                              //    (with SENS_ANAL_MONTE_CARLO_PARAM=1)
                                              //  PARAMETER_ESTIMATION:                      3
                                              //  PARAMETER_ESTIMATION MULTI-START:          3
                                              //    (with PARAM_ESTIM_RANDOM_START_VALUES=1)
                                              //  USER_DEFINED_PARAM_ESTIM                   4


SENS_ANAL_MONTE_CARLO_PARAM       0       // stochastic sens. analysis (histograms)
PARAM_ESTIM_RANDOM_START_VALUES   0       // random multi-start algorithm for param. estim.


// Multiple Stochastic Runs:

N_STOCH_RUNS                      1000    // number of runs in multiple stochastic run mode (RUN_MODE=1)
N_STOCH_RESULT_TIMESTEPS          1000    // number of time points for generation of stoch. result (mean+stddev))


// Simulation Settings:

t_0                               0       // start time of simulation
t_max                             200     // simulation runs from t=0 to t=t_max
report_step_width                 8       // only one of 'report_step_width' steps are reported into file.
global_log_level                  1       // log level: 0: everything  1: detailed  2: state variables only
ConcToParticle                    20000   // factor between concentration and particle number
StochMode                         0       // 1: purely stochastic (pure Gillespie, ODE solver not used at all)
STOCH_PARTICLE_NO_THRESHOLD       0       // 0: pure ODE solver, >0 hybrid solver with thresh. for particle number


// Model Output:

PRINT_DIFF_EQUATIONS_MUSCOD       0       // print diff. equations as C-function (right hand side, MUSCOD compatible)
PRINT_DIFF_EQUATIONS_MADONNA      1       // print diff. equations in Berkeley Madonna format
PRINT_EXP_VALUES                  0       // options for printing exp. values in MUSCOD-structure


// Sensitivity Analysis:

SENSITIVITY_CORRELATION           1       // 1: correlation of sensitivities considered in parameter estimation
SENS_PARAM_VARIATION              0.005   // relative parameter variation (numerical default method)
HISTOGRAM_DIM                     200     // number of intervals in histogram
N_MONTE_CARLO_PARAM_STEPS         10000   // number of random parameter sets
REPORT_SENS_TO_FILE               1       // sensitivities (rand. walk) are written to file after each step
DEPENDENCY_MATRIX                 0       // print dependency matrix (sens. matrix, 1: sens. above threshold, 0: else

kb_T_sensitivity_histogram        100     // kT for histograms


// Parameter Estimation:

N_PARAM_ESTIM_STARTS              10000   // number of parameter estimations (random multi-start)
N_MONTE_CARLO_SAFETY_REPORT_STEPS 10      // saving results after each N_MONTE_CARLO_SAFETY_REPORT_STEPS-interval
N_PARAMS_TO_ESTIM                 17      // number of parameters to be estimated

SENSITIVITY_THRESHOLD             0.01    // threshold for local parameter sensitivities (Theta_s) (sensitivity control)
SENSITIVITY_CORRELATION_THRESHOLD 0.98    // sensitivity correlation threshold (1-Theta_corr)
MAXIMUM_OBJ_FN_OF_FIRST_ESTIM     50000   // If the objective function of the initial values (random param. estim. mode)
                                          // is above this threshold, estimation is skipped.
MAXIMUM_OBJ_FN_OF_ANY_ESTIM       50000   // At any step: if the objective function of the values, estim. is skipped.
```

```
// Settings for ODE Solver (equivalent to Runge-Kutta solver of numerical recipes):

ODE_SOLVER_EPS                  0.00001
ODE_SOLVER_H1                   0.001
ODE_SAFETY                      0.9
ODE_PGROW                       -0.2
ODE_PSHRNK                      -0.25
ODE_ERRCON                      1.89e-5
ODE_MAXSTP                      1000000
ODE_TINY                        1.0e-30


// Hybrid Solver:

TOL_REACTION_TIME               0.000001        // max. tolerance for iterative determination of next
                                                // stoch. event
MAX_STOCH_STEPS                 10000000000     // max. number of stochastic steps
STOCH_MODE_UPDATE_INTERVAL      5               // division of reaction system into subset is rechecked after
                                                // STOCH_MODE_UPDATE_INTERVAL simulation steps.
LANGEVIN_MODE                   1               // 1: general stochastic hybrid solver GSHM (with Langevin)
                                                // 0: pure deterministic-stochastic hybrid solver
dt_max_langevin                 0.05            // max. time step width for Wiener process (Langevin)


////////////////////////////////////////////////////////////////////////////////////////////
//
// STATE VARIABLES
//
// name         initial value

MOLECULES

L               I_L             // Ligand with initial concentration I_L (parameter)
R               I_R             // Receptor with initial concentration I_R (parameter)
L_R             0               // Receptor:Ligand with initial concentration 0
FADD            I_FADD          // ....
DISCbs          0
DISCpCas8       0
DISCpCas8_2     0
...


////////////////////////////////////////////////////////////////////////////////////////////
//
// REACTIONS
//
//  A  +     B       ->     C       +     D         mode            k_ab        k_rev   ratefactor


REACTIONS

L           R           L_R           0         ELEM_REACT      K_LR            0     0
L_R         FADD        DISCbs        0         ELEM_REACT      K_DISC          0     0
DISCbs      proCas-8    DISCpCas8     0         ELEM_REACT      K_DISC_pC8      0     0

DISCpCas8   proCas-8    DISCpCas8_2   0         ELEM_REACT      K_DISC_pC8_2    0     0
DISCpCas8_2 0           Cas-8_IM      Cas-8_IM  ELEM_REACT      K_DISC_to_C8IM  0     0

proCas-3    0           Cas-3         0         RATEFACTOR      K_Cas_8_3       0     Cas8_3
proCas-3    0           Cas-3         0         RATEFACTOR      K_Cas_9_3       0     Cas9_3
proCas-7    0           Cas-7         0         RATEFACTOR      K_Cas_9_7       0     Cas9_7
proCas-2    0           Cas-2         0         RATEFACTOR      K_Cas_3_2       0     Cas3_2
...

////////////////////////////////////////////////////////////////////////////
//
//  EXPLANATION:
//
//  Mode = ELEM_REACT : A+B -> C+D with rate [A][B]*K_ab  and reverse reaction
//                              with rate [C][D]*K_rev
//
//  Mode = RATEFACTOR : A+B -> C+D with rate K_ab*ratefactor as defined in
//                              the appr. ratefactor (name RateFactorName).
////////////////////////////////////////////////////////////////////////////
```

```
////////////////////////////////////////////////////////////////////////////////////////////////
//
// PARAMETERS
//
// name             value        min      max (param. estim.)

PARAMETERS

K_LR                 0.017748     0.01     1
K_DISC_pC8           0.028354     0.01     1.5
K_DISC_to_C8IM       0.913561     0.01     1
I_FLIP               0.293669     0.02     5
I_PARP               0.050411     0.02     5
...


////////////////////////////////////////////////////////////////////////////////////////////////
//
// RATEFACTORS
//

RATE_FACTORS

RATEFACTOR Cas8_3
proCas-3     RECIPROCAL       Km_Cas_89_3
Cas-8            0                NONE

RATEFACTOR Cas9_3
proCas-3     RECIPROCAL       Km_Cas_89_3
Cas-9            0                NONE

RATEFACTOR Cas9_7
proCas-7     RECIPROCAL       Km_Cas_9_7
Cas-9            0                NONE

...

