

INAUGURAL-DISSERTATION

zur

Erlangung der Doktorwürde

der

Naturwissenschaftlich-Mathematischen Gesamtfakultät

der

RUPRECHT-KARLS-UNIVERSITÄT

HEIDELBERG

vorgelegt von

Dipl.-Math. Andreas Sommer

geboren in Trier

Tag der mündlichen Prüfung:

.....

Numerical Methods for Parameter Estimation in Dynamical Systems with Noise

with Applications in Systems Biology

Andreas Sommer

Gutachter:

Prof. Dr. Dr. h. c. mult. Hans Georg Bock

.....

Kurzbeschreibung

Die vorliegende Arbeit behandelt die Modellierung von und Parameterschätzung in dynamischen Systemen mit Fokus auf Anwendungen in der Systembiologie. In dieser interdisziplinären Forschungsarbeit zur Systembiologie von Krebserkrankungen wird ein prädiktives mathematisches Modell intrazellulärer Wechselwirkungen zweier Zytokinsignaltransduktionswege entwickelt und im Experiment bestätigt. Unerwartete Vorhersagen führen auf neue biologische Erkenntnisse. Zur Modellkalibrierung kommen klassische Methoden der Parameterschätzung in gewöhnlichen Differentialgleichungen zum Einsatz. Daran anknüpfend wird in dieser Arbeit eine neue numerische Methode zur Parameterschätzung in stochastischen Einflüssen unterliegenden dynamischen Prozessen entwickelt, theoretisch untersucht, implementiert und ihre Leistungsfähigkeit an Beispielen aus Systembiologie und Finanzwissenschaft demonstriert.

Viele Prozesse, insbesondere biologische, zeigen ein grundsätzlich gerichtetes Verhalten, das zwar gewissen Gesetzmäßigkeiten unterliegt (etwa Stoffwechselprozesse oder Signaltransduktionen), allerdings auch durch intrinsische Zufälligkeiten qualitativ und quantitativ erheblich beeinflusst wird, sodass eine rein deterministisch-mechanistische Modellierung der auftretenden Vorgänge oft nicht zielführend ist.

Eine große Klasse solcher Systeme lässt sich durch nichtlineare mehrdimensionale stochastische Differentialgleichungen (SDEs) adäquat beschreiben. Klassische stochastische Schätzverfahren, die auf (Approximation von) Übergangswahrscheinlichkeiten basieren, sind auf diese Problemklasse u. a. wegen ihres Rechenaufwands und speziellen Anforderungen an die Messdaten oft nicht oder nur eingeschränkt anwendbar.

In dieser Arbeit wird ein neues Mehrzielverfahren entwickelt, das die dem beobachteten Prozess entsprechende Realisierung der SDE stückweise durch deterministische Lösungen gewöhnlicher Differentialgleichungen (ODEs) approximiert. Diese i. d. R. un stetige Konkatenation der Trajektorien erlaubt zum einen die Wiedergabe stochastischer Effekte, zum anderen die Darstellung des Parameterschätzproblems als *deterministisches* nichtlineares Optimierungsproblem, das unter Verwendung effizienter *ableitungsbasierter* Verfahren gelöst werden kann. In der vorliegenden Arbeit kommt hierzu ein verallgemeinertes GAUSS-NEWTON-Verfahren zum Einsatz.

Die wesentlichen Ergebnisse und Resultate der vorliegenden Arbeit umfassen insbesondere:

- Es wird ein neues Verfahren zur Parameterschätzung in mehrdimensionalen nichtlinearen SDEs entwickelt, das auf stückweiser Approximation durch Lösungen von ODEs basiert. An Intervallgrenzen auftretende Sprünge werden zur Regularisierung verwendet. Unbekannte Parameter und initiale Systemzustände werden mittels einer verallgemeinerten, gewichteten Kleinste-Quadrate-Methode aus Messdaten geschätzt, die aus direkten Zustandsmessungen oder aus durch Messfunktionen beschriebenen indirekten Beobachtungen stammen, mit Fehlern behaftet und an beliebigen Zeitpunkten erfasst worden sein können. Nichtlineare Parameter- und punktweise Zustandsbeschränkungen können als Gleichungs- und Ungleichungsbedingungen formuliert werden. Die resultierenden nichtlinearen beschränkten Optimierungsprobleme werden unter Strukturausnutzung mit einem verallgemeinerten GAUSS-NEWTON-Verfahren effizient gelöst.
- Es wird ein Beweis gegeben, dass im Falle einer asymptotisch gegen unendlich gehenden Zahl von äquidistanten Mehrzielknoten die Unstetigkeiten (Sprünge) an den Intervallgrenzen gegen null gehen.
- Eine numerische Analyse offenbart die Dünnbesetztheitsstruktur der auftretenden Gleichungssysteme. Es wird gezeigt, dass die Anzahl der Nichtnulleinträge nur linear in der

Zahl der Mehrzielknoten (Zeitgitter) wächst und einer scharfen oberen Schranke genügt. Zusätzlich wird bewiesen, dass sich unter Verwendung einer geeigneten (stabilen) Zerlegung die Dünnbesetztheit erhalten lässt.

- Eine vergleichende Simulationsstudie zeigt, dass die Schätzungen sehr robust hinsichtlich der Wahl der Sprunggewichte sind. Zusätzlich wird dargelegt, welchen Einfluss die Intensität der Sprungregularisierung auf Schätzungen und approximierten Trajektorien hat.
- Es wird ein Lifting-Ansatz mit intervallweisen Parametersätzen, gekoppelt durch zusätzliche Gleichungsbedingungen, entwickelt und seine numerischen Eigenschaften analysiert. Weiter wird ein Homotopieverfahren zur Behandlung schwer lösbarer Probleme vorgeschlagen.
- Die Leistungsfähigkeit der Methode wird an Beispielen aus der Systembiologie, die jeweils unterschiedliche Aspekte beleuchten, demonstriert. Insbesondere wird gezeigt, dass sich die Methode auch zur Rekonstruktion von unbeobachteten Zuständen (engl. *hidden state estimation*), sowie zur vollständigen Rekonstruktion von Trajektorien in Zeiträumen, in denen keine Beobachtungen vorliegen, einsetzen lässt. Weiter wird ein Kriterium zur lokalen Verfeinerung des Mehrzielgitters gegeben.
- Für einen durch einen LÉVY-Sprungprozess getriebenen ORNSTEIN-UHLENBECK-Prozess (modifiziertes VAŠÍČEK-Zinsmodell) wird gezeigt, dass sich neben Gleichgewichtsniveau und Steifigkeit aus den Sprungresiduen auch die Diffusionskonstante schätzen lässt.
- Im Softwarepaket `:sfitt` wird eine effiziente Implementierung der in dieser Arbeit entwickelten Parameterschätzmethode bereitgestellt, die auf Nutzerseite eine einfache symbolische Problemformulierung erlaubt.

Teile dieser Arbeit entstanden im interdisziplinären Forschungsprojekt *SBCancer* der Helmholtz-Allianz Systembiologie und beschäftigen sich mit der Modellierung der Wechselwirkung zweier an Krebs beteiligter Zytokin-Signalkaskaden in menschlichen Hautzellen. Vom Autor wurde in enger Kooperation mit den beteiligten Biologen ein mathematisches Modell entwickelt und durch Messdaten kalibriert. Die mittels diesem Modell berechneten kontraintuitiven Vorhersagen eines vorgeschlagenen Doppelstimulationsexperimentes konnten im Labor bestätigt werden und führten auf neue biologische Erkenntnisse.

Wesentliche Beiträge und Neuerungen in dieser Arbeit sind die folgenden:

- Entwicklung eines mathematischen Modells eines *Crosstalks* der Signalwege zweier Zytokine in menschlichen Keratinozyten (Zelllinie HaCaT). Die vom Modell vorhergesagte und zuvor unbekannt nichtlineare moderierende Wirkung von GM-CSF auf den IL-6-induzierten JAK-STAT-Signalweg wurde im Labor *in vitro* nachgewiesen.
- Eine aufwendige mathematische Analyse des in der Zellbiologie häufig eingesetzten *quantitativen WESTERN-blot*-Messverfahrens zeigt, dass etablierte Normalisierungstechniken, die auf in angenommen konstanter Konzentration vorliegenden Haushaltsproteinen (engl. *housekeeping proteins*) oder auf manuell hinzugefügten Kalibrierungsproteinen beruhen, extrem anfällig für signalzerstörende statistische Artefakte sind. Zusätzlich wird gezeigt, dass sich die häufig getroffene Annahme normalverteilter Fehler nach Anwendung dieser Normalisierung i. d. R. nicht halten lässt.
- Eine auf Verstärkungsfaktoren beruhende Methode wird als Ersatz vorgeschlagen. Es werden leicht nachzuprüfende Kriterien an die Rohdaten entwickelt, anhand derer entschieden werden kann, ob die auf diese Weise prozessierten Messdaten weiterhin (approximativ) normalverteilt sind. Des Weiteren werden in einer vergleichenden Simulationsstudie die Vorteile der vorgeschlagenen Verstärkungsfaktormethode (engl. *amplification factors method*) demonstriert.

Abstract

This thesis comprises the modelling of and parameter estimation in dynamical systems, with a focus on applications in systems biology. In an interdisciplinary research project on the systems biology of cancer, we develop a predictive mathematical model of an intracellular crosstalk in cytokine signalling. Expected and unexpected predictions are confirmed in experiments and lead to new biological insights. For model calibration with measurement data, we apply well established methods for parameter estimation in ordinary differential equation models. Extending these to stochastic differential equations, we develop, analyse, and implement a new method for parameter estimation in dynamical processes with noise, and demonstrate its performance in several selected examples from systems biology and mathematical finance.

Many processes, especially in biology, obey deterministic ground rules (e.g. metabolic processes or signal transduction pathways), but may be heavily influenced by fluctuations and stochasticity inherent to the system that change its behaviour both qualitatively and quantitatively. Therefore, frequently, a deterministic description is not constructive.

A large class of such systems can be adequately described by nonlinear multi-dimensional stochastic differential equations (SDEs). Classical estimation techniques for SDEs, relying on (approximations of) transition densities, are all too often not applicable to these problems, due, inter alia, to their high computational costs and prerequisites on the measurements.

The proposed new method is based on the method of *multiple shooting*, using piecewise deterministic solutions of ordinary differential equations (ODEs) to approximate the SDE realization that corresponds to the studied process from which measurements have been taken. The generally discontinuous concatenation of ODE trajectories mimics the consequences of stochastic effects, and, further, allows to formulate the parameter estimation problem as a *deterministic* nonlinear optimization problem that can be solved with efficient *derivative-based* solution methods. In this thesis, a generalized GAUSS-NEWTON method is deployed.

Main results and contributions of this thesis are summarized in the following:

- We propose a new method for parameter estimation in nonlinear multi-dimensional SDEs, based on a piecewise approximation by solutions of ODEs. Discontinuities (jumps) occurring at the interval borders are used for regularization. Unknown parameters and initial states are estimated by a generalized weighted least squares method from data that can originate from direct complete or partial state measurements or from indirectly observed quantities. Measurement data may be afflicted with errors and arbitrarily sampled. Non-linear parameter and point constraints may be formulated as equality and inequality constraints. The resulting nonlinear constrained optimization problems are highly structured and efficiently solved using a generalized GAUSS-NEWTON method.
- We give a proof that the discontinuities at the interval borders asymptotically tend to zero if the number of equidistantly distributed shooting nodes goes to infinity.
- We show in a numerical analysis that the resulting equation systems are sparse, that the number of nonzero elements depends only linearly on the number of shooting nodes, and give sharp upper bounds. Moreover, we prove that the sparsity is maintained if an appropriate (stable) decomposition is applied.
- It is demonstrated in comparative simulation studies that the estimates are robust w.r.t. the exact choice of jump regularization weights. Moreover, the effects of jump regularization on estimates and approximated trajectories are investigated and described.

-
- A *lifting* approach with per-interval parameter sets, coupled by additional equality constraints, is developed and its numerical properties are analysed. Moreover, we propose a homotopy method for the treatment of hard problems.
 - We demonstrate the performance of the new estimation technique in examples from systems biology, each shedding some light on different aspects. Especially, we show that the method can also be used for *hidden state estimation* and for *trajectory reconstruction* in time spans without observations. Further, we derive a criterion for local grid refinement.
 - We show for an ORNSTEIN-UHLENBECK process driven by a LÉVY jump process, that, in addition to mean reversion level and mean reversion rate, also the diffusion constant may be estimated by analysing the jump residuals.
 - The software package `:sfit` is an efficient implementation of the proposed method, offering easy symbolic problem formulation to the user, from which the stochastic parameter estimation problem can be automatically built and solved.

Parts of this work emerged in the interdisciplinary research project *SBCancer* of the *Helmholtz Alliance on Systems Biology*. In close collaboration with expert biologists, we developed an mathematical model for a crosstalk of two cytokines in human skin cells that interfere in a signalling pathway frequently found aberrantly activated in cancer. After an extensive analysis of the deployed measurement data processing, the model proposed by the author of this thesis has been calibrated from experimental data. Its counter-intuitive predictions have been verified in wet lab experiments and lead to new biological insights.

Main novelties and contributions in this thesis are:

- Development of a mathematical model of the crosstalk of two cytokines in human keratinocytes (HaCaT cell line). The predicted and hitherto unknown nonlinear moderating effects of GM-CSF on the IL-6-induced JAK-STAT signalling pathway has been verified *in vitro*.
- An extensive mathematical analysis of the frequently utilized *quantitative WESTERN blotting* measurement procedure shows that established data normalization methods, relying on housekeeping proteins or manually added calibrator proteins, are prone to signal deteriorating statistical artefacts. Moreover, we show that the frequently declared assumption of normally distributed measurement errors cannot be maintained if these normalization techniques are applied.
- We propose as a remedy a normalization technique based on the calculation of amplification factors, and develop criteria for (approximate) normally distributed errors. These criteria can be easily checked using solely the raw measurement data. Moreover, we demonstrate the advantages of the proposed *amplification factors method* in a large comparative simulation study.

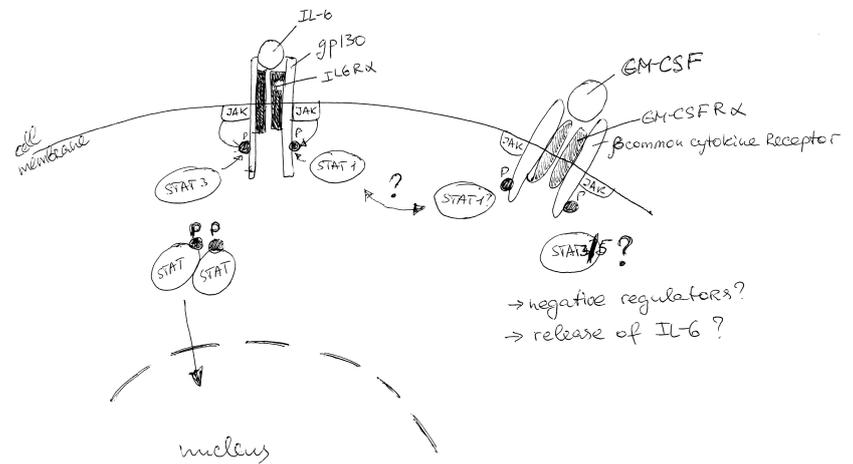
Contents

Preface	xiii
Introduction	xiii
Thesis overview	xvii
Danksagung (acknowledgement)	xxii
1 Gauß-Newton for Constrained Least Squares Problems	1
1.1 Basic overview of constrained nonlinear optimization	2
1.1.1 Optimality conditions	4
1.2 Linear least squares problems	6
1.2.1 Unconstrained linear least squares	6
1.2.2 Linear least squares with equality constraints	9
1.2.3 Linear least squares with inequality constraints	12
1.3 Generalized GAUSS-NEWTON for nonlinear least squares problems	14
1.3.1 The constrained nonlinear least squares problem (NONLIN-LSQ)	14
1.3.2 An iterative solution method based on solving linearized problems	14
1.3.3 The generalized inverse as solution operator	17
1.3.4 Globalization	19
2 Parameter Estimation in Ordinary Differential Equation Models	21
2.1 Preliminaries	21
2.1.1 Correct dynamical model	21
2.1.2 Simulated and real measurements and their errors	22
2.1.3 Maximum likelihood estimation with weighted least squares	24
2.1.4 Structural and practical identifiability of parameters	25
2.2 Boundary value problems	27
2.2.1 Single shooting	27
2.2.2 Collocation	28
2.2.3 Multiple shooting and derivative generation	29
2.3 Constrained nonlinear parameter estimation problem	35
2.3.1 Additional constraints	35
2.3.2 Constrained parameter estimation problem with underlying ODE model	35
2.4 The solver's view	37
2.5 A posteriori statistical analysis of the solution	38
3 Development of and Parameter Estimation in a Crosstalk Model of GM-CSF-mediated IL-6-induced JAK-STAT Signalling – New Biological Insights	41
3.1 Cancer	42
3.2 Biological and experimental background	43
3.2.1 The cytokines IL-6 and GM-CSF	43
3.2.2 The HaCaT-ras A5 cell line, cancer, and their connection to IL-6 and GM-CSF	43
3.2.3 Signalling pathways	44
3.2.4 Proliferation assays reveal a link between IL-6 and GM-CSF in HaCaT	47
3.2.5 Data generation by quantitative immunoblotting (WESTERN blotting)	49

3.3	Experimental settings and raw measurement data	55
3.3.1	Description of experimental procedure	56
3.3.2	Raw measurement data	56
3.4	Theoretical preliminaries for immunoblot data analysis	59
3.4.1	Approximability of a ratio of normal variables by a normal distribution	59
3.5	Data processing	63
3.5.1	Semiquantitative immunoblotting	63
3.5.2	Analysis of WESTERN blotting data	63
3.5.3	Normalization using house-keeping or recombinant proteins (calibrators)	64
3.5.4	Amplification factors method	68
3.5.5	Simulation study: Effect of normalization on signal intensity	73
3.6	Results of the IL-6 stimulation experiments: Processed data	77
3.6.1	Statistical analysis	77
3.6.2	Summary of initial experimental results	79
3.7	Predictive crosstalk model of IL-6 and GM-CSF in JAK-STAT signalling	79
3.7.1	Review of existing models	79
3.7.2	Our IL-6/GM-CSF crosstalk model	81
3.7.3	ODE model formulation	84
3.7.4	Initial cell state	84
3.7.5	Preliminary model calibration	85
3.7.6	A confirmed prediction: SOCS-3 kinetics as model validation	85
3.7.7	Re-calibration using SOCS-3 data	88
3.8	New biological insight: Dose dependency on GM-CSF	92
3.8.1	Predicted GM-CSF dose dependency	92
3.8.2	Wet lab verification of GM-CSF dose dependency	94
3.8.3	Interpretation and outlook	94
4	Introduction to Stochastic Differential Equations	97
4.1	Basic definitions and results	97
4.1.1	Random variables	97
4.1.2	Stochastic processes	103
4.2	The WIENER process and BROWNIAN motion	105
4.2.1	BROWNIAN motion or WIENER process? – A short historical note	106
4.3	Stochastic integrals and stochastic differential equations	107
4.3.1	From ordinary to stochastic differential equations	108
4.3.2	Strong and weak solutions	110
4.3.3	An existence and uniqueness result	110
4.3.4	The bridge between ODE and SDE: random ordinary differential equations (RODE)	111
4.4	Numerical integration of WIENER-driven SDEs	112
4.4.1	EULER-MARUYAMA stochastic integrator	113
4.4.2	An illustrating example of EULER-MARUYAMA integration	115
4.4.3	Further integration methods	116
4.5	LÉVY-driven SDE	118
4.5.1	A short introduction to LÉVY processes	118
4.5.2	Numerical simulation of LÉVY-driven SDE	119

5	Parameter Estimation in Stochastic Differential Equation Models	121
5.1	Introduction	121
5.1.1	The idea: ODE solutions resemble SDE solution paths on short time scales	123
5.1.2	SDE with constant diffusion by LAMPERTI transform	124
5.1.3	Restricting w.l.o.g. to time-homogeneous SDE	125
5.2	Techniques of parameter estimation in SDE	125
5.2.1	Maximum likelihood, transition densities and the FOKKER-PLANCK equation	126
5.2.2	Other estimators	128
5.2.3	Assessing estimator performance	128
5.3	On the distance between SDE and ODE solutions	129
5.3.1	Some results on WIENER processes	130
5.3.2	Distance of S-IVP and D-IVP solutions	132
5.4	A piecewise deterministic parameter estimation method	139
5.4.1	Notation and assumptions	139
5.4.2	A regularization of the multi-experiment case	142
5.4.3	Stochastic parameter estimation problem with jump regularization . . .	143
5.4.4	Imperative jump regularization: A WIENER exponential example	145
5.4.5	Choosing the jump regularization weights	147
5.4.6	A convergence result: Jumps approaching zero	148
5.5	Numerical analysis of the proposed parameter estimation method	151
5.5.1	Sparsity patterns in the linear subproblems	151
5.5.2	Derivative generation	153
5.5.3	Linear complexity growth: An upper bound on the number of nonzero elements in the combined residual vector's Jacobian $J(s, p)$	154
5.5.4	Maintaining sparsity in QR decomposition of the Jacobian $J(s, p)$. . .	155
5.5.5	Lifting with local parameters	159
5.6	Extensions	161
5.6.1	A homotopy in jump regularization weighting	161
5.6.2	Shooting grid refinement	161
6	Numerical Examples	163
6.1	The FITZHUGH-NAGUMO Oscillator	164
6.1.1	Meaning of parameters	166
6.1.2	Expected bias in parameters b and c	166
6.1.3	SDE formulation	166
6.1.4	Simulation study and a closer look at jump regularization effects	167
6.2	Calcium ion oscillations in eukariotic cells	174
6.2.1	SDE approximation and simulation	175
6.2.2	(Un-)Identifiability of parameters in partial observations	177
6.2.3	Simulation study and a closer look at reconstructing intermittent observations	178
6.3	The BISTABAER enzyme kinetics model	187
6.3.1	Model equations and parameters	187
6.3.2	SDE model	188
6.3.3	Simulation study and a closer look at grid refinement	188
6.4	The ORNSTEIN-UHLENBECK model (driven by a LÉVY process)	195
6.4.1	A simple heuristic for placing the shooting nodes	195
6.4.2	No jump regularization	197

6.4.3	Assessing the diffusion parameter by jump residuals	197
6.4.4	An example test case	198
6.4.5	Simulation study	198
6.4.6	Off-label applicability in jump processes	201
Appendix		203
A Appendix to the Crosstalk Model of GM-CSF-mediated IL-6-induced JAK-STAT Signalling		205
A.1	Biological obligation: materials and methods	205
A.1.1	Cell culture	205
A.1.2	Preparation of conditioned media and ELISA	205
A.1.3	Quantitative immunoblotting	205
A.1.4	Estimating absolute concentrations from immunoblot data	206
A.2	Tabulated data	207
B Some Results on Matrix Decompositions and Stochastic Convergence		209
B.1	QR decomposition of full rank matrices	209
B.2	The singular value decomposition (SVD)	210
B.3	Convergence of random sequences	211
C The Stochastic Integral		213
C.1	The idiosyncrasy of stochastic integrals	213
C.2	ITÔ's and STRATONOVICH's integral	214
C.2.1	Which integral to choose: STRATONOVICH or ITÔ?	215
C.3	Filtration, adaptation, and the "history of BROWNIAN motion"	215
C.4	The ITÔ integral	216
C.4.1	The ITÔ integral in 1D	217
C.4.2	Properties of the ITÔ-Integral	219
C.4.3	Extension of the ITÔ integral to n dimensions	222
S Software Package :sfit		223
S.1	GAUSS-NEWTON solver :gnsolve	223
S.2	Integrator :clradau	224
S.3	Framework :sfit	225
S.3.1	Formulating an sfit-problem	225
S.3.2	A complete example: WIENER exponential	233
List of Tables		237
List of Figures		238
Picture Credits		240
Bibliography		241
Index		251



Preface

Introduction

Modelling is a crucial term in the interdisciplinary research project SBCancer, embedded in the Helmholtz Alliance on Systems Biology, that focusses on signalling in cancer and related diseases. In its description, we read: “To promote the modelling activities [...] each of the biomedical projects is closely linked to theoretical projects.” For a fruitful collaboration across disciplines, it is essential for both sides to learn the language of the other, as identical terms all too often have different meanings. That this could lead to misunderstandings may be illustrated by a brief anecdote from my personal experience.

At the project kick-off meeting with experimental biologists, we received a thorough introduction to the biological background and to the measurement data available up to then. They told us that they have a *model* at hand, but seemed unwilling to share it. Instead, we were given a graphical illustration of the JAK-STAT signalling pathway, as it would be more important to us to understand these principles (a scan from the original piece of paper is printed above). We asked about the descriptive abilities of their *model* and were told that it mimics growth, migration and invasion of (cancerous) epithelial cells into subjacent tissue, and that it does this very well as comparisons to *in vivo* experiments with mice have shown – sounds like very sophisticated work! A second time, we made a cautious request for that well developed *model*, which again, however, just raised some eyebrows, as it would be “useless” to us. A new project’s kick-off meeting is probably the worst place for offending sensibilities, so we accepted without a further word of protest.

It finally turned out that the *model* they were speaking about was an *organotypic coculture model*: epithelial cells in a layer on top of a collagen gel populated with fibroblasts and without direct contact to the nutrient medium, imitating physiological characteristics of skin and its stroma (see the illustration in figure 3.2 on page 44). Indeed, a model with limited usage for mathematicians.

This work is about *mathematical models*, and comprises especially the modelling of a crosstalk in the intracellular JAK-STAT signal transduction pathway, triggered by external stimulation with cytokines. We develop an ordinary differential equations model, thoroughly analyse the measurement procedure and the data processing, and estimate unknown kinetic parameters, leading to a calibrated model with experimentally verified predictive power. In view of further applications, we develop a new method for parameter estimation in a class of stochastic differential equations models, analyse the numerics behind, and demonstrate its successful application on several example settings in systems biology and beyond.

Parameter estimation in differential equation models

Mathematical models form the basis of any computer-based simulation. In general, they consist of two essential parts: the *structure* describing the players and interactions in the studied system, and *parameters* that quantitatively characterize them.

Frequently, the structure is sufficiently known, e.g. reaction partners and kinetics in a biochemical process, parameters like rate constants and some or all initial system states, think of the concentrations of certain proteins in living cells, however, have to be determined from experimental data that is, except for a few cases, afflicted with errors.

The task of retrieving unknown quantities from observations, which is the central part of this thesis, is known as *parameter estimation*, *model fitting* or *model calibration*. The method of (*weighted*) *least squares* that determines unknown parameters p by minimizing the squared distance of model simulations $h_i(p)$ and M observations η_i is a well established technique for parameter estimation. If information about the variances σ_i^2 of measurement errors are available, the reciprocals of these variances are used for weighting the squared deviation between model and measurement. Using scalar quantities for ease of notation, the least squares problem

$$p^* = \arg \min_p \sum_{i=1}^M \sigma_i^{-2} \cdot (h_i(p) - \eta_i)^2 \quad (1)$$

delivers the estimate p^* for the vector of unknown parameters. It is known that the method of (weighted) least squares delivers maximum likelihood estimates if a *correct* model is used, i.e. a model that accurately describes the studied system, and if the measurement errors are independent and normally distributed.

Most dynamical processes – may it be in biology, chemistry, physics, medicine, and also in disciplines like finance, psychology, or social studies – can be mathematically modelled by differential equations that describe their temporal development.

From the large mathematical toolbox labelled with “differential equations”, the modeller has to select that tool that fits those properties of the studied system that shall be reproduced. *Ordinary differential equations* (ODEs), are suitable, if observations taken from the studied system are clearly reproducible, i.e., if the system behaves *deterministically*. Systems with intrinsic stochasticity that lead to different outcomes in each experiment repetition, are often well described by *stochastic differential equations* (SDEs).

There is a plethora of further types and subtypes of differential equations, e.g. *partial differential equations* (PDEs) are the method of choice if spatial distribution is important, *delay differential equations*, ordinary, partial, or stochastic ones, may be used to model systems where the current state depends on its history. In this thesis, we focus on systems that are described by ODEs and SDEs, primarily originating from biological settings.

The method of multiple shooting

A method for parameter estimation in ODE that has been successfully applied in many settings is BOCK’s method of *direct multiple shooting* that parametrizes the state trajectory as follows: A grid of *shooting nodes* is introduced, dividing the time horizon into separate intervals. To every node, a *shooting variable* is associated that acts as initial value for the ODE on the respective interval. Continuity of the final trajectory is guaranteed by imposing *matching conditions* that force the state values at the end of each interval to coincide with the initial values (i.e. the shooting variable) of the following one.

Main benefits of this methodology are stability and efficiency: Error propagation is interrupted at every shooting node, circumventing rampant growth of parasitic solution components, and measurement data can be used for initializing the shooting variables, accelerating the solution finding process. Furthermore, the resulting systems are highly structured and the respective initial value problems on each interval can be solved in parallel.

Development of a predictive model for a signalling crosstalk in human keratinocytes and improvements in measurement processing methods

The “classical” task of parameter estimation in *ordinary* differential equation models is still of high importance and widely used. We apply these methods in a signal transduction model that we newly developed within an interdisciplinary research project in the scope of SBCancer.

In close collaboration with expert biologists from the group “Tumour and Microenvironment” headed by Margareta M. MÜLLER at the German Cancer Research Center (DKFZ) in Heidelberg, a *crosstalk* of the two cytokines IL-6 (interleukin 6) and GM-CSF (granulocyte-macrophage colony-stimulating factor) in human skin cells has been investigated. The involved JAK-STAT signalling pathway is frequently found aberrantly activated in various cancers.

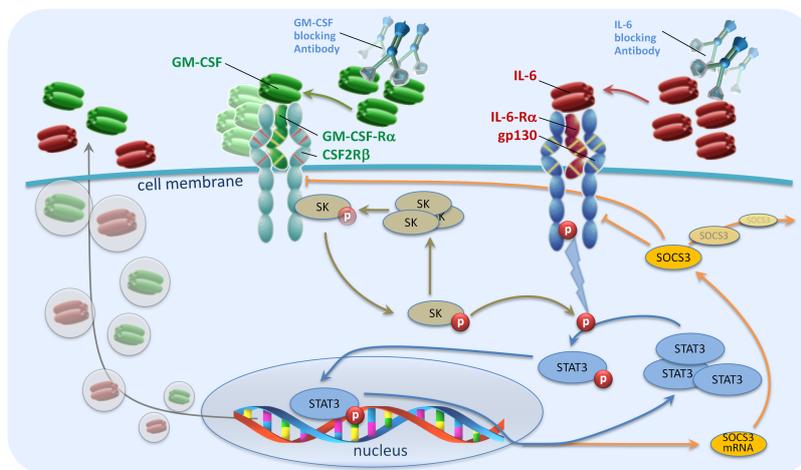


Illustration of the new *crosstalk* model

Since this crosstalk has never been investigated before – see the discussion of existing models of the JAK-STAT signalling pathways in section 3.7 – a completely new model was created from scratch by the author of this thesis, with focus on comprehensibility and interpretability.

The model was fitted to wet lab measurements originating from cell stimulation experiments with

different stimulation settings, quantified by semi-quantitative WESTERN blotting. During analysis of this data, we detected “spurious” signals that we were able to track down to the frequently applied normalization technique based on manually added *calibrator proteins*. Processing the data with the *amplification factors method* proposed by the author of this thesis eliminated these artefacts.

We give an extensive mathematical analysis of both processing methods that elucidates the made observations. We show why and to which extent the established data normalization method relying on housekeeping proteins or manually added calibrator proteins is prone to signal deteriorating statistical artefacts, and we further demonstrate that the frequently declared assumption of normally distributed measurement errors cannot be maintained in general if this normalization technique is applied. The *amplification factors method* delivers stable and reproducible results under only weak requirements, which are often fulfilled and – important for wet lab experiments – easy to verify. Furthermore, it allows the comparison of data originating from different blots and the errors can often be well approximated by a normal distribution.

We shortly mention at this place that the newly developed crosstalk model has proven its predictive capabilities in several verification experiments, and that a new co-stimulation series proposed by the author has led to new biological insights. The model predicted a hitherto unknown nonlinear dose-dependency of the IL-6 induced JAK-STAT signal on GM-CSF that has been confirmed both qualitatively and quantitatively in subsequently performed wet lab experiments. For details, we refer to chapter 3.

Extension to systems with intrinsic noise: Parameter estimation in SDEs

Most processes – not limited to biological ones – follow certain basic principles but may be influenced by intrinsic stochasticity that may significantly alter the system’s behaviour quantitatively and also qualitatively. In this case, a description with SDE is often necessary and appropriate. Informally, we may interpret a certain class of SDEs, the ones with *constant diffusion*, as ODEs “with noise”, if we compare the notation of ODE and those SDE in differential form:

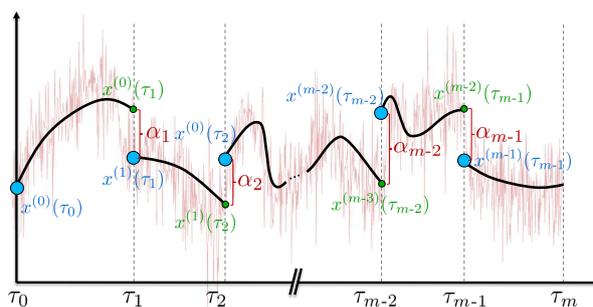
$$(ODE) \quad dx = f(t, x, p) dt \qquad (SDE) \quad dx = f(t, x, p) dt + D dW_t$$

The *drift* f depends on current time t , current state $x(t)$, and parameters p . The SDE formula on the right has an additional *diffusion* term $D dW_t$ that denotes “white noise”, the formal derivative of a WIENER process scaled with intensity D .

On small time scales, however, many of these (and other) SDE processes *resemble* ODE processes, i.e. the stochasticity is not the dominating part of the system. See, i.e., the visualizations of a bistable allosteric enzyme regulation model in figure 5.2 on page 123.

From this observation, there arises the idea to use ODEs to approximate an SDE realization. The time horizon is split into multiple intervals, in which only the drift part of the SDE, i.e. the associated ODE, is used for simulating the process, and the SDE realization is approximated by concatenating these “deterministic pieces” that describe the system for small periods of time.

This resembles the direct multiple shooting for parameter estimation in ODE, but in contrast to that method, we now do not impose matching conditions at the shooting nodes but allow discontinuities. These *jumps* may be interpreted as cumulated stochasticity of the preceding interval. It turns out that simply concatenating interval-wise ODE solutions may lead to wrong conclusions, as we show in section 5.4.4 on page 145. We can avoid that by using the jumps for regularization.



The new *piecewise deterministic approach*: Approximating an SDE trajectory with ODE solutions. The discontinuities α_i are used for regularization.

Assume the shooting grid is given by $\tau_0 < \tau_1 < \dots < \tau_m$, and let us denote the ODE solution in interval $[\tau_k, \tau_{k+1}]$ by $x^{(k)}(t) \in \mathbb{R}^{n_x}$, then instead of minimizing the least squares formulation in eq. (1), we employ a *regularized objective*:

$$p^* = \arg \min_p \sum_{i=1}^M \sigma_i^{-2} \cdot (h_i(p) - \eta_i)^2 + \sum_{k=1}^{m-1} \sum_{l=1}^{n_x} \omega_{k,l}^2 \cdot \left(x_l^{(k-1)}(\tau_k) - x_l^{(k)}(\tau_k) \right)^2 \quad (2)$$

with appropriately chosen jump regularization weights $\omega_{k,l}^2$. The simulated measurements $h_i(p)$, the *model response*, are now taken from the interval-wise trajectory approximations. A rigorous formulation is given in the definition of the constrained nonlinear parameter estimation problem 5.18 on page 144.

This *piecewise deterministic approach* allows the usage of derivative based optimization methods while maintaining the ability to reflect stochasticity intrinsic to the SDE formulation. If the parameter estimation problem (2), possibly with additional equality and inequality

constraints, is solved with a GAUSS-NEWTON method, the occurring equation systems are highly structured. This can be exploited for computational efficiency.

This new approach for parameter estimation in SDE models is subject of the fifth chapter, where we start with a detailed introduction to the idea, and also give a literature review and discussion of existing techniques for parameter estimation in SDE. Furthermore, we successfully apply this method to examples from systems biology and mathematical finance, and demonstrate that it delivers reliable and robust estimates.

An overview is also given in the subsequent *Thesis Overview* section.

Thesis overview

At this place, we present an overview about the contents of this thesis. Further details and relevant references are given in the summaries at the beginning of each chapter.

A few words on the sectioning of this thesis: The first chapter introduces and discusses a generalized GAUSS-NEWTON method for solving nonlinear least squares problems that arise in the parameter estimation problems covered in this thesis. The second chapter reviews the method of multiple shooting for solving parameter estimation problems in ODE models to have them at hand when using them in chapter three for calibrating the cytokine crosstalk model. Further, these techniques form the basis of the new parameter estimation method for SDEs developed in the fifth chapter, after introducing the theoretical foundations of SDEs and their numerical simulation in chapter four. Finally, in the sixth chapter, we demonstrate the performance of the new method in several numerical examples, each elucidating different aspects.

Chapter 1: Gauß-Newton for Constrained Least Squares Problems

The first chapter focuses on the GAUSS-NEWTON method for nonlinear least squares problems with equality and inequality constraints, as these form the basis for the new parameter estimation method for stochastic systems in chapter 5.

We recapitulate the basic theory of constrained nonlinear optimization and formulate optimality conditions in the sense of KARUSH-KUHN-TUCKER tuples. Some solution methods for (un-)constrained linear least squares are presented, also for the rank deficient case.

The generalized GAUSS-NEWTON method, based on iteratively solving linearized problems, is derived, and we show that the solution operator is a generalized inverse constructed from objective and constraints gradients. A local convergence result of BOCK is given. The chapter finishes with the presentation of two globalization methods.

Chapter 2: Parameter Estimation in Ordinary Differential Equation Models

In the second chapter, we discuss parameter estimation in ordinary differential equations, establish the notation used in subsequent chapters, and show that the method of weighted least squares delivers a maximum likelihood estimate if measurements are affected by additive and normally distributed errors. Further, we introduce the concepts of structural and practical identifiability, addressing the question whether the available measurement data contains enough information to reconstruct the unknown parameters.

Boundary value problems (BVPs) are the basis of the parameter estimation methods discussed in this thesis. We present three widespread solution methods for BVPs, single shooting, collocation, and multiple shooting, and discuss the pros and cons of these techniques.

We detail the method of multiple shooting for solving BVPs and present and discuss common techniques for generating derivatives of ODE solutions with respect to initial values needed to apply the GAUSS-NEWTON solution method described in chapter 1: *external numerical differentiation* (END, method of perturbed trajectories, finite differences, including the 50 years old but still not widely known *i-trick* of LYNESS and MOLER), *internal numerical differentiation* (IND, a principle avoiding inconsistencies by freezing adaptive components of the integrator), the *variational differential equations*, and *automatic differentiation* (AD, algorithmic differentiation), the latter implemented and used in combination with IND in the software package `:sfit` developed in this thesis.

Subsequently, we formulate the nonlinear parameter estimation problem with an underlying ODE model as an overdetermined BVP. The dynamics are treated using a multiple shooting parametrization, continuity conditions ensure a continuous solution trajectory. We further address the questions how the objective must be adjusted for correlated measurements, and how errors in the independent variables (usually time) can be treated by *total least squares*.

The chapter closes with a statistical analysis of the generated solutions, giving (linearized) confidence intervals for the estimated quantities.

Chapter 3: Development of and Parameter Estimation in a Crosstalk Model of GM-CSF-mediated IL-6-induced JAK-STAT Signalling

The third chapter contains an application in biology of the parameter estimation methods developed in chapter 2. We develop a mathematical model for an intracellular crosstalk in the so-called JAK-STAT (Janus kinase / signal transducer and activator of transcription) pathway in human skin cells, which is frequently found aberrantly activated in various cancers.

We first give an extensive introduction into the biological backgrounds of cytokines IL-6 (interleukin 6) and GM-CSF (granulocyte-macrophage colony stimulating factor), the deployed HaCaT cell line, to the JAK-STAT signalling pathway, and their connection to cancer.

Further, we describe in detail the wet lab experiments to work out possible sources of error, and investigate the processing method for data generated by the semi-quantitative WESTERN blotting procedure. Presenting and using earlier work of MARSAGLIA on the ratio of normal variables, we show that the commonly applied data normalization technique using calibrator or house-keeping proteins leads to non-normally distributed data, and is very likely to introduce severe artefacts or even totally destroy the signal of interest due to normalization.

We instead propose a different data processing method, the *amplification factors method*, and show that it delivers approximately normally distributed data suitable for maximum likelihood parameter estimation using the method of weighted least squares. As an extra benefit, the amplification factors method allows easy comparison of data generated in different experiments, which is not the case for the other technique. Moreover, we show by comparing the two data processing methods in an extensive simulation study that the amplification factors method gives better results and is much more robust towards measurement errors.

A review of existing models of IL-6 induced JAK-STAT signalling shows that many are built in great detail and heavily overparametrized for the available data. Furthermore, none of them incorporates the desired crosstalk with GM-CSF. We therefore develop an ODE-based crosstalk model from scratch, with a focus on comprehensibility and interpretability.

Using the wet lab data and the methods described in chapter 2, we estimate the unknown kinetic parameters of the model, and use the fitted model to make predictions for the kinetics of suppressors of cytokine signaling (SOCS). These were later verified in wet lab experiments.

Finally, we use the additional data of the model verification experiments to improve the parameter estimates, and use the model to predict the cell's behaviour under different co-stimulation settings that were proposed by the author. The predictions show a counter-intuitive nonlinear GM-CSF-dependent modulation of the IL-6 induced signal. The subsequently performed wet lab experiments confirmed these predictions both qualitatively and quantitatively, leading to new biological insights.

Chapter 4: Introduction to Stochastic Differential Equations

In chapter four, we first recapitulate the basic concepts of probability theory needed to introduce stochastic processes, focussing on the class of càdlàg processes, and give a definition and some properties of WIENER processes.

Moreover, we motivate and formally derive WIENER-driven stochastic differential equations, define strong and weak solutions, and give existence and uniqueness results for stochastic initial value problems (S-IVPs), based on the ITÔ theory of stochastic integrals given in appendix C. Furthermore, we discuss a third class of differential equations lying “in between” ODE and SDE: random ordinary differential equations (RODEs).

We recall some numerical integration schemes for approximating strong solutions of S-IVPs, especially the explicit and semi-implicit EULER-MARUYAMA and MILSTEIN schemes, which are used for data generation in the numerical examples chapter 6, and give their respective orders of weak and strong convergence. Also, some higher order integration schemes are discussed.

We conclude this chapter with a formal definition of and an introduction to LÉVY processes, i.e. possibly discontinuous càdlàg processes, and present numerical integration schemes, which are used in the numerical examples chapter for data generation.

Chapter 5: Parameter Estimation in Stochastic Differential Equation Models

In chapter five, we present a new method for parameter estimation in stochastic differential equations, based on a piecewise deterministic approach.

We start with an example of allosteric enzyme regulation, showing that even small stochasticity may completely alter the behaviour of a system both quantitatively and even qualitatively. Based on the observation that, for short time scales, ODE and SDE solutions do resemble each other, we outline the idea of the new parameter estimation method: Splitting the whole time horizon into multiple intervals as in the multiple shooting method for parameter estimation in ODEs presented in chapter 2, but allowing jumps at the shooting nodes and using relaxed continuity conditions for regularization.

Since this methodology only uses the drift part of the underlying SDE, we present the LAMPERTI transform that can be used to transform an SDE with state-dependent diffusion into one with constant diffusion.

We discuss existing methods for parameter estimation in SDEs, especially techniques for (approximate) maximum likelihood estimation based on transition probabilities, which rely on excessive sampling and are in general computationally very costly. We further discuss the STRATONOVICH-KÁLMÁN-BUCY filtering technique for linear problems, as well as other estimation methods, and give a note on assessing estimator performance.

In preparation to a later convergence result, we investigate the distance between solution of stochastic and deterministic initial value problems at interval boundaries and the maximum deviation over the full interval, and give bounds in expectation and mean-square for both deviations.

After these preparations, we establish the new method for parameter estimation in SDEs. We introduce the required notation, give a detailed problem description, and describe how to apply the GAUSS-NEWTON method described in chapter 1 for solving. An example on estimating initial value and exponent of a WIENER-driven exponential illustrates the method and shows that the jump regularization is indispensable for obtaining correct estimates. We further establish a connection to the methods for parameter estimation in ODE in the so-called *multi-experiment* setting with additional jump regularization, and prove that the piecewise ODE solutions converge to the (realization of the) SDE if the sizes of the shooting intervals decline to zero.

Subsequently, we perform a numerical analysis of the new method. We investigate the sparsity pattern in the linear subproblems that occur in the GAUSS-NEWTON method, and give a sharp upper bound on the number of nonzero elements in the residuals' Jacobian. We further show that the number of nonzero elements in the Jacobian grows only linearly in the number of shooting nodes, while the total number of elements grows quadratically. Moreover, we prove that sparsity in the Jacobian can be maintained if a suitable (stable) decomposition is applied, and give a sharp upper bound for the number of nonzero elements in the resulting decomposition factor.

We proceed with proposing a lifting method based on per-interval parameters, coupled by additional equality constraints, and redo the numerical analysis for the lifted problem. We conclude this chapter with a proposition for a homotopy in jump regularization that may help to treat ill-posed problems.

Strategies for choosing and adjusting the shooting grid are discussed by means of the BISTABAER example in section 6.3.3.7 on page 191. The software package `:sfit`, implementing the ideas and methods of this thesis, is described in appendix S.

Chapter 6: Numerical Examples

The final chapter presents successful applications of the new parameter estimation method in four distinct examples, elucidating different aspects. The respective sections comprise a detailed description of the studied system, a visual comparison between ODE and SDE interpretations, extensive simulation studies to assess the estimator performance, and the following investigations:

In each simulation study, 100 realizations of the SDEs are simulated with different driving WIENER processes (or LÉVY processes in the ORNSTEIN-UHLENBECK example). Model parameters are estimated using data retrieved in usually eight different scenarios constructed from combinations of full and partial observations, exact and noisy measurements, and varying jump regularization weights. All results are concisely tabulated and discussed.

The quality of the estimation method is described by the mean values of the estimates, which are found to be close to the true values in all examples, and their spread in terms of standard deviations, which are usually small, giving evidence for a reliable and robust estimation technique.

Also, common to all four examples, we describe the measurement functions, shooting grid and node initialization, as well as point and parameter constraints (if any). Further, we give information about the sparsity of the residuals' Jacobian and graphically display its sparsity structure. Moreover, we investigate certain individual properties in the respective “*a closer look*” sections.

In the FITZHUGH-NAGUMO oscillator example, prototypically describing an excitable system with intrinsic noise, we analyse and illustrate in addition to the investigations stated above the effects of different jump penalization weights on the recovered trajectories.

In the example of **calcium ion oscillations in eukariotic cells**, we first prove that the system is unidentifiable when using partial observations, and therefore select an identifiable subset of parameters for these settings in the simulation study. Further and in addition to the common tasks, we investigate the performance of our new estimation technique when having only intermittent observation, i.e. no measurement data is available for prolonged periods of time, and find that the trajectories in the unobserved areas are recovered surprisingly well (see figure 6.10e on page 186), showing that our proposed estimation technique may also be used for trajectory reconstruction.

In the **allosteric enzyme regulation model BISTABAER**, we show that the method – besides estimating the unknown parameter values – successfully reconstructs the hidden intrinsic control of the system (a bistable oscillator affected by the driving WIENER process). Moreover, we give a strategy for local refinement of the shooting grid based on local residuum analysis.

In the final example of an **ORNSTEIN-UHLENBECK mean reversion process**, driven by a discontinuous LÉVY process, we present a heuristic method for automatically placing shooting nodes using solely the available measurement data without further information. In contrast to the previous example, the jump weights are chosen as zero (as the actual trajectory is discontinuous), and we show that also the diffusion parameter can be estimated from the distribution of the jumps.

Appendix

The **appendix A** collects supplemental information about the IL-6/GM-CSF crosstalk model developed in chapter 3. It presents the *materials and methods* used in the wet lab, explains the calculation of absolute concentrations from WESTERN blot data, and gives a tabulated overview of measurement data and parameter estimates.

In **appendix B**, we give some results on QR and SVD-based matrix decompositions, and list different types of convergence of random sequences.

In **appendix C**, we derive the stochastic integrals of ITÔ and STRATONOVICH, give some basic properties and results, and introduce some further notions that we use in chapters 4 and 5.

Finally, in **appendix S**, we present the software package `:sfit` that implements the ideas and methods for parameter estimation in stochastic systems described in chapter 5, and give step-by-step instructions to set up and solve parameter estimation problems.

Some historical notes

Throughout this thesis, we give some historical notes on selected topics. We like to explicitly mention the following:

- (i) A change in notation: From KUHN-TUCKER to KARUSH-KUHN-TUCKER points, see footnote (1) on page 4
- (ii) On the invention and first FORTRAN implementation of automatic differentiation (AD), see footnotes (3) and (4) on page 34
- (iii) The “evolution” of WIENER processes over centuries from first documented observations by INGEN-HOUSZ in 1784 and the rigorous existence proof by Norbert WIENER in 1923, see section 4.2.1 and footnote (5) on page 106, and
- (iv) The STRATONOVICH-KÁLMÁN-BUCY filter and an anticipating work of THIELE from 1880, see footnote (2) on page 127

Sources for further readings are given at the mentioned places.

Danksagung

Mein erster und besonderer Dank gilt Herrn Professor Hans Georg Bock und Herrn Dr. Johannes P. Schlöder, die in der Arbeitsgruppe Simulation und Optimierung am Interdisziplinären Zentrum für Wissenschaftliches Rechnen (IWR) der Universität Heidelberg die denkbar besten Bedingungen bereitet und die denkbar beste Umgebung geschaffen haben – fachlich, wie auch menschlich.

Bedanken möchte ich mich auch bei den weiteren Mitgliedern der Arbeitsgruppe, insbesondere bei Herrn Dr. Mario S. Mommer für das Mentoring im Rahmen der Graduiertenschule und seine freundschaftliche Unterstützung auch im außeruniversitären Umfeld. Nochmals danke ich an dieser Stelle Herrn Dr. Johannes P. Schlöder, der mir über die Jahre kritischer und wertvoller Ratgeber war. Meinen Kollegen Dr. Kathrin Hatz und Dr. Simon M. Lenz ist mein herzlicher Dank für die gemeinsamen und unzähligen wissenschaftlichen Diskussionen und freundschaftlichen Gespräche gewiss.

Herrn Professor Volker Schulz, der durch seine Vorlesungen und Seminare an der Universität Trier mein Interesse am wissenschaftlichen Rechnen geweckt hat, danke ich für die Förderung während meines Studiums und darüber hinaus.

Weiter bedanke ich mich herzlich bei Frau Professorin Margareta M. Müller; zum einen für die erfolgreiche Zusammenarbeit im Projekt SBCancer, deren Ergebnisse im dritten Kapitel dieser Arbeit beschrieben sind, und zum anderen für die daraus entstandene wiederholte Übertragung der Leitung von Praktika und Vorlesungen an der Hochschule Furtwangen, die mir eigene praktische Erfahrungen in einem biotechnologischen Labor und in der Lehre ermöglicht hat.

Für die großartige gemeinsame Zeit während der Gründung und im Vorstand des *Heidelberg Chapter of SIAM* bedanke ich mich bei Dres. Dörte Beigel, Holger Diedam, Kathrin Hatz und Dennis Janka.

Weiter danke ich der Fakultät für Mathematik und Informatik der Universität Heidelberg sowie der Graduiertenschule HGS MathComp des IWR und dem Forschungsnetzwerk SBCancer der Helmholtz-Allianz für Systembiologie für die wissenschaftliche und finanzielle Förderung meiner Arbeit.

Mein größter Dank indes gilt meinen Eltern Mechthilde und Egon Sommer, deren fortwährende Unterstützung auch in schwierigen Projektphasen die inneren und äußeren Bedingungen für die Erstellung dieser Arbeit gesichert hat.

1 Gauß-Newton for Constrained Least Squares Problems

This chapter contains the mathematical background for numerically solving the problems stated in this thesis.

The first section gives a general introduction to nonlinear optimization. After the mathematical preliminaries are specified, we formulate standard optimality conditions in the sense of KARUSH-KUHN-TUCKER points.

Section two discusses linear least squares problems and their (unique) solvability criteria. Solution methods for unconstrained, equality-constrained and inequality-constrained problems (based on active-set strategies) are given, also for the rank-deficient cases.

The last section introduces nonlinear least squares problems and discusses the generalized GAUSS-NEWTON algorithm for solving them by iteratively improving an initial guess with (damped) increments generated as the solution of linearized least squares problems. For these linear least squares problems, we show that the solution operator may be written as a generalized inverse constructed from objective and constraint gradients. After a local convergence proof first stated by BOCK, we conclude the chapter with a globalization technique.

The presentation in this chapter mainly follows the descriptions in [NocedalWright2006], [Bjorck1996], and [Bock1987].

On notation

For the sake of completeness and to introduce the notation used in this thesis, we repeat the basic definitions of *nullspace* and *range* of a matrix and for *balls* in \mathbb{R}^n .

1.1 Definition (Range and nullspace)

Let $C \in \mathbb{R}^{m \times n}$. Then, we define the *range* $\mathcal{R}(C) \subseteq \mathbb{R}^m$ and the *nullspace* $\mathcal{N}(C) \subseteq \mathbb{R}^n$ as

$$\begin{aligned}\mathcal{R}(C) &:= \{y \in \mathbb{R}^m \mid \exists x \in \mathbb{R}^n : y = Ax\} \\ \mathcal{N}(C) &:= \{x \in \mathbb{R}^n \mid Cx = 0\}\end{aligned}$$

□

1.2 Definition (Open and closed balls)

For a point $x_0 \in \mathbb{R}^n$ and $\varepsilon \geq 0$, we define

1. the *open ball* around x_0 as $B^o(x_0, \varepsilon) := \{x \in \mathbb{R}^n \mid \|x - x_0\| < \varepsilon\}$
2. the *closed ball* around x_0 as $\bar{B}(x_0, \varepsilon) := \{x \in \mathbb{R}^n \mid \|x - x_0\| \leq \varepsilon\}$
3. and write $B(x_0, \varepsilon)$ for any *ball* fulfilling $B^o(x_0, \varepsilon) \subseteq B(x_0, \varepsilon) \subseteq \bar{B}(x_0, \varepsilon)$.

□

1.1 Basic overview of constrained nonlinear optimization

We start with a general formulation of constrained nonlinear optimization problems (CNLP), introducing basic notions, optimality concepts, and solution methods. The occurring functions shall be sufficiently smooth, say, C^3 . We note that most results of this section also hold for much weaker regularity assumptions.

1.3 Problem (CNLP)

Let $f: \mathbb{R}^{n_x} \rightarrow \mathbb{R}$ and $r_i: \mathbb{R}^{n_x} \rightarrow \mathbb{R}$ ($i = 1, \dots, n_{ec} + n_{ic}$),

where n_{ec} and n_{ic} denote the number of equality and inequality constraints.

Then, the *constrained nonlinear problem* is written as

$$\begin{aligned} \min_{x \in \mathbb{R}^{n_x}} \quad & f(x) \\ \text{s.t.} \quad & r_i(x) = 0 && (i = 1, \dots, n_{ec}) \\ & r_i(x) \geq 0 && (i = n_{ec} + 1, \dots, n_{ec} + n_{ic}) \end{aligned}$$

For ease of notation, we unify the equality constraints in the function r^{ec} , and the inequality constraints in the function r^{ic} :

$$\begin{aligned} r^{ec} : \mathbb{R}^{n_x} &\rightarrow \mathbb{R}^{n_{ec}} && \text{with components } r_i^{ec}(x) := r_i(x) && (i = 1, \dots, n_{ec}) \\ r^{ic} : \mathbb{R}^{n_x} &\rightarrow \mathbb{R}^{n_{ic}} && \text{with components } r_i^{ic}(x) := r_{n_{ec}+i}(x) && (i = 1, \dots, n_{ic}) \end{aligned}$$

1.4 Remark

Depending on the situation, the notations of the constraints as scalar functions r_i or as split vector valued functions r^{ec} and r^{ic} have certain advantages and drawbacks. We thus use the respective notation as appropriate. \square

1.5 Definition (Feasible point, feasible set, local and global solution)

For the constrained nonlinear problem (CNLP), let $x \in \mathbb{R}^{n_x}$.

1. We call x a *feasible point*, if x fulfills the constraints, i.e. $r^{ec}(x) = 0$ and $r^{ic}(x) \geq 0$.
2. The set $\mathcal{X} := \{x \in \mathbb{R}^{n_x} \mid r^{ec}(x) = 0 \wedge r^{ic}(x) \geq 0\}$ is called the *feasible set* for (CNLP).
3. A feasible point $x^* \in \mathcal{X}$ is called a *local solution* or *local minimizer*, if there exists an $\varepsilon > 0$ such that $f(x^*) \leq f(x)$ for all $x \in \mathcal{X} \cap B(x^*, \varepsilon)$.
If the inequality holds strictly, x^* is called a *strict local solution* or *strict local minimizer*.
4. A feasible point $x^* \in \mathcal{X}$ with $f(x^*) \leq f(x)$ for all $x \in \mathcal{X}$ is called a *global solution* or *global minimizer*. If the inequality holds strictly, x^* is called a *strict global solution* or *strict global minimizer*.

\square

1.6 Definition (Lagrangian multipliers)

The *Lagrangian* for the constrained minimization problem (CNLP) is defined as

$$\mathcal{L}(x, \lambda) := f(x) - \sum_{i=1}^{n_{ec}+n_{ic}} \lambda_i r_i(x)$$

or, equivalently,

$$\mathcal{L}(x, \lambda^{ec}, \lambda^{ic}) := f(x) - \sum_{i=1}^{n_{ec}} \lambda_i^{ec} r_i^{ec}(x) - \sum_{i=1}^{n_{ic}} \lambda_i^{ic} r_i^{ic}(x),$$

where the vector of *Lagrangian multipliers* λ is split into a part $\lambda^{ec} = (\lambda_1, \dots, \lambda_{n_{ec}})^T$ corresponding to the equality constraints and $\lambda^{ic} = (\lambda_{n_{ec}+1}, \dots, \lambda_{n_{ec}+n_{ic}})^T$ corresponding to the inequality constraints. \square

1.7 Definition (Active set)

For every $x \in \mathbb{R}^{n_x}$, each of the inequality constraints in (CNLP) is either active, $r_i^{ic}(x) = 0$, or inactive, $r_i^{ic}(x) > 0$.

1. Given a point $x \in \mathbb{R}^{n_x}$, the *active set*

$$\mathcal{A}(x) := \{i \in \{1, \dots, n_{ec} + n_{ic}\} \mid r_i(x) = 0\}$$

consists of the indices of all active constraints at this point x .

2. For notation, we split the active set into parts belonging to equality and inequality constraints:

$$\begin{aligned} \mathcal{A}^{ec}(x) &:= \{i = 1, \dots, n_{ec} \mid r_i^{ec}(x) = 0\} \\ \mathcal{A}^{ic}(x) &:= \{i = 1, \dots, n_{ic} \mid r_i^{ic}(x) = 0\} \end{aligned}$$

Thus, for a *feasible* point $x \in \mathcal{X}$, we have $\mathcal{A}^{ec}(x) = \{1, \dots, n_{ec}\}$.

\square

Before we formulate optimality conditions of constrained nonlinear problems in section 1.1.1, we introduce two common prerequisites on the regularity of feasible points.

1.8 Definition and Lemma (LICQ, MFCQ)

Let $x \in \mathcal{X}$ be a feasible point, and $\mathcal{A}(x)$ be its active set as in definition 1.7.

1. The point x fulfills the *linear independence constraint qualification* or *LICQ*, if the set of gradients of active constraints

$$\{\nabla r_i(x) \mid i \in \mathcal{A}(x)\} = \{\nabla r_i^{ec}(x) \mid i \in \mathcal{A}^{ec}(x)\} \cup \{\nabla r_i^{ic}(x) \mid i \in \mathcal{A}^{ic}(x)\}$$

is linearly independent.

2. The point x fulfills the *MANGASARIAN-FROMOVITZ constraint qualification* or *MFCQ*, if both following conditions hold:

- a) the set of equality constraint gradients $\{\nabla r_i^{ec}(x) \mid i \in \mathcal{A}^{ec}(x)\}$ is linearly independent
- b) for a certain vector $w \in \mathbb{R}^{n_x}$, it holds

$$\begin{aligned} \nabla r_i^{ec}(x)^T w &= 0 & \text{for } i \in \mathcal{A}^{ec}(x) \\ \nabla r_i^{ic}(x)^T w &> 0 & \text{for } i \in \mathcal{A}^{ic}(x) \end{aligned}$$

3. The following implication holds: LICQ \implies MFCQ.

Proof: If LICQ holds for x , we can choose w as the solution of the system

$$\begin{aligned} \nabla r_i^{ec}(x)^T w &= 0 & \text{for } i \in \mathcal{A}^{ec}(x) \\ \nabla r_i^{ic}(x)^T w &= 1 & \text{for } i \in \mathcal{A}^{ic}(x) \end{aligned}$$

and thus x fulfills the MFCQ. □

1.1.1 Optimality conditions

1.9 Theorem (Characterization of solutions: KARUSH-KUHN-TUCKER conditions)

Let x^* be a solution of (CNLP), fulfilling the LICQ or MFCQ. Then, there exist *Lagrangian multipliers* $\lambda^{ec} \in \mathbb{R}^{n_{ec}}$ and $\lambda^{ic} \in \mathbb{R}^{n_{ic}}$, such that

$$\nabla_x \mathcal{L}(x^*, \lambda^{ec}, \lambda^{ic}) = 0 \tag{1.0a}$$

$$r^{ec}(x^*) = 0 \tag{1.0b}$$

$$r^{ic}(x^*) \geq 0 \tag{1.0c}$$

$$\lambda^{ic} \geq 0 \tag{1.0d}$$

$$\lambda_i^{ic} r_i^{ic}(x^*) = 0 \quad (i = 1, \dots, n_{ic}) \tag{1.0e}$$

For a given solution x^* , the Lagrangian multipliers are unique, if the LICQ holds.

Proof: See [NocedalWright2006] chapter 12.3. □

The conditions in eqs. (1.0a) to (1.0e) are called the *KARUSH-KUHN-TUCKER*⁽¹⁾ (*KKT conditions*). Equations (1.0b) and (1.0c) ensure that x^* is a feasible point of (CNLP) and are thus called *primal feasibility conditions*. Equation (1.0a) is called *stationarity condition*, eq. (1.0d) is called *dual feasibility condition*, and eq. (1.0e) the *complementarity condition*.

⁽¹⁾ The US-American mathematician William KARUSH, 1917–1997, first formulated equivalent conditions in his master’s thesis *Minima of Functions of Several Variables with Inequalities as Side Constraints* in 1939, though his work remained unpublished. For several decades, the optimality conditions from theorem 1.9 were known as the KUHN-TUCKER conditions, after Harold William KUHN, 1925–2014, and Albert William TUCKER, 1905–1995, published their eponymous paper in 1951, unaware and independently of KARUSH’s earlier work.

Richard COTTLE gives elucidating details about *William Karush and the KKT theorem* in M. GRÖTSCHEL (ed.): *Optimization Stories*, Documenta Mathematica, Bielefeld, Germany, 2012, pp. 255–269.

As the KKT conditions contain first order derivatives, and each solution x^* fulfills them, they are also known as *first order necessary conditions*. Further, any point fulfilling the conditions (1.0a)–(1.0e) is frequently called a *KKT point* or *stationary point* of (CNLP), and the tuple $(x^*, \lambda^{ec}, \lambda^{ic})$ is named a *KKT tuple*.

1.10 Definition (Strict complementarity)

Let $(x^*, \lambda^{ec}, \lambda^{ic})$ be a KKT tuple in the sense of theorem 1.9. If for each $i = 1, \dots, n_{ic}$ in the complementarity condition eq. (1.0e) either $\lambda_i^{ic} = 0$ or $r_i^{ic}(x^*) = 0$, but not both at the same time, i.e.

$$\lambda_i^{ic} r_i^{ic}(x^*) = 0 \iff \lambda_i^{ic} = 0 \text{ xor } r_i^{ic}(x^*) = 0,$$

or, more verbose, we have $\lambda_i^{ic} = 0$ if $r_i^{ic}(x^*) > 0$, and $\lambda_i^{ic} > 0$ if $r_i^{ic}(x^*) = 0$, then we say that the *strict complementarity conditions* hold. \square

1.11 Definition (Set $V(x^*)$)

Let $(x^*, \lambda^{ec}, \lambda^{ic})$ be a KKT tuple of (CNLP). Then, we define the set $V(x^*)$ by

$$v \in V(x^*) \iff \begin{cases} \nabla r_i^{ec}(x^*)^T v = 0 & \text{for } i \in \mathcal{A}^{ec}(x^*) \\ \nabla r_i^{ic}(x^*)^T v = 0 & \text{for } i \in \mathcal{A}^{ic}(x^*) \text{ with } \lambda_i^{ic} > 0 \\ \nabla r_i^{ic}(x^*)^T v \geq 0 & \text{for } i \in \mathcal{A}^{ic}(x^*) \text{ with } \lambda_i^{ic} = 0 \end{cases}$$

\square

1.12 Lemma

Let $(x^*, \lambda^{ec}, \lambda^{ic})$ be a KKT tuple of (CNLP). Then, it holds:

$$\forall v \in V(x^*) : v^T \nabla f(x^*) = 0$$

Proof: Let $v \in V(x^*)$. Using the stationarity condition eq. (1.0a) of the Lagrangian, we get

$$v^T \nabla f(x^*) = \sum_{i \in \mathcal{A}^{ec}} \lambda_i^{ec} v^T \nabla r_i^{ec}(x^*) + \sum_{i \in \mathcal{A}^{ic}} \lambda_i^{ic} v^T \nabla r_i^{ic}(x^*),$$

and the complementarity $\lambda_i^{ec} v^T \nabla r_i^{ec}(x^*) = 0 \forall i \in \mathcal{A}^{ec}(x^*)$ and $\lambda_i^{ic} v^T \nabla r_i^{ic}(x^*) = 0 \forall i \in \mathcal{A}^{ic}(x^*)$ from definition 1.11 completes the proof. \square

1.13 Theorem (Second order necessary and sufficient conditions)

Let $(x^*, \lambda^{ec}, \lambda^{ic})$ be a KKT-tuple in the sense of theorem 1.9, and $V(x^*)$ as in definition 1.11.

1. *Second order necessary condition:*

If x^* is a solution of (CNLP) and the LICQ holds, i.e. the Lagrangian multipliers λ^{ec} and λ^{ic} are unique, then

$$v^T \nabla_{xx} \mathcal{L}(x^*, \lambda^{ec}, \lambda^{ic}) v \geq 0 \quad \text{for all } v \in V(x^*).$$

2. *Second order sufficient condition:*

If $x^* \in \mathcal{X}$ is a feasible point and there exist $\lambda^{ec} \in \mathbb{R}^{n_{ec}}$ and $\lambda^{ic} \in \mathbb{R}^{n_{ic}}$ fulfilling the KKT conditions and

$$v^T \nabla_{xx} \mathcal{L}(x^*, \lambda^{ec}, \lambda^{ic}) v > 0 \quad \text{for all } v \in V(x^*), v \neq 0,$$

then x^* is a strict local solution of (CNLP). Note that the LICQ are not required.

Proof: See [NocedalWright2006] chapter 12.4, also [Bertsekas1995] chapter 3.3. □

Most algorithms for finding a solution to problem (CNLP) rely on solving the KKT-system in theorem 1.9 for finding candidate solutions that have to be approved being an actual solution, e.g. by checking the second order sufficient conditions specified in theorem 1.13.2.

1.2 Linear least squares problems

In this section, we present basic methods for solving linear least squares problems with and without (linear) constraints, as solving linear least squares problems founds the basis for finding the solution of nonlinear problems in section 1.3.

1.2.1 Unconstrained linear least squares

We formulate the problem of unconstrained linear least squares:

1.14 Problem (LIN-LSQ)

Let $C \in \mathbb{R}^{m \times n}$, $d \in \mathbb{R}^m$. Find $x \in \mathbb{R}^n$ as the solution of

$$\min_{x \in \mathbb{R}^n} LS(x) = \frac{1}{2} \|Cx - d\|_2^2 \tag{1.1}$$

1.15 Theorem (Unique solvability of problem (LIN-LSQ))

The function $LS(x)$ in (1.1) is convex, and a unique solution exists if

$$m \geq n \quad (\text{the system is not under-determined}) \tag{1.2}$$

and $\text{rank}(C) = n \quad (C \in \mathbb{R}^{m \times n} \text{ has full column rank}) \tag{1.3}$

Proof: Convexity of $LS(x)$ is obvious. Further, any solution x^* of problem (LIN-LSQ) fulfills $\nabla LS(x^*) = C^T(Cx^* - d) = 0$, leading to the normal equations $C^T C x^* = C^T d$. If the above conditions hold, x^* is the unique solution, as $C^T C \in \mathbb{R}^{n \times n}$ is then non-singular. □

1.2.1.1 A solution method using QR decompositions

One common way [Bjorck1996] to solve unconstrained linear least squares problems that fulfill the requirements for unique solvability in theorem 1.15 is via decomposing the system matrix C , e.g. by QR factorizations or singular value decompositions.

We shortly describe a solution algorithm using an orthogonal QR decomposition. The QR decomposition of C with column pivoting is given as

$$CP = Q\tilde{R} = [Q_1 \quad Q_2] \begin{bmatrix} R \\ 0_{(m-n) \times n} \end{bmatrix} = Q_1 R$$

or equivalently

$$Q^T C P = \begin{bmatrix} R \\ 0_{(m-n) \times n} \end{bmatrix}$$

where $P \in \mathbb{R}^{n \times n}$ is a permutation matrix, $Q = [Q_1 \quad Q_2] \in \mathbb{R}^{m \times m}$ orthogonal with submatrices $Q_1 \in \mathbb{R}^{m \times n}$, $Q_2 \in \mathbb{R}^{m \times (m-n)}$, and a nonsingular upper triangular matrix $R \in \mathbb{R}^{n \times n}$.

Since multiplication with an orthogonal matrix does not change the Euclidean norm and $PP^T = I$, we can write

$$\begin{aligned} \|Cx - d\|_2^2 &= \left\| \begin{bmatrix} Q_1^T \\ Q_2^T \end{bmatrix} \cdot (C P P^T x - d) \right\|_2^2 \\ &= \left\| \begin{bmatrix} R \\ 0_{(m-n) \times n} \end{bmatrix} \cdot P^T x - \begin{bmatrix} Q_1^T d \\ Q_2^T d \end{bmatrix} \right\|_2^2 \\ &= \|R P^T x - Q_1^T d\|_2^2 + \|Q_2^T d\|_2^2 \end{aligned}$$

The second term is independent of x ; therefore the minimum of problem (LIN-LSQ) is certainly attained if x is chosen as the solution of $R P^T x = Q_1^T d$, i.e.

$$x = P R^{-1} Q_1^T d$$

which is calculated by solving $Rz = Q_1^T d$ for z , and permuting z to restore the original order, yielding $x = Pz$.

An additional benefit of the QR factorization is that the condition number of the system matrix C is not altered; therefore this method is applicable also to badly conditioned problems, where other factorizations or the solution via normal equations $C^T C x = C^T d$ might fail.

1.2.1.2 The rank-deficient case

If in the linear least squares problem (LIN-LSQ), one or both of the conditions (1.2) or (1.3) does not hold, i.e. one or both of

$$\begin{array}{ll} m < n & \text{(the system is under-determined)} \\ \text{or} & \\ \text{rank}(C) < n & \text{(C is rank deficient)} \end{array}$$

are fulfilled, then (LIN-LSQ) has usually a set of solutions with same residual values. However, there always exists a unique solution of minimum norm, as theorem 1.17 shows. Before we state that, we need to introduce the term of a *pseudo-inverse*.

1.16 Definition and Lemma (MOORE-PENROSE pseudo-inverse)

Let $C \in \mathbb{R}^{m \times n}$ be a matrix with $\text{rank}(C) = r$. The *MOORE-PENROSE pseudo-inverse* or just *pseudo-inverse* C^\dagger of C is uniquely defined by the four conditions

$$\begin{aligned} (1) \quad CC^\dagger C &= C & (2) \quad C^\dagger C C^\dagger &= C^\dagger \\ (3) \quad (CC^\dagger)^T &= CC^\dagger & (4) \quad (C^\dagger C)^T &= C^\dagger C. \end{aligned}$$

More general, a matrix C^+ fulfilling condition (2) is called a *generalized inverse*.

If $C = U\Sigma V^T = U \begin{bmatrix} \bar{\Sigma} & 0 \\ 0 & 0 \end{bmatrix} V^T$ is the SVD of C (see theorem B.4), the pseudo-inverse can be written as

$$C^\dagger = V \begin{bmatrix} \bar{\Sigma}^{-1} & 0 \\ 0 & 0 \end{bmatrix} U^T$$

with $\bar{\Sigma} \in \mathbb{R}^{r \times r}$. □

1.17 Theorem (Unique minimum-norm solution of problem (LIN-LSQ))

Let $C \in \mathbb{R}^{m \times n}$ with $\text{rank}(C) = r \leq \min\{m, n\}$, and $d \in \mathbb{R}^m$. Then, for the linear least squares problem (LIN-LSQ)

$$\min_{x \in \mathbb{R}^n} \frac{1}{2} \|Cx - d\|_2^2$$

the solution x^* given as

$$x^* = C^\dagger d = V \begin{bmatrix} \bar{\Sigma}^{-1} & 0 \\ 0 & 0 \end{bmatrix} U^T d.$$

is the unique minimum-norm solution, i.e.

$$\forall x = \arg \min_{x \neq x^*} \frac{1}{2} \|Cx - d\|_2^2 : \|x^*\|_2 < \|x\|_2.$$

Proof: Let $C = U\Sigma V^T = U \begin{bmatrix} \bar{\Sigma} & 0 \\ 0 & 0 \end{bmatrix} V^T$ the SVD of C .

We orthogonally transform x^* and the right hand side vector d , yielding

$$\begin{aligned} \bar{x} &= V^T x^* = \begin{pmatrix} \bar{x}_1 \\ \bar{x}_2 \end{pmatrix} \quad \text{with } \bar{x}_1 \in \mathbb{R}^r, \bar{x}_2 \in \mathbb{R}^{n-r} \\ \bar{d} &= U^T d = \begin{pmatrix} \bar{d}_1 \\ \bar{d}_2 \end{pmatrix} \quad \text{with } \bar{d}_1 \in \mathbb{R}^r, \bar{d}_2 \in \mathbb{R}^{m-r} \end{aligned}$$

Then:

$$\|Cx^* - d\|_2 = \|U^T(CVV^T x^* - d)\|_2 = \left\| \begin{bmatrix} \bar{\Sigma} & 0 \\ 0 & 0 \end{bmatrix} \begin{pmatrix} \bar{x}_1 \\ \bar{x}_2 \end{pmatrix} - \begin{pmatrix} \bar{d}_1 \\ \bar{d}_2 \end{pmatrix} \right\|_2 = \left\| \begin{pmatrix} \bar{\Sigma} \bar{x}_1 - \bar{d}_1 \\ -\bar{d}_2 \end{pmatrix} \right\|_2$$

and the residual norm is minimal for $\bar{x}_1 = \bar{\Sigma}^{-1} \bar{d}_1$, and independent of the choice of \bar{x}_2 . Choosing $\bar{x}_2 = 0$ is unique and minimizes $\|\bar{x}\|_2 = \|V^T x^*\|_2 = \|x^*\|_2$, since V is orthogonal. □

1.2.2 Linear least squares with equality constraints

Equipping the unconstrained least squares problem (LIN-LSQ) with linear equality constraints leads to the linearly constrained linear least squares problem (LIN-LSQ-EC), which occurs as a subproblem when solving constrained nonlinear least squares problems utilizing active-set strategies as discussed in section 1.3.

1.18 Problem (LIN-LSQ-EC)

Let $C \in \mathbb{R}^{m \times n}$, $d \in \mathbb{R}^m$, and $m \geq n$, and further $E \in \mathbb{R}^{n_{ec} \times n}$, $b \in \mathbb{R}^{n_{ec}}$, $n_{ec} < n$.

Find $x \in \mathbb{R}^n$ as the solution of

$$\begin{aligned} \min_{x \in \mathbb{R}^n} LS(x) &= \frac{1}{2} \|Cx - d\|_2^2 \\ \text{s.t. } Ex - b &= 0 \end{aligned}$$

1.19 Theorem (Unique solvability of problem (LIN-LSQ-EC))

Problem (LIN-LSQ-EC) is uniquely solvable if

$$\text{rank}(E) = n_{ec} \quad (\text{full rank of constraint matrix}) \quad (1.4)$$

$$\text{and} \quad \text{rank} \begin{bmatrix} C \\ E \end{bmatrix} = n \quad (\text{i.e. system is not under-determined}) \quad (1.5)$$

We give a constructive proof, that combines the proof of theorem 1.19 with a solution algorithm.

Proof (nullspace method):

First, compute an orthogonal basis of the nullspace of the constraint matrix, $\mathcal{N}(E)$, by calculating a QR decomposition of E^T , implicitly exploiting the full rank assumption in eq. (1.4):

$$E^T = Q \cdot \begin{bmatrix} R \\ 0 \end{bmatrix} = [Q_1 \quad Q_2] \times \begin{bmatrix} R \\ 0 \end{bmatrix}$$

with an orthogonal matrix $Q = [Q_1 \quad Q_2] \in \mathbb{R}^{n \times n}$, $Q_1 \in \mathbb{R}^{n \times n_{ec}}$, $Q_2 \in \mathbb{R}^{n \times (n - n_{ec})}$ and nonsingular triangular $R \in \mathbb{R}^{n_{ec} \times n_{ec}}$. Then, we have $\mathcal{R}(Q_2) = \mathcal{N}(E)$, because for every $y \in \mathbb{R}^{n_{ec}}$ it holds:

$$EQ_2y = [R^T \quad 0] \cdot \begin{pmatrix} Q_1^T Q_2 y \\ Q_2^T Q_2 y \end{pmatrix} = [R^T \quad 0] \cdot \begin{pmatrix} 0 \\ y \end{pmatrix} = 0$$

as Q is orthogonal. If $x \in \mathbb{R}^n$ satisfies the constraints, i.e. $Ex = b$, it can be orthogonally split:

$$x = \bar{x} + Q_2y \quad \text{with } \bar{x} = Q_1 R^{-T} b$$

for arbitrary $y \in \mathbb{R}^{n - n_{ec}}$, since $Q_2y \in \mathcal{N}(E)$. Using $Cx - d = C\bar{x} + CQ_2y - d$, it remains to solve the reduced and unconstrained system

$$\min_{y \in \mathbb{R}^{n - n_{ec}}} \frac{1}{2} \|(CQ_2)y - (d - C\bar{x})\|_2^2 \quad (1.6)$$

If we now use the full-rank condition (1.5), then for $M \in \mathbb{R}^{m+n_{ec} \times n}$ defined as

$$M := \begin{bmatrix} C \\ E \end{bmatrix} \cdot Q = \begin{bmatrix} CQ_1 & CQ_2 \\ EQ_1 & EQ_2 \end{bmatrix} = \begin{bmatrix} CQ_1 & CQ_2 \\ R^T & 0 \end{bmatrix}$$

we have $\text{rank}(M) = n$, i.e. all columns are linearly independent. It follows that $\text{rank}(CQ_2) = n - n_{ec}$, thus maximal. This ensures the unique solvability of the reduced unconstrained least squares problem (1.6), e.g. again by QR decomposition:

$$CQ_2 = \tilde{Q} \cdot \begin{bmatrix} \tilde{R} \\ 0 \end{bmatrix}$$

with orthogonal $\tilde{Q} \in \mathbb{R}^{m \times m}$ and nonsingular upper triangular $\tilde{R} \in \mathbb{R}^{(n-n_{ec}) \times (n-n_{ec})}$. Then, splitting $\tilde{Q} = [\tilde{Q}_1 \quad \tilde{Q}_2]$, $\tilde{Q}_1 \in \mathbb{R}^{m \times (n-n_{ec})}$, we get

$$y = \tilde{R}^{-1} \tilde{Q}_1^T (d - C\bar{x})$$

as the unique solution of the reduced unconstrained problem (1.6), and finally

$$x := \bar{x} + Q_2 y$$

as the unique solution of problem (LIN-LSQ-EC). □

1.2.2.1 Two rank-deficient cases and inconsistency detection

If the full-rank condition of the constraint matrix (1.4) still holds, but the second full-rank condition (1.5) does not hold, we can still compute a minimum norm solution using the above nullspace approach.

Sticking to the notation of the proof of theorem 1.19, let y be the minimum norm solution (see theorem 1.17) of the reduced unconstrained system (1.6):

$$y = (CQ_2)^\dagger (d - C\bar{x})$$

Then, since $\bar{x} \perp Q_2 y$, for $x = \bar{x} + Q_2 y$ it holds:

$$\|x\|_2^2 = \|\bar{x}\|_2^2 + \|Q_2 y\|_2^2 = \|\bar{x}\|_2^2 + \|y\|_2^2$$

i.e. with x we have found the unique minimum norm solution of problem (LIN-LSQ-EC).

If none of the full-rank conditions (1.4) and (1.5) hold, we lose the uniqueness of the solution. If it is not known in advance, whether the full-rank conditions are fulfilled or not, one might use a rank-revealing QR decomposition as in the *direct elimination method* described below, and will at least end up in the minimum-norm solution (as long as there are no inconsistent constraints).

The method of direct elimination

Let $r = \text{rank}(E) \leq n_{ec}$ be the rank of the constraint matrix and

$$EP = Q \cdot \begin{bmatrix} R_{11} & R_{12} \\ 0 & 0 \end{bmatrix}$$

be a QR decomposition with column permutation of E (see theorem B.3), with permutation matrix $P \in \mathbb{R}^{n \times n}$ and $R_{11} \in \mathbb{R}^{r \times r}$ upper triangular and nonsingular.

The constraint equation then reads as

$$\begin{bmatrix} R_{11} & R_{12} \\ 0 & 0 \end{bmatrix} P^T x - Q^T b = 0,$$

and by setting

$$\begin{aligned} \bar{x} = P^T x &= (\bar{x}_1 \quad \bar{x}_2) && \text{with } \bar{x}_1 \in \mathbb{R}^r, \bar{x}_2 \in \mathbb{R}^{n-r} \\ \bar{b} = Q^T b &= (\bar{b}_1 \quad \bar{b}_2) && \text{with } \bar{b}_1 \in \mathbb{R}^r, \bar{b}_2 \in \mathbb{R}^{n_{ec}-r} \end{aligned}$$

the constraints reduce to

$$\bar{b}_1 = [R_{11} \quad R_{12}] \cdot \bar{x} = R_{11}\bar{x}_1 + R_{12}\bar{x}_2 \implies \bar{x}_1 = R_{11}^{-1} (\bar{b}_1 - R_{12}\bar{x}_2). \quad (1.7)$$

Note that $\bar{b}_2 \neq 0$ means that the constraints are not consistent (see remark below).

Use eq. (1.7) to eliminate \bar{x}_1 from $Cx - d$, after permuting C by P , getting $\bar{C} := CP$:

$$\begin{aligned} Cx - d &= CPP^T x - d = \bar{C}\bar{x} - d = [\bar{C}_1 \quad \bar{C}_2] \cdot \begin{pmatrix} \bar{x}_1 \\ \bar{x}_2 \end{pmatrix} - d \\ &= [\bar{C}_1 \quad \bar{C}_2] \cdot \begin{pmatrix} R_{11}^{-1} (\bar{b}_1 - R_{12}\bar{x}_2) \\ \bar{x}_2 \end{pmatrix} - d \\ &= \bar{C}_1 R_{11}^{-1} (\bar{b}_1 - R_{12}\bar{x}_2) + \bar{C}_2 \bar{x}_2 - d \\ &= (\bar{C}_2 - \bar{C}_1 R_{11}^{-1} R_{12}) \bar{x}_2 - (d - \bar{C}_1 R_{11}^{-1} \bar{b}_1). \end{aligned}$$

By setting $\tilde{C} = \bar{C}_2 - \bar{C}_1 R_{11}^{-1} R_{12}$ and $\tilde{d} = d - \bar{C}_1 R_{11}^{-1} \bar{b}_1$, the original equality constrained problem (LIN-LSQ-EC) becomes equivalent to the reduced unconstrained problem

$$\min_{\bar{x}_2 \in \mathbb{R}^{n-r}} \frac{1}{2} \|\tilde{C}\bar{x}_2 - \tilde{d}\|_2^2 \quad \text{with } \tilde{C} \in \mathbb{R}^{m \times (n-r)}$$

1.20 Remark (Sequential least squares)

Inconsistent constraints $Ex = b$ become apparent as $\bar{b}_2 \neq 0$ in the direct elimination method, and can therefore be detected during the solution process. If one proceeds anyways, the solution x^* calculated by the direct elimination method fulfills the constraints in the least squares sense. That means, that x^* is the solution of the *sequential least squares* problem

$$\min_{x \in \mathcal{X}} \frac{1}{2} \|Cx - d\|_2^2 \quad \text{with } \mathcal{X} = \arg \min_{x \in \mathbb{R}^n} \frac{1}{2} \|Ex - b\|_2^2$$

□

1.2.3 Linear least squares with inequality constraints

In this section, we shortly discuss problems of the following type:

1.21 Problem (LIN-LSQ-IC)

Let $C \in \mathbb{R}^{m \times n}$, $d \in \mathbb{R}^m$, $m \geq n$, $E \in \mathbb{R}^{n_{ec} \times n}$, $b \in \mathbb{R}^{n_{ec}}$, $n_{ec} < n$, and $G \in \mathbb{R}^{n_{ic} \times n}$, $h \in \mathbb{R}^{n_{ic}}$. Find $x \in \mathbb{R}^n$ as the solution of

$$\begin{aligned} \min_{x \in \mathbb{R}^n} LS(x) &= \frac{1}{2} \|Cx - d\|_2^2 \\ \text{s.t. } Ex - b &= 0 \\ Gx - h &\geq 0 \end{aligned}$$

Imprecise statement of inequality constrained problems

From a purely theoretical point of view, these inequality constrained problems are imprecisely stated. Provided the system is solvable, at the solution point, say x^* , the i -th inequality is either “active”, i.e. $G_i x^* - h_i = 0$, thus reduced to an equality, or it is “inactive”, i.e. $G_i x^* - h_i > 0$ and thus superfluous. Therefore, in principle, inequality constrained problems can be solved using methods for equality constrained problems, and, assuming consistent constraints, the results of theorem 1.19 for (LIN-LSQ-EC) can be transferred to (LIN-LSQ-IC).

However, the set of active inequality constraints is (in general) not known in advance, and the solution methods presented in the previous section cannot be applied until the final active set has been identified.

For solving problem (LIN-LSQ-IC), the active-set identifying algorithm 1.1 as outlined in [Bock1987] can be used. The algorithm requires as initial guess a feasible point x_0 , that can be generated by applying the same algorithm to the auxiliary problem [Schloeder1983]

$$\min_x \|Ex - b\|_2^2 + \sum_{i=1}^{n_{ic}} \min\{0, G_i x - h_i\}^2$$

Once the final active set is identified, the (active) inequality constraints are treated as equality constraints, thus theorem 1.19 applies and gives a solution strategy. In the case of more than n active constraints in the solution, some must be redundant for solvability. In the (common) case of less than n active constraints in the solution, a least squares solution is to be determined following the strategies described in the preceding section 1.2.2.

Algorithm 1.1 Solving (LIN-LSQ-IC), problem 1.21

Input: feasible point x^0 , problem data C, d, E, b, G, h

INITIALIZATION:

- 1: Set iteration counter $k := 0$
- 2: Set initial active set $AS_k := \{i : G_i x^0 - h_i = 0\}$

STEP1: Determine next active constraint

- 3: Set A_k^c and a_k^c as matrix and vector of equality and active inequality constraints

$$A_k^c := \begin{bmatrix} E \\ \{G_i\}_{i \in AS_k} \end{bmatrix} \quad \text{and} \quad a_k^c := \begin{bmatrix} b \\ \{h_i\}_{i \in AS_k} \end{bmatrix}$$

- 4: Let \bar{x}_k be the solution of the equality-constrained problem

$$\min_{x \in \mathbb{R}^n} \frac{1}{2} \|Cx - d\|_2^2 \quad \text{s.t.} \quad A_k^c x - a_k^c = 0 \quad (*)$$

- 5: If $x_k - \bar{x}_k = 0$

Go to STEP2

- 6: Check if a new inequality constraint becomes active on the line from x_k to \bar{x}_k :

$$s_{max} := \max \{s \in [0, 1] \mid G_i(x_k + s(\bar{x}_k - x_k)) - h_i \geq 0, i \notin AS_k\}$$

$$x_{k+1} := x_k + s_{max}(\bar{x}_k - x_k)$$

- 7: If constraint i became active, i.e. $G_i x_{k+1} - h_i = 0$, add it to the active set:

$$AS_{k+1} := AS_k \cup \{i\}$$

$$k := k + 1$$

Go to STEP1

otherwise continue with STEP2

STEP2: x^k is a stationary point of (*)

- 8: Let $\bar{\lambda}_k^c = (\bar{\lambda}^{ec}, \{\bar{\lambda}_i^{ic}\}_{i \in AS_k})$ be the respective multipliers.

- 9: If $\bar{\lambda}_i^{ic} \geq 0 \forall i \in AS_k$

go to KKT-POINT-FOUND

- 10: Choose $i \in AS_k$ with $\lambda_i^{ic} < 0$ and remove it from the active set:

$$AS_k := AS_k \setminus \{i\}$$

- 11: Go to STEP1

KKT-POINT-FOUND:

- 12: Set $\lambda^{ec} := \bar{\lambda}^{ec}$, $\lambda_i^{ic} := \begin{cases} \bar{\lambda}_i^{ic} & \text{for } i \in AS_k \\ 0 & \text{otherwise} \end{cases}$

- 13: Return $(x_k, \lambda^{ec}, \lambda^{ic})$ as a stationary point of (LIN-LSQ-IC),
-

1.3 Generalized GAUSS-NEWTON for nonlinear least squares problems

1.3.1 The constrained nonlinear least squares problem (NONLIN-LSQ)

The statement of a constrained nonlinear least squares problem with equality and inequality constraints is given in the following problem definition:

1.22 Problem (NONLIN-LSQ)

Let $R: \mathbb{R}^{n_x} \rightarrow \mathbb{R}^{n_{LS}}$ and $r_i: \mathbb{R}^{n_x} \rightarrow \mathbb{R}$ ($i = 1, \dots, n_{ec} + n_{ic}$).

Then, the *constrained nonlinear least squares problem* is written as

$$\begin{aligned} \min_{x \in \mathbb{R}^{n_x}} LS(x) &:= \frac{1}{2} \|R(x)\|_2^2 \\ \text{s.t. } r_i(x) &= 0 && (i = 1, \dots, n_{ec}) \\ r_i(x) &\geq 0 && (i = n_{ec} + 1, \dots, n_{ec} + n_{ic}) \end{aligned} \quad (1.8)$$

The function R is called the *residual function*, and n_{ec} and n_{ic} denote the number of equality and inequality constraints, respectively.

For ease of notation, we unify the equality constraints in the function r^{ec} , and the inequality constraints in the function r^{ic} :

$$\begin{aligned} r^{ec} : \mathbb{R}^{n_x} &\rightarrow \mathbb{R}^{n_{ec}} \quad \text{with components } r_i^{ec}(x) := r_i(x) \quad (i = 1, \dots, n_{ec}) \\ r^{ic} : \mathbb{R}^{n_x} &\rightarrow \mathbb{R}^{n_{ic}} \quad \text{with components } r_i^{ic}(x) := r_{n_{ec}+i}(x) \quad (i = 1, \dots, n_{ic}) \end{aligned}$$

such that eq. (1.8) reads as

$$\begin{aligned} \min_{x \in \mathbb{R}^{n_x}} &\frac{1}{2} \|R(x)\|_2^2 \\ \text{s.t. } &r^{ec}(x) = 0 \\ &r^{ic}(x) \geq 0 \end{aligned}$$

1.3.2 An iterative solution method based on solving linearized problems

Problems of type (NONLIN-LSQ) can be solved in an iterative procedure, where an initial guess x_0 is successively updated by the solutions Δx_k (search directions) of a series of linearized problems, usually combined with a globalization technique like a line search to ensure convergence from arbitrary points:

$$x_{k+1} := x_k + \alpha_k \cdot \Delta x_k$$

These intermediate problems are generated by applying a linearization “under the norm” to the objective of (NONLIN-LSQ) as well as to the constraints:

1.23 Problem (NONLIN-LSQ-linearized)

For a given iterate $x_k \in \mathbb{R}^{n_x}$, the *search direction* $\Delta x_k \in \mathbb{R}^{n_x}$ is the solution of the linearization of (NONLIN-LSQ):

$$\begin{aligned} \min_{\Delta x_k} &\frac{1}{2} \|\nabla R(x_k) \cdot \Delta x_k + R(x_k)\|_2^2 \\ \text{s.t. } &\nabla r^{ec}(x_k) \cdot \Delta x_k + r^{ec}(x_k) = 0 \\ &\nabla r^{ic}(x_k) \cdot \Delta x_k + r^{ic}(x_k) \geq 0 \end{aligned} \quad (1.9)$$

We denote that problem by (NONLIN-LSQ-linearized(x_k)).