////////////////////////////////////////////////////////////////////////////////////////////////
//
// EXPLANATION OF RATE FACTORS:
//
// Rate factors are designed to provide felxibility in the description of reaction rates.
//
// Each line corresponds to a factor by which the reaction rate of the reaction referring
// to the rate factor is multiplied.
// A line consists of 3 elements:
//  - First element:  molecule
//  - Second element: mode (RECIROCAL OR 0)
//  - Third element:  constant (reference to parameter, given by the parameter name.
//                    'NONE' means: no further constant required
//
// The factor is computed as follows:
// mode = 0:                factor = [molecule] * constant
//                                   (if third element='NONE'(-> no const): factor = [molecule].)
// mode = RECIPROCAL:       factor = [molecule]/(const+[molecule])
//
// All factors (all lines below RATEFACTOR) are multiplied. The result is multiplied with the
// reaction rate constant of the reaction, which refers to the rate factor. This result is the
// reaction rate of the corresp. reaction !
//
// Example for Michaelis-Menten with enzyme and substr:
//
// RATEFACTOR mich_ment
// substr    RECIPROCAL   Km
// enzyme    0            NONE
//
// means:  rate = k*[enzyme][substr]/(Km+[substr])   (The reaction substr->product has to refer to
//                                                    this rate factor!)
//
////////////////////////////////////////////////////////////////////////////////////////////////


////////////////////////////////////////////////////////////////////////////////////////////////
//
// DEFINITION OF BLACK BOXES
//

FN_UNITS

// e.g. Mitochondria (Cytochrome-C-release):

BLACK_BOX_NR 1
INPUT            tBid         BclXL/2          CytoSmacStrd
OUTPUT           CytoSmacStrd CytoSmac
FNU_PARAMS   C_MITO_D_BID_BCL T_MITO_CYT_REL

...
```

```
////////////////////////////////////////////////////////////////////////////////
//
//  EXPLANATION:
//
//  BLACK_BOX_NR: number
//  INPUT:        input parameters
//  OUTPUT:       output parameters
//  FNU_PARAMS:   parameters of black box
//
////////////////////////////////////////////////////////////////////////////////



////////////////////////////////////////////////////////////////////////////////////////////////////
//
// MODEL MULTIPLIER
//
// for parameter estimation based on multiple parameters
//

MODEL_MULTIPLIER

ADDITIONAL_MODEL _slow1
L_slow1      I_L_slow1
R_slow1      I_R_slow1

...

////////////////////////////////////////////////////////////////////////////////////////////////////
//
//  EXPLANATION:
//
//  ADDITIONAL_MODEL 'identifyer':
//  'state variable'     'new parameter'
//  ...
//
//  An additional scenario with the respective identifier (here '_slow1') is added. The name of
//  each state variable (e.g.'proCas-8') is extended by this identifyer (-> e.g. 'proCas-8_slow1').
//  Experimental data referring to this scenario have contain to the new name.
//  The list of state variables and 'new parameters' defines the new scenario by indication of the
//  changes in comparison with the original scenario (here, the initial ligand concentration L_slow1
//  of the scenario '_slow1' is set to the parameter 'I_L_slow1', which has to be defined in the
//  parameter section.
//
////////////////////////////////////////////////////////////////////////////////////////////////////



////////////////////////////////////////////////////////////////////////////////////////////////////
//
// EXPERIMENTAL DATA FOR PARAMETER ESTIMATION
//
//
// TIMESERIES    moleculeName     weight   number   normalization_flag
// time data set
// value data set
// stddev data set
//
// normalization flag: that only the relative values are considered
//  (concentrations are normalized to [0..1])
// weight: weighting factor (currently not used)

EXPERIMENTAL_DATA

TIMESERIES       proCas-8         1       7       0
5        10       20       30       80       120      180      // time
1.4667   1.0667   0.5833   0.2500   0.0600   0.0433   0.010    // value
0.25     0.2      0.15     0.15     0.15     0.1      0.1      // stddev

TIMESERIES       Cas-8_IM_F       1       7       0
5        10       20       30       80       120      180      // time
0.0106   0.0325   0.0750   0.0531   0.0187   0.0056   0.0011   // value
0.015    0.02     0.03     0.025    0.02     0.01     0.01     // stddev

TIMESERIES       Cas-8            1       7       0
5        10       20       30       80       120      180      // time
0.0588   0.1562   0.2063   0.1313   0.0200   0.00625  0.00625  // value
0.02     0.04     0.04     0.03     0.02     0.02     0.02     // steddev