The connection between original and linearized problem is established in the following result:

1.24 Lemma

$(x^*, \lambda^{ec}, \lambda^{ic})$ is a KKT-tuple of the nonlinear constrained problem 1.22 (NONLIN-LSQ)
 $\iff (0, \lambda^{ec}, \lambda^{ic})$ is a KKT-tuple of the linearized problem 1.23 (NONLIN-LSQ-linearized).

Proof: The Lagrangian of (NONLIN-LSQ) is

$$\mathcal{L}(x, \lambda^{ec}, \lambda^{ic}) = \frac{1}{2}R(x)^TR(x) - (\lambda^{ec})^T r^{ec}(x) - (\lambda^{ic})^T r^{ic}(x)$$

and the Lagrangian of (NONLIN-LSQ-linearized) is, for fixed x , given as

$$\begin{aligned} \mathcal{L}^{lin}(\Delta x, \lambda^{ec}, \lambda^{ic}) = & \frac{1}{2}(\Delta x^T \nabla R(x)^T \nabla R(x) \Delta x + \Delta x^T \nabla R(x)^T R(x) + R(x)^T \nabla R(x) \Delta x + R(x)^T R(x)) \\ & - (\lambda^{ec})^T (\nabla r^{ec}(x) \Delta x - r^{ec}(x)) - (\lambda^{ic})^T (\nabla r^{ic}(x) \Delta x - r^{ic}(x)) \end{aligned}$$

The proposition follows by application of theorem 1.9 to both problems. □

1.25 Theorem (Second order sufficient condition)

Let $(x^*, \lambda^{ec}, \lambda^{ic})$ be a KKT-tuple of (NONLIN-LSQ). If

$$v^T \nabla_{xx}^2 \mathcal{L}(x^*, \lambda^{ec}, \lambda^{ic}) v > 0 \quad \forall v \in V(x^*), v \neq 0$$

then x^* is a strict local minimizer of (NONLIN-LSQ).

Proof: By theorem 1.13. □

Theorem 1.25 is a restatement of theorem 1.13, and does not demand any regularity on the candidate point x^* . However, in general it is hard to verify this second order condition on the cone $V(x^*)$. If x^* is regular, it is sufficient to verify the positive definiteness of the Hessian of the Lagrangian on a subset $\tilde{V}(x^*) \subset V(x^*)$. For the definition of that subset, we introduce a notation for combined equality and active inequality constraints.

1.26 Definition (Combined equality and (strictly) active inequality constraints: r^c, \tilde{r}^c)

We combine equality constraints and active inequality constraint into r^c , and equality constraints with *strictly* active inequality constraints into \tilde{r}^c :

$$r^c(x) := \begin{pmatrix} r^{ec}(x) \\ \{r_i^{ic}(x)\}_{i \in \mathcal{A}^{ic}(x)} \end{pmatrix}, \quad \tilde{r}^c(x) := \begin{pmatrix} r^{ec}(x) \\ \{r_i^{ic}(x)\}_{i \in \mathcal{A}^{ic}(x), \lambda_i^{ic} > 0} \end{pmatrix}$$
□

1.27 Definition (Set $\tilde{V}(x^*)$)

For a regular KKT point x^* , we define

$$\tilde{V}(x^*) := \{v \mid \nabla \tilde{r}^c(x^*) v = 0\}$$
□

1.28 Theorem (Sufficient conditions for optimal solutions of (NONLIN-LSQ))

Let $(x^*, \lambda^{ec}, \lambda^{ic})$ be a KKT-tuple of the nonlinear constrained least squares problem (NONLIN-LSQ). If the following conditions hold

(SC) *Strict Complementarity (of the inequality constraints, see definition 1.10)*

$$\lambda_i^{ic} = 0 \iff r_i^{ic}(x^*) \underset{\neq}{>} 0 \quad (1.10)$$

(CQ) *Constraint Qualification (of full rank of linearized constraint matrix, cf. eq. (1.4))*

$$\text{rank}(\nabla r^c(x)) = n_{ec} + |\mathcal{A}^{ic}(x)| \quad (1.11)$$

(PD) *Positive Definiteness (of the Hessian of the Lagrangian, cf. theorem 1.25)*

$$v^T \nabla_{xx}^2 \mathcal{L}(x^*, \lambda^{ec}, \lambda^{ic}) v > 0 \quad \forall v \in \tilde{V}(x^*), v \neq 0 \quad (1.12)$$

then, x^* is a strict local minimizer of (NONLIN-LSQ).

Proof: See [Bock1987], chapter 3.1 □

The following perturbation theorem for the linearized problem (NONLIN-LSQ-linearized) proves that, under the given regularity conditions, it is sufficient to study equality constrained problems, as the local convergence properties of the inequality constrained problem coincides with the ones of an equality constrained problem in which only the active inequalities are considered and combined with the equality constraints.

1.29 Theorem (Perturbation theorem for the linearized problem (NONLIN-LSQ-linearized))

Let $L(x_k)$ denote the linearized problem (NONLIN-LSQ-linearized(x_k)). For a point $x_k \in \mathbb{R}^{n_x}$, let $(\Delta x, \lambda^{ec}, \lambda^{ic})$ denote a KKT tuple of $L(x_k)$. Further, the sufficient conditions [SC], [CQ], and [PD] shall be fulfilled. Then it holds:

1. $(\Delta x, \lambda^{ec}, \lambda^{ic})$ is the unique KKT tuple of $L(x_k)$, and Δx is its strict minimizer.
2. The linearized problem is uniquely solvable on an environment \tilde{X} of x_k , i.e. for all $\tilde{x} \in \tilde{X}(x_k)$, the linearized problem $L(\tilde{x})$ has a unique KKT tuple $(\Delta \tilde{x}, \tilde{\lambda}^{ec}, \tilde{\lambda}^{ic})$ and $\Delta \tilde{x}$ is its strict minimizer.
3. The set $\mathcal{A}_{L(\tilde{x})}^{ic}(\Delta \tilde{x})$ of active inequality constraints of the linearized problem $L(\tilde{x})$ at the minimizer $\Delta \tilde{x}$ is the same for all $\tilde{x} \in \tilde{X}$.

Proof: See [Bock1987], Theorem 3.1.29. □

1.30 Remark

The strict complementarity (SC) condition, eq. (1.10), ensures the constancy of the active set on a whole environment of x_k . □

For the linearized problems (NONLIN-LSQ-linearized(x_k)), the following lemma gives a more practical criterion to check the positive definiteness condition (PD).

1.31 Lemma ((PD) for the linearized problem (NONLIN-LSQ-linearized))

Let the assumptions of theorem 1.29 hold. Then, the positive definiteness condition (PD) of theorem 1.28 is equivalent to a rank condition:

$$(PD) \iff \text{rank} \begin{pmatrix} \nabla R(x^*) \\ \nabla r^c(x^*) \end{pmatrix} = n_x$$

Proof: First, the Hessian of the Lagrangian of the linearized problem (NONLIN-LSQ-linearized) is given as $\nabla_{xx}^2 \mathcal{L}(x^*, \lambda^{ec}, \lambda^{ic}) = \nabla R(x^*)^T \nabla R(x^*)$, such that the positive definiteness condition reads as

$$(PD) : v^T \nabla R(x^*)^T \nabla R(x^*) v > 0 \quad \forall v \in \tilde{V}(x^*), v \neq 0$$

Second, we prove the equivalence: (PD) \iff $\text{rank} \begin{pmatrix} \nabla R(x^*) \\ \nabla r^c(x^*) \end{pmatrix} = n_x$.

“ \implies ” : Assume $\exists v \neq 0 : \begin{pmatrix} \nabla R(x^*) \\ \nabla r^c(x^*) \end{pmatrix} v = 0$. Then, $\nabla r^c(x^*) v = 0$, and thus $v \in \tilde{V}(x^*)$,

but for such v , (PD) says $v^T \nabla R(x^*)^T \nabla R(x^*) v > 0$. Contradiction!

“ \impliedby ” : For $v \in \tilde{V}(x^*)$, we have $\nabla r^c(x^*) v = 0$, thus it must hold that $\nabla R(x^*) v \neq 0$, and therefore $0 < \|\nabla R(x^*) v\|_2^2 = v^T \nabla R(x^*)^T \nabla R(x^*) v$.

Finally, the perturbation theorem 1.29 ensures $\nabla \tilde{r}^c(x^*) = \nabla r^c(x^*)$ on an environment of x^* . \square

1.3.3 The generalized inverse as solution operator

The following theorem by BOCK (cf. [Bock1987], theorem 3.1.31) shows that the solution operator for linear constrained least squares problems is itself linear and may be written as a *generalized inverse*. As stated before theorem 1.29, we can restrict ourselves to linear equality constraints without loss of generality.

1.32 Theorem (The *generalized inverse*)

Let $x_k \in \mathbb{R}^{n_x}$, and abbreviate $R_k := R(x_k)$, $\nabla R_k := \nabla R(x_k)$, $r_k^c := r^c(x_k)$, $\nabla r_k^c := \nabla r^c(x_k)$.

Consider the linear equality constrained least squares problem, originating from the linearization (NONLIN-LSQ-linearized(x_k)), with combined constraint function $r^c : \mathbb{R}^{n_x} \rightarrow \mathbb{R}^{n_c}$,

$$\begin{aligned} \min_{\Delta x} \quad & \frac{1}{2} \|\nabla R_k \cdot \Delta x + R_k\|_2^2 \\ \text{s.t.} \quad & \nabla r_k^c \cdot \Delta x + r_k^c = 0. \end{aligned}$$

Combine $F := \begin{pmatrix} R_k \\ r_k^c \end{pmatrix}$ and $J := \begin{pmatrix} \nabla R_k \\ \nabla r_k^c \end{pmatrix}$, and let (CQ) and (PD) be fulfilled on an environment $D = D(x_k)$. Then it holds:

1. The problem has a unique (strict) solution Δx .
2. There exists a linear map $J^+ : \mathbb{R}^{n_{LS} + n_c} \rightarrow \mathbb{R}^{n_x}$ with $\Delta x = -J^+ F$.
3. The solution operator J^+ fulfills $J^+ J J^+ = J^+$, i.e. it is a *generalized inverse*.
4. Both, the generalized inverse $J^+ = J^+(x)$ and $J^+(x) F(x)$ as functions of x are continuously differentiable on the environment D .
5. The generalized inverse takes the form

$$J^+ = \begin{pmatrix} \mathbb{I} & 0 \end{pmatrix} \begin{pmatrix} (\nabla R_k)^T \nabla R_k & (\nabla r_k^c)^T \\ \nabla r_k^c & 0 \end{pmatrix}^{-1} \begin{pmatrix} \nabla R_k & 0 \\ 0 & \mathbb{I} \end{pmatrix}. \quad (1.13)$$

Proof: See [Bock1987], theorem 3.1.31. \square

1.33 Remark

1. Theorem 1.32 also holds for linear constrained least squares problems that do not originate from a linearization of (NONLIN-LSQ).
2. The generalized inverse J^+ in eq. (1.13) is usually not explicitly computed and applied to F to get the increment Δx_k , as the evaluation of $\nabla R_k^T \nabla R_k$ impairs the matrix condition. Instead, the calculations are performed in terms of matrix decompositions as described in section 1.2.2. □

1.34 Theorem (Local convergence for constrained nonlinear least squares)

Let $F := \begin{pmatrix} R \\ r^c \end{pmatrix} \in C^1(D)$ and $J := \begin{pmatrix} \nabla R \\ \nabla r^c \end{pmatrix}$ denote the combined residuals and constraints function of (NONLIN-LSQ) and its Jacobian, and let J^+ denote the corresponding generalized inverse in the sense of theorem 1.32, existing on the set $D \subset \mathbb{R}^{n_x}$. Further, J and J^+ shall fulfill the LIPSCHITZ conditions

$$\begin{aligned} \|J(y)^+(J(x+t(y-x)) - J(x))(y-x)\|_2 &\leq \omega t \|y-x\|_2^2 && \text{with } \omega < \infty \\ \|(J(z)^+ - J(x)^+)(F(x) - J(x)J(x)^+F(x))\|_2 &\leq \kappa \|z-x\|_2 && \text{with } \kappa < 1 \end{aligned}$$

for all $t \in [0, 1]$ and all $x, y, z \in D$ with $x - y = J(x)^+F(x)$.

For an initial guess $x_0 \in D$ and subsequent iterates x_k , we define

$$\beta_k := \|J(x_k)^+F(x_k)\|_2 \tag{1.14}$$

$$\delta_k := \frac{\beta_k \omega}{2} + \kappa \tag{1.15}$$

$$D_0 := \bar{B}\left(x_0, \frac{\beta_0}{1 - \delta_0}\right) \subset D \tag{1.16}$$

If $\delta_0 < 1$, then it holds:

1. The full step iteration

$$x_{k+1} := x_k + \Delta x_k \quad \text{with } \Delta x_k := -J(x_k)^+F(x_k) \tag{1.17}$$

is well defined, and $x_k \in D_0 \forall k$.

2. The iteration converges, i.e.

$$x_k \longrightarrow x^* \in D_0 \text{ for } k \longrightarrow \infty \quad \text{and} \quad \Delta x^* = -J(x^*)^+F(x^*) = 0.$$

3. An *a priori* estimation of the distances to the solution can be given as

$$\|x_k - x^*\|_2 \leq \delta_0^k \frac{\beta_0}{1 - \delta_0}.$$

4. The speed of convergence is linear: $\|p_{k+1}\|_2 \leq \delta_j \|\Delta x_k\|_2$.

Proof: See [Bock1987], theorem 3.1.44. □

1.35 Remark

1. The local contraction theorem 1.34 is affine invariant.
2. The LIPSCHITZ constant ω is a measure for the curvature of the model, thus describing the strength of its nonlinearity, and measuring the size of the area, in which the linearization is a good approximation.
3. The LIPSCHITZ constant κ , termed *incompatibility constant* by BOCK, characterizes the identifiability of the nonlinear model by disturbed data.
4. BOCK proves that the fixpoint of iteration in eq. (1.17) converges to a strict local minimizer if $\kappa < 1$; thus it is stable against perturbations. Moreover, for $\kappa > 1$, the fix point x^* is still a stationary point, but not a minimizer ([Bock1987], theorem 3.1.68, *instability of minima with $\kappa(x^*) > 1$*).

□

A note on the speed of convergence

The speed of convergence depends (1) on the nonlinearity of the underlying model and (2) on the accuracy of the data. Both can be seen by comparing the true Hessian of the Lagrangian that would be used by a classical NEWTON approach to the first order Hessian approximation in the GAUSS-NEWTON algorithm.

The contribution of the least squares objective

$$\min_x LS(x) = \frac{1}{2} \|R(x)\|_2^2 \quad \text{with } R(x) = (R_1(x), \dots, R_{n_{LS}}(x))^T \quad (1.18)$$

to the Hessian of the Lagrangian reads as

$$\nabla_{xx}^2 LS(x) = \nabla R(x)^T \nabla R(x) + \sum_{i=1}^{n_{LS}} R_i(x) \nabla_{xx}^2 R_i(x) \quad (1.19)$$

The first part, $\nabla R(x)^T \nabla R(x)$ is exactly the Hessian of the Lagrangian of the linearized problem (see the proof to lemma 1.31). It is also a good approximation of the Hessian of the nonlinear problem, if (1) the nonlinearity is small, i.e. $\nabla_{xx}^2 R_i(x)$ is small near the solution, or (2) if the available data is accurate, i.e. $R_i(x)$ is small at the solution.

Thus, if these conditions hold, the GAUSS-NEWTON Hessian approximation is close to the exact Hessian, and the speed of convergence resembles a NEWTON method, i.e. up to quadratic.

1.3.4 Globalization

The local contraction theorem 1.34 proves the convergence of the full-step GAUSS-NEWTON method when starting in the proximity of the solution – an assumption frequently found violated in practical problems.

To increase the area of convergence, damped updates of the current iterate

$$x_{k+1} := x_k + \alpha_k \cdot \Delta x_k \quad (1.20)$$

are used, in which the relaxation parameter α_k is chosen in every iteration such that the iterates approach the solution.

For unconstrained problems, approaches like the classical ARMIJO-GOLDSTEIN backtracking line search ensure global convergence by using the objective itself for measuring the progress.

For constrained problems, a common way to quantify the “approaching” to the solution is by evaluating a piecewise continuously differentiable *merit function* or *level function* $T(x_k)$ that shall decrease in every step, $T(x_{k+1}) < T(x_k)$ (*monotonicity test*), and obey the *compatibility condition* of falling in the direction of the increment:

$$\left. \frac{d}{d\varepsilon} T(x_k + \varepsilon \Delta x_k) \right|_{\varepsilon \searrow 0^+} < 0 \quad (\Delta x_k \neq 0).$$

For the level function

$$T_1(x) := \frac{1}{2} \|R(x)\|_2^2 + \sum_{i=1}^{n_{ec}} \bar{\lambda}_i^{ec} |r_i^{ec}(x)| + \sum_{i=1}^{n_{ic}} \bar{\lambda}_i^{ic} |\min\{0, r_i^{ic}(x)\}|$$

with upper bounds $\bar{\lambda}_i^{ec} > \lambda_i^{ec}$ and $\bar{\lambda}_i^{ic} > \lambda_i^{ic}$ for the Lagrangian multipliers, it can be shown [Bock1987] that the series x_k converges to an KKT point of the nonlinear least squares problem (NONLIN-LSQ) for arbitrary initial guesses x_0 , if the step size α_k in eq. (1.20) is chosen as

$$\alpha_k = \arg \min_{\alpha \in [0,1]} T_1(x_k + \alpha \Delta x_k).$$

The evaluation of T_1 is based on the values of the functions R, r^{ec}, r^{ic} only (0-th order) and does not take gradient or (approximated) curvature information into account, as the solution method itself does. Thus, the direction of steepest descent of T_1 and the search direction Δx_k might be nearly perpendicular, leading to small step sizes even in the local area of contraction.

The iteratively adjusted *natural level function* [Bock1981; Deuffhard1974]

$$T^k(x) := \|J(x_k)^+ R(x)\|^2$$

does not suffer from these problems. Further, (close to) optimal step length can be computed in every iteration using estimates of the LIPSCHITZ constant ω (curvature). See [Bock1981; Bock1987; Deuffhard1975] for details.

Restrictive monotonicity test

BOCK et al. [Bock2000b] suggest the *restrictive monotonicity test* (RMT) as a damping strategy, for which they prove that the produced step length proposals do not lead to 2-cycles for a certain choice of algorithm parameters. The RMT ensures that the performed step $\alpha \Delta x_k$ does not leave the area in which $J(x_k)$ is a good approximation of the local shape.

The interpretation of the RMT as a step size control for the explicit EULER method applied to the continuous NEWTON method (DAVIDENKO differential equations, [Davidenko1953]) allows a proof of convergence not relying on classical descent properties.

We refer to the original paper [Bock2000b].

2 Parameter Estimation in Ordinary Differential Equation Models

This chapter focuses on methods for parameter estimation in deterministic models, described by ordinary differential equations.

The first section establishes the preliminaries for parameter estimation in correct dynamical (ODE) models: the error model for measurement data, a derivation of the weighted least squares objective for maximum likelihood parameter estimation, and the concepts of structural and practical identifiability.

The second section copes with (2-point) boundary value problems (BVP). After a short discussion of *single shooting* and *collocation* methods for solving BVPs, we present the method of *multiple shooting* in detail, and compare and discuss different methods for derivative generation: external and internal numerical differentiation, simultaneous solution of the variational differential equations, and automatic differentiation techniques.

In the third and fourth section, the constrained nonlinear parameter estimation problem is derived. Starting from a continuous formulation, a multiple shooting parametrization with point constraints is given, and the nomenclature used in this thesis is established.

Lastly, the fifth section closes this chapter with discussing an *a posteriori* statistical analysis of estimates.

With exception of the first section, which gives some preliminaries and introduces general concepts, and the presentation of automatic differentiation methods, this chapter is, in preparation of the new results of chapter 5, mainly based upon notation and results of [Bock1987].

2.1 Preliminaries

2.1.1 Correct dynamical model

By the term *dynamical model*, we refer to a mathematical formulation of a real world process that changes over time; may it be a chemical reaction system, changes in population, disease spreading, or the movement of mass in a gravitational field.

Though we focus on formulations based on ordinary differential equations in this thesis (and stochastic differential equations in chapter 5), we remark that it is also suitable to apply the presented techniques to static models, e.g. when a dynamical system is in equilibrium.

By the term *ODE model*, we refer to the time development of state variables of an n_x -dimensional system of ordinary differential equations, usually stated in the form of an initial value problem

$$\dot{x} = f(t, x, p), \quad x(t_0) = x_0(p), \quad x(t) \in \mathbb{R}^{n_x}, \quad p \in \mathbb{R}^{n_p}, \quad t \in [t_0, t_f], \quad (2.1)$$

that may explicitly depend on a parameter vector p . In general, the *right hand side function* $f: \mathbb{R} \times \mathbb{R}^{n_x} \times \mathbb{R}^{n_p} \rightarrow \mathbb{R}^{n_x}$ is assumed to be sufficiently smooth, usually C^3 , though piecewise LIPSCHITZ continuity is sufficient to guarantee the existence of a solution [Hairer2000].

We further assume that these equations accurately describe the investigated real-world process, in other words, we assume that the model is structurally *correct*, and only the parameter vector p is unknown.

On notation: $x(t)$ vs. $x(t; p)$ vs. $x(t; t_k, s_k, p)$ vs. $x(t; s, p)$

We denote by $x(t; p)$ the solution trajectory of initial value problems as in eq. (2.1) or boundary value problems as in eq. (2.8) for a certain choice of parameters p , explicitly denoting the dependence of the trajectory $x(t)$ on the chosen parameter vector p .

The notation $x(t; t_k, s_k, p)$ is used for solutions of initial value problems of type $\dot{x} = f(t, x, p)$ with $x(t_k) = s_k$, which will appear in the multiple shooting parametrization in problem 2.8. Furthermore, we will write $x(t; s, p)$ or also just $x(t)$ for the concatenation of solutions of such problems. If the context is clear, or there is no parameter vector, the “ p ” is omitted.

2.1.2 Simulated and real measurements and their errors

The goal of parameter estimation is to calculate the unknown parameters $p \in \mathbb{R}^{n_p}$ from *measurements* η_i taken from the real-world process described by eq. (2.1). The term *measurement* in this context might be the quantification of a certain state variable x_i at some time points t_i , or some nonlinear function depending on time, system state, and possibly some (unknown) parameters, that themselves have to be determined by the measurements.

2.1.2.1 Measurement functions h_i

Corresponding to the realization of measurements in the real-world process, the procedure of taking measurements from the simulated model can be described mathematically by introducing *measurement functions* $h_i(t, x, p)$, with

$$h_i: \mathbb{R} \times \mathbb{R}^{n_x} \times \mathbb{R}^{n_p} \rightarrow \mathbb{R},$$

depending on a known time point t_i , the system’s state $x(t_i; p)$ at this time (which itself is determined by the parameter vector p), and possibly explicitly on (some of) the parameters p . For ease of notation, we assume the measurement functions to deliver scalar data. A vector-valued function may be split into its components, resulting in multiple measurements at the same time point. We therefore allow that time points t_i might be the same for different choices of i . We also frequently call $h_i(t_i, x(t_i; p), p)$ a *model response* at time t_i .

We note that it might also be the case, that the time points, at which the measurements are taken, are not known explicitly, but defined implicitly by the zero-crossing of a function.

2.1.2.2 Measurements η_i and their error model

We denote measurements taken from the real process by the Greek letter η . Measurements are assumed to be scalar, numbered consecutively from 1 to n_M , and the measurement η_i is taken at time t_i ($i = 1, \dots, n_M$), with monotonically increasing time-points⁽¹⁾.

For a correct model described in eq. (2.1) and true parameters p^* , the i -th measurement is assumed to be affected by independent, normally distributed additive errors ε_i with zero mean and variance σ_i^2 :

$$\eta_i = h_i(t_i, x(t_i; p^*), p^*) + \varepsilon_i, \quad \varepsilon_i \sim \mathcal{N}(0, \sigma_i^2) \quad (i = 1, \dots, n_M). \quad (2.2)$$

⁽¹⁾This is just a technical assumption. In chapter 5, we will split the measurement times and values into disjoint time intervals. A monotonic order allows a simple notation in that case.

The supposed mean value of zero for the measurement error corresponds to the assumption that the measurement equipment does not suffer from systematic errors – a desirable prerequisite. For a systematic offset in the measured quantities, is it obvious how to preprocess the data, such that the above model still suits.

Common factor of data variances

If the error variances σ_i^2 are only known up to a *common factor* β_c , e.g. if only information about the relative accuracy of measurements is given, the above error model is still adequate, and the (unknown) common factor β_c may be approximated by an independent estimator given in remark 2.9(d).

Least squares for non-normally distributed data

Even in the case of a more complicated underlying error distribution, or in the case of imperfect models, the weighted least squares method described in this chapter will still deliver an estimate for the parameter values (the least squares estimate) that has shown to be of high usefulness in many settings. Only the property of being a maximum likelihood estimate might not be guaranteed anymore.

2.1.2.3 Assuming a normal distribution is often reasonable

In many (if not most) experimental settings, the experimenter has only limited information on the exact error distribution of his measuring equipment. Most likely – and hopefully – they will know about the mean and standard deviation. It can be shown (see [Bard1974], p. 20) that, if looking for a probability density function maintaining the specified mean and standard deviation, the density of the normal distribution is the one to choose, as it does introduce the least additional assumptions.

Further, the normal distribution describes many observations in nature very well, e.g. the size of humans (see section 2.1.2.3) though it is obviously not the correct one, as it assigns positive probabilities to impossible body sizes larger than, e.g., three meters and even to negative ones. However, these *tails* of the distribution are of very low measure and in many cases it is appropriate to ignore this misfit.

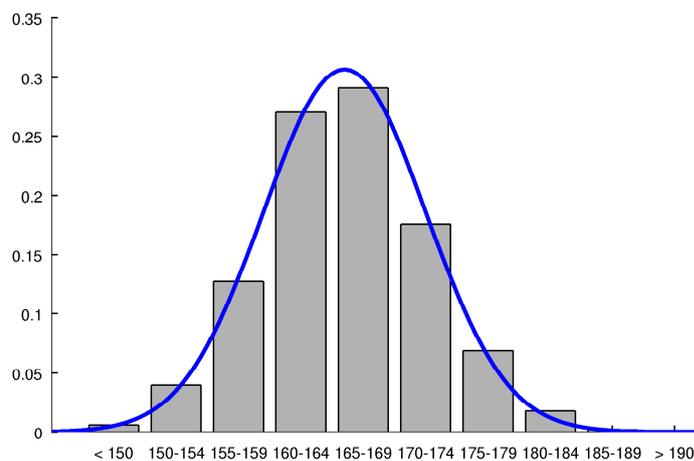


Figure 2.1: Distribution of female body size in Germany in centimeters (bar plot) and an approximating normal distribution (blue line).

Data: Sozio-oekonomisches Panel (SOEP), 2006.

At last, the *Central Limit Theorem* of LINDBERG and LÉVY states, to put it simply, that the sequence of partial sums of *any* i.i.d. random variables converges to a normally distributed random variable.

2.1.3 Maximum likelihood estimation with weighted least squares

Assuming a normal distribution of the measurement errors $\varepsilon_i \sim \mathcal{N}(0, \sigma_i^2)$, as stated in eq. (2.2), the probability density function for ε_i is given as

$$\mathbf{p}(\varepsilon_i) = \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp\left(-\frac{\varepsilon_i^2}{2\sigma_i^2}\right).$$

For pairwise uncorrelated measurement errors, $\text{corr}(\varepsilon_i, \varepsilon_j) = 0$ ($i \neq j$), i.e. independent normal random variables, the joint probability distribution of $\varepsilon = (\varepsilon_1, \dots, \varepsilon_{n_M})^T$ dissects into the product of the individual probability distributions

$$\mathbf{p}(\varepsilon) = \prod_{i=1}^{n_M} \mathbf{p}(\varepsilon_i) = (2\pi)^{-\frac{n_M}{2}} \cdot \prod_{i=1}^{n_M} \sigma_i^{-1} \exp\left(-\frac{\varepsilon_i^2}{2\sigma_i^2}\right)$$

2.1 Definition (Residuals)

For a given set of parameters p , the *residual* R_i of the i -th measurement is the difference of the actual measurement η_i and the model prediction $h_i(t_i, x_i(t_i; p), p)$:

$$R_i(p) := \eta_i - h_i(t_i, x_i(t_i; p), p)$$

and we collect them in the vector of residuals $R(p) = (R_1(p), \dots, R_{n_M}(p))^T \in \mathbb{R}^{n_M}$. □

2.2 Definition (Likelihood)

The *likelihood function of the sample* for a given set of parameters p is defined as

$$L(p) = \mathbf{p}(R(p)) = \mathbf{p}((\eta_i - h_i(t_i, x(t_i; p), p))_{i=1, \dots, n_M})$$

In the case of normally distributed measurement errors $\varepsilon_i \sim \mathcal{N}(0, \sigma_i^2)$, the likelihood reads as

$$L(p) = (2\pi)^{-\frac{n_M}{2}} \cdot \prod_{i=1}^{n_M} \sigma_i^{-1} \exp\left(-\frac{(\eta_i - h_i(t_i, x(t_i; p), p))^2}{2\sigma_i^2}\right) \quad (2.3)$$

□

For the true set of parameters p^* , the vector of residuals $R(p^*)$ equals the vector of measurement error ε , and thus the likelihood is $L(p^*) = \mathbf{p}(R(p^*)) = \mathbf{p}(\varepsilon)$. If we were able to measure without any error, i.e. $\varepsilon = 0$, the likelihood function would reach its maximum at the true parameters.

Thus, if we find a global maximizer $\bar{p} = \arg \max L(p)$ of the likelihood function, this set of parameters \bar{p} would accurately explain the observed measurements if the measurements were error-free, and thus $\bar{p} = p^*$. On the other hand, if $\varepsilon \neq 0$, maximizing the likelihood function leads to an estimate \bar{p} , that is the *most likely* set of parameters that would lead to the made observations.

In the case of independent normally distributed errors (eq. (2.2)) with known variances σ_i^2 , the formula for the likelihood is given in eq. (2.3). Instead of maximizing this product of

exponentials, we apply the monotonous logarithm, as it does not change the maximizer, but the product of exponentials may then be written as a sum of squares. By further removing constant terms, and changing maximization to minimization of the negative, we end up in the *weighted least squares* objective functional that forms the basis of this work:

$$\begin{aligned}
 \bar{p} &:= \arg \max_p \log L(p) \\
 &= \arg \max_p \log \left\{ (2\pi)^{-\frac{n_M}{2}} \cdot \prod_{i=1}^{n_M} \sigma_i^{-1} \exp \left(-\frac{(\eta_i - h_i(t_i, x(t_i; p), p))^2}{2\sigma_i^2} \right) \right\} \\
 &= \arg \max_p -\frac{n_M}{2} \log(2\pi) - \sum_{i=1}^{n_M} \log(\sigma_i) + \sum_{i=1}^{n_M} -\frac{(\eta_i - h_i(t_i, x(t_i; p), p))^2}{2\sigma_i^2} \\
 &= \arg \max_p -\frac{1}{2} \sum_{i=1}^{n_M} \frac{(\eta_i - h_i(t_i, x(t_i; p), p))^2}{\sigma_i^2} \\
 &= \arg \min_p \frac{1}{2} \sum_{i=1}^{n_M} \frac{(\eta_i - h_i(t_i, x(t_i; p), p))^2}{\sigma_i^2} \tag{2.4}
 \end{aligned}$$

The global solution \bar{p} of the above maximum likelihood problem eq. (2.4) is called the *maximum likelihood estimator* or *weighted least squares solution* of the underlying parameter estimation problem.

2.1.4 Structural and practical identifiability of parameters

There are many concepts of *identifiability* of parameters; for a detailed review we refer to the papers of COBELLI and DiSTEFANO [CobelliDiStefano1980] and MIAO et al. [Miao2011]. Here, we introduce the concepts of *structural identifiability* and *practical identifiability* for parameters in nonlinear ordinary differential equations.

Structural identifiability

The term *structural identifiability* addresses the question, whether all parameters of a nonlinear ODE may be estimated by a certain type of observation, e.g. full or partial state observations. An obvious example for an structurally *unidentifiable* system is the following exponential with unknown initial condition:

$$\dot{x}(t) = (p_1 + p_2)x(t), \quad x(t_0) = p_0, \quad t \in [t_0, t_f]. \tag{2.5}$$

Even when measuring arbitrarily many undisturbed state measurements, there is no chance in inferring the values of p_1 and p_2 without additional information, that is, p_1 and p_2 are not structurally identifiable.

On the other hand, their sum $p_s := p_1 + p_2$ together with the initial value p_0 may be determined from a two exact measurements of $x(t)$ at two distinct time points (provided $x(t)$ is not the zero function and ignoring any measurement error for the moment), i.e. the system

$$\dot{x}(t) = p_s x(t), \quad x(t_0) = p_0, \quad t \in [t_0, t_f] \tag{2.6}$$

with state observations (measurement function $h(t, x(t), p) = x(t)$) is structurally identifiable.

Practical identifiability

On the other hand, the term *practical identifiability* addresses the question, whether the amount and quality of *actual available data* is sufficient to calculate the unknown parameters. Staying at the two parameter example eq. (2.6), a *single* measurement $x(t_1)$ with $t_1 \in (t_0, t_f)$ is insufficient for determining p_s and p_0 . Assuming error-free measurements, we need at least two measurements of this type and quality at distinct time points.

In the case of noisy observations, no simple general scheme for practical identifiability may be given. The accuracy and uncertainty of parameter estimates depends on the accuracy and uncertainty of the observed data, and the minimum demands on quality of estimates depend on the respective application purpose.

Solving the inverse problem of calculating parameters from measurement data in a nonlinear ODE model, the uncertainty of the measurements propagates to the estimates in a usually nonlinear way, and may be quantified using the *FISHER information matrix*, which is the inverse of the *covariance matrix* of the estimated parameters.

Section 2.5 addresses the question of (un)certainly in terms of *confidence regions* originating from first order approximations of the covariance matrix of the estimates.

2.3 Definition (Structural global and local identifiability)

Let for an ODE model $\dot{x}(t) = f(t, x(t), p)$ an observation z of the following form be given:

$$z = z(p) = \begin{pmatrix} h(t_1, x(t_1; p), p) \\ \vdots \\ h(t_{n_M}, x(t_{n_M}; p), p) \end{pmatrix} \quad (2.7)$$

with parameter vector p taken from a parameter space $\mathcal{P} \subseteq \mathbb{R}^{n_p}$. Then, we call the system

1. *globally identifiable*, if $z(p_1) = z(p_2) \iff p_1 = p_2 \quad \forall p_1, p_2 \in \mathcal{P}$,
i.e. that identical observations are made if and only if the parameters are unique in the whole parameter space \mathcal{P} .
2. *locally identifiable*, if $z(p_1) = z(p_2) \iff p_1 = p_2 \quad \forall p_1, p_2 \in U(p^*) \subset \mathcal{P}$,
i.e. that identical observations are made if and only if the parameters are unique in a neighbourhood of a certain parameter value p^* .

We note that these definitions do not rely on any explicit initial condition of the ODE. □

Coming back to the above example of the simple exponential in eq. (2.6), the system is thus globally identifiable. As an example for an only locally identifiable system, we have a look at a trigonometric system with state observations:

$$\dot{x}(t) = \cos(t + p)$$

with solution $x(t; p) = \sin(t + p)$ and observation(s)

$$z(p) = (x(t_1; p), \dots, x(t_{n_M}; p))^T.$$

The parameter p introduces a phase shift of the sine, and since $\sin(t) = \sin(t + 2k\pi) \quad \forall k \in \mathbb{Z}$, the system is not globally identifiable. However, if we restrict the parameter space to a

neighbourhood of a certain p^* , say $U(p^*) = [p^* - 1, p^* + 1]$ we may calculate the exact value of p from a single exact measurement of the state x at an arbitrary time point. Multiple measurements will, often, also ensure *practical identifiability* in the case of noisy measurements.

2.4 Lemma (Direct testing of structural identifiability)

In the case of (nonlinear) ODE, a necessary condition for global identifiability in terms of the r.h.s. function f is given as:

$$f(t, x(t), p_1) = f(t, x(t), p_2) \implies p_1 = p_2 \quad \forall p_1, p_2 \in \mathcal{P}$$

□

2.2 Boundary value problems

Before discussing how to solve constrained parameter estimation problem as presented in section 2.3, we introduce and discuss the underlying class of *boundary value problems*.

Problems of this class specify a trajectory by a system of ordinary differential equations that has to be obeyed in a closed interval, together with some constraints on the system's states at the boundary points:

2.5 Problem (2-point boundary value problem)

$$\begin{aligned} \dot{x} &= f(t, x(t)) & t \in [t_0, t_f] \\ c(x(t_0), x(t_f)) &= 0 \end{aligned} \tag{2.8}$$

with a piecewise LIPSCHITZ-continuous r.h.s. function $f: \mathbb{R} \times \mathbb{R}^{n_x} \rightarrow \mathbb{R}^{n_x}$, differential states $x(t) \in \mathbb{R}^{n_x}$ in the time range $t \in [t_0, t_f]$, and a set of possibly coupled constraints $c: \mathbb{R}^{n_x} \times \mathbb{R}^{n_x} \rightarrow \mathbb{R}^{n_c}$ on the initial state $x(t_0)$ and the final state $x(t_f)$.

Depending on the boundary constraints $c(x(t_0), x(t_f)) = 0$, there might be no solution, several solutions or also an unique solution, e.g. $c(x(t_0), x(t_f)) = x(t_0) - x_0 = 0$ reduces the boundary value problem to an initial value problem.

2.2.1 Single shooting

Since the solution of initial value problems

$$\begin{aligned} \dot{x} &= f(t, x(t)) & t \in [t_0, t_f] \\ x(t_0) &= x_0 \end{aligned}$$

is fully determined by the initial value x_0 , and IVPs form a subgroup in the framework of boundary value problems, an intuitive, wide-spread, and easy to implement approach for solving the boundary value problem 2.5 is to iteratively adjust the initial state x_0 and (numerically) calculate the trajectory and the final state $x(t_f)$ until the boundary condition $c(x(t_0), x(t_f)) = 0$ is fulfilled with a prescribed accuracy.

This methodology of solving boundary value problems is known as *single shooting method*, the name arising from the militaristic metaphor of adjusting inclination and propellant of a gun in order to hit a distant goal. There arise diverse problems with these approaches (both militarily and mathematically; we focus on the latter).

A major difficulty is the inability of a numerical integrator to calculate a trajectory for certain initial state vectors x_0 , e.g. due to cumulated error propagation over a large time interval or exceeding step count limits. It might further happen that there does not exist a trajectory on the whole interval $[t_0, t_f]$ for certain initial values. This poses strict limits on the single shooting method.

Consider the two-dimensional example by BOCK [Bock1987] and BULIRSCH [Bulirsch1971],

$$\begin{aligned} \dot{x}_1 &= x_2 & x_1(0) &= 0 \\ \dot{x}_2 &= \mu^2 x_1 - (\mu^2 + p^2) \sin(pt) & x_2(0) &= \pi \end{aligned} \quad t \in [0, 1]$$

having for $p = \pi$ the analytical solution $x_1(t) = \sin(\pi t)$, $x_2(t) = \pi \cdot \cos(\pi t)$, independent of μ . The general solution is

$$\begin{aligned} \dot{x}_1(t) &= \sin(\pi t) + \varepsilon_1 \sinh(\mu t) + \varepsilon_2 \cosh(\mu t), & \varepsilon_1 &:= (x_2(0) - p)/\mu \\ \dot{x}_2(t) &= \pi \cos(\pi t) + \varepsilon_1 \cosh(\mu t) + \varepsilon_2 \sinh(\mu t), & \varepsilon_2 &:= x_1(0) \end{aligned}$$

and the Jacobian of the r.h.s. is $\begin{pmatrix} 0 & 1 \\ \mu^2 & 0 \end{pmatrix}$, with eigenvalues $\pm\mu$. Thus, even tiny deviations in initial values or in the parameter p are massively amplified. Even for the 64bit floating point approximation $p = 3.141592653589793$ on π , standard integrators⁽²⁾ are incapable to calculate the trajectory over the whole interval $[0, 1]$ already for moderate values of μ , say 100.

A further drawback of the single shooting approach is that information about the state trajectory, which is often available in the parameter estimation context in form of measurements, cannot be incorporated to speed up or stabilize the solution process. A frequently used remedy delivers the *collocation method* that we shortly sketch in the following section.

2.2.2 Collocation

The idea of collocation is to approximate the solution to the boundary value problem 2.5 by a linear combination of simple functions like (piecewise) polynomials or splines.

The time domain $[t_0, t_f]$ is split into a grid $G_1 := \{t_0, \dots, t_N\}$, with $t_0 < t_1 < \dots < t_N = t_f$, not necessarily equidistant. For each interval $[t_i, t_{i+1}]$, a second grid $G_2 := \{\rho_1, \dots, \rho_k\}$, with $0 \leq \rho_1 < \dots < \rho_k \leq 1$ is used to generate the $N \cdot k$ *collocation points*

$$t_{ij}^{\text{col}} := t_i + \rho_j(t_{i+1} - t_i) \quad (i = 0, \dots, N - 1; j = 1, \dots, k),$$

where the differential equation has to be fulfilled.

Let $y^{(i)}$ be a polynomial of degree k , order $k + 1$, on each interval of G_1 , defined by

$$y^{(i)}(\tau) = \sum_{s=0}^k a_s^{(i)} \tau^s \quad \text{with} \quad \tau = \frac{t - t_i}{t_{i+1} - t_i},$$

⁽²⁾e.g. the MATLAB 2013a integrator `ode45` with accuracy properties `AbsTol` set to 10^{-16} and `RelTol` set to its minimum possible value of $2.22045 \cdot 10^{-14}$ shows an error of more than 10^{10} at the final time point $t = 1$ without issuing a warning.

using the Lagrangian interpolant form here. The coefficients $a_s^{(i)} \in \mathbb{R}^{n_x}$ are determined by requiring that

1. the approximative solution fulfills the differential equation at the collocation points, i.e. $\dot{y}^{(i)}(t_{ij}^{\text{col}}) = f(t_{ij}^{\text{col}}, y^{(i)}(t_{ij}^{\text{col}})) \quad (i = 0, \dots, N-1, j = 1, \dots, k)$
2. the solution has a continuous trajectory, i.e. $y^{(i)}(t_{i+1}) = y^{(i+1)}(t_{i+1}) \quad (i = 0, \dots, N-2)$
3. the boundary conditions are fulfilled, i.e. $c(y^{(0)}(t_0), y^{(N-1)}(t_N)) = 0$.

Note that there is (formally) no integration taking place, but a solution to a nonlinear equation system has to be found. However, it can be shown that collocation is equivalent to a certain RUNGE-KUTTA method, or more formally, that collocation delivers a continuous interpolant of the points generated by this RUNGE-KUTTA method [Ascher1995]. The arising equation system is usually very large, but also exhibits a sparsity pattern that can be used for efficient solving using e.g. NEWTON-based algorithms.

In contrast to the single shooting approach, a priori knowledge about the system can (and should) be used to initialize the solution.

The choice of the number of points in the inner grid G_2 is naturally restricted by the accepted maximum degree of the piecewise polynomials. There remains the question on the choice of the grid points of G_1 . The answer strongly depends on the nonlinearity of the underlying ODE, which is in many settings not known a priori and might depend on (unknown) parameters. Adaptivity, however, is much harder to incorporate in collocation methods than in numerical integration schemes, marking a main drawback of collocation methods when solving possibly highly nonlinear differential equations.

2.2.3 Multiple shooting and derivative generation

2.2.3.1 The shooting grid and parametrization

In the single shooting approach for solving the boundary value problem 2.5, for every modification of the initial state vector x_0 at t_0 the initial value problem

$$\dot{x}(t) = f(t, x(t)), \quad x(t_0) = x_0, \quad t \in [t_0, t_f]$$

is solved by numerical integration over the whole time domain $[t_0, t_f]$. In the *multiple shooting* approach, as the name suggests, the time horizon is split by a grid, the *shooting nodes* t_k^{MS} ,

$$t_0 = t_0^{\text{MS}} < t_1^{\text{MS}} < \dots < t_{n_{\text{MS}}}^{\text{MS}} = t_f$$

into n_{MS} intervals, not necessarily all of the same size, on each of which an initial value problem is solved by numerical integration. To every time point t_k^{MS} , an initial value $s_k \in \mathbb{R}^{n_x}$ (the *shooting variables*) is associated, and on each interval the initial value problem

$$\dot{x}(t) = f(t, x(t)), \quad x(t_k^{\text{MS}}) = s_k, \quad t \in [t_k^{\text{MS}}, t_{k+1}^{\text{MS}}] \quad (k = 0, \dots, n_{\text{MS}} - 1) \quad (2.9)$$

is solved. We write the **solution on the k -th interval** as $x(t; t_k^{\text{MS}}, s_k, p)$ to visualize the dependence on the shooting node t_k^{MS} , its value s_k , and the parameter vector p . Concatenating the interval solutions $x(t; t_k^{\text{MS}}, s_k, p)$ leads to an usually discontinuous trajectory as illustrated in figure 2.2a. To ensure a continuous solution, the *continuity conditions* or *matching conditions*

$$x(t_{k+1}^{\text{MS}}; t_k^{\text{MS}}, s_k) = s_{k+1} \quad [= x(t_{k+1}^{\text{MS}}; t_{k+1}^{\text{MS}}, s_{k+1})] \quad (2.10)$$

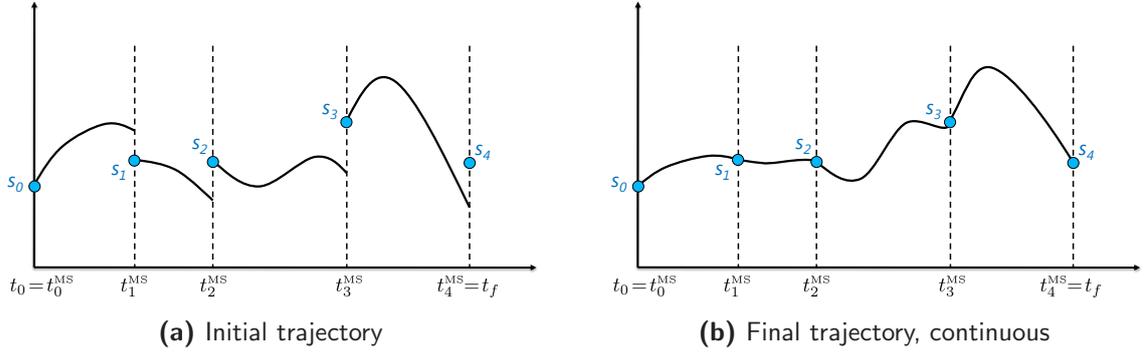


Figure 2.2: Solving boundary value problems by multiple shooting. Illustration of initial and final trajectories of a 1-dimensional boundary value problem, whose continuity and boundary value matching are enforced by eq. (2.11). The time domain has been split into four shooting intervals.

are added. Thus, having computed the interval solutions' endpoint values $x(t_{k+1}^{\text{MS}}; t_k^{\text{MS}}, s_k, p)$, the following generally nonlinear equation system in the variables s_k has to be solved:

$$F(s) = \begin{pmatrix} c(s_0, s_{n_{\text{MS}}}) \\ x(t_1^{\text{MS}}; t_0^{\text{MS}}, s_0) - s_1 \\ \vdots \\ x(t_{n_{\text{MS}}}^{\text{MS}}; t_{n_{\text{MS}}-1}^{\text{MS}}, s_{n_{\text{MS}}-1}) - s_{n_{\text{MS}}} \end{pmatrix} = 0, \quad s = \begin{pmatrix} s_0 \\ \vdots \\ s_{n_{\text{MS}}} \end{pmatrix}. \quad (2.11)$$

The concatenation of the interval solutions, as illustrated in figure 2.2b,

$$x(t) := x(t; s) := x(t; t_k^{\text{MS}}, s_k) \quad \text{for } t \in [t_k^{\text{MS}}, t_{k+1}^{\text{MS}}] \quad (2.12)$$

with $s = (s_0, \dots, s_{n_{\text{MS}}})$ is then well defined, continuous, and a solution to the boundary value problem 2.5. Note that the variables $s_{n_{\text{MS}}}$ at the last shooting node $t_{n_{\text{MS}}}^{\text{MS}}$ does not enter eq. (2.12).

The introduction of intermediate state variables $s_k \in \mathbb{R}^{n_x}$ increases the number of unknowns to $(n_{\text{MS}} + 1) \cdot n_x$, resulting in a large nonlinear system of equations to be solved, which seems undesirable at first glance. However, the special structure of the system (2.11) allows an efficient solution with a NEWTON-type method [Bock1987; Albersmeyer2010], for which derivative information is necessary (section 2.2.3.2).

The multiple shooting method combines the advantages of both the single shooting (section 2.2.1) and the collocation method (section 2.2.2), while it gets rid of their respective drawbacks. Knowledge of the underlying process, e.g. originating from experimental data, can easily be incorporated in the solution process by appropriate initialization of the state values s_k at the shooting nodes. Since the initial value problems in eq. (2.9) are fully decoupled, they can be solved independently and in parallel. Furthermore, error propagation from numerical integration is interrupted at each shooting node, thus increasing stability.

For the reasons above, we choose the multiple shooting technique for solving the problems discussed in this thesis.

2.2.3.2 Derivative generation

Solving eq. (2.11) requires the computation of derivatives of F . For that, we need to calculate the derivatives (in this context frequently called *sensitivities*) of the interval solutions $x(t_{k+1}^{\text{MS}}; t_k^{\text{MS}}, s_k)$ w.r.t. to the intermediate initial values s_k .

Precomputed analytical derivatives are available only in rare special cases, their coding is error prone and contrary to a still wide-spread belief they do not offer any advantages in terms of computation speed when compared to automatic differentiation techniques that will be discussed below.

External numerical differentiation (END)

The method of *perturbed trajectories*, mostly realized as *finite differences*, is a frequently used approach to approximate directional derivatives. Using the one-sided finite differences scheme, for a direction d , $\|d\| = 1$, and perturbation strength $h > 0$, one gets

$$\frac{d}{ds_k} x(t_{k+1}^{\text{MS}}; t_k^{\text{MS}}, s_k) \cdot d = \frac{x(t_{k+1}^{\text{MS}}; t_k^{\text{MS}}, s_k + hd) - x(t_{k+1}^{\text{MS}}; t_k^{\text{MS}}, s_k)}{h} + \mathcal{O}(h), \quad (2.13)$$

thus the evaluation of a full Jacobian on each shooting interval can be calculated with n_x additional integrations, summing up to $(n_{\text{MS}} - 1)n_x$ additional integrations in total. The error might be decreased to $\mathcal{O}(h^2)$ using central finite differences if one is willing to invest again additional n_x integrations per shooting interval.

If the integrator is available only as a black box, hence the notion *external numerical differentiation*, the method of perturbed trajectories is the only one that can be applied.

The finite differences approximation requires the integration procedure to be differentiable – an assumption that generally does not hold for error controlled integrators. Adaptive components and the usage of iterative or approximative solvers render the numerical integration non-differentiable. The externally computed derivatives are thus **inconsistent**, as the adaptive components may change for disturbed trajectories, and might not even be a rough approximation to the right ones, especially if second or higher order derivatives are to be computed.

Even if we assume a sufficiently smooth integration, there still are two sources of error: truncation and rounding.

The *truncation error* originates from the truncation of the function's TAYLOR series at first order. As exactness in eq. (2.13) only holds for $h \rightarrow 0$, a small value of h reduces the truncation error.

The *rounding error* is unavoidable in floating point arithmetic as only a finite subset of the rationals \mathbb{Q} may be represented. The rounding error shows its most malicious effects when subtracting two closely spaced values (cancellation) as it occurs in finite differencing with a small h .

For a short discussion, let $\varphi(x)$ be a sufficiently smooth real function. The one-sided finite difference approximation $D_h := \frac{\varphi(x_0+h) - \varphi(x_0)}{h}$ of the derivative $\varphi'(x_0)$ then has a maximum rounding error of $\epsilon_{\text{round}}^{\text{FD}}(h) \approx 2\epsilon_{\text{eval}}|\varphi(x_0)|h^{-1}$, and a truncation error of $\epsilon_{\text{trunc}}^{\text{FD}}(h) \approx \frac{1}{2}|\varphi''(x_0)|h$, where ϵ_{eval} denotes the relative error in the computation of φ , i.e. integrator precision in our case. Differentiating the resulting total error

$$\epsilon_{\text{total}}^{\text{FD}}(h) = \epsilon_{\text{round}}^{\text{FD}}(h) + \epsilon_{\text{trunc}}^{\text{FD}}(h) = 2\epsilon_{\text{eval}}|\varphi(x_0)|h^{-1} + \frac{1}{2}|\varphi''(x_0)|h$$

with respect to h and setting the result to zero delivers the optimum choice of the finite difference step size as

$$h^{\text{opt}} = 4\epsilon_{\text{eval}}|\varphi(x_0)/\varphi''(x_0)|$$

with an error of

$$\epsilon_{\text{total}}^{\text{FD}}(h^{\text{opt}}) = 2\sqrt{\epsilon_{\text{eval}}|\varphi(x_0)\varphi''(x_0)|}$$

in the approximation of $\varphi'(x_0)$. Thus, for a precision of ϵ in the derivative, the integrator must be accurate up to ϵ^2 , or, in other words, one loses *half of the significant*s by finite differencing.

Higher derivatives suffer even worse, rendering external numerical differentiation an computationally costly though ineffective method.

If the underlying linear algebra is capable of processing complex-valued input, for a real-valued but holomorphic function φ , there exists a remedy known as *complex step derivative approximation* or *i(maginary) trick*, based on the work of LYNESS and MOLER [LynessMoler1967], [Lyness1967]. Their findings have been reformulated by SQUIRE and TRAPP [SquireTrapp1998] to the approximation

$$\varphi'(x_0) \approx \frac{\text{Im} \varphi(x_0 + ih)}{h},$$

derived from the TAYLOR series $\varphi(x_0 + ih) = \varphi(x_0) + ih\varphi'(x_0) - h^2\frac{\varphi''(x_0)}{2!} - ih^3\frac{\varphi'''(x_0)}{3!} + \mathcal{O}(h^3)$. Taking the imaginary part and dividing by h delivers the above approximation of the first derivative with error of $\mathcal{O}(h^2)$. As the computation does not suffer from cancellation, (nearly) arbitrarily small step sizes h , e.g.. $h = 10^{-20}$ or even smaller, are admissible, effectively eliminating the truncation error. It can be “shown to be very accurate, extremely robust and surprisingly easy to implement, while retaining a reasonable computational cost”; and from the viewpoint of accuracy and robustness, it is equivalent to automatic differentiation in forward mode (see below), with slightly increased computational costs [Martins2003].

Internal numerical differentiation (IND)

If access to the integration routines is available, a technique called *internal numerical differentiation* [Bock1981; Bock1983] drastically improves the situation. The idea is to fix all adaptive components like step lengths, pivoting in the linear subproblems or other parameters of iterative or approximative solvers, thus making the integrator a differentiable mapping $I: \mathbb{R} \times \mathbb{R}^{n_x} \rightarrow \mathbb{R}^{n_x}$, accepting initial time and state as input and delivering the final state at the interval’s right boundary $(t_k^{\text{MS}}, s_k) \mapsto I(t_k^{\text{MS}}, s_k) = x(t_{k+1}^{\text{MS}}; t_k^{\text{MS}}, s_k)$.

As presented in [Bock1981], IND computes “the derivative of the internally selected discretization scheme”. When using IND with automatic differentiation (see below), rather than being a method for derivative generation, it is a principle avoiding inconsistencies between integrator evaluations for marginally changed input. The main advantage: “The method is stable in the sense that the ‘exact’ derivative of an adaptively chosen discretization scheme is generated” [Bock1981].

Freezing the integrator’s adaptive components bears the risk, that the adaptive values chosen for the nominal trajectory might not be adequate for the perturbed trajectory, e.g. the prescribed integration error tolerances might be violated. This strongly depends on the integrated function, but harsh violations are not to be expected.

A detailed discussion of IND can be found in [Albersmeyer2010; Beigel2012].

Variational differential equations

A further common approach is to solve on each shooting interval the *variational differential equations*

$$\begin{aligned} \frac{d}{dt}G(t; t_k^{\text{MS}}, s_k) &= \frac{\partial}{\partial x}f(t, x(t; t_k^{\text{MS}}, s_k)) \cdot G(t; t_k^{\text{MS}}, s_k) & t \in [t_k^{\text{MS}}, t_{k+1}^{\text{MS}}] \\ G(t; t_k^{\text{MS}}, s_k) &= \mathbb{I}_{n_x \times n_x} \end{aligned} \quad (2.14)$$

as it can be shown that $G(t; t_k^{\text{MS}}, s_k) = \frac{\partial}{\partial s_k}x(t_{k+1}^{\text{MS}}; t_k^{\text{MS}}, s_k)$ is the needed derivative (sometimes called *adjoint sensitivities*). The variational differential equations can be conjointly integrated with the original initial value problem, enlarging the system size from n_x to $(n_x + 1)n_x$ differential states, which might become computationally difficult for large systems. Care must be taken due to the fact, that the numerical solution for the initial value problem might differ in the augmented solution as the numerical integration scheme probably uses different adaptive components. Again, by using IND one overcomes these problems.

Automatic differentiation

The method of *automatic differentiation* or *algorithmic differentiation* relies on systematic usage of the chain rule. We give a glimpse on the idea in the following. For a formal introduction and detailed discussion, we refer to a definitive book on AD, e.g. the one by GRIEWANK and WALTHER [GriewankWalther2008].

We start by having a look on how a function $f: \mathbb{R}^n \rightarrow \mathbb{R}^m, (x_1, \dots, x_n) \mapsto (y_1, \dots, y_m)$ is evaluated in a computer. During the calculation of the result vector $y = f(x)$, a series of intermediate values w_i ($i = 1, \dots, N$) is generated. This series can be arranged such that

- a) $(w_1, \dots, w_n) = (x_1, \dots, x_n)$ is the input vector
- b) $w_{n+k} = \zeta_k(w_1, \dots, w_n, w_{n+1}, \dots, w_{n+k-1})$, the k -th intermediate value, depending only on previous ones
- c) $(w_{N-m+1}, \dots, w_N) = (y_1, \dots, y_m)$ the output vector.

Typically, the *elementary functions* ζ_k depend only on one or two previous variables, and consist of elementary arithmetic, trigonometric functions, exponentiation, etc., for which differentiation rules can be derived easily, i.e. ζ'_k is available.

In the *forward AD* or *forward accumulation* mode, for each intermediate w_i , an additional value w'_i is introduced to store derivative information. The values w'_1, \dots, w'_n contain the *seed*, i.e. the direction of the derivative that is to be calculated. For example, using the seed $(w'_1, \dots, w'_n) = (1, 0, \dots, 0)$ delivers the value of the derivative $\frac{\partial f}{\partial x_1}(x_1, \dots, x_n)$ in $(w'_{N-m+1}, \dots, w'_N)$. This is done by modifying the evaluation of the elementary functions ζ_k such that they also augment these intermediate values w'_i for derivative computation. For example, if $\zeta_k(a, b) = a \cdot b$ is the two-argument multiplication, the augmented elementary function then reads as $\zeta'_k(a, b, a', b') = a \cdot b' + a' \cdot b$. Table 2.1 shows how the forward AD mode is applied in the evaluation of a simple example function.

In an implementation, this can be done by rewriting the input program (AD by *source transformation*) or by augmenting the original operators in languages that support this (AD by *operator overloading*). Conditional operators (like IF) do not interfere with the above scheme, and we may calculate derivatives even for non-differential operators like MAX by rewriting $\text{MAX}(a, b)$ by $\text{IF } a < b \text{ THEN } a \text{ ELSE } b$. Care must then be taken for the case $a = b$, e.g. by ensuring that the respective derivative accumulators a' and b' coincide in this case. A rigorous

Table 2.1: Example on automatic differentiation. Evaluation of $f(x_1, x_2) = x_1 \cdot \sin(x_2) - x_1$ and the directional derivative $\frac{\partial f}{\partial x_1}(x_1, x_2) = \sin(x_2) - 1$ by automatic differentiation in forward mode. The direction $(1, 0)^T$ is put as seed in the intermediate variables (w'_1, w'_2) . While evaluating the nominal value $f(x_1, x_2)$ in w_5 , the resulting derivative value is accumulated into w'_5 at the same time.

Nominal value	Directional derivative
$f(x_1, x_2) = x_1 \cdot \sin(x_2) - x_1$	$\frac{df}{dx}(x_1, x_2) \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix}$
$w_1 = x_1$	$w'_1 = 1$
$w_2 = x_2$	$w'_2 = 0$
$w_3 = \sin(w_2)$	$w'_3 = \cos(w_2) \cdot w'_2 = \cos(x_2) \cdot 0 = 0$
$w_4 = w_1 \cdot w_3$	$w'_4 = w'_1 w_3 + w_1 w'_3 = 1 \cdot \sin(w_2) + 0 \cdot w_1 = \sin(x_2)$
$w_5 = w_4 - w_1$	$w'_5 = w'_4 - w'_1 = \sin(x_2) - 1$

approach delivering the denotational semantics based on λ -calculus and incorporating the concept of CLARKE gradients can be found in [DiGianantonioEdalat2013].

There are further variants of AD, e.g. the *backward mode*, or the propagation of TAYLOR coefficients for higher order derivatives, whose usage in the context of optimal control and parameter estimation is discussed in [Albersmeyer2010].

When using AD during the solution of an initial value problem, the principle of IND is automatically fulfilled since the derivative (or, respectively, the intermediate values needed for its compilation) is computed at the same time together with the nominal trajectory, and the computed derivatives are exact up to machine precision and consistent to the approximation of the nominal trajectory.

Although the method dates back at least to the year 1952⁽³⁾, with FORTRAN implementations already available in the 1960s⁽⁴⁾ the method is still not as widespread as one would expect regarding its advantages.

In this thesis, we implemented automatic differentiation in COMMON LISP, based on a suggestion of FATEMAN [Fateman2006] for generic arithmetic, and extended it for arbitrarily high and mixed derivatives. See also appendix S.

⁽³⁾In the proceedings of the *ACM SIGPLAN History of Programming Languages* of 1978, a transcript of the keynote address of Captain Grace Murray HOPPER, one of the builders of the first commercial large-scale computer *UNIVAC*, is conserved, where she describes how they developed and realized the idea of “analytical differentiation” on the vacuum tube based *UNIVAC* in the early 1950s. The idea was given to Harry KAHRIMANIAN, who published the work at Temple University (Philadelphia/Pennsylvania) in 1953 [Kahrmanian1953]. Around the same time, John NOLAN submitted his master’s thesis at Boston University with identical title “Analytical Differentiation on a Digital Computer” [Nolan1953], realizing automatic differentiation on the *Whirlwind I* machine at MIT Digital Computer Laboratory.

⁽⁴⁾In an article from 1964, WENGERT proposed a “procedure for automatic evaluation of total/partial derivatives of arbitrary algebraic functions” [Wengert1964] that relied on principles resembling the nowadays AD in forward mode. As appendix, the article also contains subroutines in an early FORTRAN dialect for the elementary functions addition, multiplication, and for the sine function, that automatically compute their respective derivative. With techniques like operator overloading or source code transformation not available, the programmer himself had to rewrite programs to use the augmented operators, rendering the technique a “half-automatic” process.

2.3 Constrained nonlinear parameter estimation problem

2.6 Definition (Weighted least squares objective)

Given the parameter-dependent dynamical model as initial value problem as in eq. (2.1), $\dot{x} = f(t, x, p)$, $x(t_0) = x_0(p)$, and further the corresponding measurement model as in eq. (2.2), $\eta_i = h_i(t_i, x(t_i; p^*), p^*) + \varepsilon_i$ with $\varepsilon_i \sim \mathcal{N}(0, \sigma_i^2)$ ($i = 1, \dots, n_M$), we define the *weighted least squares objective* as

$$\frac{1}{2} \sum_{i=1}^{n_M} \frac{(\eta_i - h_i(t_i, x(t_i; p), p))^2}{\sigma_i^2}$$

□

2.3.1 Additional constraints

2.3.1.1 Point constraints

In most cases, we have additional knowledge about the real-world process or the involved parameters available, that can be represented in additional pointwise equality and inequality constraints functions:

$$c^{ec}: \underbrace{\mathbb{R}^{n_x} \times \dots \times \mathbb{R}^{n_x}}_{n_{EC} \text{ times}} \times n_p \rightarrow \mathbb{R}^{n_{ec}} \quad \text{and} \quad c^{ic}: \underbrace{\mathbb{R}^{n_x} \times \dots \times \mathbb{R}^{n_x}}_{n_{IC} \text{ times}} \times n_p \rightarrow \mathbb{R}^{n_{ic}}$$

$$c^{ec}(x(t_1^{ec}; p), \dots, x(t_{n_{EC}}^{ec}; p), p) = 0 \quad c^{ic}(x(t_1^{ic}; p), \dots, x(t_{n_{IC}}^{ic}; p), p) \geq 0$$

that have to be fulfilled on specific time points $\{t_1^{ec}, \dots, t_{n_{EC}}^{ec}\}$ or $\{t_1^{ic}, \dots, t_{n_{IC}}^{ic}\}$, respectively. Equality constraints of this type may be used to force the trajectory to certain points, whereas inequality constraints might be used to prevent the trajectory entering pathological regions, or formulating restrictions on the parameters like non-negativity.

2.3.1.2 Exact measurements as point constraints

If some measurements with very high reliability (i.e. very small variance) are available, it might be beneficial for the numerical solver to treat them as equality constraints, as their inclusion into the objective functional with an extraordinary high weight would introduce the disadvantages of penalty methods mentioned earlier, leading to badly conditioned problems.

2.3.2 Constrained parameter estimation problem with underlying ODE model

2.7 Problem (Constrained ODE parameter estimation problem - continuous version)

$$\min_{x,p} \frac{1}{2} \sum_{i=1}^{n_M} \frac{(\eta_i - h_i(t_i, x(t_i; p), p))^2}{\sigma_i^2}$$

$$\text{s.t.} \quad \dot{x}(t) = f(t, x, p) \quad t \in [t_0, t_f]$$

$$x(t_0) = x_0(p)$$

$$c^{ec}(x(t_1^{ec}; p), \dots, x(t_{n_{EC}}^{ec}; p), p) = 0$$

$$c^{ic}(x(t_1^{ic}; p), \dots, x(t_{n_{IC}}^{ic}; p), p) \geq 0$$

The initial conditions on the trajectory, $x(t_0) = x_0(p)$ may also be formulated as (equality) point constraints; however, having in mind the multiple shooting parametrization, we explicitly state them here.

Using the multiple shooting parametrization described in section 2.2.3.1, we derive the parametrized version of the above parameter estimation problem, now in variables $s_0, \dots, s_{n_{\text{MS}}}$ (i.e. the state variables at the shooting nodes t_k^{MS} ($k = 0, \dots, n_{\text{MS}}$)) and the vector p of unknown parameters:

2.8 Problem (Constrained ODE parameter estimation problem - parametrized version)

$$\begin{aligned} \min_{s_0, \dots, s_{n_{\text{MS}}}, p} \quad & \frac{1}{2} \sum_{i=1}^{n_M} \frac{(\eta_i - h_i(t_i, x(t_i; s, p), p))^2}{\sigma_i^2} \\ \text{s.t.} \quad & \dot{x}(t; t_k^{\text{MS}}, s_k, p) = f(t, x, p) & t \in [t_k^{\text{MS}}, t_{k+1}^{\text{MS}}] \\ & x(t_k^{\text{MS}}; t_{k-1}^{\text{MS}}, s_{k-1}, p) - s_k = 0 & k = 1, \dots, n_{\text{MS}} \\ & c^{ec}(x(t_1^{ec}; s, p), \dots, x(t_{n_{\text{EC}}}^{ec}; s, p), p) = 0 \\ & c^{ic}(x(t_1^{ic}; s, p), \dots, x(t_{n_{\text{IC}}}^{ic}; s, p), p) \geq 0 \end{aligned}$$

where $T^{\text{MS}} = \{t_0^{\text{MS}}, \dots, t_{n_{\text{MS}}}^{\text{MS}}\}$ denotes the shooting grid, and $\{t_1^{ec}, \dots, t_{n_{\text{EC}}}^{ec}\}$ and $\{t_1^{ic}, \dots, t_{n_{\text{IC}}}^{ic}\}$ denote the grids of equality and inequality constraints, respectively. Analogously as in section 2.2.3.1, we note by $x(t; t_k^{\text{MS}}, s_k, p)$ the solution of the respective initial value problem on the k -th interval.

2.9 Remark (Condensing, correlated measurements, total least squares, common factor)

(a) **Condensing for reducing dimensionality**

The multiple shooting parametrization leads to a highly structured problem. Application of a *condensing algorithm* [Bock1987] reduces the high-dimensional system in problem 2.8 from $(n_{\text{MS}} + 1)n_x + n_p$ to only $n_x + n_p$ variables, as the (increments of) shooting variables $s_1, \dots, s_{n_{\text{MS}}}$ may be calculated using the (increment of) the shooting variable s_0 at the first node t_0^{MS} solely.

This technique also relies on the continuity condition eq. (2.10) that we are going to omit in the SDE parameter estimation chapter 2. Thus, for details on condensing, we refer to [Bock1987]. SCHLÖDER gives an extension to multi-experiment settings in [Schloeder1987].

(b) **Adjustments for correlated measurements**

The objective of the parametrized problem 2.8 may be rewritten as

$$\min_{s, p} \frac{1}{2} \|\Sigma^{-1}(\eta - h(s, p))\|^2 \quad (2.15)$$

with $\eta := (\eta_1, \dots, \eta_{n_M})^T$ being the vector of measurements, $h(s, p)$ denoting the corresponding model answer, and $\Sigma := \text{diag}\{\sigma_1, \dots, \sigma_{n_M}\}$ (see also section 2.4 for notational details).

If the data variances are correlated with positive definite correlation matrix $C \in \mathbb{R}^{n_M \times n_M}$, the objective eq. (2.15) may be used with $\Sigma^{-1} = C^{-1}$ to deliver a maximum likelihood estimate [Bard1974].

(c) **Total least squares for errors in independent variables (measurement times)**

If the measurement time points are themselves variates of a normal distribution, i.e.

$$\eta_i^t = t_i + \varepsilon_i^t \quad \varepsilon_i^t \sim \mathcal{N}\left(0, (\sigma_i^t)^2\right)$$

the method of *total least squares* may be applied by adding $\frac{1}{2} \sum_{i=1}^{n_M} \left(\frac{\eta_i^t - t_i}{\sigma_i^t}\right)^2$ to the objective in problems 2.7 and 2.8.

(d) **Estimating the common factor β_c**

If the *common factor* of the data (error) variances is not known *a priori*, then

$$b^2 := \frac{\|\Sigma^{-1}(\eta - h(s, p))\|^2}{(n_M + n_c) - (n_x + n_p)}$$

is an independent estimator for β_c [Bock1987; Bard1974]. Here, n_c denotes the sum of equality constraints and *active* inequality constraints. □

2.4 The solver's view

For a numerical solver, the parametrized problem 2.8 is translated into a general form of constrained nonlinear least squares problems that may be solved with the generalized GAUSS-NEWTON method presented in chapter 1. We introduce a vector and matrix notation of the appearing quantities that allows us a compact reformulation as problem 2.10.

$$\text{Shooting variables } s : \quad s = \begin{pmatrix} s_0 \\ \vdots \\ s_{n_{MS}} \end{pmatrix} \in \mathbb{R}^{(n_{MS}+1) \cdot n_x}, \quad s_k = \begin{pmatrix} s_{k,1} \\ \vdots \\ s_{k,n_x} \end{pmatrix} \in \mathbb{R}^{n_x} \quad (2.16a)$$

$$\text{Parameter vector } p : \quad p = \begin{pmatrix} p_1 \\ \vdots \\ p_{n_p} \end{pmatrix} \in \mathbb{R}^{n_p} \quad (2.16b)$$

$$\begin{array}{l} \text{Measurements } \eta: \\ \text{Measurement weights } \Sigma: \end{array} \quad \eta = \begin{pmatrix} \eta_1 \\ \vdots \\ \eta_{n_M} \end{pmatrix}, \quad \Sigma = \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_{n_M} \end{bmatrix} \quad (2.16c)$$

$$\text{Model response } h : \quad h(s, p) = \begin{pmatrix} h_1(x(t_1; s, p), p) \\ \vdots \\ h_{n_M}(x(t_{n_M}; s, p), p) \end{pmatrix} \quad (2.16d)$$

2.10 Problem (Constrained nonlinear least squares problem)

Collecting the shooting variables s and the vector of unknown parameters p in the vector $z := (s_0, \dots, s_{n_{MS}}, p) \in \mathbb{R}^{n_z}$ with dimension $n_z := (n_{MS} + 1)n_x + n_p$, and using the notations from eq. (2.16), we define:

$$\begin{aligned}
 R(z) &= \Sigma^{-1}(h(z) - \eta) && \text{vector of weighted residuals} \\
 r^{ec}(z) &= \begin{pmatrix} c_1^{ec}(x(t_1^{ec}; s, p), \dots, x(t_{n_{EC}}^{ec}; s, p), p) \\ \vdots \\ c_{n_{ec}}^{ec}(x(t_1^{ec}; s, p), \dots, x(t_{n_{EC}}^{ec}; s, p), p) \\ x(t_1^{MS}; t_0^{MS}, s_0, p) - s_1 \\ \vdots \\ x(t_{n_{MS}}^{MS}; t_{n_{MS}-1}^{MS}, s_{n_{MS}-1}, p) - s_{n_{MS}} \end{pmatrix} && \begin{array}{l} \text{(original equality constraints)} \\ \text{vector of equality constraints} \\ \text{(continuity conditions)} \end{array} \\
 r^{ic}(z) &= \begin{pmatrix} c_1^{ic}(x(t_1^{ic}; s, p), \dots, x(t_{n_{IC}}^{ic}; s, p), p) \\ \vdots \\ c_{n_{ic}}^{ic}(x(t_1^{ic}; s, p), \dots, x(t_{n_{IC}}^{ic}; s, p), p) \end{pmatrix} && \text{vector of inequality constraints} \\
 n_{rec} &= n_{ec} + n_{MS}n_x && \text{dimension of } r^{ec} \\
 n_{ric} &= n_{ic} && \text{dimension of } r^{ic}
 \end{aligned}$$

such that $R: \mathbb{R}^{n_z} \rightarrow \mathbb{R}^{n_M}$, $r^{ec}: \mathbb{R}^{n_z} \rightarrow \mathbb{R}^{n_{rec}}$, and $r^{ic}: \mathbb{R}^{n_z} \rightarrow \mathbb{R}^{n_{ric}}$.

Then, the *constrained nonlinear least squares problem* may be written as

$$\begin{aligned}
 \min_{z \in \mathbb{R}^{n_z}} \quad & \frac{1}{2} \|R(z)\|^2 \\
 \text{s.t.} \quad & r^{ec}(z) = 0 \\
 & r^{ic}(z) \geq 0
 \end{aligned} \tag{2.17}$$

2.5 A posteriori statistical analysis of the solution

Since the vector of measurements η is a random variable, also the solution of the parameter estimation problem 2.8 or the equivalent problem 2.10 will be a random variable.

Thus, the solution of a parameter estimation problem is incomplete if solely an estimation on the parameters is delivered, but no information about the confidence in the estimated quantities is provided.

For the analysis here, we assume that problem 2.10 has been successfully solved by the generalized GAUSS-NEWTON algorithm presented in section 1.3 with solution z^* that is not rank-deficient. Combining the equality and *active* inequality constraints in r^c , problem 2.10 then reads as

$$\begin{aligned}
 \min_{z \in \mathbb{R}^{n_z}} \quad & \frac{1}{2} \|R(z)\|^2 \\
 \text{s.t.} \quad & r^c(z) = 0
 \end{aligned} \tag{2.18}$$

with $z \in \mathbb{R}^{n_z}$, $R: \mathbb{R}^{n_z} \rightarrow \mathbb{R}^{n_M}$, $r^c: \mathbb{R}^{n_z} \rightarrow \mathbb{R}^{n_c}$.

As in section 1.3, we investigate the linearization at the solution z^* , i.e. the problem

$$\begin{aligned} \min_{\Delta z_*} \quad & \frac{1}{2} \|\nabla R_* \cdot \Delta z_* + R_*\|^2 \\ \text{s.t.} \quad & \nabla r_*^c \cdot \Delta z_* + r_*^c = 0 \end{aligned}$$

with $R_* = R(z^*)$, $\nabla R_* = \nabla R(z^*)$, $r_*^c = r^c(z^*)$, $\nabla r_*^c = \nabla r^c(z^*)$. As in theorem 1.32, we further combine

$$F = \begin{pmatrix} R_* \\ r_*^c \end{pmatrix} \quad \text{and} \quad J = \begin{pmatrix} \nabla R_* \\ \nabla r_*^c \end{pmatrix}$$

and denote by

$$J^+ = \begin{pmatrix} \mathbb{I} & 0 \end{pmatrix} \begin{pmatrix} (\nabla R_*)^T \nabla R_* & (\nabla r_*^c)^T \\ \nabla r_*^c & 0 \end{pmatrix}^{-1} \begin{pmatrix} \nabla R_* & 0 \\ 0 & \mathbb{I} \end{pmatrix}$$

the generalized inverse of J (theorem 1.32).

The following result for the linearized system delivers a linear approximation on the covariance of a computed solution z^* .

2.11 Lemma (Covariance of the estimates)

Let z^* be the solution of problem 2.10, derived by the generalized GAUSS-NEWTON algorithm from section 1.3. Using the notation above, we get for the expectation and variance of the increment $\Delta z = -J^+ F$, i.e. for the solution of the linearized problem:

$$\begin{aligned} \mathbb{E}[\Delta z] &= 0 \\ \text{Cov}[\Delta z, \Delta z] &= J^+ \begin{pmatrix} \mathbb{I} & 0 \\ 0 & 0 \end{pmatrix} (J^+)^T \end{aligned} \quad (2.19)$$

Proof: First, using the error model in eq. (2.2), i.e. $\eta_i = h_i(z^*) + \varepsilon_i$ with measurement errors $\varepsilon_i \sim \mathcal{N}(0, \sigma_i^2)$, we have $\mathbb{E}[\varepsilon] = 0$ and thus

$$\begin{aligned} \mathbb{E}[\Delta z] &= \mathbb{E} \left[-J^+ \begin{pmatrix} R_* \\ r_*^c \end{pmatrix} \right] = -J^+ \begin{pmatrix} \mathbb{E}[R_*] \\ \mathbb{E}[r_*^c] \end{pmatrix} \\ &= -J^+ \begin{pmatrix} \mathbb{E}[\Sigma^{-1}(\eta - h(z^*))] \\ 0 \end{pmatrix} = -J^+ \begin{pmatrix} \Sigma^{-1} \mathbb{E}[\varepsilon] \\ 0 \end{pmatrix} = 0 \end{aligned}$$

Second, since $r_*^c = r^c(z_*) = 0$, and since r_*^c is not a random variable, we have $\mathbb{E}[R_*(r_*^c)^T] = 0$ and $\mathbb{E}[r_*^c(r_*^c)^T] = 0$. We further get using $\Sigma^2 = \text{Cov}[\varepsilon, \varepsilon] = \mathbb{E}[\varepsilon \varepsilon^T]$ and the symmetry of the covariance that

$$\mathbb{E}[R_* R_*^T] = \mathbb{E} \left[\Sigma^{-1} (h(z^*) - \eta) (h(z^*) - \eta)^T \Sigma^{-T} \right] = \Sigma^{-1} \mathbb{E}[\varepsilon \varepsilon^T] \Sigma^{-1} = \mathbb{I}$$

and thus finally

$$\begin{aligned} \text{Cov}[\Delta z, \Delta z] &= \mathbb{E}[\Delta z \Delta z^T] = \mathbb{E}[J^+ F F^T (J^+)^T] \\ &= J^+ \mathbb{E} \left[\begin{pmatrix} R_* R_*^T & R_*(r_*^c)^T \\ (R_*(r_*^c)^T)^T & r_*^c (r_*^c)^T \end{pmatrix} \right] (J^+)^T = J^+ \begin{pmatrix} \mathbb{I} & 0 \\ 0 & 0 \end{pmatrix} (J^+)^T \end{aligned}$$

□

2.12 Remark

If the covariances σ_i of the measurements η_i are only known up to a *common factor* β_c , the covariance matrix of Δz reads as

$$\text{Cov}[\Delta z, \Delta z] = \beta_c J^+ \begin{pmatrix} \mathbb{I} & 0 \\ 0 & 0 \end{pmatrix} (J^+)^T$$

□

Approximation of the region of confidence

We conclude this chapter by giving a linear approximation (confidence intervals) on the *indifference region* $G_N(\alpha)$,

$$G_N(\alpha) := \left\{ z \mid r^c(z) = 0, \|R(z)\|^2 - \|R(z^*)\|^2 \leq \gamma^2(\alpha) \right\}$$

itself being an approximation on the $(100 \cdot \alpha)\%$ confidence region, based on the GAUSS-NEWTON solution z^* of eq. (2.18). Here,

$$\gamma^2(\alpha) := \chi_{n_z - n_c}^2(1 - \alpha)$$

denotes the quantile of the χ^2 distribution to probability α with $n_z - n_c$ degrees of freedom.

A linear approximation to $G_n(\alpha)$ is given as

$$\begin{aligned} G_L(\alpha) &= \left\{ z \mid r^c(z^*) + \nabla r^c(z^*)(z - z^*) = 0, \right. \\ &\quad \left. \|R(z^*) + \nabla R(z^*)(z - z^*)\|^2 - \|R(z^*)\|^2 \leq \gamma^2(\alpha) \right\} \\ &= \left\{ z \mid z = z^* - J^+ \begin{pmatrix} \delta w \\ 0 \end{pmatrix}, \delta w \in \mathbb{R}^{n_M}, \|\delta w\|^2 \leq \gamma^2(\alpha) \right\}, \end{aligned}$$

from which the following *confidence intervals* for the parameter estimates may be derived:

2.13 Theorem (Confidence intervals)

Let $\theta_i := \gamma(\alpha) \cdot \sqrt{C_{ii}}$, where $C_{ii} = C_{ii}(z^*)$ denotes the i -th element of the main diagonal of the covariance matrix in eq. (2.19). Then,

$$G_L(\alpha) \subset G^* := \bigtimes_{i=1}^{n_z} [z_i^* - \theta_i]$$

i.e. the linear confidence region approximation G_L is enclosed by the hyperrectangle G^* .

Proof: [Bock1987], pp. 136–138. □

Note that theorem 2.13 gives confidence intervals for both the parameter vector p and the trajectory, described by the shooting variables s .

A detailed discussion on the statistical analysis of parameter estimates in the context of linear and nonlinear constrained least squares problems based on the generalized inverse J^+ as well as a comprehensive derivation of the formulas above is also found in [Lenz2014].

3 Development of and Parameter Estimation in A Crosstalk Model of GM-CSF-mediated IL-6-induced JAK-STAT Signalling

In this chapter we develop and experimentally verify a mathematical model for an intracellular *crosstalk* in human skin cells. Keratinocytes are stimulated by the two cytokines IL-6⁽¹⁾ and GM-CSF⁽²⁾ and the intracellular response of the JAK-STAT3⁽³⁾ signalling pathway, which is frequently found aberrantly activated in various cancers, is quantified by Western blotting.

Wet lab data is carefully analysed and a normalization method is proposed that – in contrast to other frequently used normalization techniques – allows using the methods developed in chapter 2 for *maximum likelihood* parameter estimation.

Using these methods, we verify model predictions and gain new biological insights that support the identification of new therapeutic intervention sites in cancer therapy.

The starting section 3.1 addresses the connection of the investigated signal transduction pathway to cancer diseases.

Section 3.2 gives an introduction to the biological backgrounds of cytokine-induced cellular signalling, presents previous experimental observations from proliferation assays, and describes and illustrates in detail the procedures and steps of the Western blotting technique for cellular protein quantification.

The third section 3.3 describes the four distinct experimental set-ups (three chosen by expert biologists and one subsequently proposed by the author) and present the raw measurement data.

In section 3.4, we present some methods and criteria of MARSAGLIA for approximating the ratio of normal random variates by a normal distribution.

Section 3.5 addresses the statistical analysis of the generated immunoblot data. It is shown that a widespread data normalization technique using additional measurements of so-called calibrator proteins is incompatible to the assumption of normally distributed measurement errors. Further, it is shown in both an *in silico* study and in the actual *in vitro* measurements that normalization by calibrator proteins is prone to heavily disturb the actual data and induces severe artefacts. As an alternative, a different normalization technique, the *amplification factors method* that requires no additional experimental effort, is proposed and analysed. We further develop conditions under which the measurement error follows an approximate normal distribution and give criteria for the experimenter that can be easily verified using the experimental data.

⁽¹⁾Interleukin 6, details in section 3.2.1

⁽²⁾Granulocyte-Macrophage Colony-Stimulating Factor, details in section 3.2.1

⁽³⁾Janus Kinase - Signal Transducer and Activator of Transcription 3, details in section 3.2.3

These methods are applied to the measurement data in section 3.6. Using WELCH’s two-tailed t-test, it is further shown that the observed GM-CSF-moderated intensity variations in the IL-6-induced signal is statistically significant.

In section 3.7, we start with a review of existing models for IL-6-induced JAK-STAT signalling. As none of the literature models incorporates GM-CSF signalling, we develop a new *crosstalk model*. Based on mass action kinetics, the presented ODE model proposes a supportive kinase as the link between the two signalling pathways. Devised with a focus on a small number of parameters to improve identifiability, the model is capable of qualitatively and quantitatively reproducing the observations with a unique parameter set. Predictions on the time course of a negative pathway regulator are successfully used for model verification by experiments.

Finally, in section 3.8, a **new biological insight proves the predictive power** of our model. We predict a nonlinear dose-dependency of the IL-6-induced signal on GM-CSF and demonstrate it in *in vitro* experiments, showing that GM-CSF is a potent regulator of a signalling pathways frequently found persistently activated in cancer cells, and as such it is a target for therapeutic intervention.

All wet lab experiments have been performed in the group “Tumour and Microenvironment” of Prof. Dr. Margareta M. MÜLLER⁽⁴⁾ at the German Cancer Research Center (DKFZ) in Heidelberg. The obligatory “materials and methods” section is found in appendix A.1.

Parts of the results presented in this chapter have been submitted as [Sommer2014] (accepted).

3.1 Cancer

Cancer is the name of a family of related diseases. Common to all is the unchecked growth of cells, the natural replacement of aged and damaged cells, triggered by inter- and intracellular communication, is disturbed.

Altered or impaired cytokine signalling and inflammation is frequently observed in cancer [Seruga2008], especially the pro-inflammatory properties of IL-6-type cytokines have been linked to tumourigenesis [Putoczki2010]. Persistent JAK-STAT3 activation by IL-6 and the IL-6-type cytokine IL-1 1 is known to maliciously affect the reproduction of intestinal epithelial cells, and promotes gastric tumourigenesis [Ernst2008]. Impaired JAK-STAT3 signalling is also found in various other cancers [Buettner2002].

Regulators of JAK-STAT3 signalling are therefore an **important target for therapeutic intervention in cancer**, thus requiring the identification of new regulators and their regulation mechanisms. In this study, we determined GM-CSF as such a potent regulator.

A review of the role of (disturbed) JAK-STAT signalling in cancer is given, e.g., by YU et al. [Yu2014].

⁽⁴⁾Now: Faculty of Mechanical and Medical Engineering, Hochschule Furtwangen University



Figure 3.1: Biological assembly image of human IL-6 and GM-CSF as published in the RCSB Protein Data Bank

3.2 Biological and experimental background

3.2.1 The cytokines IL-6 and GM-CSF

Cytokines are small proteins that are involved in many cellular signalling processes in multicellular organisms. They are released by cells and influence the behaviour of other cells or themselves by binding to certain receptors. Cytokines play an important role in the immune system, growth regulation and differentiation of certain cell types, as well as in inflammatory processes. Disordered cytokine signalling is ubiquitously found in cancer [Seruga2008].

3.2.1.1 IL-6

Interleukin 6 (IL-6) is a multifunctional cytokine that is secreted by many cell types like fibroblasts, endothelial cells, keratinocytes, macrophages, T and B lymphocytes as well as various tumour cell lines [vanSnick1990]. It shows lots of opposite qualities, as it acts as both pro- and anti-inflammatory protein. IL-6 also plays a two-sided role in cancer. It exhibits pro-tumourigenic abilities as well as anti-tumourigenic properties, the latter for instance by stimulating the anti-tumour activity of macrophages and preventing apoptosis in neutrophils, thus increasing their cytotoxicity in tumours [Tripathi2003].

3.2.1.2 GM-CSF

Granulocyte-macrophage colony-stimulating factor (GM-CSF) is a type I interleukin like IL-6, and has pleiotropic functions. It is crucial in formation of blood cellular components, and modulates differentiation and proliferation of many hematopoietic cells and their precursors [deGroot1998]. Further, GM-CSF acts as an autocrine (i.e. self-stimulating) signalling molecule in squamous cell carcinoma (SCC), where it stimulates growth and progression of tumour cells, and it is frequently found in certain types of solid tumours, e.g. head and neck SCC [Ninck2003].

3.2.2 The HaCaT-ras A5 cell line, cancer, and their connection to IL-6 and GM-CSF

All experiments have been performed using an *in vitro* carcinogenesis model of human skin keratinocytes. These HaCaT⁽⁵⁾ cells, originating from keratinocytes that have been extracted from the proximity of a melanoma of a 62-year old male patient in the 1980s, have been

⁽⁵⁾human adult low calcium high temperature keratinocytes; the cell line's name reflects the original processing conditions in the lab

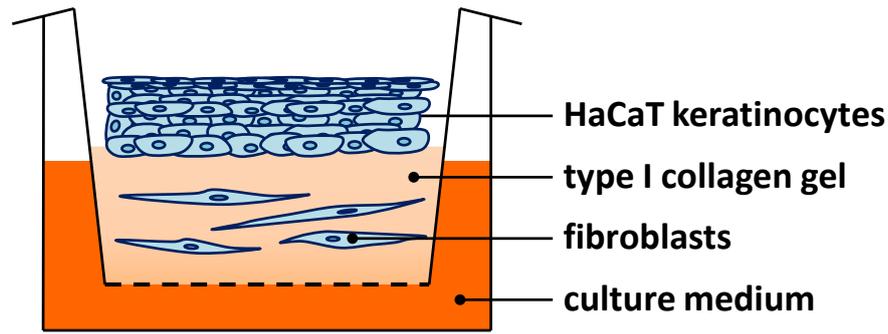


Figure 3.2: 3D organotypic co-culture model as stromal equivalent. Air exposed growth of epithelial tumour cells (HaCaT A5 keratinocytes) on top of a type 1 collagen gel. Like in human skin, the gel is pervaded by fibroblasts. A nutrient solution is located below the collagen gel without direct contact to the top layer of keratinocytes.

shown to be immortal, non-tumourigenic, and exhibiting a normal differentiation capability [Boukamp1988]. Due to these properties, HaCaT cells are used as a model for the investigation of skin cancer *in vitro*.

A stable transfection of HaCaT with a the human ras oncogene [Fusenig1998] leads to a tumourigenic phenotype. One of the resulting cell lines, HaCaT-ras A5, shows benign non-invasive tumour growth in 3D organotypic co-culture (OTC) models (schema in figure 3.2 on this page) and in living mice (figure 3.3 on the next page, A-5C3 Control)) [Lederle2011; Depner2014].

Additional transfection of HaCaT-ras A5 to constitutively overexpression and production of IL-6 transforms the benign phenotype to a malignant one *in vivo* (figure 3.3 on the facing page). With increasing IL-6 expression, also the invasive abilities of the tumour increase. By application of an IL-6 neutralizing antibody, invasion of tumour cells into subjacent tissue can be prevented (figure 3.3 on the next page, right), thus proving the facilitating effects of IL-6 on tumour progression.

Further, it has been shown that GM-CSF is also a potent activator of tumour progression in both the HaCaT model of skin and in SCC of living patients [Mueller2001]. In the OTC model, presence of GM-CSF results in invasive growth and tumour cell proliferation, whereas *in vivo* the observed tumour growth is only transient [Obermueller2004].

3.2.3 Signalling pathways

One possibility for cells to react on environmental changes is via cytokine signalling. Cytokines are secreted by cells in order to change the behaviour of certain target cells. The binding of a cytokine to a cell surface receptor is recognized for example by conformational changes of the receptor proteins that modify the activity of the intracellular parts of the receptor complex and subsequently triggering intracellular processes. Binding of cytokines to a receptor may activate, cease or modify specific signalling pathways, both independent of and depending on ligand concentration.

For instance, the intracellular parts of a ligand-bound receptor may become activated by phosphorylation of certain amino acids of the receptor protein, in most cases occurring on

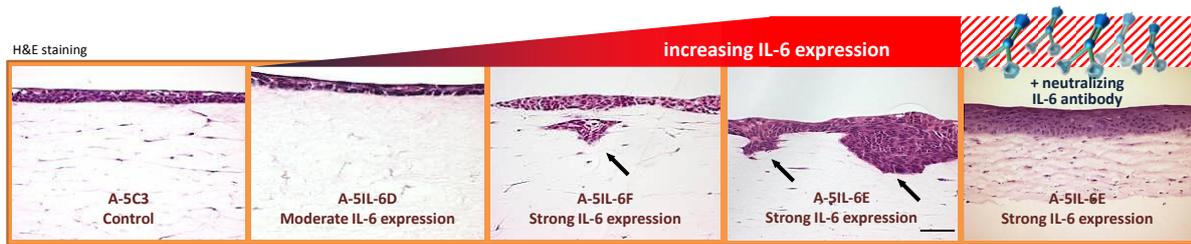


Figure 3.3: IL-6 facilitates tumour progression. HaCaT-ras A5 cells were stably transfected to constitutively overexpress IL-6. Observation of growth behaviour, proliferation and invasion potential in the OTC-Model (see figure 3.2 on the facing page) shows that IL-6 overexpression induces tumour cell invasion into the collagen gel. This effect can be blocked with an IL-6 neutralizing antibody (right), proving IL-6 plays an essential role in malign degeneration. OTC H&E staining pictures kindly provided by Margareta M. MÜLLER, see picture credits on page 240.

serine, threonin, tyrosine, or histidine residues, which become thereby binding sites of further intracellular signalling molecules.

In many cases, a whole network of interior cell proteins is involved in transmitting the external stimulus from the receptor to a target place inside, possibly forwarding the information into the nucleus to initiate transcription of target genes. Subsequent protein neogenesis enables the cell to react in a specific way to the signal.

These signalling pathways sometimes involve hundreds (or more) of substances, of which many are not even identified yet. Further, there is no one-to-one relation between cytokines and triggered signalling pathways; usually, more than one single pathway is activated by a certain cytokine. Signalling pathways are cell-type specific (as different cell types may have different receptors), and also dosage of stimulus as well as environmental conditions decide about the response upon a certain signal.

In the case of IL-6 and GM-CSF, it is known that both cytokines trigger at least two common signalling pathways, namely the MAPK/ERK⁽⁶⁾ and the JAK-STAT⁽⁷⁾ pathway, of which we will focus on the latter.

3.2.3.1 IL-6-induced JAK-STAT signalling pathway

The predominant pathway that is activated by IL-6 is the JAK-STAT signalling pathway, to which we give an introduction here. IL-6 first binds to its extracellular α receptor IL-6R α ⁽⁸⁾ before the IL-6-IL-6R α -complex is recruited to the exterior part of the gp130 receptor. The stoichiometry (hexameric, tetrameric, or other forms) is still under discussion [Groetzinger1999].

Once the ligand is bound, gp130 dimerises and initiates the intracellular signalling cascade by recruiting receptor associated Janus kinases (JAK), followed by a mutual phosphorylation of JAK and the cytoplasmic part of gp130. STAT proteins can bind to the now activated receptor

⁽⁶⁾Mitogen Activated Protein Kinase / Extracellular-signal Regulated Kinases, also known as Ras-Raf-MEK-ERK pathway, a multi-tiered intracellular signalling cascade that communicates an extra-cellular signal to the nucleus and DNA; plays an important role in cellular differentiation, growth, apoptosis, and many other processes [Pearson2001]

⁽⁷⁾JAK: Janus Kinase, STAT: Signal Transducer and Activator of Transcription, details on this pathway are given in the following sections and the reviews [Rawlings2004; Heinrich1998]

⁽⁸⁾ The α receptor is not involved in the intracellular signal transduction, but binding to it is a requirement for the subsequent binding to the cytokine receptor gp130. Other names for IL-6R α are CD126 or gp80.

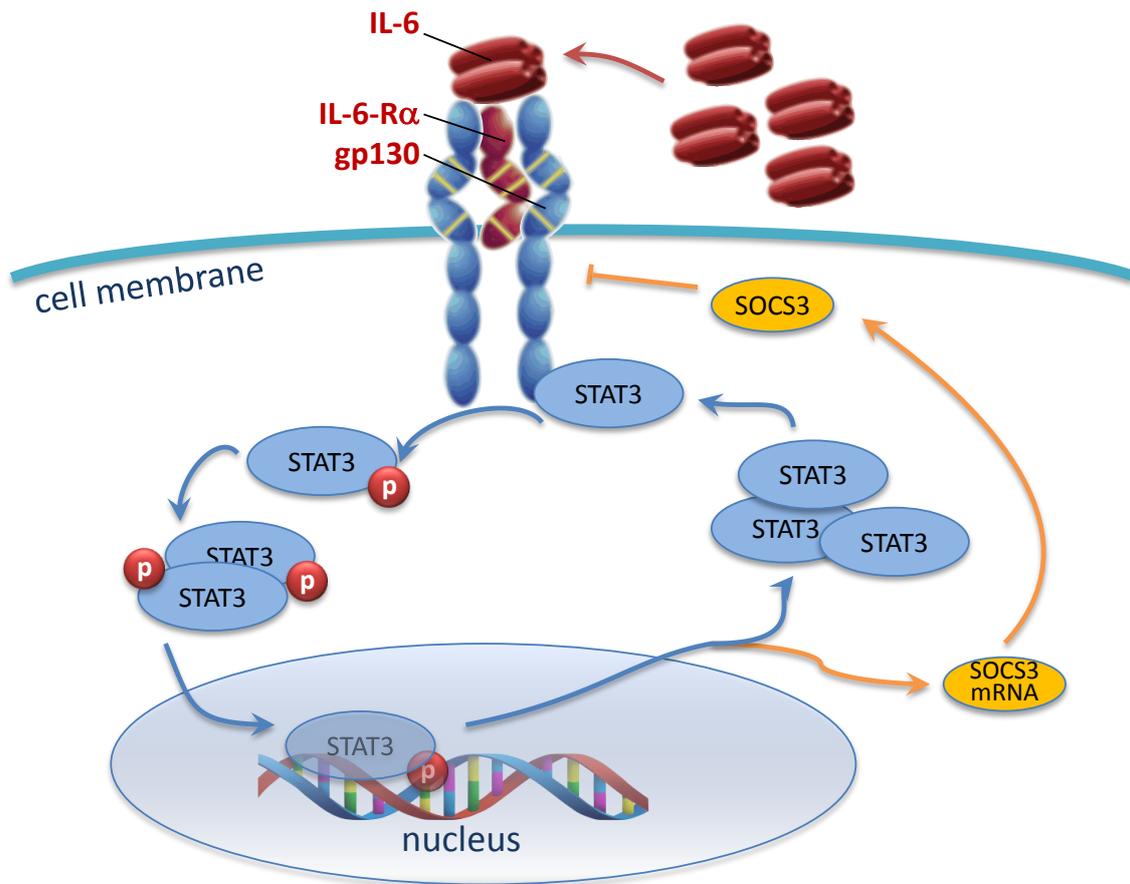


Figure 3.4: Simplified IL-6-induced JAK-STAT signalling pathway. The ligand IL-6 activated the receptor gp130 by binding to it. STAT-3 molecules can bind to the activated receptor, get phosphorylated and form dimers, which are able to translocate into the nucleus. There, after inducing the transcription of SOCS-3 mRNA, they get dephosphorylated, leave the nucleus and join the cytoplasmic pool of monomeric STAT-3. The newly produced SOCS-3 mRNA is also exported to the cytoplasm and translated into SOCS-3 protein, which can bind to the active receptor and cease the signalling cascade.

complex (mostly STAT-3 and STAT-1), become tyrosine-phosphorylated, and translocate into the nucleus, where they regulate the transcription of target genes. Negative regulation of the transcriptional activity of nuclear STAT is done by members of the PIAS⁽⁹⁾ family. After dephosphorylation, STAT returns into the cytoplasm.

The predominant ceasing factor of JAK-STAT signalling is the SOCS⁽¹⁰⁾ family of feedback inhibitors (in case of STAT-3 signalling, mostly SOCS-3), which interrupt the phosphorylation of gp130, JAK and STAT molecules and thus initiate termination of signalling.

A simplified view on the pathway is depicted in figure 3.4 on the current page. Detailed information on IL-6 and JAK-STAT signalling pathways can be found e.g. in the reviews [Heinrich1998; Rawlings2004].

⁽⁹⁾Protein Inhibitor of Activated STAT

⁽¹⁰⁾ Suppressor Of Cytokine Signalling

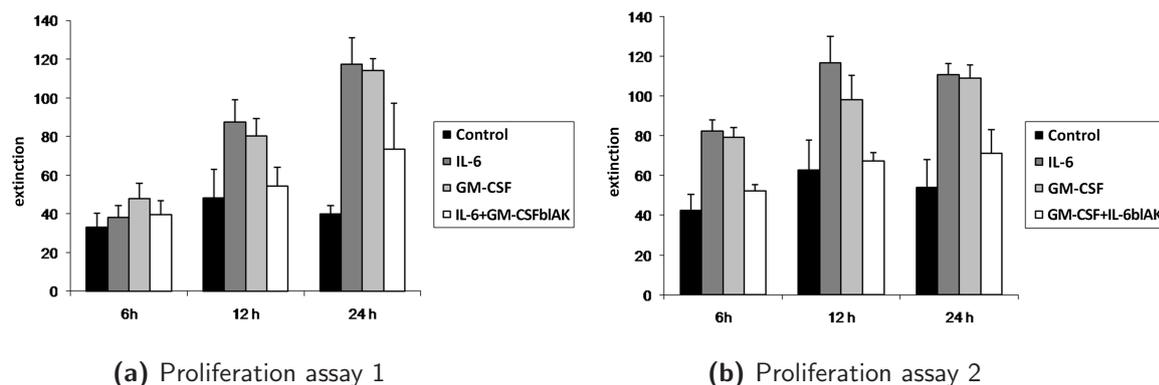


Figure 3.5: Proliferation of HaCaT in different stimulation scenarios

Proliferation of unstimulated HaCaT (both assays, leftmost bars, control), upon stimulation with IL-6 and GM-CSF (both assays, middle bars), upon stimulation with IL-6 while blocking GM-CSF with antibodies (section 3.2.4, rightmost bar), and upon stimulation with GM-CSF while blocking IL-6 with antibodies (section 3.2.4, rightmost bar).

The term *extinction* refers to photometric measurements. It is proportional to the concentration of an agent in solution. Figures kindly provided by Margareta M. MÜLLER, see picture credits on page 240.

3.2.3.2 GM-CSF-induced JAK-STAT signalling

Similar to IL-6, the GM-CSF ligand first binds to its cytokine-specific α chain GM-CSFR α . In a second step, this complex of ligand and α receptor binds the common β chain (βc , CD131, CSF2RB), a signal transducing receptor that is shared between multiple cytokines. Again, binding of a ligand results in activation of receptor associated kinases, followed by phosphorylation of distinct STAT molecules (predominantly STAT-5). Tyrosine phosphatases and the family of SOCS proteins have been described as negative feedback regulators in GM-CSF signalling.

The reviews [deGroot1998; MartinezMoczygemba2003] give detailed information on cytokine signalling via the common β chain.

3.2.4 Proliferation assays reveal a link between IL-6 and GM-CSF in HaCaT

Both factors, IL-6 and GM-CSF, stimulate the proliferation of HaCaT cells *in vitro* with similar outcome when applied separately (figure 3.5). After 24 hours, the number of cells roughly doubled (compare the stimulation bars of IL-6 and GM-CSF in figure 3.5 to the control bars).

Stimulation with one agent while blocking the other using appropriate antibodies shows a different outcome (rightmost white bars in section 3.2.4). Cell proliferation is only slightly raised compared to control, and considerably lower compared to unblocked stimulation with the other agent.

These observations indicate a **reciprocal interaction** between the two factors and point to a connection of the signalling pathways of IL-6 and GM-CSF. Further experiments as described in section 3.3 demonstrate interactions in the JAK-STAT signalling pathways.

Table 3.1: Individual and shared components in JAK-STAT signalling.

IL-6 and GM-CSF share some signalling components. The mark \checkmark means that the respective component takes an active role in the respective signalling pathway, whereas (\checkmark) in the transducer section codes a less intense activation. The mark $-$ means that the respective component is not or only weakly involved.

		IL-6	GM-CSF
α receptors	IL-6R α	\checkmark	$-$
	GM-CSFR α	$-$	\checkmark
β receptors	gp130	\checkmark	$-$
	β c	$-$	\checkmark
tyrosine kinases	JAK-1	\checkmark	$-$
	JAK-2	\checkmark	\checkmark
	TYK-2	\checkmark	$-$
signal transducers	STAT-1	(\checkmark)	$-$
	STAT-3	\checkmark	(\checkmark)
	STAT-5 a/b	(\checkmark)	\checkmark
signal suppressors	SOCS-3	\checkmark	\checkmark

3.2.4.1 Known interconnections in GM-CSF and IL-6-induced JAK-STAT signalling

On the level of signal detection, IL-6 signals through its specific α receptor IL-6R α and the common gp130 receptor, while GM-CSF uses GM-CSFR α and the common β chain. At receptor level, no shared components are currently known.

At the level of tyrosine kinases, IL-6 stimulation leads to an activation of JAK-1, JAK-2, and TYK-2, while GM-CSF triggers the phosphorylation of JAK-2, rendering this specific tyrosine kinase a possible interconnection point of both pathways. As a shared component, there might be a concurrency situation upon co-stimulation with IL-6 and GM-CSF, which is indeed observed (see table 3.3b). However, if JAK-2 had been a limiting factor, blocking GM-CSF while simultaneously stimulating with IL-6 would not have resulted in the observed lowered pSTAT-3 signal (see table 3.3c). In further conclusion, these observations also make it very unlikely that the interconnection is due to other, – possibly unknown – shared and thus limiting factors.

At the level of signal transducers, IL-6 is known to activate STAT-1, STAT-3, and STAT-5 a/b, while predominantly STAT-3 is used as transducer. GM-CSF activates STAT-3 and STAT-5 a/b, where the latter are the main transducer molecules. Obviously, STAT-3 is a common signal transducer of both cytokines. To give one result in advance: Although both cytokines signal via STAT-3, we did not observe an increased STAT-3 activation in the IL-6 and GM-CSF co-stimulation experiments compared to sole IL-6 stimulation. In fact, we made opposite observations (table 3.3b, see also section 3.6.2).

Table 3.1 summarizes JAK-STAT pathway component activation.

3.2.5 Data generation by quantitative immunoblotting (WESTERN blotting)

Immunoblotting or Western⁽¹¹⁾ blotting names the procedure of transferring proteins from an electrophoresis gel onto a membrane, on which the proteins can be further processed and detected.

In what follows, we give a description how the experimental data that was used in this project has been produced. Having a view on the complexity on the involved processes, one gets an impression about possible sources of error.

Though the presentation of experimental data as a time series might seem like samples were taken from a single cell population after different amount of time, in fact, there are *multiple* sets of cells in dishes, at least one for each time point, and they all differ in cell count. Each Petri dish receives a stimulus individually.

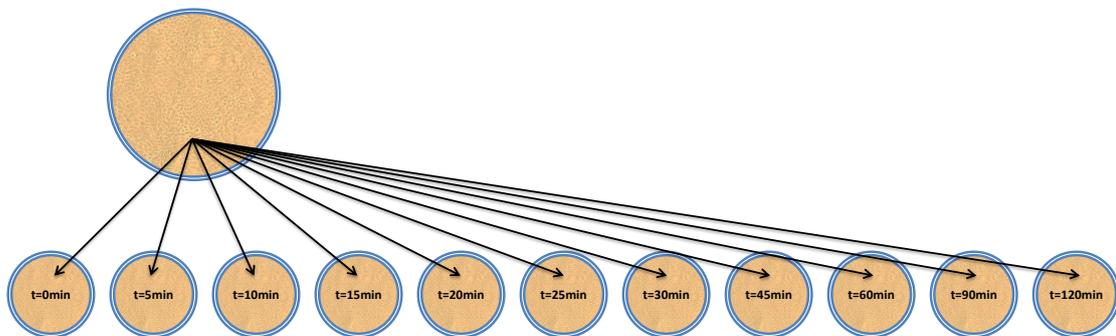
We will describe and illustrate the individual steps of the experimental procedure in the following sections 3.2.5.1 to 3.2.5.3.

3.2.5.1 Step 1: Cell preparation, stimulation, and lysis

1a) Seeding and growing of cells in multiple wells (splitting)

For each intended time point, a Petri dish with an (ideally fixed) amount of HaCaT cells gets prepared and incubated with a nutrient solution until confluency⁽¹²⁾ is reached.

Before stimulation, the nutrient solution is removed and replaced by a starvation medium in order to synchronize cell cycles and to cease active intracellular signalling pathways.

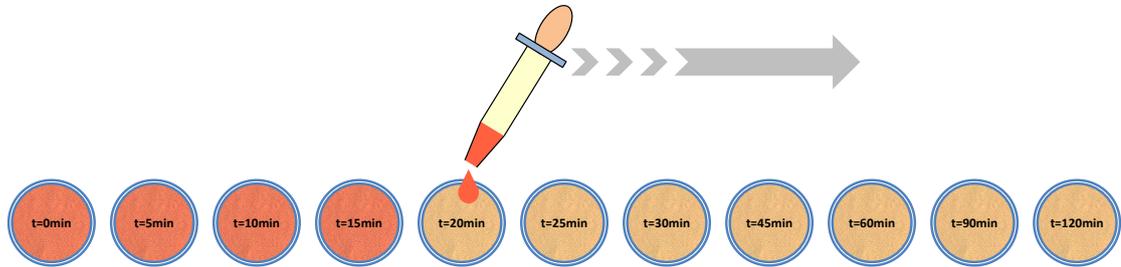


⁽¹¹⁾In the 1970s, Sir Edwin Mellor SOUTHERN developed a method for analysing DNA by separation according to molecule size and transfer to a membrane (Edwin M. Southern: “Detection of specific sequences among DNA fragments separated by gel electrophoresis”. *J. Mol. Biol.* 98 (Nov. 1975), pp. 503–517). The technique was finally named after its inventor. In allusion to the remaining geographical directions, technically similar RNA and protein detection techniques have been given the names of *Northern* and *Western* blotting. The term *Eastern* blotting is not consistently used but refers to similar techniques.

⁽¹²⁾Confluency is the state when the dish surface is completely covered by a single layer of cells.

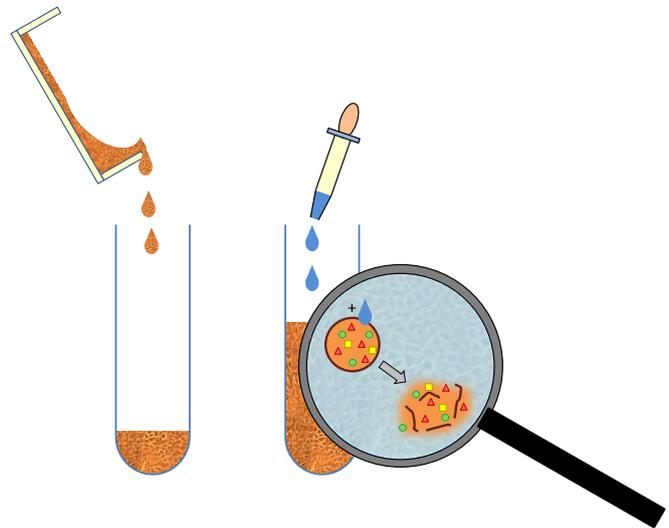
1b) Stimulation

Stimulation is performed by substituting the starvation medium by a conditioned medium containing the cytokines (IL-6 or GM-CSF) in each dish. At least one dish containing unstimulated cells serves as control and reference.



1c) Harvesting the cells and subsequent lysis

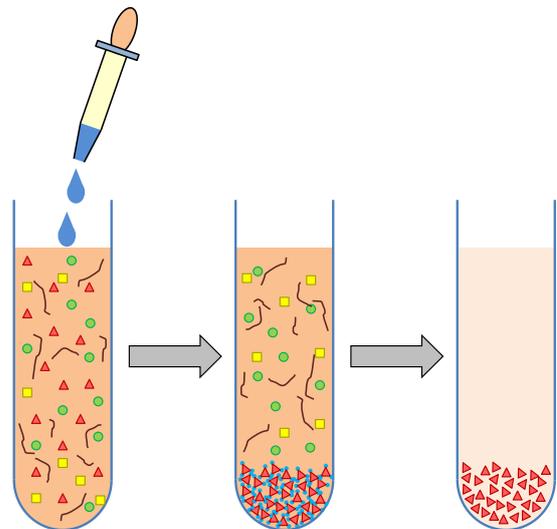
After certain periods of time (time points 5, 10, 15, 20, 25, 30, 45, 60, 90, and 120 minutes are chosen according to expert knowledge) the cells of each plate are washed to remove the stimulating agent and put on ice to stop ongoing interactions. Cell lysates are then centrifuged for separating the cytoplasmic fraction from the more heavy-weight cell parts like the nuclei, and the supernatant is collected.



3.2.5.2 Step 2: Immunoprecipitation and gel electrophoresis

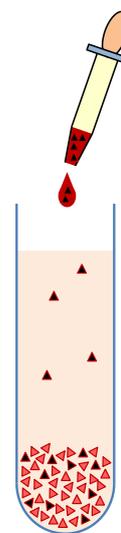
2a) Immunoprecipitation to concentrate proteins of interest

To separate the proteins of interest from other substances (purification), specific antibodies are added to the supernatant. The resulting antibody-antigen complexes bind to admixed insoluble beads of agarose or sepharose in the micrometer range and can be extracted by centrifugation (purification by immunoprecipitation). Further treatment removes the beads, short term boiling results in protein denaturation.



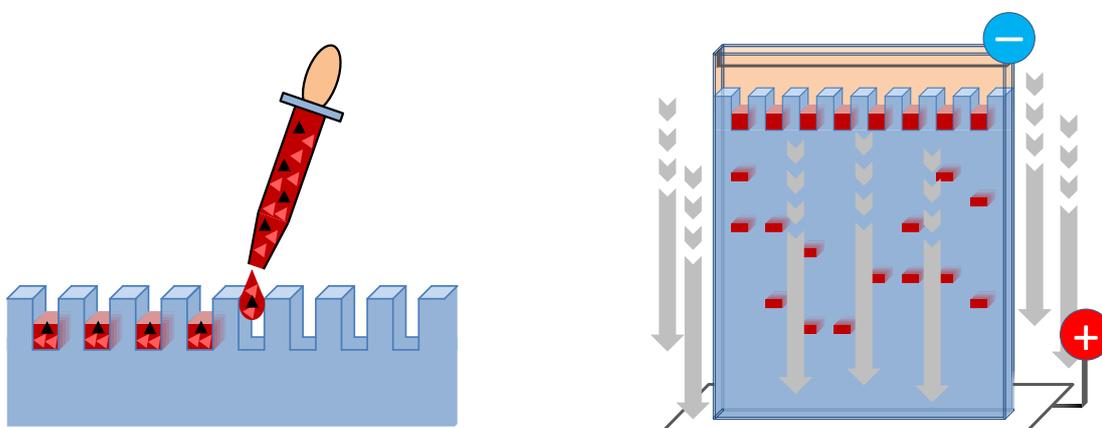
2b) Adding recombinant proteins for calibration (optional)

Further, specially labelled (e.g. GST-marked) recombinant proteins may be added to serve as an internal standard. The idea behind the addition of a recombinant protein is to make the processing error trackable. The recombinant proteins undergo all subsequent processes identically as the proteins of interest. Since the concentration of the recombinant proteins is known, their fluctuations can be used for correcting processing errors. See section 3.5.2 for possible issues in this approach.



2c) Loading of gels with identical amount of proteins, electrophoresis

The concentrated proteins are loaded onto different lanes of an SDS-polyacrylamide electrophoresis gel (SDS-PAGE⁽¹³⁾) in an arbitrary order. The random placement breaks the error correlation between neighbouring time points, such that errors in the gel do not transfer in a correlated way to the signal (figure 3.7 on page 54).



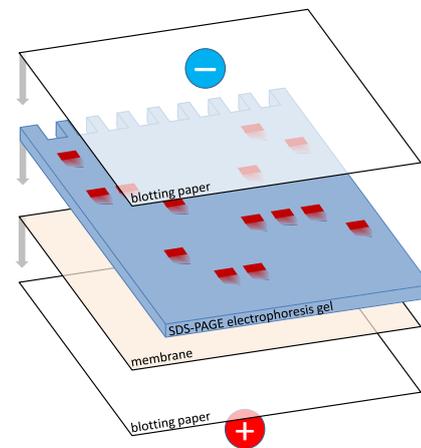
The gel is placed into a cooling bath, and an electrical current is applied for a couple of hours (figure 3.6 on page 53). The negatively charged SDS linearizes and envelopes the denatured proteins, allowing them to move through the gel, finally resulting in a separation by molecule weight, since larger molecules move slower as they are retained by the gel. Different electrical charges of the individual proteins are balanced by the highly negatively charged SDS.

⁽¹³⁾Sodium Dodecyl Sulfate PolyAcrylamide Gel Electrophoresis

3.2.5.3 Step 3: Immunoblotting, immunolabelling and evaluation

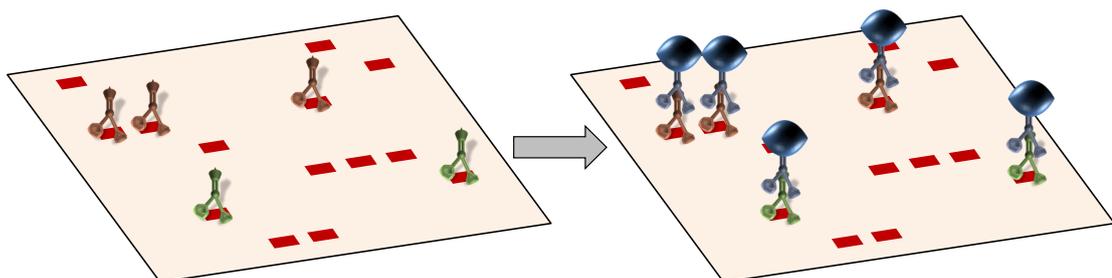
3a) Immunoblotting to membrane

To extract the proteins, the gel is positioned upon a nitrocellulose membrane and a second, now perpendicular electrical current is applied, in which the proteins move directed to the anode. After this process, called *blotting*, the membrane surface contains a flat image of the former protein distribution in the gel.



3b) Application of primary and secondary antibodies (immunolabelling)

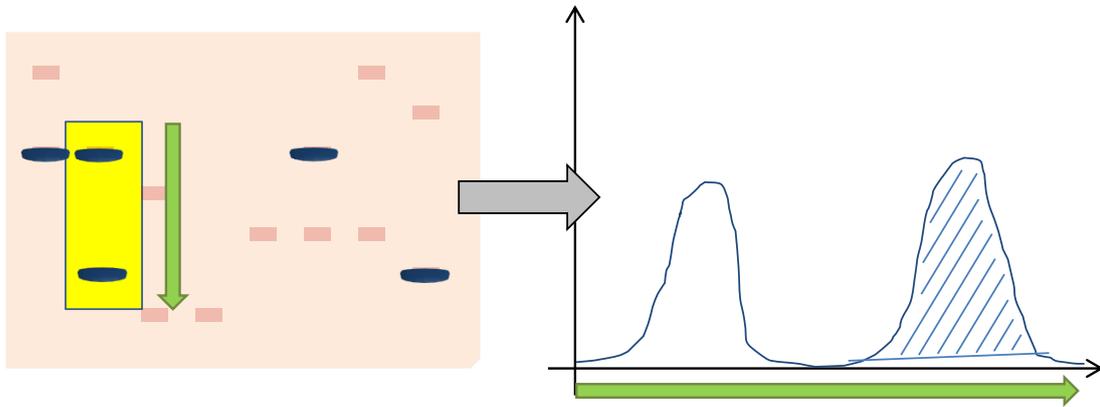
The membrane is incubated with specific primary antibodies that exclusively bind to their target proteins. By washing the membrane, excess unbound antibodies are removed. A secondary antibody, which binds exclusively to the primary one, is added for detection of the target protein. The secondary antibody can be labelled by radioactive substances or chemiluminescent markers, as in the experiments in our case. The emitted light can be detected by photographic films or (semi-)automated machines like LumiImager⁽¹⁴⁾.



⁽¹⁴⁾Roche Applied Science, Mannheim

3c) Manual or automated evaluation (e.g. LumiImager, ImageJ)

Light intensity is (over a certain range) proportional to protein amount, but the proportionality constant depends on gel, antibodies, and overall experimental procedure, and is therefore unknown. Thus, data originating from different gels cannot be compared without further processing, see section 3.5 for details.



To detect a second, different protein that is also present on the same gel, e.g. the GST-labelled recombinant proteins manually added before immunoprecipitation, the antibodies can be removed from membrane by incubation with a *stripping buffer solution* (consisting, e.g., of Tris-HCl, SDS and β -mercaptoethanol, the latter detaches the antibodies). After stripping, the membrane can be reprobed with different primary and secondary antibodies.

Further processing of measurement data is described in section 3.5.

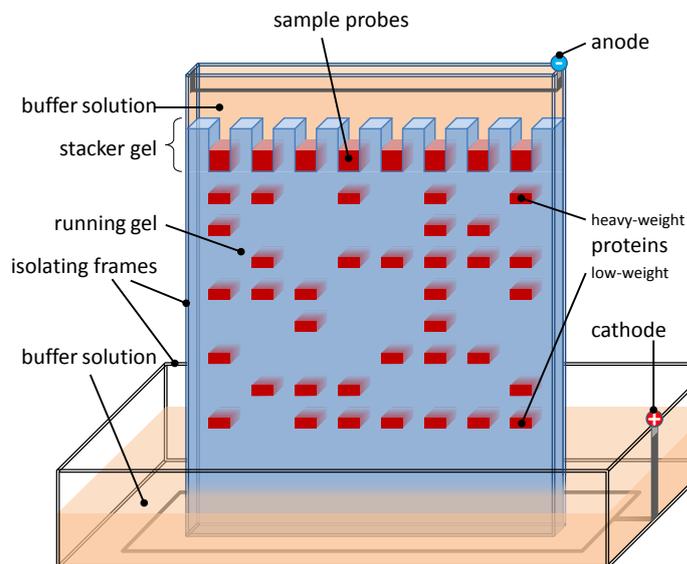


Figure 3.6: Sketch of gel electrophoresis. The running gel (blue) is framed between two isolating plastic planes. Sample probes are put into the pockets of the stacker gel atop the running gel. The whole frame is put into a buffer solution, and an electrical current is applied between anode and cathode. The buffer solution ensures linearization and electrical neutrality of the sample proteins, which then can move through the gel pores in the electrical field and are thus separated according to their molecular size. For constant temperatures, the buffer solution may be externally cooled.

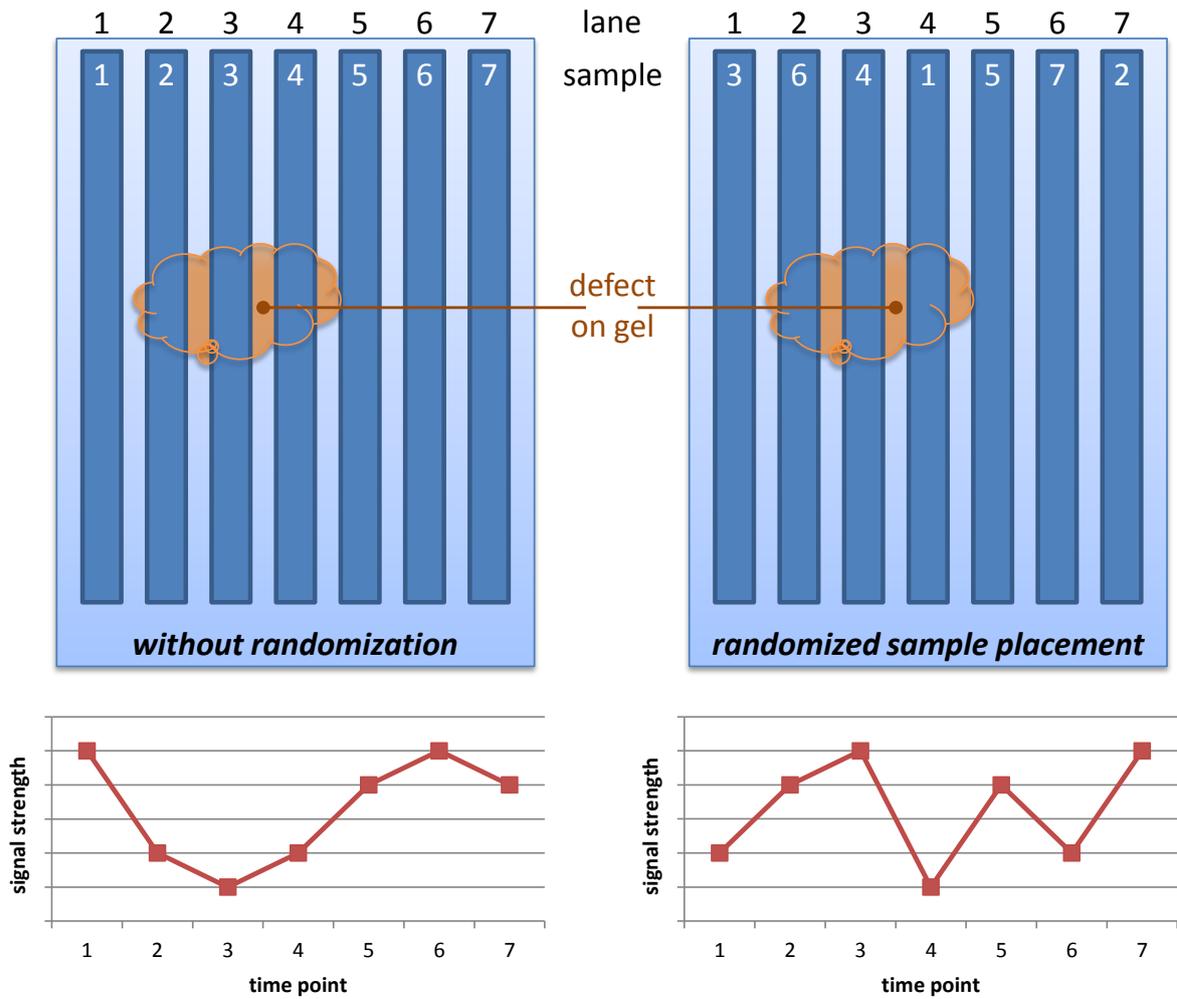


Figure 3.7: Randomization breaks time correlation in errors. Without randomized sample placement, gel defects (e.g. entrapped air) may induce a time correlated error, mimicking an actual signal (left). Randomization breaks error correlation. (*illustrative graphic*)

Table 3.2: Abbreviations of experimental settings.

<i>Experimental Setting</i>	<i>Abbreviation</i>
(*) Stimulation with IL-6	<i>IL6+</i>
(*) Stimulation with IL-6 while blocking GM-CSF	<i>IL6+/GMCSF-</i>
(**) Stimulation with both IL-6 and GM-CSF	<i>IL6+/GMCSF+</i>
Unstimulated	<i>unstim</i>

(*) stimulation setting chosen by expert biologists

(**) co-stimulation setting proposed by the author

3.3 Experimental settings and raw measurement data

The proliferation experiments described in section 3.2.4 indicate an intracellular connection between IL-6 and GM-CSF signalling pathways. In the first proliferation assay, the two IL-6 stimulation experiments draw special attention to themselves.

On the one hand, HaCaT keratinocytes have been stimulated with IL-6 solely; on the other hand, stimulated with IL-6 while blocking GM-CSF with appropriate antibodies at the same time. All other experimental parameters have been kept fixed. The observation that the proliferation rate of the solely IL-6 stimulated cells was significantly higher than the unstimulated control is not surprising, as IL-6 is known as a proliferative factor for many years.

Cells that have been stimulated with IL-6 with simultaneous blockade of GM-CSF, however, still show an increased growth compared to control, but, interestingly, reveal a significantly lower proliferation compared to sole IL-6 stimulation.

This strongly suggests a *Crosstalk* of the involved intracellular signalling pathways.

To systematically explore the interactions of IL-6 and GM-CSF signalling pathways in HaCaT keratinocytes, the JAK-STAT pathway is a good candidate as it is shared by the signalling of both cytokines. Especially the IL-6-induced STAT-3 activation has been shown to be positively correlated to proliferation of HaCaT keratinocytes, and an inhibition of this signalling molecule also inhibits cell proliferation [Nici2014].

As the proliferation assay shows the pro-proliferative effects of GM-CSF on HaCaT cells, the author suggested to include a fourth experimental setting, in which the keratinocytes are stimulated with both IL-6 and GM-CSF.

Simplifying the referencing, we introduce the following abbreviations for the four investigated experimental settings:

- *IL6+* for sole IL-6 stimulation,
- *IL6+/GMCSF-* for IL-6 stimulation with simultaneous blockade of GM-CSF
- *IL6+/GMCSF+* for co-stimulation with both factors IL-6 and GM-CSF
- *unstim* for non-stimulated control experiments.

Table 3.2 summarizes these scenarios.

3.3.1 Description of experimental procedure

Due to laboratory restrictions, no more than 11 samples can be blotted on the same gel. The time points have been chosen by expert knowledge, close-meshed at the first half of the observed time span of two hours. Samples are probed after 0, 5, 10, 15, 20, 25, 30, 45, 60, 90, 120 minutes.

For each time point, well plates with approximately 2 million cells are prepared and starved for 24 hours to cease down proliferation and to synchronize cell cycle.

At the beginning of each time series, the starvation medium is removed and – with exception of the unstimulated control sample for $t = 0min$ and the *unstim* setting – replaced with a stimulation medium containing either IL-6 (experiment *IL6+*), IL-6 and a GM-CSF blocking antibody (experiment *IL6+/GMCSF-*), or both IL-6 and GM-CSF (experiment *IL6+/GMCSF+*).

At the respective time points, the stimulation medium is removed, the reaction is stopped, and cells are lysed, centrifuged and the supernatant (containing the cytoplasmic fraction) is kept on ice for further processing.

After samples for all time points have been taken, STAT-3 proteins are purified and detected as described in section 3.2.5.

The four experiments are repeated in triplicates, yielding 131 values of raw data⁽¹⁵⁾.

3.3.2 Raw measurement data

Table 3.3 lists the raw data obtained by quantitative immunoblotting upon the four different experimental settings. The raw BLU data is illustrated in figure 3.8 on page 58.

Since the BLU measurement values are given in relative units that are incomparable between different experiments, they allow only a qualitative interpretation. In all three stimulation experiments (*IL6+*, *IL6+/GMCSF+*, *IL6+/GMCSF-*, figure 3.8, rows 1–3), a steep rise in cytoplasmic pSTAT-3 levels after stimulation can be observed, reaching a peak after 15–30 minutes, and subsequently monotonously declining to basal levels at around 90–120 minutes after stimulation.

In unstimulated cells (*unstim*, figure 3.8, last row), a stable level of pSTAT-3 is detected in the cytosol.

For further analysis, these raw measurements were processed as described in section 3.5.

⁽¹⁵⁾ The 120-minutes data point in the first *unstim* replicate has not been taken (see table 3.3d), thus there are $131 = 3 \cdot 11 \cdot 4 - 1$ data points.

Table 3.3: Time course data of pSTAT-3 upon four different stimulation settings. Measurement values are in relative units (BLU), thus neither comparable between experiments nor between replicates without further processing (see section 3.5). Times are given in minutes after stimulation. In 3.3d, rep #1, there was no sample taken at time point 120.
 mean: arithmetic mean value · sd: sample standard deviation (with BESSEL's correction to ensure unbiased sample variance) · cv: coefficient of variation

Time	rep #1	rep #2	rep #3
0	25736	28128	87794
5	243000	348000	1060000
10	759000	1030000	2910000
15	839000	1160000	2100000
20	1190000	875000	3080000
25	1370000	1110000	4190000
30	1370000	716000	3690000
45	805000	545000	2360000
60	464000	352000	1780000
90	89580	104000	416000
120	31718	76681	131000
mean	653367	576801	1982254
sd	518285	426229	1427361
cv	0.7933	0.7390	0.7201

(a) Time course of pSTAT-3 upon stimulation with IL-6
IL6+

Time	rep #1	rep #2	rep #3
0	9315	380000	5200000
5	66842	1210000	10500000
10	75450	2500000	6670000
15	112000	4770000	63700000
20	156000	7010000	26000000
25	176000	6520000	85300000
30	148000	4220000	33200000
45	71949	2000000	20300000
60	23389	1430000	13300000
90	31244	210000	8450000
120	14696	45387	5410000
mean	80444	2754126	25275455
sd	59782	2504477	26374448
cv	0.7432	0.9094	1.0435

(b) Time course of pSTAT-3 upon stimulation with IL-6 and GM-CSF
IL6+/GMCSF+

Time	rep #1	rep #2	rep #3
0	440000	9670000	218000
5	1840000	40900000	106000
10	3610000	46200000	1290000
15	4410000	57700000	2840000
20	4350000	40000000	1160000
25	5780000	49600000	2300000
30	4590000	54900000	4700000
45	4550000	54900000	1530000
60	2120000	48200000	345000
90	1630000	15800000	109000
120	488000	20000000	52693
mean	3073455	39806364	1331881
sd	1834203	16914776	1465458
cv	0.5968	0.4249	1.1003

(c) Time course of pSTAT-3 upon stimulation with IL-6 while blocking GM-CSF
IL6+/GMCSF-

Time	rep #1	rep #2	rep #3
0	165000	11400000	5130000
5	164000	10400000	4120000
10	137000	11300000	6910000
15	199000	13000000	5210000
20	183000	8670000	6830000
25	219000	14300000	7100000
30	194000	13300000	5110000
45	185000	10600000	8500000
60	216000	11100000	5590000
90	175000	10100000	5270000
120	–	10200000	4890000
mean	183700	11306364	5878182
sd	25002	1637268	1282746
cv	0.1361	0.1448	0.2182

(d) Time course of pSTAT-3 in unstimulated cells
unstim

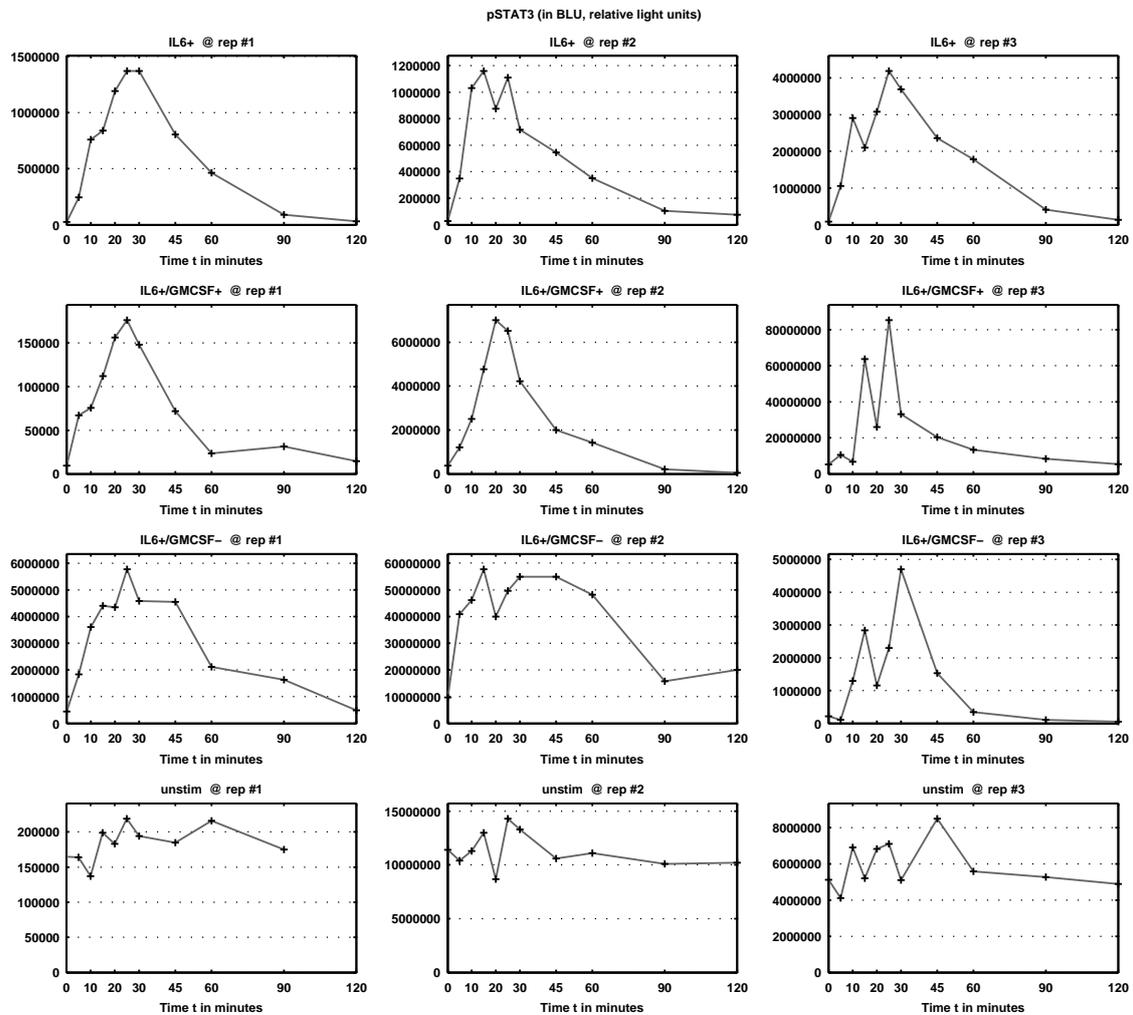


Figure 3.8: BLU data of cytosolic pSTAT-3 measurements in four experimental settings. Depicted here are the linearly interpolated raw BLU values of cytosolic pSTAT-3, which show similar qualitative behaviour in all three stimulation experiments (*IL6+*, *IL6+/GMCSF+*, and *IL6+/GMCSF-*), but cannot be compared quantitatively without further processing (see section 3.5). In the unstimulated control experiments (*unstim*, bottom row), the basal level of cytosolic pSTAT-3 shows only small fluctuations.

3.4 Theoretical preliminaries for immunoblot data analysis

3.4.1 Approximability of a ratio of normal variables by a normal distribution

MARSAGLIA, in his paper from 1965 [Marsaglia1965] and the follow-up paper 40 years later [Marsaglia2006], investigates the distribution of the *ratio of normal variables*, as stated in the paper titles. He develops the density function of such ratios, and gives theoretical and numerical evidence for when approximation of these ratios by a normal distribution is appropriate.

We carve out some parts that we need to analyse the data obtained in the experiments described in section 3.2.5.

The main idea is to transform a ratio of any jointly distributed normal variates W and Z , possibly correlated, into a standard form $(a + X)/(b + Y)$ with (non-negative) translations a, b and two independent standard normally distributed random variables X, Y .

3.1 Proposition (Transformation into standard form $(a + X)/(b + Y)$)

Let $Z \sim \mathcal{N}(\mu_Z, \sigma_Z^2)$, $W \sim \mathcal{N}(\mu_W, \sigma_W^2)$ be two jointly distributed normal variables, and let denote $\rho := \text{Cov}[Z, W]/(\sigma_Z \sigma_W)$ their correlation coefficient. Let further be $X, Y \sim \mathcal{N}(0, 1)$ two independent standard normally distributed random variables, and choose

$$r := \frac{\sigma_W}{k \sigma_Z \sqrt{1 - \rho^2}}, \quad s := \rho \frac{\sigma_Z}{\sigma_W}, \quad a := \frac{\mu_Z - \rho \frac{\mu_W}{\sigma_W}}{k \sqrt{1 - \rho^2}}, \quad b := \frac{\mu_W}{\sigma_W}, \quad k \in \{-1, 1\}$$

where k shall be chosen to guarantee $\text{sign}(a) = \text{sign}(b)$.

Then, it holds:

1. $Z - sW$ and W are independent normal variables
2. $\text{Var}[Z - sW] = \sigma_Z^2(1 - \rho^2)$
3. $\frac{Z}{W} \sim \frac{1}{r} \left(\frac{a + X}{b + Y} \right) + s$ and $r \left(\frac{Z}{W} - s \right) = r \frac{Z - sW}{W} \sim \frac{a + X}{b + Y}$

Remark: Since the distributions of $(a + X)/(b + Y)$ and $(-a + X)/(-b + Y)$ are the same, we can postulate the *non-negativity of translation constants a and b* .

Proof:

1. It holds: $\mathbb{E}[(Z - sW)W] = \mathbb{E}[ZW] - s\mathbb{E}[W^2]$

$$= \mathbb{E}[Z]\mathbb{E}[W] + \text{Cov}[Z, W] - \rho \frac{\sigma_Z}{\sigma_W} (\mathbb{E}[W]^2 + \text{Var}[W])$$

$$= \mu_Z \mu_W + \text{Cov}[Z, W] - \frac{\text{Cov}[Z, W]}{\sigma_W^2} (\mu_W^2 + \sigma_W^2)$$

$$= \mu_Z \mu_W - \text{Cov}[Z, W] \frac{\mu_W^2}{\sigma_W^2} = \mu_Z \mu_W - \rho \frac{\sigma_Z}{\sigma_W} \mu_W^2$$

$$= \mathbb{E}[Z]\mathbb{E}[W] - s\mathbb{E}[W]^2 = \mathbb{E}[Z - sW]\mathbb{E}[W]$$

2. Using the bilinearity of covariance and the definitions of ρ and s , we get

$$\begin{aligned} \text{Var}[Z - sW] &= \text{Var}[Z] + \text{Var}[-sW] + 2 \cdot \text{Cov}[Z, -sW] \\ &= \sigma_Z^2 + s^2 \sigma_W^2 - 2s \cdot \text{Cov}[Z, W] = \sigma_Z^2(1 - \rho^2). \end{aligned}$$

3. We prove the second statement, as the first immediately follows. The numerator of r is the standard deviation of W , and the denominator is the standard deviation of $Z - sW$ (except for the sign k). Thus, multiplying $(Z - sW)/W$ by r results in a random variable $(Z - sW)/(k\sigma_Z\sqrt{1 - \rho^2}) \sim \mathcal{N}(a, 1)$ in the numerator, as its mean is $\mathbb{E}[Z - sW]/(k\sigma_Z\sqrt{1 - \rho^2}) = (\mu_Z - \mu_W\rho\sigma_Z/\sigma_W)/(k\sigma_Z\sqrt{1 - \rho^2}) = a$. Further, we have a random variable $W/\sigma_W \sim \mathcal{N}(b, 1)$ in the denominator. Centering the random variables in numerator and denominator completes the proof. \square

In his 2006 paper, MARSAGLIA investigates for which values of a and b the standard form ratio $T := (a + X)/(b + Y)$ can be approximated by a normal distribution. The resulting distribution T does not possess any moments (as the respective integrals do not exist), nevertheless, for certain values of a and b , mean and standard deviation of an approximating normal distribution can be computed.

3.2 Lemma

Let a, b be non-negative real numbers, and $X, Y \sim \mathcal{N}(0, 1)$ two independent standard normally distributed random variables, and $T := (a + X)/(b + Y)$. Then, it holds:

1. The probability density of T is given as

$$f_T(t) = \frac{e^{-\frac{1}{2}(a^2+b^2)}}{\pi(1+t^2)} \left[1 + qe^{\frac{1}{2}q^2} \cdot \int_0^q e^{-\frac{1}{2}x^2} dx \right], \quad q := \frac{b+at}{\sqrt{1+t^2}}. \quad (3.1)$$

It is thus a convex combination

$$f_T(t) = kf_1(t) + (1-k)f_2(t), \quad \text{with } k := e^{-\frac{1}{2}(a^2+b^2)}$$

of the unimodal CAUCHY density f_1 and a bimodal density f_2

$$f_1(t) = \frac{1}{\pi(1+t^2)}, \quad f_2(t) = \frac{q \int_0^q e^{-\frac{1}{2}(x^2-q^2)} dx}{\pi(1+t^2) + (e^{\frac{1}{2}(a^2+b^2)} - 1)}.$$

2. For $a < 2.25$ and $b > 4$, the distribution of T can be well approximated by a normal distribution with mean value μ and variance σ^2 given as

$$\begin{aligned} \mu &= \frac{a}{(1.01b - 0.2713)} \\ \sigma^2 &= (a^2 + 1)/(b^2 + 0.108b - 3.795) - \mu^2 \end{aligned} \quad (3.2)$$

Proof (sketch):

1. Expressions for the cumulative distribution function $F(t) = P(T < t)$ can be derived using the bivariate normal distribution function $L(h, k, \rho) = P[\xi \geq h, \eta \geq k]$ with standard normal variables ξ and η with $\text{Cov}[\xi, \eta] = \rho$ and connective properties of L and NICHOLSON'S V function $V(h, q) = \int_0^h \int_0^{qx/h} \varphi(x)\varphi(y) dy dx$, with φ denoting the density of the normal distribution. Having the CDF $F(t)$ of T , the density function $f_T(t)$ is found by differentiation.

2. The density $f_T(t)$ of T is a mixture of the unimodal CAUCHY density $f_1(t) = (\pi(1 + t^2))^{-1}$ that is independent of a and b , and a second, bimodal density $f_2(t)$. For values of $a \leq 1$, the resulting density is always unimodal. When the asymptotic value $\tilde{a} \approx 2.256058904$ is exceeded, the resulting density is always bimodal. For values $a \in (1, \tilde{a})$, the shape of the density depends on b . There is a separating curve in the a - b -plane that is implicitly given by $\frac{d^2 f(t)}{dt^2} = \frac{df(t)}{dt} = 0$. Points (a, b) lying left of this curve result in a unimodal, points lying right of it in a bimodal density, though in many bimodal cases, the second mode can be considered negligible for many practical settings. Figure 3.9 displays an approximation on this separating curve.

Conditioning the denominator of $T = (a + X)/(b + Y)$ by $Y > -4$ (this holds with a probability greater than 99.9968%) ensures that all (conditioned) moments $\mathbb{E}[T^i | Y > -4] = \mathbb{E}[(a + X)^i | Y > -4] \cdot \mathbb{E}[1/(b + Y)^i | Y > -4]$ (X, Y independent) exist.

Using $\mathbb{E}[A|B] = \mathbb{E}[\mathbb{1}_B A] / P(B)$, the conditional expectation (definition and lemma 4.13) in the denominator can be computed as

$$\mathbb{E}[1/(b + Y)^i | Y > -4] = \frac{\int_{-4}^{\infty} \varphi(y)/(b + y)^i dy}{\int_{-4}^{\infty} \varphi(y) dy}.$$

The formulas are derived by fitting curves to point evaluations of the first two conditioned moments $\mathbb{E}[1/(b + Y)^i | Y > -4]$ ($i = 1, 2$) at certain values of b .

For a full proof, we refer to [Marsaglia1965] and [Marsaglia2006], where also a numerically obtained approximation of the uni-/bimodal separating curve is given. \square

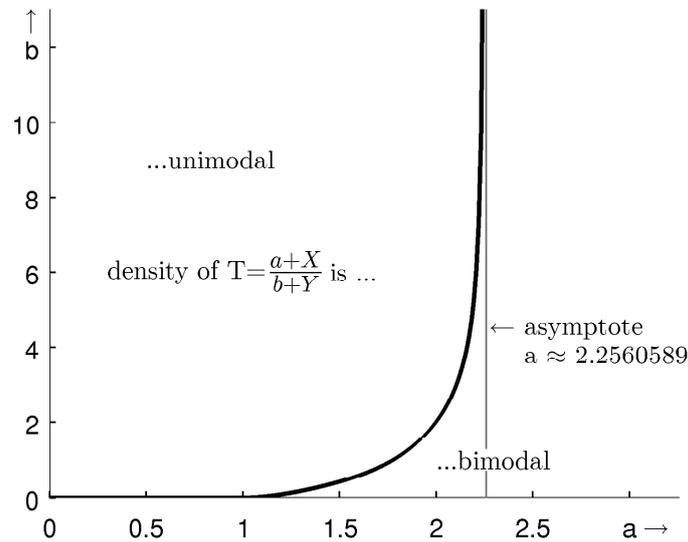


Figure 3.9: Curve separating between uni- and bimodal density of $T = (a + X)/(b + Y)$. For coordinates (a, b) to the left of the curve, the ratio distribution density of T is unimodal; for coordinates lying right of the curve, it is bimodal. The separating curve is characterized by $f''(t) = f'(t) = 0$. This figure adapted from [Marsaglia2006] shows the numerical approximation for the separating curve given therein: $b(a) = \frac{18.621 - 63.411a + 84.041a^2 - 54.668a^3 + 17.716a^4 - 2.2986a^5}{2.256058904 - a}$. For picture credits see page 240.

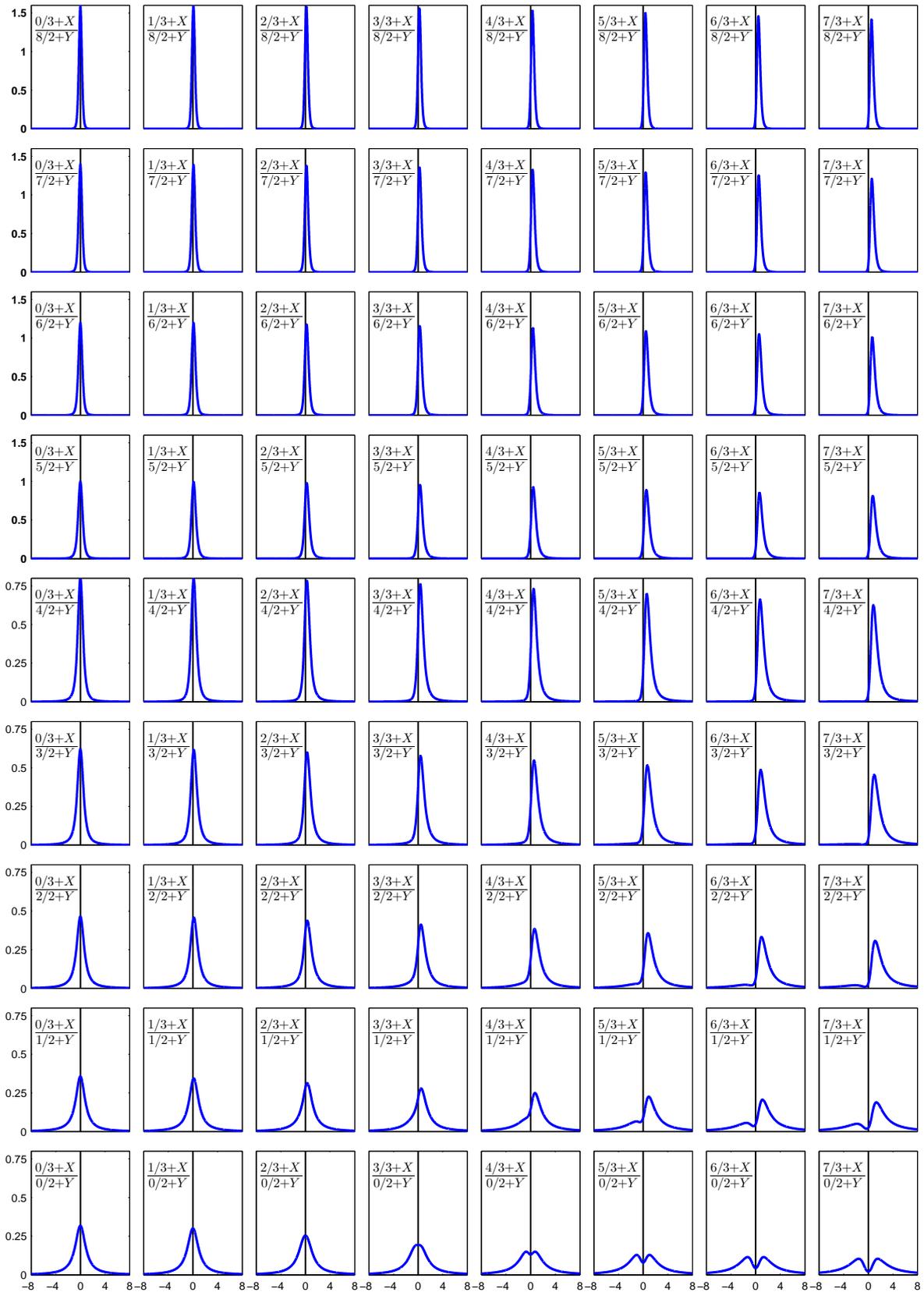


Figure 3.10: Densities of $T = (a + X)/(b + Y)$ for distinct values of a and b . X and Y are independent, standard normal variates. The density function of T is given in eq. (3.1). Values of a cover the range of 0 to the asymptotic value ≈ 2.25 , where unimodal distributions are possible. Values of b were chosen to sample the interval $[0, 4]$, with 4 being the lower bound for approximability of the ratio T by a normal distribution. The depicted densities represent the possible shapes of the density function of the ratio T . Figure adapted from [Marsaglia1965], see page 240 for picture credits.

3.5 Data processing

3.5.1 Semiquantitative immunoblotting

Since the development of the DNA blots by SOUTHERN more than 40 years ago, blotting methods have developed from qualitative to quantitative methods. Chemiluminescence detection by charge-coupled devices strongly increased sensitivity and linear range compared to x-ray films [Martin1994]. However, the resulting *light units*⁽¹⁶⁾ are still a relative measure, that depends on parameters like the electrophoresis gel, used antibodies, etc., and cannot directly be related to absolute concentrations. As a consequence, data from different blots cannot be compared without further processing⁽¹⁷⁾.

The most obvious remedy is to add a dilution series of the protein of interest (or a recombinant analogon) on the gel. However, space in gels is restricted, sometimes to only 7 or 11 lanes, rendering this strategy inappropriate for time course studies.

In-place signal normalization and calibration techniques have been proposed [Schilling2005; Schilling2005a], where housekeeping or manually added (recombinant) proteins are used for signal normalization and correction. The idea is, that the added proteins undergo the same processing in gel electrophoresis, blotting and detection such that their signal will reflect the processing errors as their concentration is assumed constant. However, for these types of **normalization techniques**, a very stable reference normalizer or calibrator signal is required, and they may introduce strong artificial artefacts as we work out in section 3.5.3. Moreover, they tend to change the underlying error model in an disadvantageous way.

We propose to use a quite simplistic technique called *amplification factors method* (section 3.5.4) that calculates the x-fold increase in signal strength compared to a single(!) reference point. Under only weak requirements, which are often fulfilled and – important for wet lab experiments – easy to verify, this method **delivers stable and reproducible results**.

3.5.2 Analysis of WESTERN blotting data

Let n_M be the number of measurement time points, and n_{rep} the number of experiment replicates⁽¹⁸⁾. For the analysis, we assume that the cellular concentrations $\hat{\nu}_i$ only depend on time t_i ($i = 1, \dots, n_M$) and are equal for all replicates j ($j = 1, \dots, n_{rep}$), i.e. that the cells always behave identically upon stimulation and observed fluctuations are only due to measurement errors⁽¹⁹⁾.

Let $\eta_i^{(j)} = \hat{\eta}_i^{(j)} + \varepsilon_i^{(j)}$ be the measurement value in relative light units, taken by quantitative immunoblotting, corresponding to time t_i at the j -th replicate. Here, $\hat{\eta}_i^{(j)} = c^{(j)}\hat{\nu}_i$ denotes the hypothetical true (but inaccessible) measurement value at time point t_i , and $c^{(j)}$ is the unknown blot/antibody-dependent proportionality constant for the true concentration $\hat{\nu}_i$.

⁽¹⁶⁾Often called BLU, *Boehringer Light Units*, named after a provider of commercial solutions. Nowadays, the abbreviation is more frequently read as *Bioluminescence Light Unit*, or the acronym RLU (*Relative Light Unit*) is used.

⁽¹⁷⁾The Isogen Life Science company, a Dutch supplier of bio labs, warns in its documentation: “Compare within the same blot. Due to transfer and handling differences, only compare proteins on the same blot and not between blots.” Isogen Life Science: *Important Factors in Performing Quantitative Western Blots*. Nov. 2013. URL: <http://www.isogen-lifescience.com/blotting/western-blotting-documentation/important-factors-in-performing-quantitative-western-blots> (visited on 08/08/2016)

⁽¹⁸⁾A replicate is an identical experiment blotted using a different gel.

⁽¹⁹⁾While the assumption of deterministic behaviour is questionable for a single cell, it is justified for the population mean that is observed in the described experiments.

The error in the measurements is modelled as an additive⁽²⁰⁾ normally distributed random variable $\varepsilon_i^{(j)} \sim \mathcal{N}(0, (\sigma_i^{(j)})^2)$. In the following two sections 3.5.3 and 3.5.4, we analyse and discuss two techniques for processing the raw measurement data.

3.5.3 Normalization using house-keeping or recombinant proteins (calibrators)

House-keeping proteins are believed to be expressed at a constant level and are therefore often used for normalization and error correction, as they undergo identical experimental processing⁽²¹⁾.

The usage of antibodies with a weaker affinity to the protein of interest might result in a high background signal, due to cross reactions to the epitopes of other proteins. The same problems occur if no specific antibody is available, or when using generic phosphotyrosine antibodies. These are cases, where proteins cannot be quantified by immunoblotting directly, thus requiring a purification of the protein of interest by immunoprecipitation in advance (see section 3.2.5.2).

Since *house-keeping proteins* like GAPDH⁽²²⁾ or β -actin⁽²³⁾ are lost during this purification step, *recombinant proteins* are manually added prior to immunoprecipitation and later used as normalizers.

A *recombinant protein* is an artificially produced variant of the target protein, with a slightly different molecular weight but sharing the same epitopes. Thus they (are believed to) behave identically and their error will reflect the processing error. As their concentration is known, their deviation from a constant signal can be used to correct the signal of interest. Normalization strategies using recombinant proteins are described e.g. in [Schilling2005; Schilling2005a].

As we will see, however, this method may also lead to spurious outcomes, if the calibrator signal is too noisy ([Schilling2005a] also discusses this case). Further, it changes the statistics of the data as it involves the calculation of a ratio of two random variables.

3.5.3.1 Normalization affects error distribution

We start with a theoretical analysis by having a look at the distribution of data normalized by this method. This applies to both using house-keeping or recombinant proteins.

Again, we write $\eta_i^{(j)}$ for the i -th measurement on blot/membrane j of the signal of interest, and denote by $\eta_{i,\text{norm}}^{(j)}$ the respective measurement value of the normalizer in the same lane.

A signal measurement $\eta_i^{(j)}$ is thus normalized as

$$\tilde{\eta}_i^j := \frac{\eta_i^{(j)}}{\eta_{i,\text{norm}}^{(j)}} = \frac{\hat{\eta}_i^{(j)} + \varepsilon_i^{(j)}}{\hat{\eta}_{i,\text{norm}}^{(j)} + \varepsilon_{i,\text{norm}}^{(j)}}, \quad (3.3)$$

i.e. the normalized measurements $\tilde{\eta}_i^j$ are indeed ratios of normal variates.

⁽²⁰⁾By allowing the variances $\sigma_i^{(j)}$ to depend on the time index i , also normally distributed multiplicative noise is covered with this formulation, since $\eta = a\hat{\eta}$, $a \sim \mathcal{N}(1, \sigma^2)$ can be written as $\eta = a\hat{\eta} = (1+\varepsilon)\hat{\eta} = \hat{\eta} + \varepsilon\hat{\eta} = \hat{\eta} + \tilde{\varepsilon}$ with $\tilde{\varepsilon} \sim \mathcal{N}(0, \hat{\eta}^2\sigma^2)$

⁽²¹⁾This assumption is only partly valid: The processing differs in general at least at the detection step, as different primary antibodies are used

⁽²²⁾Glyceraldehyde 3-phosphate dehydrogenase, an enzyme facilitating the break down of glucose in glycolysis

⁽²³⁾Multifunctional protein forming microfilaments

In section 3.4.1, we gave criteria to check whether these ratios can be reasonably approximated by a normal distribution. Since the measurements of the signal $\eta_i^{(j)}$ differ in time and blot, it is impossible to estimate their moments unless one is willing to put several probes of the same time point t_i on the same blot j , that would rather not be practically performed⁽²⁴⁾. Due to this, a direct examination of the ratio in eq. (3.3) cannot be done in general.

For a negative result concerning MARSAGLIA's findings about approximability by a normal distribution, it is sufficient to have information about the normalizer. As its concentration is assumed constant (house-keeping or manually added), their signal is expected to be constant per blot/membrane, and we can use the n_M measurements $\eta_{i,\text{norm}}^{(j)}$ ($i = 1, \dots, n_M$) of the normalizer signal and approximate the true normalizer signal on blot/membrane j by their mean value:

$$\hat{\eta}_{\text{norm}}^{(j)} \approx \eta_{\text{norm}}^{(j)} := \frac{1}{n_M} \sum_{i=1}^{n_M} \eta_{i,\text{norm}}^{(j)} \quad \text{approximation of true normalizer signal}$$

and we might further get estimates of their variance and standard deviation.

For analysing the statistics of normalized measurements, we use the notation of section 3.4.1, keeping both time-point-index i and blot-index j fixed, and set:

$$\begin{aligned} Z &:= \eta_i^{(j)} = \hat{\eta}_i^{(j)} + \varepsilon_i^{(j)} && \text{signal with error} \\ W &:= \eta_{i,\text{norm}}^{(j)} = \hat{\eta}_{i,\text{norm}}^{(j)} + \varepsilon_{i,\text{norm}}^{(j)} && \text{normalizer with error} \\ \mu_Z &&& \text{mean of } \eta_i^{(j)}, \text{ not accessible} \\ \sigma_Z &&& \text{standard deviation of } \eta_i^{(j)}, \text{ not accessible} \\ \mu_W &:= \hat{\eta}_{\text{norm}}^{(j)} && \text{mean value of normalizer} \\ \sigma_W &:= \sigma_{\text{norm}}^{(j)} && \text{standard deviation of normalizer} \end{aligned}$$

such that $\tilde{\eta}_i^j = \frac{\eta_i^{(j)}}{\eta_{i,\text{norm}}^{(j)}} = \frac{Z}{W}$.

Since signal of interest and normalizer are assumed uncorrelated in our case, the equations for the constants a and b in proposition 3.1 reduce to

$$a := \frac{\mu_Z}{\sigma_Z} = \frac{1}{CV_Z} \quad \text{and} \quad b := \frac{\mu_W}{\sigma_W} = \frac{1}{CV_W},$$

where CV_Z and CV_W again denote the coefficients of variation of the signal Z and the normalizer W respectively. From the conditions given by lemma 3.2 for approximability by a normal distribution,

$$a < 2.25 \quad \text{and} \quad b > 4$$

we get the following condition on the coefficient of variation of the normalizer:

$$4 < b = \frac{1}{CV_W} \implies CV_W < \frac{1}{4}, \quad (3.4)$$

thus requiring a ‘‘reasonably stable’’ normalizer signal.

⁽²⁴⁾This would require the availability of large (and therefore expensive) electrophoresis gels. Even if these are available, one would rather use the extra space for a finer time resolution or simultaneous evaluation of multiple replicates instead of resampling probes of the same lysate on several lanes of the gel. Also, in cases where the amount of protein to be quantified is very limited, there might be just enough protein available for a single probe.

Table 3.4: Measurements of recombinant STAT-3 have been performed in six experiments (see table A.1). With the exception of #7, none of the measurement series fulfills the sufficient condition for approximability by a normal distribution given in eq. (3.4).

Data origin: A, B, C: *IL6+* · D, E: *IL6+/GMCSF-* · F: *IL6+/GMCSF+* · G: *unstim*

mean: arithmetic mean value · sd: sample standard deviation (with BESSEL's correction to ensure unbiased sample variance) · cv: coefficient of variation

time	Measurements of recombinant STAT-3						
	A	B	C	D	E	F	G
0	16278	68696	210000	438000	593000	331000	571000
5	52645	33238	455000	252000	1260000	430000	332000
10	58513	39359	431000	380000	879000	244000	495000
15	45761	64062	388000	288000	402000	207000	487000
20	29875	41054	234000	436000	1000000	390000	562000
25	34018	56970	292000	277000	908000	479000	436000
30	26978	17466	241000	164000	179000	438000	486000
45	20286	78551	177000	403000	700000	379000	764000
60	30328	40846	257000	289000	1580000	190000	517000
90	18766	68145	270000	359000	533000	424000	524000
120	13963	86380	173000	108000	871000	308000	517400
mean	31583	54070	284364	308545	809545	347273	517400
sd	14960	21209	98106	107557	393716	99010	104465
cv	0.4737	0.3922	0.3450	0.3486	0.4863	0.2851	0.2019

On the other hand, if the condition $b > 4$ is not fulfilled, from figure 3.10, one may conclude by visual inspection of the densities of the ratio $T = Z/W$ that small values of a might still lead to approximability by a normal distribution. However, requesting small values of $a = CV_Z^{-1}$ is equivalent to requiring a large coefficient of variation CV_Z of the signal itself – a condition that is most objectionable. A further theoretical investigation in the case $b < 4$ about approximability of T by a normal distribution is therefore dispensable for our purposes.

Table 3.4 lists the measurements of recombinant STAT-3 protein that was manually added in fixed and known concentration to serve as normalizer. With coefficients of variations of 0.47, 0.39, 0.35, 0.35, 0.49, 0.29, and 0.20, all but the last one clearly exceed the limit of $CV_W < 1/4$.

Moreover, since the distribution of the normalized data $T = Z/W$ is quite awkward and may even be bimodal, it is hard to derive an objective function that ensures a maximum likelihood estimation of the unknowns.

3.5.3.2 Normalization may lead to spurious signals

While the mentioned results of non-approximability by a normal distribution are of rather theoretical nature, figure 3.11 shows the possible practical distortive effects of normalization, as they were observed while using recombinant STAT-3 proteins as normalizer.

The first column of figure 3.11 depicts the phospho-STAT-3 measurements upon different IL-6 stimulation settings. The second columns shows the normalized signal, where the recombinant proteins were used to “correct” the phospho-STAT-3 measurements. In both columns, the measurements are given relative to time point 0 (x-fold amplification), and the axes were chosen to be equal. The signal of the normalizer (recombinant STAT-3 protein, in BLU) is

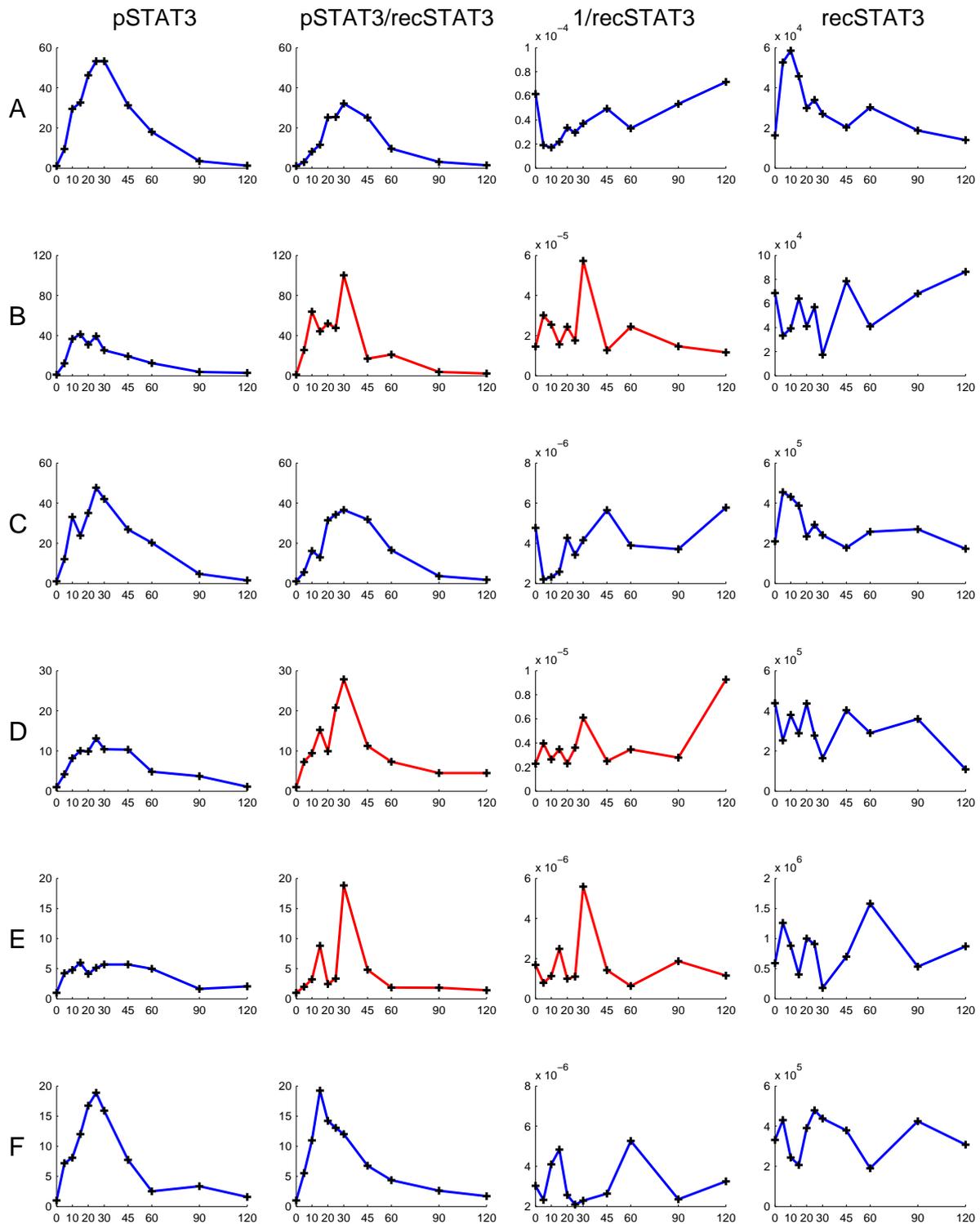


Figure 3.11: Possible destructive effects of normalization. Linearly interpolated measurements of phospho-STAT-3 signal (1st column), the normalized signal (2nd column) the reciprocal of the normalizer (3rd column) and the normalizer itself (4th column). For the first and second column, amplification factors relative to time point 0 are depicted (column 2: after normalization). Units in columns 3 and 4 are BLU^{-1} and BLU , respectively.

In red: The normalized signal is clearly shaped by the normalizer. Further, the normalization by recombinant STAT-3 might lead to tremendous differences in maximum amplification factors (peak levels in columns 1 and 2). Data: A, B, C: $IL6+/+$ · D, E: $IL6+/GMCSF-/-$ · F: $IL6+/GMCSF+/+$

given in the last column. Ideally, a straight line of a constant value would be obtained, as the recombinant protein is added in identical amounts to each probe. The third column is the reciprocal of the normalizer (in BLU^{-1}).

The red coloured graphs document the **spurious signals introduced by this normalization technique**. The phospho-STAT-3 signals in **B**, **D**, and **E** are heavily distorted by normalization, and are clearly shaped by the normalizer signal. In **E**, the shape of the normalized signal is virtually identical to the reciprocal of the normalizer, i.e. **we would get highly similar kinetics even if we totally neglect the phospho-STAT-3 measurements!**

Moreover, as one sees by comparing first to second column, the normalization drastically affects the maximum amplification factors. The settings **A**, **B**, and **C** are the three replicates of sole IL-6 stimulation, setting *IL6+*, and all three replicates show a maximum amplification factor between 41 and 53-fold, whereas the maximum amplification factors of their normalized counterparts vary between 32 and 100-fold.

Only in 2 out of these 6 experiments, in **C** and in **F**, the normalization did neither destroy the signal's shape nor had severe distortive effects on the signal levels.

Due to these disappointing results, the recombinant STAT-3 proteins were not further used in the remaining experiments.

3.5.3.3 Leaving the linear range: saturation effects

What has not been discussed so far is the problem of saturation when using house-keeping proteins as normalizers. Often, the target proteins are available in low concentration only. In this case, the experimental layout is constructed in a way to have these low concentrated proteins lying in the linear range of detection.

However, as most house-keeping proteins are quite abundant, their signal might leave the linear range, as saturation effects occur. While it might be, in theory, possible to cope with this by including MICHAELIS-MENTEN-like kinetics in the normalization procedure, this is practically not accomplishable.

3.3 Conclusion

Normalization using house-keeping or recombinant proteins changes the error statistics and may introduce spurious signals. Furthermore, saturation effects may result in leaving the linear range of detection.

3.5.4 Amplification factors method

We write

$$\hat{m}_i = \frac{\hat{\eta}_i^{(j)}}{\hat{\eta}_0^{(j)}} = \frac{\hat{\nu}_i}{\hat{\nu}_0} \quad (\forall j)$$

for the unknown true amplification factors, which we approximate on a per-blot basis from measurement data. We mark a reference measurement by an index of 0 (e.g. corresponding to the measurement at the initial time point t_0), and the approximation of \hat{m}_i on the j -th blot is given by

$$m_i^{(j)} = \frac{\eta_i^{(j)}}{\eta_0^{(j)}} = \frac{\hat{\eta}_i^{(j)} + \varepsilon_i^{(j)}}{\hat{\eta}_0^{(j)} + \varepsilon_0^{(j)}} = \hat{m}_i \cdot \frac{\hat{\eta}_0^{(j)}}{\hat{\eta}_0^{(j)} + \varepsilon_0^{(j)}} + \frac{\varepsilon_i^{(j)}}{\hat{\eta}_0^{(j)} + \varepsilon_0^{(j)}}, \quad (3.5)$$

where we have a *multiplicative bias* ϵ_m and an *additive error* ϵ_a

$$\epsilon_m(j) := \hat{\eta}_0^{(j)} / (\hat{\eta}_0^{(j)} + \epsilon_0^{(j)}) \quad (3.6)$$

$$\epsilon_a(i, j) := \epsilon_i^{(j)} / (\hat{\eta}_0^{(j)} + \epsilon_0^{(j)}) \quad (3.7)$$

The multiplicative bias ϵ_m is a per-blot property, and only depends on the measurements of the reference point, but not on other time points, whereas the additive error ϵ_a changes by blot and time point. Both quantities are ratios of normally distributed variables.

We first analyse the **additive error** ϵ_a , using the measurements of cytosolic pSTAT-3 in unstimulated cells (table 3.3d), as these don't manifest any biological dynamic. For data generated from blot j for a time point i , using the same notation as in section 3.4.1, and assuming that correlation of measurement errors has been avoided by randomization (section 3.2.5.2), we have:

$$\frac{Z}{W} := \frac{\epsilon_i^{(j)}}{\hat{\eta}_0^{(j)} + \epsilon_0^{(j)}}, \quad \mu_Z = 0, \quad \mu_W = \hat{\eta}_0^{(j)}, \quad \sigma_Z = \sigma_i^{(j)}, \quad \sigma_W = \sigma_0^{(j)}, \quad \rho = 0. \quad (3.8)$$

The exact values of μ_W , σ_W , and σ_Z are not available; the best we can do is to estimate them using the arithmetic mean and sample standard deviation (with BESSEL's correction applied to ensure unbiased sample variance). Omitting the index j denoting the replicate, we define for a fixed blot j :

$$\bar{\mu}_W := \frac{1}{n_M} \sum_{i=1}^{n_M} \eta_i, \quad \bar{\sigma}_W := \sqrt{\frac{1}{(n_M - 1)} \sum_{i=1}^{n_M} (\eta_i - \mu_W)^2}, \quad \bar{\sigma}_Z := \sqrt{\frac{1}{(n_M - 1)} \sum_{i=1}^{n_M} \eta_i^2} \quad (3.9)$$

approximating $\mu_W \approx \bar{\mu}_W$, $\sigma_W \approx \bar{\sigma}_W$, $\sigma_Z \approx \bar{\sigma}_Z$.

Using the unstimulated pSTAT-3 data, we further get:

$$\begin{aligned} r = r_i^j &:= \frac{\sigma_W}{\sigma_Z \sqrt{1 - \rho^2}} = \frac{\sigma_0^{(j)}}{\sigma_i^{(j)}} && \text{(varying for time indices } i \text{ and blots } j) \\ s = s_i^j &:= \rho \frac{\sigma_Z}{\sigma_W} && \text{(fix for all time indices } i \text{ and blots } j) \\ a = a_i^j &:= \frac{\mu_Z - s\mu_W}{\sigma_Z \sqrt{1 - \rho^2}} && \text{(fix for all time indices } i \text{ and blots } j) \\ b = b_i^j &:= \frac{\mu_W}{\sigma_W} = \frac{\hat{\eta}_0^{(j)}}{\sigma_0^{(j)}} \approx \frac{\bar{\mu}_W}{\bar{\sigma}_W} && \text{(varying only for blots } j) \end{aligned}$$

The shape determining translation constants a and b are independent of the time indices i , and fixed per blot j . Using the BLU measurement values of the *unstim* experiment from table 3.3d, we get:

$$\begin{aligned} a^{(1)} &= 0, & a^{(2)} &= 0, & a^{(3)} &= 0 \\ b^{(1)} &\approx 7.35, & b^{(2)} &\approx 6.91, & b^{(3)} &\approx 4.58 \end{aligned} \quad (3.10)$$

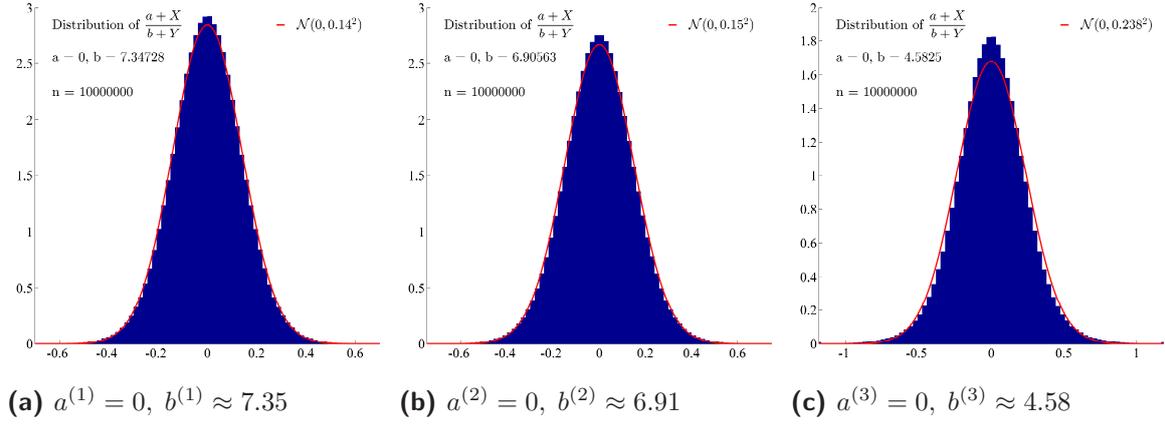


Figure 3.12: Distribution of additive error ϵ_a . Dark areas mark normalized histograms resulting of ten million evaluations of $T = (a + X)/(b + Y)$ with $X, Y \sim \mathcal{N}(0, 1)$. Cumulants μ and σ^2 of the approximating normal distribution (red line) are calculated using formula (3.2).

which, for b , are the reciprocals of the coefficients of variation of the respective unstimulated experiment, and the conditions for approximability by a normal distribution ($a < 2.25, b > 4$) (see section 3.4.1) are fulfilled.

The standard deviations $\sigma_i^{(j)}$ of the measurements $\eta_i^{(j)}$ in relative light units are available as estimates in eq. (3.9) only for the unstimulated setting. For the other experiments, there is only exactly one sample for each time point on every blot, prohibiting to estimate $\sigma_i^{(j)}$ from data. However, it seems reasonable to assume that, for higher signal intensities, the variance will also increase, but the coefficient of variation will be constant. Since the normalizing constants $b^{(j)}$ (eq. (3.10)) are the reciprocals of the coefficients of variation we can extrapolate values of $b_i^{(j)}$ for the stimulation experiments e.g. by the worst-case value 4.58 (corresponding to a CV of 0.2182 as observed in table 3.3d, rep #3, or less pessimistic to the mean value of the $b^{(j)}$ above.

This leads us to the following conclusion:

3.4 Conclusion

The additive error term ϵ_a is approximately normally distributed.

For the **multiplicative bias** $\epsilon_m = \hat{\eta}_0^{(j)} / (\hat{\eta}_0^{(j)} + \varepsilon_0^{(j)})$ of the amplification factors $m_i^{(j)}$, the methods developed by MARSAGLIA cannot be applied: since the standard deviation of the (constant) numerator is zero, it cannot be transformed to the standard form.

To investigate the distribution of the multiplicative bias, we abbreviate

$$\mu := \hat{\eta}_0^{(j)}, \quad \sigma := \sigma_0^{(j)}, \quad \tilde{X} := \varepsilon_0^{(j)} \sim \mathcal{N}(0, \sigma^2), \quad \tilde{Y} := \epsilon_m = \frac{\mu}{\mu + \tilde{X}}$$

Thus, we have

$$\tilde{Y} \sim Y := \frac{1}{X}, \quad \text{if } X \sim \mathcal{N}\left(1, (\sigma/\mu)^2\right),$$

having the same distribution as the multiplicative bias.

Conditioning the denominator with $X > 0$ (even for the worst observed coefficient of variation in our case of measuring pSTAT-3 in unstimulated cells, see table 3.3d, we have $P(X < 0) \approx 2 \cdot 10^{-6}$) we can derive the cumulative distribution function of Y as:

$$\begin{aligned} F_Y(x) &= P(Y \leq x) = P\left(\frac{1}{X} \leq x\right) = P\left(X \geq \frac{1}{x}\right) = 1 - P\left(X \leq \frac{1}{x}\right) = 1 - F_X\left(\frac{1}{x}\right) \\ &= \frac{1}{2} - \frac{1}{2} \operatorname{erf}\left(\frac{\frac{1}{x} - \mu}{\sqrt{2\sigma^2}}\right) = \frac{1}{2} - \frac{1}{\sqrt{\pi}} \int_0^{\frac{\frac{1}{x} - \mu}{\sqrt{2\sigma^2}}} e^{-t^2} dt. \end{aligned}$$

Taking the derivative w.r.t. x yields the probability density of Y as

$$f_Y(x) = \frac{d}{dx} F_Y(x) = \frac{1}{\sqrt{2\pi\sigma^2}x^2} e^{-\left(\frac{\frac{1}{x} - \mu}{\sqrt{2\sigma^2}}\right)^2}. \quad (3.11)$$

For measuring the similarity of two probability densities, we use the *overlapping coefficient*, as given in the following definition.

3.5 Definition (Overlapping Coefficient (OVL))

For two real probability density functions $f(x)$ and $g(x)$, the *overlapping coefficient* (OVL) is defined as

$$\text{OVL} = \int_{-\infty}^{\infty} \min\{f(x), g(x)\} dx. \quad (3.12)$$

□

Obviously, $\text{OVL} \in [0, 1]$, and $\text{OVL} = 1$ denotes a full overlap of two densities (i.e. same distribution). Note that the overlapping coefficient holds the relation $\text{OVL} = 1 - D$ to the widespread *dissimilarity index* $D = \frac{1}{2} \int_{-\infty}^{\infty} |f(x) - g(x)| dx$.

As INMAN and BRADLEY denote:

“The *overlapping coefficient* possesses three notable advantages as a measure of the agreement between two distributions. First, it provides a common approach for the measure of the similarity of these distributions in any distributional setting. Second, OVL is based on a simple, easily comprehended concept of the agreement or similarity of probability distributions. Third, the invariance of OVL under appropriate transformation makes this measure of agreement attractive from the standpoint of computation and estimation.” [InmanBradley1989]

Especially its invariance under continuously differentiable transformations in change of variables renders the overlapping coefficient a suitable measure for our case.