...
```

Appendix B

Model Definition

In this chapter, the model presented in Section 5.2 (Fig. 5.3) is defined in more detail. The definition consists of the reaction network, the parameters and the definition of black boxes. Although many parameter sets were found, for which the model simulation leads to equivalent results, the parameter set providing the best fit with experimental data is presented here for reasons of reproducibility. All diagrams shown in Chapter 9 are based on this parameter set.

B.1 Reaction Schema and Parameters

In this section, the reaction network is defined by a list of all reactions, a list of all molecules and other state variables including initial values and the list of parameters. The parameter list also contains the parameters for the black boxes.

B.1.1 Reaction Schema

Reactions	Rate constants and other parameters[1)]
Elementary Reactions	**k, k_b**
L+R -> L:R [2)]	k_LR
L:R+FADD -> DISCbs	k_DISCbs
DISCbc + procaspase-8 -> DISCbs:procaspase-8	k_DISC_procas8
DISCbs:procaspase-8 + procaspase-8 -> DISCbs:procaspase-8:procaspase-8	k_DISC_procas8_2
DISCbs:procaspase-8:procaspase-8 -> DISCbs:p43/p41	k_DISC_to_cas8_IM
DISCbs:p43/p41 -> DISCbs + caspase-8[3)]	k_cas8_IM_to_cas8
DISCbs + c-FLIPL -> DISCbs:c-FLIPL	k_DISC_FLIP
DISCbs:c-FLIPL + procaspase-8 -> DISCbs:c-FLIPL:procaspase-8	k_DISC_procas8_2
DISCbs:procaspase-8 + c-FLIPL -> DISCbs:c-FLIPL:procaspase-8	k_DISC_FLIP
DISCbs:c-FLIPL:procaspase-8 -> blockedDISCbs + p43/p41	k_DISC_FLIP_to_cas8_IM
DISCbs + c-FLIPS -> blockedDISCbs	k_DISC_FLIP
DISCbs:procaspase-8 + c-FLIPS -> blockedDISCbs	k_DISC_FLIP
caspase-3 + IAP <-> caspase-3:IAP	k_cas39_IAP, kb_cas_IAP
caspase-7 + IAP <-> caspase-7:IAP	k_cas7_IAP, kb_cas_IAP
caspase-9 + IAP <-> caspase-9:IAP	k_cas39_IAP, kb_cas_IAP
Cytochrome C(released) + Apaf-1 <-> Apaf-1:Cytochrome C	k_ApafCyto, kb_ApafCyto
Apaf-1:Cytochrome C + procaspase-9 -> Apoptosome	k_Apoptosom
Apoptosome -> Apaf-1:Cytochrome C + caspase-9	kb_Apoptosom
Smac/Diablo + IAP <-> IAP:Smac/Diablo	k_Smac_IAP, kb_Smac_IAP
Enzymatic reaction / enzyme	**k, K_m[4)]**
procaspase-3 -> caspase-3 / caspase-8	k_cas_8_3, Km_cas_89_3
procaspase-3 -> caspase-3 / caspase-9	k_cas_9_3, Km_cas_89_3
procaspase-2 -> caspase-2 / caspase-3	k_cas_3_2, Km_cas_3_269
procaspase-6 -> caspase-6 / caspase-3	k_cas_3_6, Km_cas_3_269
procaspase-8 -> caspase-8 / caspase-6	k_cas_6_8, Km_cas_6_8
procaspase-7 -> caspase-7 / caspase-8	k_cas_8_7, Km_cas_8_7
procaspase-7 -> caspase-7 / caspase-9	k_cas_9_7, Km_cas_9_7
Bid -> tBid / caspase-2	k_cas2_Bid, Km_cas28_Bid
Bid -> tBid / caspase-8	k_cas8_Bid, Km_cas28_Bid
Apoptosome -> Apaf-1:Cytochrome C + caspase-9 / caspase-3	k_cas_3_9, Km_cas_3_269
PARP -> cPARP / caspase-3 [5)]	k_cas36_apop_activity, Km_cas367_apop_activity
PARP -> cPARP / caspase-7	k_cas7_apop_activity, Km_cas367_apop_activity
Black boxes	**Parameters**
Mitochondria	
See supp. online material (black boxes).	C_MITO_DELTA_BID_BCL T_MITO_RELEASE
Degradation	
See supp. online material (black boxes).	K_DEGRAD K_APOP_ACTIVITY_FLUCT C_APOP_EFF_ACTIVITY K_DEGRAD_PARP K_STEADY_FLUCT

[1)] The parameters refer to supp. table 2

[2)] Example: d[L:R]/dt= k_LR*[L]*[R] - k_DISCbs*[L:R]*[FADD]

[3)] caspase-8 = active form of caspase-8 $(p18/p10)_2$

[4)] $k*[E_t] = V_{max}$

[5)] The PARP cleavage runs in parallel to the 'black box' degradation process.

B.1.2 Molecules and State Variables

Molecule Nr.	Molecule	Initial Value
1	L	I_L
2	R	I_R
3	L:R	0
4	FADD	I_FADD
5	DISCbs	0
6	DISC_procas-8	0
7	DISC_procas-8_2	0
8	blocked DISCbs	0
9	cFLIP_L	I_cFLIP
10	cFLIP_S	I_cFLIP
11	DISC_FLIP	0
12	DISC_FLIP_procas-8	0
13	caspase-8_IM_FINAL	0
14	procas-8	I_procas-8
15	caspase-8_IM	0
16	caspase-8	0
17	procas-2	I_procas-2
18	caspase-2	0
19	procas-3	I_procas-3
20	caspase-3	0
21	procas-6	I_procas-6
22	caspase-6	0
23	procas-7	I_procas-7
24	caspase-7	0
25	procas-9	I_procas-9
26	caspase-9	0
27	Apoptsome	0
28	Apaf-1:Cytochrome C	0
29	Cytochrome C/Smac, released	0
30	caspase-3:IAP	0
31	caspase-7:IAP	0
32	caspase-9:IAP	0
33	IAP	I_IAP
34	Smac/Diablo:IAP	0
35	Bid	I_Bid
36	tBid	0
37	BclXL/2	I_BclXL/2
38	Apaf-1	I_Apaf-1
39	Cytochrome C/Smac, stored	I_Cytchr_Smac_strd
40	PARP	I_PARP
41	cPARP	0
Other quantities		
42	x_apop_activitiy	

B.1.3 Parameters

Param Nr.	Parameter	Value
1	k_LR	1,60E+09
2	k_DISCbs	8,39E+11
3	k_DISC_procas8	2,56E+09
4	k_DISC_procas8_2	7,43E+11
5	k_DISC_to_cas8_IM	9,14E-01
6	I_cFLIP	6,50E-08
7	k_DISC_FLIP_to_cas8_IM	1,11E-02
8	k_DISC_FLIP	1,88E+11
9	k_cas8_IM_to_cas8	4,71E+00
10	k_cas2_Bid	2,17E-01
11	k_cas8_Bid	2,14E+00
12	C_MITO_DELTA_BID_BCL	0,00E+00
13	Km_cas28_Bid	4,01E-06
14	k_cas_9_7	1,13E-03
15	Km_cas_3_269	5,56E-08
16	Km_cas_9_7	6,66E-08
17	k_cas_9_3	1,01E+00
18	k_cas_3_6	7,01E-02
19	k_cas_3_2	1,81E-01
20	k_cas_6_8	1,00E-01
21	Km_cas_6_8	2,21E-08
22	k_ApafCyto	1,92E+10
23	kb_ApafCyto	6,02E-02
24	k_Apoptosom	1,85E+09
25	kb_Apoptosom	1,61E-03
26	k_cas_3_9	2,14E-01
27	T_MITO_RELEASE	7 min
28	k_cas39_IAP	4,58E+13
29	k_cas7_IAP	9,37E+11
30	kb_cas_IAP	1,30E-03
31	k_Smac_IAP	4,77E+10
32	kb_Smac_IAP	5,83E-02
33	k_cas_8_3	1,90E+00
34	Km_cas_89_3	8,90E-08
35	k_cas_8_7	1,26E-09
36	Km_cas_8_7	1,24E-07
37	K_DEGRAD_PARP	1,45E-02
38	Km_cas367_apop_activity	5,60E-06
39	k_cas36_apop_activity	1,28E+01
40	k_cas7_apop_activity	2,56E+02
41	C_APOP_EFF_ACTIVITY	9,22E+00
42	I_Bid	2,32E-07
43	I_Apaf-1	9,14E-07
44	I_Cytchr_Smac_strd	5.11E-07[1)]
45	I_BclXL/2	1.13E-08[1)]
46	I_IAP	1,22E-08
47	I_PARP	1,12E-08
48	I_R	4,43E-07
49	I_L	1.99E-06, 7.96E-08 (5µg, 200ng)
50	I_FADD	2,97E-07
51	I_procas-8	4,43E-07
52	I_procas-3	1,12E-07
53	I_procas-7	1,88E-08
54	I_procas-2	1,58E-07
55	I_procas-9	2,45E-07
56	K_DEGRAD	0.891, 0.184[2)]
57	K_DEGRAD_steady	1,30E-02
58	K_DEGRAD_deathSub	4,47E-03

Units:
First order rate constants and K_DEGRAD values: 1/min
Second order rate constants (e.g. A+B->C): 1/(M*min)
Michaelis-Menten: k-values: 1/min, Km-values: M
Initial concentrations: mol

[1)] The concentrations of Cytochrome C, Smac/Diablo, Bcl-XL and Bcl-2 were replaced by two 'effective' concentrations (see supp. online material, black boxes) - the initial concentration is therefore not identical with the real one.
[2)] These constants refer to the fast and the slow activation scenario, since according to experimental observations, the degradation constant of the black box varies for different activation scenarios.

B.2 Functional Units

The definition of the functional subunits (black boxes), namely the degradation system and the mitochondrial Cytochrome-C release is given here. The biological motivation is provided in Section 5.2.2.