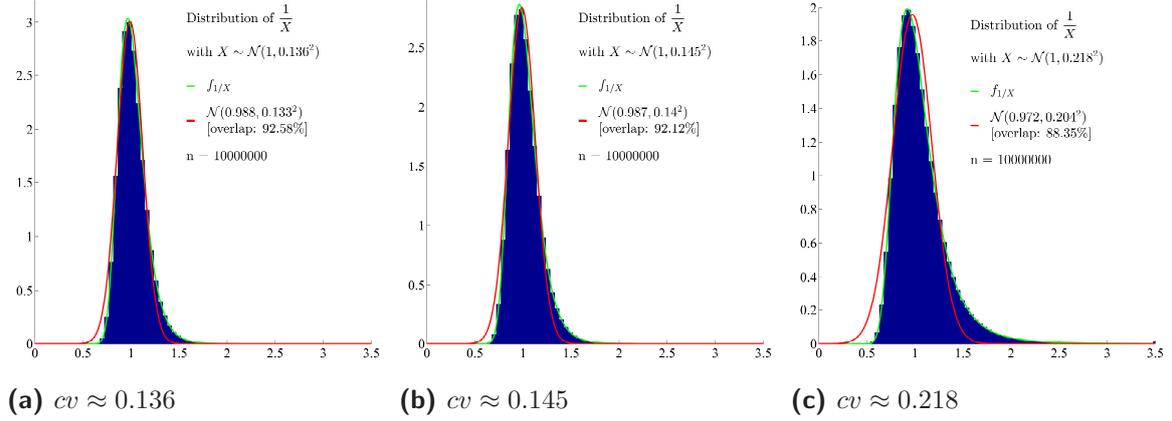


Figure 3.13: Distribution of multiplicative bias ϵ_m . Filled area marks normalized histogram resulting of ten million evaluations of $T = 1/X$ with $X \sim \mathcal{N}(1, \tilde{\sigma}^2)$. The theoretically derived density $f_{1/X}$ (see eq. (3.11)) is printed as green line. Coefficients of variation of pSTAT-3 in the unstimulated setting (table 3.3d) are used as standard deviation $\tilde{\sigma}$. Approximating normal distributions (red line) are derived by maximizing the overlap (integral of minimum) of the normal density function and $f_{1/X}$.

Figure 3.13 depicts the distribution of ϵ_m . It shows some positive skewness, and with overlapping coefficients of about 90%, approximability by normal distributions with a mean value close to 1 and a small variance seems feasible.

There remains the questions, whether the shape of the distribution of the **measured amplification factors** is formed more by the slightly skewed multiplicative bias or by the (close to) normally distributed additive error. To answer this, a simulation study was done.

From the wet lab data (table 3.3a), the maximum amplification factor due to stimulation with IL-6 can be determined to lie around 50. For the simulation study, $\hat{m} = 50$ was used as the experimentally non-accessible true amplification factor. Multiplicative bias ϵ_m and additive error ϵ_a were calculated as described in the previous sections. Figure 3.14 shows the simulation result.

With overlapping coefficients higher than 95% to a normal distribution with mean close to 50, we come to the following conclusion towards the distribution of amplification factors:

3.6 Conclusion

The amplification factors method delivers approximately normally distributed quantities, whose mean is close to the true amplification factor.

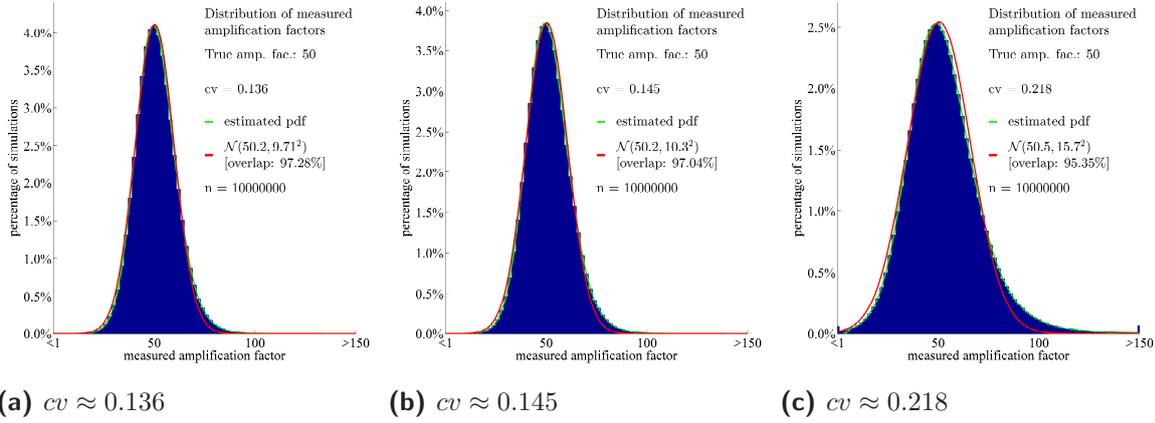


Figure 3.14: Distribution of simulated amplification factors. Filled area marks normalized histogram resulting of ten million evaluations of simulated amplification factors. True amplification factor was $m = 50$. Coefficients of variation (cv) of pSTAT-3 in the unstimulated setting (table 3.3d) are used as standard deviation σ_0 of the reference points, whereas $m\sigma_0$ was used in the amplified signal, resulting in the same coefficients of variation.

The green lines indicate the estimated probability density derived by nonparametric kernel smoothing using normal kernel functions. Approximating normal distributions (red line) are derived by maximizing the overlap (integral) of the normal density function and the estimated probability density function.

Note: This simulation takes into account only a single measurement of the maximum amplification factor. For whole experiment simulations see figure 3.15.

3.5.5 Simulation study: Effect of normalization on signal intensity

To further investigate the effects of normalization using house-keeping or recombinant proteins, we perform an *in silico* study on an artificially generated signal and normalizer signal.

We take a look at the signal intensity in terms of amplification factors on the noisy measurements of the artificial signal itself, and of the normalized signal, processed by using the artificial normalizer measurements.

3.5.5.1 Generation of the artificial signal

We apply a piecewise linear time transformation to a modified sine function to generate an artificial signal that is remarkably similar to the experimentally observed kinetics in the *IL6+* setting (compare section 3.6 on page 78 to the artificial signal in figure 3.15b).

The underlying sine function that we use as signal generator is given as

$$f(x) = 24.5 \cdot (1 + \sin(x + 1.5\pi)) + 1. \quad (3.13)$$

Further, the continuous piecewise linear time transformation,

$$T(t) = \begin{cases} \frac{\pi}{25}t & t \in [0, 30] \\ \frac{\pi}{75}t + \frac{4}{5}\pi & t \in [30, 60] \\ \frac{\pi}{150}t + \frac{6}{5}\pi & t \in [60, 120] \end{cases} \quad (3.14)$$

delivers a bijective mapping from the time domain $[0, 120]$ into the interval $[0, 2\pi]$.

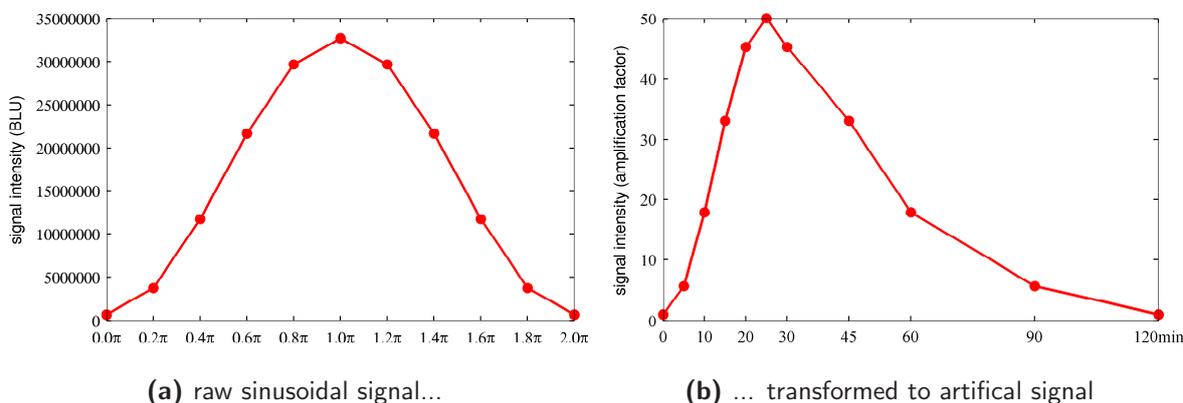


Figure 3.15: Raw and time-transformed artificial signal.

The sinusoidal signal in (a) is generated by evaluating eq. (3.13) at time points $\tilde{t}_i = i/5 \cdot \pi$ ($i = 0..10$), equidistantly sampling the interval $[0, 2\pi]$ for a whole period.

The artificial signal (b) is formed by applying the time-transformation T (eq. (3.14)) to the measurement time points $\{0, 5, 10, 15, 20, 25, 30, 45, 60, 90, 120\}$.

To generate the artificial signal, for each time point $t_i \in \{0, 5, 10, 15, 20, 25, 30, 45, 60, 90, 120\}$, the signal generator function f is evaluated at $f(T(t_i))$ and scaled⁽²⁵⁾ by a fixed random number $k \sim \mathcal{U}(500000, 1000000)$, independent of i . Figure 3.15 shows both the artificial signal generated by the procedure above (b) and the underlying sinusoidal base signal (a).

The artificial normalizer signal is generated by choosing a number $\tilde{k} \sim \mathcal{U}(500000, 1000000)$ and assigning it to all sample time points.

Both, the artificial and the normalizer signals are subsequently distorted by a multiplicative normally distributed noise. It is obvious, that this distortion certainly changes the observed maximum amplification factor. Indeed, it might also influence the time point where it is observed, as, by chance, the (truly) highest signal might be distorted in a way that it is lower than the signal observed at other time points.

3.5.5.2 Simulation results

As figure 3.16 shows, both expected types of errors occur. Ideally, i.e. without any error in the measurements, a single bar at 50-fold containing 100% of the simulations would be observed, and the time point histogram in the right column of figure 3.16 would consist of a single bar at time point 25min.

Obviously, applying the normalization to the sinusoidal signal before calculating the amplification factors results in a much broader distribution of the observed maximum amplification factor.

For a noise level of 10%, both mean maximum amplification factors of the **signal itself** (52.1-fold) and of the **(normalized signal)** (54.3-fold) are close to the true value of 50-fold; however, their standard deviations differ considerably (6.7 vs. 10.0). For higher noise levels, the mean value of the normalized signal deviates much faster from the true value than the mean value of the unprocessed signal (noise level of 20%: 57.8-fold vs. 65.6-fold, noise level of 30%: 65-fold vs. 78.1-fold). For the standard deviation, these effects are even worse (noise level of 20%: 15.5 vs. 25.3, noise level of 30%: 26 vs. 39.2).

⁽²⁵⁾ Multiplying by k does not affect the subsequent calculations, but scales the raw artificial signal into a range frequently observed in BLU data.

The impact of measurement errors on the time point at which the maximum signal intensity is detected is demonstrated in the rightmost pictures in figure 3.16. If solely the signal itself is used to calculate the amplification factor (depicted as green boxes ) , with increasing noise level, the time point at which the maximum signal is observed spreads: For a noise level of 0.1, the correct time point is identified in more than two thirds of all cases (figure 3.16a); this declines to still more 40% if a noise level of 0.3 (figure 3.16c) is simulated.

Again, the situation is worse if the signal is at first normalized (red boxes ): Even for small disturbances ($CV = 0.1$), the time point of the maximum signal intensity is correctly detected in only about half of all cases. For larger noise levels, this already declines to less than 40% ($CV = 0.2$) or roughly 30% ($CV = 0.3$). In the case of the highest simulated noise, two out of five measurements will show the maximum level more than 10 minutes away from the true time point – in contrast to only 5% if the signal is not normalized.

3.7 Conclusion

Normalization with a house-keeping or manually added calibrator protein may massively disturb the intensity and time point of the peak signal already at moderate noise levels. Under same noise intensities, using solely the raw signal – as proposed in the amplification factors method – delivers much better estimates on peak amplification and peak time point (see figure 3.16).

We finish this small study with the remark that the underlying sampling is also influencing the observation distribution. For example, a thinner sampling around the time point of the maximum signal level (25min) would result in a narrower distribution of the detected time points. Taking more samples generally results in a broadened histogram. For the simulation study, we used the same sampling as in the real experiments (section 3.3).

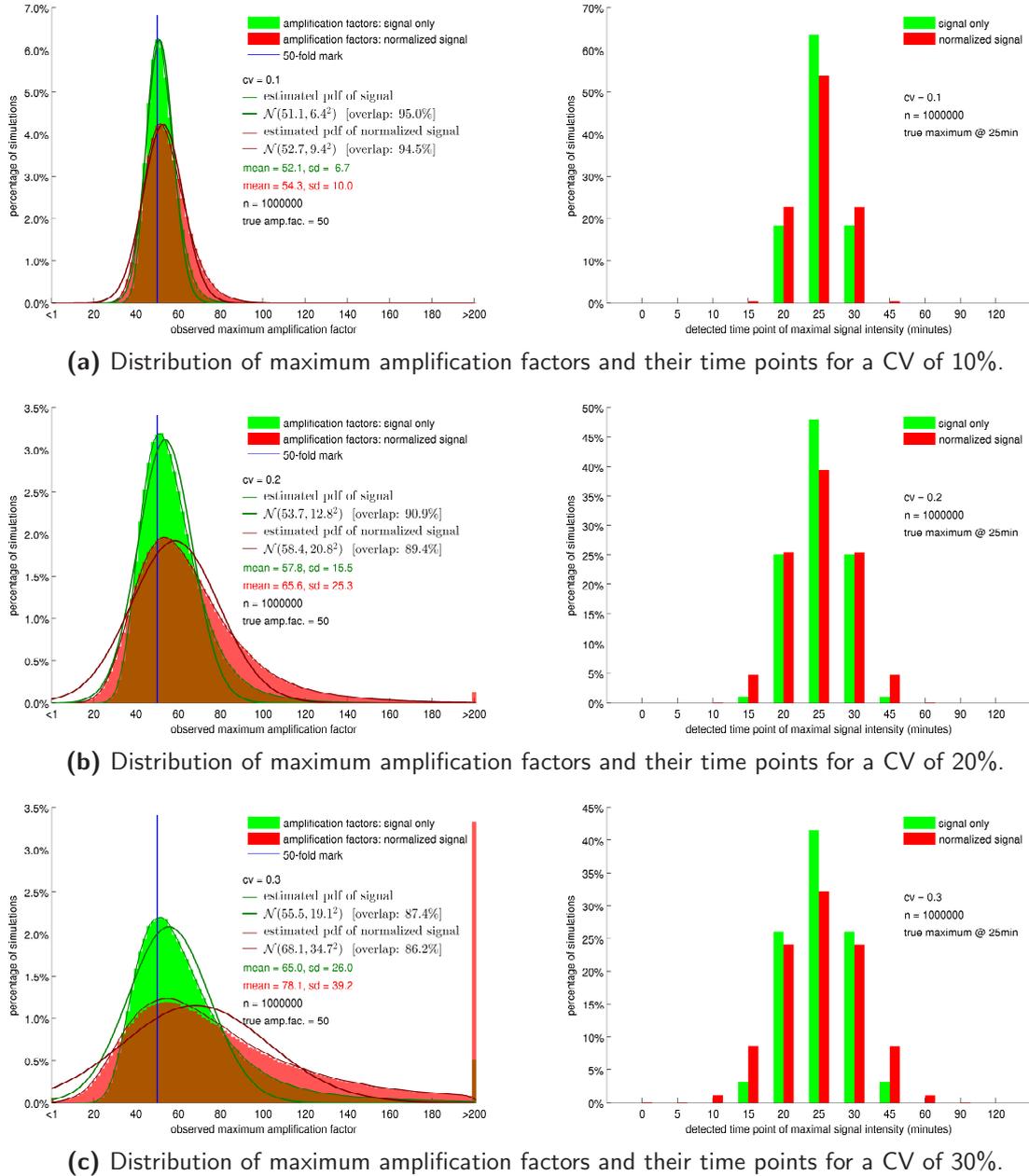


Figure 3.16: Normalization effects on distribution of maximum amplification factors and their time point for different levels of noise: (a) CV = 0.1, (b) CV = 0.2, CV = 0.3.

In this simulation study, 1000000 artificial sinusoidal signals with a peak amplification factor of 50 reached at 25min (see figure 3.15b) as well as the same amount of artificial normalizer signals were created; both distorted by identically distributed noise to simulate measurement errors.

Shown are histograms of the observed maximum amplification factors with (■) and without (■) prior normalization. Without any disturbance, a single bar at the 50-fold mark would be observed. From both histograms, an estimated probability density was derived by nonparametric kernel smoothing using normal kernel functions (thin green — and red — lines), to which normal distributions were fitted by maximizing the overlap (integral) of the respective estimated probability density function and the normal density (thick green — and red — lines). The values of *mean*, *sd* (in green colour), *mean*, *sd* (in red colour) denote the empirical mean and standard deviations for each series. Amplification factors exceeding the interval (1, 200) were excluded from all calculations.

The rightmost figures show the distribution of the time points, where the maximum signal was observed. The symmetry around the true time point 25min is due to the symmetry of the underlying sine function.

3.6 Results of the IL-6 stimulation experiments: Processed data

From the BLU measurement data given in table 3.3, molar concentrations have been calculated using the amplification factors of pSTAT-3 as described in appendix A.1.4. Figure 3.17 gives a graphical representation of the cytosolic phospho-STAT-3 time course upon the four experimental settings. Depicted are the mean values (black plus signs, +) and their linear interpolation (gray line, —). The error bars originate from triplicate measurements and show the measurements' standard deviations for each time point. For comparison, the results of sole IL-6 stimulation (section 3.6, *IL6+*) are copied to the other stimulation settings in light gray tones.

In all three stimulation experiments (sections 3.6 to 3.6), a kinetic of similar quality is observed. An initial rise of the signalling phospho-STAT-3 immediately follows the stimulation and reaches a peak after around 25 to 30 minutes. The signal then slowly declines, reaching basal levels after 90–120 minutes again. Contrary to the similar qualitative behaviour, the outcomes differ significantly in their quantitative characteristic.

Sole stimulation with IL-6 (section 3.6, *IL6+*) results in a strong cellular response in cytosolic phospho-STAT-3. The peak is reached at about 25 minutes after stimulation, showing an approximately 47-fold increase of pSTAT-3 in the cytoplasm compared to basal level at the initial time point $t = 0min$.

In contrast to that, co-stimulation with GM-CSF (section 3.6, *IL6+/GMCSF+*) leads to a significantly lower response in cytosolic pSTAT-3. The peak is still reached after 25–30 minutes, however, only a 12-fold increase of the activated transducing protein is observed.

Similar observations are made if secreted GM-CSF is blocked by appropriate antibodies while maintaining the IL-6 stimulation (section 3.6, *IL6+/GMCSF-*). While the peak of pSTAT-3 in the cytosol is still achieved 25 minutes after stimulation, its level is again significantly lower compared to the sole IL-6 stimulation setting: only a 18-fold increase after stimulation is observed.

In unstimulated cells (section 3.6, *unstim*), a stable basal level of cytosolic phospho-STAT-3 is observed, varying no more than 1.3-fold (note the different scale).

3.6.1 Statistical analysis

To check whether the observed deviation in maximum cytosolic phospho-STAT-3 levels in the co-stimulation and blocking experiments compared to sole IL-6 stimulation are due to the different stimulation setting, and not due to random perturbations, we performed two-tailed t-tests (WELCH's t-test, [Welch1947]). Population variances have been estimated from the unbiased sample variance (with applied BESSEL's correction) retrieved from triplicate measurements.

Table 3.5 gives an overview of the respective peak amplification factors of the phospho-STAT-3 signal in every replicate of each experimental setting.

With p-values of less than 0.0049 for the *IL6+/GMCSF-* setting and less than 0.011 for the *IL6+/GMCSF+* co-stimulation setting, we can conclude that these observations are very unlikely to be a random event but clearly resulting from the different stimulation settings. The null hypotheses that the maximum amplification factors of *IL6+* and *IL6+/GMCSF+*, or *IL6+* and *IL6+/GMCSF-*, originate of independent observations with same mean (no assumptions on the variances, as they might be influenced by different blots) have to be withdrawn with a significance level of more than 98%.

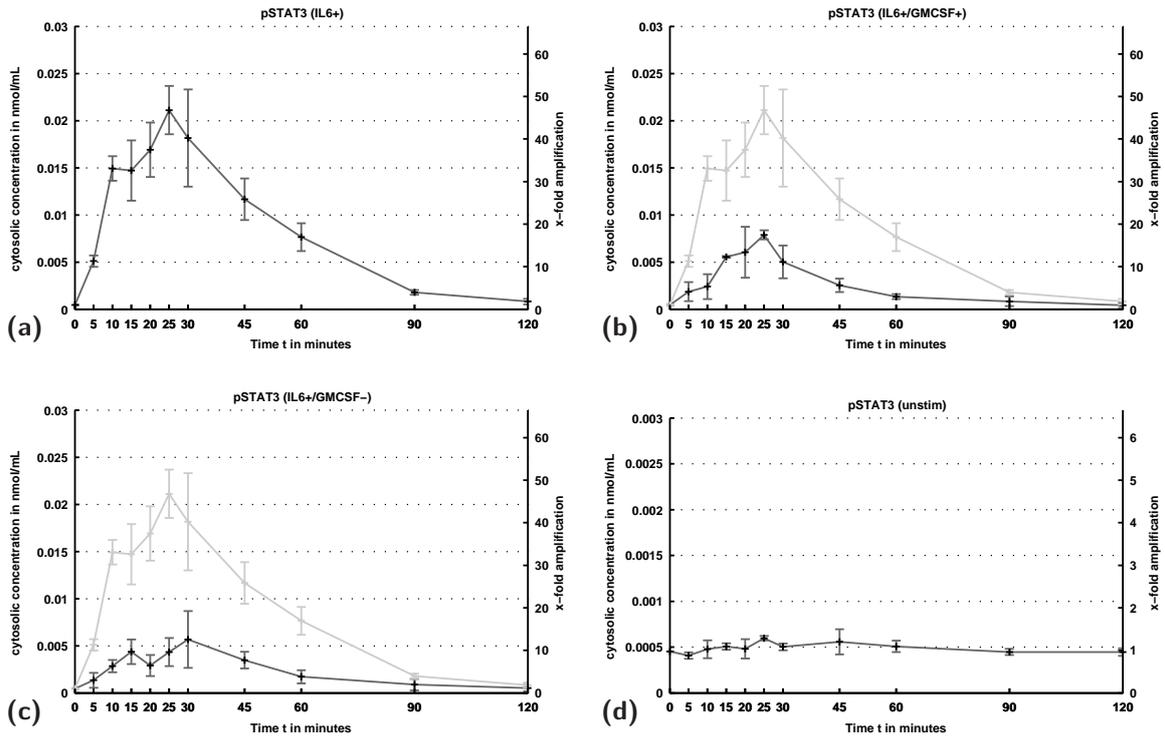


Figure 3.17: Measurements of cytosolic pSTAT-3 upon four experimental settings. Linearly interpolated mean values of triplicate measurements, with error bars from empirical standard deviation. For comparison, the course of sole IL-6 stimulation (a) is copied to (b) and (c) in light gray colours. (a) Stimulation with sole IL-6 results in a 47-fold rise of cytoplasmic phosphorylated STAT-3 level relative to time $t = 0min$. (b) Simultaneous stimulation with both IL-6 and GM-CSF leads to a significantly lowered response compared to sole IL-6 stimulation, and results in an only 16-fold increase in cytoplasmic pSTAT-3 level. (c) With blockade of GM-CSF with appropriate antibodies, IL-6 stimulation leads only to a 14-fold increase in cytoplasmic pSTAT-3 concentration. (d) In unstimulated cells, a low but stable base-level of phosphorylated STAT-3 is observed (note the different scale).

Table 3.5: Statistical analysis of pSTAT-3 peak levels.

Amplification factors (peak-to-unstimulated, rounded to nearest integer for replicates) upon different stimulation settings. *IL6+*: Sole IL-6 stimulation, *IL6+/GMCSF-*: IL-6 stimulation with simultaneous GM-CSF blockade, *IL6+/GMCSF+*: Co-stimulation with IL-6 and GM-CSF. For the unstimulated control *unstim*, fluctuation factors around the mean are given. Standard deviation (SD) is with applied BESSEL's correction.

The p-values denote the probability that the changes in the amplification factors compared to sole IL-6 stimulation are not due to the different stimulation setting.

experiment	amplification factors (peak to unstimulated)					p-value
	rep #1	rep #2	rep #3	mean	SD	
<i>IL6+</i>	53x	41x	48x	47.4x	6.0x	—
<i>IL6+/GMCSF-</i>	13x	6x	22x	13.6x	7.8x	< 0.0049
<i>IL6+/GMCSF+</i>	19x	18x	16x	17.9x	1.3x	< 0.011
<i>unstim</i>	0.7x–1.2x	0.8x–1.3x	0.7x–1.4x	1.0x	—	—

We thus formulate the following finding:

3.8 Conclusion

The observations of lowered response in IL-6-induced phospho-STAT-3 in the setting of co-stimulation with IL-6 and GM-CSF, $IL6+/GMCSF+$, as well as in the setting of IL-6 stimulation with simultaneous blocking of GM-CSF, $IL6+/GMCSF-$, compared to sole IL-6 stimulation is with very high probability caused by the different stimulation scenarios and not due to random events.

3.6.2 Summary of initial experimental results

Qualitatively, STAT-3 activation upon all IL-6 stimulation settings (i.e. $IL6+$, $IL6+/GMCSF+$, and $IL6+/GMCSF-$) is as expected and as described in literature: A steep rise peaking around 15–25 minutes after stimulation followed by a slow decline back to basal levels due to pathway down regulation.

However, the amplification factors approach reveals large quantitative differences in STAT-3 activation upon sole IL-6 stimulation ($IL6+$) and the GM-CSF-blocking ($IL6+/GMCSF-$) settings, as the latter shows a much weaker STAT-3 activation.

While both cytokines, IL-6 and GM-CSF, have been shown to stimulate cell proliferation (figure 3.5) and to activate STAT-3 ([Rawlings2004; Valdembri2002]), the co-stimulation setting $IL6+/GMCSF+$ proposed by the author did not show the expected increased STAT-3 activation compared to sole IL-6 stimulation. In fact, we made the counter-intuitive observation that co-stimulation also results in a significantly lowered STAT-3 activation.

3.7 A quantitative and predictive model of a crosstalk of IL-6 and GM-CSF in JAK-STAT signalling in HaCaT A5 benign tumour keratinocytes

3.7.1 Review of existing models

There exist a number of mathematical models for the JAK-STAT pathway [Yamada2003; Ghosh2011; Guerriero2009; Singh2006; Soebiyanto2007; Sun2008; Swameye2003], that address signalling in a variety of cell types induced by different cytokines (IL-6 signalling in hepatocytes [Singh2006], erythropoietin (EPO) signalling in BaF3 cells [Swameye2003], unified leukaemia inhibitory factor (LIF) and oncostatin M (OSM) signalling via gp130 in MCF-7 cells [Guerriero2009]) or pure *in silico* experiments (EPO signalling [Sun2008], interferon (IFN) signalling [Yamada2003; Soebiyanto2007]). Small-sized simplified models consisting of only a handful of formal species and parameters [Sun2008; Swameye2003] stand in contrast to detailed ones with dozens of species [Yamada2003; Guerriero2009; Singh2006], and most models are of a qualitative or semi-quantitative nature.

Our model addresses a crosstalk in the signalling of two cytokines that signal via distinct receptor classes. Models describing signalling cascades induced by more than a single source of stimulation are rare. One example is the JAK-STAT model of GUERRIERO et al.

[Swameye2003], which includes the competition between LIF and OSM that can both signal via a gp130/LIFR receptor complex and uses computational costly stochastic GILLESPIE simulations.

YAMADA et al. [Yamada2003] perform *in silico* studies with a detailed model for IFN signalling with 31 formal species and 51 parameters, but their model is too large for the available data, and it is unclear whether the manually tuned parameters therein can be conferred to other models.

SOEBIYANTO et al. [Soebiyanto2007] applied model reduction to the model of YAMADA et al., but using their simplified model for our purpose is cast into doubt by the authors' report that their simulation results contradict literature data for other cell lines.

SUN et al. [Sun2008] and SWAMEYE et al. [Swameye2003] use interpolated time-course measurements of the receptor as a "driving function" (control, input), restricting the model to the specific cell type and stimulation setting used for recording the input function.

GHOSH et al. [Ghosh2011] investigated cell proliferation during cancer metastasis. They address *intercellular* communications and therefore combine the whole IL-6-induced *intracellular* signalling in a single equation, derived from a minimalistic model, ignoring nuclear transcription and neogenesis of any inhibitory protein like SOCS-3.

The IL-6 stimulation model of STEVEN et al. [Steven2009] (see figure 3.18) is based on the work of SINGH et al. [Singh2006], and incorporates intracellular crosstalks between the JAK-STAT adopted from YAMADA's [Yamada2003] work and the mitogen-activated protein kinase (MAPK/ERK) pathways adopted from SCHOEBERL et al. [Schoeberl2002], but both triggered by a sole stimulus of IL-6. STEVEN et al. present an extensive mathematical model consisting of 65 ordinary differential equations and 111 (unknown) parameters that describes the intracellular reactions. Yet, the vast number of parameters introduces a significant risk of overfitting the comparatively sparse data that is available, making it difficult to use this model by extending it with the GM-CSF signalling part.

Summarizing, there are a bunch of existing models, but suffering from identical problems. All models show a certain degree overparametrization, are on the one hand too detailed for our purposes (see, e.g. figure 3.18), and do not contain essential pathway connections on the other hand.

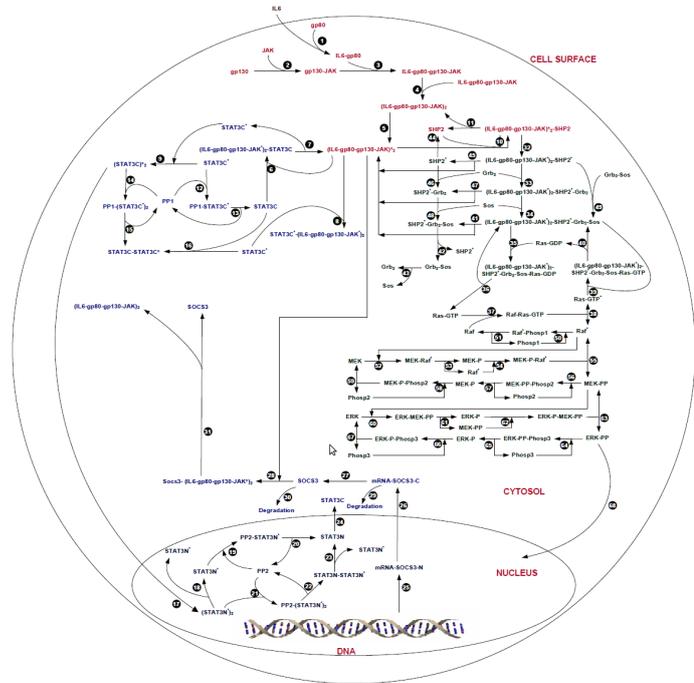


Figure 3.18: JAK-STAT model by STEVEN et al. [Steven2009]. The model consists of 65 formal species, one control (IL-6 stimulation), and 111 parameters that are manually set to "reasonable" values.

3.7.2 Our IL-6/GM-CSF crosstalk model

Figure 3.19 gives a detailed graphical representation of our crosstalk model.

Upstream of the signalling cascade are the receptors for IL-6 and GM-CSF, which can only bind the appropriate proteins that are continuously secreted by the cells and can additionally be given as external stimuli. We focus on the essential players on both the receptor activation and the intracellular signal transduction levels to avoid extensive overparametrization of the model. As the receptor formation stoichiometry of the IL-6 receptor has not yet been finally determined [Schroers2005], both the IL-6 receptor and its β receptor gp130 have been unified to a single “IL-6 receptor complex” species in the model, and the same simplification was applied to the GM-CSF signalling components. Ligand binding to these complexes was modelled to be sufficient for receptor activation, and the activation rates for both IL-6 and GM-CSF receptor complexes were modelled to be equal.

An activated IL-6 receptor complex phosphorylates STAT-3 molecules and transforms them into a signalling competent species, implicitly modelling dimerisation of activated STAT-3. We modelled the activation of STAT-3 to be facilitated by a supportive kinase (SK) that is activated by the ligand-bound GM-CSF receptor. Activated STAT-3 may enter the nucleus and starts transcription of SOCS-3 mRNA, gets dephosphorylated by nuclear phosphatases and

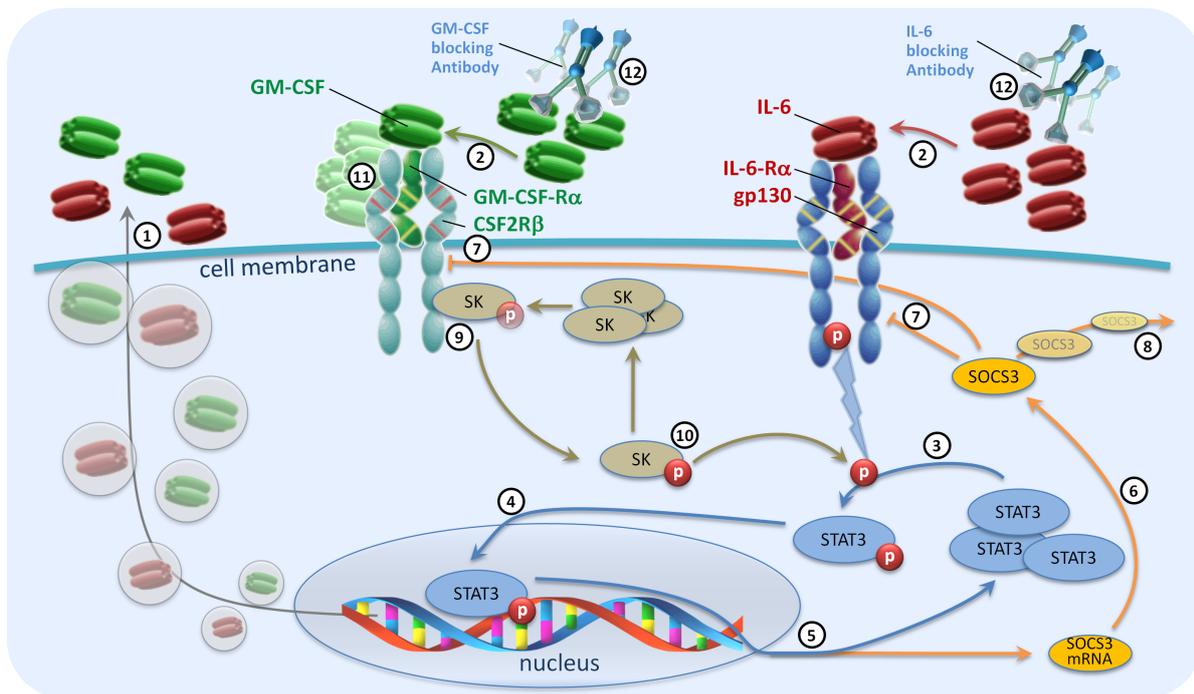


Figure 3.19: The crosstalk model. ① Continuous production and secretion of IL-6 and GM-CSF ② Receptor activation through ligand binding ③ IL-6-induced phosphorylation of STAT-3 ④ Nuclear import of activated STAT-3 protein ⑤ Transcription of SOCS-3 mRNA and export of deactivated STAT-3 protein ⑥ Translation of SOCS-3 in the cytoplasm ⑦ SOCS-3 binds to active receptors thus cancelling signal transduction ⑧ Degradation of SOCS-3 ⑨ SK associates to and gets activated on active receptor ⑩ Activated SK enhances IL-6-induced STAT-3 signalling ⑪ Overstimulation with GM-CSF inhibits SK activation ⑫ Cytokines can be blocked with appropriate antibodies

Table 3.6: List of model species in our IL-6/GM-CSF crosstalk model.

<i>Var.</i>	<i>Species Name</i>	<i>Meaning</i>
x_1	GMCSF	auto-/paracrine GM-CSF
x_2	GMCSF_RC	α and β receptor complex for GM-CSF
x_3	GMCSF_LRC	ligand bound active GM-CSF receptor
x_4	GMCSF_LRC_blocked	overstimulated GM-CSF receptor complex
x_5	IL6	auto-/paracrine IL-6
x_6	IL6_RC	α and β receptor complex for IL-6
x_7	IL6_LRC	ligand bound active IL-6 receptor
x_8	STAT3	cytoplasmic STAT-3
x_9	pSTAT3	cytoplasmic phospho-STAT-3
x_{10}	pSTAT3_nuc	nuclear phospho-STAT-3
x_{11}	SK	supporting kinase
x_{12}	aSK	active supporting kinase
x_{13}	SOCS3_mRNA	SOCS-3 mRNA in cytoplasm
x_{14}	SOCS3	SOCS-3 in cytoplasm
x_{15}	IL6_LRC_SOCS3	SOCS-3 deactivated IL-6 receptor
x_{16}	GMCSF_LRC_SOCS3	SOCS-3 deactivated GM-CSF receptor

leaves the nucleus again, refilling the cytoplasmic pool of STAT-3. The translation process of SOCS-3 mRNA in the cytoplasm was simplified by the assumption that 10 proteins are built from one mRNA. The translated protein then binds to active receptor complexes, eventually ceasing signal transduction.

We addressed modelling the transient behaviour of phospho-STAT-3 in HaCaT keratinocytes using solely mass action kinetics, since MICHAELIS-MENTEN type kinetics are applicable only if the amount of substrate is much higher than the amount of enzyme, a condition that is usually not fulfilled in signal transduction pathways (see, e.g., [Yamada2003]).

3.7.2.1 Model species and reaction scheme

Our model consists of 16 formal species, denoted either by a textual abbreviation in type-writer letters (e.g. pSTAT3) or by a corresponding indexed variable (e.g. x_9). Species names in the model are chosen corresponding to cellular molecules or subunits, though simplified. The acronym RC is short for *receptor complex*, combining all involved receptor subunits into one virtual species. Analogously, LRC denotes a *ligand receptor complex*, coding a signalling competent and ligand-bound active signalling receptor complex. A complete list of species is given in table 3.6.

We formulated the model in 15 biochemical reactions, which we found to be a minimum set describing the IL-6/GM-CSF crosstalk and JAK-STAT signalling cascade with the simplifications described at the beginning of this section. All reactions are modelled as irreversible biochemical reactions that can be translated into ordinary differential equations using the law of mass action.

Table 3.7 lists the involved reactions in our crosstalk model. In table 3.8, the resulting mathematical model is given in terms of a system of ordinary differential equations. The derivation of initial values is discussed in section 3.7.4.

3.7.2.2 Comprehensibility and interpretability as crucial points in our modelling

Concerning the fact, that there are only cytosolic phospho-STAT-3 and SOCS-3 measurement data available for model calibration (see sections 3.7.5 to 3.7.7), the model is quite detailed, though most of its formal species cannot be directly accessed by wet lab experiments. It is therefore also afflicted with a certain degree of overparametrization.

While there exist efficient model reduction techniques (model order reduction, MOR) that reduce the complexity of large-scale dynamical systems by finding lower-dimensional approximations producing the same input-output relations, we refuse to utilize these methods.

From our point of view, while building mathematical models of biological pathways, it is of most importance that the mathematical model must stay *understandable*, i.e. there must be a clear mapping between formal model species to *real* molecules or functional subsystems in the cell, and from formal model parameters to *real* properties of biological processes like reaction rates.

The application of automatic model order reduction techniques might deliver better estimates in a reduced model (better in the sense of reduced parameter uncertainty), but at the cost of losing the indispensable mapping between model and reality.

Table 3.7: Reactions in the crosstalk model. Underlying reactions in our IL-6/GM-CSF crosstalk model. The educts GMCSF and IL6 consist of autocrine and external cytokine, with exception of the GM-CSF receptor overstimulation reaction, where only externally applied cytokine is involved.

<i>Description</i>	<i>Par.</i>	<i>Reaction equation</i>
Autocrine IL-6 and GM-CSF production		
GM-CSF production	p_1	\Rightarrow GMCSF
IL-6 production	p_1	\Rightarrow IL6
Ligand receptor interaction		
activation of GM-CSF receptor	p_2	$\text{GMCSF} + \text{GMCSF_RC} \Rightarrow \text{GMCSF_LRC}$
activation of IL-6 receptor	p_2	$\text{IL6} + \text{IL6_RC} \Rightarrow \text{IL6_LRC}$
GM-CSF receptor overstimulation	p_3	$\text{GMCSF_LRC} + \text{GMCSF} \Rightarrow \text{GMCSF_LRC_blocked}$
Activation of supporting kinase and catalysis of STAT-3 activation		
SK activation by GMCSF_LRC	p_4	$\text{SK} + \text{GMCSF_LRC} \Rightarrow \text{aSK} + \text{GMCSF_LRC}$
aSK catalyses pSTAT-3 formation	p_5	$\text{aSK} + \text{IL6_LRC} + \text{STAT3} \Rightarrow \text{aSK} + \text{IL6_LRC} + \text{pSTAT3}$
SOCS-3 interactions		
SOCS-3 decay	p_6	$\text{SOCS3} \Rightarrow$
SOCS-3 deactivates IL6_LRC	p_7	$\text{IL6_LRC} + \text{SOCS3} \Rightarrow \text{IL6_LRC_SOCS3}$
SOCS-3 deactivates GMCSF_LRC	p_7	$\text{GMCSF_LRC} + \text{SOCS3} \Rightarrow \text{GMCSF_LRC_SOCS3}$
IL-6 signal transduction via JAK-STAT and SOCS-3 induction		
STAT-3 phosphorylation IL6_LRC	p_8	$\text{IL6_LRC} + \text{STAT3} \Rightarrow \text{IL6_LRC} + \text{pSTAT3}$
pSTAT-3 translocation into nucleus	p_9	$\text{pSTAT3} \Rightarrow \text{pSTAT3_nuc}$
nuclear export of STAT-3	p_{10}	$\text{pSTAT3_nuc} \Rightarrow \text{STAT3}$
induction/export of SOCS-3 mRNA	p_{11}	$\text{pSTAT3_nuc} \Rightarrow \text{STAT3} + \text{SOCS3_mRNA}$
translation of SOCS-3	p_{12}	$\text{SOCS3_mRNA} \Rightarrow 10 * \text{SOCS3}$

3.7.3 ODE model formulation

Using the law of mass action to describe the reactions listed in table 3.7, we formulated the following nonlinear ordinary differential equation (ODE) model of the IL-6/GM-CSF crosstalk model in HaCaT-ras A5 keratinocytes, where u_1 and u_2 denote the external stimuli of GM-CSF and IL-6 respectively.

Table 3.8: ODE model equations.

GMCSF :	$\dot{x}_1 = +p_1 - p_2x_1x_2$
GMCSF_RC :	$\dot{x}_2 = -p_2(u_1 + x_1)x_2$
GMCSF_LRC :	$\dot{x}_3 = +p_2(u_1 + x_1)x_2 - p_3x_3u_1 - p_7x_3x_{14}$
GMCSF_LRC_blocked :	$\dot{x}_4 = +p_3x_3u_1$
IL6 :	$\dot{x}_5 = +p_1 - p_2x_5x_6$
IL6_RC :	$\dot{x}_6 = -p_2(u_2 + x_5)x_6$
IL6_LRC :	$\dot{x}_7 = +p_2(u_2 + x_5)x_6 - p_7x_7x_{14}$
STAT3 :	$\dot{x}_8 = -p_5x_7x_8x_{12} - p_8x_7x_8 + (p_{10} + p_{11})x_{10}$
pSTAT3 :	$\dot{x}_9 = +p_5x_7x_8x_{12} + p_8x_7x_8 - p_9x_9$
pSTAT3_nuc :	$\dot{x}_{10} = +p_9x_9 - (p_{10} + p_{11})x_{10}$
SK :	$\dot{x}_{11} = -p_4x_3x_{11} + p_5x_7x_8x_{12}$
aSK :	$\dot{x}_{12} = +p_4x_3x_{11} - p_5x_7x_8x_{12}$
SOCS3_mRNA :	$\dot{x}_{13} = +p_{11}x_{10} - p_{12}x_{13}$
SOCS3 :	$\dot{x}_{14} = -p_7(x_7x_{14} + x_3x_{14}) + 10p_{12}x_{13} - p_6x_{14}$
IL6_LRC_SOCS3 :	$\dot{x}_{15} = +p_7x_7x_{14}$
GMCSF_LRC_SOCS3 :	$\dot{x}_{16} = +p_7x_3x_{14}$

3.7.4 Initial cell state

The diameter of HaCaT A5 cells was determined with a CASY cell counter (Millipore, Billerica, MA) to be around 20 μm . Thus, the average HaCaT A5 cell volume was estimated to be $4.2 \cdot 10^{-6}$ μL , assuming a spherical cell shape.

The total number of gp130 molecules on a single cell was determined to be roughly $2.0 \cdot 10^4$ by quantitative immunoprecipitation, resulting in a molar concentration of about $7.96 \cdot 10^{-3}$ nmol/ μL . The total number of STAT-3 molecules in HaCaT keratinocytes was determined by the same methodology to be $3.7 \cdot 10^5$, resulting in a molar concentration of about 0.146 nmol/mL.

Using the amplification factors method, the initial concentration of cytosolic phospho-STAT-3 was calculated as $4.51 \cdot 10^{-3}$ nmol/mL.

Mass concentration of GM-CSF (8 pmol/mL per million cells) and IL-6 (6–14 pmol/mL per million cells) secreted by unstimulated cells has been determined by ELISA and both species were initialized with $0.4 \cdot 10^{-6}$ nmol/mL in the model, corresponding to 10 pmol/mL in the model for an assumed molar mass of 25 kDa both.

Initial concentration of SK was taken from the paper [RamisConde2009], where the concentration of the candidate kinase Src is given to be $3.3 \cdot 10^{-2}$ nmol/mL.

The number of molecules of the GM-CSF β receptor (common β chain) is set to one tenth of gp130 to approximately 2000 receptors, which is in good concordance to the range reported in literature (1518–3783 in [Williams1988], 1058–2304 in [Chiba1990]).

Section 3.7.4 summarizes the values that have been used to initialize the model.

Table 3.9: Initial values for the crosstalk model.

<i>Var.</i>	Species Name	<i>Initial Value (nmol/mL)</i>
x_1	GMCSF	$0.4 \cdot 10^{-6}$
x_2	GMCSF_RC	$7.957934 \cdot 10^{-4}$
x_3	GMCSF_LRC	0
x_4	GMCSF_LRC_blocked	0
x_5	IL6	$0.4 \cdot 10^{-6}$
x_6	IL6_RC	$7.957934 \cdot 10^{-3}$
x_7	IL6_LRC	0
x_8	STAT3	$1.454126 \cdot 10^{-1}$
x_9	pSTAT3	$4.511127 \cdot 10^{-4}$
x_{10}	pSTAT3_nuc	0
x_{11}	SK	$3.3 \cdot 10^{-2}$
x_{12}	aSK	0
x_{13}	SOCS3_mRNA	0
x_{14}	SOCS3	0
x_{15}	IL6_LRC_SOCS3	0
x_{16}	GMCSF_LRC_SOCS3	0

3.7.5 Preliminary model calibration

The phospho-STAT-3 measurements upon the experimental conditions *IL6+*, *IL6+/GMCSF+*, *IL6+/GMCSF-* and *unstim* were used to calibrate the model by applying the parameter estimation techniques described in chapter 2.

As figure 3.20 shows, the model is able to reproduce the measurements of phospho-STAT-3 in the cytoplasm both qualitatively and quantitatively. The calibrated model captures the cellular behaviour in all three stimulation settings *IL6+*, *IL6+/GMCSF+*, *IL6+/GMCSF-* and in the unstimulated control experiment *unstim* with a *single set* of parameters.

In the simulations, peak levels of cytosolic phospho-STAT-3 are attained about 20–30 minutes after stimulation. This is in good concordance to the experimental observations. Also, the simulated peak levels are close to the actual measurements, as far as the measurements' time grid permits a conclusion.

Table 3.10 lists the parameter values that have been the result of the parameter estimation, and have been used to create the simulations in figure 3.20. Also compare the final results in table 3.12 and figure 3.23.

3.7.6 A confirmed prediction: SOCS-3 kinetics as model validation

The least a model should be able to do is to explain (or, say, reproduce) already existent data. As shown in the previous section, our model complies with this requirement. We now take a step forward and use the model to predict the kinetics of other parts of the studied pathway, that we did not have observed until now.

We choose to forecast and measure the SOCS-3 protein in the cytosol, that is induced after STAT-3 activation (see section 3.2.3.1). This protein is chosen for three reasons: First, there is a one-to-one correspondence between the “real” SOCS-3 molecule in the living cell and in our HaCaT keratinocyte model. Second, its basal level is low, and protein neogenesis

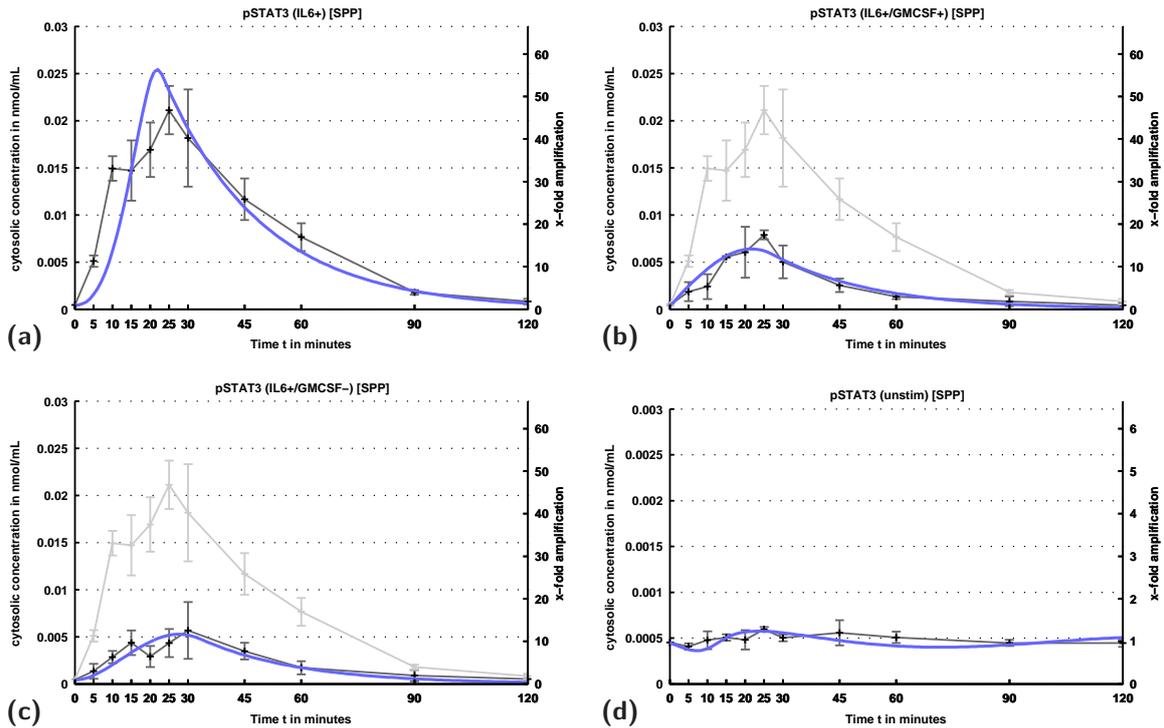


Figure 3.20: Preliminary model calibration using solely pSTAT-3 measurements.

The calibrated model reproduces the measurements of cytosolic phospho-STAT-3 in the three stimulation experiments (a) *IL6+*, (b) *IL6+/GMCSF+*, and (c) *IL6+/GMCSF-*, as well as in the unstimulated control experiment *unstim* (d), both quantitatively and qualitatively, with a unique set of parameters.

The pictures show the simulated pSTAT-3 time course (thick blue line) of the respective experimental setting, together with the linearly interpolated mean values of triplicate measurements, with error bars from empirical standard deviation. For comparison, the course of sole IL-6 stimulation (a) is copied to (b) and (c) in light gray colours. Note the different scaling in (d).

For model calibration, the cytosolic phospho-STAT-3 measurements (see table 3.3) have been used exclusively. Compare section 3.7.6 and figure 3.23, where also SOCS-3 measurements, which have been taken in subsequent experiments, are considered.

is inducible by IL-6 stimulation. And last, and this point is essential from the experimenter's point of view, it can be detected using well established laboratory protocols.

Using the preliminary parameter set given in table 3.10, we simulate the time course of cytosolic SOCS-3 in the sole IL-6 stimulation setting *IL6+*. Section 3.7.6 visualizes the prediction of cytosolic SOCS-3 concentration after stimulation with IL-6. While, formally, the model gives a forecast in molar concentration, we adopt a cautious attitude and suggest to interpret the prediction in a qualitative manner. What can be said, is, that up to 15–20 minutes after IL-6 stimulation, only a negligible amount of SOCS-3 protein will be produced, but for the following 20 minutes, a fast rise of protein concentration is predicted, with a peak level between 45 and 60 minutes after stimulation. Then, SOCS-3 protein concentration is slowly declining to medium levels until the end of the observation horizon of 120 minutes.

To validate our model, we perform a further set of IL-6 stimulation experiments (*IL6+*) and quantify the induced SOCS-3 protein by quantitative immunoblotting (section 3.2.5). As section 3.7.6 shows, the actual measurements corroborate the predicted kinetics and validate our model.

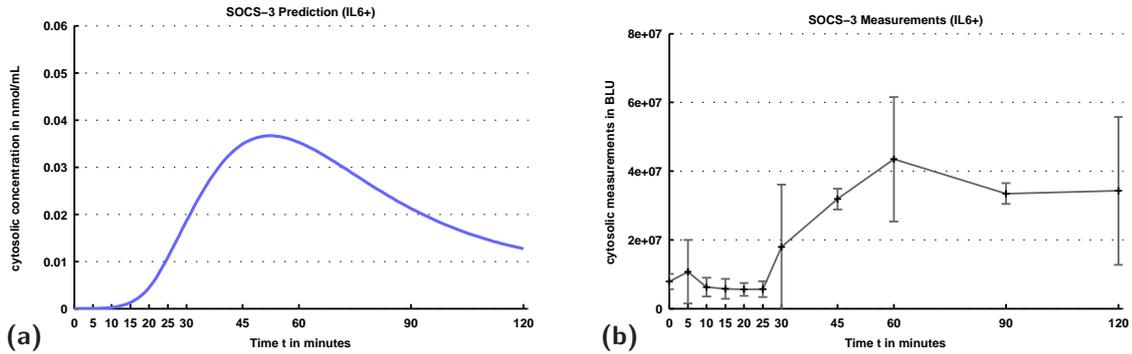


Figure 3.21: Prediction and measurements of cytosolic SOCS-3 upon IL-6 stimulation.

(a) Forecast of cytosolic SOCS-3 concentration in the *IL6+* experimental setting.

(b) Measurements of cytosolic SOCS-3 in relative light units (BLU) by quantitative immunoblotting.

The picture on the right shows linearly interpolated mean values of triplicate measurements, with error bars originating from empirical standard deviation.

Table 3.10: Kinetic parameters used for SOCS-3 prediction.

Estimates obtained from parameter estimation using data of triplicate measurements of cytosolic phospho-STAT-3 in the four experimental settings *IL6+*, *IL6+/GMCSF+*, *IL6+/GMCSF-*, and *unstim*.

The parameter estimates above were used for predicting the time course of cytosolic SOCS-3 protein upon IL-6 stimulation (see section 3.7.6).

Par.	Estimate	Unit	Meaning
p_1	$5.77 \cdot 10^{-7}$	$\frac{nmol}{mL \cdot min}$	production rate of IL-6 and GM-CSF
p_2	$2.2 \cdot 10^1$	$\frac{mL}{nmol \cdot min}$	activation rate of IL-6 receptor and GM-CSF receptor upon ligand binding
p_3	$1.00 \cdot 10^5$	$\frac{mL}{nmol \cdot min}$	blockade of GM-CSF receptor upon overstimulation
p_4	$7.27 \cdot 10^4$	$\frac{mL}{nmol \cdot min}$	activation rate of SK by active GM-CSF receptor
p_5	$2.84 \cdot 10^4$	$\frac{mL^2}{nmol^2 \cdot min}$	SK enhanced STAT-3 activation on active IL-6 receptor
p_6	$1.00 \cdot 10^{-1}$	min^{-1}	degradation rate of SOCS-3 in cytosol
p_7	$3.04 \cdot 10^3$	$\frac{mL}{nmol \cdot min}$	deactivation rate of IL-6 and GM-CSF receptor by SOCS-3
p_8	$5.08 \cdot 10^{-1}$	$\frac{mL}{nmol \cdot min}$	STAT-3 phosphorylation rate on active IL-6 receptor
p_9	$3.83 \cdot 10^{-2}$	min^{-1}	translocation of phospho-STAT-3 into nucleus
p_{10}	$4.16 \cdot 10^{-2}$	min^{-1}	export rate of nuclear STAT-3 into cytoplasm
p_{11}	$4.76 \cdot 10^{-2}$	min^{-1}	induction and export rate of SOCS-3 mRNA by nuclear phospho-STAT-3
p_{12}	$1.40 \cdot 10^{-1}$	min^{-1}	translation rate of SOCS-3 mRNA in cytoplasm (10 proteins from 1 mRNA)

Table 3.11: Time course data of cytosolic SOCS-3 upon IL-6 stimulation.
Measurement values in relative units (BLU).

Time	rep #1	rep #2	rep #3
0	8700000	4870000	10100000
5	1480000	7410000	23300000
10	2440000	8600000	7780000
15	2560000	5140000	9580000
20	3540000	5110000	8050000
25	4370000	8910000	3690000
30	8540000	1950000	43300000
45	30600000	36000000	29000000
60	63200000	47700000	19500000
90	29300000	34800000	36300000
120	31000000	9690000	62100000

Time course of SOCS-3 in setting *IL6+*.

3.7.7 Re-calibration using SOCS-3 data

Detection and quantification of SOCS-3 protein is done in the same way as the detection of pSTAT-3 by using (semi-)quantitative immunoblotting, yielding relative light units (BLU) proportional to the molecular concentration but with an unknown and blot-dependent proportionality constant (parameter p_{SOCS}).

Since we do not have a reliable estimate for the initial SOCS-3 concentration, we cannot use the amplification factor method (see section 3.5.4) to calculate absolute concentrations from the measurements. We therefore implement a traditional method and introduce a scaling parameter that maps the BLU values onto molar concentration, and try to estimate this scaling parameter from measurement data. Introducing three separate scaling parameters, one for each blot as the pure teaching would require, would increase the uncertainty in all parameters without any benefits for the model calibration. The restriction to a single scaling parameter for the three replicates is justified by the observation, that the BLU measurements are not only in the same order of magnitude, but also coincide in the nominal values for most time points.

The additional quantitative SOCS-3 measurements can be used to further improve the estimates of the kinetic parameters. Table 3.12 lists the parameter values of the re-calibrated model, together with their standard error (diagonal elements of the covariance matrix approximation), computed by the methods presented in section 2.5 on page 38.

Figure 3.23 on the next page compares the cytosolic phospho-STAT-3 data to the simulation of the re-calibrated model. By visual inspection, it is very close to the preliminary fit not using SOCS-3 data as in figure 3.20, while there are some subtle changes. The most prominent difference is observed in the simulated pSTAT-3 time course upon sole IL-6 stimulation (setting *IL6+*). The peak amplification factor is reduced from 57-fold in the preliminary towards 52-fold in the final fit, and thus moved closer to the experimentally observed mean factor of 47-fold. In figure 3.22, the model simulation of cytosolic SOCS-3 after model re-calibration is shown to be in good concordance to the wet lab measurements.

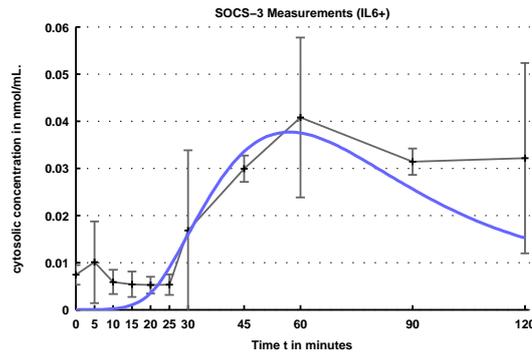


Figure 3.22: Model fit of cytosolic SOCS-3 upon IL-6 stimulation.

Simulated time course of IL-6-induced cytosolic SOCS-3 (blue line), and linearly interpolated mean values of triplicate measurements, with error bars originating from empirical standard deviation.

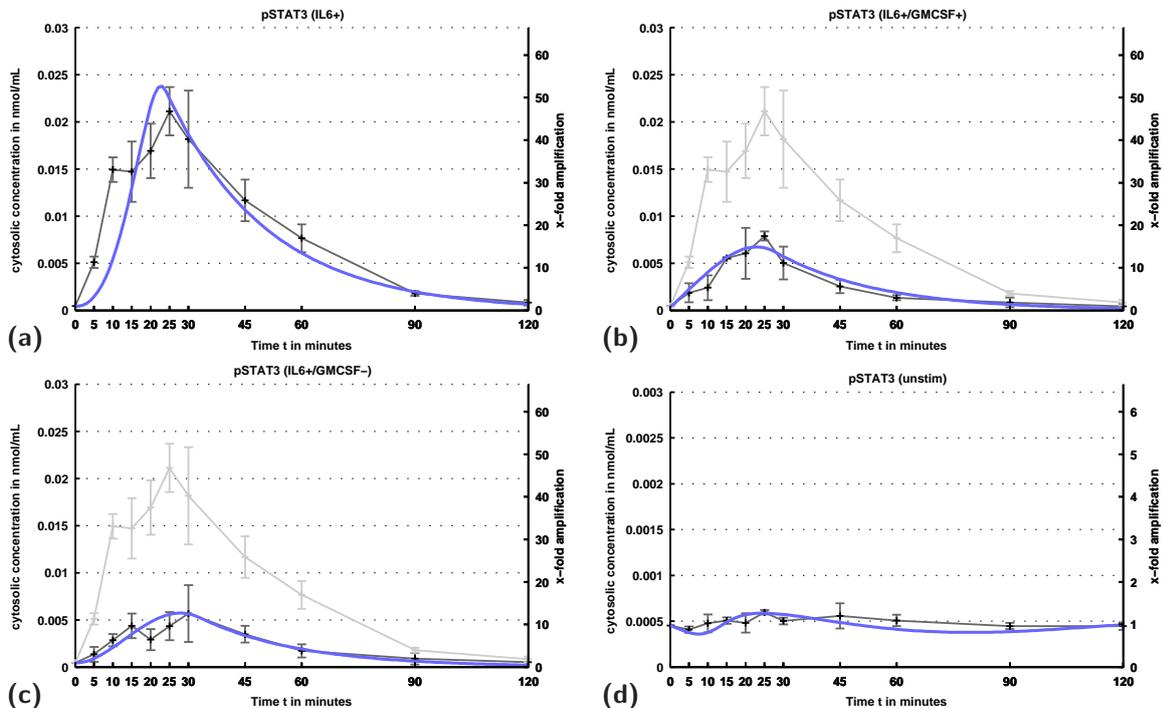


Figure 3.23: Final model calibration using pSTAT-3 and SOCS-3 measurements from four experimental settings.

The calibrated model reproduces the measurements of cytosolic phospho-STAT-3 in the three stimulation experiments *IL6+* (a), *IL6+/GMCSF+* (b), and *IL6+/GMCSF-* (c), as well as in the unstimulated control experiment *unstim* (d), both qualitatively and quantitatively.

The pictures show the simulated pSTAT-3 time course (thick blue line) of the respective experimental setting, together with the linearly interpolated mean values of triplicate measurements, with error bars from empirical standard deviation. For comparison, the course of sole IL-6 stimulation (a) is copied to (b) and (c) in light gray colours. Note the different scale in (d).

For model calibration, the cytosolic phospho-STAT-3 measurements (see table 3.3) as well as the SOCS-3 measurements (table 3.11) have been used. Compare figure 3.20 for a preliminary fit using solely pSTAT-3 measurements.

Table 3.12: Final estimates of kinetic parameters.

Estimates obtained from parameter estimation using the data of triplicate measurements of cytosolic phospho-STAT-3 in the four experimental settings *IL6+*, *IL6+/GMCSF+*, *IL6+/GMCSF-*, and *unstim*, as well as the data of triplicate measurements of cytosolic SOCS-3 in the *IL6+* setting.

The mark *indeterminate* denotes estimated parameters that have a large confidence interval because they have practically no measurable influence on the observed quantities. The parameter marked with * showed an ever increasing value in parameter estimation without affecting the result and was therefore fixed to a “sufficiently large” value. The parameter marked with ** (SOCS degradation) is highly coupled with the parameter describing the SOCS transcription (a higher degradation rate can be counterbalanced with a higher transcription rate) and is thus fixed to the denoted value. All values are rounded to 3 digits; non-rounded ones are found in the appendix (table A.2). SE: Standard Error.

Par.	Estimate	Unit	SE	Meaning
p_1	$5.77 \cdot 10^{-7}$	$\frac{nmol}{mL \cdot min}$	<i>indeterminate</i>	production rate of IL-6 and GM-CSF
p_2	$1.82 \cdot 10^1$	$\frac{mL}{nmol \cdot min}$	9.6 %	activation rate of IL-6 receptor and GM-CSF receptor upon ligand binding
p_3	$1.00 \cdot 10^5$	$\frac{mL}{nmol \cdot min}$	<i>indeterminate*</i>	blockade of GM-CSF receptor upon overstimulation
p_4	$7.27 \cdot 10^4$	$\frac{mL}{nmol \cdot min}$	<i>indeterminate</i>	activation rate of SK by active GM-CSF receptor
p_5	$3.17 \cdot 10^4$	$\frac{mL^2}{nmol^2 \cdot min}$	39.8 %	SK enhanced STAT-3 activation on active IL-6 receptor
p_6	$1.00 \cdot 10^{-1}$	min^{-1}	<i>indeterminate**</i>	degradation rate of SOCS-3 in cytosol
p_7	$3.00 \cdot 10^3$	$\frac{mL}{nmol \cdot min}$	30.2 %	deactivation rate of IL-6 and GM-CSF receptor by SOCS-3
p_8	$6.13 \cdot 10^{-1}$	$\frac{mL}{nmol \cdot min}$	10.2 %	STAT-3 phosphorylation rate on active IL-6 receptor
p_9	$3.75 \cdot 10^{-2}$	min^{-1}	5.3 %	translocation of phospho-STAT-3 into nucleus
p_{10}	$2.15 \cdot 10^{-2}$	min^{-1}	43.4 %	export rate of nuclear STAT-3 into cytoplasm
p_{11}	$3.97 \cdot 10^{-2}$	min^{-1}	17.3 %	induction and export rate of SOCS-3 mRNA by nuclear phospho-STAT-3
p_{12}	$1.40 \cdot 10^{-1}$	min^{-1}	<i>indeterminate</i>	translation rate of SOCS-3 mRNA in cytoplasm (10 proteins from 1 mRNA)
p_{SOCS}	$9.38 \cdot 10^{-10}$	$\frac{nmol}{mL \cdot BLU}$	11.5 %	scaling constant for SOCS-3 measurement data

Table 3.13: Phospho-STAT-3 peak levels in measurements, preliminary, and final calibrated model. Maximum cytosolic phospho-STAT-3 concentration upon the three stimulation experiments *IL6+*, *IL6+/GMCSF+*, and *IL6+/GMCSF-*, together with their respective time points and amplification factors. The peak time ranges of the measurements consists of the time points of the samples taken one before and one after the maximum. For the simulations, the peak time points rounded to minutes are given. The peak concentrations of the measurement is the mean of triplicate measurements with empirical standard deviation, the same holds for the peak amplification factors of the measurements. The simulated data displayed at *prelim. fit* and *final fit* was created using the model with the preliminary parameter set given in table 3.10 and the final parameter set given in table 3.12. The simulation of the *IL6+/GMCSF-* setting shows a slightly delayed time peak compared to the other two stimulation settings (peak level is attained 28 minutes after stimulation compared to 22 minutes). This seems to be supported by measurement data, however, due to the restricted time resolution, this cannot be stated to be a significant finding.

<i>IL6+</i>	<i>unit</i>	<i>measurement</i>	<i>prelim. fit</i>	<i>final fit</i>
peak time	min	20–30	≈ 22	≈ 23
peak concentration	pmol/mL	21.12 ± 2.55	25.42	23.76
peak mean amplification factor ^(*)	x-fold	46.8 ± 5.7	56.4	52.7
mean peak amplification factor ^(**)	x-fold	47.4 ± 6.0	56.4	52.7
<i>IL6+/GMCSF+</i>				
peak time	min	20–30	≈ 22	≈ 23
peak concentration	pmol/mL	7.89 ± 0.47	6.40	6.72
peak mean amplification factor	x-fold	17.5 ± 1.0	14.2	14.9
mean peak amplification factor	x-fold	17.9 ± 1.3	14.2	14.9
<i>IL6+/GMCSF-</i>				
peak time	min	25–45	≈ 28	≈ 28
peak concentration	pmol/mL	5.66 ± 3.00	5.26	5.72
peak mean amplification factor	x-fold	12.6 ± 6.7	11.7	12.7
mean peak amplification factor	x-fold	13.6 ± 7.8	11.7	12.7

(*) The *peak mean* amplification factor is the maximum amplification factor observed over the mean values of each time point · (**) The *mean peak* amplification factor denotes the mean value of the maximum amplification factors per replicate (which might occur at different time points).

3.8 New biological insight: Dose dependency on GM-CSF in IL-6-induced phospho-STAT-3 signalling

A calibrated model allows us to make simulated experiments, and to make predictions about the actual outcome, with the only cost being the computational effort. For a medium-sized model like our HaCaT cell model, even this is negligible. On the other hand, real wet lab experiments can easily cost hundreds of Euros each.

As seen from the *IL6+/GMCSF+* experiments (section 3.6, table 3.3a), co-stimulation with IL-6 and GM-CSF leads to a lower response in cytosolic phospho-STAT-3 compared to the experiments with sole IL-6 stimulation, in which the (very low) physiological GM-CSF level was maintained.

We therefore decided to predict the kinetics of cytosolic phospho-STAT-3 upon co-stimulation with a fixed concentration of IL-6 (100ng/mL) and different doses of additional GM-CSF (concentrations of 10pg/mL, 100pg/mL, 1ng/mL, 10ng/mL, and 100ng/mL) covering the range between physiological and maximum experimental concentrations, to investigate whether there is an “optimal” level of additional GM-CSF that results in a maximum pSTAT-3 response – an intuitive expectation concerning the facts that both co-stimulation with as well as blockade of GM-CSF leads to a decreased IL-6-induced STAT-3 activation.

3.8.1 Predicted GM-CSF dose dependency

Figure 3.24 depicts the predicted time course of cytosolic phospho-STAT-3 upon co-stimulation with a fixed amount of IL-6 and different doses of GM-CSF. For comparison, the time courses upon sole IL-6 stimulation as well as with simultaneous blockade of GM-CSF and the unstimulated control (see section 3.7.7) are re-printed. As one can see from the image, a negative correlation between additional GM-CSF and phospho-STAT-3 response is predicted. None of the simulated co-stimulation experiments shows a higher response in STAT-3 activation than in the sole IL-6 stimulation setting. A somehow counter-intuitive prediction, that was confirmed when the actual wet lab experiments were performed (see figure 3.26).

As figure 3.25 shows, the maximum level of signalling competent and active IL-6-bound gp130 ligand-receptor-complexes is reached 15 minutes after stimulation in the co-stimulation experiment *IL6+/GMCSF+*. This time point was chosen for the experimental verification of the predicted nonlinear dependency of IL-6-induced phospho-STAT-3 on GM-CSF concentration.

The model predicted an 22-fold rise in cytosolic phospho-STAT-3 15 minutes after stimulation with 100ng/mL IL-6 and 100pg/mL of GM-CSF. After co-stimulation with 10ng/mL GM-CSF, the simulation yields an 16.8-fold increase in cytosolic phospho-STAT-3. A simulated co-stimulation with 1ng/mL results in an 15.6-fold increase of phospho-STAT-3 in the cytosol, whereas a co-stimulation with only 100pg/mL of additional GM-CSF is predicted to yield an only 12.4-fold increase in the cytosolic concentration of phospho-STAT-3.

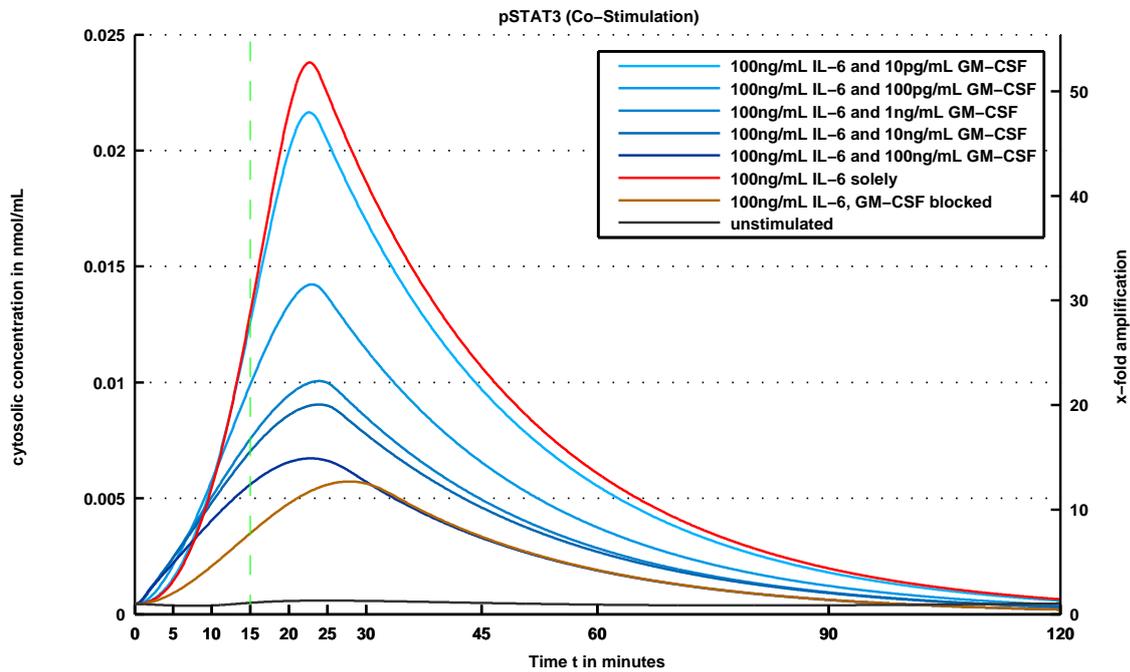


Figure 3.24: Prediction of time course in co-stimulation experiments. Depicted are the predicted time courses of cytosolic phospho-STAT-3 upon co-stimulation with IL-6 (100ng/mL) and different doses of GM-CSF (shades of blue). For comparison, the time-course of the unstimulated control (black), sole IL-6 stimulation (red), as well as IL-6 stimulation while blocking GM-CSF (brown) is displayed. The time point chosen for wet lab confirmation is illustrated as green dashed line.

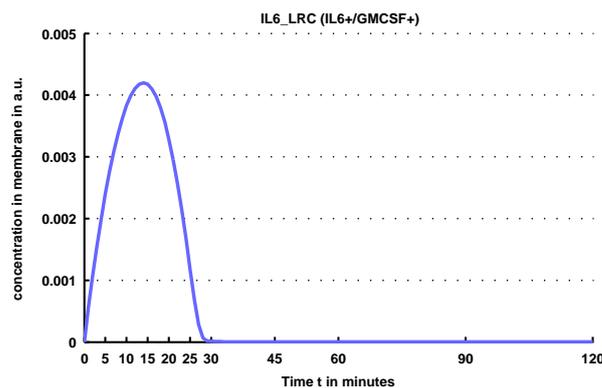


Figure 3.25: Predicted time course of activated IL-6/gp130 ligand receptor complex. Simulation of the activated ligand-receptor-complex in the co-stimulation experiment *IL6+/GMCSF+*. The maximum level of active signalling IL-6-bound gp130 receptors is reached 15 minutes after stimulation with 100ng/mL of IL-6 and GM-CSF each.

3.8.2 Wet lab verification of GM-CSF dose dependency in IL-6-induced phospho-STAT-3 signalling

The outcome of the co-stimulation experiments series is depicted in figure 3.26. Each group of bars belongs to prediction and wet lab measurements of the amplification factors in cytosolic phospho-STAT-3 for a different experimental setting. From left to right, the concentration of GM-CSF increases, whereas the IL-6 concentration is constant. For comparison, also the measurements of the *IL6+* setting (sole IL-6 stimulation keeping the physiological GM-CSF level) as well as the ones of the *IL6+/GMCSF-* setting (sole IL-6 stimulation while blocking GM-CSF) are given.

The model predicted amplification factors after 15 minutes are depicted as light blue boxes (□) in figure 3.26. Qualitatively, the measurements depicted as dark green boxes (■) are in good concordance with the prediction.

By delaying “model time” by only 3 minutes, we actually get a good *quantitative* concordance between prediction and outcome. The predictions of amplification factors of cytosolic phospho-STAT-3 12 minutes after stimulation are depicted as blue boxes (■) in figure 3.26. The displayed amplification factors are mean values originating from triplicate measurements.

Co-stimulation with a GM-CSF concentration of 100ng/mL yields a 11.3-fold increase in cytosolic pSTAT-3, close to the prediction of a 10.5-fold increase. The observed rise to an 11.6-fold level of pSTAT-3 in the co-stimulation experiment with 10ng/mL GM-CSF confirms the predicted 12.7-fold rise. Furthermore, with a prediction of 13.5-fold increase and an observed 13.1-fold rise in cytosolic pSTAT-3 concentration, the model forecasts the outcome pretty well. Last, in the experiment with 100pg/mL additional GM-CSF, where a 16.4-fold increase is predicted, we observe a 14.9-fold increase in cytosolic pSTAT-3.

3.9 Conclusion

Our mathematical modelling revealed a nonlinear dose dependency on GM-CSF in IL-6-induced phospho-STAT-3 signalling.

We predicted and experimentally verified (see figure 3.26) that the physiological GM-CSF level is correlated to highest activation levels of cytosolic phospho-STAT-3 levels upon IL-6 stimulation.

3.8.3 Interpretation and outlook

What might be a next step? Our model enables us to simulate, e.g., knockout studies and other interventions to specifically manipulate cellular levels of pSTAT-3. Such a study, however, can not be solved but only be assisted by mathematical means, as it heavily relies on expert knowledge of the biological interconnections to other pathways, proteins and signal transducers inside and outside of cells.

Elevated levels of pSTAT-3 are known to be present in many human cancer types [JohnstonGrandis2011]. Further, the usage of GM-CSF in cancer therapy has been and currently still is investigated, see i.e. [Kanerva2013]. As mentioned in the introduction (sections 3.1 and 3.2 on page 42ff), aberrant and persistent activation of pSTAT-3 is frequently found in cancer [Ernst2008; Buettner2002; Yu2014].

With our model, we have established a connection between the two cytokines GM-CSF and IL-6 in JAK-STAT signalling, based on and supported by *in vitro* experiments and data. We were able to prove the mediating effect of GM-CSF on IL-6-induced pSTAT-3, and developed and validated a quantitative model for the occurring interactions, that may be used to predict and strategically investigate the mediating effects of GM-CSF on pSTAT-3 activation, making it possible to formulate and address new questions.

The discovered nonlinear influence of GM-CSF on IL-6-induced pSTAT-3 signalling gives a promising target for the development of new therapeutic intervention methods, and further research can be assisted by mathematical modelling.

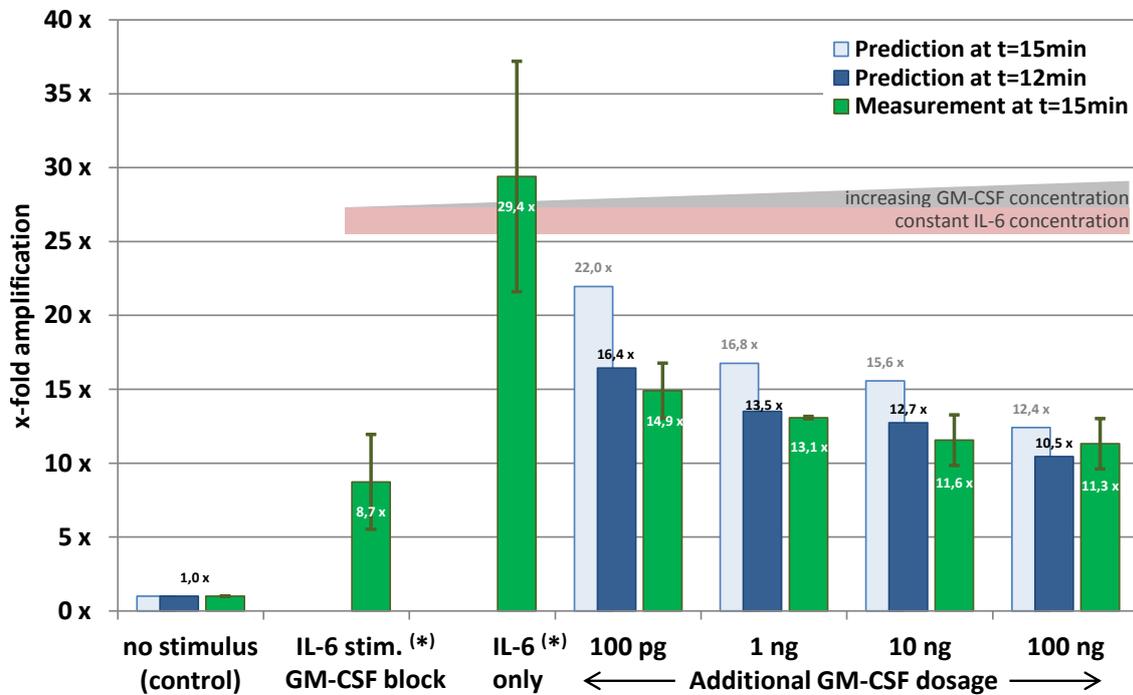


Figure 3.26: Dose dependency on GM-CSF in IL-6-induced phospho-STAT-3 signalling.

Phospho-STAT-3 induction upon co-stimulation with 100 ng/mL IL-6 and different doses of additional GM-CSF (increasing from left to right), as well as an unstimulated control (first column) and phospho-STAT-3 induction upon IL-6 stimulation while blocking GM-CSF with appropriate antibodies (second column). Model prediction (□ light blue) shows a negative correlation between *additional* GM-CSF stimulus and IL-6-induced cytosolic phospho-STAT-3. This negative correlation has been confirmed in the wet lab experiments (■ dark green). Delaying “model time” by 3 minutes gives a good quantitative concordance between prediction (■ blue) and actual measurements.

In the stimulation experiments, measurements are taken 15 minutes after stimulation. Depicted are mean values from triplicate experiments with empirical standard deviation as error bars.

(*) The *IL6+/GMCSF-* experiment (second column) corresponds to an overall GM-CSF level of 0 (zero) pg/mL, as GM-CSF is blocked by appropriate antibodies. In the *IL6+* experiment (third column), the physiological GM-CSF level of about 10 pg/mL is maintained. Thus, there is experimental evidence that the physiological level of GM-CSF is correlated to highest activation of cytosolic phospho-STAT-3 upon IL-6 stimulation.

How to read the figure: For the co-stimulation experiment with 100ng/mL IL-6 and 1ng/mL GM-CSF (fifth column), the model predicted a 16.8-fold rise in cytosolic phospho-STAT-3 15 minutes after stimulation. The experimental verification yielded an actual mean value of 13.1-fold increase. Delaying the model time by 3 minutes results in a good quantitative concordance of 13.5-fold increase (model prediction at 12 minutes after stimulation).

4 Introduction to Stochastic Differential Equations

In this chapter, we recapitulate the basics of measure theory and stochastic processes that are required to introduce the concept of WIENER processes. Though these results may be found in any textbook on probability, we restate them not only for the sake of completeness, but also to establish the notation used in this thesis.

Following the introduction by [Oksendal1998], we develop stochastic differential equations by modifying an ODE population growth example, affecting it by some noise. A formal introduction and definition of the underlying stochastic ITÔ⁽¹⁾ integrals is given in appendix C.

We further discuss numerical integration schemes, especially the EULER-MARUYAMA stochastic integrator, and their convergence properties.

To conclude this chapter, we give an introduction to SDE driven by LÉVY processes with jumps and their numerical integration.

Main references for the presented results in this chapter are the textbooks of ØKSENDAL [Oksendal1998], KLOEDEN and PLATEN [KloedenPlaten1995], and PROTTER [Protter2004].

4.1 Basic definitions and results

4.1.1 Random variables

4.1 Definition (σ -Algebra)

Let Ω be a set. A family $\mathcal{A} \subseteq \mathcal{P}(\Omega)$ of subsets of Ω , is called a σ -algebra, if it holds

$$\text{i) } \Omega \in \mathcal{A} \quad \text{ii) } A \in \mathcal{A} \implies A^c \in \mathcal{A} \quad \text{iii) } A_n \in \mathcal{A} \quad \forall n \in \mathbb{N} \implies \bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$$

where $A^c := \Omega \setminus A$ denotes the *complement* of A . □

4.2 Lemma

Let \mathcal{A} be a σ -algebra on the set Ω . Then, it holds:

1. $\emptyset \in \mathcal{A}$
2. $A_n \in \mathcal{A} \quad \forall n \in \mathbb{N} \implies \bigcap_{n=1}^{\infty} A_n \in \mathcal{A}$
3. $A, B \in \mathcal{A} \implies A \setminus B \in \mathcal{A}$
4. The intersection of any two σ -algebras is again a σ -algebra. □

⁽¹⁾ Kiyosi ITÔ (伊藤 清), 1915–2008, Japanese mathematician and pioneer of stochastic integration and stochastic differential equations. While the Hepburn romanization of his name is “Kiyoshi ITÔ” (with a macron), he himself preferred the spelling ITÔ with a circumflex (Kunrei-shiki romanization).

4.3 Definition (Generating σ -algebra)

Let $\Sigma \subseteq \mathcal{P}(\Omega)$ be a family of subsets of Ω . The smallest σ -algebra on Ω that contains Σ is called the σ -algebra generated by Σ , and is written as:

$$\sigma(\Sigma) := \bigcap \{ \mathcal{A} \in \mathcal{P}(\Omega) \mid \Sigma \subset \mathcal{A} \text{ and } \mathcal{A} \text{ is a } \sigma\text{-algebra} \}. \quad \square$$

4.4 Definition (Measurable space, measure, measure space, probability space, completeness)

Let \mathcal{A} be a σ -algebra on Ω .

1. The pair (Ω, \mathcal{A}) is called a *measurable space*, and the sets $A \in \mathcal{A}$ are called the *\mathcal{A} -measurable sets of Ω* .
2. A function $\mu : \mathcal{A} \rightarrow [0, \infty]$ is called a *measure* on the measurable space (Ω, \mathcal{A}) , if

$$(i) \quad \mu(\emptyset) = 0 \qquad (ii) \quad \mu \left(\bigcup_{n=1}^{\infty} A_n \right) = \sum_{n=1}^{\infty} \mu(A_n) \text{ for disjoint sets } A_n \in \mathcal{A}.$$

Then, for $A \in \mathcal{A}$, the value $\mu(A)$ is called the *measure* of A .

3. The triple $(\Omega, \mathcal{A}, \mu)$ is called a *measure space*.
4. A measure space $(\Omega, \mathcal{A}, \mathbb{P})$ with $\mathbb{P}(\Omega) = 1$ is called a *probability space*, and the measure \mathbb{P} is called a *probability measure*. Further, for $A \in \mathcal{A}$, the value $\mathbb{P}(A)$ is called the *probability* of A .
5. A probability space is said to be *complete*, if any subset of a set with probability 0 is included in the underlying σ -algebra, i.e. $\forall B \subset A : A \in \mathcal{A}, \mathbb{P}(A) = 0 \implies B \in \mathcal{A}$ □

A probability space $(\Omega, \mathcal{A}, \mathbb{P})$ is the underlying mathematical construct of probability theory to model a random experiment. Here, the set Ω is the *sample space* that contains all possible *outcomes*, whereas the σ -algebra \mathcal{A} denotes a set of events that we are interested in. An *event* is a set of one or multiple outcomes⁽²⁾. Finally, the probability measure \mathbb{P} endows each event $A \in \mathcal{A}$ with a certain probability $\mathbb{P}(A)$.

4.5 Definition (Measurable function, generated σ -algebra)

1. Let $f : \Omega_1 \rightarrow \Omega_2$ be a map between two measurable spaces $(\Omega_1, \mathcal{A}_1)$ and $(\Omega_2, \mathcal{A}_2)$. Then, f is called a *measurable function*, if the preimage of every measurable set in $(\Omega_2, \mathcal{A}_2)$ is measurable in $(\Omega_1, \mathcal{A}_1)$, i.e. if

$$f^{-1}(A) \in \mathcal{A}_1 \quad \text{for every } A \in \mathcal{A}_2,$$

and then we write $f : (\Omega_1, \mathcal{A}_1) \rightarrow (\Omega_2, \mathcal{A}_2)$.

2. Let $g : \Omega_1 \rightarrow \Omega_2$ be a map, and \mathcal{A}_2 be a σ -algebra on Ω_2 . Then we define $\sigma(g)$, the *σ -algebra generated by g* , by

$$\sigma(g) := \{g^{-1}(A) : A \in \mathcal{A}_2\}. \quad \square$$

⁽²⁾ Using the common “throwing once a dice” example, the sample space of possible outcomes may be chosen as $\Omega = \{1, 2, 3, 4, 5, 6\}$. The event “throwing a three” is then described by the set $\{3\}$, the event “throwing an even number” by $\{2, 4, 6\}$.

4.6 Lemma

Let $g : \Omega_1 \rightarrow \Omega_2$ be a map, \mathcal{A}_2 be a σ -algebra on Ω_2 , and $\sigma(g)$ the σ -algebra generated by g . Then, it holds:

1. $\sigma(g)$ is a σ -algebra on Ω_1 .
2. $\sigma(g)$ is the smallest σ -algebra on Ω_1 , such that $g : (\Omega_1, \sigma(g)) \rightarrow (\Omega_2, \mathcal{A}_2)$ is measurable.

□

4.7 Definition (Topological space, topology)

Let \mathcal{X} be a set, and $\mathcal{T} \subseteq \mathcal{P}(\mathcal{X})$ a family of subsets of \mathcal{X} . The pair $(\mathcal{X}, \mathcal{T})$ is called a *topological space*, if

- i) $\mathcal{X} \in \mathcal{T}$ and $\emptyset \in \mathcal{T}$
- ii) $A, B \in \mathcal{T} \implies A \cap B \in \mathcal{T}$
- iii) $\mathcal{S} \subset \mathcal{T} \implies \bigcup_{A \in \mathcal{S}} A \in \mathcal{T}$

Then, \mathcal{T} is called a *topology* on \mathcal{X} and we may write it as the set of all open subsets of \mathcal{X} , i.e. $\mathcal{T} = \{X \subseteq \mathcal{X} \mid X \text{ is an open set}\}$.

□

4.8 Definition (BOREL σ -algebra)

Let $(\mathcal{X}, \mathcal{T})$ be a topological space.

1. The *BOREL σ -algebra* \mathbb{B} is the σ -algebra that is generated by the open sets of \mathcal{X} :

$$\mathbb{B} := \mathbb{B}(\mathcal{X}) := \sigma(\{A \subset \mathcal{X} \mid A \text{ is open}\}),$$

i.e. the BOREL σ -algebra is the smallest σ -algebra containing the open sets, and an element $B \in \mathbb{B}$ is called a *BOREL set*.

2. As a convention, a topological space is assumed to be endowed with its BOREL σ -algebra \mathbb{B} , if not otherwise specified.
3. We write \mathbb{B}^n for the BOREL sets in \mathbb{R}^n , and \mathbb{B} for the BOREL sets in \mathbb{R} .
4. A measurable function $X : (\Omega, \mathcal{A}) \rightarrow (\mathbb{R}^n, \mathbb{B}^n)$ is called BOREL measurable.
5. Let $(\mathbb{R}^n, \mathbb{B}^n, \lambda)$ be a measure space. If the measure λ has the property

$$\lambda([a_1, b_1] \times [a_2, b_2] \times \dots \times [a_n, b_n]) = (b_1 - a_1) \cdot (b_2 - a_2) \cdot \dots \cdot (b_n - a_n)$$

then it is called the *LEBESGUE-BOREL measure*⁽³⁾.

□

⁽³⁾ The LEBESGUE measure is the complete measure built from the LEBESGUE-BOREL measure by adding to \mathbb{B}^n all sets $A \subset \mathbb{R}^n$ with $\underline{A} \subset A \subset \overline{A}$, where $\underline{A}, \overline{A} \in \mathbb{B}^n$ and $\lambda(\overline{A} \setminus \underline{A}) = 0$. The corresponding sets are called LEBESGUE measurable and form the LEBESGUE σ -algebra (which is in some sense much larger than the BOREL σ -algebra).

In what follows, the triple $(\Omega, \mathcal{A}, \mathbb{P})$ denotes a complete probability space if not otherwise specified.

4.9 Definition and Lemma (Random variable, induced probability measure)

1. Let $(\Omega_1, \mathcal{A}_1, \mathbb{P})$ be a complete probability space and $(\Omega_2, \mathcal{A}_2)$ be a measurable space. A measurable map $X : (\Omega_1, \mathcal{A}_1) \rightarrow (\Omega_2, \mathcal{A}_2)$ is then called a $(\Omega_2, \mathcal{A}_2)$ -valued *random variable* on Ω_1 .
2. For abbreviation, if $X : (\Omega_1, \mathcal{A}_1) \rightarrow (\mathbb{R}^n, \mathbb{B}^n)$, instead of talking about an X as an \mathbb{R}^n -valued random variable on Ω_1 , we denote it for short as a *random variable*.
3. Let $X : (\Omega_1, \mathcal{A}_1) \rightarrow (\mathbb{R}^n, \mathbb{B}^n)$ be a random variable. Then, the *distribution* of X is defined by $\mu_X(B) := \mathbb{P}(X^{-1}(B))$ for $B \in \mathbb{B}^n$, and μ_x is a probability measure on $(\mathbb{R}^n, \mathbb{B}^n)$.