B.2.1 Degradation System

The degradation subsystem (black box 1) is dependent on:

- concentration of caspase-3, -6 and -7 (executioner caspases),
- concentration of PARP.

It affects the following state variables:

- concentration of PARP and cPARP,
- concentration of all molecules due to degradation.

Function

The virtual state variable called x_{apop} quantifies the 'apoptotic activity' (Section 5.2.2). It is assumed that the velocity of the cell's degradation is directly influenced by this activity. It is also assumed that the activity itself is caused by active caspase-3, -6, and -7 and that the increase of the activity runs in parallel to the experimentally observable PARP cleavage. The PARP cleavage and the increase of the apoptotic activity are described as enzymatic processes. In addition, it is assumed that the activity itself is subject to a slow decay. Thus the time evolution of x_{apop} is defined as

$$\begin{aligned}\frac{dx_{\text{apop}}}{dt} &= \frac{k_{\text{cas36\_apop\_activity}} \cdot (1 - x_{\text{apop}}) \cdot x_{\text{caspase}-3}}{K_{\text{m, cas367\_apop\_activity}} + (1 - x_{\text{apop}})} \\ &+ \frac{k_{\text{cas36\_apop\_activity}} \cdot (1 - x_{\text{apop}}) \cdot x_{\text{caspase}-6}}{K_{\text{m, cas367\_apop\_activity}} + (1 - x_{\text{apop}})} \\ &+ \frac{k_{\text{cas7\_apop\_activity}} \cdot (1 - x_{\text{apop}}) \cdot x_{\text{caspase}-7}}{K_{\text{m, cas367\_apop\_activity}} + (1 - x_{\text{apop}})} \\ &- K_{\text{apop\_activity\_decay}} \cdot x_{\text{apop}}\end{aligned}$$

with $x_{\text{apop}} = 0$ for $t = 0$. Note that x_{apop} is normalized to $[0; 1]$. The introduced parameters correspond to the constants of a Michaelis-Menten equation (Section 2.1.3).
The velocity of the degradation process depends on x_{apop} influencing all molecule concentrations. A quadratic relation between x_{apop} and the degradation is assumed in order to account for a superproportional influence of the apoptotic activity. In addition, a steady

degradation of all active caspases like the caspase-8 $(p18/p10)_2$-complex is assumed regardless of the apoptotic activity. The influence of the degradation subsystem on the molecule concentration is therefore given by

$$f_i^1(\mathbf{x}, x_{\text{apop}}) = -x_i \cdot k_{\text{degrad}} \cdot x_{\text{apop}}^2 - \begin{cases} k_{steady\_degrad} \cdot x_i & : \text{ for active caspase complexes,} \\ 0 & : \text{ otherwise.} \end{cases}$$

The degradation constant k_{degrad} was assumed to be the same for all molecules[1].

B.2.2 Mitochondrial Cytochrome-C Release

The mitochondrial subsystem (black box 2) describes the Cytochrome-C release triggered upon an increase of tBid. See Section 5.2.2 for details. The release is dependent on:

- Bcl-Xl/Bcl-2 (anti-apoptotic substrates),
- tBid (pro-apoptotic substrates).

It affects the concentration of:

- Cytochrome-C and Smac/Diablo, stored in mitochondria,
- Cytochrome-C and Smac/Diablo, released.

Function

As soon as tBid reaches a certain level in comparison with the anti-apoptotic substrate concentration of Bcl-Xl and Bcl-2[2], cytochrome-C release is triggered:

> IF $x_{\text{tBid}}(t) > x^*_{\text{Bcl-Xl/Bcl-2}}(t) + C_{\Delta\text{bid,bcl}}$ AND release has not been triggered yet
> THEN $t_{\text{trigger}} = t$

Based on experimental observations [93, 94], the subsequent release is defined by

$$\begin{aligned} x_{\text{cytochr-x,stored}}(t) &= x_{\text{cytochr-x,stored}}(0) \cdot f_{\text{release}}(t - t_{\text{trigger}}), \\ x_{\text{Smac/Diablo,stored}}(t) &= x_{\text{Smac/Diablo,stored}}(0) \cdot f_{\text{release}}(t - t_{\text{trigger}}). \\ \frac{dx_{\text{cytochr-x,released}}}{dt} &= -\frac{dx_{\text{cytochr-x,stored}}}{dt} \\ \frac{dx_{\text{Smac/Diablo,released}}}{dt} &= -\frac{dx_{\text{Smac/Diablo,stored}}}{dt} \end{aligned}$$

The function $f_{\text{release}}(t)$ describes a smooth and complete release within a time interval of length $T_{\text{mito\_release}}$ (see also [34]).

[1] The degradation of x_{PARP} is modelled separately by introduction of $k_{\text{degrad\_PARP}}$, since x_{PARP} is an active part of the degradation process itself.

[2] For computation of the threshold, the concentrations of Bcl-Xl, Bcl-2 and tBid are not directly comparable. The threshold is therefore defined as follows: $x_{\text{tBid}} > const_1 \cdot x_{\text{Bcl-Xl}} + const_2 \cdot x_{\text{Bcl-2}} + C_{\Delta\text{bid,bcl}}$. Since the absolute amounts of Bcl-Xl, Bcl-2 and their anti-apoptotic efficiency are not known, the expression $(const_1 \cdot x_{\text{Bcl-Xl}} + const_2 \cdot x_{\text{Bcl-2}})$ was replaced by an 'efficient' concentration $x^*_{\text{Bcl-Xl/Bcl-2}}$.

Appendix C

Exeprimental Procedures

The experiments presented in this thesis were performed by our collaboration partners Inna Lavrik and Simone Stösser from the lab of Peter H. Krammer at the German Cancer Research Center. For completeness, the experimental procedures of [34] are summarized here. An overview about the methods can be found in [123, 124].

C.1 Cell Cultures and Reagents

To standardize the assays, SKW 6.4 cells were taken from the logarithmic growth phase. To ensure the linear relation between the antigen and the strength of the signal in the western blot, serial dilutions of recombinant proteins or cell lysates were probed with various antibodies. Quantification of chemiluminescence showed good linearity in proportion to the amount of an antigen ([34], supp. Fig. 3). Thus, the following western blot experiments were performed using the same concentrations of the antibodies.

The human B-lymphoblastoid cell line SKW 6.4 was cultured in RPMI 1640 medium supplemented with 10% fetal calf serum, 50 $\mu g/ml$ gentamycin and 5 mM HEPES. Anti-APO-1 (anti-CD95) is an agonistic monoclonal antibody (IgG3) recognizing an epitope on the extracellular part of APO-1 (CD95/Fas) [135]. The C15 mAb (mouse IgG2b) recognizes the p18 subunit of caspase-8 [136].

Apoptosis Assays

Anti-APO-1 antibody was added to SKW 6.4 cells in a concentration range from 5 $\mu g/ml$ to 0.1 ng/ml. The density of cells was 10^6 cells/ml. Samples were incubated at 37^oC for various periods of time. After incubation the samples were divided into three parts. The first part was subjected to analysis of cell viability by Flow Cytometry, the second one was probed for caspase activity and the third one was analyzed by western blot.

Flow Cytometry Analysis

The percentage of viable cells was determined by FSC/SSC using a FACSscan Cytometer (BD). A minimum of 10.000 cells per sample was analyzed.

Western Blot

For western blot analysis the cells were lysed in buffer A (20 mM Tris/HCl, pH 7.4, 1 % Triton X-100, 10 % glycerol, 150 mM NaCl, 1 mM PMSF and 1 $\mu g/ml$ of leupeptin, antipain, chymostatin and pepstatin A) for 15 min. on ice and centrifuged (15 min, 14,000 g). Postnuclear supernatant equivalents of $0.5 \cdot 10^6$ cells or 25 μg of protein as determined by the BCA method (Pierce) were separated on 12% SDS-PAGE and blotted onto a nitrocellulose membrane (Amersham). The sample corresponding to each time point was loaded on each gel twice to estimate the error value. The western blot procedure was performed as described previously [76]. Blots were quantified using the LumiImager TMF1 system and Lumi Analyst software Version 3.0 (Boehringer Mannheim GmbH, Mannheim, Germany). For calibration of LumiImager System we used serial dilutions of the corresponding recombinant proteins (from 100 $fmol$ to 10 $pmol$) or cell lysates (from 1 μg to 50 μg). The measurements were always performed in the region of a linear relation between the amount of antigen and the signal strength. The light signal was measured in Boehringer Light Units (BLU) and in relative amounts. The BLU value calculated is the absolute integration value of the band that was evaluated. The standard deviation was calculated from up to three independent experiments.

Caspase Activity Assays

Cytosolic lysates were incubated with 50M site-specific tetrapeptide substrates (z-IETD-AFC for caspase-8, z-DEVD-AFC for caspase-3) in a caspase assay buffer B (50 mM HEPES, 100 mM NaCl, 10 mM dithiothreitol, 0.1% (w/v) CHAPS, 10% (w/v) sucrose, pH 7.4) in a final volume of 200 μl. The release of the fluorogenic group AFC was determined after 1h of incubation at 37^oC by a microplate fluorescence reader Wallach 1420 (Perkin Elmer) at the excitation wavelength of $405nm$ and emission wavelength of $535nm$.

Cycloheximide experiments

Cycloheximide (CHX) was used at a concentration of 10 $\mu g/ml$ for two hours before the addition of anti-APO-1.

Antibodies and Recombinant Proteins

Anti-APO-1 (anti-CD95) is an agonistic monoclonal antibody (IgG3) recognizing an epitope on the extracellular part of APO-1 (CD95/Fas) [135]. The C15 mAb (mouse IgG2b) recognizes the p18 subunit of caspase-8 (Scaffidi et al., 1997). The NF6 mAB (mouse IgG1) recognizes N-terminal part of c-FLIP (Scaffidi et al., 1997). Anti-caspase-3 (clone 19) and

anti-caspase-2 (clone 35) mABs were from BD Transduction Laboratories. Anti-caspase-7 polyclonal antibody and anti-caspase-9 (clone F-7) mAb were from Santa Cruz Biotechnology (Santa Cruz, CA). Anti-Actin mAb (clone AC-40) was from Sigma (Saint Louis, Missouri). Anti-Bid polyclonal antibody was from Biosource International (Nivelles, Belgium). The HRPO-conjugated goat anti-rabbit IgG was from Santa Cruz Biotechnology (Santa Cruz, CA). The HRPO-conjugated goat anti-mouse IgG1, IgG2a and IgG2b were from Southern Biotechnology Associates (Birmingham, AL). Recombinant human tBid, caspase-3, caspase-9, caspase-8 were from Apotech (Lausanne, Switzerland). Recombinant Caspase-2 was from Sigma (Saint Louis, Missouri).

C.2 Standard Deviation

For each activation scenario, time series of molecule concentrations were measured. These experiments were repeated several times (at least 3 times for most important molecules). However, the quality of the biological experiments varied in some cases. In order to avoid artifacts caused by outliers and noisy data, the parameter estimation was based on time series of the experiment with the highest quality in these cases, instead of averaging data of different qualities. Nevertheless, the standard deviations σ_{ij} were based on deviations between the different experiments. For most molecules, the relative standard deviation was about 20% on average. For data points, which are close to zero after start of degradation, a larger relative standard deviation was chosen. Otherwise, even small differences between simulation and experiment of these points would already lead to unreasonably high contributions to the least square term, misleading the parameter estimation.

Bibliography

[1] J. D. Murray. Mathematical Biology (Springer, Berlin, 2002).

[2] H. Kitano. Systems biology: a brief overview. Science 295, 1662 (2002).

[3] B. Alberts, D. Bray, J. Lewis, M. Raff, K. Roberts, J. D. Watson. Molecular Biology of the Cell (Garland, New York, 2002).

[4] A. G. Gilman. Overview of the alliance for cellular signaling. Nature 420,703 (2002).

[5] R. Hofestädt and S. Thelen. Quantitative Modeling of Biochemical Networks. In silico Biology 1, 39 (1998).

[6] W. Reisig. Petri Nets, An Introduction (Springer, Berlin, 1985).

[7] P. J. Goss and J. Peccoud. Quantitative modeling of stochastic systems in molecular biology by using stochastic Petri nets. PNAS 95, 6750 (1998).

[8] S. Wolfram. Cellular Automata and Complexity: Collected Papers. (Addison-Wesley, Reading, MA, 1994).

[9] D. B. Kell, H. V. Westerhoff. Metabolic control theory: its role in microbiology and biotechnology. FEMS Microbiol. Rev. 39, 305 (1986).

[10] R. Heinrich and S. Schuster. The regulation of cellular systems (Chapman & Hall, New York, 1996).

[11] C. H. Schilling, S. Schuster, B. O. Palsson, R. Heinrich. Metabolic pathway analysis: basic concepts and scientific applications in the post-genomic era. Biotechnol. Prog. 15, 296 (1999).

[12] M. E. Csete and J. C. Doyle. Reverse engineering of biological complexity. Science 295, 1664 (2002).

[13] D. A. Lauffenburger. Cell signaling pathways as control modules: complexity for simplicity? PNAS 97, 5031 (2000).

[14] D. Garfinkel and B. Hess. Metabolic Control Mechanisms. J. Biol. Chem. 239, 971 (1964).

[15] D. Garfinkel. The role of computer simulation in biochemistry. Comput. Biomed. Res. 2: i-ii (1968).

[16] U. S. Bhalla and R. Iyengar. Emergent properties of networks of biological signaling pathways. Science 283, 381 (1999).

[17] P. Mendes. Biochemistry by numbers: simulation of biochemical pathways with Gepasi 3. Trends Biochem. Sci. 22, 361 (1997).

[18] P. Mendes. GEPASI: A software package for modelling the dynamics, steady states and control of biochemical and other systems. Comput. Applic. Biosci. 9, 563 (1993).

[19] M. Tomita, K. Hashimoto, et al. E-CELL: Software environment for whole cell simulation. Bioinformatics, 15, 72 (1999).

[20] B. E. Shapiro, A. Levchenko, E. D. Mjolsness. Automatic Model Generation with Applications to MAP-Kinase Pathways, in Foundations of Systems Biology (MIT Press, Cambridge, Massachusetts, 2001).