□

4.10 Definition (Indicator function)

Let Ω be a set, $A \subseteq \Omega$ a subset, and $x \in \Omega$ an element. The *indicator function* $\mathbb{1}_A : \Omega \rightarrow \{0, 1\}$ is defined as

$$\mathbb{1}_A(x) = \begin{cases} 1 & (x \in A) \\ 0 & (x \notin A) \end{cases}$$

□

4.11 Definition (Stochastic independence)

1. A finite set $\mathcal{A}_1, \dots, \mathcal{A}_n$ of sub- σ -algebras of \mathcal{A} is *stochastically independent*, if for any $A_i \in \mathcal{A}_i$, it holds:

$$\mathbb{P}(A_1 \cap A_2 \cap \dots \cap A_n) = \mathbb{P}(A_1) \cdot \mathbb{P}(A_2) \cdot \dots \cdot \mathbb{P}(A_n)$$

An infinite set of sub- σ -algebras is said to be (*mutually*) *stochastically independent*, if any finite subset is stochastically independent.

2. Random variables X_1, X_2, \dots are *stochastically independent*, if the σ -algebras generated by them, $\sigma(X_1), \sigma(X_2), \dots$ are stochastically independent.
3. Events $A_1, A_2, \dots \subseteq \Omega$ are *stochastically independent*, if the σ -algebras generated by the respective indicator functions, $\sigma(\mathbb{1}_{A_1}), \sigma(\mathbb{1}_{A_2}), \dots$, are (*mutually*) stochastically independent.

This corresponds to the definition from elementary probability, that two events A_1, A_2 are stochastically independent, if $\mathbb{P}(A_1 \cap A_2) = \mathbb{P}(A_1) \cdot \mathbb{P}(A_2)$.

The adverbs *mutually* and *stochastically* are often omitted.

□

The following compact definitions and results of the *expectation* and *conditional expectation* of random variables are taken from [Oksendal1998] and extended to our need:

4.12 Definition and Lemma (Expectation)

Let $X : (\Omega, \mathcal{A}) \rightarrow (\mathbb{R}^n, \mathbb{B}^n)$ be a random variable with distribution μ_X .

1. If $\int_{\Omega} |X(\omega)| d\mathbb{P}(\omega) < \infty$, then the number

$$\mathbb{E}[X] := \int_{\Omega} X(\omega) d\mathbb{P}(\omega) = \int_{\mathbb{R}^n} x d\mu_X(x)$$

is called the *expectation* of X w.r.t. the measure \mathbb{P} .

2. Let $f : \mathbb{R}^n \rightarrow \mathbb{R}$ be BOREL measurable with $\int_{\Omega} |f(X(\omega))| d\mathbb{P}(\omega) < \infty$. Then:

$$\mathbb{E}[f(X)] := \int_{\Omega} f(X(\omega)) d\mathbb{P}(\omega) = \int_{\mathbb{R}^n} f(x) d\mu_X(x)$$

3. Let $X_1, X_2 : \Omega \rightarrow \mathbb{R}$ be two independent, real-valued random variables with $\mathbb{E}[|X_1|] < \infty$, $\mathbb{E}[|X_2|] < \infty$. Then, the expectation of their product $X_1 X_2$ is

$$\mathbb{E}[X_1 X_2] = \mathbb{E}[X_1] \cdot \mathbb{E}[X_2]$$

Note: The integrals are LEBESGUE integrals. □

4.13 Definition and Lemma (Conditional expectation)

Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space, $X : \Omega \rightarrow \mathbb{R}^n$ a random variable with $\mathbb{E}[|X|] < \infty$, and $\mathcal{H} \subset \mathcal{A}$ be a sub- σ -algebra. Then, the *conditional expectation of X given \mathcal{H}* is a random variable $\mathbb{E}[X|\mathcal{H}] : \Omega \rightarrow \mathbb{R}^n$ that fulfills the following conditions:

- a) $\mathbb{E}[X|\mathcal{H}]$ is \mathcal{H} -measurable
- b) $\mathbb{E}[\mathbb{1}_H \cdot \mathbb{E}[X|\mathcal{H}]] = \mathbb{E}[\mathbb{1}_H \cdot X]$ for all $H \in \mathcal{H}$.

The conditional expectation is almost-surely unique. □

Defining a measure μ on \mathcal{H} by $\mu(H) := \int_H X d\mathbb{P}$ ($H \in \mathcal{H}$), then μ is absolutely continuous w.r.t. $\mathbb{P}|_{\mathcal{H}}$ (i.e. $\forall H \in \mathcal{H} : \mathbb{P}(H) = 0 \implies \mu(H) = 0$), thus, by the RADON-NIKODÝM theorem, there exists a \mathcal{H} -measurable function F on Ω , such that $\mu(H) = \int_H F d\mathbb{P} \forall H \in \mathcal{H}$, and F is almost-surely unique. This F is the conditional expectation.

An equivalent formulation of definition and lemma 4.13b) is given as

$$\int_H \mathbb{E}[X|\mathcal{H}] d\mathbb{P} = \int_H X d\mathbb{P} \quad (\forall H \in \mathcal{H}).$$

Thus, the conditional expectation can be interpreted as the RADON-NIKODÝM derivative of μ w.r.t. \mathbb{P} .

Without proof, we give some basic properties of the conditional expectation:

4.14 Lemma (Some properties of the conditional expectation)

Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space, $X: \Omega \rightarrow \mathbb{R}^n$ a random variable with $\mathbb{E}[|X|] < \infty$, and $\mathcal{H} \subset \mathcal{A}$ be a sub- σ -algebra. Additionally, let $Y: \Omega \rightarrow \mathbb{R}^n$ be a second random variable with existing expectation and $a, b \in \mathbb{R}$. Then, it holds:

1. The conditional expectation is linear, i.e.

$$\mathbb{E}[aX + bY|\mathcal{H}] = a \cdot \mathbb{E}[X|\mathcal{H}] + b \cdot \mathbb{E}[Y|\mathcal{H}]$$
2. $\mathbb{E}[\mathbb{E}[X|\mathcal{H}]] = \mathbb{E}[X]$
3. $\mathbb{E}[X|\mathcal{H}] = X$ if X is \mathcal{H} -measurable
4. $\mathbb{E}[X|\mathcal{H}] = \mathbb{E}[X]$ if X is independent of \mathcal{H}
5. $\mathbb{E}[X^T Y|\mathcal{H}] = Y^T \mathbb{E}[X|\mathcal{H}]$ if Y is \mathcal{H} -measurable

□

4.15 Definition and Lemma (Normal or GAUSSian distribution, POISSON distribution)

Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a complete probability space.

1. Let $X: \Omega \rightarrow \mathbb{R}^n$ be a random variable. X is said to be *GAUSSian distributed* or *normally distributed* with expectation μ and covariance matrix Σ , if its density function is given as

$$f_X(x) = \frac{1}{(2\pi)^{\frac{n}{2}} (\det \Sigma)^{\frac{1}{2}}} \exp\left(-\frac{1}{2}(x - \mu)^T \Sigma^{-1}(x - \mu)\right)$$

and we write $X \sim \mathcal{N}(\mu, \Sigma)$.

2. In the special case that a normally distributed random variable X is centered ($\mu = 0 \in \mathbb{R}^n$) and the components are uncorrelated ($\Sigma = \mathbb{I}_{n \times n}$), the density function reduces to:

$$f_X(x) = \frac{1}{(2\pi)^{\frac{n}{2}}} \exp\left(-\frac{1}{2}\|x\|^2\right)$$

3. Let $X: \Omega \rightarrow \mathbb{N}_0$ be an integer-valued (discrete) random variable. X is said to be *POISSON distributed* with parameter λ , if its probability mass function is given as

$$P_X(k) = \begin{cases} e^{-\lambda} \cdot \frac{\lambda^k}{k!} & k \in \mathbb{N}_0 \\ 0 & \text{otherwise} \end{cases}$$

and we write $X \sim \mathcal{Poi}(\lambda)$.

4. For a POISSON distributed random variable $X \sim \mathcal{Poi}(\lambda)$, it holds:

$$\mathbb{E}[X] = \lambda \quad \text{and} \quad \text{Var}[X] = \lambda.$$

□

4.1.2 Stochastic processes

Stochastic processes can be defined as measurable mappings from any probability space into arbitrary measurable spaces. Here, however, we restrict ourselves to \mathbb{R}^n -valued stochastic processes.

4.16 Definition (Stochastic process)

Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a complete probability space, and T an index set (interpreted as interval of time). $X_t: \Omega \rightarrow \mathbb{R}^n$ shall be a random variable for every $t \in T$.

The parametrized collection of random variables

$$\{X_t\}_{t \in T}$$

is called a *stochastic process*, and often written as X_t for short.

Usually, we choose w.l.o.g. $T = \mathbb{R}^+ := [0, \infty)$. □

A stochastic process $\{X_t\}_{t \in T}$ can be interpreted as a mapping

$$X: \Omega \times T \rightarrow \mathbb{R}^n \quad \text{with} \quad (\omega, t) \mapsto X_t(\omega) = X(\omega, t).$$

such that, for each fixed time point $t \in T$, the map

$$\omega \mapsto X_t(\omega) \quad (\omega \in \Omega)$$

is a random variable, and for a fixed choice of $\omega \in \Omega$, the function

$$t \mapsto X_t(\omega) \quad (t \in T)$$

is a certain *realization* or (*sample*) *path* of the stochastic process.

One must be precise when comparing stochastic processes, as the value of $X_t(\omega)$ depends on $t \in T$ and $\omega \in \Omega$, with T and Ω both being uncountable sets in general.

4.17 Definition (Version, modification, indistinguishability)

Let $\{X_t\}_{t \geq 0}$ and $\{\tilde{X}_t\}_{t \geq 0}$ be two stochastic processes on $(\Omega, \mathcal{A}, \mathbb{P})$, i.e. $X: (\omega, t) \mapsto X_t(\omega)$ and $\tilde{X}: (\omega, t) \mapsto \tilde{X}_t(\omega)$.

1. \tilde{X}_t is a *version* or *modification* of X_t , if

$$\forall t \geq 0: \mathbb{P}(X_t = \tilde{X}_t) = \mathbb{P}(\{\omega \in \Omega \mid X_t(\omega) = \tilde{X}_t(\omega)\}) = 1$$

2. \tilde{X}_t and X_t are *indistinguishable*, if

$$\mathbb{P}(\{\omega \in \Omega \mid \forall t \geq 0: \tilde{X}_t(\omega) = X_t(\omega)\}) = 1$$

i.e., X_t and \tilde{X}_t have almost surely the same sample paths.

An equivalent characterization is: X_t and \tilde{X}_t are indistinguishable, if there exists a subset $A \subset \Omega$ with $\mathbb{P}(A) = 1$ and $X_t(\omega) = \tilde{X}_t(\omega) \forall \omega \in A$ and all $t \geq 0$. □

4.18 Example (Version, modification, indistinguishability)

Let $T \sim \mathcal{U}(0,1)$ be a uniformly distributed random variable, $X_t := \mathbb{1}_{\{t=T\}}$, and $\tilde{X}_t := 0$ the zero process. Then, for all t , we have $\mathbb{P}(X_t \neq 0) = \mathbb{P}(T = t) = 0$, or, equivalently, $\forall t : P(X_t = \tilde{X}_t) = 1$, i.e.. X_t is a version of \tilde{X}_t . But obviously, X_t is not indistinguishable from the zero process \tilde{X}_t , as its sample path $t \mapsto X_t(\omega)$ always takes the value 1 for one certain time point for every $\omega \in \Omega$. \square

Throughout this thesis, we will deal only with processes that fulfill the *càdlàg* property. As a pleasant side effect, for *càdlàg* processes, being a version and being indistinguishable coincides:

4.19 Definition and Lemma (*Càdlàg – continue à droite, limite à gauche*)

1. Let $I \subseteq \mathbb{R}$. A function $f: I \rightarrow \mathbb{R}^n$ is called *càdlàg*, if it is “continuous on the right” and has “limits on the left”, i.e.
 - a) the right limit $f(a^+) := \lim_{x \rightarrow a^+} f(x)$ exists and $f(a^+) = f(a)$
 - b) the left limit $f(a^-) := \lim_{x \rightarrow a^-} f(x)$ exists.
2. A stochastic process X_t on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ is called *càdlàg*, if almost all its sample paths $t \mapsto X_t(\omega)$ are *càdlàg*.
3. Let X_t and \tilde{X}_t be *càdlàg* stochastic processes. If \tilde{X}_t is a modification/version of X_t , then X_t and \tilde{X}_t are indistinguishable.

Proof: \tilde{X}_t is a version of X_t , thus $\mathbb{P}(X_t = \tilde{X}_t) = 1$ for all $t \geq 0$.

Let $A = \{\omega \in \Omega \mid \exists t \geq 0 : \tilde{X}_t(\omega) \neq X_t(\omega)\}$. Then, since \tilde{X}_t and X_t are *càdlàg*, we have $\mathbb{P}(A) = \mathbb{P}\left(\bigcup_{q \in \mathbb{Q}^+} \{\omega \in \Omega \mid \tilde{X}_q(\omega) \neq X_q(\omega)\}\right) \leq \sum_{q \in \mathbb{Q}^+} \mathbb{P}\left(\{\omega \in \Omega \mid \tilde{X}_q(\omega) \neq X_q(\omega)\}\right) = 0$. \square

We introduce two frequently referenced properties on the increments of stochastic processes.

4.20 Definition (independent and stationary increments)

Let X_t be a stochastic process.

1. X_t has *independent increments*, if for all $0 \leq t_0 < t_1 < t_2 < \dots < t_n$ the increments $X_{t_1} - X_{t_0}, X_{t_2} - X_{t_1}, \dots, X_{t_n} - X_{t_{n-1}}$ are independent.
2. X_t has *stationary increments*, if the process $X_t - X_s$ has the same distribution as X_{t-s} for all $0 \leq s < t < \infty$.

\square

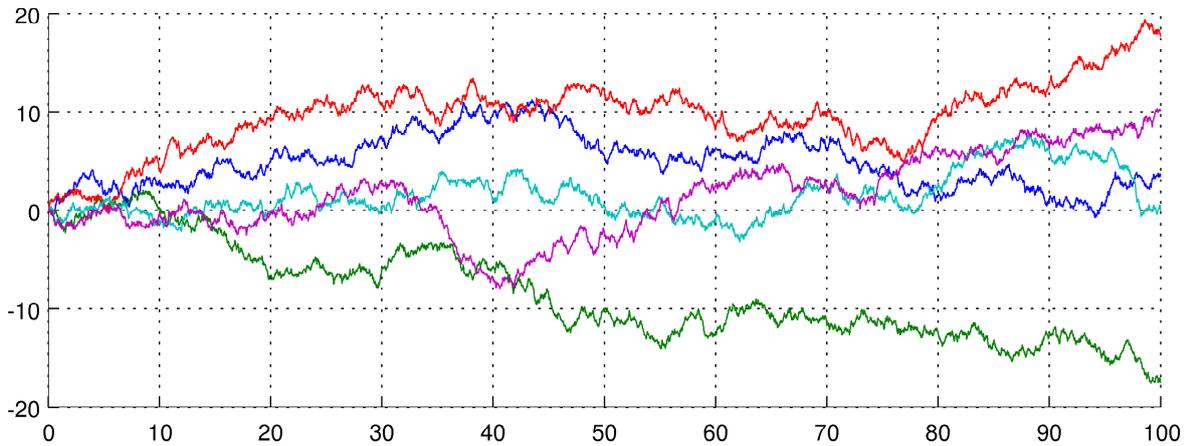


Figure 4.1: Five example realizations of standard WIENER processes in the time range $[0, 100]$, generated by the explicit EULER-MARUYAMA scheme developed in section 4.4.1 with time steps of $\Delta t = 0.01$.

4.2 The WIENER process and BROWNIAN motion

We introduce the notion of *WIENER process* and give a short survey of its development starting from the observation of particle movement in BROWNIAN motion.

4.2.1 Definition (WIENER process, 1D)

Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space. A stochastic process $\{W_t\}_{t \geq 0}$, $W_t : \Omega \rightarrow \mathbb{R}$, is called a *1D standard WIENER process*, if

- i) $W_0 = 0$ with probability 1 (\mathbb{P} -almost surely),
- ii) for two time points $0 \leq s < t$, the *increments* $W_t - W_s$ are stationary and normally distributed with mean zero and variance $t - s$, i.e. $W_t - W_s \sim \mathcal{N}(0, t - s)$,
- iii) for each finite set of time points $\{t_1, \dots, t_m\}$, the increments

$$W_{t_1} - W_{t_0}, W_{t_2} - W_{t_1}, \dots, W_{t_m} - W_{t_{m-1}}$$
 are stochastically independent,
- iv) for $\omega \in \Omega$, the realization $X_t(\omega)$ is continuous with probability 1 (\mathbb{P} -almost surely).

Figure 4.1 shows the trajectories of five 1D standard WIENER processes. □

The continuity requirement in definition iv) is not mandatory. KOLMOGOROV's continuity theorem says that if there exist constants $a, b > 0$ and $c < \infty$ for a stochastic process $\{X_t\}_{t \geq 0}$, such that $\mathbb{E}[|X_t - X_s|^a] \leq c|t - s|^{1+b} \forall t, s \geq 0$, then there exists a continuous *modification* \tilde{X}_t of X_t , i.e. \tilde{X}_t is continuous and $\mathbb{P}(X_t = \tilde{X}_t) = 1 \forall t \in [0, \infty)$.

It can be shown that such constants a, b, c do exist for a WIENER process W_t (see, e.g., [Krylov2002]), and thus we can always construct a continuous version \tilde{W}_t that is even indistinguishable from W_t due to its càdlàg property.

4.22 Definition and Lemma (Drift and volatility, 1D)

Let W_t be a standard WIENER process, and $\mu, D \in \mathbb{R}$. The stochastic process

$$X_t := \mu t + DW_t$$

is called a WIENER process (or BROWNIAN motion) with *drift* μ and *volatility* D .

For $0 \leq s < t$, the distribution of the increments is given as

$$X_t - X_s \sim \mathcal{N}(\mu(t-s), D^2(t-s)).$$

□

4.2.1 BROWNIAN motion or WIENER process? – A short historical note

Introducing the BROWNIAN motion as a synonym for a WIENER process actually is a reversal of history. It was BROWN⁽⁴⁾ who studied the movement of small particles ejected from grains of pollen suspended in water during the year 1827⁽⁵⁾. He described a “rapid oscillatory motion” [Brown1828] of those particles in his (initially unpublished) manuscript *A brief account of microscopical observations made in the months of June, July and August, 1827, on the particles contained in the pollen of plants; and on the general existence of active molecules in organic and inorganic bodies*. BROWN stated, that this motion cannot be due to the vitality of living organisms – a belief widely spread at this time and also BROWN’s initial guess – as he made identical observations with dead materials.

It was Christian WIENER⁽⁶⁾ who showed in 1863 by experiment that the quick and jerky motion of small particles suspended in liquids is due to the motion of the liquids’ molecules [Wiener1863].

It is not obvious that a stochastic process with the properties of definition 4.21 exists. It took 60 years after Christian WIENER’s experimental demonstration, until his namesake Norbert WIENER⁽⁷⁾ – inspired by work of BACHELIER, EINSTEIN, SMOLUCHOWSKI, PERRIN, and others – developed a rigorous mathematical proof using new techniques and results of LEBESGUE and BOREL in measure theory [Wiener1923]. Rigorous proofs can also be found for example in [IkedaWatanabe1981] or [Oksendal1998].

⁽⁴⁾Robert BROWN, 1773–1858, Scottisch botanist

⁽⁵⁾Johann INGEN-HOUSZ, 1730–1799, Dutch physiologist and botanist and the discoverer of photosynthesis described a similar movement of charcoal dust in ethanol decades earlier than BROWN in 1784 in his *Bemerkungen über den Gebrauch des Vergrößerungsglases* (Remarks on the usage of the magnifying glass, [IngenHousz1784]):

„Um es klar einzusehen, daß man sich aus Mangel der Aufmerksamkeit in seinem Urtheile hierüber betrügen könnte, darf man nur in den Brennpunct eines Mikroskops einen Tropfen Weingeist sammt etwas gestoßener Kohle setzen; man wird diese Körperchen in einer verwirrten beständigen und heftigen Bewegung erblicken, als wenn es Thierchen wären, die sich reissend unter einander fortbewegen.“

— To see it clearly that one may deceive oneself in one’s verdict due to a lack of attentiveness, one only has to put a drop of ethanol together with a little ground charcoal into the focus of a microscope; one will find those corpuscles in a confused, persistent and boisterous motion, as if they were animalcules rapidly moving among themselves.

The claim for discovering what is known today as BROWNIAN motion might thus be attributed to INGEN-HOUSZ, though his experimental set-up also resembles the BÉNARD experiment. Possibly, the movements observed by INGEN-HOUSZ might also have been due to RAYLEIGH-BÉNARD convection.

Peter HÄNGGI has collated many “historical items and surveys” about BROWNIAN motion on his web page at University of Augsburg: <http://www.physik.uni-augsburg.de/theo1/hanggi/History/BM-History.html>

⁽⁶⁾Ludwig Christian WIENER, 1826–1896, German mathematician, physicist und philosopher

⁽⁷⁾Norbert WIENER, 1894–1964, US-American mathematician and philosopher, and considered the originator of *cybernetics*

We list some basic properties of WIENER processes. For proofs, we refer to the literature, e.g. [Oksendal1998; Protter2004].

4.23 Lemma (Distribution of the WIENER process)

Let W_t be a standard WIENER process. Then, it holds:

1. $W_t \sim \mathcal{N}(0, t)$, that is $\mathbb{P}(W_t \in (a, b)) = \frac{1}{\sqrt{2\pi t}} \int_a^b e^{-\frac{x^2}{2t}} dx$.
2. W_t and $\Delta W_{t+s,s} := W_{t+s} - W_s$ are identically distributed.
3. $\mathbb{E}[W_t] = 0$. □

4.24 Definition and Lemma (WIENER process, nD)

Let $W_t^{(1)}, W_t^{(2)}, \dots, W_t^{(n)}$ be stochastically independent standard WIENER processes. Then, the stochastic process

$$W_t := \left(W_t^{(1)}, W_t^{(2)}, \dots, W_t^{(n)} \right)^T$$

is called *n-dimensional standard WIENER process*, and for $0 \leq s < t$, the increments $\Delta W_{t,s} = W_t - W_s \in \mathbb{R}^n$ are stochastically independent and $\mathcal{N}(0, (t-s)\mathbb{I}_n)$ normally distributed. □

4.25 Definition and Lemma (Drift and Variance, nD)

Let W_t be an *n-dimensional standard WIENER process*, and $\mu \in \mathbb{R}^n$, $D \in \mathbb{R}^{n \times n}$. The stochastic process

$$X_t := \mu t + DW_t$$

is called an *n-dimensional WIENER process* (or *BROWNIAN motion*) with *drift* μ and *volatility* or *variance* DD^T . Thus, we have

$$X_t - X_s \sim \mathcal{N}((t-s)\mu, (t-s)DD^T).$$

for the increments $X_t - X_s$ and $0 \leq s < t$. □

4.3 Stochastic integrals and stochastic differential equations

In section 4.4 that copes with numerical integration of stochastic differential equations, we will use a discrete reformulation of a stochastic initial value problem (S-IVP), and will develop a simple numerical integration scheme (the *explicit EULER-MARUYAMA*) along the way.

Before that, we specify which type of solutions we are looking for and give some results on the existence and uniqueness of solutions to stochastic initial value problems.

An introduction to the theory of the underlying stochastic integrals of ITO is found in appendix C.

4.3.1 From ordinary to stochastic differential equations

Inspired by the introduction of ØKSENDAL [Oksendal1998], we consider a classical population growth model.

4.26 Problem (Population Growth)

Let $N(t)$ denote the size of a population (in continuous units) at time t , with an initial population of $N_0 \geq 0$. Further, let $a(t)$ be the growth rate of the population at a certain time point t .

Then, we can describe the time evolution of population size by the initial value problem

$$\frac{dN}{dt}(t) = a(t)N(t), \quad N(0) = N_0 \quad (4.1)$$

The growth function is usually not completely known, but can it can be modelled deterministically by knowledge or assumptions on the environment, e.g. nutrition conditions. However, even if we know a lot about the system, there remains some *randomness*, such that the growth rate might be modelled as the sum of a deterministic function $\bar{a}(t)$ that is affected by some noise:

$$a(t) = \bar{a}(t) + \text{“noise”}$$

such that the above problem reads as

$$\frac{dN}{dt}(t) = (\bar{a}(t) + \text{“noise”}) \cdot N(t) = \bar{a}(t)N(t) + N(t) \cdot \text{“noise”}$$

The question arises, how to model the “noise”, and what its mathematical interpretation is. More generally, one would like to study equations of the form

$$\frac{dX_t}{dt} = f(t, X_t) + g(t, X_t) \cdot \text{“noise”}$$

or

$$\frac{dX_t}{dt} = f(t, X_t) + g(t, X_t) \cdot Z_t$$

with given *drift* function f , *volatility* function g , and a suitable “noise” process Z_t .

We might introduce the term *white noise process* for Z_t at this point. It is a generalized stochastic process that can be interpreted as a probability measure on the space of tempered distributions $\mathcal{S}'([0, \infty))$ ⁽⁸⁾, and we could finally interpret the BROWNIAN motion W_t as an integral of the white noise process:

$$W_t = \int_0^t dZ_t \quad \text{and thus, formally,} \quad \frac{dW_t}{dt} = Z_t.$$

⁽⁸⁾The SCHWARTZ space $\mathcal{S}([0, \infty)) = \{\varphi \in C^\infty([0, \infty)) \mid \forall \alpha, \beta \in \mathbb{N}_0^n : \sup_{x \in \mathbb{R}^n} |x^\alpha D^\beta \varphi(x)| < \infty\}$ is the space of rapidly falling C^∞ -functions. A continuous linear mapping $f: \mathcal{S}([0, \infty)) \rightarrow \mathbb{C}$ is called a *tempered distribution*, and the set of all those mappings is $\mathcal{S}'([0, \infty))$, i.e. the dual space of $\mathcal{S}([0, \infty))$.

However, the latter equation states Z_t being the time derivative of a WIENER process, which is nondifferentiable almost everywhere. For solving stochastic differential equations, fortunately, we do not need to study white noise processes first. Instead, we interpret SDEs as a formal writing for stochastic integral equations, which can then be analysed and solved with methods from the ITÔ calculus.

Analogously to solving an ordinary differential equation (more precisely, the initial value problem)

$$\frac{dx}{dt}(t) = f(t, x(t)), \quad x(t_0) = x_{t_0}$$

by calculating the respective integral

$$x(t) = x_{t_0} + \int_{t_0}^t f(s, x(s)) ds,$$

and transferring this ansatz to a (preliminary) *stochastic differential equation* (again, more precisely, a *stochastic initial value problem*, S-IVP)

$$dX_t = f(t, X_t) dt + g(t, X_t) dW_t, \quad X_{t_0} = x_{t_0}$$

leads to the *stochastic integral equation*

$$X_t = x_{t_0} + \int_{t_0}^t f(s, X_s) ds + \int_{t_0}^t g(s, X_s) dW_s$$

with a BROWNIAN motion or WIENER process W_t .

The first integral can be interpreted as an “ordinary” LEBESGUE-integral, if the drift function f and its arguments are continuous almost everywhere. The second one, however, has to be treated with care, as it is an integral with respect to a WIENER measure. In appendix C, a short theoretical introduction to this kind of stochastic integrals is given.

We note that the initial value x_{t_0} can also be random, though we will focus on a deterministic initial state in the following discussion.

Similar to the deterministic case for non-autonomous ODE, any n -dimensional time-inhomogeneous SDE

$$dX_t = f(t, X_t) dt + g(t, X_t) dW_t, \quad X_{t_0} = x_{t_0}$$

can be written as a time-homogeneous one by transformation into an $(n + 1)$ -dimensional system with an appropriate modification of the drift and diffusion functions f and g . Whereas in the deterministic setting, the addition of the ODE $\frac{d}{dt}x_{n+1} = 1$ with initial condition $x_{n+1}(t_0) = t_0$ can be used as a substitute for the time variable, in the stochastic setting, an additional differential state with drift $f \equiv 1$ and diffusion $g \equiv 0$ is added. Without loss of generality, it would thus be sufficient to study time-homogeneous SDEs only.

However, for numerical research, it is common and computationally more efficient, to cope with an explicit time-dependence. Thus and for the sake of simplicity, we investigate w.l.o.g. processes with initial time point $t_0 = 0$ and end time $t_f = T$.

We formally restate a multi-dimensional stochastic initial value problem, where the underlying ITÔ integral is driven by a WIENER process. An introduction to the ITÔ calculus is found in appendix C. Conditions for existence and uniqueness of solutions are given in section 4.3.2.

4.27 Definition (SDE, S-IVP)

Let W_t be an n_W -dimensional WIENER process, and let further be given a vector-valued function $f: \mathbb{R} \times \mathbb{R}^{n_x} \rightarrow \mathbb{R}^{n_x}$ and a matrix-valued function $g: \mathbb{R} \times \mathbb{R}^{n_x} \rightarrow \mathbb{R}^{n_x \times n_W}$.

1. We call f the *drift*, and g the *diffusion* function.
2. A *stochastic initial value problem (S-IVP)* is of the form

$$dX_t = f(t, X_t) dt + g(t, X_t) dW_t, \quad X_0 = x_0, \quad t \in [0, T] \quad (4.2)$$

with initial state $x_0 \in \mathbb{R}^{n_x}$.

3. The associated *stochastic integral equation* reads as

$$X_t = x_0 + \int_0^t f(s, X_s) ds + \int_0^t g(s, X_s) dW_s \quad (4.3)$$

4. If the diffusion function g is a constant, independent of X_t , i.e. it has a representation as a matrix in $\mathbb{R}^{n_x \times n_W}$, the system in eq. (4.2) is called an SDE with *additive noise*. Otherwise, the SDE is said to have *multiplicative noise*.

For the broader class of LÉVY-driven SDE, where the driving WIENER process W_t is substituted by a LÉVY process, see section 4.5. □

4.3.2 Strong and weak solutions

When we look for a solution X_t of the S-IVP $dX_t = f(t, X_t) dt + g(t, X_t) dW_t$, $X_0 = x_0$, there are two kind of solutions, depending on whether the driving WIENER process W_t is given or not.

In the first case, informally spoken, the WIENER process (or the underlying probability space) is given and we try to find a solution process X_t to the S-IVP, which is then called a *strong solution*. This kind of solution can be seen as being fully determined by the initial value and the values of the driving WIENER process.

In the second case, we are looking for a pair (X_t, W_t) , or a process X_t and an underlying probability space $(\Omega, \mathcal{A}, \mu)$, that fulfills the S-IVP. This pair will be called a *weak solution*, if it exists.

In the setting of estimating parameters from measured quantities, we have some measurements of the stochastic process available (or of functions thereof), i.e. the driving WIENER process has already been fixed and we are not free to choose one. Therefore, we are only interested in strong solutions.

4.3.3 An existence and uniqueness result

The following theorem gives sufficient conditions on the drift and diffusion coefficient functions such that a solution to an SDE exists. These are rather strong and clearly not necessary.

4.28 Theorem (Existence and uniqueness theorem for stochastic differential equations)

Let $f : [0, T] \times \mathbb{R}^{n_x} \rightarrow \mathbb{R}^{n_x}$ be the drift function and $g : [0, T] \times \mathbb{R}^{n_x} \rightarrow \mathbb{R}^{n_x \times n_w}$ be the diffusion function of the stochastic initial value problem

$$dX_t = f(t, X_t) dt + g(t, X_t) dW_t, \quad X_0 = x_0$$

on the interval $[0, T]$. Then, there exists a unique time-continuous solution $X_t(\omega)$, if it holds:

- (1) (measurability): f and g are measurable functions
- (2) (LIPSCHITZ condition in space): There exists a constant $L_1 > 0$, such that

$$|f(t, x) - f(t, y)| + |g(t, x) - g(t, y)| \leq L_1|x - y|$$

for all $x, y \in \mathbb{R}^{n_x}$ and $t \in [0, T]$

- (3) (linear growth bound): There exists a constant $L_2 > 0$, such that

$$|f(t, x)| + |g(t, x)| \leq L_2(1 + |x|)$$

where we set $|g|^2 := \sum_{i=1}^m \sum_{j=1}^n |g_{ij}|^2$ for the matrix-valued diffusion function g .

Further, the solution process X_t fulfills $\mathbb{E} \left[\int_0^T |X_t|^2 dt \right] < \infty$. □

The above theorem holds for random initial states, with x_0 being a measurable random variable such that $\mathbb{E}[|x_0|^2] < \infty$, i.e. the initial state must only be of finite variance. For a proof, see [Oksendal1998], theorem 5.2.1.

Further, the global LIPSCHITZ condition 4.28(2) might be weakened to a local one, and it can even be omitted if one is not interested in uniqueness. The linear growth bound 4.28(3) ensures that $|X_t|$ does not go to infinity in finite time, and can thus be weakened if we are willing to accept that there might only exist solutions on a smaller interval, whose extent will be usually depending on the initial value x_0 .

4.3.4 The bridge between ODE and SDE: random ordinary differential equations (RODE)

We shortly mention a class of differential equations located somewhere between ODE and SDE: *random ordinary differential equations (RODE)*. A common interpretation of RODE is to see them as *pathwise* ODEs that are driven by a (known) stochastic process in their r.h.s. function. They have been studied extensively in the 1970s [Strand1970; Soong1973; Doss1977; Sussmann1978], and are still receiving attraction [ImkellerSchmalfuss2001; AsaiKloeden2016], as they deliver means to numerically solve (i.e. retrieve a certain realization) SDE using ODE techniques.

4.29 Definition (Random Ordinary Differential Equation)

Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space, $Z_t : [0, T] \times \Omega \rightarrow \mathbb{R}^{n_z}$ be a stochastic process with continuous sample paths, and $f : \mathbb{R}^{n_z} \times \mathbb{R}^{n_x} \rightarrow \mathbb{R}^{n_x}$ continuous. Then, the system

$$\dot{x}(t) = f(x(t), Z_t(\omega)) \tag{4.4}$$

is called a *random ordinary differential equation (RODE)*. □

There is a close relationship between SDE, RODE, and ODE, as eq. (4.4) denotes an ODE for fixed $\omega \in \Omega$. Further, by allowing arbitrary ω , the driving process Z_t puts us in the stochastic setting.

WONG and ZAKAI et al. [WongZakai1965] showed the convergence of ODE to SDE driven by a WIENER process W_t by using smooth approximations W_t^n that converge to W_t . It was first shown by SUSSMANN [Sussmann1978] that a certain class of SDE can be solved by application of ODE integration schemes to *all* sample paths of the intrinsic stochastic process, though this is, from a numerical point of view, a rather theoretical result.

An RODE with a WIENER process in the r.h.s. function can be written as SDE in a quite obvious way (restricted here to the scalar case, but the principle holds for any dimension of the state variable x and the WIENER process W_t):

$$\dot{x}(t) = f(x(t), W_t) \iff d \begin{pmatrix} X_t \\ Z_t \end{pmatrix} = \begin{pmatrix} f(X_t, Z_t) \\ 0 \end{pmatrix} dt + \begin{pmatrix} 0 \\ 1 \end{pmatrix} dW_t.$$

On the other hand, SDE may be reformulated as RODE as DOSS and SUSSMANN have shown in the late 1970s for a certain form of noise [Doss1977; Sussmann1978]. This result was extended a couple of years ago by IMKELLER and SCHMALFUSS [ImkellerSchmalfuss2001] to all SDE with finite dimension. The following scalar example with additive noise shows the idea of how this can be done. Using an ORNSTEIN-UHLENBECK process Z_t that satisfies the SDE $dZ_t = -Z_t dt + dW_t$, the equivalence of RODE and SDE becomes obvious:

$$dX_t = f(X_t) dt + dW_t \iff \dot{z}(t) = f(z(t) + Z_t) + Z_t \text{ with } z(t) := X_t - Z_t$$

since $Z_0 = 0$ holds for the ORNSTEIN-UHLENBECK process Z_t , and $X_0 = z(0)$ for the initial condition, we have $z(t) = X_t - Z_t = X_0 - Z_0 + \int_0^t f(X_s) + Z_s ds = z(0) + \int_0^t f(z(s) + Z_s) + Z_s ds$, so z is (pathwise) differentiable.

Thus, in the case of WIENER processes, SDE and RODE may be considered equivalent.

4.4 Numerical integration of WIENER-driven SDEs

Formally, as stated above, the stochastic initial value problem (S-IVP) (4.2) is actually interpreted as the stochastic integral equation (4.3) with a driving BROWNIAN motion W_t .

This integral equation might, in rare cases, be analytically solved by methods of the ITÔ calculus, resulting in a stochastic process X_t . We instead consider a discrete reformulation of the stochastic initial value problem

$$dX_t = f(t, X_t) dt + g(t, X_t) dW_t, \quad X_0 = x_0, \quad (4.5)$$

by replacing the time differential dt by increments $\Delta\tau_k$ and the stochastic differential dW_t by increments (differences) of the BROWNIAN motion W_t :

$$\tilde{X}_{k+1} - \tilde{X}_k = \Delta\tilde{X}_k = f(\tau_k, \tilde{X}_k)\Delta\tau_k + g(\tau_k, \tilde{X}_k)\Delta W_k, \quad \tilde{X}_0 = x_0, \quad (4.6)$$

where $\tilde{X}_k := \tilde{X}_{\tau_k}$ is the discrete solution at time τ_k ; further $\Delta\tau_k := \tau_{k+1} - \tau_k$ denotes the k -th time increment, and $\Delta W_k := W_{k+1} - W_k := W_{\tau_{k+1}} - W_{\tau_k}$ the corresponding increment of the WIENER process. For simplicity, we assume an equidistant time grid $\tau_k := k \cdot \frac{T}{N}$ ($k = 0, \dots, N$).

One can show that this discretization converges to (a version of) the stochastic solution process X_t , both in the *strong* and *weak* sense; two terms that we are going to define next.

4.30 Definition (Weak and strong convergence)

Let $X_t = X_t(\omega)$ be a stochastic process on $[0, T]$, fulfilling the S-IVP

$$dX_t = f(t, X_t) dt + g(t, X_t) dW_t, \quad X_0 = x_0$$

and let $\tilde{X}_k = \tilde{X}_k(\omega) = \tilde{X}_{\tau_k}(\omega)$ be a numerical approximation on X_{τ_k} on a time grid $\tau_k := k \cdot \Delta\tau$ ($k = 0, \dots, N$) with $\Delta\tau := \frac{T}{N}$.

1. The numerical method is said to have *strong order of convergence* ρ if there exists a constant $C > 0$ such that

$$\mathbb{E}\left[|X_{\tau_k} - \tilde{X}_k|\right] \leq C \cdot (\Delta\tau)^\rho$$

at any fixed $k \in \{0, \dots, N\}$ and sufficiently small $\Delta\tau$.

Thus, the order of *strong convergence* quantifies the rate at which the *mean of errors* decreases for step sizes $\Delta\tau \rightarrow 0$.

2. The numerical method is said to have *weak order of convergence* ρ if there exists a constant $C > 0$ such that

$$|\mathbb{E}[X_{\tau_k}] - \mathbb{E}[\tilde{X}_k]| \leq C \cdot (\Delta\tau)^\rho$$

at any fixed $k \in \{0, \dots, N\}$ and sufficiently small $\Delta\tau$.

Thus, the order of *weak convergence* quantifies how fast the *error of the means* reduces. \square

4.4.1 EULER-MARUYAMA stochastic integrator

The integration scheme sketched in eq. (4.6) is called the *EULER-MARUYAMA*⁽⁹⁾ or *stochastic EULER* scheme. In the general multi-dimensional setting, where the n -dimensional stochastic process X_t depends on m standard WIENER processes, the scheme for the stochastic initial value problem on the time domain $[0, T]$

$$dX_t = f(t, X_t) dt + g(t, X_t) dW_t, \quad X_0 = x_0, \quad t \in [0, T]$$

with drift function $f: [0, T] \times \mathbb{R}^n \rightarrow \mathbb{R}^n$, diffusion function $g: [0, T] \times \mathbb{R}^n \rightarrow \mathbb{R}^{m \times n}$, initial value $x_0 \in \mathbb{R}^n$, and constant time increments $\Delta\tau := \frac{T}{N}$ ($N \in \mathbb{N}$), can be written as

$$\tilde{X}_{k+1} := \tilde{X}_k + f(\tau_k, \tilde{X}_k)\Delta\tau + g(\tau_k, \tilde{X}_k)\Delta W_k, \quad \tilde{X}_0 := x_0 \quad (4.7)$$

with \tilde{X}_k being the approximation of X_{τ_k} at time points $\tau_k := k \cdot \Delta\tau$ ($k = 0, \dots, N$). As the approximation at τ_{k+1} only depends on the values at τ_k , the stochastic EULER-MARUYAMA method is an *explicit* integration scheme that extends the classical explicit EULER method for ordinary differential equations by a stochastic term.

The increments ΔW_k are derived from the driving m -dimensional WIENER process W_t , i.e. $\Delta W_k := W_{\tau_{k+1}} - W_{\tau_k}$. Since the component functions of W_t are mutually independent scalar WIENER processes, the i -th component of increment ΔW_k can be chosen as a normally distributed random number with zero mean and variance $\Delta\tau$, such that $\Delta W_k \sim \mathcal{N}(0, \Delta\tau \cdot \mathbb{I}_{m \times m})$.

Algorithm 4.1 gives an implementation of the EULER-MARUYAMA scheme in pseudo-code.

⁽⁹⁾ Leonhard EULER, 1707–1783, Swiss mathematician, physicist, astronomer, and engineer
Gishiro MARUYAMA (丸山 儀四郎), 1916–1986, Japanese mathematician

Algorithm 4.1 EULER-MARUYAMA stochastic integrator

Integrates the S-IVP $dX_t = f(t, X_t) dt + g(t, X_t) dW_t$, $X_0 = x_0$ by performing N EULER-MARUYAMA steps of length Δt .

Input: $f: \mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R}^n$, $g: \mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R}^{n \times m}$, $t_0 \in \mathbb{R}$, $x_0 \in \mathbb{R}^n$, $\Delta t \in \mathbb{R}$, $N \in \mathbb{N}$

Output: discrete trajectory $X(0:N)$

```

1: function EULERMARUYAMA( $f, g, t_0, x_0, \Delta t, N$ )
2:    $X(0) := x_0$ 
3:   for  $k = 0$  to  $N - 1$  do
4:     Choose  $\Delta W \in \mathbb{R}^m$  with  $\Delta W_i \sim \mathcal{N}(0, \Delta t)$  and  $\Delta W_i, \Delta W_j$  independent for  $i \neq j$ 
5:      $X(k + 1) := X(k) + f(t_0 + k \cdot \Delta t, X(k)) \cdot \Delta t + g(t_0 + k \cdot \Delta t, X(k)) \cdot \Delta W$ 
6:   end for
7:   return  $X$ 
8: end function

```

As with every numerical method, the question arises whether an EULER-MARUYAMA approximation generated by algorithm 4.1 actually converges to the true solution process. The following theorem answers this question, and additionally gives a uniform error estimate over the whole interval $[0, T]$.

4.31 Theorem (Strong convergence of the EULER-MARUYAMA scheme)

Let X_t be solution process of the S-IVP

$$dX_t = f(t, X_t) dt + g(t, X_t) dW_t, \quad X_0 = x_0$$

with drift and diffusion functions fulfilling a LIPSCHITZ condition in space

$$|f(t, x) - f(t, y)| + |g(t, x) - g(t, y)| \leq L_1 |x - y|,$$

a linear growth condition

$$|f(t, x)| + |g(t, x)| \leq L_2(1 + |x|),$$

and a coupled condition

$$|f(s, x) - f(t, x)| + |g(s, x) - g(t, x)| \leq L_3(1 + |x|) \cdot |s - t|^{\frac{1}{2}}$$

for all time points $s, t \in [0, T]$ and vectors $x, y \in \mathbb{R}^n$, and some constants L_1, L_2, L_3 independent of the integration stepsize Δt .

Then, for $\tilde{X}(T)$ being the EULER-MARUYAMA approximation at time T , generated by algorithm 4.1 with constant step size $\Delta \tau$, it holds

$$\mathbb{E}[X_T - \tilde{X}(T)] \leq C \cdot (\Delta \tau)^{0.5}$$

for a constant $C > 0$, independent of $\Delta \tau$, i.e. the EULER-MARUYAMA scheme is strongly convergent with order 0.5. \square

For a proof, see, [KloedenPlaten1995], theorem 10.2.2, which is formulated in a more general sense, as it also allows the initial value x_0 to be a random variate. One can further show that the EULER-MARUYAMA scheme has a weak order of convergence of 1.

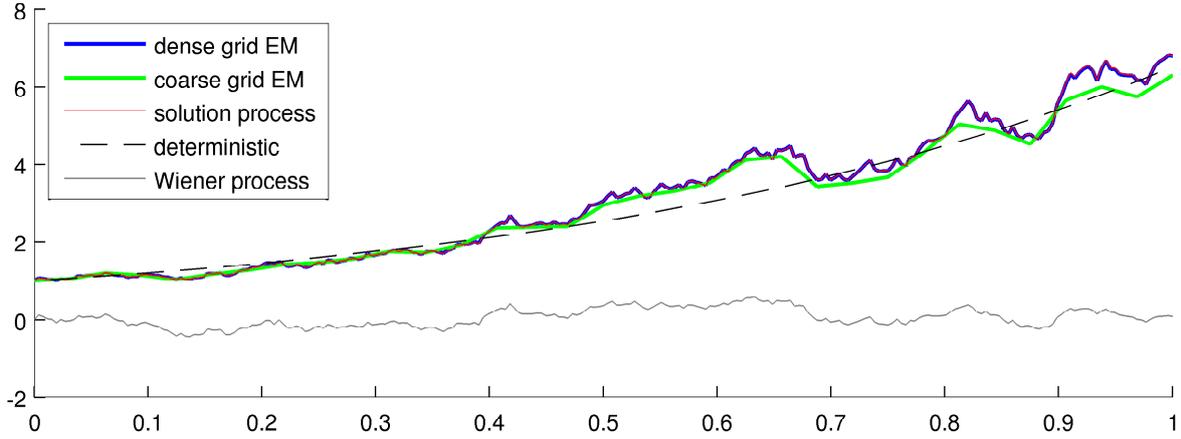


Figure 4.2: Comparing EULER-MARUYAMA approximations to a strong solution of an S-IVP.

Two EULER-MARUYAMA approximations of the solution of the S-IVP $dX_t = \mu X_t dt + DX_t dW_t$, $X_0 = 1$, $\mu = 2$, $D = 0.5$, with stepsize 2^{-8} (blue line, —) and with stepsize 2^{-5} (green line, —). A linear interpolation is performed between the distinct evaluation points.

The dense grid EULER-MARUYAMA approximation is on this scale practically indistinguishable from the trajectory of the exact solution process $X_t = X_0 \cdot \exp((\mu - \frac{1}{2}D^2)t + DW_t)$ (thin red line, —).

For comparison, the deterministic counterpart $X(t) = X_0 \cdot \exp(\mu t)$ is added (thin dashed black line, --).

In the lower part of the figure, the driving WIENER process that is used in the solution process and in the EULER-MARUYAMA approximations is printed (thin gray line, —).

4.4.2 An illustrating example of EULER-MARUYAMA integration

As an illustrating example, we have a look at the 1-dimensional time-homogeneous S-IVP

$$dX_t = \mu X_t dt + DX_t dW_t, \quad X_0 = x_0 \quad (4.8)$$

with constant drift $f(t, x) = \mu x$, $\mu > 0$, constant diffusion $g(t, x) = Dx$, $D > 0$, and given initial value x_0 . For this simple S-IVP, using the ITÔ calculus, we can give a closed expression for the solution process:

$$X_t = x_0 \cdot \exp((\mu - \frac{1}{2}D^2)t + DW_t). \quad (4.9)$$

Once we have fixed the driving WIENER process W_t , this enables us to compare the true solution process X_t given above to approximations \tilde{X}_t generated by the EULER-MARUYAMA method with different step sizes.

Figure 4.2 compares two approximative solutions of the S-IVP (4.8) generated by EULER-MARUYAMA approximations with different step sizes to the solution process given in eq. (4.9). Both the approximations and the solution process share the same driving WIENER process. The dense grid EULER-MARUYAMA approximation is indistinguishable (visually, not in the meaning of stochastic processes) from the solution process at this scale. Indeed, the maximum pointwise error is around 0.6% for this specific realization. A linear interpolation between the point evaluations is performed to get a continuous trajectory.

4.4.3 Further integration methods

Besides the described EULER-MARUYAMA algorithm, there exist further methods for pathwise simulations of stochastic processes, differing in their convergence properties inter alia. To give one example, we establish the MILSTEIN scheme for stochastic integration.

The EULER-MARUYAMA scheme, restricted to the 1-dimensional case, reads as

$$X_{k+1} := X_k + f\Delta\tau_k + g\Delta W_k$$

where we omitted the arguments of drift and diffusion functions, i.e. $f = f(\tau_k, X_k)$, $g = g(\tau_k, X_k)$, and noted X_k for the approximation instead of \tilde{X}_k for increased readability. It is the simplest method in the class of strong TAYLOR approximations. Similar to deterministic integration, other methods can be constructed by truncating the TAYLOR series of the solution, for example the method proposed by MILSTEIN⁽¹⁰⁾ [Milstein1975], that extends the EULER-MARUYAMA scheme by a single term:

$$X_{k+1} := X_k + f\Delta\tau_k + g\Delta W_k + \frac{1}{2}gg'((\Delta W_k)^2 - \Delta\tau_k)$$

where the prime denotes the derivative with respect to the state variable: $g'(t, x) = \frac{dg}{dx}(t, x)$. This scheme, the *explicit MILSTEIN method*, can be shown to increase the *strong* order of convergence to 1.0, while maintaining weak order of 1.0, however, its calculation requires the evaluation of the first derivative of the diffusion function. From the view point of convergence order, the MILSTEIN method is the actual stochastic counterpart of the explicit EULER method for ODEs.

Obviously, for SDEs with additive noise, i.e. the diffusion does not depend on the state X_t , the MILSTEIN and the EULER-MARUYAMA scheme coincide as the derivative g' vanishes.

(Semi-)Implicit integration schemes

The previously described EULER-MARUYAMA and MILSTEIN schemes are *explicit* schemes, as the value at the next time step is determined by the value of the previous time point. Similar to the deterministic setting, implicit integration schemes incorporate the unknown new state X_{k+1} on both sides of the equation.

Fully implicit stochastic integration methods suffer from the fact that the generated approximations will not have finite absolute moments, so a strong convergence analysis cannot be performed. See [KloedenPlaten1995], chapter 12, for details. Semi-implicit methods, however, contain the new iterate X_{k+1} only in their drift function, while maintaining the term X_k in the diffusion function.

The (semi-)implicit one-dimensional counterparts of the above two schemes are given as the *implicit EULER scheme*

$$X_{k+1} := X_k + f(\tau_{k+1}, X_{k+1})\Delta\tau_k + g(\tau_k, X_k)\Delta W_k$$

and the *implicit MILSTEIN scheme*

$$X_{k+1} := X_k + f(\tau_{k+1}, X_{k+1})\Delta\tau_k + g(\tau_k, X_k)\Delta W_k + \frac{1}{2}g(\tau_k, X_k)g'(\tau_k, X_k)((\Delta W_k)^2 - \Delta\tau_k).$$

⁽¹⁰⁾ Grigory Noichowitsch MILSTEIN (former transcription: Mil'shtejn or Mil'shtein), Russian mathematician.

In fact, a whole family of implicit schemes can be created by using a convex combination $\alpha f(\tau_{k+1}, X_{k+1}) + (1 - \alpha)f(\tau_k, X_k)$ of implicit ($\alpha = 1$) and explicit ($\alpha = 0$) scheme.

Depending on the choice of the coefficients, the implicit EULER method achieves a strong order of convergence of 0.5, the implicit MILSTEIN method reaches 1.0 – both corresponding to their fully explicit counterparts.

The main advantage of implicit methods is, like in the deterministic case, their superior behaviour on stiff systems. *Stiffness* in the context of SDEs is characterized by the LYAPUNOV exponents of the (linearized) system. Widely spread LYAPUNOV exponents refer to widely differing time scales in the solution, analogously to the real parts of the coefficient matrix' eigenvalues in linear ODEs. We refer to [KloedenPlaten1995] for the topic of stability in SDEs.

Higher order methods

We conclude this section by noting that there exist higher order methods that, however, become computationally very costly and are often difficult to implement.

For example, the implicit order 2.0 strong TAYLOR scheme for the 1-dimensional time-homogeneous STRATONOVICH S-IVP, taken from [KloedenPlaten1995], reads as

$$\begin{aligned} X_{k+1} := & X_k + \frac{1}{2}(\underline{f}(X_{k+1}) + \underline{f}(X_k))\Delta\tau_k + g\Delta W_k + \underline{f}g'(\Delta W_k\Delta\tau_k - \Delta Z_k) \\ & + \underline{f}'g(\Delta Z_k - \frac{1}{2}\Delta W_k\Delta\tau_k) + \frac{1}{2}gg'(\Delta W_k)^2 \\ & + \frac{1}{3!}g(gg')(\Delta W_k)^3 + \frac{1}{4!}g(g(gg'))(\Delta W_k)^4 \\ & + \underline{f}(gg')'J_{(0,1,1)} + g(\underline{f}g')'J_{(1,0,1)} + g(\underline{f}g)'(J_{(1,1,0)} - \frac{1}{4}(\Delta W_k)^2\Delta\tau_k), \end{aligned}$$

where we have waived the argument X_k to the drift and diffusion functions, and $\underline{f} := f - \frac{1}{2}gg'$. The random variate $\Delta Z_k := \int_{\tau_k}^{\tau_{k+1}} \int_{\tau_k}^{s_2} dW_{s_1} ds_2$ is normally distributed with mean $\mathbb{E}[\Delta Z_k] = 0$, variance $\text{Var}[\Delta Z_k] = \frac{1}{3}(\Delta\tau_k)^3$, and covariance $\mathbb{E}[\Delta Z_k\Delta W_k] = \frac{1}{2}(\Delta\tau_k)^2$. To make it more readable, the above written functional J abbreviates multiple STRATONOVICH integrals:

$$J(j_1, j_2, j_3) := \int_{\tau_k}^{\tau_{k+1}} \int_{\tau_k}^{\tau_{k+1}} \int_{\tau_k}^{s_2} \circ dW_{s_1}^{j_1} \circ dW_{s_2}^{j_2} \circ dW_{s_3}^{j_3},$$

with the convention $dW_s^0 = ds$.

Thus, in every step $k \rightarrow k + 1$ in the above scheme, derivatives up to third order of the diffusion function and multiple STRATONOVICH integrals have to be evaluated, which makes this approach computationally costly.

As a side note: The only reason why we switched here to the STRATONOVICH version is that it is *more convenient to write* compared to the ITÔ version.

The (implicit) EULER or MILSTEIN methods already reach strong order of convergence of 0.5 and 1.0, respectively. In practical settings it might be questionable whether the additional effort for reaching a strong order of 2.0 is beneficial. Although the above formulas simplify in special setting like additive noise, it is often more convenient and computationally more efficient to use lower order methods instead, not even speaking about the error-prone implementation that are hard to check due to their stochastic nature.

4.5 LÉVY-driven SDE

The theory of stochastic integrals in appendix C and of stochastic differential equations discussed in section 4.3 is suited to WIENER processes *driving* the SDE. A more general class of driving processes is the one of LÉVY processes, for which the theory may be extended accordingly (see, e.g., [Protter2004; Bichteler2002]).

4.5.1 A short introduction to LÉVY processes

4.32 Definition (LÉVY process)

Let X_t be a stochastic process with $X_0 = 0$ a.s. Then X_t is called a LÉVY process, if the following conditions hold:

- i) the sample paths of X_t are càdlàg a.s.,
- ii) X_t has independent increments,
- iii) X_t has stationary increments. □

4.33 Remark

The requirement in definition 4.32 of a LÉVY processes being càdlàg can be substituted by requiring continuity in probability, i.e. for $\epsilon > 0$ and $t \geq 0$, it holds $\lim_{h \rightarrow 0} \mathbb{P}(|X_{t+h} - X_t| > \epsilon) = 0$ (or, equivalently, $\lim_{s \rightarrow t} X_s = X_t$ in probability).

One can then show that for every LÉVY process there exists a càdlàg version of it. For a proof, see e.g. [Protter2004]. □

4.34 Definition and Lemma (Counting process, (compound) POISSON process)

1. A stochastic process $\{N_t\}_{t \geq 0}$ is called a *counting process*, if the following conditions hold:
 - a) $N_t \in \mathbb{N}_0$ for all $t \geq 0$
 - b) $N(t) \geq N(s)$ for $t \geq s$

A counting process, as the name suggests, counts certain “events” – whatever they may be – that have been occurred up to and including time t . Thus, the increment

$$\Delta N((s, t]) := N(t) - N(s)$$

is the number of events occurring in the interval $(s, t]$.

2. A counting process $\{N_t\}_{t \geq 0}$ is called a *POISSON process* with rate (or intensity) $\lambda > 0$, if the following conditions hold:
 - a) $N_0 = 0$ a.s.
 - b) the process has independent increments,
 - i.e. $\Delta N((s, s+t])$ has the same distribution as $\Delta N((0, t])$ for all $t > s \geq 0$.
 - c) the increments are POISSON distributed
 - i.e. $\Delta N((s, s+t]) \sim \mathcal{Poi}(\lambda(t-s))$ for all $t > s \geq 0$.

3. Let N_t be a POISSON process with rate $\lambda > 0$, and $D^{(i)}$ ($i \in \mathbb{N}$) independent and identically distributed random variables with common cumulative distribution function F_D . Further, $D^{(i)}$ shall be independent of N_t for all $i \in \mathbb{N}$ and $t \geq 0$. Then,

$$X_t := \sum_{i=1}^{N_t} D^{(i)} \quad (4.10)$$

is called a *POISSON process* with rate λ and jump size distribution F_D .

4. Every POISSON process and every compound POISSON process is a LÉVY process.

Proof: Let N_t be a POISSON process as above. Then, $\Delta N((s, t]) \sim \mathcal{Poi}(\lambda(t - s))$ and also $\Delta N((0, t - s]) \sim \mathcal{Poi}(\lambda(t - s))$, i.e. the increments are stationary. Further, the increments are also independent by definition and N_t is càdlàg by construction, thus N_t is a LÉVY process.

Let X_t be a compound POISSON process as above. Then, by construction, it is càdlàg and has independent increments. For $0 \leq s < t$, we have (a) $X_t - X_s = \sum_{i=1}^{N_t} D_i - \sum_{i=1}^{N_s} D_i = \sum_{i=N_s+1}^{N_t} D_i$ and (b) $X_{t-s} = \sum_{i=1}^{N_{t-s}} D_i$. The number of jumps in (a) is $N_t - N_s$ which has the same distribution as N_{t-s} in (b), since the POISSON process N_t has stationary increments, and thus also $X_t - X_s$ and X_{t-s} have the same distribution, as the D_i are i.i.d. Thus, X_t is a LÉVY process. \square

Figure 4.3 shows two realizations of LÉVY processes: A POISSON process with rate $\lambda = 5.0$ and a compound POISSON process with rate $\lambda = 5.0$ and standard-normally distributed jumps.

4.35 Remark (LÉVY-KHINTCHINE representation of LÉVY processes)

We note that every LÉVY process Z_t may be decomposed

$$Z_t = b(t) + aW_t + M_t$$

into its deterministic drift $b(t)$, a scaled WIENER process W_t , and a jump process M_t being a superposition of independent POISSON processes, with M_t independent of W_t , $W_0 = 0$, and $M_0 = 0$. \square

4.5.2 Numerical simulation of LÉVY-driven SDE

Some numerical integration schemes for SDE with WIENER-driven diffusion are described in section 4.4. There, we deal with ITÔ integrals w.r.t. WIENER processes, for which TAYLOR expansions are available, giving rise to several integration schemes with (in principle) arbitrary convergence rate [KloedenPlaten1995], with the EULER-MARUYAMA and MILSTEIN methods being the most prominently used ones.

In this section, we give a short introduction to a closely related numerical integration method for LÉVY-driven SDE that follows ideas analogous to the explicit EULER-MARUYAMA method for WIENER-driven SDE.

We employ appropriate smoothness conditions on f and g (e.g. standard LIPSCHITZ conditions) as well as a sufficient integrability condition on the driving m -dimensional LÉVY process Z_t (e.g. square integrability, $\mathbb{E}[|Z_t|^2] < \infty$) for strong solutions to exist.

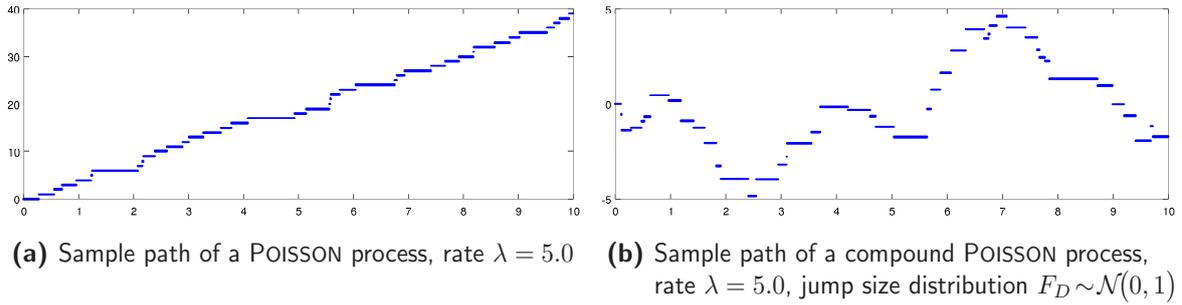


Figure 4.3: Sample paths of (compound) POISSON processes.

Let $f: [0, T] \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ be the deterministic drift function, and $g: [0, T] \times \mathbb{R}^n \rightarrow \mathbb{R}^{m \times n}$ be a coefficient function for the driving LÉVY process Z_t . Then

$$\tilde{X}_{k+1} := \tilde{X}_k + f(\tau_k, \tilde{X}_k) \Delta\tau + g(\tau_k, \tilde{X}_k) \Delta Z_k, \quad \tilde{X}_0 := x_0 \quad (4.11)$$

delivers an EULER approximation \tilde{X}_k of X_{τ_k} for the S-IVP

$$dX_t = f(t, X_t) dt + g(t, X_{t-}) dZ_t, \quad X_0 = x_0 \quad (4.12)$$

with initial value x_0 , constant time increments $\Delta\tau := \frac{T}{N}$ ($N \in \mathbb{N}$), and $\tau_k := k \cdot \Delta\tau$ for $k = 0, \dots, N$. Note that since Z_t is càdlàg, the evaluation of the coefficient function g in the S-IVP occurs at the left limits, “right before” a possible jump.

4.36 Remark

It would be mathematically sufficient to formulate a LÉVY-driven SDE as $dX_t = g(X_{t-}) dZ_t$ due to the LÉVY-KHINTCHINE representation of Z_t (see remark 4.35). \square

4.37 Remark

If we are able to compute (simulate) the increments of the LÉVY process $\Delta Z_k := Z_{\tau_{k+1}} - Z_{\tau_k}$ exactly, the method is called a *genuine EULER method*.

However, for many LÉVY processes, simulation of exact increments is computationally hard. Often, approximations $\Delta \tilde{Z}_k$ on the increments are used, introducing a second source of error apart from the unavoidable discretization error. In this case, the method is frequently called an *approximative EULER method*. \square

Under mild assumptions on the first moment of Z_t , and still requiring a finite second moment of Z_t , in addition to sufficient smoothness of f and g , it can be shown that the genuine EULER method has an error of $\mathcal{O}(\Delta\tau)$, see [ProtterTalay1997] and [DereichHeidenreich2011]. JACOD et al. [Jacod2005] give precise error bounds on both the genuine and approximate EULER methods.

5 Parameter Estimation in Stochastic Differential Equation Models

In this chapter, we present and analyse a new method for parameter estimation in stochastic differential equations, based on a piecewise deterministic approach.

In the first section, we give an introduction into the topic of parameter estimation in SDE, describe our new approach in words, and formulate some technical assumptions.

Section 2 outlines existing estimation techniques for parameter estimation in continuous-time SDE with discretely sampled observations. References are given within the discussion.

The third section presents some results (with proofs) on the distance of solutions of stochastic initial value problems, S-IVP, to be defined in eq. (5.10), to the corresponding solutions of deterministic initial value problems, D-IVP, defined in eq. (5.11). For both, distance at the interval end as well the maximum distance throughout the interval, upper bounds in expectation and mean-square are given.

In section 4, we present our new approach, the *piecewise deterministic parameter estimation method* for SDE in detail. After introducing notation and basic assumptions, mainly to ensure existence of strong solutions to the SDE, the new method is derived on the basis of the multiple shooting technique for parameter estimation in ODE presented in chapter 2. The continuity condition is replaced by a carefully weighted *jump regularization term* in the objective, allowing for a discontinuous trajectory and thus mimicking the stochasticity of an SDE formulation. It is shown that this regularization is necessary in order to get correct parameter estimates. Further, results from the third section are used to prove that the jumps asymptotically converge to zero if the number of equidistantly chosen shooting intervals goes to infinity.

Section 5 gives an numerical analysis of the proposed method. The sparsity pattern of the Jacobian of the *combined residual vector*, composed of measurement and jump residuals, is investigated, and it is shown that the number of nonzero elements grows only linearly in the number of shooting intervals. Further, we prove that the sparsity is maintained under HOUSEHOLDER based decomposition techniques. The section concludes with the proposition and analysis of a lifting technique based on interval-wise decoupling of parameters.

In section 6, we propose two extensions: an homotopy ansatz for pathological problems and a grid refinement strategy that is elaborated in more detail in the numerical examples chapter 6.

We also refer to appendix S for a discussion of the software package `:sfit` that implements the presented parameter estimation method.

5.1 Introduction

In many settings, e.g. in cellular biology, occurring processes happen in a “mostly” deterministic way, so to speak. For example, in enzyme kinetics, the reactions follow clear principles: reaction partners are known and if there are enough molecules of each occurring species, modelling with ODE is appropriate (at least for a certain time horizon).

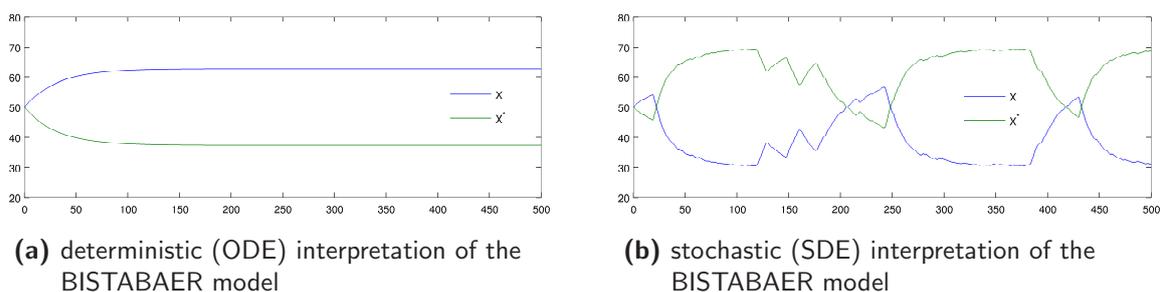


Figure 5.1: Stochasticity in the BISTABAER model completely alters its behaviour.

The ODE system (a) quickly approaches the steady state, whereas in the SDE interpretation (b), ongoing transitions between two steady states are observed. See section 6.3 for details.

However, this is not always the case. As an illustrative example, we have a look at the *allosteric regulation*⁽¹⁾ of an enzyme, which might switch between its inactive form X and active form X^* due to allosterically binding by its activator or inhibitor, respectively.

Figure 5.1 shows the time course of a simplified allosteric regulation model. In a deterministic interpretation, the system quickly reaches its (quasi-)steady state, as can be seen on section 5.1, whereas the stochastic version reveals a (random) switching between two stable steady states.

In both settings, a hidden control (the bistable system $\dot{L} = L(1 - L^4)$) is implemented that influences the amount of activator (effector) concentration (details are given in section 6.3). In the ODE interpretation, this control system quickly reaches and remains on a (locally) stable steady state, allowing the enzyme model to run towards its equilibrium. In the stochastic setting, the bistable control system gets permanent input from a WIENER process. While the control system tries to maintain its stability, a displacement by the driving WIENER process that is high enough might bring the control system into the contraction area of the second steady state, resulting in a different concentration of the activator. As a consequence, the equilibrium point of the enzyme model is changed.

However, the underlying *dynamics* of this system are – in a way – deterministic: If we knew the state of the (stochastically driven) control system L , the resulting kinetics can be calculated using an ODE integrator to arbitrary precision.

In other cases, e.g. in mathematical finance, some general principles of interest rates are known or assumed. Properties like (exponential) *mean reversion* to a level μ with reversion rate θ can be described by a simple deterministic model

$$\dot{x}(t) = \theta(\mu - x(t)),$$

however, the actual sample paths are stochastic, possibly not even continuous (see section 6.4).

As soon as there are stochastic parts in a system, modelling with ODE is, in general, inappropriate, as the system is likely to show behaviour that one is unable to describe using pure ODE formulations, albeit the main occurring processes possess deterministic properties.

⁽¹⁾ The activity of an enzyme might be regulated by one or more *effector proteins*, that bind to the enzyme at its *allosteric site* (not the active site). An *activator* might stabilize an enzyme in its active form, whereas the binding of an *inhibitor* might change the enzyme's conformation in a way that the active site becomes inaccessible to the substrate.

On notation: X_t vs. $x(t)$

Throughout this and the following chapter, capital letters with indices, e.g. X_t , denote stochastic processes, especially solutions of an S-IVP

$$dX_t = f(t, X_t) dt + g(t, X_t) dW_t, \quad X_{t_0} = x_0$$

whereas the corresponding ODE solution, i.e. the solution of the D-IVP

$$\dot{x}(t) = f(t, x(t)), \quad x(t_0) = x_0$$

is denoted by the respective small letter with arguments set in parenthesis, $x(t)$.

5.1.1 The idea: ODE solutions resemble SDE solution paths on short time scales

When the influence of a driving stochastic process W_t is not “too strong”, we intuitively expect a solution X_t of an **SDE with constant diffusion D** , i.e. with diffusion function $g \equiv D$,

$$dX_t = f(X_t) dt + D dW_t \quad X_{t_0} = x_0$$

to be close to the solution $x(t)$ of the **corresponding ODE** with same initial value

$$\dot{x}(t) = f(x(t)) \quad x(t_0) = x_0$$

at least for a small period of time. Indeed, this can be shown to true; estimates of the maximum distance and of the distance at the interval and are given in section 5.3.

We can also make this observation in the introductory example of an allosteric enzyme regulation model. Though the deterministic and stochastic interpretation differ qualitatively over the whole time domain, the ODE and SDE trajectories stay closely together over a short time span as depicted in figure 5.2.

If we know the state of the stochastically influenced system at a specific time point t , say X_t , and simulate the corresponding ODE in the interval $[t, t + \Delta t]$ with this state as initial condition, we may expect to stay “close” to the SDE solution, where the admissible length Δt obviously depends on the driving process’s activity in that interval. Thus, we may try to approximate the solution process X_t over the whole time domain by a number of discontinuously concatenated ODE solutions.

The finer we choose the grid of ODE approximations, the smaller the gaps eventually become, approaching zero in the limit, as we will show in section 5.4.6.

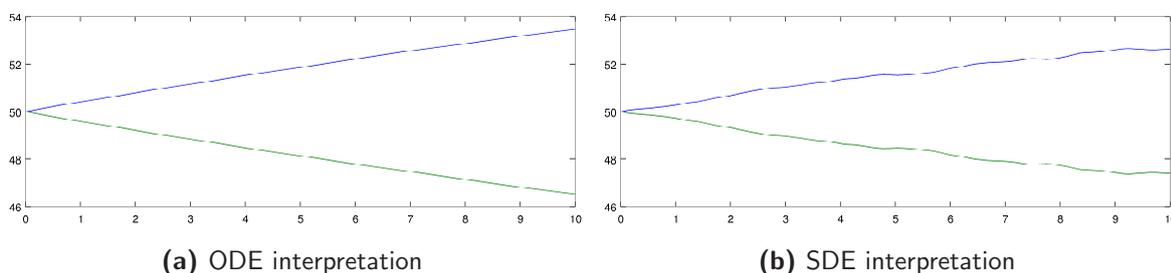


Figure 5.2: Short term similarity of ODE and SDE interpretation of the BISTABAER model.

This picture is a detail enlargement of figure 5.1. For a short time scale, the ODE and SDE trajectories are very similar. The impact of the driving WIENER process manifests as a small jitter in the SDE interpretation.

Clearly, if we already have the stochastic solution process X_t (i.e. a certain realization of it), there is no need to approximate it any more. In the context of parameter estimation, we (usually) do not have a continuous observation of the system, but rather some measurement values of its state or functions thereof.

Now, the above observation gives rise to the following idea: Dissect the time horizon into small intervals, and determine an ODE solution on each interval, such that their (discontinuous) concatenation is close to the observations. Further, since the (unknown) stochastic solution process is continuous, choose the initial conditions on each interval in a way such that the discontinuities at the interval borders become small.

The above description resembles the method of multiple shooting for parameter estimation in ODE (section 2.2.3). What is most important, is that this ansatz gives us access to derivatives and allows the application of gradient-based optimization methods, although the approximated stochastic process X_t is nowhere differentiable in general.

We will see, that this idea is not just a pious hope, as section 5.3 gives some general results on the convergence of WIENER-driven SDE to ODE solutions. Further, section 5.4.6 gives a convergence result for the proposed parameter estimation method.

The precise formulation of the *parameter estimation method with jump regularization* is done in section 5.4.

5.1.2 SDE with constant diffusion by LAMPERTI transform

In this thesis, the focus lies on SDE with constant diffusion. Under the conditions of theorem 4.28 (existence and uniqueness of SDE solutions), a 1-dimensional SDE

$$dX_t = f(t, X_t) dt + g(t, X_t) dW_t \quad (t \in [0, T])$$

with state-dependent diffusion may be transformed into an SDE with constant diffusion $D = 1$, using the *LAMPERTI transform* $\mathfrak{L}(t, X_t)$ [Iacus2008; MollerMadsen2010; LuschgyPages2006], which is based on the ITÔ formula (see, e.g. [Oksendal1998], Theorems 4.1.2 and 4.2.1):

$$Z_t := \mathfrak{L}(t, X_t) := \int_{\xi}^{X_t} \frac{1}{g(t, x)} dx \quad (5.1)$$

with an arbitrary value ξ from the state space of X_t . We note that the LAMPERTI transformation in eq. (5.1) is bijective if $g(t, X_t) > 0 \forall (t, X_t)$, as for every $t \in [0, T]$, $x \mapsto \mathfrak{L}(t, x)$ is continuous and strictly increasing [LuschgyPages2006].

The transformed process Z_t solves the SDE

$$dZ_t = \left[\frac{d}{dt} \mathfrak{L}(t, \mathfrak{L}^{-1}(t, Z_t)) + \frac{f(\mathfrak{L}^{-1}(t, Z_t))}{g(\mathfrak{L}^{-1}(t, Z_t))} - \frac{1}{2} \frac{d}{dx} g(t, \mathfrak{L}^{-1}(t, Z_t)) \right] dt + 1 dW_t$$

from which, after simulation, the original process may be reconstructed as $X_t = \mathfrak{L}^{-1}(t, Z_t)$. For time-independent diffusion $g(X_t)$, the SDE for Z_t simplifies to

$$dZ_t = \left[\frac{f(\mathfrak{L}^{-1}(Z_t))}{g(\mathfrak{L}^{-1}(Z_t))} - \frac{1}{2} \frac{dg}{dx}(\mathfrak{L}^{-1}(Z_t)) \right] dt + 1 dW_t .$$

The application of the LAMPERTI transform is limited by the fact that we need to be able to compute its inverse; however, this is possible for quite general classes of diffusion processes [MollerMadsen2010].

A general multivariate version of the LAMPERTI transform is currently not available [MollerMadsen2010], but MØLLER and MADSEN give a *multivariate LAMPERTI transform* for the class of SDEs that are of the form

$$dX_t = f(t, X_t) dt + g(t, X_t)M(t) dW_t \quad \text{with} \quad g(t, X_t) = \begin{pmatrix} g_1(t, X_{1,t}) & & \\ & \ddots & \\ & & g_n(t, X_{n,t}) \end{pmatrix}$$

with state variable $X_t \in \mathbb{R}^n$, an n -dimensional WIENER process $W_t \in \mathbb{R}^n$, a matrix function $M(t) \in \mathbb{R}^{n \times n}$ of the time variable t , and a diagonal diffusion matrix $g(t, X_t) \in \mathbb{R}^{n \times n}$, $g(t, X_t) = \text{diag} \{g_1(t, X_{1,t}), \dots, g_n(t, X_{n,t})\}$, such that the state-dependent parts of the diffusion are not influencing across components. Then, the 1-dimensional LAMPERTI transform, eq. (5.1), can be applied component-wise.

We refer to [MollerMadsen2010] for details.

5.1.3 Restricting w.l.o.g. to time-homogeneous SDE

Using the “standard trick” of adding time as an extra dimension to the state vector, augmenting its dimension by 1, we can always transform time-inhomogeneous SDE into time-homogeneous, so it is sufficient to study time-homogeneous SDE (and the corresponding autonomous ODE).

5.2 Techniques of parameter estimation in SDE

There exists a large variety of estimation methods for SDE. Most methods are specific to different classes of problems: e.g. state-independent vs. state-dependent diffusion, time-homogeneous vs. time-dependent diffusion, continuous vs. discrete-time observations, equidistant vs. arbitrary sampling times, direct state observation vs. indirect/partial observation, exact vs. noisy measurements, scalar vs. multi-dimensional processes, to name a few, and many methods are tailored to problems belonging to certain combinations of these classes.

An overview of methods for state and parameter estimation in stochastic differential equations can be found e.g. in [Singer2002; Sorensen2004; Cysne2004] and the references therein.

Here, we present some work on estimation techniques for nonlinear SDE models with time-discrete observations. The general parametrized continuous-time SDE model with discrete-time measurements taken at time points t_i is frequently written as

$$dX_t = f(t, X_t, p) dt + g(t, X_t, p) dW_t \quad X_{t_0} = x_0(p) \quad (t \in [t_0, t_f]) \quad (5.2)$$

$$\eta_i = h_i(t_i, X_{t_i}, p) + \varepsilon_i, \quad (5.3)$$

with (nonlinear) drift f and diffusion g (see definition 4.27).

Frequently, only time-homogeneous SDEs

$$dX_t = f(X_t, p) dt + g(X_t, p) dW_t \quad X_{t_0} = x_0(p) \quad (t \in [t_0, t_f]) \quad (5.4)$$

are considered (cf. section 5.1.3).

5.2.1 Maximum likelihood, transition densities and the FOKKER-PLANCK equation

Most estimation methods rely on transition densities that describe the probability distribution of the transition from one system state at a certain time into another at a later time. If these transition densities are known, the likelihood function can be calculated and subsequently a maximum likelihood estimator is derived. Under certain regularity conditions ([WongHajek2012], Chapter 4), the transition densities may be calculated by solving a functional partial differential equation, the so-called FOKKER-PLANCK equation 5.7 (see below).

5.2.1.1 Transition probability densities

We introduce the required notions for the special case of error-free full state observations, i.e. $\eta_i = h(t_i, X_{t_i}, p) = X_{t_i}$. To keep the writing short, we have here used one single measurement function $h(t, X_t, p)$ delivering possibly vector-valued measurements.

Writing $\mathbb{P}(X_t \in B | X_s = y)$ for the *conditional probability* of observing X_t lying in a measurable set B at time t , given that $X_s = y$ at a previous time $s < t$, we can note the *transition (probability) density* (or *conditional (probability) density*) as

$$\mathbf{p}(t, x | s, y; p) = \frac{\partial}{\partial x} \mathbb{P}(X_t \leq x | X_s = y; p)$$

that depends on the parameter vector p . Interpreting the initial condition $X_{t_0} = x_0(p)$ as an observation, we further write $X = \{X_{t_0}, X_{t_1}, \dots, X_{t_N}\}$ for the collected observations at time points t_0, t_1, \dots, t_N , and write for short

$$\mathbf{p}(X_{t_{i+1}} | X_{t_i}; p) := \mathbf{p}(t_{i+1}, X_{t_{i+1}} | t_i, X_{t_i}; p)$$

for the transition densities. Using the MARKOV property (ITÔ processes are “memoryless”), the probability distribution of the observations X can be written as

$$\mathbf{p}(X; p) = \mathbf{p}(X_{t_N}, \dots, X_{t_1}, X_{t_0}; p) = \mathbf{p}(X_{t_0}; p) \cdot \prod_{i=0}^{N-1} \mathbf{p}(X_{t_{i+1}} | X_{t_i}; p).$$

5.2.1.2 Likelihood function and maximum likelihood estimator

We can now define the *likelihood function* (compare definition 2.2) as

$$L(p; X) := \mathbf{p}(X; p) \tag{5.5}$$

and define the *maximum likelihood estimator* \hat{p}_{ML} in the usual manner:

$$\hat{p}_{\text{ML}} := \arg \max_p \log L(p; X) = \arg \max_p \log \mathbf{p}(X_{t_0}; p) + \sum_{i=0}^{N-1} \log \mathbf{p}(X_{t_{i+1}} | X_{t_i}; p) \tag{5.6}$$

5.2.1.3 FOKKER-PLANCK equation

The required transition density $\mathbf{p}(t, x | s, y; p)$ is a solution of the *FOKKER-PLANCK equation* (also known as *KOLMOGOROV's forward equation of diffusion*), a functional partial differential equation describing the time evolution of a probability density under drift $f(t, X_t, p)$ and

diffusion $g(t, X_t, p)$, as in the SDE eq. (5.2):

$$\frac{\partial}{\partial t} \mathbf{p}(t, x | t_0, x_0; p) = - \sum_i \frac{\partial}{\partial x_i} [f_i(t, x, p) \cdot \mathbf{p}(t, x | t_0, x_0; p)] \quad (5.7)$$

$$+ \frac{1}{2} \sum_{i,j} \frac{\partial^2}{\partial x_i \partial x_j} [\mathfrak{D}_{ij}(t, x, p) \cdot \mathbf{p}(t, x | t_0, x_0; p)] \quad (5.8)$$

with initial condition $\mathbf{p}(t_0, x | t_0, x_0; p) = \delta(x - x_0)$, a DIRAC delta function, and a diffusion tensor $\mathfrak{D} := gg^T$ [WongHajek2012].

5.2.1.4 Computing the maximum likelihood estimator from transition densities

To compute the maximum likelihood estimator \hat{p}_{ML} from the transition densities, the FOKKER-PLANCK equation 5.7 must be solved repeatedly in a nonlinear optimization algorithm. For linear drift and state-independent diffusion, the obtained densities are GAUSSIAN [Singer2002]; but analytical solutions are only available for very restricted special cases of scalar SDEs. In the general multivariate case, the FOKKER-PLANCK equation is solved approximately, e.g. by finite-difference or finite-element based methods [Pichler2013], matrix continued-fractions [Risken1996], Metropolis Monte Carlo (MC) methods [Kikuchi1991], and other techniques.

Another frequently used technique for approximating the transition densities $\mathbf{p}(X_{t_{i+1}} | X_{t_i}; p)$ is the *MARKOV chain Monte Carlo (MCMC) method* [Pedersen1995; Elerian2001].

We note that, as LO points out, many of these “discretized maximum likelihood” estimators, which are based on a discretization of the underlying SDE (mostly using the EULER-MARUYAMA method, see section 4.4.1), are not necessarily consistent ([Lo1998], Example 3). However, whether the strength of this inconsistency (bias) is meaningful depends on the studied process and must be decided w.r.t. the concrete application.

For linear problems, the well known *STRATONOVICH-KÁLMÁN-BUCY*⁽²⁾ *filtering* techniques are considered the method of choice [Singer2004]. For linear SDE with discrete observations, a compact introduction to KÁLMÁN filtering is found in [Singer2004]. A recent

⁽²⁾ Rudolf Emil KÁLMÁN, born 1930, Hungarian-American engineer and mathematician
 Ruslan Leontievich STRATONOVICH, 1930–1997, Russian physicist and probabilist
 Richard Snowden BUCY, born 1935, American mathematician

Though the mentioned “filter” was named after KÁLMÁN, many of his results were achieved in cooperation with Richard BUCY, and frequently their paper [KalmanBucy1961] is given as original reference for the KÁLMÁN-BUCY filter.

Hardly noticed in the Western world, STRATONOVICH developed a more general nonlinear filter earlier [Stratonovich1960], that contained the linear KÁLMÁN-BUCY filter as a special case. Also to mention is the work of Peter SWERLING, 1929–2000, American radar theoretician and statistician, who developed a similar algorithm for estimating satellite orbits [Swerling1959].

Interestingly, already decades earlier in 1880, Thorvald Nicolai THIELE, 1838–1910, a Danish astronomer, actuary and mathematician, developed a recursive least squares based regression method for noisy observations of the astronomical geodesy problem of determining the distance from Copenhagen (Denmark) to Lund (Sweden). He observed that the errors in a time series of data *appear* to be systematic (but recognized that no systematic correction is possible) and used a BROWNIAN motion to describe the accumulating errors. In [Thiele1880; Thiele1880a], he “proposes a model consisting of a sum of a regression component, a Brownian motion and a white noise for his observations, although he does not use these terms himself. He solves the problem of estimating the regression coefficients and predicting the values of the Brownian motion by the method of least squares and gives an elegant recursive procedure for carrying out the calculations. The procedure is nowadays known as Kalman filtering” [Lauritzen1981].

review of the KÁLMÁN filter and a novel derivation from NEWTON’s method for root finding is also found in [Humpherys2012].

For nonlinear SDE with small sampling intervals $\Delta t_i = t_{i+1} - t_i$, the transition densities may be approximated by conditional GAUSSIAN densities by linearizing the exact moment equations (*extended KÁLMÁN filter (EKF)*, or *second-order nonlinear filter (SNF)*, [Jazwinski1970]) or linearizing the SDE itself using ITÔ’s lemma as in the *local linearization technique* of SHOJI and OZAKI [ShojiOzaki1997; ShojiOzaki1998]. Also to mention in this context is the *prediction error estimation method* [Ljung2002; LjungSoderstrom1985].

Common to all these methods is that they rely on excessive sampling and are therefore **computationally very costly**.

5.2.2 Other estimators

HANSEN presents the *generalized method of moments (GMM)* [Hansen1982] that minimizes a function on the moments of the samples, derived from orthogonality conditions. HANSEN and SCHEINKMAN [HansenScheinkman1995] give a recipe to formulate these conditions for continuous-time MARKOV processes, i.e. also suitable for SDEs, with discrete-time observations. However, this technique relies on complete state observations at any point in time, and thus, it’s applicability is limited to certain cases.

When using *martingale estimation functions (MEF)* [KesslerSorensen1999], the transition densities are not necessarily approximated by (conditional) GAUSSIAN densities, but may have more complex shapes. KESSLER and SORENSEN give criteria under which the developed estimators are consistent. However, the actual construction of these MEF is problem-dependent, as it requires the eigenfunctions of the generator $G_p := f(x, p) \frac{d}{dx} + \frac{1}{2} g(x, p) \frac{d^2}{dx^2}$ (a differential operator) of the SDE model in eq. (5.4).

There exist plenty of other estimation techniques, e.g. based on *indirect inference* [Gourieroux1993; BianchiCleur1996] and *auxiliary models* (e.g. the *efficient method of moments (EMM)* of GALLANT and TAUCHEN [GallantTauchen1996]). In *nonparametric approaches*, conditional densities are approximated by kernel density estimates, possibly without assuming a certain functional form of drift and diffusion functions [BandiPhillips2003], or, as in [Shoji2002; Shoji2003], the coefficients of the TAYLOR series expansion of the drift are estimated. We refer to the literature.

5.2.3 Assessing estimator performance

Comparisons of estimators are not frequently found in literature. The few comparisons that have been done are often restricted to special classes of SDEs that can also be solved analytically to rate the estimator performance. For a discussion of the difficulties and pitfalls in performance assessment, we refer to [Cysne2004].

One of the few actual evaluations of estimator performance is found in [ShojiOzaki1997]. They investigate five different estimators based on maximum likelihood and GMM on the basis of two scalar SDEs, one linear system with state-dependent diffusion, $dX_t = (\alpha_1 + \beta_1 X_t) dt + \sigma_1 X_t dW_t$, and a nonlinear system with constant diffusion, $dX_t = \alpha_2 X_t^3 dt + \sigma_2 dW_t$, with dense and equidistantly sampled, error free, direct observations over the whole time horizon as required by many estimations methods.

The new method we propose in this thesis aims at solving complex multi-dimensional problems with several parameters, unknown initial system state, and arbitrarily sampled, with

intermittent, partial and/or indirect observations, which may additionally be afflicted with measurement errors, like the problems in the numerical examples chapter 6. For these types of problems, no established test suites and protocols exist. Such a test suite has to be developed in future work.

5.3 On the distance between SDE and ODE solutions

We have seen in the introductory example in section 5.1 and figure 5.1 that the addition of even small stochasticity (or “intrinsic noise”) may completely alter the model behaviour seen over the whole time domain.

For smaller time scales, as depicted for the same example in figure 5.2, SDE and ODE solutions are quite similar. There arises thus the question to quantify the deviation between an SDE model and its deterministic ODE counterpart.

To keep the writing short, but w.l.o.g.⁽³⁾, we restrict ourselves here to time-homogeneous SDEs (ODE: autonomous) with constant diffusion. Further, we require the drift function $f: \mathbb{R}^{n_x} \rightarrow \mathbb{R}^{n_x}$ (ODE: r.h.s. function) to be LIPSCHITZ continuous with constant L , i.e.

$$\exists L > 0 : \quad |f(x) - f(y)| \leq L|x - y|, \quad (5.9)$$

to ensure the existence of strong solutions of the SDE (ODE: unique solutions), see theorem 4.28. Moreover, we omit the explicit notation of dependence on a parameter vector $p \in \mathbb{R}^{n_p}$.

By X_t , we denote the (strong) solution process⁽⁴⁾ of the *stochastic initial value problem*

$$\text{S-IVP:} \quad dX_t = f(X_t) dt + D dW_t, \quad X_0 = X_0^{(s)} \quad (5.10)$$

and by $x(t)$ the (unique) solution of the corresponding *deterministic initial value problem*

$$\text{D-IVP:} \quad dx(t) = f(x(t)) dt, \quad x(0) = X_0^{(d)} \quad (5.11)$$

in the time domain $[0, T]$ with non-random initial values $X_0^{(s)}$ and $X_0^{(d)}$, constant diffusion matrix $D \in \mathbb{R}^{n_x \times n_w}$, and an n_w -dimensional WIENER process W_t .

We can write the solutions X_t of the S-IVP as

$$\text{S-IVP:} \quad X_t = X_0^{(s)} + \int_0^t f(X_s) ds + \int_0^t D dW_s \quad (t \in [0, T]) \quad (5.12)$$

and the solution $x(t)$ of the D-IVP as

$$\text{D-IVP:} \quad x(t) = X_0^{(d)} + \int_0^t f(x(s)) ds. \quad (t \in [0, T]) \quad (5.13)$$

⁽³⁾ see section 5.1.2 and section 5.1.3

⁽⁴⁾ i.e., driven by a certain realization of the driving WIENER process W_t

Using this notation, we define the *difference process* Δ_t , measuring the distance between SDE solution X_t and ODE solution $x(t)$:

5.1 Definition (Difference Process)

Let X_t be a strong solution process to the S-IVP (5.10), and $x(t)$ the solution of the corresponding D-IVP (5.11). Then, the stochastic process

$$\Delta_t := X_t - x(t) \tag{5.14}$$

is called the *difference process* of X_t and $x(t)$. □

5.2 Definition (Maximum Process)

For a stochastic process Z_t , we define

$$Z_t^* := \sup_{0 \leq s \leq t} |Z_s|$$

as the *maximum process* of Z_t . □

We note that throughout this chapter, for an n -dimensional vector $x \in \mathbb{R}^n$, we write $|x| := \|x\|_2$. The following lemma is frequently used:

5.3 Lemma (JENSEN's inequality for concave functions)

Let $g: \mathbb{R} \rightarrow \mathbb{R}$ be a concave function, and X an integrable real-valued random variable. Then, it holds:

$$\mathbb{E}[g(X)] \leq g(\mathbb{E}[X]) \tag{5.15}$$

Proof: See, e.g., [Chung2000], Theorem 9.1.4. □

5.3.1 Some results on WIENER processes

Before starting the analysis of the difference process Δ_t , we give some results about WIENER processes.

5.4 Lemma (ITÔ integral over a constant matrix)

Let $W_t = (W_t^{(1)}, \dots, W_t^{(n_W)})^T$ be an n_W -dimensional WIENER process, and $D \in \mathbb{R}^{n_x \times n_W}$. Then:

$$\int_0^T D dW_s = D \cdot W_T \tag{5.16}$$

Proof: We first show the result for a one-dimensional standard WIENER process $W_t \in \mathbb{R}$ and $D \in \mathbb{R}$. Using definition C.4 for the ITÔ integral, we have for any partition $0 = \tau_0 < \dots < \tau_m = T$:

$$\int_0^T D dW_s = \sum_{j=0}^{m-1} D \cdot \Delta W_j = D \sum_{j=0}^{m-1} (W_{\tau_{j+1}} - W_{\tau_j}) = D \cdot W_T.$$

Thus, for a matrix-valued diffusion constant $D \in \mathbb{R}^{n_x \times n_W}$ and an n_W -dimensional WIENER process $W_t \in \mathbb{R}^{n_W}$, using $\int_0^T D_{ij} dW_s^{(j)} = D_{ij} \cdot W_T^{(j)}$, and using definition C.14 for the multi-dimensional ITÔ-integral, we have

$$\int_0^T D dW_s = \begin{pmatrix} \sum_{j=1}^{n_W} \int_0^T D_{1j} dW_s^{(j)} \\ \vdots \\ \sum_{j=1}^{n_W} \int_0^T D_{n_x j} dW_s^{(j)} \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^{n_W} D_{1j} \cdot W_T^{(j)} \\ \vdots \\ \sum_{j=1}^{n_W} D_{n_x j} \cdot W_T^{(j)} \end{pmatrix} = D \cdot W_T. \quad \square$$

5.5 Lemma (Some expectations on the WIENER process)

Let $W_t = (W_t^{(1)}, \dots, W_t^{(n_W)})^T$ be an n_W -dimensional WIENER process.

1. For the expectation of the (squared) absolute value of W_t , it holds:

$$(a) \quad \mathbb{E}[|W_t|] \leq \sqrt{n_W t} \quad \text{and} \quad (b) \quad \mathbb{E}[|W_t|^2] = n_W t$$

and for a one-dimensional WIENER process, i.e. $n_W = 1$, we have $\mathbb{E}[|W_t^{(1)}|] = \sqrt{\frac{2t}{\pi}}$. We note that the above bound on $\mathbb{E}[|W_t|]$ is sharp for $n_W \rightarrow \infty$.

2. For the expectation of the running (squared) absolute supremum of W_t , it holds:

$$(a) \quad \mathbb{E}\left[\sup_{0 \leq s \leq t} |W_s|\right] \leq 2\sqrt{n_W t} \quad \text{and} \quad (b) \quad \mathbb{E}\left[\sup_{0 \leq s \leq t} |W_s|^2\right] \leq 4n_W t$$

3. For a matrix $D \in \mathbb{R}^{n_x \times n_W}$ and a compatible matrix norm $\|\cdot\|$, it holds:

$$\mathbb{E}[|DW_t|^2] \leq \|D\|^2 n_W t$$

Proof: Using JENSEN's inequality (lemma 5.3), DOOB's maximal inequality for continuous martingales (lemma C.11), the linearity of expectation, and $t = \text{Var}[W_t^{(j)}] = \mathbb{E}[|W_t^{(j)}|^2]$, we find:

1. It holds:

$$(a) \quad \mathbb{E}[|W_t|] = \mathbb{E}\left[\sqrt{\sum_{j=1}^{n_W} |W_t^{(j)}|^2}\right] \leq \sqrt{\mathbb{E}\left[\sum_{j=1}^{n_W} |W_t^{(j)}|^2\right]} = \sqrt{\sum_{j=1}^{n_W} \mathbb{E}[|W_t^{(j)}|^2]} = \sqrt{n_W t}.$$

$$(b) \quad \mathbb{E} \left[|W_t|^2 \right] = \mathbb{E} \left[\sum_{j=1}^{n_W} |W_t^{(j)}|^2 \right] = \sum_{j=1}^{n_W} \mathbb{E} \left[|W_t^{(j)}|^2 \right] = n_W t$$

Further, for $n_W = 1$, using $W_t^{(1)} \sim \mathcal{N}(0, t)$ and $f(x) = \frac{1}{\sqrt{2\pi t}} \cdot e^{-\frac{x^2}{2t}}$, we get

$$\mathbb{E} \left[|W_t^{(1)}| \right] = \int_{-\infty}^{\infty} |x| f(x) dx = 2 \int_0^{\infty} x f(x) dx = \sqrt{\frac{2t}{\pi}} \quad (\leq \sqrt{t})$$

2. For the expectation of the running supremum, we first show (b):

$$(b) \quad \mathbb{E} \left[\sup_{0 \leq s \leq t} |W_s|^2 \right] \leq \mathbb{E} \left[\sum_{j=1}^{n_W} \sup_{0 \leq s \leq t} |W_s^{(j)}|^2 \right] \stackrel{\substack{\mathbb{E} \text{ is} \\ \text{linear}}}{=} \sum_{j=1}^{n_W} \mathbb{E} \left[\sup_{0 \leq s \leq t} |W_s^{(j)}|^2 \right]$$

$$\stackrel{\substack{W_s^{(j)} \\ \text{indep.}}}{=} n_W \cdot \mathbb{E} \left[\sup_{0 \leq s \leq t} |W_s^{(1)}|^2 \right] \stackrel{\substack{\text{lemma} \\ C.11}}{=} n_W \cdot \left(\frac{2}{2-1} \right)^2 \cdot \mathbb{E} \left[|W_s^{(1)}|^2 \right] = 4n_W t$$

$$(a) \quad \mathbb{E} \left[\sup_{0 \leq s \leq t} |W_s| \right] = \mathbb{E} \left[\sup_{0 \leq s \leq t} \left(\sum_{j=1}^{n_W} |W_s^{(j)}|^2 \right)^{\frac{1}{2}} \right] = \mathbb{E} \left[\left(\sup_{0 \leq s \leq t} \sum_{j=1}^{n_W} |W_s^{(j)}|^2 \right)^{\frac{1}{2}} \right]$$

$$\leq \mathbb{E} \left[\sup_{0 \leq s \leq t} \sum_{j=1}^{n_W} |W_s^{(j)}|^2 \right]^{\frac{1}{2}} \stackrel{(b)}{\leq} 2\sqrt{n_W t}$$

3. Using the linearity of expectation and the compatibility of the matrix norm, we get

$$\mathbb{E} \left[|DW_t|^2 \right] \leq \mathbb{E} \left[\|D\|^2 \cdot \sum_{j=1}^{n_W} (W_t^{(j)})^2 \right] \leq \|D\|^2 \cdot \sum_{j=1}^{n_W} \mathbb{E} \left[(W_t^{(j)})^2 \right] = \|D\|^2 n_W t$$

□

5.3.2 Distance of S-IVP and D-IVP solutions

We formulate some results for one-dimensional, time-homogeneous (autonomous) systems. These results will be generalized to arbitrary finite dimension later on.

Though most of these distance results are already known, the focus in literature usually lies on estimating the probabilities for large deviations (*large deviation theory for stochastic processes* based on the work of FREIDLIN and WENTZELL [FreidlinWentzell1998]). We therefore give proofs suited to the purposes here.

We start with some results that we'll rely upon in later proofs.

5.6 Lemma (A quadratic inequality)

For any $a_i \in \mathbb{R}^d$ ($i = 1, \dots, n$; $d \in \mathbb{N}$), it holds:

$$\left| \sum_{i=1}^n a_i \right|^2 \leq 2^{n-1} \sum_{i=1}^n |a_i|^2 \quad (5.17)$$

Proof: We first proof the result for scalar values $\alpha_i \in \mathbb{R}$ ($i = 1, \dots, n$) by induction:

Case $n = 1$ is obvious. For $n = 2$, from $|a_1|^2 - 2|a_1||a_2| + |a_2|^2 = ||a_1| - |a_2||^2 \geq 0$, by addition of $|a_1|^2 + 2|a_1||a_2| + |a_2|^2$, it follows $2|a_1|^2 + 2|a_2|^2 \geq |a_1|^2 + 2|a_1||a_2| + |a_2|^2 = (|a_1| + |a_2|)^2 \geq |a_1 + a_2|^2$. Let now eq. (5.17) hold for $n - 1$. Then:

$$\left| \sum_{i=1}^n a_i \right|^2 = \left| \sum_{i=1}^{n-1} a_i + a_n \right|^2 \stackrel{n=2}{\leq} 2 \left(\left| \sum_{i=1}^{n-1} a_i \right|^2 + |a_n|^2 \right) \stackrel{i.h.}{\leq} 2 \cdot 2^{n-2} \sum_{i=1}^{n-1} |a_i|^2 + 2|a_n|^2 \leq 2^{n-1} \sum_{i=1}^n |a_i|^2 .$$

Now, for vectors $\alpha_i \in \mathbb{R}^d$ ($i = 1, \dots, n$), we have:

$$\left| \sum_{i=1}^n a_i \right|^2 = \sum_{j=1}^d \left| \sum_{i=1}^n a_{i,j} \right|^2 \stackrel{\text{scalar}}{\leq} \sum_{j=1}^d 2^{n-1} \sum_{i=1}^n |a_{i,j}|^2 = 2^{n-1} \sum_{i=1}^n \sum_{j=1}^d |a_{i,j}|^2 = 2^{n-1} \sum_{i=1}^n |a_i|^2 . \quad \square$$

5.7 Lemma (GRONWALL-BELLMAN inequality)

Let u and β be continuous, non-negative functions on $I := [t_0, t_f]$, and α be a continuous, positive function and monotonously increasing function on I . Then, the following implication holds:

$$u(t) \leq \alpha(t) + \int_{t_0}^t \beta(s)u(s) ds \quad (t \in [t_0, t_f]) \quad \implies \quad u(t) \leq \alpha(t)e^{\int_{t_0}^t \beta(s) ds} \quad (t \in [t_0, t_f]) .$$

If, as a special case, $\alpha(t) \equiv a > 0$ and $\beta(t) \equiv b \geq 0$ are constants and $[t_0, t_f] = [0, T]$, this reads as

$$u(t) \leq a + b \int_0^t u(s) ds \quad (t \in [0, T]) \quad \implies \quad u(t) \leq ae^{bt} \quad (t \in [0, T]) .$$

Proof: See, e.g., [Pachpatte1998], Theorem 1.3.1. □

5.8 Lemma (CAUCHY-BUNYAKOVSKY-SCHWARZ inequality)

For square integrable real-valued functions $f, g: \mathbb{R} \rightarrow \mathbb{R}$, it holds

$$\left| \int_{\mathbb{R}} f(x)g(x) dx \right|^2 \leq \int_{\mathbb{R}} |f(x)|^2 dx \cdot \int_{\mathbb{R}} |g(x)|^2 dx .$$

Proof: See [Bouniakowsky1859]. □

5.9 Lemma (Distance between SDE and ODE solution - 1D)

Assume that the S-IVP (5.10) and D-IVP (5.11) are one-dimensional and time-homogeneous or autonomous, respectively. Then, it holds: The difference Δ_t at time t of the solution X_t of

S-IVP and the solution $x(t)$ of the corresponding D-IVP is bounded by

$$|\Delta_t| \leq \left(|X_0^{(s)} - X_0^{(d)}| + D|W_t| \right) \cdot e^{Lt}, \quad (t \in [0, T]) \quad (5.18)$$

where L denotes the LIPSCHITZ constant of the drift function f , and D is the SDE's diffusion. Note that $|X_0^{(s)} - X_0^{(d)}| = |\Delta_0|$ is the initial deviation of SDE and ODE solution.

Proof: Using $\int_0^t dW_s = W_t$ and the LIPSCHITZ continuity of f , we get:

$$\begin{aligned} |\Delta_t| &= |X_t - x(t)| \\ &= \left| (X_0^{(s)} - X_0^{(d)}) + \int_0^t f(X_s) - f(x(s)) ds + \int_0^t D dW_s \right| \\ &\leq |X_0^{(s)} - X_0^{(d)}| + \int_0^t |f(X_s) - f(x(s))| ds + D|W_t| \\ &\leq |X_0^{(s)} - X_0^{(d)}| + L \cdot \int_0^t |X_s - x(s)| ds + D|W_t| \\ &= |X_0^{(s)} - X_0^{(d)}| + D|W_t| + L \cdot \int_0^t |\Delta_s| ds \end{aligned} \quad (5.19)$$

such that we can use the GRONWALL-BELLMAN inequality (lemma 5.7) to get

$$|\Delta_t| \leq \left(|X_0^{(s)} - X_0^{(d)}| + D|W_t| \right) \cdot e^{Lt} = (|\Delta_0| + D|W_t|) \cdot e^{Lt}.$$

□

5.10 Corollary (Maximum deviation between SDE and ODE solution - 1D)

For the maximum difference between X_t and $x(t)$ on the whole interval $[0, T]$, it holds:

$$\Delta_T^* = \sup_{0 \leq t \leq T} |\Delta_t| \leq \left(|X_0^{(s)} - X_0^{(d)}| + DW_T^* \right) \cdot e^{LT}. \quad (5.20)$$

where again L denotes the LIPSCHITZ constant of the drift function f .

Proof: From inequality (5.19), using the monotonicity of the maximum process, we get:

$$\begin{aligned} \Delta_t^* &= \sup_{0 \leq s \leq t} |\Delta_s| \leq \sup_{0 \leq s \leq t} (|\Delta_0| + D|W_s|) + \sup_{0 \leq s \leq t} L \cdot \int_0^s |\Delta_r| dr \\ &\leq (|\Delta_0| + DW_t^*) + L \cdot \int_0^t \Delta_r^* dr \end{aligned}$$

for all $t \in [0, T]$. Application of the GRONWALL-BELLMAN inequality gives the desired result:

$$\Delta_t^* \leq (|\Delta_0| + DW_t^*) \cdot e^{Lt} \quad \forall t \in [0, T].$$

□

5.11 Corollary (Difference between SDE and ODE solution with same initial value - 1D)

If S-IVP (5.10) and D-IVP (5.11) share the same non-random initial value $X_0^{(s)} = X_0^{(d)} = x_0$, then it holds for the difference process:

$$|\Delta_t| \leq D|W_t| \cdot e^{Lt} \quad (t \in [0, T]) \quad \text{and} \quad \Delta_T^* \leq DW_T^* \cdot e^{LT} .$$

Proof: Immediately follows from lemma 5.9 and corollary 5.10. \square

5.12 Lemma (Distance and maximum deviation between SDE and ODE solution - nD)

Let W_t be an n_W -dimensional WIENER process, and let

$$\begin{aligned} X_t &= X_0^{(s)} + \int_0^t f(X_s) ds + \int_0^t D dW_s && \text{with } X_0^{(s)} \in \mathbb{R}^{n_x}, W_t \in \mathbb{R}^{n_W}, D \in \mathbb{R}^{n_x \times n_W} \\ x(t) &= X_0^{(d)} + \int_0^t f(x(s)) ds && \text{with } X_0^{(d)} \in \mathbb{R}^{n_x} \end{aligned}$$

be the solutions of S-IVP (5.10) and of D-IVP (5.11), W.l.o.g. assuming that S-IVP and D-IVP are time-homogeneous or autonomous, respectively. Then, we have:

1. For any compatible matrix norm $\|\cdot\|$, it holds

$$|\Delta_t| \leq (|\Delta_0| + \|D\| |W_t|) \cdot e^{Lt}, \quad (t \in [0, T]) \quad (5.21)$$

where L denotes the LIPSCHITZ constant of the drift function f , and $\Delta_0 = X_0^{(s)} - X_0^{(d)}$.

2. For the maximum deviation between the SDE and ODE solution in $[0, T]$, it holds

$$\Delta_T^* \leq (|\Delta_0| + \|D\| W_T^*) \cdot e^{LT} . \quad (5.22)$$

3. If S-IVP and D-IVP share the same initial value, i.e. $|\Delta_0| = |X_0^{(s)} - X_0^{(d)}| = 0$, we have

$$\Delta_T^* \leq \|D\| W_T^* \cdot e^{LT} \quad (\text{for } \Delta_0 = 0) . \quad (5.23)$$

Proof: 1. is shown as in lemma 5.9, using now lemma 5.4 to get $\int_0^t D dW_s = DW_t$, and the compatibility $|DW_t| \leq \|D\| |W_t|$ of the matrix norm. 2. and 3. are proved analogously to corollary 5.10, using 1. and again the compatibility of the matrix norm. \square

The above results deliver an estimate on the difference process Δ_t between the SDE solution X_t and the ODE solution $x(t)$ at any time $t \in [0, T]$, and therefore also an upper bound on the maximum deviation Δ_T^* over the whole time domain $[0, T]$.

If the time domain is small, i.e. for $T \rightarrow 0$, the exponentials in eqs. (5.22) and (5.23) go to 1, but the maximum WIENER process W_T^* has a positive probability of exceeding any given value $\delta > 0$, though “intuitively”, one expects it to be small, as $W_0^* = W_0 = 0$. However, the bounds given in lemma 5.12 are themselves random variates and as such not very useful.

Finally, in the following theorems, we prove the stronger result that the solutions stay close to each other in expectation and in mean square, that is – in this case – intuition is not misleading (as it frequently is in probability theory).

5.13 Theorem (Convergence of SDE and ODE solutions in expectation)

Let $W_t = (W_t^{(1)}, \dots, W_t^{(n_W)})^T$ be an n_W -dimensional WIENER process. With the same notation and prerequisites as in lemma 5.12, it holds for the expected difference between S-IVP and D-IVP solutions X_t and $x(t)$:

$$\mathbb{E}[|\Delta_t|] \leq (|\Delta_0| + \|D\| \sqrt{n_W t}) \cdot e^{Lt}$$

and for the maximum difference:

$$\mathbb{E}[\Delta_T^*] \leq (|\Delta_0| + 2\|D\| \sqrt{n_W T}) \cdot e^{LT}$$

with L being the LIPSCHITZ constant of the drift function f and D the SDE's diffusion matrix. Further,

$$\mathbb{E}[|\Delta_t|] \xrightarrow{t \rightarrow 0} 0 \quad \text{if} \quad X_0^{(s)} = X_0^{(d)}.$$

i.e. for identical initial values, the SDE and ODE solutions converge in expectation for $t \rightarrow 0$.

Proof: Using the linearity of expectation in eq. (5.21) or eq. (5.22), respectively, and applying lemma 5.5, i.e. $\mathbb{E}[|W_t|] \leq \sqrt{n_W t}$ or $\mathbb{E}[W_t^*] \leq 2\sqrt{n_W t}$.

Further, for same initial values, $|\Delta_0| = 0$ and $\|D\| \sqrt{n_W t} \cdot e^{Lt} \rightarrow 0$ ($t \rightarrow 0$). □

5.14 Lemma (An integral estimate)

For X_t and $x(t)$, the solutions of S-IVP (5.10) and D-IVP (5.11), respectively, it holds:

$$\left| \int_0^t f(X_s) - f(x(s)) \, ds \right|^2 \leq t \cdot \int_0^t |f(X_s) - f(x(s))|^2 \, ds$$

Proof: The CAUCHY-BUNYAKOVSKY-SCHWARZ inequality (lemma 5.8), applied component-wise, gives

$$\left| \int_0^t f(X_s) - f(x(s)) \, ds \right|^2 \leq \int_0^t |f(X_s) - f(x(s))|^2 \, ds \cdot \int_0^t 1 \, ds = t \cdot \int_0^t |f(X_s) - f(x(s))|^2 \, ds$$

□

5.15 Theorem (Convergence of SDE and ODE solutions in mean square)

Let $W_t = (W_t^{(1)}, \dots, W_t^{(n_W)})^T$ be an n_W -dimensional WIENER process. With the same notation and prerequisites as in lemma 5.12, it holds

$$\mathbb{E}[|\Delta_t|^2] \leq \left(4|\Delta_0|^2 + 4\|D\|^2 n_W t \right) \cdot e^{4t^2 L^2}$$

where D is the SDE's diffusion matrix, $\|\cdot\|$ denotes any compatible matrix norm and L is the LIPSCHITZ constant of the drift function f .

Further,

$$\mathbb{E}\left[|\Delta_t|^2\right] \xrightarrow{t \rightarrow 0} 0 \quad \text{if } X_0^{(s)} = X_0^{(d)}.$$

i.e. for identical initial values, the SDE and ODE solutions converge in mean square for $t \rightarrow 0$.

Proof: It holds for the expected squared deviation:

$$\begin{aligned} \mathbb{E}\left[|X_t - x(t)|^2\right] &= \mathbb{E}\left[\left|X_0^{(s)} - X_0^{(d)} + \int_0^t f(X_s) - f(x(s)) ds + DW_t\right|^2\right] \\ &\stackrel{\text{lemma 5.6}}{\leq} \mathbb{E}\left[4|X_0^{(s)} - X_0^{(d)}|^2 + 4\left|\int_0^t f(X_s) - f(x(s)) ds\right|^2 + 4|DW_t|^2\right] \\ &\stackrel{\substack{\mathbb{E} \text{ is} \\ \text{linear}}}{=} 4 \underbrace{\mathbb{E}\left[|X_0^{(s)} - X_0^{(d)}|^2\right]}_{\substack{=|X_0^{(s)} - X_0^{(d)}|^2 \\ \text{(non-random)}}} + 4 \mathbb{E}\left[\left|\int_0^t f(X_s) - f(x(s)) ds\right|^2\right] + 4 \mathbb{E}\left[|DW_t|^2\right] \\ &\stackrel{\substack{\text{lemmata} \\ 5.14 + 5.5}}{\leq} 4|X_0^{(s)} - X_0^{(d)}|^2 + 4t \cdot \mathbb{E}\left[\int_0^t |f(X_s) - f(x(s))|^2 ds\right] + 4t \cdot \|D\|^2 n_W \\ &\stackrel{\substack{\text{FUBINI,} \\ f \text{ is LIPSCHITZ}}}{\leq} 4|X_0^{(s)} - X_0^{(d)}|^2 + 4tL^2 \cdot \int_0^t \mathbb{E}\left[|X_s - x(s)|^2\right] ds + 4t \cdot \|D\|^2 n_W \\ &= \underbrace{4|X_0^{(s)} - X_0^{(d)}|^2 + 4t \cdot \|D\|^2 n_W}_{=: \alpha(t)} + \int_0^t \underbrace{4tL^2}_{=: \beta(s)} \cdot \underbrace{\mathbb{E}\left[|X_s - x(s)|^2\right]}_{=: u(s)} ds \end{aligned}$$

with monotonically increasing $\alpha(t)$, such that the GRONWALL-BELLMAN inequality (lemma 5.7) gives with α , β , and u as above:

$$\mathbb{E}\left[|X_t - x(t)|^2\right] \leq \underbrace{(4|X_0^{(s)} - X_0^{(d)}|^2 + 4t \cdot \|D\|^2 n_W)}_{\rightarrow 0 \text{ (} t \rightarrow 0)} \cdot \underbrace{e^{4t^2 L^2}}_{\rightarrow 1 \text{ (} t \rightarrow 0)} \xrightarrow{t \rightarrow 0} 4|X_0^{(s)} - X_0^{(d)}|^2$$

tending to zero if the initial values $X_0^{(s)}$ and $X_0^{(d)}$ coincide. \square

A similar result is obtained for the maximum mean square deviation between S-IVP and D-IVP solution *over the whole interval* $[0, T]$, as given in the next theorem.

5.16 Theorem (Maximum distance of SDE and ODE solutions in mean square)

Let $W_t = (W_t^{(1)}, \dots, W_t^{(n_W)})^T$ be an n_W -dimensional WIENER process. With the same notation and prerequisites as in lemma 5.12, it holds for the maximal mean square distance between S-IVP and D-IVP solution over the whole interval $[0, T]$:

$$\mathbb{E} \left[|\Delta_T^*|^2 \right] = \mathbb{E} \left[\sup_{0 \leq t \leq T} |\Delta_t|^2 \right] \leq \left(4|X_0^{(s)} - X_0^{(d)}|^2 + 16 \|D\|^2 n_W T \right) \cdot e^{4T^2 L^2} \quad (5.24)$$

where D is the SDE's diffusion matrix, $\|\cdot\|$ denotes a compatible matrix norm and L is the LIPSCHITZ constant of the drift function f .

Proof: First, $|\Delta_T^*|^2 = \left(\sup_{0 \leq t \leq T} |\Delta_t| \right)^2 = \sup_{0 \leq t \leq T} |\Delta_t|^2$, as $|\Delta_t| \geq 0$. Further, for every $t \in [0, T]$:

$$\begin{aligned} \mathbb{E} \left[\sup_{0 \leq s \leq t} |\Delta_s|^2 \right] &= \mathbb{E} \left[\sup_{0 \leq s \leq t} \left| X_0^{(s)} - X_0^{(d)} + \int_0^s f(X_r) - f(x(r)) dr + \int_0^s D dW_r \right|^2 \right] \\ &\stackrel{\text{lemma 5.6}}{\leq} \mathbb{E} \left[\sup_{0 \leq s \leq t} \left\{ 4|X_0^{(s)} - X_0^{(d)}|^2 + 4 \left| \int_0^s f(X_r) - f(x(r)) dr \right|^2 + 4 \left| \int_0^s D dW_r \right|^2 \right\} \right] \\ &\stackrel{\text{lemma 5.14}}{\leq} \mathbb{E} \left[4|\Delta_0|^2 + \sup_{0 \leq s \leq t} 4s \int_0^s |f(X_r) - f(x(r))|^2 dr + \sup_{0 \leq s \leq t} 4|DW_s|^2 \right] \\ &\stackrel{s \leq t}{\leq} 4|\Delta_0|^2 + \mathbb{E} \left[\sup_{0 \leq s \leq t} 4t \int_0^s |f(X_r) - f(x(r))|^2 dr \right] + \mathbb{E} \left[\sup_{0 \leq s \leq t} 4\|D\|^2 |W_s|^2 \right] \\ &\stackrel{\text{lemma 5.5}}{\leq} 4|\Delta_0|^2 + \mathbb{E} \left[\sup_{0 \leq s \leq t} 4t \int_0^s L^2 |X_r - x(r)|^2 dr \right] + 4\|D\|^2 \cdot 4n_W t \\ &\stackrel{f \text{ LIPSCHITZ}}{\leq} 4|\Delta_0|^2 + 16\|D\|^2 n_W t + 4L^2 t \cdot \mathbb{E} \left[\int_0^t \sup_{0 \leq s \leq r} |\Delta_s|^2 dr \right] \\ &\stackrel{\text{FUBINI}}{=} \underbrace{4|\Delta_0|^2 + 16\|D\|^2 n_W t}_{=: \alpha(t)} + \int_0^t \underbrace{4L^2 r}_{=: \beta(r)} \underbrace{\mathbb{E} \left[\sup_{0 \leq s \leq r} |\Delta_s|^2 \right]}_{=: u(r)} dr . \end{aligned}$$

With continuous, positive, monotonically increasing α and continuous β and u as above, the GRONWALL-BELLMAN inequality (lemma 5.7) delivers eq. (5.24), completing the proof. \square

5.4 A piecewise deterministic parameter estimation method

In this section, we first establish the notation needed for the formulation of the stochastic parameter estimation problem in section 5.4.3 on page 143. The new method will be derived on the basis of jump regularization in the ODE parameter estimation problem discussed in chapter 2.

The same notation will be used for the numerical analysis in the subsequent section 5.5.

5.4.1 Notation and assumptions

5.4.1.1 Stochastic differential equation model

In what follows, we assume that observations originate from a *single realization* of a system whose dynamics follow a stochastic differential equation with *additive noise*:

$$dX_t = f(t, X_t, p) dt + D dW_t \quad (t \in [t_0, t_f])$$

with *drift* $f: [t_0, t_f] \times \mathbb{R}^{n_x} \times \mathbb{R}^{n_p} \rightarrow \mathbb{R}^{n_x}$, fulfilling the prerequisites of theorem 4.28 to ensure the existence of strong solutions, and constant *diffusion* $D \in \mathbb{R}^{n_x \times n_w}$ with an n_w -dimensional driving WIENER process $W_t \in \mathbb{R}^{n_w}$.

The initial state X_{t_0} is not necessarily known. Further, we assume that the (stochastic) noise is “small” or “non-dominating”, in the interpretation that, for a small time-horizon, the system is well approximated by its corresponding ODE

$$dx = f(t, x(t), p) dt \quad (t \in [\tau, \tau + \Delta\tau]).$$

5.4.1.2 Shooting grid \mathbf{T}^{MS}

The time horizon $[t_0, t_f]$ is partitioned in n_{MS} intervals, not necessarily of the same size, at the time points $t_0 = t_0^{\text{MS}} < t_1^{\text{MS}} < \dots < t_{n_{\text{MS}}}^{\text{MS}} = t_f$, collected into the *shooting grid*

$$\mathbf{T}^{\text{MS}} = \{t_0^{\text{MS}}, \dots, t_{n_{\text{MS}}}^{\text{MS}}\}.$$

As a convention, we say interval k starts at the *shooting node* (time point) t_k^{MS} , i.e. we start counting the intervals at 0 (zero).

5.4.1.3 Shooting variables s_k

The initial values (states) $s_k \in \mathbb{R}^{n_x}$ at the shooting nodes t_k ($k = 0, \dots, n_{\text{MS}} - 1$) are combined into the vector s :

$$s = \begin{pmatrix} s_0 \\ \vdots \\ s_{n_{\text{MS}}-1} \end{pmatrix} \in \mathbb{R}^{n_{\text{MS}} \cdot n_x}, \quad s_k = \begin{pmatrix} s_{k,1} \\ \vdots \\ s_{k,n_x} \end{pmatrix} \in \mathbb{R}^{n_x} \quad (k = 0, \dots, n_{\text{MS}} - 1).$$

Note that there is no shooting variable at the last node $t_{n_{\text{MS}}}^{\text{MS}}$.

5.4.1.4 Parameter vector p

The (unknown) parameters of the deterministic drift term $f(t, x(t), p)$ and of the measurement functions $h_i(t_i, x(t_i; p), p)$ are combined in the vector p :

$$p = (p_1, \dots, p_{n_p}) \in \mathbb{R}^{n_p} .$$

5.4.1.5 Measurement times t_i and selector functions \underline{l} , \bar{l} , ι , κ

We assume that measurements of the stochastic process have been taken at not necessarily distinct time point t_i , which we store in the *multiset*

$$\mathbb{T}^M := \{t_1, \dots, t_{n_M}\}.$$

We further define the *selector functions* for the k -th shooting interval ($k = 0, \dots, n_{\text{MS}} - 1$)

$$\underline{l}(k) := \min \{i \in \{1, \dots, n_M\} \mid t_k^{\text{MS}} \leq t_i < t_{k+1}^{\text{MS}}\} \quad (5.25a)$$

$$\bar{l}(k) := \begin{cases} \max \{i \in \{1, \dots, n_M\} \mid t_k^{\text{MS}} \leq t_i < t_{k+1}^{\text{MS}}\} & \text{for } k = 0, \dots, n_{\text{MS}} - 2 \\ \max \{i \in \{1, \dots, n_M\} \mid t_k^{\text{MS}} \leq t_i \leq t_{k+1}^{\text{MS}}\} & \text{for } k = n_{\text{MS}} - 1 \end{cases} \quad (5.25b)$$

$$\iota(k) := \{i \mid \underline{l}(k) \leq i \leq \bar{l}(k)\} \quad (5.25c)$$

$$\kappa(t) := \min \{k \in \{0, \dots, n_{\text{MS}} - 1\} \mid t_k^{\text{MS}} \leq t\} \quad (5.25d)$$

that give the first index, eq. (5.25a), and the last index, eq. (5.25b), of the measurement times t_i that lie in the k -th shooting interval, as well as the whole set of indices i that belong to measurements in the k -th interval, eq. (5.25c). The last selector function in eq. (5.25d) returns for a given time point t the respective shooting interval in which this time point falls.

We mention that the above definition of the selector functions \underline{l} and \bar{l} implies that measurements taken at an inner shooting node t_k^{MS} are always associated with the k -th interval that begins at this node. This will be discussed in section 5.4.1.10 on page 142.

5.4.1.6 Measurement function h_i , measurements η_i , measurement variances σ_i^2

As in the ODE case, we assume having n_M scalar *measurements* $\eta_i \in \mathbb{R}$ ($i = 1, \dots, n_M$), taken at the not necessarily distinct time points t_i . The measurements may be affected by normally distributed and independent measurement errors $\varepsilon_i \sim \mathcal{N}(0, \sigma_i^2)$ ($i = 1, \dots, n_M$).

If X_t is the *true* stochastic process, we may write the measurements as

$$\eta_i = h_i(t_i, X_{t_i}, p) + \varepsilon_i, \quad \varepsilon_i \sim \mathcal{N}(0, \sigma_i^2) \quad (i = 1, \dots, n_M)$$

using the scalar *measurement functions* $h_i: \mathbb{R} \times \mathbb{R}^{n_x} \times \mathbb{R}^{n_p} \rightarrow \mathbb{R}$.

We denote by $h^{(k)}$ the vector of all measurements in the k -th shooting interval, i.e.

$$h^{(k)} = h^{(k)}(s_k, p) = \begin{pmatrix} h_{\underline{l}(k)}(t_{\underline{l}(k)}, x(t_{\underline{l}(k)}; p, t_k^{\text{MS}}, s_k), p) \\ \vdots \\ h_{\bar{l}(k)}(t_{\bar{l}(k)}, x(t_{\bar{l}(k)}; p, t_k^{\text{MS}}, s_k), p) \end{pmatrix} \in \mathbb{R}^{|\iota(k)|} \quad (k = 0, \dots, n_{\text{MS}} - 1)$$

and collect them in the complete observation vector h :

$$h = h(s, p) = \left(h^{(0)}(s_0, p)^T, \dots, h^{(n_{\text{MS}}-1)}(s_{n_{\text{MS}}-1}, p)^T \right)^T \in \mathbb{R}^{n_M} .$$

Analogously, we collect the measurements in the k -th shooting interval into $\eta^{(k)}$, combine all measurements in the vector η , and do the same combination for the respective standard deviations in a per-interval-wise manner in $\Sigma^{(k)}$ and in total in Σ :

$$\begin{aligned}\eta^{(k)} &:= (\eta_{\underline{t}(k)}, \dots, \eta_{\bar{t}(k)})^T \in \mathbb{R}^{|\iota(k)|}, & \eta &:= \left((\eta^{(0)})^T, \dots, (\eta^{(n_{\text{MS}}-1)})^T \right)^T \in \mathbb{R}^{n_M} \\ \Sigma^{(k)} &:= \text{diag} \{ \sigma_{\underline{t}(k)}, \dots, \sigma_{\bar{t}(k)} \} \in \mathbb{R}^{|\iota(k)| \times |\iota(k)|}, & \Sigma &:= \text{diag} \{ \Sigma^{(0)}, \dots, \Sigma^{(n_{\text{MS}}-1)} \} \in \mathbb{R}^{n_M \times n_M}\end{aligned}$$

for $k = 0, \dots, n_{\text{MS}} - 1$.

5.4.1.7 Trajectory $x(t)$ or $x(t; s, p)$

For $t \in [t_k^{\text{MS}}, t_{k+1}^{\text{MS}}]$, we denote by $x(t; t_k^{\text{MS}}, s_k, p)$ the *trajectory* on the closed interval $[t_k^{\text{MS}}, t_{k+1}^{\text{MS}}]$, i.e. $x(t; t_k^{\text{MS}}, s_k, p)$ is the unique solution of the IVP

$$\dot{x}(t) = f(t, x(t), p), \quad x(t_k^{\text{MS}}) = s_k, \quad t \in [t_k^{\text{MS}}, t_{k+1}^{\text{MS}}] \quad (k = 0, \dots, n_{\text{MS}} - 1).$$

Similar as in the ODE case in eq. (2.12), we define $x(t) = x(t; s, p)$ to be the concatenation of the *interval solutions*:

$$x(t; s, p) := x(t; t_k^{\text{MS}}, s_k, p) \quad \text{for } t \in [t_k^{\text{MS}}, t_{k+1}^{\text{MS}}] \quad (k = 0, \dots, n_{\text{MS}} - 1) \quad (5.26)$$

and set $x(t_{n_{\text{MS}}}^{\text{MS}}; s, p) := x(t_{n_{\text{MS}}}^{\text{MS}}; t_{n_{\text{MS}}-1}^{\text{MS}}, s_{n_{\text{MS}}-1}, p)$ at the last shooting node. Doing so ensures that $x(t; s, p)$ is càdlàg on $[t_0, t_f]$. Also see section 5.4.1.10 for the values at the shooting nodes.

5.4.1.8 Stochastic jumps α_k and jump regularization weights ω_k^2

Anticipating definition 5.17, we denote by α_k the discontinuity at the *inner* shooting grid point t_k^{MS} ($k = 1, \dots, n_{\text{MS}} - 1$), and call it *stochastic jump*, i.e.

$$\alpha_k := x(t_k^{\text{MS}}; t_{k-1}^{\text{MS}}, s_{k-1}, p) - s_k \in \mathbb{R}^{n_x} \quad (k = 1, \dots, n_{\text{MS}} - 1)$$

and combine them into the vector α :

$$\alpha = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_{n_{\text{MS}}-1} \end{pmatrix} \in \mathbb{R}^{(n_{\text{MS}}-1) \cdot n_x}, \quad \alpha_k = \begin{pmatrix} \alpha_{k,1} \\ \vdots \\ \alpha_{k,n_x} \end{pmatrix} \in \mathbb{R}^{n_x} \quad (k = 1, \dots, n_{\text{MS}} - 1).$$

We endow every jump $\alpha_{k,l}$ with a *jump regularization weight* $\omega_{k,l}^2 \in \mathbb{R}_0^+$, and build the weighting matrices

$$\begin{aligned}\Omega^{(k)} &:= \text{diag} \{ \omega_{k,1}, \dots, \omega_{k,n_x} \} \in \mathbb{R}^{n_x \times n_x} \quad (k = 1, \dots, n_{\text{MS}} - 1) \\ \Omega &:= \text{diag} \{ \Omega^{(1)}, \dots, \Omega^{(n_{\text{MS}}-1)} \} \in \mathbb{R}^{((n_{\text{MS}}-1)n_x) \times ((n_{\text{MS}}-1)n_x)}.\end{aligned}$$

5.4.1.9 Equality and inequality point constraints c^{ec} and c^{ic}

As in section 2.3.1.1, *point constraints* are described by

$$c^{ec}(x(t_1^{ec}; p), \dots, x(t_{n_{\text{EC}}}^{ec}; p), p) = 0 \quad \text{and} \quad c^{ic}(x(t_1^{ic}; p), \dots, x(t_{n_{\text{IC}}}^{ic}; p), p) \geq 0$$

with dimensions $c^{ec}(\cdot) \in \mathbb{R}^{n_{ec}}$ and $c^{ic}(\cdot) \in \mathbb{R}^{n_{ic}}$.

5.4.1.10 Measurements at inner shooting nodes

The piecewise assembled deterministic trajectory $x(t)$ ($t \in [t_0, t_f]$), defined in eq. (5.26), is, in general, discontinuous at the inner shooting nodes $\{t_1^{\text{MS}}, \dots, t_{n_{\text{MS}}-1}^{\text{MS}}\}$.

For measurements taken at these nodes thus arises the question which shooting interval they should be associated with, i.e., shall we compute the residual by comparing the measurement with the simulated measurement taken on the left limit $\lim_{t \nearrow t_k^{\text{MS}}} x(t; t_{k-1}^{\text{MS}}, s_{k-1}, p)$ or on the right limit $\lim_{t \searrow t_k^{\text{MS}}} x(t; t_k^{\text{MS}}, s_k, p) = s_k$?

Since the original (unknown) SDE trajectory is càdlàg, we associate, as a convention, measurements taken at an inner shooting node t_k^{MS} with the k -th interval; that means we compare it to the simulated measurement taken at the right limit $\lim_{t \searrow t_k^{\text{MS}}} x(t; t_k^{\text{MS}}, s_k, p) = s_k$.

This choice is also reflected in the definition of the selector functions \underline{t} and \bar{t} in eq. (5.25).

5.4.2 A regularization of the multi-experiment case

As already mentioned in remark 2.9 on page 36, SCHLÖDER gives an extension of the ODE parameter estimation problem 2.8 for the multi-experiment case, in a general and comprehensive formulation with multiple trajectories $x^{(k)}$ describing different experiments or experimental set-ups, and governed by differential equations $\dot{x}^{(k)}(t) = f^{(k)}(t, x^{(k)}(t), p)$, i.e. each experiment might follow different kinetics [Schloeder1987].

Omitting the parts for stationarity and jump⁽⁵⁾ conditions, the multi-experiment problem for parameter estimation with data from m_d distinct experiments reads as

$$\begin{aligned} \min_{x,p} \quad & \frac{1}{2} \sum_{k=0}^{m_d-1} \left\| r^{(k)} \left(x^{(k)}(t_1^{(k)}), \dots, x^{(k)}(t_M^{(k)}), p \right) \right\|_2^2 \\ \text{s.t.} \quad & \dot{x}^{(k)} = f^{(k)}(t, x^{(k)}, p) \quad t \in [t_0^{(k)}, t_f^{(k)}] \\ & c^{ec} \left(x^{(k)}(t_1^{(k)}), \dots, x^{(k)}(t_M^{(k)}), p \right) = 0 \\ & c^{ic} \left(x^{(k)}(t_1^{(k)}), \dots, x^{(k)}(t_M^{(k)}), p \right) \geq 0 \end{aligned} \quad (5.27)$$

The ODE parameter estimation problem 2.7 is a special case of eq. (5.27) with $m_d = 1$. We also may interpret parameter estimation problems in SDE as a multi-experiment problem, in which the individual experiments are run sequentially one after another, i.e. the SDE time horizon $[t_0, t_f]$ is split by the grid points $t_0 = t_0^{\text{MS}} < t_1^{\text{MS}} < \dots < t_{n_{\text{MS}}}^{\text{MS}} = t_f$ into $m_d = n_{\text{MS}}$ intervals, such that the intervals' boundaries in eq. (5.27) are given as $t_0^{(k)} = t_k^{\text{MS}}$ and $t_f^{(k)} = t_{k+1}^{\text{MS}}$ for $k = 0, \dots, n_{\text{MS}} - 1$. The individual experiments follow the same kinetic, i.e. they share the same r.h.s. function $f^{(k)}(t, x, p) \equiv f(t, x, p) \forall k$.

⁽⁵⁾ The jump conditions in the multi-experiment formulation of [Schloeder1987] describe discontinuous transitions between different steady states x^s and require stationarity, i.e. a vanishing r.h.s. function $f(t, x^s, p) = 0$, as well as an instability, i.e. $\frac{d}{dx} f(t, x^s, p)$ must be singular. This kind of jumps is different from the ones considered in this thesis.

The multi-experiment formulation also allows for discontinuities in the k -th experiment trajectory, described by $x^{(k)}(\tau_j^+) = x^{(k)}(\tau_j^-) + S(\tau_j^k, x^{(k)}(\tau_j^k), p)$ with a known update function S at time points τ_j^k (given explicitly or characterized implicitly by $Z(\tau_j^k, x^{(k)}(\tau_j^k), p) = 0$). However, the original SDE trajectory is continuous and the update function S is unavailable.

For intervals without observations, this leads into singular problems, since the trajectory, say on the k -th interval, $x^{(k)}(t)$ may not be computed solely from the differential equation (i.e. without, for example, an initial value). One remedy would be to abstain a trajectory in intervals without observations, but that loses information: we know that the state value at the end of the k -th interval is a good initial value for the subsequent interval $k + 1$. And this information should be propagated to subsequent intervals.

Thus, to avoid singularity but maintain and use the available information, we may add a **jump regularization term** to the objective of eq. (5.27), formulated as a sum of least squares,

$$\frac{1}{2} \sum_{k=1}^{n_{\text{MS}}-1} \left\| \Omega \cdot (x^{(k-1)}(t_k^{\text{MS}}) - x^{(k)}(t_k^{\text{MS}})) \right\|_2^2$$

with some suitable weighting matrix $\Omega \in \mathbb{R}^{n_x \times n_x}$, forcing the initial values of each interval to stay “nearby” the final value of the preceding interval.

We do this for all intervals, as we approximate the SDE realization by solutions of ODEs, leading to the *stochastic parameter estimation problem* formulated in the next section.

5.4.3 Stochastic parameter estimation problem with jump regularization

In contrast to the ODE parameter estimation described in chapter 2, in the setting of parameter estimation in a WIENER-driven SDE, we do not implement continuity conditions but we allow discontinuities at the shooting nodes t_k^{MS} , as outlined in the previous sections.

These “jumps” collect and correct for the deviations between the solutions of the SDE and the corresponding ODE (compare section 5.3). To give them a name, we formulate:

5.17 Definition (Stochastic jumps and jump regularization weights)

Suppose a multiple-shooting parametrization with shooting variables s_k at nodes $t_0^{\text{MS}}, \dots, t_{n_{\text{MS}}}^{\text{MS}}$ is given. We then call the discontinuities at the (inner) shootings nodes $t_1^{\text{MS}}, \dots, t_{n_{\text{MS}}-1}^{\text{MS}}$

$$\alpha_k := \alpha_k(p, s_{k-1}, s_k) := x(t_k^{\text{MS}}; t_{k-1}^{\text{MS}}, s_{k-1}, p) - s_k \in \mathbb{R}^{n_x} \quad (k = 1, \dots, n_{\text{MS}} - 1) \quad (5.28)$$

the *stochastic jumps*. By definition, α_k denotes the jump at shooting node t_k^{MS} .

By an additional index l , we denote the respective component:

$$\alpha_{k,l} = [x(t_k^{\text{MS}}; t_{k-1}^{\text{MS}}, s_{k-1}, p) - s_k]_l \quad (l = 1, \dots, n_x) \quad (5.29)$$

With every jump $\alpha_{k,l}$ we associate a *jump regularization weight* $\omega_{k,l}^2 \in \mathbb{R}_0^+$, or *jump weight* for short, building the weighting matrices

$$\begin{aligned} \Omega^{(k)} &:= \text{diag} \{ \omega_{k,1}, \dots, \omega_{k,n_x} \} \in \mathbb{R}^{n_x \times n_x} \quad (k = 1, \dots, n_{\text{MS}} - 1) \\ \Omega &:= \text{diag} \{ \Omega^{(1)}, \dots, \Omega^{(n_{\text{MS}}-1)} \} \in \mathbb{R}^{((n_{\text{MS}}-1)n_x) \times ((n_{\text{MS}}-1)n_x)}. \end{aligned}$$

Figure 5.3 gives an illustration of the jumps α_k ; also see figure 5.6. □

We may now formulate the constrained nonlinear stochastic parameter estimation problem with point constraints as follows:

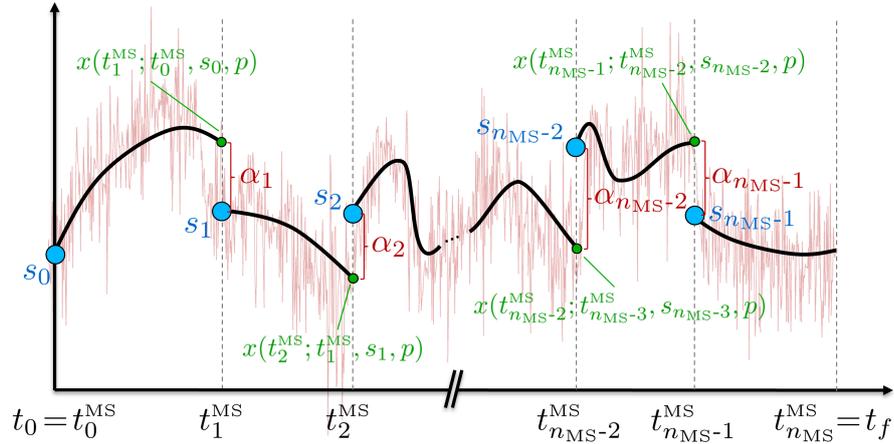


Figure 5.3: Illustration of stochastic jumps α_k at the inner shooting nodes t_k^{MS} ($k = 1, \dots, n_{\text{MS}} - 1$). The respective shooting variables s_k are given as blue dots; the left limits $x(t_k^{\text{MS}}; t_{k-1}^{\text{MS}}, s_{k-1}, p)$ of the càdlàg trajectory $x(t; s, p)$ are marked with small green dots. The sizes of the stochastic jumps α_k are highlighted by red braces. Also see figure 5.6.

5.18 Problem (Constrained nonlinear stochastic parameter estimation problem)

$$\begin{aligned} \min_{s_0, \dots, s_{n_{\text{MS}}-1}, p} \quad & \frac{1}{2} \sum_{i=1}^{n_M} \sigma_i^{-2} \cdot (h_i(t_i, x(t_i; s, p), p) - \eta_i)^2 + \frac{1}{2} \sum_{k=1}^{n_{\text{MS}}-1} \sum_{l=1}^{n_x} \omega_{k,l}^2 \cdot \alpha_{k,l}^2 \\ \text{s.t.} \quad & \dot{x}(t; t_k^{\text{MS}}, s_k, p) = f(t, x, p) \quad t \in [t_k^{\text{MS}}, t_{k+1}^{\text{MS}}) \\ & c^{ec}(x(t_1^{ec}; s, p), \dots, x(t_{n_{\text{EC}}}^{ec}; s, p), p) = 0 \\ & c^{ic}(x(t_1^{ic}; s, p), \dots, x(t_{n_{\text{IC}}}^{ic}; s, p), p) \geq 0 \end{aligned}$$

with $\alpha_k := x(t_k^{\text{MS}}; t_{k-1}^{\text{MS}}, s_{k-1}, p) - s_k$ as defined in definition 5.17.

Using the vector notation from section 5.4.1, the objective of the above parameter estimation problem 5.18 may be rewritten as

$$\min_{s, p} \quad \frac{1}{2} \|\Sigma^{-1}(h(s, p) - \eta)\|_2^2 + \frac{1}{2} \|\Omega \cdot \alpha\|_2^2.$$

Combining the individual parts of the objective into the *combined residual vector* F ,

$$F(s, p) := \begin{pmatrix} \Sigma^{-1}(h(s, p) - \eta) \\ \Omega \cdot \alpha \end{pmatrix} \in \mathbb{R}^{n_M + (n_{\text{MS}}-1)n_x} \quad (5.30)$$

and unifying the equality point constraints into $d^{ec}(s, p) := c^{ec}(x(t_1^{ec}; s, p), \dots, x(t_{n_{\text{EC}}}^{ec}; s, p), p)$, and the inequality constraints into $d^{ic}(s, p) := c^{ic}(x(t_1^{ic}; s, p), \dots, x(t_{n_{\text{IC}}}^{ic}; s, p), p)$, we end up in the problem

$$\begin{aligned} \min_{s, p} \quad & \frac{1}{2} \|F(s, p)\|_2^2 \\ \text{s.t.} \quad & d^{ec}(s, p) = 0 \\ & d^{ic}(s, p) \geq 0 \end{aligned} \quad (5.31)$$

from the viewpoint of a numerical solver.

The problem in eq. (5.31) is of the same type as the discretized constrained nonlinear least squares problem 2.10 in the ODE case and may thus be solved using the GAUSS-NEWTON method as described in chapter 1.

We shortly mention a **second interpretation** of the jump regularization term in problem 5.18. Driving the jump regularization weights $\omega_{k,l}^2$ to infinity corresponds to solving the ODE parameter estimation problem 2.8 with a penalty method, where the continuity conditions, eq. (2.10), are shifted into the objective (not discussing the arising numerical difficulties here).

High jump regularization weights may thus be used to generate trajectories with smaller jumps, which might be beneficial in some cases, e.g. as in the numerical examples for the FITZHUGH-NAGUMO oscillator, section 6.1.

5.4.4 Imperative jump regularization: A WIENER exponential example

The question whether it is necessary to include the jump regularization term in the objective instead of omitting them completely shall be shortly addressed here. On first sight, it might seem reasonable to view the interval solutions as independent problems, which are only coupled by the common parameters. If there is plenty of measurement data containing enough and the right information (e.g. many full-state measurements with small measurement error), this, indeed, may succeed. However, the following example shows that this approach might also fail even in the case of error-free full-state measurements, and that the inclusion of the jump regularization term in the objective leads to good estimates.

We investigate a WIENER-driven exponential, given as

$$dX_t = pX_t dt + D dW_t, \quad X_0 = 1.0, \quad p = 0.25, \quad D = 2.0, \quad t \in [0, 10], \quad (5.32)$$

whose deterministic counterpart has the solution $x(t) = X_0 \cdot e^{pt}$.

Figure 5.4 shows simulation results for the above system. On the left, in (a), the deterministic interpretation without diffusion is shown. On the right, in (b), a certain SDE solution together with the underlying realization of the driving WIENER process is depicted.

Notably is the fact that this realization of W_t drives the system into negative states, giving the impression of a negative initial value.

From this SDE solution, we take measurements at points $\{k+0.5, k+0.55 \mid k = 0, \dots, 9\}$, from which the parameter p and initial state X_0 shall be recovered. The time horizon is equidistantly split in 10 shooting intervals; all jump regularization weights ω_k^2 are set to 1.0. The values s_k at the shooting nodes and the parameter p , in this example 11 unknowns, are estimated by solving the parameter estimation problem 5.18 using the 20 state measurements.

Table 5.1 gives a description of the experimental set-up and shows the estimation results. The method we propose in this thesis recovers both, the parameter p and the initial state X_0 . Removing the jump regularization gives a much better fit in terms of residual reduction (a residual norm of approx. 0.31 compared to 10.2 for exact measurements) but leads to improper estimates.

The visualization in figure 5.5 sheds light on the cause. Without jump regularization, the coupling between the interval solution is only by the parameter p , i.e. the state variables at the shooting nodes are independent from the state values at the previous interval's end, and can be freely chosen to minimize the residual, leading to a trajectory with large jumps at the grid points. Moreover, by comparing figure 5.5c and figure 5.5d, and the respective results in table 5.1, we see that **without jump regularization, the results are not robust** to measurement noise: In both cases the residuals are small, but the estimates differ immensely.

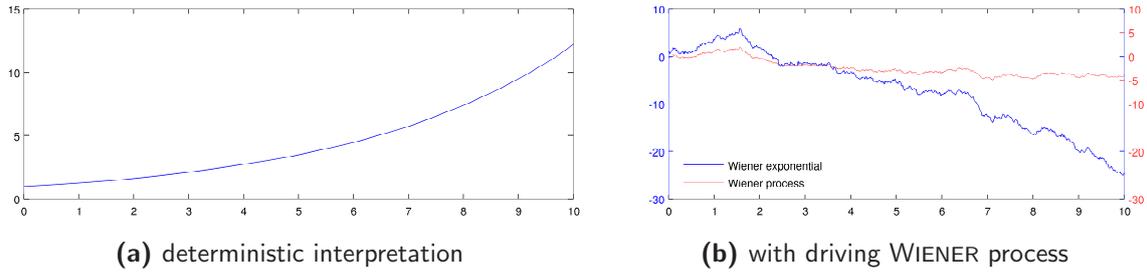


Figure 5.4: Trajectories of a WIENER exponential. Deterministic and stochastic interpretation of the exponential system described in eq. (5.32), with $p = 0.25$ and initial value $X_0 = 1.0$. The impact of the driving WIENER process, displayed as second graph in the right figure, is clearly visible in the WIENER exponential and manifests in jitter and even switching signs, changing it into a negativ exponential despite a positiv initial value.

Also, given for comparison, fitting the deterministic exponential $x(t) = X_0 \cdot e^{pt}$ is not convincing. The residual norm is high, the estimate for p is poor, and the estimate of the initial state X_0 has the wrong sign. Clearly, as this realization of the WIENER process drives the system into the negative halfspace, the initial value for a continuous exponential has to be negative, i.e. the initial state $X_0 = 1.0$ cannot be recovered.

As can be seen in figures 5.5a and 5.5b, the new method with jump regularization is able to deliver good estimates, even in situations where other methods fail, and further robustly approximates the trajectory for exact and noisy measurements in this WIENER exponential example. We refer to chapter 6 for further successful applications of the proposed method.

We show in appendix S.3.2, how this problem may be set up and solved using the software package `:sfit` developed in this thesis.

Table 5.1: Estimation results for the WIENER exponential with and without jump regularization for the system described by eq. (5.32). In the first column, exact state measurements, taken at time points $\{k+0.5, k+0.55 \mid k = 0, \dots, 9\}$ have been used for the estimation. In the second column, every datapoint was additively disturbed by a random value drawn from a normal distribution with zero mean and a variance of 0.25% of the respective measurement value. The shooting node variables at time points $\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$ were initialized by the temporally most proximate measurement data; the initial guess for the parameter was $p_0 = 0.5$. Measurement variances σ_i and jump regularization weights ω_k^2 were chosen as 1.0.

The letters (a)–(e) refer to the respective picture in figure 5.5; the values R and J refer to the 2-norm of the residual vector and of the weighted jumps, respectively. Values rounded. Also see appendix S.3.2.

$\mathbf{p}^* = 0.25$ $\mathbf{X}_0^* = 1.0$	true values	exact measurements		with measurement error			
		fig.	estimate	Res./Jmp.	fig.	estimate	Res./Jmp.
with jump regularization		(a)	$\mathbf{p} = 0.2896$ $\mathbf{X}_0 = 1.1841$	$R = 10.2$ $J = 10.5$	(b)	$\mathbf{p} = 0.2629$ $\mathbf{X}_0 = 1.0429$	$R = 20.4$ $J = 14.8$
without jump regularization		(c)	$\mathbf{p} = 0.5133$ $\mathbf{X}_0 = 0.7317$	$R = 0.31$ $J = 68.7$	(d)	$\mathbf{p} = -1.3901$ $\mathbf{X}_0 = 1.5294$	$R = 2.85$ $J = 1651.7$
continuous trajectory		(e)	$\mathbf{p} = 0.3461$ $\mathbf{X}_0 = -0.8410$	$R = 61.9$ $J \approx 10^{-10}$	(f)	$\mathbf{p} = 0.3371$ $\mathbf{X}_0 = -0.9192$	$R = 82.6$ $J \approx 10^{-10}$

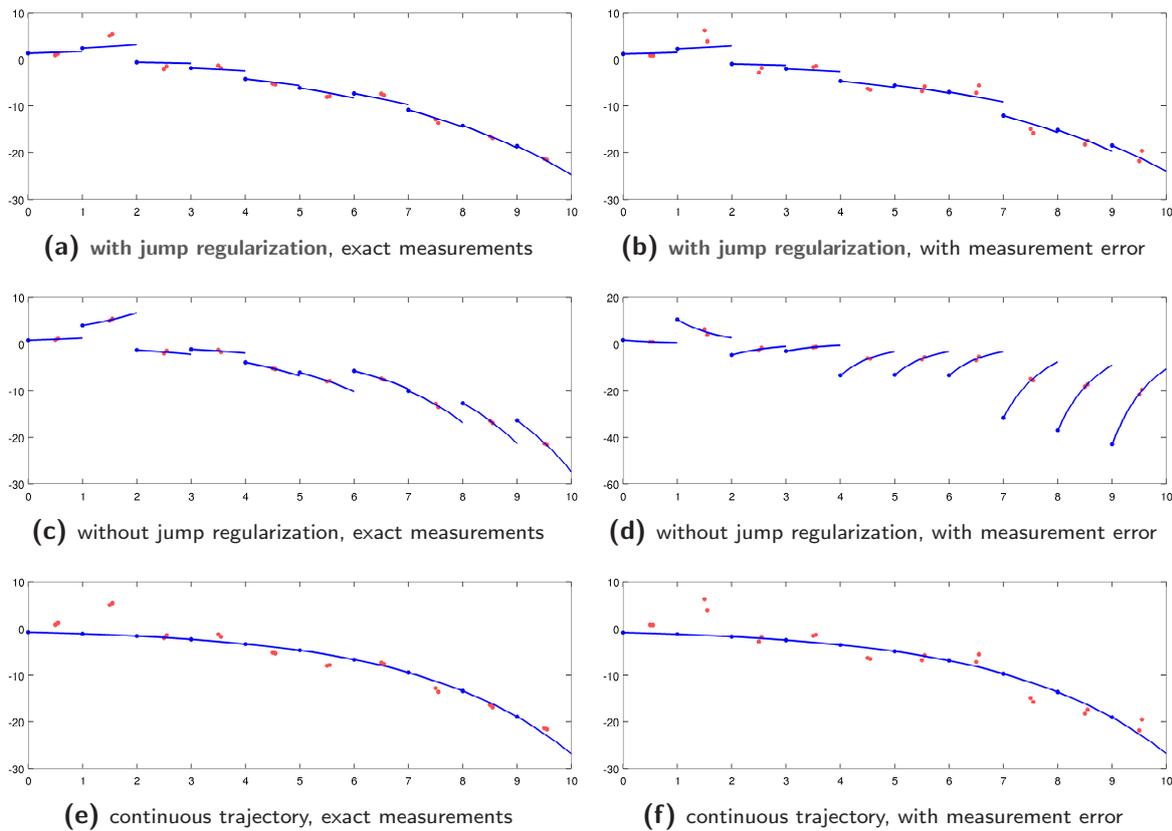


Figure 5.5: Fitted exponential trajectory with and without jump regularization for the cases of the WIENER exponential parameter estimation problem as described in table 5.1. In (a) and (b), the fit originating from the new method with jump regularization is depicted. The discontinuities at the shooting nodes are clearly visible. Though the residual is much smaller in the settings (c) and (d) without jump regularization, the resulting estimates are worthless. The continuous trajectory in (e) and (f), apparently giving a good fit, also delivers unsatisfactory results. See table 5.1 for details.

The new method with jump regularization, shown in (a) and (b), successfully recovers both the kinetic parameter and the initial state value.

Fitted trajectory as blue line (—), shooting nodes marked with a blue dot (\bullet), measurements marked with a red dot (\bullet).

In appendix S.3.2 we show how to set up and solve this problem using the software package `:sfit`.

5.4.5 Choosing the jump regularization weights

The inclusion of jumps as regularization term in the objective as well as the choice of their respective weights strongly depends on the concrete problem. Main factors are:

1. the dynamics of the system
2. the strength of diffusion
3. the realization of the driving process
4. the number of measurements in the respective interval
5. the variance of the measurement error
6. the length of the shooting intervals.

Choosing the jump regularization weights currently requires knowledge about and experience with the examined system.

Furthermore, one can adjust the regularization according to the purpose: higher weights lead to a “less discontinuous” fitted trajectory, i.e. with smaller jumps at the shooting nodes (cf. figure 6.5 on pages 172 and 173).

Investigations on example problems show that the estimation results are quite robust and not very sensitive to the choice of weights over at least one order of magnitude (see chapter 6: Numerical Examples, page 163ff).

5.4.6 A convergence result: Jumps approaching zero

The jump regularization in the parameter estimation problem 5.18,

$$\frac{1}{2} \sum_{k=1}^{n_{\text{MS}}-1} \sum_{l=1}^{n_x} \omega_{k,l}^2 \cdot \alpha_{k,l}^2 \quad \text{with} \quad \alpha_k := x(t_k^{\text{MS}}; t_{k-1}^{\text{MS}}, s_{k-1}, p) - s_k$$

raises some questions about what happens asymptotically if the number of shooting intervals n_{MS} tends to infinity: Does the sum (series) converge? Is $\sum_{l=1}^{n_x} \alpha_{k,l}^2$ a null sequence? This last question is of particular importance, as a positive answer implies asymptotically vanishing stochastic jumps.

Piecewise solutions

For the analysis, as in section 5.3, we assume w.l.o.g. that the SDE/ODE system is time-homogeneous/autonomous, and the time domain is $[0, T]$.

We split the time domain in n_{MS} intervals of length $h = T/n_{\text{MS}}$. Similar to the definitions in eqs. (5.12) and (5.13), we introduce *interval solutions* $X_t^{(k)}$ of S-IVP and $x^{(k)}(t)$ of D-IVP on the intervals $[t_k^{\text{MS}}, t_{k+1}^{\text{MS}}]$, defined by the **shooting grid**

$$\mathbb{T}^{\text{MS}} = \{t_0^{\text{MS}}, t_1^{\text{MS}}, \dots, t_{n_{\text{MS}}}^{\text{MS}}\} \quad \text{with} \quad t_k^{\text{MS}} = kh \quad (k = 0, \dots, n_{\text{MS}}) \quad \text{and} \quad h = \frac{T}{n_{\text{MS}}}.$$

For time-homogeneous SDE and autonomous ODE, we can change the integration limits of the interval solutions such that each integration interval $[t_k^{\text{MS}}, t_{k+1}^{\text{MS}})$ is shifted to $[0, h)$. We can do this also for integrals w.r.t. the WIENER process, since the WIENER increments are independent.

Further, we do not write the dependence on the parameter vector p in this analysis.

5.19 Definition (Interval solutions $X_t^{(k)}$ and $x^{(k)}(t)$ of S-IVP and D-IVP)

Assuming w.l.o.g. that the underlying SDE (ODE) is time-homogeneous (autonomous), let the time domain $[t_0, t_f]$ of the original parameter estimation problem be shifted to $[0, T]$ with $T = t_f - t_0$, and split in n_{MS} intervals of length $h = T/n_{\text{MS}}$.

We denote by $X_t^{(k)}$ the solution of the S-IVP on the k -th interval with initial state $X_{0,k}^{(s)}$,

$$X_t^{(k)} = X_{0,k}^{(s)} + \int_0^t f(X_s^{(k)}) ds + \int_0^t D dW_s \quad (t \in [0, h))$$

Analogously, we write $x^{(k)}(t)$ for the deterministic solution of the D-IVP on the k -th interval with initial state $X_{0,k}^{(d)}$:

$$x^{(k)}(t) = X_{0,k}^{(d)} + \int_0^t f(x^{(k)}(s)) ds \quad (t \in [0, h])$$

Lastly, we write analogously to definition 5.1

$$\Delta_t^{(k)} := X_t^{(k)} - x^{(k)}(t)$$

for the difference process on the k -th interval. \square

The SDE solution is continuous (“has a continuous version”), i.e.

$$X_{0,k}^{(s)} = X_h^{(k-1)} \quad (k = 1, \dots, n_{\text{MS}} - 1) \quad (5.33)$$

and the SDE solution over the whole time domain $[0, T]$ is thus the (continuous) concatenation of the above interval solutions, i.e. we write the complete solution process X_t as

$$X_t := X_{t-t_k^{\text{MS}}}^{(k)} \quad \text{for } t \in [t_k^{\text{MS}}, t_{k+1}^{\text{MS}})$$

and, analogously, we write

$$x(t) := x^{(k)}(t - t_k^{\text{MS}}) \quad \text{for } t \in [t_k^{\text{MS}}, t_{k+1}^{\text{MS}})$$

for the piecewise continuous deterministic approximation $x(t)$. Formally, we define the values at the right interval boundary $t_{n_{\text{MS}}}^{\text{MS}}$ as $X_{t_{n_{\text{MS}}}^{\text{MS}}} := X_h^{(n_{\text{MS}}-1)}$ and $x(t_{n_{\text{MS}}}^{\text{MS}}) := x^{(n_{\text{MS}}-1)}(h)$.

Figure 5.6 illustrates the above definitions.

Theorem 5.15 gives the result

$$\mathbb{E} \left[|\Delta_t^{(k)}|^2 \right] \rightarrow 0 \quad (t \rightarrow 0) \quad \text{if } X_0^{(s)} = X_0^{(d)}, \quad (5.34)$$

i.e. if the initial values of S-IVP and D-IVP coincide. In this case, we can identify the jumps $\mathcal{A}_k = x^{(k-1)}(t_k^{\text{MS}}; t_{k-1}^{\text{MS}}, X_{0,k-1}^{(s)}) - X_{0,k}^{(s)}$ with evaluations of the difference process $\Delta_t^{(k-1)}$ at the end of the interval $[0, h)$, i.e.

$$|\Delta_h^{(k-1)}| = |X_h^{(k-1)} - x^{(k-1)}(h)| = |X_{0,k}^{(s)} - x^{(k-1)}(h)| = |\mathcal{A}_k|, \quad (5.35)$$

as $X_h^{(k-1)} = X_{0,k}^{(s)}$ due to the continuity of the process X_t , eq. (5.33).

To analyse the properties of the parameter estimation method proposed in this thesis, and the behaviour of the sum/series $\sum_{k=1}^{n_{\text{MS}}-1} \sum_{l=1}^{n_x} \mathcal{A}_{k,l}^2$ for $n_{\text{MS}} \rightarrow \infty$, we have to understand \mathcal{A}_k now as random variables⁽⁶⁾.

⁽⁶⁾ when analysing the *method*, we have to consider the whole set of SDE trajectories, not a certain realization.

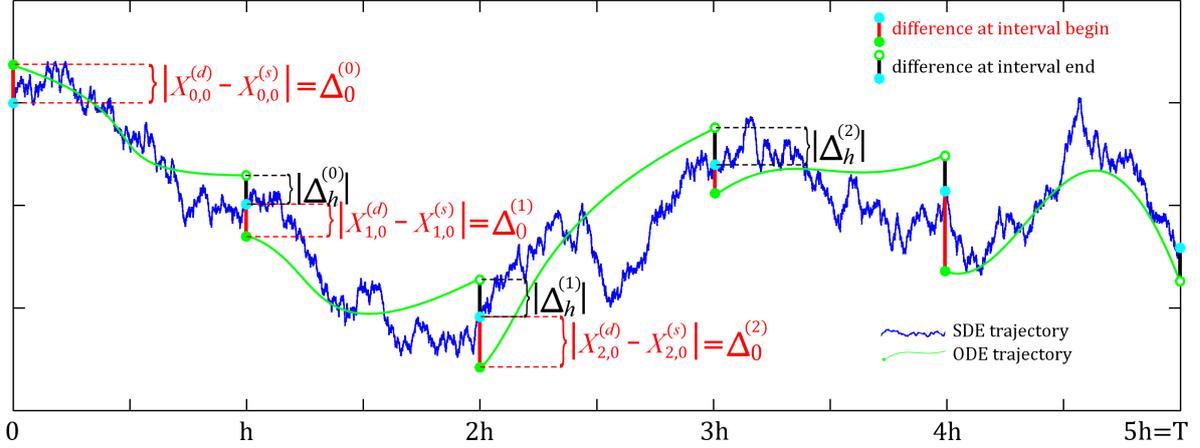


Figure 5.6: Illustration of interval solutions (illustrative image of quantities).

The continuous S-IVP solution trajectory is shown as blue solid line (—), the piecewise D-IVP solutions as green solid lines (—). Initial values $X_{0,k}^{(s)}$ of the interval S-IVPs are marked with a filled cyan dot (\bullet), and the respective D-IVP initial states $X_{0,k}^{(d)}$ by a filled green dot (\bullet). D-IVP states at the interval ends are marked by an empty green dot (\circ). Explicitly depicted are the first three differences at interval start $\Delta_0^{(k)} = |X_{0,k}^{(d)} - X_{0,k}^{(s)}|$ and at interval end $\Delta_h^{(k)}$.

For a single interval $[0, T]$, i.e. $n_{\text{MS}} = 1$, theorem 5.15 gives

$$\mathbb{E} \left[|\Delta_T|^2 \right] \leq (4|X_0^{(s)} - X_0^{(d)}| + 4\|D\|^2 n_W T) \cdot e^{4T^2 L^2}$$

where L is the drift function's LIPSCHITZ constant, D denotes the SDE's diffusion matrix, and n_W is the dimension of the driving WIENER process. Dividing the interval $[0, T]$ in n_{MS} intervals, we get

$$\begin{aligned} \mathbb{E} \left[\sum_{k=1}^{n_{\text{MS}}} |\Delta_{T/n_{\text{MS}}}^{(k)}|^2 \right] &\leq \sum_{k=1}^{n_{\text{MS}}} (4|X_{0,k}^{(s)} - X_{0,k}^{(d)}|^2 + 4\|D\|^2 n_W \frac{T}{n_{\text{MS}}}) \cdot e^{4(T/n_{\text{MS}})^2 L^2} \\ &= \underbrace{\left(4 \sum_{k=1}^{n_{\text{MS}}} |\Delta_0^{(k)}|^2 + 4n_{\text{MS}} \|D\|^2 n_W \frac{T}{n_{\text{MS}}} \right)}_{=0 \text{ by eq. (5.34)}} \cdot e^{4(T/n_{\text{MS}})^2 L^2} \\ &= 4\|D\|^2 n_W T \cdot e^{4(T/n_{\text{MS}})^2 L^2} \longrightarrow 4\|D\|^2 n_W T \quad (n_{\text{MS}} \rightarrow \infty) \end{aligned}$$

i.e. the summed up squared differences at the interval borders have a finite expectation, for which we can give a bound that is solely determined by the intensity of the diffusion and the size of the full time domain.

By eq. (5.35), we can identify $|\alpha_k| = |\Delta_h^{(k-1)}|$ for $k = 1, \dots, n_{\text{MS}} - 1$, and it immediately follows

$$\mathbb{E} \left[\sum_{k=1}^{n_{\text{MS}}-1} \sum_{l=1}^{n_x} |\alpha_{k,l}|^2 \right] \leq 4\|D\|^2 n_W T \cdot e^{4(T/n_{\text{MS}})^2 L^2} \longrightarrow 4\|D\|^2 n_W T \quad (n_{\text{MS}} \rightarrow \infty)$$

i.e. the components $\alpha_{k,l}$ of the stochastic jumps are null sequences.

5.5 Numerical analysis of the proposed parameter estimation method

5.5.1 Sparsity patterns in the linear subproblems

When solving the stochastic parameter estimation problem 5.18 with the GAUSS-NEWTON method, there arises a special structure in the Jacobian of the linearized systems due to the special form of the objective. This sparsity pattern could (and should) be exploited in the determination of the increment Δx_k in the linearized subproblems, section 1.3.2.

The structure of the Jacobian of the constraints strongly depends on the concrete problem and may vary from dense (for highly interconnected constraints across many shooting variables) to very sparse (e.g. in the case of simple bound constraints or only local constraints within the shooting intervals).

We focus in section on the frequent case of unconstrained problems (or problems with simple bound constraints, as the affected variables can be easily eliminated from the system in advance), and investigate the structure that occurs during computation of the increment Δx_k as solution of the linearized systems (see section 1.3.2). All these findings also hold in the constrained case for the upper part of the system matrix in the linearized subproblems, belonging to the combined residual vector of problem 5.18.

In the following, we analyse the structure of the *combined residual vector* $F(s, p)$ defined in eq. (5.30), which consists of the weighted residuals $F_1 = F_1(s, p) := \Sigma^{-1}(h(s, p) - \eta)$ and the weighted stochastic jumps $F_2 = F_2(s, p) := \Omega \cdot \alpha$ that we both split up into the parts belonging to the individual shooting intervals:

$$F = \begin{pmatrix} F_1 \\ F_2 \end{pmatrix}, \quad F_1 := \begin{pmatrix} F_1^{(0)} \\ \vdots \\ F_1^{(n_{\text{MS}}-1)} \end{pmatrix} \in \mathbb{R}^{n_M}, \quad F_2 := \begin{pmatrix} F_2^{(1)} \\ \vdots \\ F_2^{(n_{\text{MS}}-1)} \end{pmatrix} \in \mathbb{R}^{(n_{\text{MS}}-1)n_x} \quad (5.36)$$

where

$$\begin{aligned} F_1^{(k)} &:= (\Sigma^{(k)})^{-1} \cdot (h^{(k)}(s_k, p) - \eta^{(k)}) \in \mathbb{R}^{|\iota(k)|} \quad (k = 0, \dots, n_{\text{MS}} - 1) \\ F_2^{(k)} &:= \Omega^{(k)} \cdot (x(t_k^{\text{MS}}; t_{k-1}^{\text{MS}}, s_{k-1}, p) - s_k) \in \mathbb{R}^{n_x} \quad (k = 1, \dots, n_{\text{MS}} - 1) \end{aligned} \quad (5.37)$$

denote the contributions from the respective shooting intervals.

The Jacobian $J = J(s, p)$ thus has the form

$$J(s, p) = \frac{dF(s, p)}{d(s, p)} = \begin{pmatrix} \frac{dF_1}{ds_0} & \dots & \frac{dF_1}{ds_{n_{\text{MS}}-1}} & \frac{dF_1}{dp} \\ \frac{dF_2}{ds_0} & \dots & \frac{dF_2}{ds_{n_{\text{MS}}-1}} & \frac{dF_2}{dp} \\ \vdots & & \vdots & \vdots \\ \frac{dF_1^{(n_{\text{MS}}-1)}}{ds_0} & \dots & \frac{dF_1^{(n_{\text{MS}}-1)}}{ds_{n_{\text{MS}}-1}} & \frac{dF_1^{(n_{\text{MS}}-1)}}{dp} \\ \frac{dF_2^{(1)}}{ds_0} & \dots & \frac{dF_2^{(1)}}{ds_{n_{\text{MS}}-1}} & \frac{dF_2^{(1)}}{dp} \\ \vdots & & \vdots & \vdots \\ \frac{dF_2^{(n_{\text{MS}}-1)}}{ds_0} & \dots & \frac{dF_2^{(n_{\text{MS}}-1)}}{ds_{n_{\text{MS}}-1}} & \frac{dF_2^{(n_{\text{MS}}-1)}}{dp} \end{pmatrix} = \begin{pmatrix} \frac{dF_1^{(0)}}{ds_0} & \dots & \frac{dF_1^{(0)}}{ds_{n_{\text{MS}}-1}} & \frac{dF_1^{(0)}}{dp} \\ \frac{dF_1^{(1)}}{ds_0} & \dots & \frac{dF_1^{(1)}}{ds_{n_{\text{MS}}-1}} & \frac{dF_1^{(1)}}{dp} \\ \vdots & & \vdots & \vdots \\ \frac{dF_1^{(n_{\text{MS}}-1)}}{ds_0} & \dots & \frac{dF_1^{(n_{\text{MS}}-1)}}{ds_{n_{\text{MS}}-1}} & \frac{dF_1^{(n_{\text{MS}}-1)}}{dp} \\ \frac{dF_2^{(1)}}{ds_0} & \dots & \frac{dF_2^{(1)}}{ds_{n_{\text{MS}}-1}} & \frac{dF_2^{(1)}}{dp} \\ \vdots & & \vdots & \vdots \\ \frac{dF_2^{(n_{\text{MS}}-1)}}{ds_0} & \dots & \frac{dF_2^{(n_{\text{MS}}-1)}}{ds_{n_{\text{MS}}-1}} & \frac{dF_2^{(n_{\text{MS}}-1)}}{dp} \end{pmatrix} \quad (5.38)$$

Using the definitions of $F_1^{(k)}$ and $F_2^{(k)}$, and writing $\mathbb{O}_{m \times n}$ for the $m \times n$ zero matrix, we get for the individual derivatives of the residual functions:

$$\frac{dF_1^{(k)}}{ds_j} = \begin{cases} (\Sigma^{(k)})^{-1} \cdot \frac{d}{ds_k} h^{(k)}(s_k, p) & \text{for } j = k \\ \mathbb{O}_{|\iota(k)| \times n_x} & \text{for } j \neq k \end{cases} \quad (k = 0, \dots, n_{\text{MS}} - 1)$$

$$\frac{dF_1^{(k)}}{dp} = (\Sigma^{(k)})^{-1} \cdot \frac{d}{dp} h^{(k)}(s_k, p) \quad (k = 0, \dots, n_{\text{MS}} - 1)$$

and for the jumps:

$$\frac{dF_2^{(k)}}{ds_j} = \begin{cases} -\Omega^{(k)} & \text{for } j = k \\ \Omega^{(k)} \cdot \frac{d}{ds_{k-1}} x(t_k^{\text{MS}}; t_{k-1}^{\text{MS}}, s_{k-1}, p) & \text{for } j = k - 1 \\ \mathbb{O}_{n_x \times n_x} & \text{otherwise} \end{cases} \quad (k = 1, \dots, n_{\text{MS}} - 1)$$

$$\frac{dF_2^{(k)}}{dp} = \Omega^{(k)} \cdot \frac{d}{dp} x(t_k^{\text{MS}}; t_{k-1}^{\text{MS}}, s_{k-1}, p) \quad (k = 1, \dots, n_{\text{MS}} - 1)$$

i.e. the Jacobian J is indeed very sparse:

$$J = \begin{array}{|c|c|c|c|c|} \hline \frac{dF_1^{(0)}}{ds_0} & & & & \frac{dF_1^{(0)}}{dp} \\ \hline & \frac{dF_1^{(1)}}{ds_1} & & & \frac{dF_1^{(1)}}{dp} \\ & & & & \vdots \\ & & & \frac{dF_1^{(n_{\text{MS}}-1)}}{ds_{n_{\text{MS}}-1}} & \frac{dF_1^{(n_{\text{MS}}-1)}}{dp} \\ \hline \frac{dF_2^{(1)}}{ds_0} & -\Omega^{(1)} & & & \frac{dF_2^{(1)}}{dp} \\ \hline & \frac{dF_2^{(2)}}{ds_1} & -\Omega^{(2)} & & \frac{dF_2^{(2)}}{dp} \\ & & & & \vdots \\ & & & \frac{dF_2^{(n_{\text{MS}}-1)}}{ds_{n_{\text{MS}}-2}} & -\Omega^{(n_{\text{MS}}-1)} & \frac{dF_2^{(n_{\text{MS}}-1)}}{dp} \\ \hline \end{array} \quad (5.39)$$

We explicitly point to the fact that the jump regularization weights matrices $\Omega^{(k)}$ are also very sparse, since they are diagonal matrices. Figure 5.7 depicts the sparsity pattern from two estimation problems discussed in the numerical examples section 6.2.

More structure from problem-dependent properties and lifting

More structure may be induced depending on the specific problem: Measurement functions depending on only a subset of all states create zeros in the respective parts of $\frac{d}{ds_k} F_1^{(k)}$. Local parameters, i.e. parameters whose influence is restricted to certain intervals, generate zeros in both $\frac{d}{dp} F_1^{(k)}$ and $\frac{d}{dp} F_2^{(k)}$ of the remaining intervals. A thereby inspired **lifting reformulation** leading to a complete interval-wise decoupling of the objective will be addressed in section 5.5.5.

We remark that the pattern displayed in eq. (5.39) and figure 5.7 is thus a “worst case” setting in terms of density of the Jacobian.

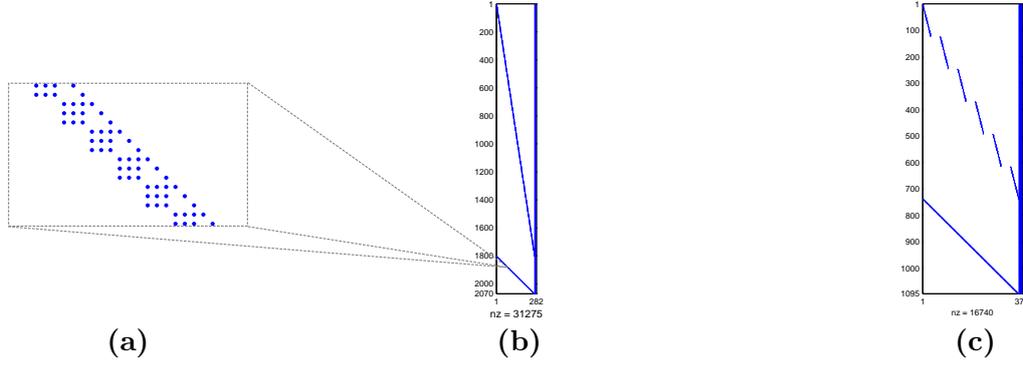


Figure 5.7: Sparsity patterns of the combined residual vector's Jacobian in the calcium oscillator examples from section 6.2. The full Jacobian $J(s, p)$ always has the size $(n_M + (n_{MS} - 1)n_x) \times n_{MS}n_x + n_p$. In (b), full state measurements are taken in equidistant intervals, thus $\frac{dF_1}{ds}$ has block-diagonal structure. With $n_x = 3$, $n_p = 12$, $n_{MS} = 90$, $n_M = 1803$, the combined residual vector's Jacobian has $1803 + 89 \cdot 3 = 2070$ rows and $3 \cdot 90 + 12 = 282$ columns. Only 31275 of 583740 elements are nonzero, i.e. less than 5.4%. Note that due to the special structure of measurements, not the maximum number of $31317 = 1803 \cdot (3 + 12) + 89 \cdot (3 \cdot 3 + 3 + 3 \cdot 12)$ is observed. See section 6.2.3.1 for a description of the measurement scheme.

In (a), a detail enlargement of the sparsity pattern belonging to the stochastic jumps of (b) is depicted. The 3×3 blocks originate from $\frac{d}{ds_{k-1}} F_2^{(k)}$, whereas the diagonal matrices on the right of them are $\frac{d}{ds_k} F_2^{(k)} = -\Omega^{(k)}$.

In (c), the full state observations have been done intermittently, i.e. there are shooting intervals with no measurements, resulting in the multiple block-diagonal structure in the upper left part from $\frac{dF_1}{ds}$. Here, $n_x = 3$, $n_p = 12$, $n_{MS} = 120$, $n_M = 738$. See section 6.2.3.8 for a description of the measurement scheme.

Sparsity pattern for constrained problems

In the setting of constrained problems (with the exception of simple bounds on the variables, which can be easily eliminated from the linearized system, and constraints restricted on single intervals), an additional horizontal band is added below the Jacobian in eq. (5.39), such that the system matrix \bar{J} in the linearized problems has the form

$$\bar{J} = \begin{pmatrix} J \\ \boxed{} \end{pmatrix}$$

with J as in eq. (5.39). The sparsity pattern of the block below J depends on the actual constraints and may vary from very sparse (if only the constraints for the local parameters are present) to dense.

5.5.2 Derivative generation

The upper part of the Jacobian $J(s, p)$, i.e. $\frac{d}{d(s,p)} F_1$, requires the calculation of the measurement function derivatives:

$$\frac{dh^{(k)}}{ds_k} = \frac{\partial h^{(k)}}{\partial x} \cdot \frac{\partial x}{\partial s_k} \quad \text{and} \quad \frac{dh^{(k)}}{dp} = \frac{\partial h^{(k)}}{\partial x} \cdot \frac{\partial x}{\partial p} + \frac{\partial h^{(k)}}{\partial p}.$$

Depending on the implementation and software framework (deprecating finite difference methods here), one might (a) get the sensitivities $\frac{\partial x}{\partial s_k}$ and $\frac{\partial x}{\partial p}$ via the variational differential equa-

tions, eq. (2.14), and determine the partial derivatives of the measurement function, $\frac{\partial h^{(k)}}{\partial x}$ and $\frac{\partial h^{(k)}}{\partial p}$ analytically or via AD, or (b) retrieve the total derivatives $\frac{dh^{(k)}}{ds_k}$ and $\frac{dh^{(k)}}{dp}$ via an AD-capable integrator while computing the simulated measurements. In the software package `sfit` implemented in this thesis, the latter approach is used.

5.5.3 Linear complexity growth: An upper bound on the number of nonzero elements in the combined residual vector's Jacobian $J(s, p)$

The maximum number of nonzero elements in the Jacobian $J(s, p)$ of the combined residual vector, eq. (5.39), that may be achieved if all measurements have non-vanishing sensitivities and no local parameters are present, may be calculated from the number of states n_x , the number of shooting intervals n_{MS} , the number of parameters n_p , and the number of measurements n_M .

Given a certain system, the number of state variables n_x and parameters n_p stay fixed, and the question arises, how the complexity of the linear subproblems develops in terms of measurements n_M and especially in terms of an increasing number of shooting intervals n_{MS} , which might be needed to ensure a sufficient quality of approximation of the stochastic process's realization.

The total size of the Jacobian $J(s, p) \in \mathbb{R}^{(n_M + n_x(n_{MS} - 1)) \times (n_x n_{MS} + n_p)}$ is the sum of its four quadrant derivatives, which have the dimensions

$$\begin{aligned} \frac{dF_1}{ds} &\in \mathbb{R}^{n_M \times n_x n_{MS}} & \frac{dF_1}{dp} &\in \mathbb{R}^{n_M \times n_p} \\ \frac{dF_2}{ds} &\in \mathbb{R}^{(n_{MS} - 1)n_x \times n_x n_{MS}} & \frac{dF_2}{dp} &\in \mathbb{R}^{(n_{MS} - 1)n_x \times n_p} \end{aligned}$$

such that the total number of elements of $J(s, p)$, denoted as $n_{\text{total}}(J)$, is

$$\begin{aligned} n_{\text{total}}(J) &= (n_M + (n_{MS} - 1)n_x) \cdot (n_{MS}n_x + n_p) \\ &= n_{MS}^2 n_x^2 + n_{MS}(n_x n_p - n_x^2) + n_{MS} n_M n_x + n_M n_p - n_x n_p \\ &= \mathcal{O}(n_{MS}^2) + \mathcal{O}(n_{MS} n_M) + \mathcal{O}(n_M) \end{aligned}$$

that is, the total number of elements in the Jacobian is linear in the number of measurements n_M but grows *quadratically* in the number of shooting intervals n_{MS} .

Number of nonzero elements in $J(s, p)$ grows only linearly

Due to the special structure of the Jacobian, depicted in eq. (5.39), we have for the (maximum) number of nonzero elements nnz_{max} of the respective parts of $J(s, p)$:

$$\begin{aligned} \text{nnz}_{\text{max}}\left(\frac{dF_1}{ds}\right) &= n_M n_x & \text{nnz}_{\text{max}}\left(\frac{dF_1}{dp}\right) &= n_M n_p \\ \text{nnz}_{\text{max}}\left(\frac{dF_2}{ds}\right) &= (n_{MS} - 1)n_x + n_x = n_{MS} n_x & \text{nnz}_{\text{max}}\left(\frac{dF_2}{dp}\right) &= (n_{MS} - 1)n_x n_p \end{aligned}$$

so we get for the total (maximum) number of nonzero elements of the Jacobian:

$$\begin{aligned} \text{nnz}_{\text{max}}(J(s, p)) &= n_M n_x + n_x n_{MS} + n_M n_p + n_x n_{MS} n_p - n_x n_p \\ &= n_M(n_x + n_p) + n_{MS} n_x(1 + n_p) - n_x n_p = \mathcal{O}(n_{MS}) + \mathcal{O}(n_M), \end{aligned}$$

that is, the number of nonzero elements in $J(s, p)$ grows only *linearly* in the number of shooting intervals n_{MS} .

5.5.4 Maintaining sparsity in QR decomposition of the Jacobian $J(s, p)$

In the following, we stick to the assumption of absence of additional point constraints, as they might change the sparsity properties of the linear problems. Simple bound constraints on the optimization variables do not interfere with the analysis here. We will see that, using a *sparse and Q -less QR decomposition*, the sparsity property of J will be maintained and transferred to the decomposition factor R (definition 5.20 and theorem 5.21).

When solving the (unconstrained) stochastic parameter estimation problem 5.18 with the GAUSS-NEWTON method of section 1.3, linear subproblems

$$\min_{\Delta x} \|J\Delta x + F\| \quad (5.40)$$

have to be solved, where the *combined residual vector* $F = F(s, p) = (F_1(s, p)^T, F_2(s, p)^T)^T$ (see eq. (5.30)) consists of the weighted measurement residuals and the weighted stochastic jumps at the current iterate $x = (s, p)$, and $J = J(s, p) = \frac{dF}{d(s, p)}$ is the respective Jacobian.

Having a (*reduced or economy size*) QR decomposition (with pivoting) of $J \in \mathbb{R}^{m \times n}$, $m > n$, i.e. $J = QR$ with orthogonal matrix $Q \in \mathbb{R}^{m \times n}$ and upper triangular matrix $R \in \mathbb{R}^{n \times n}$, the increment Δx may be numerically stably calculated as

$$\Delta x = -R^{-1}Q^TF. \quad (5.41)$$

Rank-deficiency of J may be detected and handled by using a rank-revealing QR decomposition as presented in section 1.3. We note that this problem is mathematically equivalent to solving the (numerically unstable) normal equations $J^TJ\Delta x = -J^TF$.

The matrix Q is usually dense, but its explicit computation is avoidable as Q^TF may be efficiently computed while decomposing J .

Due to the special structure of the Jacobian J , eq. (5.39), the sparsity pattern of R is independent of the number of measurements n_M and only determined by the number of state variables n_x , the number of parameters n_p , and the number of shooting intervals n_{MS} .

5.20 Definition (Decomposition factor R)

We call the matrix $R = Q^TJ \in \mathbb{R}^{(n_{MS}n_x + n_p) \times (n_{MS}n_x + n_p)}$ in eq. (5.41), stably computed by a QR decomposition, the *decomposition factor* $R^{(7)}$. \square

5.21 Theorem (A sharp upper bound on the number of nonzero elements in R)

By using a HOUSEHOLDER QR decomposition of J , we ensure that the decomposition factor R always has *at most*

$$\begin{aligned} \text{nnz}_{\max}(R) &= (n_{MS} - 1)n_x^2 + \frac{1}{2}n_x(n_x + 1)n_{MS} + n_{MS}n_xn_p + \frac{1}{2}n_p(n_p + 1) \\ &< n_{MS}n_x^2 + n_{MS}n_xn_p + n_p^2 \end{aligned} \quad (5.42)$$

nonzero elements, while in total R has

$$n_{\text{total}}(R) = (n_{MS}n_x + n_p)^2 = n_{MS}^2n_x^2 + 2n_{MS}n_xn_p + n_p^2$$

elements, i.e. the occupancy is as low as $\mathcal{O}(\frac{1}{n_{MS}})$.