[21] H. M. Sauro, D. A. Fell. SCAMP: A Metabolic Simulator and Control Analysis Program. Mathl. Comput. Modelling, 15, 15-28 (1991).

[22] B. N. Kholodenko, O. V. Demin, G. Moehren, J. B. Hoek. Quantification of short term signaling by the epidermal growth factor receptor. J. Biol. Chem. 274, 30169 (1999).

[23] B. Schoeberl, C. Eichler-Jonsson, E. D. Gilles, G. Müller. Computational modeling of the dynamics of the MAP kinase cascade activated by surface and internalized EGF receptors. Nature Biotech 20, 370 (2002).

[24] P. Deuflhard. Numerical treatment of inverse problems in differential and integral equations (Birkhäuser, Basel, 1983).

[25] P. Mendes and D.B. Kell. Non-linear optimization of biochemical pathways: applications to metabolic engeneering and parameter estimation. Bioinformatics 14, 869 (1998).

[26] H. G. Bock. Numerical treatment of inverse problems in chemical reaction kinetics. In K. H. Ebert, P. Deuflhard, W. Jäger (Eds.), Modelling of Chemical Reaction Systems, 8, 102 (Springer, New York, 1981).

[27] I. Swameye, T. G. Müller, J. Timmer, O. Sandra, U. Klingmüller. Identification of nucleocytoplasmic cycling as a remote sensor in cellular signaling by data-based modeling. PNAS 100, 1028 (2003).

[28] M. Fussenegger, J. Bailey, J. Varner. A mathematical model of caspase function in apoptosis. Nature Biotech. 18, 768 (2000).

[29] U. Alon, M. G. Surette, N. Barkai, S. Leibler. Robustness in bacterial chemotaxis. Nature 397, 168 (1999).

[30] N. Barkai, S. Leibler. Robustness in simple biochemical networks. Nature 387, 913 (1997).

[31] J. M. Carlson, J. C. Doyle. Complexity and robustness. PNAS 99, 2538 (2002).

[32] A. Kremling, E. D. Gilles. The organization of metabolic reaction networks: II. Signal processing in hierarchical structured functional units. Metabolic Engineering, 3, 138-150 (2001).

[33] A. Varma, M. Morbidelli, H. Wu. Parametric Sensitivity in Chemical Systems (Cambridge University Press, New York, 1999).

[34] M. Bentele, I. Lavrik, M. Ulrich, S. Stösser, D.W. Heermann, H. Kalthoff, P.H. Krammer, R. Eils. Mathematical modeling reveals threshold mechanism in CD95-induced apoptosis. Journal of Cell Biology, Vol. 166, 839 (2004).

[35] N. Le Novre and T. S. Shimizu. STOCHSIM: modelling of stochastic biomolecular processes. Bioinformatics 17, 575 (2001).

[36] D. T. Gillespie. Exact stochastic simulation of coupled chemical reactions. J. Phys. Chem. 81, 2340-2361 (1977).

[37] N. G. van Kampen. Stochastic processes in physics and chemistry (North Holland, 1981).

[38] P. Krammer. CD95's deadly mission in the immune system. Nature 407, 789 (2000).

[39] M. E. Peter, P. H. Krammer. The CD95(APO-1/Fas) DISC and beyond. Cell Death Differ. 10, 26 (2003).

[40] A. Arkin, J. Ross, H. H. McAdams. Stochastic Kinetic Analysis of a Developmental Pathway Bifurcation in Phage-l Escherichia coli. Genetics 149, 1633 (1998).

[41] P. Erdi, J. Todt, in Mathematical models of chemical reactions (Princeton University Press, 1989)

[42] C. W. Gear. Numerical Initial Value Problems in Ordinary Differential Equations, Chapter 2 (Englewood Cliffs, 1971).

[43] P. Deuflhard, F. Bornemann. Numerische Mathematik II. Integration gewöhnlicher Differentialgleichungen (de Gruyter: Berlin, New York, 1994).

[44] P. Deuflhard, F. Bornemann. Scientific Computing with Ordinary Differential Equations. In: Applied Mathematics 42 (Springer, New York, 2002).

[45] J. Stoer, R. Bulirsch. Introduction to Numerical Analysis (Springer, New York, 2002).

[46] G. D. Byrne, A. C. Hindmarsh. Stiff ODE Solvers: A Review of Current and Coming Attractions. J. Comp. Phys., 70, 1 (1987).

[47] L. R. Petzold, U. M. Ascher. Computer Methods for Ordinary Differential Equations and Differential-Algebraic Equations (Society for Industrial & Applied Mathematics, 1998).

[48] P. W. Atkins. Physical Chemistry (Oxford University Press, 2001).

[49] L. Stryer. Biochemistry (WH Freeman, New York, 1995).

[50] D. A. McQuarrie. Stochastic Approach to Chemical Kinetics. J. Appl. Prob., 4, 413 (1967).

[51] D. T. Gillespie. A General Method for Numerically Simulating the Stochastic Time Evolution of Coupled Chemical Reactions. J. Comp. Phys. 22, 403 (1976).

[52] D. T. Gillespie. A rigorous derivation of the chemical master equation. Physica A 188, 404 (1992).

[53] C. W. Gardiner. Handbook of Stochastic Methods for Physics, Chemistry and the Natural Sciences (Springer, Berlin, 1985).

[54] H. Risken. The Fokker-Planck Equation: Methods of Solution and Applications (Springer, Berlin, 1989).

[55] P. E. Kloeden, E. Platen. Numerical Solution of Stochastic Differential Equations (Springer, Berlin, 1992).

[56] J. L. Doob. Stochastic Processes (Wiley, New York, 1953).

[57] D. T. Gillespie. The multivariate Langevin and Fokker-Planck equations. Am. J. Phys. 64, 1246 (1976).

[58] D. T. Gillespie. The chemical Langevin equation. J. Chem. Phys. 113, 297 (2000).

[59] M. Bentele and R. Eils. General stochastic hybrid method for the simulation of chemical reaction processes in cells. CMSB04, Lecture Notes in Computer Science, Springer, Heidelberg, in press (2004).

[60] D. W. Heermann. Computer Simulation Methods in Theoretical Physics (Springer, Heidelberg, 1990).

[61] M. A. Gibson and J. Bruck. Efficient Exact Stochastic Simulation of Chemical Systems with Many Species and Many Channnels. J. Phy. Chem. 104, 1876 (2000).

[62] K. Burrage and T. Tian. Poisson Runge-Kutta methods for Chemical Reaction Systems. Proceedings of the Hong Kong Conference on Scientific Computing (2003).

[63] B. D. Ripley. The computer generation of random variables: a tutorial. Int. Statist. Rev. 51, 301 (1983).

[64] B. D. Ripley. Stochastic Simulation, 237pp (Wiley, New York, 1987).

[65] R. Y. Rubinstein. Simulation and the Monte-Carlo Method (Wiley, New York, 1981).

[66] B. J. Morgan. Elements of Simulation (Chapman & Hall,London, 1984).

[67] T. R. Kiehl, R. M. Mattheyses, M. K. Simmons. Hybrid simulation of cellular behavior. Bioinformatics 20, 316 (2004).

[68] G. Evan and T. Littlewood. A matter of life and cell death. Science 281, 1317 (1998).

[69] S. Nagata. Fas ligand-induced apoptosis. Annu. Rev. Genet. 33, 29 (1999).

[70] A. Ashkenazi and V. Dixit. Apoptosis control by death and decoy receptors. Curr. Opin. Cell Biol. 11, 255 (1999).

[71] G. S. Salvesen. Caspases: opening the boxes and interpreting the arrows. Cell Death Differ. 9, 3 (2002).

[72] N. Thornberry, and Y. Lazebnik. Caspases: Enemies Within. Science 281, 1312 (1998).

[73] N. N. Danial and S. J. Korsmeyer. Cell Death: Critical Control Points. Cell 116, 205 (2004).