⁽⁷⁾ Since $J^TJ = (QR)^T(QR) = R^TR$, the matrix R is the CHOLESKY factor of J^TJ , that one would compute when solving the normal equations. In this thesis, the normal equations are nowhere used or solved.

We remind that the matrices R_k are non-square in general, though the above picture might give this impression. The row permutation is not necessary, as for any permutation matrix P we have

$$J = QR = P^T PQR \iff PJ = PQR \iff \tilde{J} = \tilde{Q}R$$

with permuted system matrix $\tilde{J} := PJ$, orthogonal $\tilde{Q} := PQ$ and same decomposition factor R , but the application of the row permutation described in eq. (5.45) allows an easier formulation of the HOUSEHOLDER matrices that will be applied in the QR decomposition for finding the decomposition factor R .

The first column j_1 of the row-permuted Jacobian \tilde{J} has the sparsity pattern

$$j_{1,l} = \begin{cases} * & \text{for } l = 1, \dots, n_x + |\iota(0)| \\ 0 & \text{for } l \geq n_x + |\iota(0)| + 1 \end{cases} \quad (5.46)$$

and its HOUSEHOLDER vector $v_1 := j_1 - \|j_1\| e_1$ has the same.

Thus, writing $n_J := (n_{\text{MS}} - 1)n_x + n_M$ for the number of rows of J , we see that the first HOUSEHOLDER matrix that will be applied to \tilde{J} has the form

$$H_1 = \mathbb{I}_{n_J \times n_J} - 2 \frac{v_1 v_1^T}{v_1^T v_1} = \begin{bmatrix} \tilde{H}_1 & \mathbb{O} \\ \mathbb{O} & \mathbb{I} \end{bmatrix} \in \mathbb{R}^{n_J \times n_J}$$

with a generally dense matrix $\tilde{H}_1 \in \mathbb{R}^{n_{H_1} \times n_{H_1}}$ and $n_{H_1} = n_x + |\iota(0)|$. Here, \mathbb{O} and \mathbb{I} denote zero and identity matrices of suitable dimensions.

Due to that structure, the application of H_1 to \tilde{J} leads to a fill-in in the first $|\iota(0)|$ rows in the columns $n_x + 1, \dots, 2n_x$ of $H_1 \tilde{J}$, i.e. in the block right above D_1 in eq. (5.45).

In general, the l -th HOUSEHOLDER matrix for $l = 1, \dots, (n_{\text{MS}} - 1)n_x$, embedded in the space $\mathbb{R}^{n_J \times n_J}$, has the form:

$$H_l = \begin{bmatrix} \mathbb{I}_1^{(l)} & & \\ & \tilde{H}_l & \\ & & \mathbb{I}_2^{(l)} \end{bmatrix} \in \mathbb{R}^{n_J \times n_J} \quad \text{with dimensions} \quad \begin{array}{l} \mathbb{I}_1^{(l)} \in \mathbb{R}^{n_1^{(l)} \times n_1^{(l)}} \\ \tilde{H}_l \in \mathbb{R}^{n_{H_l} \times n_{H_l}} \\ \mathbb{I}_2^{(l)} \in \mathbb{R}^{n_2^{(l)} \times n_2^{(l)}} \end{array}$$

and

$$n_1^{(l)} = l - 1, \quad n_{H_l} = n_x + \sum_{r=0}^{\lfloor \frac{l-1}{n_x} \rfloor} |\iota(r)| - ((l-1) \bmod n_x), \quad n_2^{(l)} = n_J - n_1^{(l)} - n_{H_l}.$$

The identity matrix $\mathbb{I}_1^{(l)}$ of dimension $(l-1)$ embeds the lower-dimensional HOUSEHOLDER matrix $\begin{bmatrix} \tilde{H}_l \\ \mathbb{I}_2^{(l)} \end{bmatrix}$ into the $\mathbb{R}^{n_J \times n_J}$. The matrix \tilde{H}_l is dense in general; the identity matrix $\mathbb{I}_2^{(l)}$ at the lower right part is responsible for maintaining the sparsity.

The successive application of the HOUSEHOLDER matrices H_l while transforming \tilde{J} in upper triangular form always causes a fill-in of only $\sum_{r=0}^{k-1} |\iota(r)|$ rows above the blocks D_k in eq. (5.45), leading to the sparsity structure depicted in eq. (5.43).

The remaining HOUSEHOLDER matrices \tilde{H}_l for $l = (n_{\text{MS}} - 1)n_x + 1, \dots, (n_{\text{MS}} - 1)n_x + n_p$, working on the lower parts of the right band of P_k^R and P_k^S blocks, are usually dense but do not interfere with the sparsity of $H_{(n_{\text{MS}}-1)n_x} \cdot \dots \cdot H_1 \cdot \tilde{J}$ as they only act on the last $n_J - l$ columns and rows.

Figure 5.8 gives an illustration. □

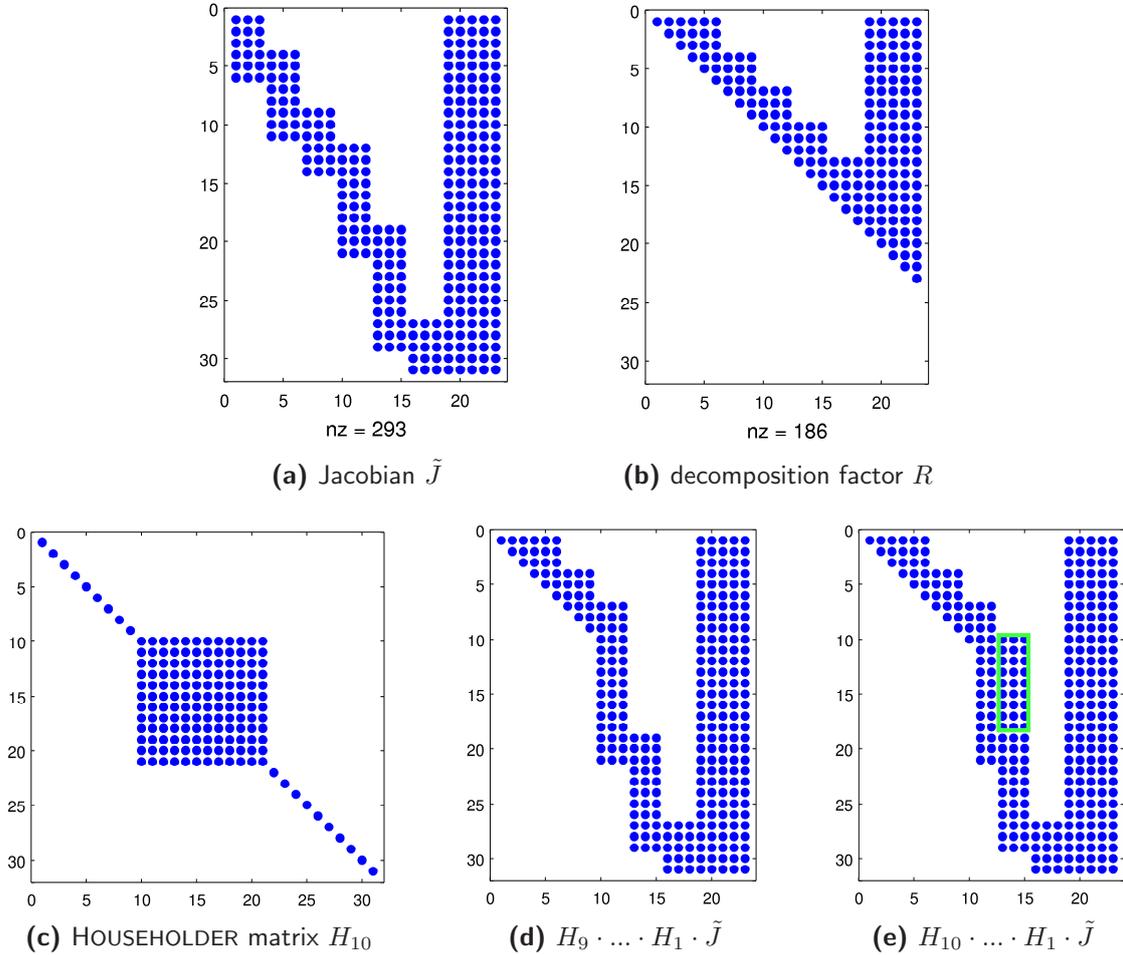


Figure 5.8: Examples of sparsity patterns of Jacobian J and decomposition factor R for an unconstrained problem with $n_{MS} = 6$, $n_x = 3$, $n_p = 5$, and $\{3, 2, 0, 4, 5, 2\}$ scalar measurements in the respective shooting intervals.

(a) shows the sparsity pattern of the row-permuted Jacobian \tilde{J} in eq. (5.45), while (b) displays the sparsity pattern after an HOUSEHOLDER QR decomposition has been applied. The upper quadratic part is the decomposition factor R , whose maximum number of nonzero elements $\text{nnz}_{\max}(R) = 186$ is independent of the number of measurements and only depends on n_{MS} , n_x , and n_p , see eq. (5.42).

The series at the lower part depicts exemplarily the fill-in effect of (intermediate) HOUSEHOLDER matrices: (c) shows the sparsity structure of the tenth HOUSEHOLDER matrix H_{10} , whose upper left identity is due to the embedding to the correct space, and whose lower right identity is due to the sparsity structure of the intermediate matrix $H_9 \cdot \dots \cdot H_1 \cdot \tilde{J}$ that is displayed in (d); also see section 5.5.4. The fill-in effect of its application is seen in (e), showing the intermediate matrix $H_{10} \cdot \dots \cdot H_1 \cdot \tilde{J}$, where the fill-in has been highlighted in a green box.

5.5.5 Lifting with local parameters

While in the stochastic parameter estimation problem 5.18, $p \in \mathbb{R}^{n_p}$ denotes a unique set of parameters, one might introduce *per interval* parameters $p_k \in \mathbb{R}^{n_p}$ for $k = 0, \dots, n_{\text{MS}} - 1$, and add linear constraints of the form

$$p_{k-1} - p_k = 0 \quad (k = 1, \dots, n_{\text{MS}} - 1) \quad (5.47)$$

ensuring that the per interval parameters coincide in the solution. Mathematically, the additional $(n_{\text{MS}} - 1)n_p$ variables are determined by these additional constraints, so the degree of freedom does not change, and the “lifted” problem is equivalent to the original one.

However, the enlarged search space gives additional structure to the (block-)sparse linear systems. The lifted problem (compare section 5.4.3) may be formulated as:

5.22 Problem (Constrained nonlinear stochastic parameter estimation problem with local parameters)

$$\begin{aligned} \min_{\substack{s_0, \dots, s_{n_{\text{MS}}-1}, \\ p_0, \dots, p_{n_{\text{MS}}-1}}} & \frac{1}{2} \sum_{i=1}^{n_M} \sigma_i^{-2} \cdot (h_i(t_i, x(t_i; s, p), p_{\kappa(t_i)}) - \eta_i)^2 + \frac{1}{2} \sum_{k=1}^{n_{\text{MS}}-1} \sum_{l=1}^{n_x} \omega_{k,l}^2 \cdot \alpha_{k,l}^2 \\ \text{s.t.} & \quad \dot{x}(t; t_k^{\text{MS}}, s_k, p_k) = f(t, x, p_k) \quad t \in [t_k^{\text{MS}}, t_{k+1}^{\text{MS}}) \\ & \quad x(t_k^{\text{MS}}; t_k^{\text{MS}}, s_k, p_k) = s_k \quad k = 0, \dots, n_{\text{MS}} - 1 \\ & \quad \tilde{c}^{ec}(x(t_1^{ec}; s, p), \dots, x(t_{n_{\text{EC}}}^{ec}; s, p), p) = 0 \\ & \quad c^{ic}(x(t_1^{ic}; s, p), \dots, x(t_{n_{\text{IC}}}^{ic}; s, p), p) \geq 0 \end{aligned}$$

with $\alpha_k := x(t_k^{\text{MS}}; t_{k-1}^{\text{MS}}, s_{k-1}, p_{k-1}) - s_k$ being the stochastic jumps at the inner shooting nodes, $p = (p_0^T, \dots, p_{n_{\text{MS}}-1}^T)^T \in \mathbb{R}^{n_{\text{MS}}n_p}$ denoting the combined vector of the per-interval parameters p_k , and \tilde{c}^{ec} combines the original point constraints c^{ec} and the linear constraints from eq. (5.47) for the local parameters.

With functions F_1 and F_2 as in eqs. (5.36) and (5.37), adjusted to the lifted problem 5.22, and modifying the definition of P_k^R and P_k^S in eq. (5.44) to

$$P_k^R := \frac{dF_1^{(k)}}{dp_k} \quad (k = 0, \dots, n_{\text{MS}} - 1) \quad \text{and} \quad P_k^S := \frac{dF_2^{(k+1)}}{dp_k} \quad (k = 0, \dots, n_{\text{MS}} - 2)$$

the sparsity pattern of the combined residual vector’s Jacobian in the linear systems then has the form (compare section 5.5.1):

$$J(s, p) = \frac{dF(s, p)}{d(s, p)} = \left(\begin{array}{c|c} \frac{dF_1}{ds} & \frac{dF_1}{dp} \\ \hline \frac{dF_2}{ds} & \frac{dF_2}{dp} \end{array} \right) = \left(\begin{array}{c|c} \frac{dF_1}{ds_0} \dots \frac{dF_1}{ds_{n_{\text{MS}}-1}} & \frac{dF_1}{dp_0} \dots \frac{dF_1}{dp_{n_{\text{MS}}-1}} \\ \hline \frac{dF_2}{ds_0} \dots \frac{dF_2}{ds_{n_{\text{MS}}-1}} & \frac{dF_2}{dp_0} \dots \frac{dF_2}{dp_{n_{\text{MS}}-1}} \end{array} \right)$$

The matrix $J(s, p)$ has now dimension $(n_M + (n_{MS} - 1)n_x) \times (n_{MS}(n_x + n_p))$; and we note again that the matrices R_k and P_k^R are not square in general. The equality conditions eq. (5.47) and the stochastic jumps establish the connection between the interval-wise problems in the linearized subproblems.

We remark that a globalization technique as described in chapter 1 often benefits from an initial infeasibility in the parameter constraints eq. (5.47), which is maintained as long as no full steps are accepted. Further, the local area of contraction is often increased. For a general discussion on lifting, we refer to [AlbersmeyerDiehl2010].

The sparsity structure in the QR decomposition of the Jacobian of the lifted problem, however, is disturbed by both the additional two block (half-)diagonals of P_k^R and P_k^S and the equality conditions on the per-interval parameters p_k . Still, the resulting decomposition factor $R^{\text{lifted}} \in \mathbb{R}^{(n_{MS}(n_x + n_p)) \times (n_{MS}(n_x + n_p))}$ is sparse – but with a more complex structure – as it can be shown by similar considerations as in section 5.5.4. It has the general form:

We note that there exist some edge cases of measurement time distribution within the shooting intervals such that the decomposition factor R^{lifted} exhibits more zeros in the lower right block diagonal.

5.6 Extensions

5.6.1 A homotopy in jump regularization weighting

The objective in the stochastic parameter estimation problem 5.18 (see page 144),

$$\min_{s_0, \dots, s_{n_{MS}-1}, p} \frac{1}{2} \sum_{i=1}^{n_M} \sigma_i^{-2} \cdot (h_i(t_i, x(t_i; s, p), p) - \eta_i)^2 + \frac{1}{2} \sum_{k=1}^{n_{MS}-1} \sum_{l=1}^{n_x} \omega_{k,l}^2 \cdot \alpha_{k,l}^2, \quad (5.49)$$

may be interpreted as a **multi-objective optimization functional**: Both the unknown parameters shall be recovered (by the first weighted least squares term), and the concrete realization of the underlying SDE's trajectory shall be approximated (by jump regularization in the second least squares term).

As shown in the numerical examples section, the proposed method is robust against moderate changes in the jump regularization. For low jump regularization weights, however, the computed solution might still recover the data – i.e. the first least squares term is close to zero – but the thereby computed parameters might be wrong, as we have seen in the WIENER exponential example in section 5.4.4 on page 145.

Also, for problems with complex dynamics, and depending on the available data quality and quantity, finding a **reasonable initialization** for the shooting node variables can be a hard subproblem – especially in the case of intermittent observations with long time periods without any measurement.

With the results of section 5.3 (On the distance between SDE and ODE solutions) in mind, it can be beneficial to initially stay “close” to the deterministic (ODE) trajectory. This can be achieved by increased jump regularization weights. However, the whole idea of the proposed parameter estimation method is thwarted if artificially high regularization weights force a strong “binding” to the ODE trajectory.

For pathological cases, we propose to **iteratively adjust the jump weights** by a homotopy strategy, by using the objective

$$\min_{s_0, \dots, s_{n_{MS}-1}, p} \frac{1}{2} \sum_{i=1}^{n_M} \sigma_i^{-2} \cdot (h_i(t_i, x(t_i; s, p), p) - \eta_i)^2 + \mu \cdot \frac{1}{2} \sum_{k=1}^{n_{MS}-1} \sum_{l=1}^{n_x} \omega_{k,l}^2 \cdot \alpha_{k,l}^2, \quad (5.50)$$

instead of the original one in eq. (5.49), with an homotopy parameter $\mu \in [1, \infty)$. An infinite value $\mu = \infty$, is equivalent to forcing *continuity conditions* (eq. (2.10)) on the trajectory, as in the case of multiple shooting in ODE parameter estimation (section 2.2.3). The lower bound value $\mu = 1$ corresponds to the intended jump regularization weights.

By gradually driving the homotopy parameter μ from an initially large value down to 1 *during the optimization*, the optimizer has more control over the trajectory, progressively allowing higher gaps at the shooting nodes. This may also speed up convergence for non-pathological problems.

5.6.2 Shooting grid refinement

If one is in the lucky case of abundant high quality measurements, e.g. densely sampled temperature data with low variances, one is able to identify areas where grid refinement would be beneficial (see figure 6.14a on page 194 for an example). While manual refinement is always

possible, this can be also done in an automated way: Intervals with considerably large residuals may be split, and the respective jump regularization weights should be adapted. A similar approach can be used if extraordinary stochastic jumps are observed.

We discuss a strategy based on this methodology in the BISTABAER example in section 6.3.3.7 on page 191.

5.7 Software package `:sfit`

The parameter estimation method described in this chapter has been implemented in the software package `:sfit`, that is presented in appendix S.

6 Numerical Examples

This chapter applies the methodology developed in chapter 5 to examples from biology and finance. The results are appealing even for a naïve equidistant choice of the shooting grid. We study the effects of this choice in the BISTABAER model example (section 6.3) in some detail.

We further show that the proposed method may also be successfully applied to LÉVY-driven SDEs with jumps, and give an example for which we are also able to estimate a diffusion parameter.

The first example is the **FITZHUGH-NAGUMO oscillator**; a classical model for neuronal firing, with two states and four parameters. Starting from the original ODE model, which rests on its steady state without any excitation, an SDE formulation is retrieved that is affected by noise in one component, leading to irregular “firing” of the modelled neuron (i.e. the oscillator entering its limit cycle). The SDE model is used for artificial data generation.

A simulation study is performed for estimating the 4 model parameters from the generated data. The intention of this series is to determine accuracy and standard deviation of the estimation, to check for biased estimates, and to investigate the effect of different choices of jump regularization weights. Inter alia, it is shown that higher jump regularization weights are beneficial if the measurements are affected by additional measurement noise, and in partial observation settings.

The second example copes with a **model for calcium oscillations in eukariotic cells**. The system consists of a three dimensional state space, governed by 12 kinetic parameters. An SDE formulation is retrieved, from which data for subsequent parameter estimation series is generated, similar to the settings in the FITZHUGH-NAGUMO example. It is shown that the system is unidentifiable for the partial observation of calcium only. A further parameter estimation study is thus performed on a subset of parameters that are identifiable by sole calcium measurements.

Further, for an example with intermittent observations, it is shown that the proposed parameter estimation method is not only able to recover the kinetic parameters, but also to regenerate intermediate (non-observed) parts of the trajectory with surprising accuracy.

As third example from biology, a **bistable allosteric enzyme regulator model** is investigated. Internally driven by an (unobserved) bistable oscillator, which is driven by a WIENER process in the SDE, the observed trajectories significantly differ in every realization.

Again, a parameter estimation series is performed for retrieving the two kinetic parameters of the model from both error-free and noisy measurements. Further, it is shown that the hidden internal control may be reconstructed.

The surprising finding of a lowered estimation accuracy when using error-free data is investigated, explained, and a remedy in terms of either adjusting the shooting grid or the jump regularization weights is given.

The last example originating from finance proves the applicability of the proposed parameter estimation method beyond its original design. Successful parameter estimation (with low uncertainty) is shown for the ORNSTEIN-UHLENBECK or VAŠÍČEK model for the time evolution of interest rates.

This 1-dimensional stochastic process is driven by a pure jump LÉVY process. Using an automated shooting grid generation heuristic solely based on the available measurement data, the mean reversion level and mean reversion rate parameters of the SDE's drift term are well recovered.

It is further shown that the process' volatility (a diffusion parameter!) may be estimated by analysing the distribution of the stochastic jumps α_k .

6.1 The FITZHUGH-NAGUMO Oscillator

The *FITZHUGH-NAGUMO (FHN) oscillator* is a two-species model of an excitable system, derived by Richard FITZHUGH⁽¹⁾ [FitzHugh1961] as a simplification of the original HODGKIN-HUXLEY model [HodgkinHuxley1952] describing the initiation and propagation of action potential in neuronal cells. Jin-ichi NAGUMO⁽²⁾, together with Suguru ARIMOTO and Shuji YOSHIZAWA derived the corresponding circuit model [Nagumo1962] for electronic simulation on analog computers.

Nowadays known as FHN oscillator, the model was originally named *BONHÖFFER-VAN DER POL (BVP) model* by its inventor, as it contains the VAN DER POL oscillator [vander-Pol1926] for certain parameter values and “is also qualitatively similar to that proposed by Bonhoeffer (1941, 1948, 1953) and by Bonhoeffer and Langhammer (1948) to describe the Ostwald-Lillie iron wire model of nerve. These authors drew comparable phase planes, but specified no equations” [FitzHugh1961].

The original BVP model derived by FITZHUGH, with an external stimulus $\iota(t)$ is given as

$$\dot{y}_1 = \gamma(y_2 + y_1 - y_1^3/3 + \iota(t)) \quad \dot{y}_2 = -(y_1 - \alpha + \beta y_2)/\gamma$$

and can be transformed by coordinate transformation and re-parametrization into the system

$$\dot{x}_1 = x_1 - zx_1^3 - x_2 + I(t) \quad \dot{x}_2 = a(x_1 + b + cx_2) \quad (6.1)$$

where now $I(t)$ acts as external stimulus. Without any excitation, the FHN has a stable node or focus, which can be derived by computing the nullclines of eq. (6.1):

$$0 = x_1 - zx_1^3 - x_2 \quad 0 = a(x_1 + b + cx_2)$$

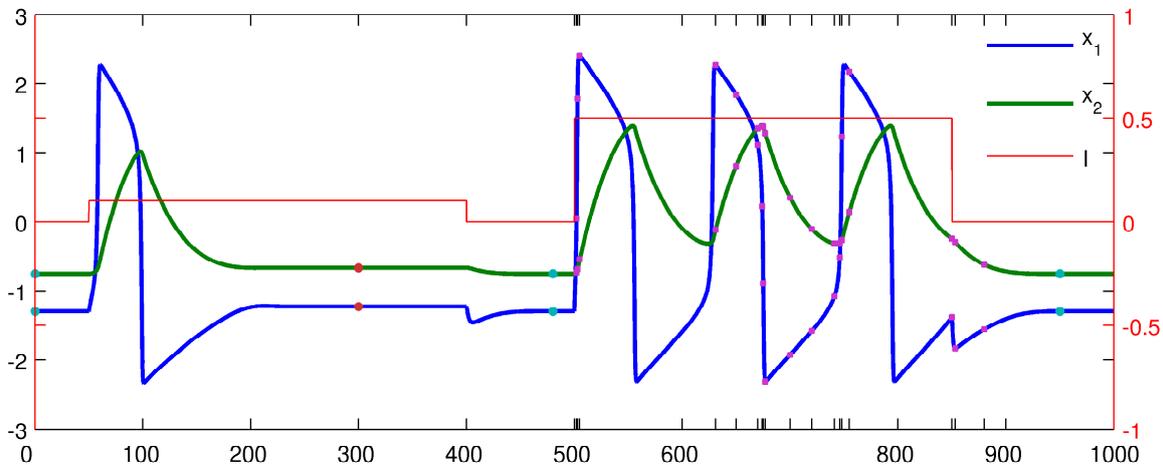
showing that there is no chance of estimating the value of parameter a by measuring only the steady state values. For $cz \neq 0$, the cubic equation has a single real solution, delivering the **nullclines** as:

$$x_1 = \frac{1}{3\sqrt[3]{2cz}} \cdot k - \sqrt[3]{2}(-c-1) \cdot \frac{1}{k} \quad x_2 = x_1 - zx_1^3 \quad (6.2)$$

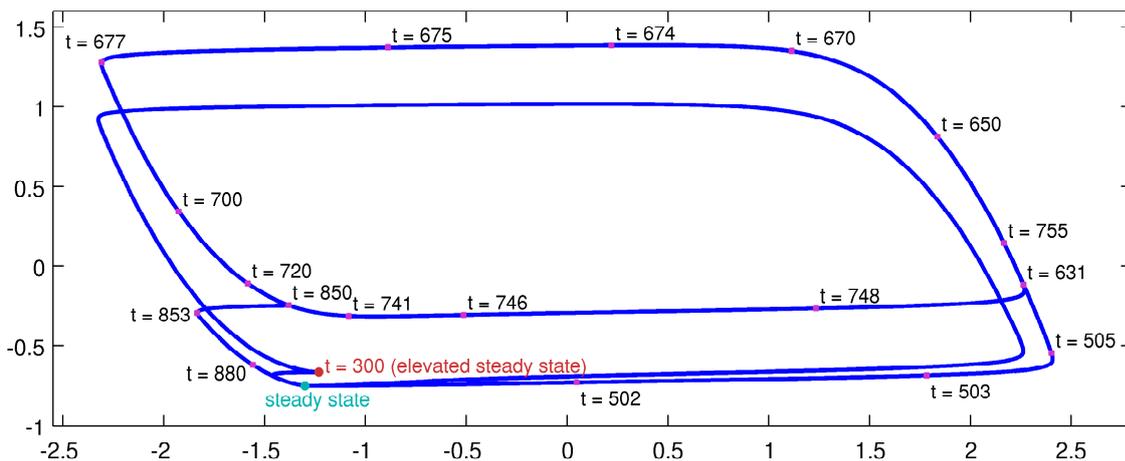
with $k = \left(\sqrt{729b^2c^4z^4 + 108(-c-1)^3c^3z^3 + 27bc^2z^2} \right)^{\frac{1}{3}}$.

⁽¹⁾ Richard FITZHUGH, 1922-2007, American biophysicist

⁽²⁾ Jin-Ichi NAGUMO, 1926-1999, Japanese physicist



(a) Time course of stimulated FITZHUGH-NAGUMO oscillator.



(b) Phase portrait of the above FITZHUGH-NAGUMO oscillator.

Figure 6.1: Time course and phase portrait of the FHN oscillator upon excitation by a piecewise constant function $I(t)$ (red line). While initially being in the steady state (cyan circle), the medium sized first stimulation at $t = 50$ leads to a single excitation of the oscillator, followed by resting on an *elevated* steady state (red circle). As soon as the first stimulation stops ($t = 400$), the system “falls back” towards its original steady state.

The strong second stimulation at $t = 500$ forces the system to enter a limit cycle as long as the external stimulus is present, again relaxing to the steady state as soon as the excitation vanishes ($t = 850$).

In this simulation, the kinetic parameters were chosen as $a = 0.02$, $b = 0.7$, $c = -0.8$, $z = 0.25$, with initial values being very close to the steady state – see eqs. (6.2) and (6.3).

To display the time course in the phase portrait, some time points are marked in both plots (magenta squares).

6.1.1 Meaning of parameters

As a phenomenological description, parameter a determines the frequency of firing (excitation) upon a constant lasting external impulse $I(t) = i_0$, while parameter z determines the maximum amplitude upon any (reasonable) external excitation. The remaining parameters b and c are coupled and both influence the excitability of the FHN model in an absolute (b) and relative (c) way, and both influence its ability of successive firing upon a lasting external impulse.

In figure 6.2, the sensitivities of the steady state around the true parameter values are depicted. Comparing the sensitivity of the steady state value on the parameters b , c , and d , it becomes obvious that the steady state value is mostly determined by z , and b and c have opposite effects in about the same magnitude.

6.1.2 Expected bias in parameters b and c

With the estimation of parameter values of b and c in mind, the last finding together with the formula for the steady state value of the FITZHUGH-NAGUMO oscillator, eq. (6.2), implies that these two parameters are expected to be biased if the steady state is permanently disturbed by a WIENER-driven noise.

We will find this presumption confirmed by the estimation series in section 6.1.4. Also compare the elevated steady state upon excitation depicted in the phase portrait in figure 6.1b.

6.1.3 SDE formulation

For this numerical study, we choose the following parameter values, initial conditions, and time domain:

$$\begin{aligned} a = 0.02, \quad b = 0.7, \quad c = -0.8, \quad z = 0.25, \quad D = 0.1 \\ x_1(t_0) = -1.300, \quad x_2(t_0) = -0.7506, \quad [t_0, t_f] = [0, 1000] \end{aligned} \quad (6.3)$$

with initial conditions being close (up to 4 digits) to the steady state. Further, we set the external stimulus function I in the first component to be a WIENER process, such that the FHN oscillator model can be written as SDE with drift f and diffusion g :

$$\begin{aligned} dX_t = f(X_t, p)dt + g(X_t, p)dW_t \\ f(X_t, p) = \begin{pmatrix} x_1 - zx_1^3 - x_2 \\ a(x_1 + b + cx_2) \end{pmatrix}, \quad g(X_t, p) = \begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix} \end{aligned} \quad (6.4)$$

with all kinetic parameters collected in the vector $p = (a, b, c, z)^T$, and D denoting the diffusion parameter.

Figure 6.3a shows the time course of an FHN oscillator with parameter values as in eq. (6.3) for a certain realization of a driving WIENER process. The WIENER noise is clearly visible in the first component. Due to the oscillator's damping properties, the impact on the second component is smoothed out and much smaller. A deterministic interpretation, i.e. as ODE without diffusion, for the same set of parameters is depicted in figure 6.3b, where we observe no changes in the system's state, as without any external excitation, the excitable system stays in its steady state.

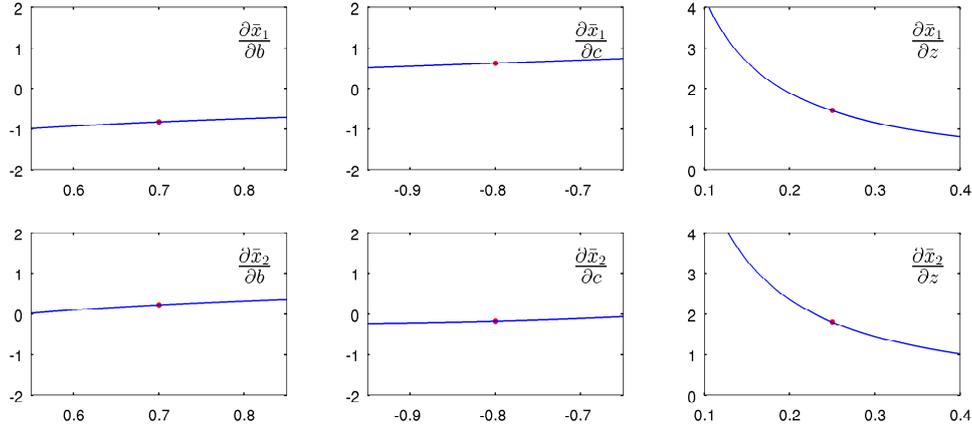


Figure 6.2: Steady state sensitivities of the FITZHUGH-NAGUMO oscillator with respect to the parameters b , c , and z . Depicted are the derivatives of the steady state value (\bar{x}_1, \bar{x}_2) , see eq. (6.2), around the nominal values $b = 0.7, c = -0.8, z = 0.25$, varying one while keeping the others fixed (the steady state value does not depend on parameter a). One can see that parameter z has the highest impact on the steady state value at these parameter values.

6.1.4 Simulation study and a closer look at jump regularization effects

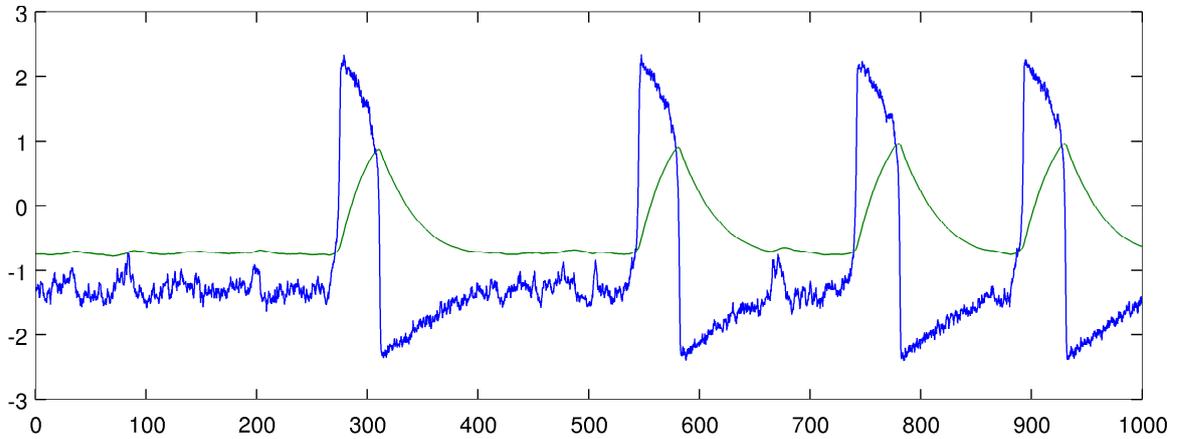
To investigate the performance of the estimation method described in section 5.4, we make a test series consisting of 4 observation scenarios – full and partial state observations, exact and noisy measurements – and two choices of jump regularization weights to investigate their impact on the estimation.

The resulting 8 scenarios are:

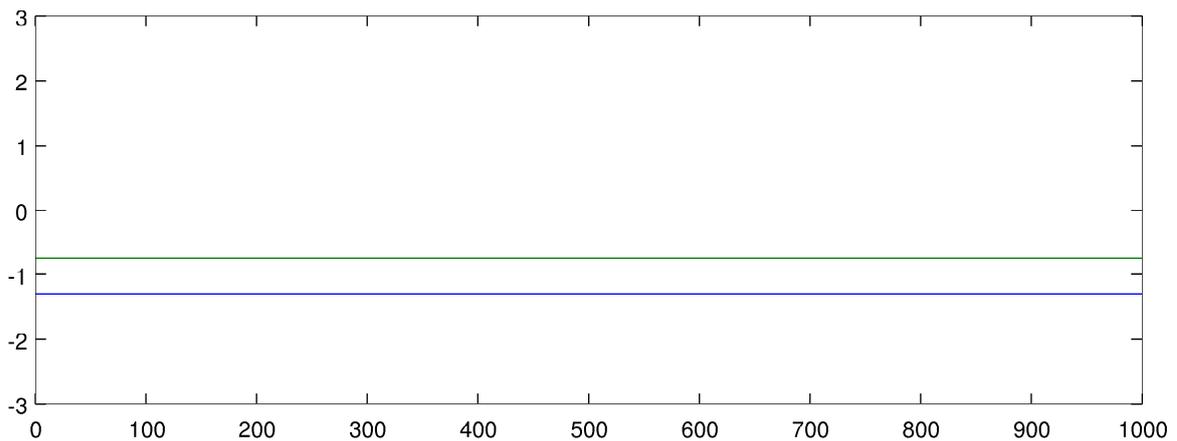
- (A) exact measurements, full observation, jump weight 1.0
- (B) exact measurements, full observation, jump weight 10.0
- (C) exact measurements, partial observation, jump weight 1.0
- (D) exact measurements, partial observation, jump weight 10.0
- (E) noisy measurements, full observation, jump weight 1.0
- (F) noisy measurements, full observation, jump weight 10.0
- (G) noisy measurements, partial observation, jump weight 1.0
- (H) noisy measurements, partial observation, jump weight 10.0

For each setting, 100 realizations⁽³⁾ of the stochastic FITZHUGH-NAGUMO oscillator eq. (6.4) driven by a WIENER process acting on the first component with diffusion coefficient $D = 0.1$, and kinetic parameter values given in eq. (6.3) were made, and the parameters therein estimated using the method proposed in section 5.4.

⁽³⁾ generated by a stochastic EULER scheme with step size 10^{-3}



(a) stochastic (SDE) interpretation of the FITZHUGH-NAGUMO model



(b) deterministic (ODE) interpretation of the same FITZHUGH-NAGUMO model

Figure 6.3: Simulation of the FITZHUGH-NAGUMO oscillator with a driving WIENER process effective on the first component. In the SDE interpretation, the first component (blue) of the FITZHUGH-NAGUMO oscillator is disturbed by a driving standard WIENER process, visible as a “noisy” trajectory for that component. The stability properties of the FITZHUGH-NAGUMO oscillator push both components towards the steady state, unless the displacement by the WIENER process is too high; in that case, the oscillator traverses its limit cycle and returns towards its steady state again.

The ODE interpretation (which is equivalent to an SDE interpretation with zero diffusion), shows no activity at all, since the oscillator remains at its steady state.

The kinetic parameters are $a = 0.02$, $b = 0.7$, $c = -0.8$, $z = 0.25$, diffusion parameter $D = 0.1$; initial values were chosen close to the steady state. Both simulations are generated by an Euler scheme with stepsize 10^{-3} .

6.1.4.1 Measurement functions and weights

All settings share the same sample interval of 5 time units, resulting in 201 full or partial state observations. In the case of partial observation, only the first component x_1 is measured. In the “noisy” scenarios, normally distributed noise with zero mean and a standard deviation of 0.1 (about 10% of the steady state value) is added to the measurements.

That is, we have as measurement functions⁽⁴⁾:

$$\begin{aligned} h_i(x(t_i)) &= x(t_i) \in \mathbb{R}^2 \quad (i = 1, \dots, 201) \quad \text{for scenarios (A), (B), (E), (F)} \\ h_i(x(t_i)) &= x_1(t_i) \in \mathbb{R}^1 \quad (i = 1, \dots, 201) \quad \text{for scenarios (C), (D), (G), (H)} \end{aligned}$$

with an equidistant measurement grid $T^M = \{0, 5, 10, 15, \dots, 1000\}$. Note that we have omitted writing the dependence on the parameter vector p .

In the undisturbed settings, the measurement weights are chosen as 1.0, in the “noisy” scenarios, the reciprocal of the above standard deviation is used⁽⁵⁾.

6.1.4.2 Shooting grid and node initialization

The time domain $[0, 1000]$ is divided into 50 evenly sized intervals, i.e. the shooting grid is

$$T^{MS} = \{t_0^{MS}, \dots, t_{50}^{MS}\} \quad \text{with } t_0^{MS} = 0, \quad t_{50}^{MS} = 1000, \quad t_k^{MS} = 20k \quad (k = 1, \dots, 49)$$

and the shooting node variables s_k ($k = 0, \dots, 49$) are initialized with their temporally most proximate measurement.

In the partial observation settings (C), (D), (G), (H), the unobserved species x_2 is initialized with its approximate steady state value -0.7506 .

6.1.4.3 Initial parameter guess and stopping criterion for the GAUSS-NEWTON solver

The initial guess of the parameters is set to 50% of the true values (see table 6.1), ensuring that the local area of contraction of the GAUSS-NEWTON method is left, thus globalization takes effect. We remark that also for more distant as well as randomized initial guesses, convergence to the solution is observed.

The optimization is stopped when the maximum norm of the search direction $\|\Delta x_k\|_\infty$ (see section 1.3 on page 14) falls below 10^{-3} .

6.1.4.4 Constraints on optimization variables

The following constraints (beyond reachable values for the FITZHUGH-NAGUMO oscillator) on the *state variables* at the shooting nodes are set:

$$x_1, x_2 \in [-4, 4].$$

The following constraints on the *kinetic parameters* ensure the right sign of the parameters:

$$a \in [0.001, 0.1], \quad b \in [0.01, 2], \quad c \in [-2, -0.01], \quad z \in [0.01, 1].$$

In the test series, no constraints are active in the solutions.

⁽⁴⁾ we use a vector-valued measurement function here solely for the sake of convenient notation. Using scalar measurement functions h_i as required in section 2.1.2, we might write $h_i(x(t_i)) = \begin{cases} x_1(t_i) & \text{if } i \in 2\mathbb{N} - 1 \\ x_2(t_i) & \text{if } i \in 2\mathbb{N} \end{cases}$ with a multiset $T^M = \{0, 0, 5, 5, 10, 10, \dots, 1000, 1000\}$ holding the (now non-unique) measurement times.

⁽⁵⁾ this increases the impact of measurements (in the least squares objective) in the noisy scenarios.

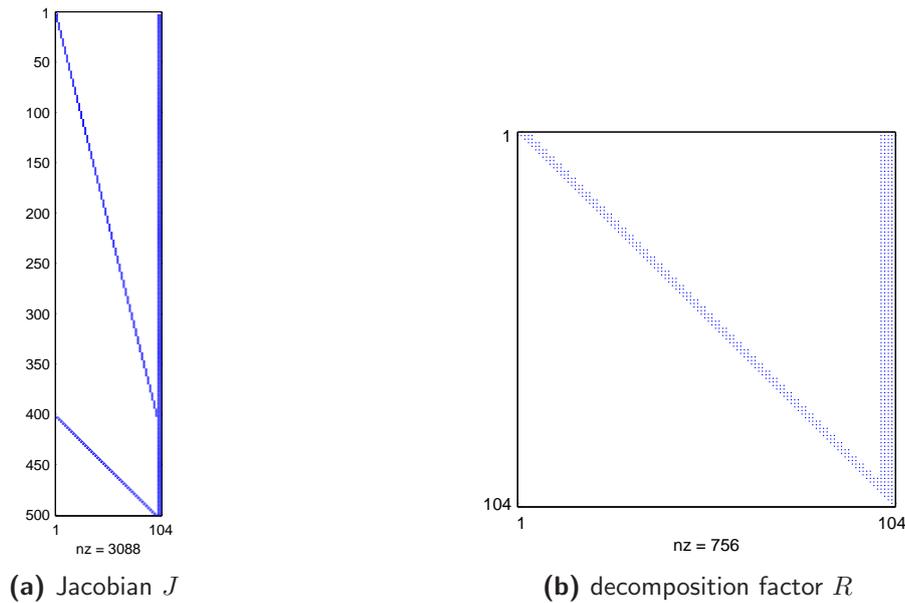


Figure 6.4: Sparsity pattern of the combined residual vector's Jacobian - FITZHUGH-NAGUMO oscillator with full state measurements.

(a) Jacobian J with dimension 500×104 (52000 elements), nonzero elements: 3088 (5.9%)

(b) decomposition factor R with dimension 104×104 (10816 elements), nonzero elements: 756 (7.0%)

See section 5.5.1 for details on the sparsity pattern.

6.1.4.5 Sparsity pattern of the combined residual vector's Jacobian

Figure 6.4 shows the sparsity pattern of the the combined residual vector's Jacobian J , i.e. the system matrix of the linearized problem without constraints, as well as the sparsity pattern of its decomposition factor that may be used for solving the linearized problem.

With a total dimension of 500×104 , only 3088 (5.9%) out of 52000 elements are nonzero. The decomposition factor $R \in \mathbb{R}^{104 \times 104}$ has roughly the same low occupancy rate of 7.0%.

See section 5.5.1 for details.

6.1.4.6 Results of the test series

The results of the test series are shown in table 6.1. As one would expect, having exact full state measurements (scenarios **(A)** and **(B)**) gives the best estimation results. For parameters a , b , and z , the mean estimate is very close (0.3–2.6% relative error) to the true parameter values, and the standard deviation of the parameter estimates is satisfying.

As already foreshadowed in section 6.1.2, parameters b and c that describe the steady state value of the non-excited FITZHUGH-NAGUMO oscillator, show a small bias in most of the testing scenarios, because of the enduring excitation by the driving WIENER process and thus moving the steady state value to a (varying) elevated steady state (see figure 6.1b). Parameters a and z may be recovered fairly well in all experimental settings.

For partial noisy observations (scenarios **(G)** and **(H)**), stronger jump regularization leads to considerable improvement, as both the relative error of the estimates as well as their variance is reduced (see the lower part of table 6.1; also compare the similar findings in the calcium oscillation example in the next section).

Table 6.1: Estimation test series: FITZHUGH-NAGUMO oscillator. Results of parameter estimation on 100 independent simulations (in each setting) of a FITZHUGH-NAGUMO oscillator, whose first component is affected by a WIENER process, eq. (6.4). For a discussion of the bias in parameters b and c , see section 6.1.2. Initial guess for the kinetic parameters was 50% of the true values. Values rounded to 3 digits.

exact observations without measurement error						
true parameter name value	scenario (A) full observation all jump weights 1.0			scenario (B) full observation all jump weights 10.0		
	estimate \pm SD	(SD%)	RelErr%	estimate \pm SD	(SD%)	RelErr%
a 0.02	0.021 \pm 0.001	(3.8%)	2.6%	0.020 \pm 0.001	(6.6%)	2.0%
b 0.7	0.698 \pm 0.021	(3.0%)	0.3%	0.689 \pm 0.059	(8.5%)	1.6%
c -0.8	-0.728 \pm 0.057	(7.8%)	9.0%	-0.692 \pm 0.085	(12.2%)	13.5%
z 0.25	0.251 \pm 0.009	(3.4%)	0.5%	0.253 \pm 0.027	(10.8%)	1.2%
true parameter name value	scenario (C) partial observation all jump weights 1.0			scenario (D) partial observation all jump weights 10.0		
	estimate \pm SD	(SD%)	RelErr%	estimate \pm SD	(SD%)	RelErr%
a 0.02	0.021 \pm 0.001	(4.0%)	6.0%	0.021 \pm 0.001	(3.1%)	3.3%
b 0.7	0.645 \pm 0.068	(10.6%)	7.9%	0.670 \pm 0.039	(5.8%)	4.2%
c -0.8	-0.782 \pm 0.092	(11.8%)	2.3%	-0.713 \pm 0.063	(8.9%)	10.9%
z 0.25	0.240 \pm 0.013	(5.5%)	3.9%	0.236 \pm 0.013	(5.3%)	5.7%
SD%: standard deviation of the estimate · RelErr%: relative deviation of the estimate from the true parameter value						
noisy observations with measurement error $\epsilon \sim \mathcal{N}(0, 0.1^2)$						
true parameter name value	scenario (E) full observation all jump weights 1.0			scenario (F) full observation all jump weights 10.0		
	estimate \pm SD	(SD%)	RelErr%	estimate \pm SD	(SD%)	RelErr%
a 0.02	0.020 \pm 0.001	(6.7%)	0.8%	0.021 \pm 0.002	(7.7%)	4.0%
b 0.7	0.686 \pm 0.077	(11.2%)	2.0%	0.682 \pm 0.064	(9.4%)	2.5%
c -0.8	-0.673 \pm 0.108	(16.0%)	15.9%	-0.770 \pm 0.120	(15.6%)	3.7%
z 0.25	0.250 \pm 0.009	(3.7%)	0.1%	0.255 \pm 0.018	(7.1%)	1.9%
true parameter name value	scenario (G) partial observation all jump weights 1.0			scenario (H) partial observation all jump weights 10.0		
	estimate \pm SD	(SD%)	RelErr%	estimate \pm SD	(SD%)	RelErr%
a 0.02	0.021 \pm 0.003	(14.1%)	6.4%	0.021 \pm 0.001	(6.8%)	7.3%
b 0.7	0.509 \pm 0.165	(32.4%)	27.3%	0.652 \pm 0.094	(14.4%)	6.9%
c -0.8	-0.887 \pm 0.238	(26.8%)	10.9%	-0.783 \pm 0.136	(17.4%)	2.1%
z 0.25	0.227 \pm 0.020	(8.6%)	9.2%	0.237 \pm 0.019	(7.9%)	5.1%
SD%: standard deviation of the estimate · RelErr%: relative deviation of the estimate from the true parameter value						

6.1.4.7 A closer look at a single realization

Figure 6.5 shows the fitted trajectories for a certain realization of a WIENER-driven FITZHUGH-NAGUMO oscillator for each of the eight testing scenarios discussed in section 6.1.4.

As one can see by comparing figure 6.5a and figure 6.5b, smaller jump weights allow a better reproduction of the influences of the driving WIENER process, while stronger jump regularization leads to a “more steady” trajectory (in both settings, exact measurements are used). Depending on the user’s intention, each of them might be more suitable.

The figures 6.5c to 6.5h show detail enlargements of the approximate interval $[280, 620]$ of the fitted trajectories in the respective scenarios. The interval contains one transit in the limit cycle and a subsequent stay around the steady state.

Especially in the “noisy” scenarios, i.e. with additional measurement noise, the application of higher jump regularization weights delivers a much “smoother” trajectory, which would be beneficial if state estimation is a user goal.

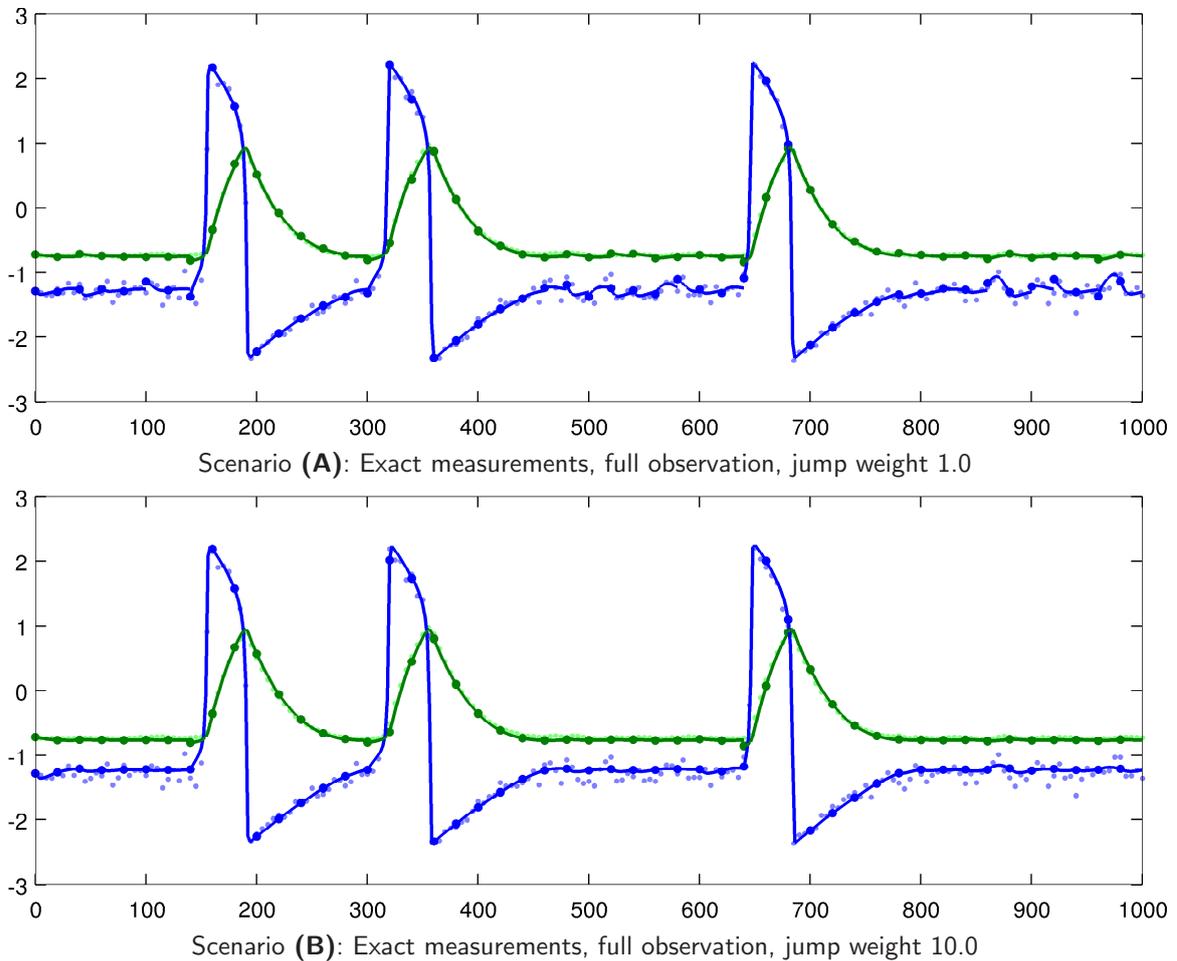
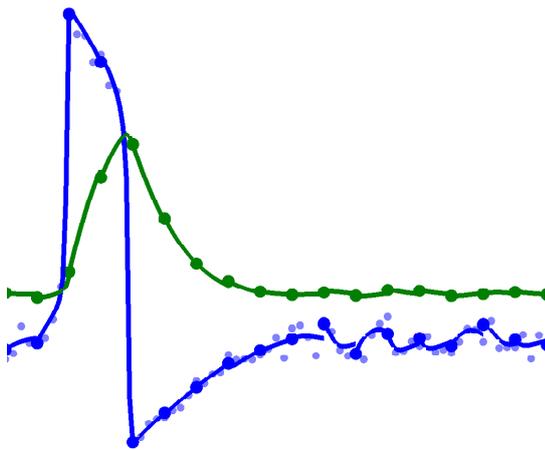


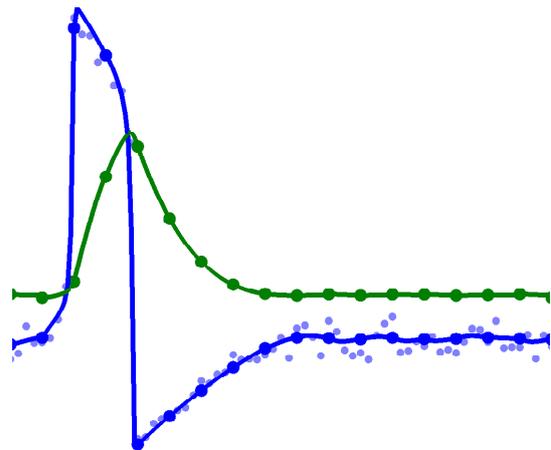
Figure 6.5: FITZHUGH-NAGUMO example fit (continues on the facing page)

Fitted trajectories for scenarios (A) and (B) for a certain realization of the FITZHUGH-NAGUMO oscillator. While in the low jump weight scenario (A) the trajectory “mimics” the driving WIENER process affecting the first component, the higher jump weights in (B) act as regularization, leading to a “more continuous” trajectory. The detail enlargements for scenarios (C) to (H) on the next page show this more clearly.

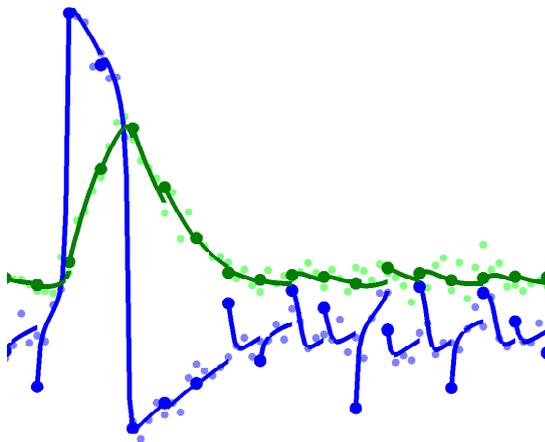
Fitted trajectory of component x_1 as blue line (—), of component x_2 as green line (—), shooting nodes as dots in the respective colors (\bullet , \bullet). Measurements of x_1 as small light blue dot (\bullet), of x_2 as small light green dot (\bullet).



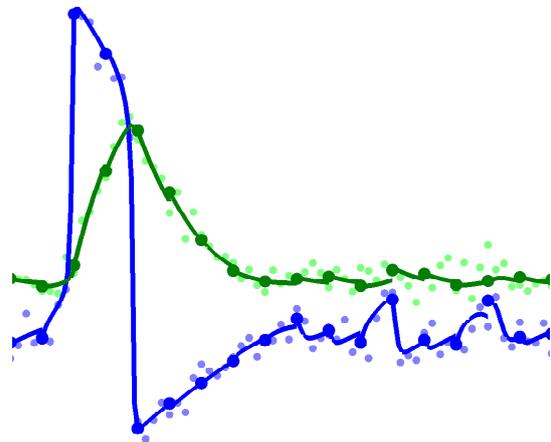
Scenario (C): Exact measurements, partial observation, jump weight 1.0



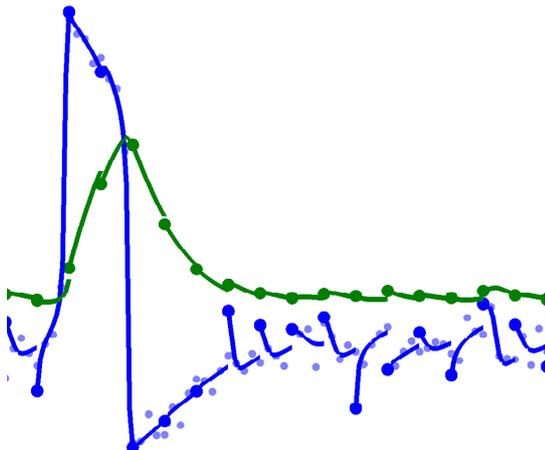
Scenario (D): Exact measurements, partial observation, jump weight 10.0



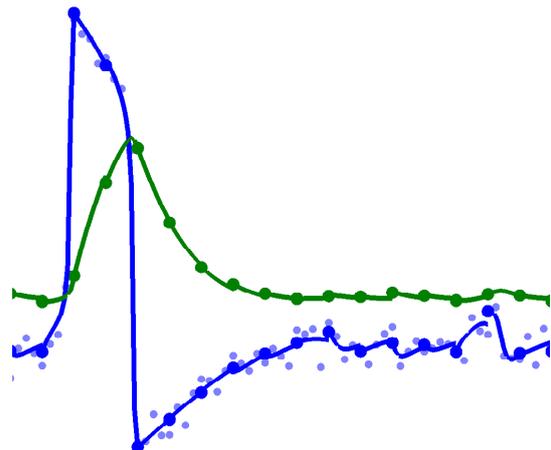
Scenario (E): Noisy measurements, full observation, jump weight 1.0



Scenario (F): Noisy measurements, full observation, jump weight 10.0



Scenario (G): Noisy measurements, partial observation, jump weight 1.0



Scenario (H): Noisy measurements, partial observation, jump weight 10.0

300 400 500 600

300 400 500 600

Figure 6.5 (continued): Detail enlargement of the time interval $[280, 620]$ of the respective scenarios.

6.2 Calcium ion oscillations in eukariotic cells

Cells of animals and plants use changes in free cytoplasmic calcium concentration for information processing and gene expression regulation. These changes in free calcium concentration may be initiated by diverse kinds of external stimuli and intracellular conditions. Oscillations of free calcium concentration have first been described by WOODS et al. [Woods1986] three decades ago. The finding that the qualitative type of oscillation (“bursts” or “periodic spikes”) depends on the stimulation agonist has led to extensive studies of this cellular information processing system.

KUMMER et al. developed a compact deterministic ODE model that is able to simulate both types of oscillation by changing the kinetic parameters of the system [Kummer2000]. In a subsequent paper [Kummer2005], they describe the “transition from stochastic to deterministic behavior” depending on the overall count of particles involved.

The general scheme of calcium signalling in eucariotic cells is the following: Binding of an agonist to a cellular surface receptor triggers the activation of a specific subunit (the α subunit) of a certain type of receptor-bound intracellular *G-protein*⁽⁶⁾, which in turn activates PLC⁽⁷⁾. This induces hydrolysis of the membrane lipid PIP₂⁽⁸⁾ finally releasing IP₃⁽⁹⁾. IP₃ may now bind to membrane receptors of the endoplasmic reticulum which triggers the opening of calcium channels and releasing high amounts of calcium (Ca) from intracellular stores into the cytoplasm. By a mechanism called *calcium-induced calcium release*, this first release of calcium is intensified and propagates the calcium signal.

In her 2005 paper, KUMMER compares the deterministic interpretation of the ODE system eq. (6.5) to a stochastic interpretation based on GILLESPIE simulations [Gillespie1976]. She shows that the two interpretations qualitatively diverge for low particle numbers. This makes the calcium oscillator an interesting study object for the new parameter estimation method. Parameter estimation on this model has also been studied by ZIMMER [Zimmer2015].

For a detailed discussion of the biological background, the derivation of this model, and the meaning of its kinetic parameters, we refer to the original papers of KUMMER et al. and the references therein.

We summarize the “core model” and investigate the applicability of our new parameter estimation method. The model consists of only three species:

- 1) G_α , the concentration of active subunits of the G-protein
- 2) PLC, the concentration of active phospholipase C
- 3) Ca, the amount of free calcium in the cytoplasm

⁽⁶⁾ *G-proteins* or *guanine nucleotide-binding proteins* denote a family of intracellular proteins that are ubiquitously coupled to cell surface or transmembrane receptors, playing a crucial role in signal transduction

⁽⁷⁾ *phosphoinositid-phospholipase C* is a family of signal transducers, transmitting an external stimulus from the membrane located G-proteins into the interior of the cell

⁽⁸⁾ *phosphatidylinositol 4,5-bisphosphate*, a family of phospholipids located mostly at the interior side of the cell membrane

⁽⁹⁾ *inositol 1,4,5-trisphosphate*, a secondary messenger molecule that leaves the membrane and diffuses through the cell, transducing signals

whose interactions are described by the ODE system

$$\begin{aligned}\frac{dG_\alpha}{dt} &= p_1 + p_2 \cdot G_\alpha - \frac{p_3 \cdot G_\alpha \cdot \text{PLC}}{p_4 + G_\alpha} - \frac{p_5 \cdot G_\alpha \cdot \text{Ca}}{p_6 + G_\alpha} \\ \frac{d\text{PLC}}{dt} &= p_7 \cdot G_\alpha - \frac{p_8 \cdot \text{PLC}}{p_9 + \text{PLC}} \\ \frac{d\text{Ca}}{dt} &= p_{10} \cdot G_\alpha - \frac{p_{11} \cdot \text{Ca}}{p_{12} + \text{Ca}}\end{aligned}\tag{6.5}$$

As parameter values, we choose the following values for the kinetic parameters (values from [Kummer2000], figure 7, with an adapted value for p_2 ensuring a limit cycle (regular oscillations) in the deterministic interpretation, see figure 6.6a on the next page; original value: $p_2 = 2.9259$)

$$\begin{aligned}p_1 &= 0.212 & p_5 &= 4.88 & p_9 &= 29.09 \\ p_2 &= 2.95 & p_6 &= 1.18 & p_{10} &= 13.58 \\ p_3 &= 1.52 & p_7 &= 1.24 & p_{11} &= 153.0 \\ p_4 &= 0.19 & p_8 &= 32.24 & p_{12} &= 0.16\end{aligned}\tag{6.6}$$

and initial concentrations

$$G_\alpha(0) = \text{PLC}(0) = \text{Ca}(0) = 0.01 .\tag{6.7}$$

6.2.1 SDE approximation and simulation

Exact stochastic simulations of the system eq. (6.5) would require time consuming GILLESPIE simulations. Here, we use an SDE approximation for data generation. For a detailed discussion of the applicability of this technique, we refer to [Gillespie2000] and the references therein (especially to the work of KURTZ in the 1970s regarding the underlying *chemical LANGEVIN equation*).

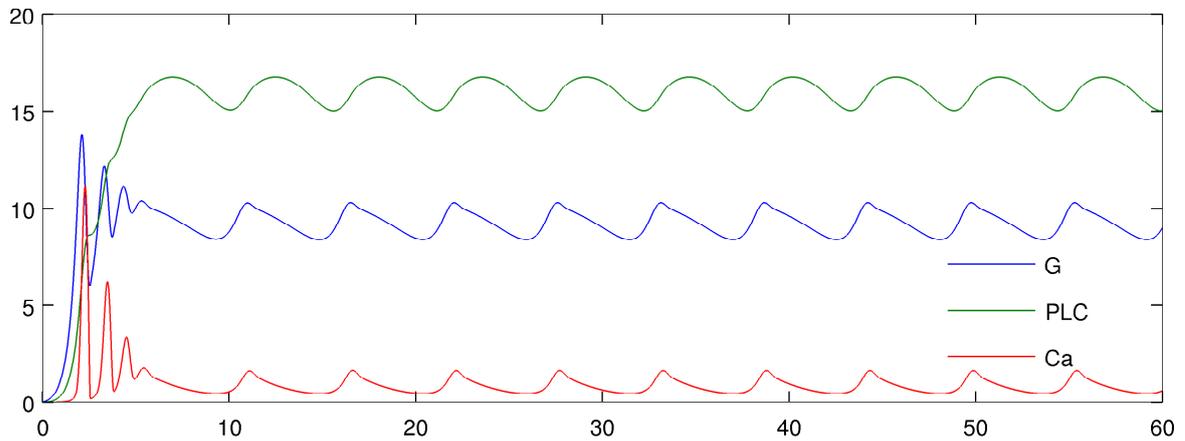
To generate artificial data, we formulate the SDE model with drift function f as described by the ODE model, and constant diffusion function g acting solely on the PLC component with intensity $D = 0.2$ and a (formally) three dimensional driving WIENER process W_t :

$$dX_t = f(X_t, p) dt + g(X_t, p) dW_t\tag{6.8}$$

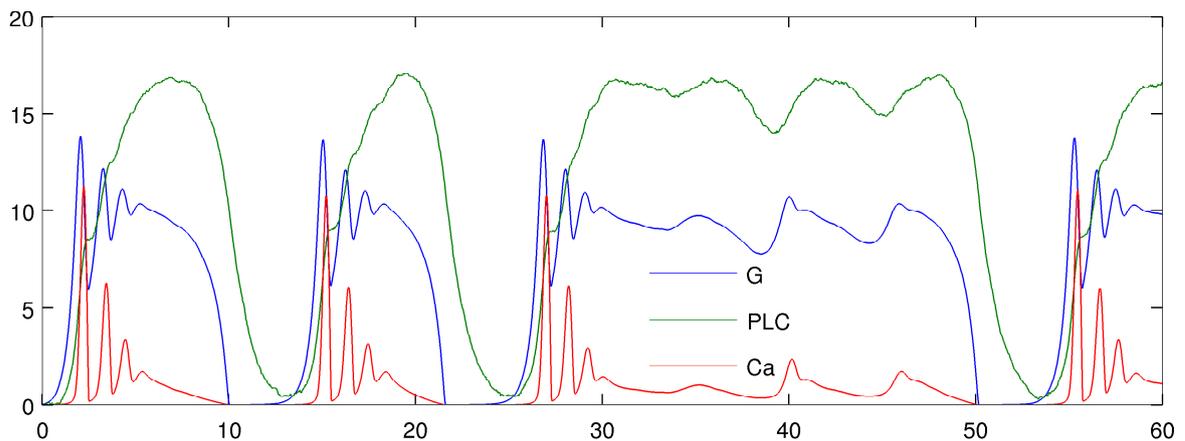
with f as in eq. (6.5), and $g(X_t, p) \equiv \begin{pmatrix} 0 & 0 & 0 \\ 0 & D & 0 \\ 0 & 0 & 0 \end{pmatrix}$, $D = 0.2$

where $p = (p_1, p_2, \dots, p_{12})^T$ denotes the vector of unknown parameters and initial concentrations as specified in eq. (6.7).

Figure 6.6 compares a deterministic interpretation (i.e., with $D = 0$) with a certain stochastic realization of the calcium oscillator SDE. While the deterministic system, after an initial burst, enters a stable limit cycle, the stochastic analogon shows repeated spiking of calcium concentration and partially prolonged phases of low calcium levels.



(a) deterministic (ODE) interpretation of the same calcium oscillator model



(b) stochastic (SDE) interpretation of the calcium oscillator model

Figure 6.6: Simulation of the calcium oscillator with a driving WIENER process effective on the PLC component. The ODE interpretation (which is equivalent to an SDE interpretation with zero diffusion), enters a stable limit cycle after an initial burst.

Small disturbances by the driving WIENER process in the SDE interpretation lead to a qualitatively different behaviour. Kinetic parameters are given in eq. (6.6), diffusion parameter $D = 0.2$; initial values as in eq. (6.7).

Both simulations are generated by an EULER scheme with stepsize 10^{-3} .

6.2.2 (Un-)Identifiability of parameters in partial observations

The model (6.5) as developed by KUMMER is identifiable if full state measurements are available. However, the (formal) model species G_α and PLC combine several individual players of the cells and are thus not accessible in biological experiments. The only measurable species is Ca, the concentration of free calcium in the cytosol.

Unfortunately, this partial observability leads to structural unidentifiability of the model, as the following calculation shows. Changing the amount of G_α and/or PLC by a constant factor may always be compensated by a different choice of parameters.

We change the amount of active G-protein by a factor of g to $\tilde{G}_\alpha := g \cdot G_\alpha$ and the concentration of phospholipase C by a factor of k to $\tilde{\text{PLC}} := k \cdot \text{PLC}$. With these new species, we get:

$$\begin{aligned} \frac{d\tilde{G}_\alpha}{dt} &= g \frac{dG_\alpha}{dt} = g \cdot \left(p_1 + p_2 \frac{\tilde{G}_\alpha}{g} \right) - g \cdot \frac{p_3 \cdot \frac{\tilde{G}_\alpha}{g} \cdot \frac{\tilde{\text{PLC}}}{k}}{p_4 + \frac{\tilde{G}_\alpha}{g}} - g \cdot \frac{p_5 \cdot \frac{\tilde{G}_\alpha}{g} \cdot \text{Ca}}{p_6 + \frac{\tilde{G}_\alpha}{g}} \\ &= gp_1 + p_2 \cdot \tilde{G}_\alpha - \frac{\frac{g}{k} p_3 \cdot \tilde{G}_\alpha \cdot \tilde{\text{PLC}}}{gp_4 + \tilde{G}_\alpha} - \frac{gp_5 \cdot \tilde{G}_\alpha \cdot \text{Ca}}{gp_6 + \tilde{G}_\alpha} \\ \frac{d\tilde{\text{PLC}}}{dt} &= k \frac{d\text{PLC}}{dt} = kp_7 \cdot \frac{\tilde{G}_\alpha}{g} - k \cdot \frac{p_8 \cdot \frac{\tilde{\text{PLC}}}{k}}{p_9 + \frac{\tilde{\text{PLC}}}{k}} \\ &= \frac{k}{g} p_7 \cdot \tilde{G}_\alpha - \frac{kp_8 \cdot \tilde{\text{PLC}}}{kp_9 + \tilde{\text{PLC}}} \\ \frac{d\text{Ca}}{dt} &= \frac{1}{g} p_{10} \cdot \tilde{G}_\alpha - \frac{p_{11} \cdot \text{Ca}}{p_{12} + \text{Ca}} \end{aligned}$$

Thus, choosing the scaled parameters

$$\begin{aligned} \tilde{p}_1 &:= g \cdot p_1, & \tilde{p}_2 &:= p_2, & \tilde{p}_3 &:= \frac{g}{k} \cdot p_3, & \tilde{p}_4 &:= g \cdot p_4, \\ \tilde{p}_5 &:= g \cdot p_5, & \tilde{p}_6 &:= g \cdot p_6, & \tilde{p}_7 &:= \frac{k}{g} \cdot p_7, & \tilde{p}_8 &:= k \cdot p_8, \\ \tilde{p}_9 &:= k \cdot p_9, & \tilde{p}_{10} &:= g^{-1} \cdot p_{10}, & \tilde{p}_{11} &:= p_{11}, & \tilde{p}_{12} &:= p_{12} \end{aligned}$$

we get a structurally identical model as eq. (6.5). This means, we can compensate for higher and lower levels of G_α and PLC by adjusting the kinetic parameters accordingly while *maintaining* the level of Ca.

As consequence, the direct test for structural identifiability from lemma 2.4 fails, showing that the system is **structurally and hence practically not identifiable** from sole Ca measurements (as long as we do not force initial conditions, which are unknown here anyways, or fix some parameters).

We note, that this simple transformation is not necessarily the only one; there might exist other, more complicated choices of adjusted parameter values that can compensate for different levels of G_α and PLC.

6.2.3 Simulation study and a closer look at reconstructing intermittent observations

To investigate the performance of the parameter estimation technique developed in section 5.4 on the calcium oscillator problem, we make a study in which we estimate parameters in 100 distinct realizations. Since the complete set of parameters is structurally not identifiable for partial (Ca only) observations as we have shown in section 6.2.2, we study the following four full state observation scenarios for the full set of unknown parameters:

- (A) exact measurements, full state observation, jump weight 1.0
- (B) exact measurements, full state observation, jump weight 10.0
- (C) noisy measurements, full state observation, jump weight 1.0
- (D) noisy measurements, full state observation, jump weight 10.0

Further, we manually identified a subset of parameters that can be estimated from sole calcium measurements and try to estimate this subset in the scenarios

- (E) exact measurements, partial observation, jump weight 1.0, selected parameter set
- (F) exact measurements, partial observation, jump weight 10.0, selected parameter set
- (G) noisy measurements, partial observation, jump weight 1.0, selected parameter set
- (H) noisy measurements, partial observation, jump weight 10.0, selected parameter set

For each setting, 100 realizations⁽¹⁰⁾ of the stochastic calcium model driven by a WIENER process acting on the PLC component with diffusion coefficient $D = 0.2$ and kinetic parameter values given in eq. (6.6) were made and the generated artificial observations were used to estimate the parameters.

6.2.3.1 Measurement functions, weights, and noise

All settings share the same sample interval of 0.1 time units, resulting in 601 full or partial state observations. In the case of partial observation, only the species Ca (free calcium in the cytosol) is measured. That is, we have as measurement functions⁽¹¹⁾:

$$\begin{aligned} h_i(x(t_i)) &= x(t_i) \in \mathbb{R}^3 \quad (i = 1, \dots, 601) \quad \text{for scenarios (A) to (D)} \\ h_i(x(t_i)) &= x_3(t_i) \in \mathbb{R}^1 \quad (i = 1, \dots, 601) \quad \text{for scenarios (E) to (H)} \end{aligned}$$

with an equidistant measurement grid $T^M = \{0, 0.1, 0.2, 0.3, \dots, 60\}$. Note that we have omitted writing the dependence on the parameter vector p , and set $x(t) := (G_\alpha(t), \text{PLC}(t), \text{Ca}(t))^T$.

Artificial measurement noise and measurement weights

In the “noisy” scenarios, a *relative* measurement error is added, since the concentration of each species varies in several orders of magnitude. To generate the noisy data, every evaluation of the measurement function (applied to a simulation with true parameters) is disturbed by a random value chosen from a normal distribution with zero mean and a variance adjusted to 1% of the measurement value (vector valued measurements are processed componentwise).

This procedure results in a standard deviation of each (scalar component of the) measurement of 10%; a quite high disturbance. The variances are stored and used in the weighted least squares part of the objective as measurement weights. In the undisturbed settings, the measurement weights are chosen as 1.0.

⁽¹⁰⁾ generated by a stochastic EULER scheme with step size 10^{-3}

⁽¹¹⁾ We use vector-valued measurement functions solely for the ease of notation. See the respective footnote (4) on page 169 for the FITZHUGH-NAGUMO measurement functions.

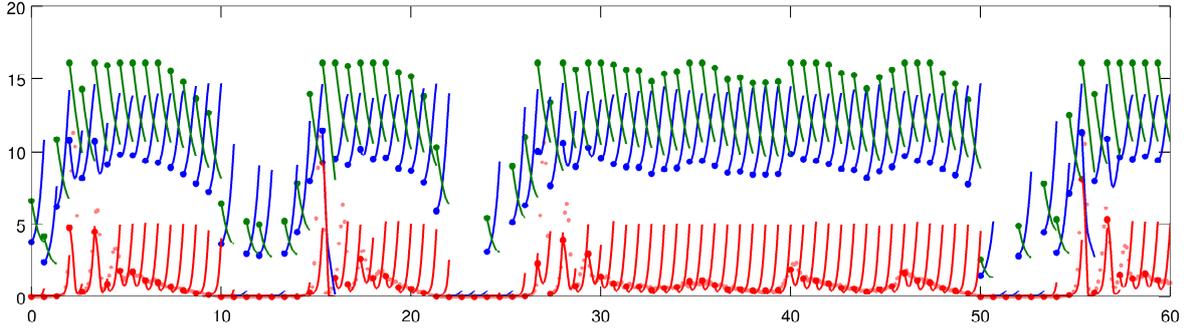


Figure 6.7: Initial trajectory in parameter estimation scenario (G) for a certain realization.

Initial trajectory after shooting node initialization as described in section 6.2.3.2. The applied initialization method brings the node values for the unobserved species in a “physiologically realistic” range (compare the solution trajectory of this problem as given in figure 6.9). Trajectory of component G_α as blue line (—), of component PLC as green line (—), of component Ca as red line (—). Shooting nodes as dots in respective colors (\bullet , \bullet , \bullet). Measurements of Ca as small light red dot (\bullet).

6.2.3.2 Shooting grid and node initialization

The interval $[0, 60]$ is divided into 90 evenly sized intervals, such that

$$T^{\text{MS}} = \{t_0^{\text{MS}}, \dots, t_{90}^{\text{MS}}\} \quad \text{with} \quad t_0^{\text{MS}} = 0, \quad t_{90}^{\text{MS}} = 60, \quad t_k^{\text{MS}} = \frac{2}{3}k \quad (k = 1, \dots, 89)$$

and the shooting node variables s_k ($k = 0, \dots, 89$) are initialized by measurement data.

In the full state observation scenarios (A),(B),(C),(D), the node values are initialized by the temporally most proximate measurement data, or 0 (zero), if the measurement value is negative (which might be due to the applied measurement noise).

In the partial observation scenarios (E),(F),(G),(H), node values for the component Ca are initialized from measurement data in the same way as above. The unobserved species are initialized by an empirically derived inverse approximation:

Let c denote the measurement value of the observed Ca species. Then, initialize

$$\begin{aligned} \text{Ca} & \text{ as } \max\{0, c\} \\ G_\alpha & \text{ as } \max\{0, \log(10000 \cdot \text{Ca})\} \\ \text{PLC} & \text{ as } 1.75 \cdot \max\{0, \log(10000 \cdot \min\{1, \text{Ca}\})\} \end{aligned}$$

where we define $\log(0) := -\infty$. Figure 6.7 shows an example trajectory using this initialization.

6.2.3.3 Initial parameter guess and stopping criterion for the GAUSS-NEWTON solver

The initial guess of the parameters is set to 50% of the true values (see table 6.2), ensuring that the local area of contraction of the GAUSS-NEWTON method is left, thus globalization takes effect. We remark that also for more distant as well as randomized initial guesses, convergence to the solution is observed.

The optimization is stopped when the maximum-norm of the search direction $\|\Delta x_k\|_\infty$ (see section 1.3 on page 14) falls below $5 \cdot 10^{-3}$.

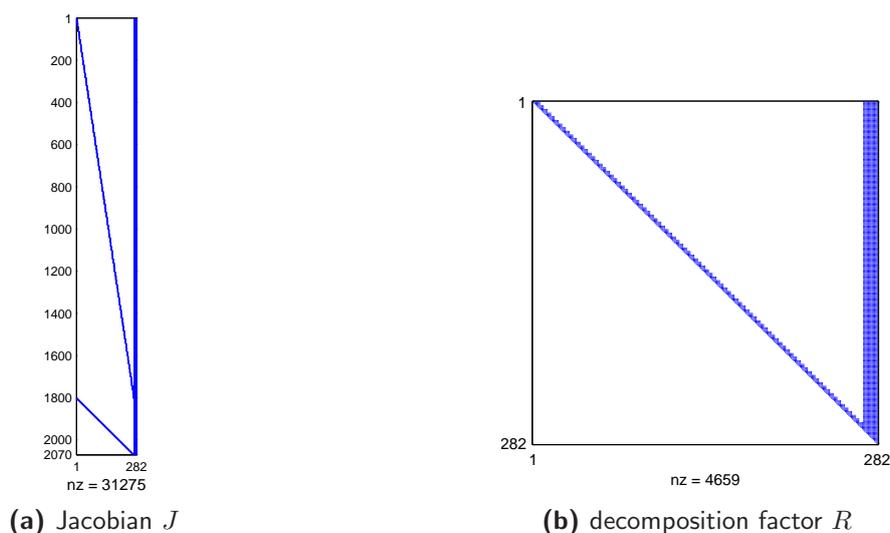


Figure 6.8: Sparsity pattern of the combined residual vector's Jacobian - Calcium oscillator with full state measurements.

(a) Jacobian J with dimension 2070×282 (583740 elements), nonzero elements: 31275 (5.4%)

(b) decomposition factor R with dimension 282×282 (79524 elements), nonzero elements: 4659 (5.9%)

See section 5.5.1 for details on the sparsity pattern.

6.2.3.4 Constraints on optimization variables

To avoid negative protein concentrations and unphysiological states during solution, the following constraints on the *state variables* at the shooting nodes are set:

$$G_\alpha, \text{PLC}, \text{Ca} \in [0, 50].$$

To avoid backward-occurring reactions and keeping the parameters in physiological ranges, the following constraints on the *kinetic parameters* are set:

$$\begin{aligned} p_1 &\in [0.01, 1.0] & (\text{true value: } 0.212), & p_7 &\in [0.1, 10.0] & (\text{true value: } 1.24), \\ p_2 &\in [0.1, 10.0] & (\text{true value: } 2.95), & p_8 &\in [1.0, 100.0] & (\text{true value: } 32.24), \\ p_3 &\in [0.1, 10.0] & (\text{true value: } 1.52), & p_9 &\in [1.0, 100.0] & (\text{true value: } 29.09), \\ p_4 &\in [0.01, 1.0] & (\text{true value: } 0.19), & p_{10} &\in [1.0, 50.0] & (\text{true value: } 13.58), \\ p_5 &\in [0.1, 10.0] & (\text{true value: } 4.88), & p_{11} &\in [10.0, 500.0] & (\text{true value: } 153.0), \\ p_6 &\in [0.1, 10.0] & (\text{true value: } 1.18), & p_{12} &\in [0.01, 10.0] & (\text{true value: } 0.16). \end{aligned}$$

In the test series, no parameter constraints are active in the solutions. As one would expect, some state variables attain the lower bound 0 (zero), see figure 6.9.

6.2.3.5 Sparsity pattern of the combined residual vector's Jacobian

Figure 6.8 shows the sparsity pattern of the combined residual vector's Jacobian J , i.e. the system matrix of the linearized problem without constraints, as well as the sparsity pattern of its decomposition factor that may be used for solving the linearized problem.

With a total dimension of 2070×282 , only 31275 (5.4%) out of 583740 elements are nonzero. The decomposition factor $R \in \mathbb{R}^{282 \times 282}$ has roughly the same low occupancy rate of 5.9%. See section 5.5.1 for details.

Table 6.2: Estimation test series: Calcium oscillator (full observations). Results of parameter estimation on 100 independent simulations (in each setting) of the calcium oscillator, whose PLC component is affected by a WIENER process, eq. (6.8). For jump weight 1.0, the initial guess for the kinetic parameters was 50% of the true values. For jump weight 10.0, the results for jump weight 1.0 were used as initial guess. Values rounded to 5 digits.

full observations of G_α , PLC, Ca — exact measurements						
true parameter name value	scenario (A) jump weights 1.0			scenario (B) jump weights 10.0		
	estimate \pm SD	(SD%)	RelErr%	estimate \pm SD	(SD%)	RelErr%
p_1 0.212	0.2167 \pm 0.0274	(12.68%)	2.23%	0.2307 \pm 0.0730	(31.66%)	8.86%
p_2 2.95	2.9527 \pm 0.0078	(0.27%)	0.09%	2.9970 \pm 0.0460	(1.54%)	1.60%
p_3 1.52	1.5206 \pm 0.0050	(0.33%)	0.04%	1.5461 \pm 0.0234	(1.52%)	1.72%
p_4 0.19	0.1845 \pm 0.0166	(9.02%)	2.88%	0.1656 \pm 0.0503	(30.40%)	12.84%
p_5 4.88	4.7982 \pm 0.0417	(0.87%)	1.68%	5.0628 \pm 0.2058	(4.07%)	3.75%
p_6 1.18	0.9788 \pm 0.0834	(8.53%)	17.05%	1.4615 \pm 0.3825	(26.18%)	23.86%
p_7 1.24	1.2383 \pm 0.0105	(0.85%)	0.13%	1.2416 \pm 0.0134	(1.08%)	0.14%
p_8 32.24	31.940 \pm 1.2954	(4.06%)	0.93%	32.138 \pm 1.5105	(4.70%)	0.32%
p_9 29.09	28.744 \pm 1.7553	(6.11%)	1.19%	28.909 \pm 2.0603	(7.13%)	0.62%
p_{10} 13.58	13.559 \pm 0.0269	(0.20%)	0.15%	13.381 \pm 0.1848	(1.38%)	1.46%
p_{11} 153.0	152.72 \pm 0.2953	(0.19%)	0.18%	150.99 \pm 1.8837	(1.25%)	1.31%
p_{12} 0.16	0.1596 \pm 0.0004	(0.28%)	0.23%	0.1591 \pm 0.0056	(3.55%)	0.55%

full observations of G_α , PLC, Ca — noisy measurements ^a						
true parameter name value	scenario (C) jump weights 1.0			scenario (D) jump weights 10.0		
	estimate \pm SD	(SD%)	RelErr%	estimate \pm SD	(SD%)	RelErr%
p_1 0.212	0.1924 \pm 0.0373	(19.40%)	9.22%	0.1945 \pm 0.0332	(17.08%)	8.24%
p_2 2.95	2.9688 \pm 0.0397	(1.34%)	0.64%	2.9743 \pm 0.0340	(1.14%)	0.82%
p_3 1.52	1.5317 \pm 0.0235	(1.54%)	0.77%	1.5347 \pm 0.0193	(1.26%)	0.97%
p_4 0.19	0.2037 \pm 0.0337	(16.55%)	7.22%	0.2002 \pm 0.0281	(14.06%)	5.42%
p_5 4.88	4.9647 \pm 0.3462	(6.97%)	1.74%	4.8907 \pm 0.2221	(4.54%)	0.22%
p_6 1.18	1.3008 \pm 0.7602	(58.45%)	10.24%	1.1303 \pm 0.4123	(36.48%)	4.21%
p_7 1.24	1.2296 \pm 0.0353	(2.87%)	0.83%	1.2360 \pm 0.0139	(1.13%)	0.32%
p_8 32.24	31.686 \pm 3.4231	(10.80%)	1.72%	32.084 \pm 1.4260	(4.44%)	0.48%
p_9 29.09	28.712 \pm 4.5299	(15.78%)	1.30%	28.993 \pm 1.9850	(6.85%)	0.33%
p_{10} 13.58	13.425 \pm 0.2658	(1.98%)	1.14%	13.491 \pm 0.2490	(1.85%)	0.65%
p_{11} 153.0	151.32 \pm 2.8687	(1.90%)	1.10%	152.16 \pm 2.6209	(1.72%)	0.54%
p_{12} 0.16	0.1596 \pm 0.0033	(2.12%)	0.24%	0.1608 \pm 0.0030	(1.90%)	0.53%

SD%: standard deviation of the estimate · RelErr%: relative deviation of the estimate from the true parameter value

^anormally distributed measurement noise with zero mean and standard deviation of 10% of the respective measurement value (multiplicative noise), see section 6.2.3.1

Table 6.3: Estimation test series: Calcium oscillator (partial observations). Results of parameter estimation on 100 independent simulations (in each setting) of the calcium oscillator, whose PLC component is affected by a WIENER process, eq. (6.8). As the full set of parameters is unidentifiable when using partial observations of Ca only, an identifiable subset of parameters was estimated and the remaining were kept fixed to their true values. For jump weight 1.0, the initial guess for the kinetic parameters was 50% of the true values. For jump weight 10.0, the results for jump weight 1.0 were used as initial guess. Values rounded to 5 digits.

		partial observations of Ca only			—	exact measurements		
true parameter name value	scenario (E) jump weights 1.0				scenario (F) jump weights 10.0			
	estimate ± SD	(SD%)	RelErr%		estimate ± SD	(SD%)	RelErr%	
p_7 1.24	1.2328 ± 0.0534	(4.34%)	0.58%		1.2712 ± 0.0638	(5.02%)	2.52%	
p_9 29.09	29.445 ± 2.5795	(8.76%)	1.22%		27.989 ± 2.5813	(9.22%)	3.78%	
p_{10} 13.58	13.507 ± 0.0364	(0.27%)	0.53%		13.476 ± 0.1205	(0.89%)	0.76%	
p_{11} 153.0	152.22 ± 3.6164	(2.38%)	0.51%		153.83 ± 4.3694	(2.84%)	0.54%	
p_{12} 0.16	0.1585 ± 0.0036	(2.30%)	0.88%		0.1577 ± 0.0041	(2.64%)	1.39%	

		partial observations of Ca only			—	noisy measurements ^a		
true parameter name value	scenario (G) jump weights 1.0				scenario (H) jump weights 10.0			
	estimate ± SD	(SD%)	RelErr%		estimate ± SD	(SD%)	RelErr%	
p_7 1.24	1.1462 ± 0.1089	(9.50%)	7.56%		1.2273 ± 0.0332	(2.71%)	1.02%	
p_9 29.09	33.917 ± 5.4398	(16.04%)	16.60%		29.640 ± 1.3839	(4.67%)	1.89%	
p_{10} 13.58	13.511 ± 0.2851	(2.11%)	0.50%		13.495 ± 0.0777	(0.58%)	0.62%	
p_{11} 153.0	146.48 ± 7.3676	(5.03%)	4.26%		151.58 ± 2.5347	(1.67%)	0.92%	
p_{12} 0.16	0.1588 ± 0.0081	(5.10%)	0.71%		0.1586 ± 0.0029	(1.83%)	0.83%	

SD%: standard deviation of the estimate · RelErr%: relative deviation of the estimate from the true parameter value

^anormally distributed measurement noise with zero mean and standard deviation of 10% of the respective measurement value (multiplicative noise), see section 6.2.3.1

6.2.3.6 Results of the test series

Tables 6.2 and 6.3 give the parameter estimation results of the test series for full observations in scenarios (A)–(D) and for partial observations scenarios (E)–(H), respectively. For all tested scenarios, the overall performance of the proposed method is very satisfying.

In the full state observation scenario (A) with exact measurement data and jump regularization weights of 1.0, the mean parameter estimates are very accurate: with the exception of parameter p_6 , they have less than 3% error, and also show a small standard deviation in the set of 100 realizations. The stronger jump regularization in scenario (B) leads to less accurate but still satisfactory estimates with less than 2% error in 8 out of 12 parameters; again parameter p_6 shows an enlarged error. Thus, for exact measurements, the increased jump weights are less beneficial.

In contrast to that, the estimates in the noisy full state observation scenarios significantly benefit from stronger jump regularization. 9 out of 12 parameter estimates have a relative error of less than 1% in the strong jump regularization scenario **(D)**, with a maximum error of about 8% in the remaining three parameters. Also, the estimates show fairly small variations in the test series, again with exception of parameter p_6 whose standard deviation lies around 36%.

Estimates for noisy measurements clearly benefit from higher jump weights

As table 6.3 shows, stronger jump regularization leads to better estimates in the noisy scenarios (compare **(G)** to **(H)**). Better in two ways: the mean estimates are much closer to the true parameter values (less than 2% error in contrast to a maximum relative error of 17% for weaker jump regularization), and also the standard deviation of the estimates are considerably reduced (below 5% in contrast to about 16%).

However, for undisturbed observations **(E)** and **(F)**, this does, again, not hold: while the standard deviation of the estimates are approximately identical, the relative errors of the mean estimates are considerably increased, but still remain below 4%.

Very similar findings are discussed in the BISTABAER model in section 6.3 on page 187.

6.2.3.7 A closer look at a single realization

In figure 6.9, fitted trajectories of the calcium oscillator model to measurements of the scenarios **(A)**, **(C)**, **(E)**, **(G)** (all with jump regularization weights 1.0).

Base of all measurements is a certain realization of the calcium oscillator; therefore the overall kinetics is identical in all fits. However, the employed measurement data differs: full state or partial observations, exact or disturbed measurements (see the captions in figure 6.9). Comparing the trajectories, we see that they differ mainly in the jumps at the nodes:

- With exact measurement data, scenarios **(A)** and **(E)**, the trajectory resembles a continuous one, the jumps at the shooting nodes small; too small to see at this level of detail. Note that the original trajectory is indeed continuous, as the driving WIENER process is continuous.
- In the fits for disturbed measurement data, scenarios **(C)** and **(G)**, we observe bigger jumps, as the optimization technique uses these additional degrees of freedom to cope for the measurement noise.