[74] S. Nagata. Apoptosis by Death Factor. Cell 88, 355 (1997).

[75] I. Lavrik, A. Krueger, et al. The active caspase-8 heterotetramer is formed at the CD95 DISC. Cell Death Differ. 10, 144 (2003).

[76] Krueger, A., S. Baumann, P. H. Krammer, S. Kirchhoff. FLICE-Inhibitory Proteins: Regulators of Death Receptor-Mediated Apoptosis. Mol. Cell. Biol. 21, 8247 (2001).

[77] U. Fischer, R. U. Janicke, K. Schulze-Osthoff. Many cuts to ruin: a comprehensive update of caspase substrates. Cell Death Differ. 10, 76 (2003).

[78] S. Nagata, H. Nagase, K. Kawane, N. Mukae, H. Fukuyama. Degradation of chromosomal DNA during apoptosis. Cell Death Differ. 10, 108 (2003).

[79] C. Scaffidi, S. Fulda, et al. Two CD95 (APO-1/Fas) signaling pathways. The EMBO Journal 17, 1675 (1998).

[80] I. Schmitz, H. Walczak, P. H. Krammer, M. E. Peter. Differences between CD95 type I and tpye II cells detected with the CD95 ligand. Cell Death Differ. 6, 821 (1999).

[81] O. von Ahsen, N. J. Waterhouse, T. Kuwana, D. D. Newmeyer, D. R. Green. The 'harmless' release of cytochrome c. Cell Death Differ. 7, 1192 (2000).

[82] H. Zou, R. Yang, J. Hao, J. Wang, C. Sun, et al. Regulation of the Apaf-1/Caspase-9 Apoptosome by Caspase-3 and XIAP. J. Biol. Chem. 278, 8091 (2003).

[83] G. S. Salvesen, C. S. Duckett. Apoptosis: IAP proteins: blocking the road to death's door. Nature Reviews Molecular Cell Biology 3, 401 (2002).

[84] D. T. Chao, S. J. Korsmeyer. BCL-2 Family: Regulators of Cell Death. Annu. Rev. Immunol. 16, 395 (1998).

[85] C. S Mitsiades, N. Mitsiades, V. Poulaki, R. Schlossman, M. Akiyama, et al. Activation of NFκ-B and upregulation of intracellular anti-apoptotic proteins via the IGF-1/Akt signaling in human multiple myeloma cells: therapeutic implications. Oncogene 21, 5673 (2002).

[86] A. A. Beg and D. Baltimore. An Essential Role for NF-B in Preventing TNF-Induced Cell Death. Science 274, 782 (1996).

[87] S. Westphal, H. Kalthoff. Apoptosis: targets in pancreatic cancer. Mol. Cancer. 2, 6 (2003).

[88] F. Schacherer, C. Choi, U. Götze, M. Krull, S. Pistor, E. Wingender. The TRANSPATH signal transduction database: a knowledge base on signal transduction networks. Bioinformatics 17, 1053 (2001).

[89] www.biobase.de.

[90] J. P. Medema. Receptors to die for. Cell Death and Differentiation 10, 1 (2003).

[91] J. Schlessinger, A. Ullrich. Growth factor signaling by receptor tyrosine kinases. Neuron 9, 383 (1992).

[92] G. Hämmerlin, K. H. Hoffmann. Numerische Mathematik (Springer, Berlin, Heidelberg 1994).

[93] J. C. Goldstein, N. J. Waterhouse, et al. The coordinate Release of Cytochrome C during Apoptosis is rapid, complete and kinetically invariant. Nature Cell Biology 2, 156 (2000).

[94] M. Luetjens, D. Kögel, C. Reimertz, H. Düssmann, A. Renz, et al. Multiple Kinetics of Mitochondrial Cytochrome c Release in Drug-Induced Apoptosis. Mol. Pharmacol. 60, 1008 (2001).

[95] P. M. Frank. Introduction to system sensitivity theory (Academic, New York, 1978).

[96] R. Tomovic and M. Vukobratovic. General Sensitivity Theory. (Elsevier, New York, 1972).

[97] V. V. Fedorov, P. Hackl. Model-Oriented Design of Experiments (Springer, New York, 1997).

[98] H. Rabitz, M. Kramer, D. Dacol, Sensitivity analysis in chemical kinetics. Ann. Rev. Phys. Chem. 34, 419 (1983).

[99] D. A. Fell. Metabolic control analysis: a survey of its theoretical and experimental development. Biochem. J. 286, 313 (1992).

[100] E. Meir, G. von Dassow, E. Munro, G. M. Odell. Robustness, flexibility, and the role of lateral inhibition in the neurogenic network. Current Biology, 12, 778 (2002).

[101] M. A. Savageau. Parameter sensitivity as a criterion for evaluating and comparing the performance of biochemical systems. Nature 229, 542 (1971).

[102] L. Ma, P. A. Iglesias. Quantifying robustness of biochemical network models. BMC Bioinformatics 3, 38 (2002).

[103] M. Morohashi, A. E. Winn, M. T. Borisuk, H. Bolouri, J. Doyle, H. Kitano. Robustness as a Measure of Plausibility in Models of Biochemical Networks. Journal of Theoretical Biology 216, 19 (2002).

[104] H. R. Stennicke, G. S. Salvesen. Catalytic properties of the caspases. Cell Death Differ. 6, 1054 (1999).

[105] A. Kremling, K. Jahreis, J. W. Lengeler, E. D. Gilles. The organization of metabolic reaction networks: A signal-oriented approach to cellular models. Metabolic Engineering, 2, 190 (2000).

[106] I. Ioslovich, P. O. Gutman, I. Seginer. Dominant Parameter Selection in The Marginally Identifiable Case. Journal of Mathematics and Computers in Simulation, 65, 127 (2004).

[107] G. Li, C. Rosenthal, H. Rabitz. High Dimensional Model Representations. J. Phys. Chem. A 105, 7765 (2001).

[108] N. Gershenfeld. The Nature of Mathematical Modeling (Cambridge University Press, Cambridge, UK, 1999).

[109] L. Ljung. System Identification Theory for the User, Prentice-Hall, Englewood Cliffs, NJ (1999).

[110] K. Levenberg. A method for the solution of certain problems in least squares. Quart. Appl. Math. 2, 164 (1944).

[111] D. Marquardt. An algorithm for least-squares estimation of nonlinear parameters. SIAM J. Appl. Math. 11, 431 (1963).

[112] D. M. Bates and D. G. Watts. Nonlinear Regression and its Applications. (Wiley, New York, 1988).

[113] S. Kirkpatrick, C. D. Gelatt, M.P. Vecchi. Science 220, 671 (1983).

[114] H. G. Bock. Randwertproblemmethoden zur Parameteridentifizierung in Systemen nichtlinearer Differentialgleichungen. PhD thesis, University Bonn, Bonner Mathematische Schriften Nr. 183 (1987).

[115] D. B. Leineweber. The theory of MUSCOD in a nutshell. Diploma thesis, Universitt Heidelberg, IWR-Preprint 96-19 (1996).

[116] S. Mesecke. Quantitative Modellierung und Simulation von Signaltransduktion. Diploma thesis, Universitt Hannover (2004).

[117] I. Vacheva, M. Bentele, R. Eils. Optimal experiment design for discriminating between competing signal transduction models, in preparation.

[118] S. Koerkel, Numerische Methoden für Optimale Versuchsplanungsproblem bei nichtlinearen DAE-Modellen. PhD thesis, IWR, Uni Heidelberg (2003).

[119] M. Bentele and R. Eils. Apoptosis. In: Systems Biology. Springer, Heidelberg, in press (2004).

[120] W. Paul, D. W. Heermann, R. C. Desai. Implementation of a Random Number Generator in OCCAM. J. Comput. Phys. 82, 487 (1989).

[121] M. Matsumoto, T. Nishimura. Mersenne Twister: A 623-dimensionally equidistributed uniform pseudorandom number generator. ACM Trans. on Modeling and Computer Simulation 8, 3 (1998).

[122] G. Lewis, W. H. Payne. Generalized Feedback Shift Register Pseudorandom Number Algorithm. J. ACM 20, 456 (1973).

[123] C. A. Janeway, P. Travers, M. Walport, M. Shlomchik. Immunobiology (Garland Publishing, New York, 2001).

[124] F. Lottspeich, H. Zorbas. Bioanalytik (Spektrum Akademischer Verlag, Heidelberg, Berlin, 1998).

[125] A. Krueger, I. Schmitz, S. Baumann, P. H. Krammer, S. Kirchhoff. Cellular FLICE-inhibitory protein splice variants inhibit different steps of caspase-8 activation at the CD95 death-inducing signaling complex. J. Biol. Chem. 276, 20633 (2001).

[126] J. Silke, P. Ekert, C. Day, C. Hawkins, M. Baka, J. Chew. Direct inhibition of caspase 3 is dispensable for the anti-apoptotoc activity of XIAP. EMBO Journal. 20, 3123 (2001).

[127] L. E. French, J. Tschopp. Thyroiditis and hepatitis: Fas on the road to disease. Nature Med. 3, 385 (1997).

[128] O. Micheau. Cellular FLICE-inhibitory protein: an attractive therapeutic target? Expert. Opin. Ther. Targets. 7, 559 (2003).

[129] A. L. Barabasi, R. Albert. Emergence of scaling in random networks. Science 286, 509 (1999).

[130] H. Jeong, B. Tombor, R. Albert, Z. N. Oltvai, A. L. Barabasi. The large-scale organization of metabolic networks. Nature 407, 651 (2000).

[131] R. Albert, H. Jeong, A. L. Barabasi. Error and attack tolerance of complex networks. Nature 406, 378 (2000).

[132] K. S. Brown, J. P. Sethna. Statistical mechanical approaches to models with many poorly known parameters. Phys. Rev. E 68, 21904 (2003).

[133] M. Ulrich , M. Bentele, I. Lavrik, P. Krammer, R. Eils. Spatial modeling of DISC-formation and caspase-8 activation in CD95-induced apoptosis, in preparation.

[134] I. Lavrik, M. Bentele, A. Golks, R. Eils, P. H. Krammer. CD95 signaling under threshold conditions reveals different mechanism, in preparation.

[135] B. C. Trauth, C. Klas, A. M. Peters, S. Matzku, P. Moller, W. Falk, K. M. Debatin, P.H. Krammer. Monoclonal antibody-mediated tumor regression by induction of apoptosis. Science 245, 301 (1989).

[136] C. Scaffidi, J. P. Medema, P. H. Krammer, M.E. Peter. FLICE is predominantly expressed as two functionally active isoforms, caspase-8/a and caspase-8/b. J. Biol. Chem. 272, 26958 (1997).

List of Publications

W. Tvarusko, **M. Bentele**, T. Misteli, R. Rudolf, C. Kaether, D.L. Spector, H.H. Gerdes, R. Eils. Time-resolved analysis and visualization of dynamic processes in living cells. *PNAS*, Vol. 96, 7950 (1999).

M. Bentele and R. Eils. General stochastic hybrid method for the simulation of chemical reaction processes in cells. *In:* CMSB04, Lecture Notes in Computer Science, Springer, Heidelberg, *in press* (2004).

M. Bentele, I. Lavrik, M. Ulrich, S. Stösser, D.W. Heermann, H. Kalthoff, P.H. Krammer, R. Eils. Mathematical modeling reveals threshold mechanism in CD95-induced apoptosis. *Journal of Cell Biology*, Vol. 166, 839 (2004).

M. Bentele and R. Eils. Systems Biology of Apoptosis. *In:* Systems Biology, Springer, Heidelberg, *in press* (2004).

I. Vacheva, **M. Bentele**, R. Eils. Optimal experiment design for discriminating between competing signal transduction models, *in preparation.*

M. Bentele, I. Lavrik, R. Eils. Model-based analysis of stochastic effects in CD95-induced apoptosis, *in preparation.*

I. Lavrik, **M. Bentele**, A. Golks, R. Eils, P. H. Krammer. CD95 signaling under threshold conditions reveals different mechanism, *in preparation*

M. Ulrich , **M. Bentele**, I. Lavrik, P. Krammer, R. Eils. Spatial modeling of DISC-formation and caspase-8 activation in CD95-induced apoptosis, *in preparation.*

Selected Talks

Large-scale Simulation of Signal Transduction in Cells using Hybrid Models, *International Conference on High Performance Scientific Computing 2003*, Hanoi, Vietnam

Stochastic Simulation of large Signal Transduction Systems in Cells using Hybrid Methods, *DECHEMA 2003*, Munich, Germany

General stochastic hybrid method for the simulation of chemical reaction processes in cells, *Computational Methods in Systems Biology 2004*, Paris, France

Quantitative Approaches for Large-Scale Modeling and Simulation of Signal Transduction, *International Workshop on Parameter Estimation and Optimal Design of Experiments 2004*, Heidelberg, Germany

New quantitative approaches for modeling and simulation of large signal transduction networks, *International Conference on Systems Biology 2004*, Heidelberg, Germany

Invited Talks at the University Heidelberg (Interdisciplinary Center for Scientific Computing) and at the University Bielefeld, Germany.

Acknowledgements

I would like to thank Prof. Dieter W. Heermann for the supervision and the support of this thesis and for inspiring me with the topic of computer simulation methods by his interesting lectures. Prof. Roland Eils deserves many thanks for the great opportunities he gave me in his research group at the DKFZ, for his strong interest in this thesis, for his commitment and steady support, for many helpful discussions and valuable contacts to people in this field of research.

Many thanks to Dr. Inna Lavrik, Simone Stösser and Prof. Peter H. Krammer for the excellent collaboration and for the outstanding work on the experimental side. I also would like to thank Prof. Holger Kalthoff, Dr. Anna Trauzold, Matthias Kühl, Dr. Sabine Westphal and Dr. Christian Röder from the University of Kiel for their experimental data and for contributions during the initiation of this project.

I want to express my gratitude to Dr. Johannes Schlöder, Dr. Dirk Lebiedz, Ulrich Brandt-Pollmann, Prof. Georg Bock and Prof. Willi Jäger from the Interdisciplinary Center for Scientific Computing for many important discussions and their support concerning the parameter estimation. Many thanks to Dr. Jens Timmer from the University of Freiburg for all the discussions and his reviews.

I also would like to thank all my colleagues for the friendly atmosphere and the cooperativeness in our group; in particular Markus Ulrich for his contributions, Hauke Busch for his critical reviews and Ivayla Vacheva, Rainer König, Sven Mesecke, Stefan Hezel and Dietmar Volz for many fruitful discussions. Special thanks to Marco Weismüller for his help and his contributions in early stages of this project. Furthermore, I want to thank Christina Grosch and Karlheinz Gross for their support and the professional organization and IT-infrastructure of the research group. I also want to thank Accenture GmbH for giving me the opportunity to take a sabbatical and to do this thesis. Last but not least, I would like to thank my parents for their support.