If one requires for any reason a “more continuous” trajectory, increasing the jump regularization would be an effective strategy. We refer to the FITZHUGH-NAGUMO example in section 6.1.4.7, especially to the detail enlargements in figure 6.5 on page 173, where this effect is depicted.

6.2.3.8 Recovering trajectories in intermittent observations

As a final example with respect to the calcium oscillation model, we investigate two test cases with intermittent full state observations: once with exact measurements and once with noisy measurements. As in the previous example, the measurement noise follows the statements given in section 6.2.3.1, i.e. the a noise level of 10% of the measurement values.

In contrast to the previously discussed test series with equidistant measurements throughout the whole time domain $[0, 60]$, the measurements are sampled intermittently. Starting at time point 0, data acquisition occurs every 0.125 time units, running for 5 time units. After a pause of 6 time units, data is sampled again for 5 units at the same time increments of 0.125, and so on, resulting in 246 measurement time points in total. In figure 6.10, the data

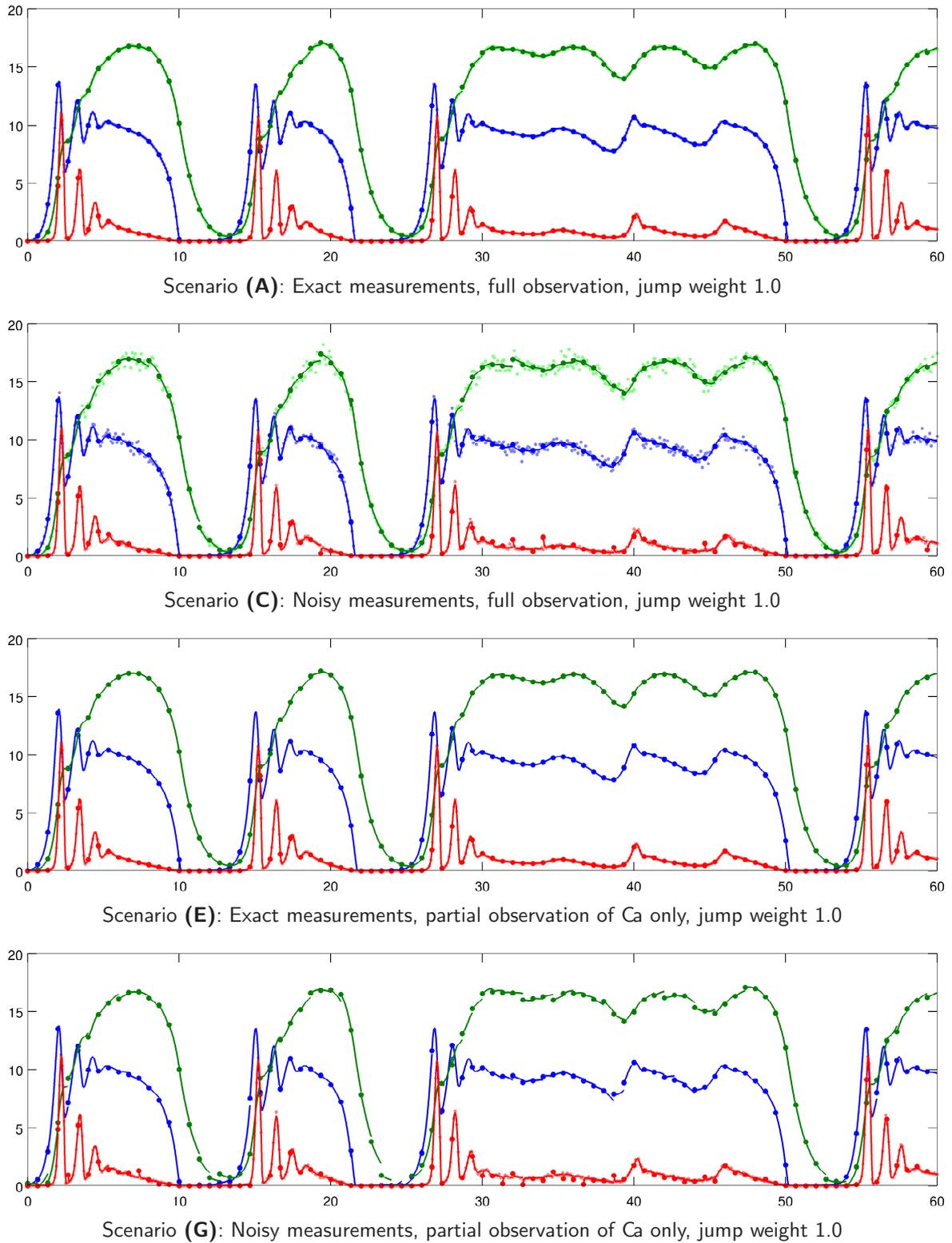


Figure 6.9: Calcium oscillator example fit. Fitted trajectories of the four scenarios with jump weight 1.0 listed in section 6.2.3 for a certain realization of the calcium oscillator.

Fitted trajectory of component G_α as blue line (—), of component PLC as green line (—), of component Ca as red line (—). Shooting nodes as dots in the respective colors (\bullet , \bullet , \bullet). Measurements of G_α as small light blue dot (\bullet), of PLC as small light green dot (\bullet), of Ca as small light red dot (\bullet).

acquisition intervals are marked in yellow color. Figures 6.10a and 6.10b depict the collected measurement data.

Having in mind the findings from the calcium test series in section 6.2.3.6, the jump weights are chosen as 1.0 in the exact measurement setting, and as 10.0 in the noisy data setting. The time domain $[0, 60]$ is divided into 120 intervals of the same size; node initialization is done exactly as described in section 6.2.3.2.

Figures 6.10c and 6.10c show the trajectories of the fits using exact and noisy measurements. For comparison, the original realization is given in figure 6.10e. In both settings, the original trajectory is surprisingly well recovered – also in those areas, where no measurement data was available (white background).

Trajectory and parameter reconstruction

As shown in table 6.4, the parameter values are recovered within the accuracy described in the test series results of section 6.2.3.6.

The well reproduction of the mostly unobserved calcium oscillations in time span $[38, 44]$ may be attributed to the fact that it is already initiated at the end of measurement interval #4. Quite impressive is the reproduction of the “bump” around $t = 53$, which lies in an completely unobserved interval (compare again the samples in figures 6.10a and 6.10b). Surely, without any data from the adjacent measurement intervals, such a close to perfect reconstruction would be improbable. However, since there is no direct data available in the unobserved time spans, the **trajectory reconstruction of this WIENER-driven stochastic calcium oscillator** can be attributed mainly to the jump regularization technique proposed in this thesis.

Table 6.4: Estimation results with intermittent observations. Results of parameter estimation on a certain realization of the calcium oscillator, whose PLC component is affected by a WIENER process, eq. (6.8). The initial guess for the kinetic parameters was 50% of the true values. Values rounded to 4 digits after the decimal point.

test case with intermittent data acquisition				
true parameter name value	exact measurements jump weights 1.0		noisy measurements jump weights 10.0	
	estimate	RelErr%	estimate	RelErr%
p_1 0.212	0.2583	(21.84%)	0.2261	(6.67%)
p_2 2.95	2.9455	(0.15%)	2.9961	(1.56%)
p_3 1.52	1.5148	(0.34%)	1.5421	(1.45%)
p_4 0.19	0.1559	(17.95%)	0.1840	(3.14%)
p_5 4.88	4.8006	(1.63%)	5.0503	(3.49%)
p_6 1.18	0.9962	(15.57%)	1.3707	(16.16%)
p_7 1.24	1.2326	(0.59%)	1.2036	(2.94%)
p_8 32.24	32.7563	(1.60%)	33.2263	(3.06%)
p_9 29.09	30.1799	(3.75%)	32.0433	(10.15%)
p_{10} 13.58	13.5736	(0.05%)	13.3017	(2.05%)
p_{11} 153.0	152.9277	(0.05%)	150.4137	(1.69%)
p_{12} 0.16	0.1600	(0.02%)	0.1628	(1.73%)

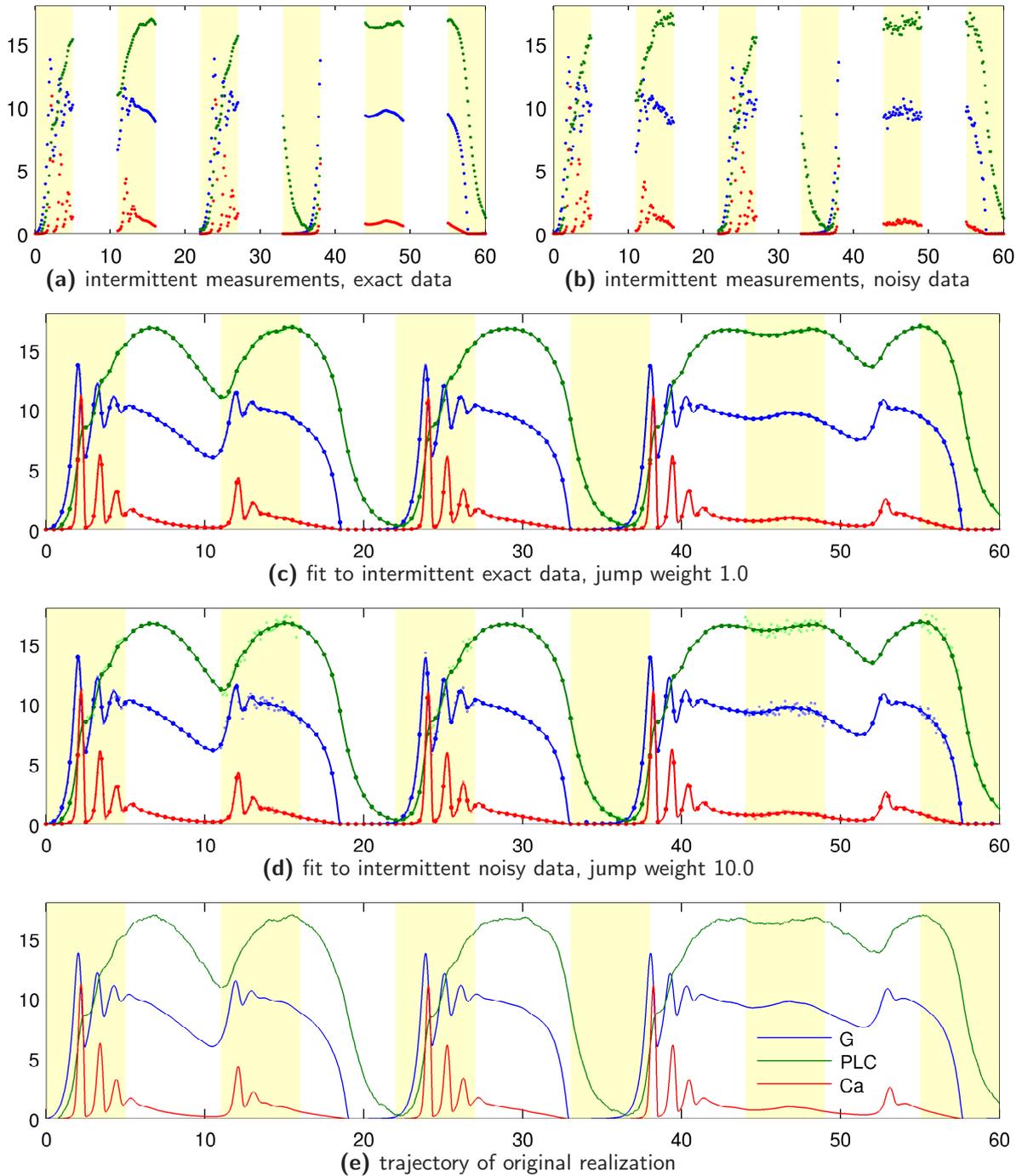
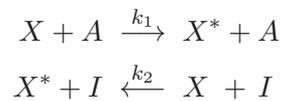
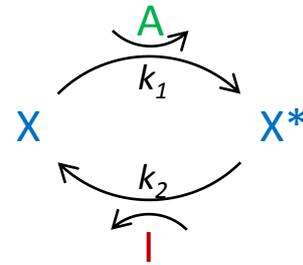


Figure 6.10: Calcium oscillator fit to intermittent data. Fitted trajectories based on intermittent full state observations of a certain realization of the calcium oscillator. See section 6.2.3.8 for a discussion. Trajectory of component G_α as blue line (—), of component PLC as green line (—), of component Ca as red line (—). Shooting nodes as dots in the respective colors (\bullet , \bullet , \bullet). Measurements of G_α as small light blue dot (\bullet), of PLC as small light green dot (\bullet), of Ca as small light red dot (\bullet). The yellow rectangles mark the intervals where measurement data has been sampled.

6.3 The BISTABAER enzyme kinetics model

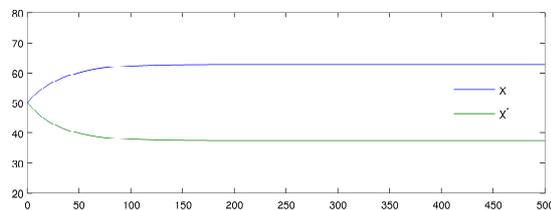
We come back to the introductory example in figure 5.1 on page 122, reprinted in figure 6.11. The underlying equations describe an allosteric enzyme regulation model whose activator efficacy depends on an intrinsic control that may vary between two locally stable steady states (resulting in "high" and "low" activator efficacy).

The enzyme X may exist in two conformations: an *inactive form* X and an *activated form* X^* . The transition between these two states is mediated by two proteins acting as *allosteric regulators*: the *activator* A and the *inhibitor* I , which are both present in constant concentration. This may be depicted as in the reaction graph to the right.

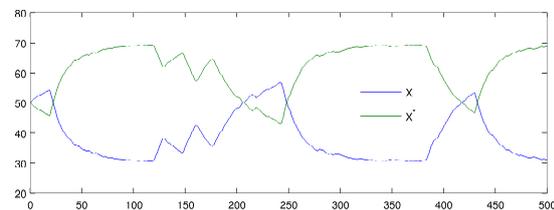


The activator's efficacy depends on an intrinsic control L ; the effective reaction speed of the activation step is modelled as $k_1^{eff} := k_1 \cdot (A + L)$. The internal control L is described by the bistable differential equation $\frac{dL}{dt} = L(s^2 - L^4)$ with an unstable steady state $L = 0$ and two (real) locally stable states at $L = \pm\sqrt{s}$ (for $s > 0$; in the example, we choose $s := 1$).

In the deterministic interpretation, the control L of the allosteric regulator protein approaches its steady state concentration very fast (or stays on its unstable steady state if no perturbation occurs), driving also the BISTABAER model into a steady state as depicted in figure 6.11a. Introducing stochasticity in the system in terms of a driving WIENER process acting on the control L leads to a fundamentally changed behaviour of the system (figure 6.11b).



(a) deterministic (ODE) interpretation



(b) stochastic (SDE) interpretation

Figure 6.11: Stochasticity in the BISTABAER model completely alters its behaviour. The ODE system (a) quickly approaches the steady state, whereas in the SDE interpretation (b) transitions are observed.

6.3.1 Model equations and parameters

The underlying model equations, reduced to a minimal set, read as:

$$\begin{aligned} \frac{dX}{dt} &= -k_1^{eff} X A + k_2 X^* I \\ \frac{dX^*}{dt} &= k_1^{eff} X A - k_2 X^* I \\ \frac{dL}{dt} &= L(1 - L^4) \end{aligned} \quad k_1^{eff} := k_1 \cdot (A + L) \quad (6.9)$$

with initial conditions, time domain, and system constants

$$\begin{aligned} X(t_0) = 50, \quad X^*(t_0) = 50, \quad L(t_0) = -1.0 \\ [t_0, t_f] = [0, 500], \quad A := 1.7, \quad I := 5 \end{aligned} \quad (6.10)$$

and kinetic parameters

$$k_1 := 0.01, \quad k_2 := 0.004 \quad (6.11)$$

used for the simulations in figure 6.11 and the estimation test series in section 6.3.3.

6.3.2 SDE model

As depicted in the introductory text, a WIENER process acts on the hidden internal control L . To generate artificial data, we formulate the SDE model with drift function f as described by the ODE model, and constant diffusion function g acting solely on the internal control L with intensity $D = 0.5$ and a (formally) three dimensional driving WIENER process W_t :

$$dZ_t = f(Z_t, p) dt + g(Z_t, p) dW_t \quad (6.12)$$

where $Z_t = (X(t), X^*(t), L(t))^T$ is the state vector and $p = (k_1, k_2)^T$ denotes the vector of unknown kinetic parameters. The drift function f is defined by the ODE model in eq. (6.9), and the diffusion function g is defined as

$$g(Z_t, p) := \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & D \end{pmatrix}, \quad \text{with } D = 0.5.$$

Initial concentrations are as specified in eq. (6.10).

6.3.3 Simulation study and a closer look at grid refinement

To investigate the performance of the parameter estimation technique developed in section 5.4 on the BISTABAER problem, we study the following two scenarios for estimating the unknown kinetic parameters k_1 and k_2 :

- (A) exact measurements, partial state observation of X and X^*
- (B) noisy measurements, partial state observation of X and X^*

The jump weights are all set to 1.0.

For each setting, 100 realizations⁽¹²⁾ of the allosteric regulation model eq. (6.12) driven by a WIENER process acting on the hidden internal control L with diffusion coefficient $D = 0.5$, and kinetic parameter values given in eq. (6.11) were made and the generated artificial observations were used to estimate the parameters.

6.3.3.1 Measurement functions, weights, and noise

Measurements are available for the inactive and active form of the protein, i.e. we can measure the model species X and X^* . Dropping the dependence on the parameters, the measurement

⁽¹²⁾generated by a stochastic EULER scheme with step size 10^{-2}

functions⁽¹³⁾ thus are:

$$h_i(x(t_i)) = (x_1(t_i), x_2(t_i))^T \in \mathbb{R}^2 \quad (i = 1, \dots, 1001)$$

where the state of the system eq. (6.12) is combined in the vector $x(t) = (X(t), X^*(t), L(t))^T$.

In both scenarios, samples are taken every 0.5 time units, resulting in an overall amount of 1001 measurements of X and X^* . In the “noisy” scenarios, normally distributed noise with zero mean and a standard deviation of 2.5 is added to the measurements, leading to quite high disturbances (see figure 6.13b). Measurement weights are chosen as 1.0 in the undisturbed scenario; in the “noisy” scenario, the reciprocal of the above standard deviation is used.

6.3.3.2 Shooting grid and node initialization

The time domain $[0, 500]$ is (naïvely, see the following remark) equidistantly partitioned into 100 shooting intervals, thus the shooting grid is given as

$$T^{\text{MS}} = \{5(i-1) : i = 1, \dots, 100\}.$$

The shooting node values for X and X^* are initialized with the temporally most proximate measurement values; the component for the internal control L is always initialized with -1.0 .

A remark on the naïve choice of the shooting grid

Having in mind the (hidden) bistable internal control L , the above choice is clearly not optimal. Ideally, one would start a new shooting interval whenever the internal control L swaps and approaches a different steady state⁽¹⁴⁾. However, this BISTABAER example is ideal to investigate the **robustness** of the proposed parameter estimation method for a non-optimal choice of the shooting grid (and, besides, the required information is not directly available from the measurements). We will see how this affects the estimation in the discussion in sections 6.3.3.6 and 6.3.3.7.

6.3.3.3 Initial parameter guess and stopping criterion for the GAUSS-NEWTON solver

The initial guess of the parameters is set to 50% of the true values (see table 6.5), ensuring that the local area of contraction of the GAUSS-NEWTON method is left, thus globalization takes effect. We remark that also for more distant as well as randomized initial guesses, convergence to the solution is observed.

The optimization is stopped when the maximum-norm of the search direction $\|\Delta x_k\|_\infty$ (see section 1.3 on page 14) falls below $5 \cdot 10^{-2}$.

6.3.3.4 Constraints on optimization variables

To avoid negative protein concentrations and unphysiological states during solution, we set some constraints on the state variables at the shooting nodes. Further, we employ some restrictions on the admissible space for the kinetic parameters to avoid backward-occurring reactions and keeping their values in physiological ranges:

$$X, X^* \in [0, 2000], \quad L \in [-2, 2], \quad k_1, k_2 \in [10^{-9}, 10^{-1}].$$

⁽¹³⁾ We use vector-valued measurement functions solely for the ease of notation. See the respective footnote (4) on page 169 for the FITZHUGH-NAGUMO measurement functions.

⁽¹⁴⁾ Due to the driving WIENER process W_t , the internal control L will vary around a steady state value for small disturbances (increments) of W_t , while a sufficiently large excitation might push it towards the other locally stable steady state’s attractive region

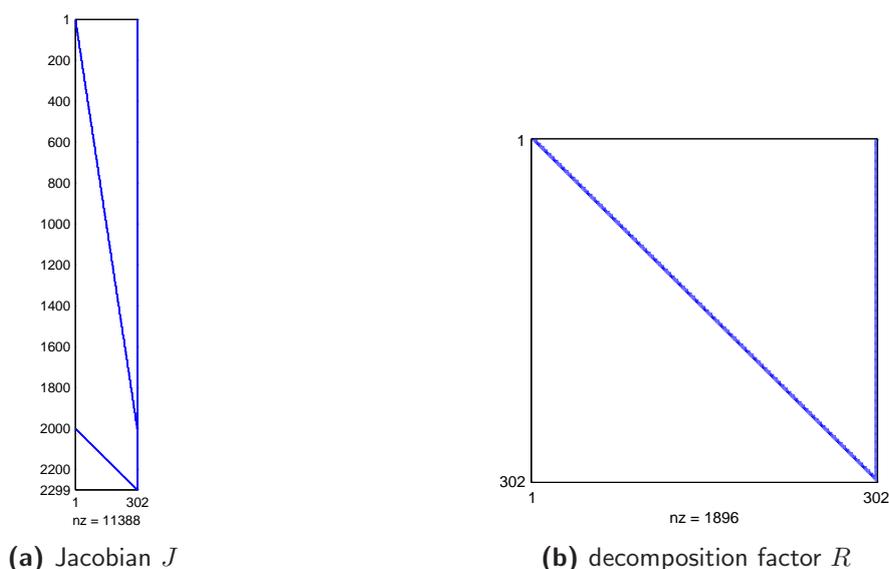


Figure 6.12: Sparsity pattern of the combined residual vector's Jacobian - BISTABAER model .
 (a) Jacobian J with dimension 2299×302 (694298 elements), nonzero elements: 11388 (1.6%)
 (b) decomposition factor R with dimension 302×302 (91204 elements), nonzero elements: 1896 (2.1%)
 See section 5.5.1 for details on the sparsity pattern.

In the test series, no parameter constraints are active in the solutions. In some rare cases⁽¹⁵⁾, the optimization variable for the state of the internal control L attains it's bound.

6.3.3.5 Sparsity pattern of the combined residual vector's Jacobian

Figure 6.12 shows the sparsity pattern of the combined residual vector's Jacobian J , i.e. the system matrix of the linearized problem without constraints, as well as the sparsity pattern of its decomposition factor that may be used for solving the linearized problem.

With a total dimension of 2299×302 , only 11388 (1.6%) out of 694298 elements are nonzero. The decomposition factor $R \in \mathbb{R}^{302 \times 302}$ has roughly the same low occupancy rate of 2.1%. See section 5.5.1 for details.

6.3.3.6 Results of the test series

Table 6.5 shows the results of the test series. In both settings, the true values of $k_1 = 0.01$ and $k_2 = 0.004$ of the kinetic parameters are reproduced: with exact observations to $k_1 = 0.00910$ and $k_2 = 0.00362$, with noisy observations to $k_1 = 0.01003$ and $k_2 = 0.00396$. In both settings, the standard deviation of the estimated parameter values does not exceed 10%.

At first glance, it might be surprising that the estimation for noisy observation is much better than for exact estimations (less than 1% error in the noisy scenario in contrast to 9% in the exact scenario). There are (at least) two reasons: First, this is due to the *ad hoc* choice of the shooting grid that collides with the stochastic changes of the internal control L , which is

⁽¹⁵⁾ This is only observed in the first shooting interval. As there occurs no jump regularization at $t = 0$ (it's the *first* interval), the solver might choose a large absolute value for L to reduce the residual in the measurements of X and X^* . Due to the high attractiveness of the steady states of L , this choice does not induce an increased jump size at the subsequent shooting node.

switching between the two attractive regions of its locally stable steady states (see the remark in section 6.3.3.2 and the discussion in section 6.3.3.7). Second, there is no clear choice of measurement weights when the error variance is zero as it is for exact measurements.

One can see from the detail enlargements in figure 6.14a, that for exact measurements, the improper choice of the shooting grid leads to increased residuals when the internal control L switches towards a different steady state in the interior of a shooting interval. While, in principle, this effect is also present in the noisy setting, it is strongly dominated by the measurement noise (figure 6.14b).

As a remark: When being in the lap of luxury of having exact measurement values, it would be easy to generate a perfect shooting grid for this BISTABAER example by ensuring that shooting nodes are placed whenever the measurements of X and X^* change their direction, as this marks the swap towards a different steady state in the hidden control L . However, even the *ad hoc* choice of the equidistant shooting grid delivers reasonable approximations on the kinetic parameters.

6.3.3.7 A closer look at a single realization

Figure 6.13 shows the fitted trajectories for a certain realization of the stochastic BISTABAER system eq. (6.12) for the two test scenarios described in section 6.3.3.

From visual inspection, the fit on the species X and X^* looks satisfying; though a more careful look reveals some relatively large deviations, e.g. at the shooting node at time $t = 240$ in figure 6.13a. These misfits are found also, but weaker, on several other places where the

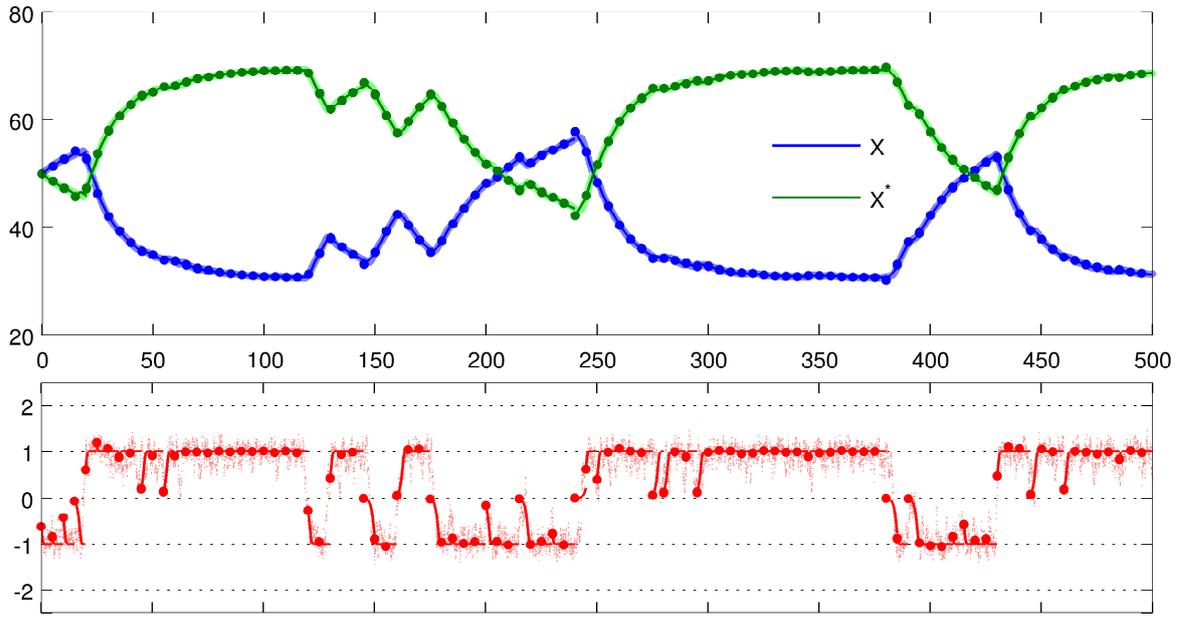
Table 6.5: Estimation test series: BISTABAER allosteric regulation model. Results of parameter estimation on 100 independent simulations (in each setting) of the BISTABAER model, whose internal control L is affected by a WIENER process, eq. (6.12). Initial guess for the kinetic parameters was 50% of the true values. Values rounded to 4 digits.

exact observations without measurement error				
true parameter		scenario (A)		
name value		observation of X and X^*		
		estimate \pm SD	(SD%)	RelErr%
k_1	0.01	0.00910 \pm 0.00095	(10.40%)	8.99%
k_2	0.004	0.00362 \pm 0.00017	(4.67%)	9.46%

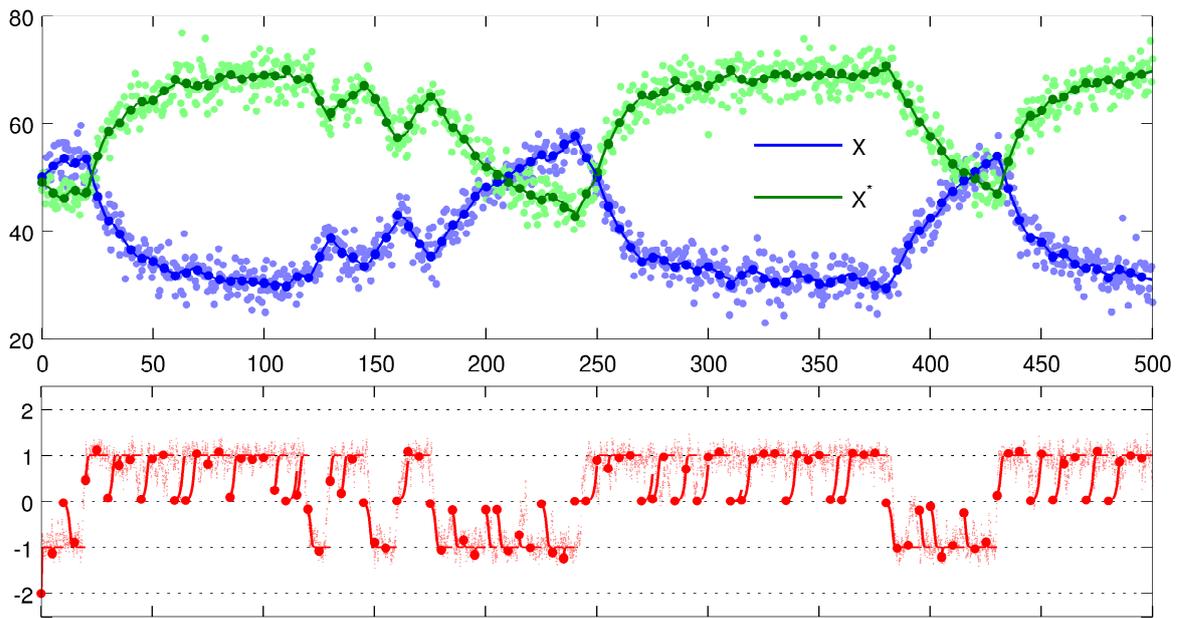
noisy observations with error $\epsilon \sim \mathcal{N}(0, 2.5^2)$				
true parameter		scenario (B)		
name value		observation of X and X^*		
		estimate \pm SD	(SD%)	RelErr%
k_1	0.01	0.01003 \pm 0.00079	(7.86%)	0.28%
k_2	0.004	0.00396 \pm 0.00031	(7.89%)	1.01%

SD%: standard deviation of the estimate

RelErr%: relative deviation of the estimate from the true parameter value



(a) Fit of species X and X^* to exact data in scenario (A) (top), recovered intrinsic control L (bottom)



(b) Fit of species X and X^* to noisy data in scenario (B) (top), recovered intrinsic control L (bottom)

Figure 6.13: Example fit of the BISTABAER model with exact and noisy measurements, together with reconstructed intrinsic control L . See section 6.3.3.7 for a discussion, and figure 6.14a for detail enlargements.

Fitted trajectory of X as blue line (—), of X^* as green line (—), of L as red line (—). Shooting nodes as dots in the respective colors (\bullet , \bullet , \bullet). Measurements of X as small light blue dots (\circ), of X^* as small light green dots (\circ), of L as small light red dots (\circ).

Note that no samples of L were used while fitting.

trajectory shows a kink, e.g. in the time range [110, 155]. The detail enlargement of this time span in figure 6.14a discloses that one reason is the *ad hoc* choice of the shooting grid.

On these intervals, the deterministic approximation may not cope with a switching internal control, if the transition occurs in the interior of a shooting interval. This becomes obvious if we focus on the three marked shooting intervals in figure 6.14a. For minimizing the objective in problem 5.18, an “intermediate” trajectory is chosen (more precise: parameters that generate this trajectory), minimizing both the residuals and the jumps at the shooting nodes.

Optimizing the shooting grid

In this example with dense and error-free measurements, the shooting grid might be adjusted after a first solution has been computed, or even during the solution process. Unsuitable shooting intervals may be detected by two markers: (1) an enlarged residual compared to “low residual” intervals, and (2) larger jumps at the beginning or end. Already a bisection of the respective intervals would strongly improve the fitting capabilities and may be easily automated. Surely, this requires an appropriate adaptation of the jump regularization weights of the divided interval. Without using further knowledge about the underlying mechanics, a natural (not necessarily optimal) adjustment of the jump weights for a centrally divided interval would be a factor of 1/2.

Indeed, adding 10 nodes at the intervals with highest residuals reduces the error in the estimate of k_1 by 11.0% and of k_2 by 15.9% (see table 6.6)

Reconstruction of the hidden intrinsic control L

As the bottom parts of figures 6.13a and 6.13b show, the hidden bistable control L , which is continuously disturbed by the driving WIENER process W_t is nicely reproduced.

Since there is no measurement data of L available, the only direct contribution of L to the objective function occurs via the jump regularization term in the objective.

An interesting observation: In scenario **(A)** of exact measurements, the recovered intrinsic control L lies rather stable around its steady states, whereas the in the noisy scenario **(B)**, it shows higher fluctuations. This is an artefact of the used measurement weights: The data is accurate, so their variance is 0, implying positive infinite weights. For standard least squares parameter estimation in ODE, one could effectively use an arbitrary but common weighting factor for the residuals. Actually, in this example, the residual weights were chosen as 1.0, effectively assuming a variance of 1.0 of the measurements. As a consequence, the jump regularization has an increased impact on the estimation. In practical settings, exact measurements do not occur, thus using the reciprocals of the (then available) measurement error’s variances in the weighted least squares part of the objective in problem 5.18 will avoid this hassle completely.

Nevertheless, the above observation gives hints for improving the estimation in the exact scenario: It might be beneficial for this example to use lower (or zero) regularization weights for jumps in the hidden control L (equivalent effects are obtained by increasing the residual weights (i.e. lowering their variance values). This is indeed successful, as a recalculation shows: For this specific example, setting the regularization weights for jumps in L to zero leads to an error reduction in the estimates of k_1 and k_2 by 21.6% and 10.0%, respectively (see table 6.6).

The doubling of shooting nodes combined with disabled jump regularization reduces the error by even 71.1% and 62.4%, at the cost of increased computational effort.

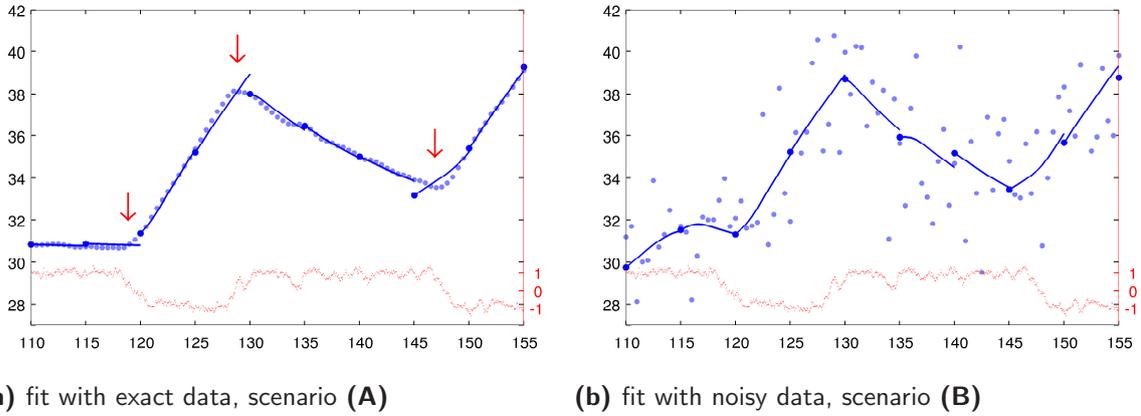


Figure 6.14: Detail enlargement of a BISTABAER fit of the time interval $[110, 155]$. The *ad hoc* choice of the shooting grid (a) is not adapted to the fluctuations of the intrinsic control L (red dots), leading to enlarged residuals and jumps in the respective intervals. In the noisy scenario (b), the measurement noise dominates. Both figures originate from the same BISTABAER realization. Visible here are the fitted trajectory of X as blue line (—) with respective shooting nodes as blue dots (•) and measurements of X as light blue dots (◐). The samples of the intrinsic control L are given for comparison as small light red dots (◐), with scale on the right axis.

Table 6.6: Improved estimates by adjusting grid and jump regularization weights. The table shows the effects of grid and jump weight adjustments discussed in section 6.3.3.7 for the BISTABAER realization depicted in figure 6.13a. The *error reduction* is given relative to the naïve grid with 101 equidistant nodes and regularization weights of 1.0 for jumps in the L component. The parameters' true values are $k_1 = 0.01$, $k_2 = 0.004$. Values rounded to 3 digits.

setting		estimate		error reduction [†]	
grid,	L jump weights	k_1	k_2	k_1	k_2
naïve ^a	1.0	0.00902	0.00365	—	—
adjusted ^b	1.0	0.00913	0.00370	11.0%	15.9%
doubled ^c	1.0	0.00912	0.00370	9.5%	16.2%
naïve	0.0	0.00924	0.00368	21.6%	10.0%
adjusted	0.0	0.00935	0.00376	33.2%	32.3%
doubled	0.0	0.00972	0.00386	71.1%	62.4%

[†]compared to default setting (a) with naïve grid and L jump weights 1.0

^anaïve grid by dividing the time horizon in 100 intervals of the same size

^bmanually adjusted grid from (a), with additional grid points at timepoints 18.51, 118.1, 128.1, 147.1, 218.1, 243.1, 378.1, 383.1, 413.1, 432.1

^cnaïve grid by dividing the time horizon in 200 intervals of the same size

6.4 The ORNSTEIN-UHLENBECK model (driven by a LÉVY process)

Though the proposed parameter estimation method is developed for recovering parameters in the drift term of the underlying SDE, under certain conditions, it is also capable to recover parameters in the diffusion term. In this example, besides some drift parameters, it successfully recovers the diffusion constant by an analysis of the jump distribution.

The classical ORNSTEIN-UHLENBECK⁽¹⁵⁾ process (OU process) [UhlenbeckOrnstein1930] is a stationary GAUSSIAN process, driven by a WIENER process, and defined as the solution of the stochastic differential equation

$$dX_t = \theta(\mu - X_t) dt + D dW_t$$

where μ is the *mean reversion level*, θ the *mean reversion rate*, D the *volatility*, and dW_t a standard WIENER process. In financial applications, the OU process is known as VAŠÍČEK model [Vasicek1977] for time evolution of interest rates.

A more general class of stochastic processes is the one of ORNSTEIN-UHLENBECK type processes, in which the driving WIENER process of the classical OU process is substituted by a LÉVY process (see section 4.5 on page 118) and that may be written as

$$dX_t = \theta(\mu - X_t) dt + D dZ_t$$

where now Z_t is the driving (LÉVY-)process.

ORNSTEIN-UHLENBECK processes and processes of ORNSTEIN-UHLENBECK type have a wide range of application in physics (e.g. the HOOKEan spring), biology (molecular biology), and especially finance (financial econometrics, [Barndorff2001]), where OU-type processes with non-negative increments from a *background driving LÉVY process* (BDLP) guarantee positive (nominal) interest rates. However, since the European debt crisis, reality also proved negative nominal interest rates plausible.

Figure 6.15a shows a realization of an ORNSTEIN-UHLENBECK process with initial value and parameters

$$X_0 = 12.0, \quad \mu = 16.0, \quad \theta = 2.2, \quad D = 4.5,$$

driven by a compound POISSON process⁽¹⁶⁾ with intensity $\lambda = 7.5$ and standard normally distributed increments.

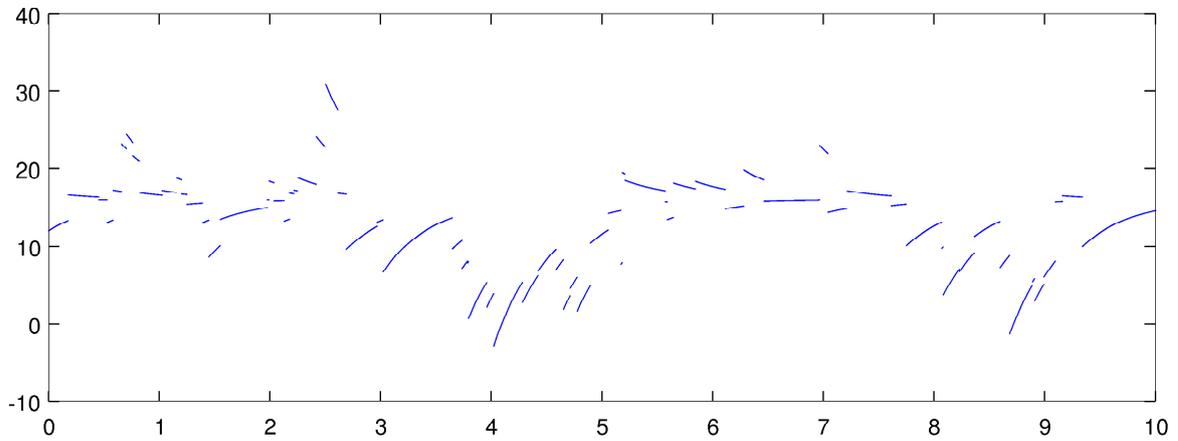
6.4.1 A simple heuristic for placing the shooting nodes

The choice of a suitable shooting grid is crucial for this type of problem. An equidistant grid will most likely lead to wrong results because fixed-sized shooting intervals spanning over one or even multiple jump times of the BDLP will result in large residuals even for true initial states and true parameters. Obviously, the shooting grid should consist of all the jump times of the background driving LÉVY process, since the trajectory between the discrete jumps is continuous.

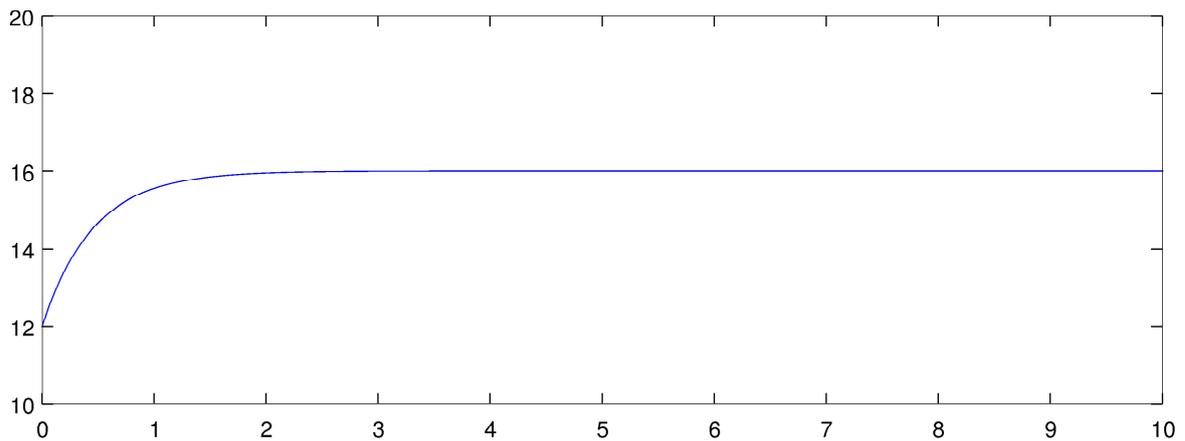
For a densely sampled trajectory without any measurement error, detection of jumps and thus placement of shooting nodes may be done using a simple heuristic: If the value of two successive measurements exceeds a certain limit Δ_{min} (that clearly depends on the jumpsize distribution of the BDLP), a shooting node is placed in between them.

⁽¹⁵⁾ Leonard Salomon ORNSTEIN, 1880–1941, Dutch physicist
George Eugene UHLENBECK, 1900–1988, US-American physicist

⁽¹⁶⁾ A compound POISSON process is a continuous-time LÉVY process whose jump times follow a POISSON process and whose increments follow an arbitrary distribution that is independent of the jump time POISSON process. See section 4.5 on page 118.



(a) stochastic (SDE) interpretation of the ORNSTEIN-UHLENBECK model



(b) deterministic (ODE) interpretation of the same ORNSTEIN-UHLENBECK model

Figure 6.15: Realisation of an ORNSTEIN-UHLENBECK process with a driving compound POISSON process. In the SDE interpretation, the POISSON distributed jumps of the driving compound POISSON process Z_t are clearly visible, arbitrarily pushing the ORNSTEIN-UHLENBECK process away from its mean reversion level $\mu = 16.0$. The process' mean reversion rate $\theta = 2.2$ determines the speed of reaching the mean reversion level; the processes diffusion parameter $D = 4.5$ determines the influence by the driver. The driving compound POISSON process Z_t has intensity 7.5 with standard normally distributed increments.

The ODE interpretation is equivalent to an SDE interpretation with zero diffusion, showing an exponential attraction towards the mean reversion level.

Both simulations are generated by an Euler scheme with stepsize 10^{-3} .

Let η_i be the measurement at time t_i . Then, the shooting grid consists of initial time t_0 , final time point t_f , and the interior time points are chosen by the heuristic

$$\text{if } |\eta_{i+1} - \eta_i| \geq \Delta_{min}, \text{ add } \frac{t_i + t_{i+1}}{2} \text{ to the shooting grid.} \quad (6.13)$$

The above heuristic only depends on the measurements, i.e., when sampled adequately, only on the BDLP, and not on any further knowledge of the process realization. In our example of a BDLP with standard normally distributed jump sizes and a diffusion coefficient of $D = 4.5$, a value of $\Delta_{min} = 0.5$ will detect more than 90% of jumps when error-free measurements are available. We note that the choice of Δ_{min} depends on (1) the jump size distribution of the BDLP and the diffusion coefficient D , (2) the sample interval, (3) the parameters of the system (as, e.g., a higher mean reversion rate θ leads to bigger jumps for fixed sample times), and (4) the noise level of the measurements.

6.4.2 No jump regularization

Since there are actual jumps in the ORNSTEIN-UHLENBECK process, it is not considered constructive to include a jump regularization term in the objective of the optimization problem, but to retain a discontinuous trajectory. Thus, the jump weights in the objective of the parameter estimation problem 5.18 are chosen as zero, $\omega_{k,l}^2 = 0$. In this setting, the parameter estimation problem is structurally identical to the so-called *multi-experiment setting* for which efficient solution methods have been developed by SCHLÖDER [Schloeder1987].

6.4.3 Assessing the diffusion parameter D by jump residuals

Though the estimation technique presented in this thesis only addresses parameters in the drift function, there is in certain cases the possibility to estimate diffusion parameters *en passant*. In this specific example, the stochasticity lies in the driving compound POISSON process, that introduces jumps into the trajectory. Thus, the distribution of these discontinuities delivers an approximation on the diffusion parameter D .

If we knew the times at which the driving process jumps, we would have the perfect grid for placing the shooting nodes. Then, under the assumption of exact and sufficiently many state observations, the fitted trajectory's discontinuities α_k at these nodes are exactly the jumps of the driving compound POISSON process. Thus, although these two assumptions will not hold in general, we might still use information about the distribution of the jumps α_k – more precisely their standard deviation – as an approximation on the diffusion parameter D .

Two remarks on the estimation of the diffusion parameter

(1) The simple heuristic for placing the shooting nodes presented in eq. (6.13) cannot detect jumps of the driving compound POISSON process smaller than the detection limit Δ_{min} . Thus, if we have exact measurements, even in the case of time-continuous observations, the above described procedure for estimating D will deliver an over-estimation (in mean), because small jumps below the threshold are left out.

(2) Application of heuristic (6.13) to a set of noisy observations for generating the shooting grid is expected to deliver as well an over-estimation of D , if the detection threshold Δ_{min} is chosen too large, for the same reason as above. For small values of Δ_{min} , it will deliver an under-estimation of D , as some measurement noise from the tails of the errors' normal distribution will be mistakenly detected as jumps of the driving process.

6.4.4 An example test case

Three estimation settings for a single realization of the compound POISSON-driven ORNSTEIN-UHLENBECK process with initial state and parameters as in eq. (6.13) are depicted in figure 6.16:

- (a) with exact measurement data and a shooting grid established by heuristic (6.13) with $\Delta_{min} = 0.5$
- (b) with an additive normally distributed measurement noise with variance 1.0 and the same shooting grid as in (a) (unrealistic set-up, for comparison)
- (c) with the same noise as in (b) and a shooting grid generated by heuristic (6.13) with $\Delta_{min} = 3.0$.

For the test case (a), as one can see in figure 6.16a, the shooting grid built by the simple heuristic eq. (6.13) delivers a **nearly perfect fit**, and the parameters are recovered when using error-free measurements: $\theta = 2.1902$ (true value: 2.2), $\mu = 15.981$ (true value: 16.0).

For test case (b), figure 6.16b depicts the fit when using measurement data disturbed by an additive normally distributed measurement error with zero mean and variance 1.0. As the same shooting grid as in (a) was used (which is not available in general) the recovered trajectory is close to the original one. Parameters are still well recovered, only slightly worse than in the undisturbed case: $\theta = 2.2537$, $\mu = 15.833$.

Finally, the fit of scenario (c), which is the most realistic scenario, is shown in figure 6.16c. An elevated jump detection level of $\Delta_{min} = 3.0$ for shooting grid generation by heuristic eq. (6.13) is necessary due to the added measurement error. It can be seen that some shooting nodes are newly created, e.g. in the time span $[7, 8]$, while others vanished, as in the interval $[0.2, 1.2]$. The parameters are still nicely recovered: $\theta = 2.2800$, $\mu = 15.621$.

Recovery of the diffusion parameter D

Further, in each of the above settings, also the diffusion parameter D can be recovered. Calculating the standard deviation of the jumps of the respective trajectory (i.e. of the values α_k on the inner shooting nodes, see definition 5.17) delivers a fairly accurate estimate on the true value $D^* = 4.5$ of the diffusion parameter: (a) $D = 4.5896$, (b) $D = 4.6256$, (c) $D = 4.5554$.

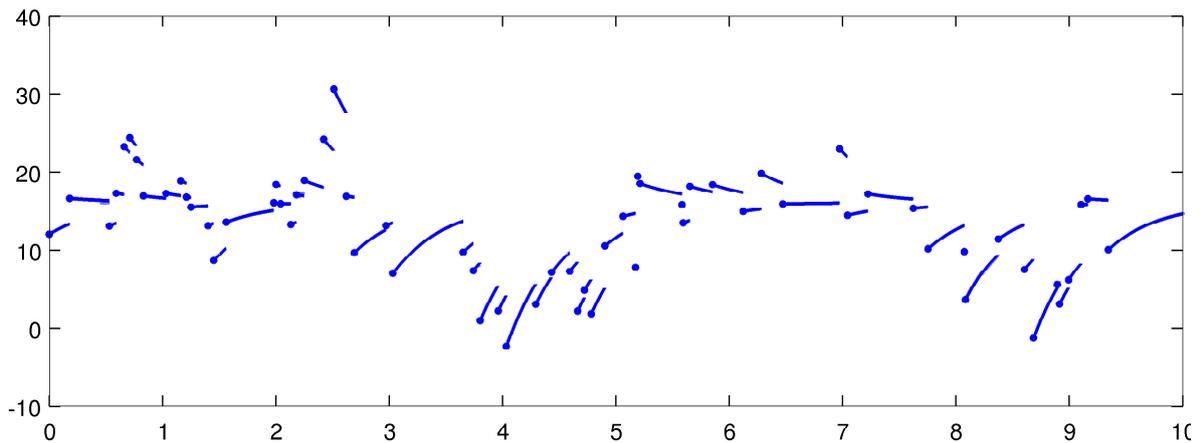
6.4.5 Simulation study

We investigate the performance of the new parameter estimation method using 100 independent realizations⁽¹⁷⁾ of a compound POISSON-driven ORNSTEIN-UHLENBECK process, with the three scenarios of

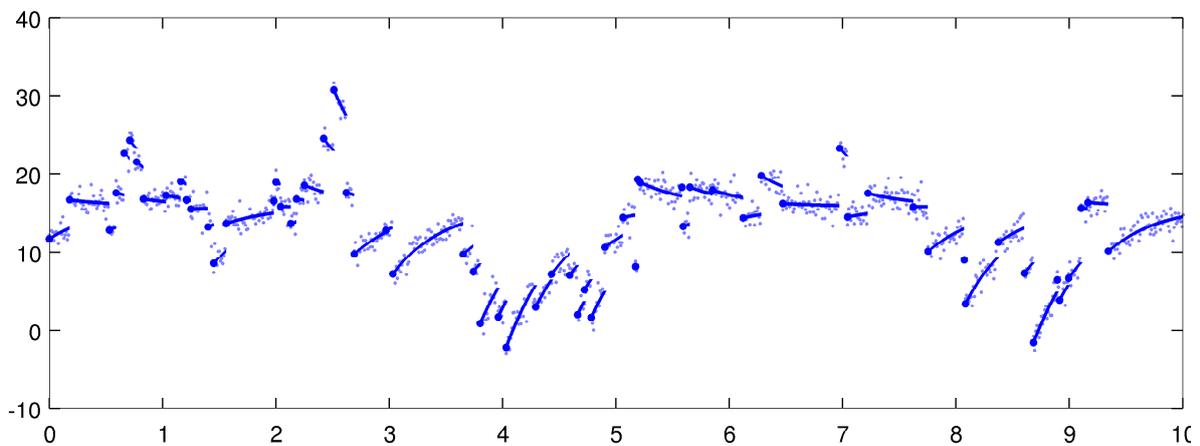
- (A) no measurement error, shooting grid generated by heuristic (6.13) with $\Delta_{min} = 0.5$
- (B) additive normal noise with zero mean and standard deviation 1.0, and same shooting grid as in (A)
- (C) same noise as in (B) but with a new shooting grid generated by heuristic (6.13) with $\Delta_{min} = 3.0$.

No jump regularization is included in the objective as stated earlier, as the driving compound POISSON process is a true jump process.

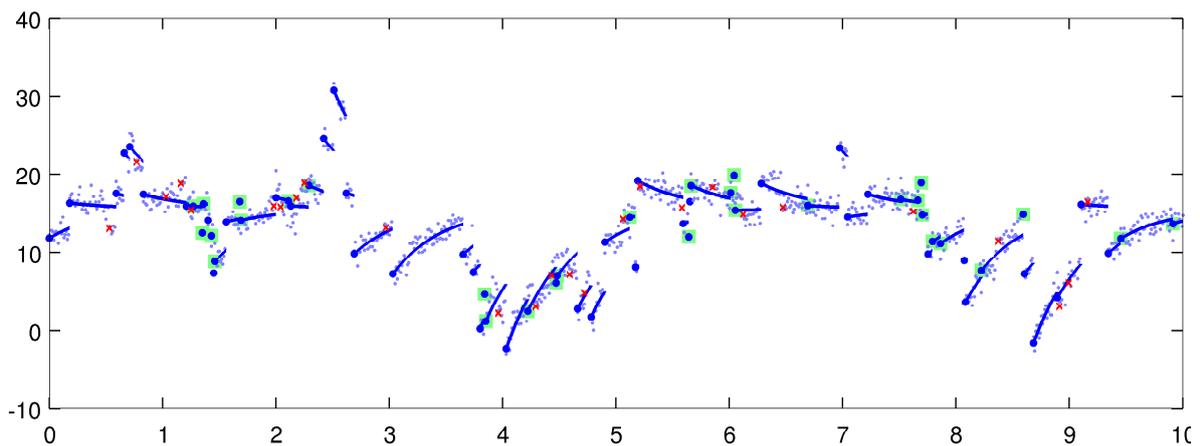
⁽¹⁷⁾ generated by an EULER scheme with step size 10^{-3}



(a) Trajectory of fitted process, error-free measurements, nodes by heuristic (6.13) with $\Delta_{min} = 0.5$



(b) Fit for noisy measurements using the same grid as in (a)



(c) Fit for noisy measurements using a new shooting grid produced by heuristic (6.13) with $\Delta_{min} = 3.0$

Figure 6.16: Example of fitted trajectories of an ORNSTEIN-UHLENBECK process.

In (b), the same shooting grid as in (a) was used, whereas in (c) a new grid was generated using the heuristic (6.13) on the disturbed measurements.

In (c), new grid points are highlighted in green \blacksquare and removed grid points are marked with a red \times .

In all settings, jump regularization weights were chosen as zero, since the driving compound POISSON process has jumps.

6.4.5.1 Measurement functions and weights

The measurement function is the identity function on the single state variable. Measurement weights are set to 1.0 in both cases of exact and noisy measurements, as the variance for each measurement is constant.

6.4.5.2 Shooting grid and node initialization

The shooting grid T^{MS} is built by heuristic (6.13) with Δ_{min} as given in the scenario descriptions. Node values are initialized by the temporally most proximate measurement data.

6.4.5.3 Initial parameter guess and stopping criterion for the GAUSS-NEWTON solver

The initial guess of the parameters is set to 50% of the true values (see table 6.7), ensuring that the local area of contraction of the GAUSS-NEWTON method is left, thus globalization takes effect. We remark that also for more distant as well as randomized initial guesses, convergence to the solution is observed.

The optimization is stopped as soon as the maximum norm of the search direction $\|\Delta x_k\|_\infty$ (see section 1.3 on page 14) falls below 10^{-6} .

6.4.5.4 Constraints on optimization variables

The following constraints on the *state variable* at the shooting nodes and on the *parameters* were used:

$$X_t \in [-1000, 1000], \quad \theta \in [0.1, 10], \quad \mu \in [0.1, 30].$$

In the test series, no constraints are active in the solutions.

6.4.5.5 Sparsity pattern of the combined residual vector's Jacobian

Figure 6.17 shows the sparsity pattern of the combined residual vector's Jacobian J , i.e. the system matrix of the linearized problem without constraints, as well as the sparsity pattern of its decomposition factor that may be used for solving the linearized problem.

With a total dimension of 1001×74 , only 3001 (4.1%) out of 74074 elements are nonzero. The decomposition factor $R \in \mathbb{R}^{74 \times 74}$ has roughly the same low occupancy rate of 4.0%.

See section 5.5.1 for details.

6.4.5.6 Results

Table 6.7 shows the estimation results. The mean estimates of the kinetic parameters μ and θ are very close to the true values in both scenarios **(A)** and **(B)** that share the same shooting grid. When using exact measurement data **(A)**, their standard deviations are small; the artificial measurement noise in **(B)** enlarges this uncertainty slightly.

In scenario **(C)**, where a new shooting grid is built, the mean reversion level μ is still reproduced, but with an standard deviation enlarged to about 10%. The mean reversion rate θ seems to have a systematic error, though the true value lies still within one standard deviation of the estimate. As the noise is exactly the same in **(C)** as in **(B)**, the quite accurate estimation in **(C)** proves that this is due to the very simple heuristic for jump detection, and may be obliterated by better adjusted shooting grids. See the remarks in section 6.4.3.

The estimates of the diffusion constant D depend on the shooting discretization and might also be improved by a more accurate heuristic than the one in eq. (6.13). However, though the diffusion constant is not at all included in the parameter estimation algorithm, it can be approximated fairly well by calculating the mean of the standard deviations of all discontinuities (the α_k) of the individual fitted trajectories.

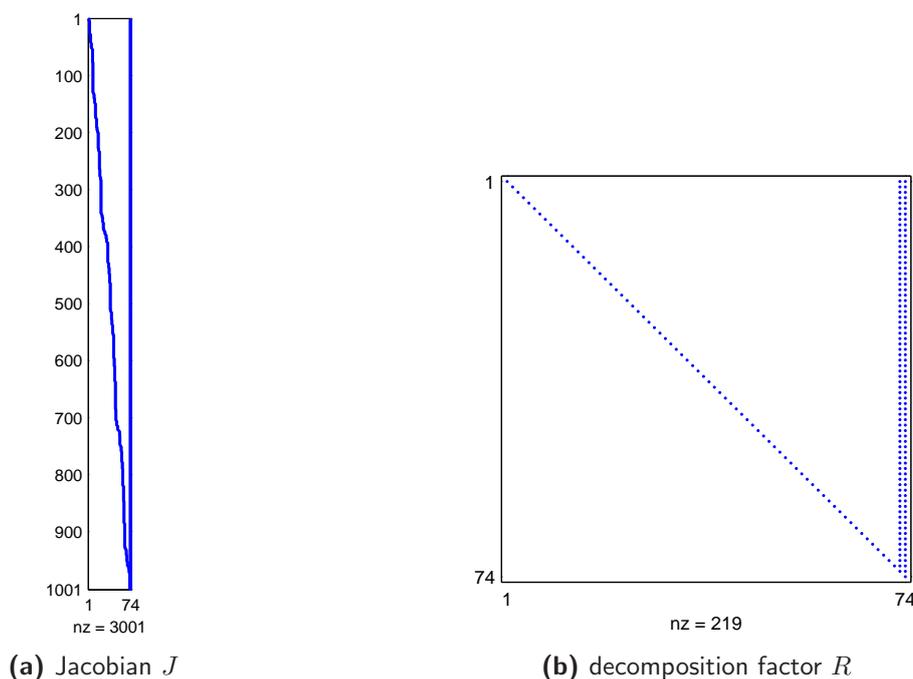


Figure 6.17: Sparsity pattern of the combined residual vector's Jacobian - OU model

(a) Jacobian J with dimension 1001×74 (74074 elements), nonzero elements: 3001 (4.1%)

(b) decomposition factor R with dimension 74×74 (5476 elements), nonzero elements: 219 (4.0%)

See section 5.5.1 for details on the sparsity pattern.

6.4.6 Off-label applicability in jump processes

This example of the ORNSTEIN-UHLENBECK process illustrates the applicability of the proposed parameter estimation method also in the setting of discontinuous jump processes, by setting the jump regularization weights ω_k^2 to zero. Further, depending on the problem type, analysing the distribution of the stochastic jumps α_k allows the inference of the diffusion parameter D , that may be estimated as the standard deviation of the jumps at the inner shooting nodes.

Table 6.7: Estimation test series: ORNSTEIN-UHLENBECK process. Results of parameter estimation on 100 independent realizations of an ORNSTEIN-UHLENBECK process driven by a compound POISSON process. Initial guess for the parameters μ and θ is 50% of the true value. Parameter D is estimated as the standard deviation of the jumps α_k at the inner shooting nodes. Values rounded to 4 digits.

Scenario (A): exact observations ^a						
parameter		estimation results, no jump regularization				
name	value	estimate	\pm	SD	(SD%)	RelErr%
μ	16.0	15.82	\pm	0.187	(1.2%)	1.1%
θ	2.2	2.181	\pm	0.028	(1.3%)	0.8%
D	4.5	4.733	\pm	0.440	(9.3%)	5.2%

Scenario (B): noisy observations, same grid ^b						
parameter		estimation results, no jump regularization				
name	value	estimate	\pm	SD	(SD%)	RelErr%
μ	16.0	15.82	\pm	0.238	(1.5%)	1.1%
θ	2.2	2.183	\pm	0.085	(3.9%)	0.8%
D	4.5	4.778	\pm	0.434	(9.1%)	6.2%

Scenario (C): noisy observations, new grid ^c						
parameter		estimation results, no jump regularization				
name	value	estimate	\pm	SD	(SD%)	RelErr%
μ	16.0	16.12	\pm	1.383	(8.6%)	0.8%
θ	2.2	1.912	\pm	0.427	(22.3%)	13.1%
D	4.5	4.763	\pm	0.465	(9.8%)	5.8%

^ano measurement error, shooting grid by heuristic (6.13) with $\Delta_{min} = 0.5$

^bmeasurement error $\epsilon \sim \mathcal{N}(0, 1.0^2)$, same shooting grid as in (a)

^cmeasurement error $\epsilon \sim \mathcal{N}(0, 1.0^2)$, shooting grid by heuristic (6.13) with $\Delta_{min} = 3.0$

SD%: standard deviation of the estimate

RelErr%: relative deviation of the estimate from the true parameter value

Appendix

A Appendix to the Crosstalk Model of GM-CSF-mediated IL-6-induced JAK-STAT Signalling

A.1 Biological obligation: materials and methods

As common practice, the *materials and methods* section briefly summarizes the utilized cell lines, reagents, laboratory equipment, and protocols. We include this section at this place for the sake of completeness.

A.1.1 Cell culture

The cell line used in this study is the benign tumourigenic HaCaT-ras cell line A5 [Mueller2001; Fusenig1998]. HaCaT A5 was cultivated in 4x modified Eagle's medium (MEM), 10% fetal calf serum (FCS) and 200 µg/mL neomycin (PAA, Colbe, Germany). Cells were passaged at a split ratio of 1:10, routinely tested for mycoplasma contamination as described in [Stacey1997] and always found to be negative.

A.1.2 Preparation of conditioned media and ELISA

Confluent HaCaT A5 cells were starved for 24 h in 4xMEM without FCS and then stimulated with recombinant human IL-6 (100 ng/mL; R&D Systems). Conditioned medium was harvested 5, 10, 15, 20, 25, 30, 45, 60, 90 and 120 min post stimulation, centrifuged for 5 min at 10000 rpm and stored in aliquots at -80°C. The cells were trypsinized and counted.

Secretion of GM-CSF was measured by ELISA using Quantikine Immunoassay kits (R&D Systems, Minneapolis, MN). Samples were tested in duplicates and experiments repeated twice.

A.1.3 Quantitative immunoblotting

Confluent HaCaT A5 cells were starved for 24 h in 4xMEM without FCS and then treated as follows: (i) stimulated solely with recombinant human IL-6 (100 ng/mL), setting *IL6+*, or (ii) stimulated solely with recombinant human GM-CSF (100 ng/mL), or (iii) stimulated with a combination of both factors (100 ng/mL each), setting *IL6+/GMCSF+*, or (iv) treated with IL-6 (100 ng/mL) combined with GM-CSF blocking antibody (2 µg/mL; MAB215), setting *IL6+/GMCSF-*, for 5, 10, 15, 20, 25, 30, 45, 60, 90 and 120 min with unstimulated cells as control. All factors and antibodies were purchased from R&D Systems.

After stimulation, cells were washed with ice-cold TBS (10 mM Tris pH 7.4, 150 mM NaCl) and lysed in 500 µL ice-cold NP-40 lysis buffer (150 mM NaCl, 20 mM Tris pH 7.4, 1 mM EDTA, 1 mM ZnCl₂, 1 mM MgCl₂, 1 mM Na₃VO₄, 10% Glycerin, 1% NP-40, protease inhibitor cocktail (Roche Diagnostics GmbH, Mannheim, Germany)).

Cell lysates were rotated for 20 min at 4°C, then centrifuged at 5000 rpm for 5 min at 4°C.

For immunoprecipitation (IP) the supernatant containing the cytoplasmic fraction was incubated on a rotating wheel over night at 4°C with antibodies and 5% Protein A-Sepharose CL-4B (GE Healthcare, Uppsala, Sweden). The following antibodies were used: anti-STAT-3, rabbit polyclonal (No. 9132, Cell Signaling Technology, Inc, MA); anti-gp130, rabbit polyclonal (No. sc-655, Santa Cruz Biotechnology, CA); anti-SOCS-3, mouse monoclonal (Invitrogen, Darmstadt Germany).

After incubation, the samples were washed with lysis buffer, eluted by boiling at 95°C in Laemmli sample buffer for 5 min, separated by SDS-PAGE and transferred to nitrocellulose membranes. Membranes were blocked in TBS-T (TBS, 0.2% Tween-20) with 2% BSA (PAA, Colbe, Germany) for 1 h at room temperature and incubated with primary antibody over night at 4°C. Primary antibodies: anti-phospho-STAT-3 (Tyr705), rabbit polyclonal (No. 9131, Cell Signaling Technology); anti-phosphotyrosine, mouse monoclonal (clone 4G10, No. 05-321, Millipore, MA); anti-SOCS-3, rabbit polyclonal (Abcam, Cambridge, UK). Antibodies were removed by treating the blots with β -mercaptoethanol and SDS as described in [Klingmueller1995]. Reprobes for quantification of the total protein were performed with the antibodies that were used for immunoprecipitation. Antibody binding was detected with horseradish peroxidase conjugated secondary antibodies (Protein A-HRP, GE Healthcare) and ECL reagent (GE Healthcare) according to the manufacturer's instructions. The signal was quantified by LumiImager using LumiAnalyst software (Roche Diagnostics). Experiments were performed in triplicates.

For determination of endogenous protein levels of gp130, defined amounts of the respective calibrators [Schilling2005] were added to cell lysates prior to immunoprecipitation. After quantification of immunoblotting signals, calibration curves were calculated by regression analysis of the dilution series of added recombinant calibrator proteins and concentrations could be estimated in absolute molecule numbers per cell as described in [Schilling2005].

A.1.4 Estimating absolute concentrations from immunoblot data

Signal quantification of phospho-STAT-3 by LumiImager results in relative Boehringer Light Units (BLU), which are not comparable between different blotting gels. To achieve comparability and to convert the BLU values into absolute concentrations, data was processed as described below. For each experiment, every BLU value in a time course was divided by the respective value at time $t = 0$ min. The resulting data denote “x-fold amplification” in cytosolic phospho-STAT-3 signal after stimulation, allowing the comparison of data obtained in different experiments.

The total number of STAT-3 molecules in HaCaT cells was determined by quantitative immunoprecipitation. To calculate the number of phosphorylated STAT-3 molecules we used literature data, where the maximum answer in phospho-STAT-3 upon IL-6 stimulation is described to be about 14.5% of total STAT-3 protein [Bohl2009]. Thus, the amplification factors were then used to calculate the base phosphorylation level of cytoplasmic STAT-3 before stimulation using the formula

$$\frac{\text{total number of molecules} \cdot \text{percentage of activation}}{\text{mean amplification factor on IL-6 stimulation}}$$

resulting in reproducible and statistically significant results.

To estimate absolute concentrations from the BLU measurement data, we used the method for calculating phospho-STAT-3 molecules described above, and estimated the average volume of HaCaT keratinocytes by measuring the cell diameter with a CASY cell counter (Millipore, Billerica, MA) and assuming a spherical shape.

Initial molar concentrations for the remaining species were calculated from measured mass concentrations (ELISA) according to the following formula:

$$\text{molar concentration (in nmol/mL)} = \frac{\text{mass concentration (in ng/mL)}}{\text{molar mass (in Dalton)}}$$

under consideration of the volume of the medium or a single HaCaT cell, respectively.

A.2 Tabulated data

Table A.1 lists the complete experimental data. For parameter estimation, only the pSTAT-3 measurements were used after processing as described in section 3.5. Data analysis was also done for recombinant (recSTAT3) and total (totSTAT3) pSTAT-3 (see section 3.5.3).

Table A.2 lists the exact values of the kinetic parameters for our model after parameter estimation. See section 3.7.3, and also the text in table 3.12 for details.

Table A.1: Time course data upon four different stimulation settings. Measurement values are in relative units (BLU), thus neither comparable between experiments nor between replicates. Times are given in minutes after stimulation. The mark “-” means, that the respective experiment has not been conducted. See section 3.5 for the applied data processing.

Setting	Time	Replicate #1			Replicate #2			Replicate #3		
		pSTAT3	recSTAT3	totSTAT3	pSTAT3	recSTAT3	totSTAT3	pSTAT3	recSTAT3	totSTAT3
unstimulated	0	165000	571000	841000	11400000	-	-	5130000	-	-
	5	164000	332000	715000	10400000	-	-	4120000	-	-
	10	137000	495000	1030000	11300000	-	-	6910000	-	-
	15	199000	487000	1100000	13000000	-	-	5210000	-	-
	20	183000	562000	1130000	8670000	-	-	6830000	-	-
	25	219000	436000	1090000	14300000	-	-	7100000	-	-
	30	194000	486000	1110000	13300000	-	-	5110000	-	-
	45	185000	764000	1290000	10600000	-	-	8500000	-	-
	60	216000	517000	1250000	11100000	-	-	5590000	-	-
	90	175000	524000	997000	10100000	-	-	5270000	-	-
	120	183700	517400	1055300	10200000	-	-	4890000	-	-
	IL6+	0	25736	16278	206000	28128	68696	165000	87794	210000
5		243000	52645	429000	348000	33238	168000	1060000	455000	-
10		759000	58513	408000	1030000	39359	285000	2910000	431000	-
15		839000	45761	378000	1160000	64062	303000	2100000	388000	-
20		1190000	29875	224000	875000	41054	202000	3080000	234000	-
25		1370000	34018	284000	1110000	56970	261000	4190000	292000	-
30		1370000	26978	234000	716000	17466	159000	3690000	241000	-
45		805000	20286	170000	545000	78551	254000	2360000	177000	-
60		464000	30328	238000	352000	40846	183000	1780000	257000	-
90		89580	18766	270000	104000	68145	228000	416000	270000	-
120		31718	13963	162000	76681	86380	321000	131000	173000	-
IL6+/GMCSF-		0	440000	438000	1010000	9670000	593000	-	218000	-
	5	1840000	252000	638000	40900000	1260000	-	106000	-	-
	10	3610000	380000	613000	46200000	879000	-	1290000	-	-
	15	4410000	288000	833000	57700000	402000	-	2840000	-	-
	20	4350000	436000	823000	40000000	1000000	-	1160000	-	-
	25	5780000	277000	211000	49600000	908000	-	2300000	-	-
	30	4590000	164000	424000	54900000	179000	-	4700000	-	-
	45	4550000	403000	547000	54900000	700000	-	1530000	-	-
	60	2120000	289000	666000	48200000	1580000	-	345000	-	-
	90	1630000	359000	622000	15800000	533000	-	109000	-	-
	120	488000	108000	211000	20000000	871000	-	52693	-	-
	IL6+/GMCSF+	0	9315	331000	771000	380000	-	-	5200000	-
5		66842	430000	917000	1210000	-	-	10500000	-	2660000
10		75450	244000	718000	2500000	-	-	6670000	-	673000
15		112000	207000	689000	4770000	-	-	63700000	-	16700000
20		156000	390000	612000	7010000	-	-	26000000	-	1070000
25		176000	479000	712000	6520000	-	-	85300000	-	20000000
30		148000	438000	886000	4220000	-	-	33200000	-	9210000
45		71949	379000	771000	2000000	-	-	20300000	-	11300000
60		23389	190000	787000	1430000	-	-	13300000	-	11200000
90		31244	424000	839000	210000	-	-	8450000	-	20500000
120		14696	308000	685000	45387	-	-	5410000	-	18300000

Table A.2: Final estimates of parameters in the crosstalk model – non-rounded values. Non-rounded parameter values retrieved for the model in section 3.7.3 after parameter estimation. See table 3.12 for explanations. SE: Standard Error.

Par.	Estimate	Unit	SE	Meaning
p_{01}	$5.77211608 \cdot 10^{-7}$	$\frac{nmol}{mL \cdot min}$	—*	production rate of IL-6 and GM-CSF
p_{02}	$1.81723897 \cdot 10^1$	$\frac{mL}{nmol \cdot min}$	9.6 %	activation rate of IL-6 receptor and GM-CSF receptor upon ligand binding
p_{03}	$1.00000000 \cdot 10^5$	$\frac{mL}{nmol \cdot min}$	—*	blockade of GM-CSF receptor upon overstimulation
p_{04}	$7.27351250 \cdot 10^4$	$\frac{mL}{nmol \cdot min}$	—*	activation rate of SK by active GM-CSF receptor
p_{05}	$3.16609443 \cdot 10^4$	$\frac{mL^2}{nmol^2 \cdot min}$	39.8 %	SK enhanced STAT-3 activation on active IL-6 receptor
p_{06}	$1.00000000 \cdot 10^{-1}$	min^{-1}	—*	degradation rate of SOCS-3 in cytosol
p_{07}	$3.00482209 \cdot 10^3$	$\frac{mL}{nmol \cdot min}$	30.2 %	deactivation rate of IL-6 and GM-CSF receptor by SOCS-3
p_{08}	$6.12798523 \cdot 10^{-1}$	$\frac{mL}{nmol \cdot min}$	10.2 %	STAT-3 phosphorylation rate on active IL-6 receptor
p_{09}	$3.75073537 \cdot 10^{-2}$	min^{-1}	5.3 %	translocation of phospho-STAT-3 into nucleus
p_{10}	$2.15261866 \cdot 10^{-2}$	min^{-1}	43.4 %	export rate of nuclear STAT-3 into cytoplasm
p_{11}	$3.96962821 \cdot 10^{-2}$	min^{-1}	17.3 %	induction and export rate of SOCS-3 mRNA by nuclear phospho-STAT-3
p_{12}	$1.40267745 \cdot 10^{-1}$	min^{-1}	—*	translation rate of SOCS-3 mRNA in cytoplasm (10 proteins from 1 mRNA)
p_{SOCS}	$9.40267745 \cdot 10^{-10}$	min^{-1}	11.5 %	scaling constant for SOCS-3 measurement data

* parameter was fixed to specified value

B Some Results on Matrix Decompositions and Stochastic Convergence

B.1 QR decomposition of full rank matrices

As discussed in section 1.2, linear least squares problems LIN-LSQ

$$\min_{x \in \mathbb{R}^n} \frac{1}{2} \|Cx - d\|_2^2 \quad \text{with } C \in \mathbb{R}^{m \times n}, d \in \mathbb{R}^m \quad (\text{B.1})$$

(unconstrained like above or with linear equality constraints like problem LIN-LSQ-EC on page 9) can be solved using decompositions of the system matrices.

We summarize some results on QR decompositions of matrices; proofs can be found on most textbooks of linear algebra.

B.1 Theorem (QR decomposition of full-rank matrices)

Let $C \in \mathbb{R}^{m \times n}$, $m \geq n$, $\text{rank}(C) = n$. Then there exist an orthogonal matrix $Q \in \mathbb{R}^{m \times m}$ and a nonsingular upper triangular matrix $R \in \mathbb{R}^{n \times n}$ such that C can be decomposed as:

$$C = Q \cdot \begin{bmatrix} R \\ 0_{(m-n) \times n} \end{bmatrix} \quad \square$$

By choosing the signs of the columns of Q in the right way, we can ensure positive diagonal entries in R , and derive the following uniqueness result of the QR decomposition:

B.2 Theorem (Uniqueness of the QR decomposition)

Let $C \in \mathbb{R}^{m \times n}$, $m \geq n$, $\text{rank}(C) = n$, and $C = Q \cdot \begin{bmatrix} R \\ 0 \end{bmatrix}$ with orthogonal $Q \in \mathbb{R}^{m \times m}$ and nonsingular upper triangular $R \in \mathbb{R}^{n \times n}$ as in theorem B.1. Further, let Q be partitioned as

$$Q = [Q_1 \quad Q_2], \quad Q_1 \in \mathbb{R}^{m \times n}, \quad Q_2 \in \mathbb{R}^{m \times (m-n)}.$$

Then we have

$$C = [Q_1 \quad Q_2] \cdot \begin{bmatrix} R \\ 0 \end{bmatrix} = Q_1 R \quad \implies \quad Q_1 = CR^{-1},$$

so Q_1 is uniquely determined. □

The calculation of Q and R can be done using elementary orthogonal transformations, like GIVENS rotations, HOUSEHOLDER transformations, or GRAM-SCHMIDT orthogonalization, or variants of these. See, e.g., [GolubVanLoan1996].

The QR decomposition can also be applied to rank-deficient matrices $C \in \mathbb{R}^{m \times n}$ with $\text{rank}(C) < n$, which includes the setting of the system matrix C having more columns than rows, i.e. $m < n$, as special case.

B.3 Theorem (QR decomposition of rank-deficient matrices)

Let $C \in \mathbb{R}^{m \times n}$ with $\text{rank}(C) = r$. Then there exist an orthogonal matrix $Q \in \mathbb{R}^{m \times m}$, a nonsingular upper triangular matrix $R_{11} \in \mathbb{R}^{r \times r}$, and a permutation matrix $P \in \mathbb{R}^{n \times n}$ such that the matrix C can be decomposed as:

$$CP = Q \cdot \begin{bmatrix} R_{11} & R_{12} \\ 0_{(m-r) \times r} & 0_{(m-r) \times (m-r)} \end{bmatrix} \quad \square$$

For a fixed permutation P , the matrices R_{11} and R_{12} are unique if we fix the signs of the diagonal elements of R_{11} (e.g. all positive). In that case, also the first r columns of Q are uniquely determined.

B.2 The singular value decomposition (SVD)

By choosing the right orthogonal coordinate systems, every matrix can be diagonalized by finding its *singular value decomposition* (SVD). Though more expensive than QR decompositions, it offers much more insight into the underlying structures, and is used in many fields of application, like pattern recognition, time-series analysis, image processing and for solving least squares problems, especially in the rank-deficient case.

Though the formulation as well as applications for complex matrices exist, we restrict ourselves here to the SVD for real matrices. The complex version reads identically when replacing \mathbb{R} by \mathbb{C} , “orthogonal” by “unitary”, and “transposition” by “conjugate transposition”. For proofs, we refer to, e.g., [GolubVanLoan1996].

B.4 Theorem (Singular Value Decomposition)

Let $C \in \mathbb{R}^{m \times n}$ with $\text{rank}(C) = r$. Then there exist an orthogonal matrix $U \in \mathbb{R}^{m \times m}$, an orthogonal matrix $V \in \mathbb{R}^{n \times n}$, and a diagonal matrix $\Sigma \in \mathbb{R}^{m \times n}$, such that

$$C = U\Sigma V^T, \quad \Sigma = \begin{bmatrix} \bar{\Sigma} & 0 \\ 0 & 0 \end{bmatrix} \in \mathbb{R}^{m \times n}, \quad \bar{\Sigma} = \begin{bmatrix} \sigma_1 & & 0 \\ & \ddots & \\ 0 & & \sigma_r \end{bmatrix} \in \mathbb{R}^{r \times r}$$

with nonzero *singular values* $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > 0$. The column vectors u_i of $U = (u_1, \dots, u_m)$ are the *left singular vectors* of C , and the column vectors v_i of $V = (v_1, \dots, v_n)$ are the *right singular vectors* of C . □

B.5 Theorem (Pseudo-inverse by SVD)

Let $C \in \mathbb{R}^{m \times n}$ with $\text{rank}(C) = r$, and $C = U\Sigma V^T$ a singular value decomposition as in theorem B.4. Then, the *pseudo-inverse* or *MOORE-PENROSE pseudo-inverse* of C is given as

$$C^\dagger = V\Sigma^\dagger U^T, \quad \Sigma^\dagger = \begin{bmatrix} \bar{\Sigma}^{-1} & 0 \\ 0 & 0 \end{bmatrix} \in \mathbb{R}^{n \times m}, \quad \bar{\Sigma}^{-1} = \begin{bmatrix} \sigma_1^{-1} & & \\ & \ddots & \\ & & \sigma_r^{-1} \end{bmatrix} \in \mathbb{R}^{r \times r}$$

Note that, while the dimension of Σ is $m \times n$, its pseudoinverse Σ^\dagger has dimension $n \times m$. □

B.3 Convergence of random sequences

Here, we shortly summarize some common types of convergence of random sequences and give some characterizations without proof. We have, again, $(\Omega, \mathcal{A}, \mathbb{P})$ a probability space, and random sequences $X_n(\omega) \rightarrow X(\omega)$ ($n \rightarrow \infty$)

(I) *Pointwise convergence, or sure convergence:*

$$\lim_{n \rightarrow \infty} X_n(\omega) = X(\omega) \quad \forall \omega \in \Omega \quad (\text{B.2})$$

(II) *Convergence with probability 1, or almost sure convergence, or strong convergence:*

$$\mathbb{P}\left(\lim_{n \rightarrow \infty} X_n = X\right) = \mathbb{P}\left(\{\omega \in \Omega \mid \lim_{n \rightarrow \infty} X_n(\omega) = X(\omega)\}\right) = 1 \quad (\text{B.3})$$

(III) *Mean square convergence, or convergence in L^2 :*

$$\mathbb{E}[X^2] < \infty, \quad \mathbb{E}[X_n^2] < \infty \quad \forall n, \quad \text{and} \quad \lim_{n \rightarrow \infty} \mathbb{E}[|X_n - X|^2] = 0 \quad (\text{B.4})$$

(IV) *Convergence in probability, or stochastic convergence:*

$$\lim_{n \rightarrow \infty} \mathbb{P}\left(\{\omega \in \Omega \mid |X_n(\omega) - X(\omega)| \geq \varepsilon\}\right) = 0 \quad \forall \varepsilon > 0 \quad (\text{B.5})$$

(V) *Convergence in distribution, or convergence in law, or weak convergence:*

$$\lim_{n \rightarrow \infty} F_{X_n}(x) = F_X(x) \quad \text{at all continuity points } x \quad (\text{B.6})$$

with F_X and F_{X_n} denoting the cumulative distribution functions of X and X_n .

Equivalently in terms of densities:

(VI) *Weak convergence:*

$$\lim_{n \rightarrow \infty} \int_{-\infty}^{\infty} f(x) dF_{X_n}(x) = \int_{-\infty}^{\infty} f(x) dF_X(x) \quad (\text{B.7})$$

for all test functions $f : \mathbb{R} \rightarrow \mathbb{R}$, continuous and with bounded support.

If the distribution functions possess densities p_n and p , then this is equivalent to

$$\lim_{n \rightarrow \infty} \int_{-\infty}^{\infty} f(x) p_n(x) dx = \int_{-\infty}^{\infty} f(x) p(x) dx$$

Further, we have the following implications (this is not a complete list):

- 1) almost sure convergence (II)
 - \implies convergence in probability (IV)
 - \implies convergence in distribution (V)
- 2) mean square convergence (III)
 - \implies convergence in probability (IV)
 - \implies convergence in distribution (V)

C The Stochastic Integral

This introduction to the stochastic integral is based on chapter 3 of the book of Bernt ØKSENDAL [Oksendal1998]. We assume that the reader is familiar to the basic definitions from probability theory in section 4.1.

In what follows, W_t is a standard WIENER process (BROWNIAN motion) on the time domain $[0, \infty)$.

We restrict ourselves to real valued functions and conclude with a definition of an n -dimensional stochastic integral. Most results extend analogously to multi-dimensional functions, but require much more theoretical preparation. We refer to [Oksendal1998], [Bichteler2002], and [Protter2004].

C.1 The idiosyncrasy of stochastic integrals

Let $f: [0, T] \times \Omega \rightarrow \mathbb{R}$ be a stochastic process. It is a natural approach to approximate $f(\cdot, \omega)$ by a sum of piecewise constant functions

$$f(t, \omega) \approx \sum_{j=0}^{m-1} f(t_j, \omega) \cdot \mathbb{1}_{[\tau_j, \tau_{j+1})}(t) \quad (\text{C.1})$$

using a time grid $0 = \tau_0 < \tau_1 < \dots < \tau_m = T$ and supporting points $t_j \in [\tau_j, \tau_{j+1}]$.

One might try to define a stochastic integral with respect to a WIENER process W_t similarly to the RIEMANN-STIELTJES integral as:

$$\int_0^T f(t, \omega) dW_t(\omega) := \lim_{m \rightarrow \infty} \sum_{j=0}^{m-1} f(t_j, \omega) \cdot (W_{\tau_{j+1}} - W_{\tau_j}). \quad (\text{C.2})$$

However, this definition brings us into serious trouble, as the following example shows:

Let $f(t, \omega) = W_t(\omega)$ be a WIENER process. Then, the functions

$$f_1(t, \omega) = \sum_{j=0}^{m-1} W_{\tau_j}(\omega) \cdot \mathbb{1}_{[\tau_j, \tau_{j+1})}(t)$$

$$f_2(t, \omega) = \sum_{j=0}^{m-1} W_{\tau_{j+1}}(\omega) \cdot \mathbb{1}_{[\tau_j, \tau_{j+1})}(t)$$

would be two approximations of f that only differ in the choice of the supporting points t_j . As $|\tau_{j+1} - \tau_j| \rightarrow 0$ ($m \rightarrow \infty$), both approximations converge to f .

For f_1 , we have chosen to evaluate f at the left boundary of the interval $[\tau_j, \tau_{j+1}]$, and at the right boundary for f_2 , leading to respective integrals

$$\begin{aligned} \int_0^T f_1(t, \omega) dW_t(\omega) &= \lim_{m \rightarrow \infty} \sum_{j=0}^{m-1} W_{\tau_j}(\omega) \cdot (W_{\tau_{j+1}} - W_{\tau_j}) \\ \int_0^T f_2(t, \omega) dW_t(\omega) &= \lim_{m \rightarrow \infty} \sum_{j=0}^{m-1} W_{\tau_{j+1}}(\omega) \cdot (W_{\tau_{j+1}} - W_{\tau_j}). \end{aligned}$$

Assuming the linearity of the integral (as we seriously do not want to construct a nonlinear integral), for any partition $0 = \tau_0 < \tau_1 < \dots < \tau_m = T$, using the fact that a WIENER process has independent increments, one gets

$$\begin{aligned} \mathbb{E} \left[\int_0^T f_1(t) dW_t \right] &= \lim_{m \rightarrow \infty} \sum_{j=0}^{m-1} \mathbb{E} [W_{\tau_j} \cdot (W_{\tau_{j+1}} - W_{\tau_j})] \\ &= \lim_{m \rightarrow \infty} \sum_{j=0}^{m-1} \mathbb{E} [W_{\tau_j} W_{\tau_{j+1}}] - \mathbb{E} [W_{\tau_j}^2] = \lim_{m \rightarrow \infty} \sum_{j=0}^{m-1} (\tau_{j+1} - \tau_j) = 0 \end{aligned}$$

and

$$\begin{aligned} \mathbb{E} \left[\int_0^T f_2(t) dW_t \right] &= \lim_{m \rightarrow \infty} \sum_{j=0}^{m-1} \mathbb{E} [W_{\tau_{j+1}} \cdot (W_{\tau_{j+1}} - W_{\tau_j})] \\ &= \lim_{m \rightarrow \infty} \sum_{j=0}^{m-1} \mathbb{E} [W_{\tau_{j+1}}^2] - \mathbb{E} [W_{\tau_j} W_{\tau_{j+1}}] = \lim_{m \rightarrow \infty} \sum_{j=0}^{m-1} (\tau_{j+1} - \tau_j) = T \end{aligned}$$

The above example illustrates that, in spite of both functions f_1 and f_2 being good approximations of W_t , the value of the integral depends on the choice where the integrand is evaluated, and these values do not converge no matter how we chose the partition of the time domain – a less desired property of an integral. As ØKSENDAL writes:

“This reflects the fact that the variations of the paths of W_t are too big to enable us to define the integral in the RIEMANN-STIELTJES sense. [...] In particular, the total variation of the path is infinite, a.s.” [Oksendal1998]

C.2 ITÔ's and STRATONOVICH's integral

As exemplified in section C.1, the (expected) value of a stochastic integral

$$\int_0^T f(t, \omega) dW_t(\omega) = \lim_{m \rightarrow \infty} \sum_{j=0}^{m-1} f(t_j, \omega) \cdot (W_{\tau_{j+1}}(\omega) - W_{\tau_j}(\omega)).$$

depends on the choice of the supporting points t_j , of which there are two popular choices:

- 1) $t_j := \tau_j$, the left boundary of the interval $[\tau_j, \tau_{j+1}]$, leads to the *ITÔ stochastic integral*, which is usually written as

$$\int_0^T f(t, \omega) dW_t$$

- 2) $t_j := \frac{1}{2}(\tau_j + \tau_{j+1})$, the middle point of the interval $[\tau_j, \tau_{j+1}]$, leads to the *STRATONOVICH stochastic integral*, usually written as

$$\int_0^T f(t, \omega) \circ dW_t .$$

Both choices for t_j are useful in different settings of theory and practical calculation. For any continuously differentiable function $g(t, W)$, they can be easily converted into each other by the formula

$$\int_0^T g(t, W_t) \circ dW_t = \frac{1}{2} \int_0^T \frac{\partial g}{\partial W}(t, W_t) dt + \int_0^T g(t, W_t) dW_t .$$

C.2.1 Which integral to choose: STRATONOVICH or ITÔ?

In many settings, especially in biology, one would prefer the ITÔ interpretation, as it “does not look into the future”. The advantage of the STRATONOVICH integral is that it leads to an “ordinary” chain rule, whereas the chain rule of the ITÔ integral is different. Thus, as soon as variable changes occur, the STRATONOVICH integral leads to “less complicated” formulas than the ITÔ version. Another point is, that ITÔ integrals are martingales (see definition C.10), whereas STRATONOVICH integrals are not, which makes the first one easier to compute.

However, as both integrals can be converted into each other for a wide class of integrands, it is usually sufficient to investigate only one of them. For a deeper discussion, see [Oksendal1998], chapter 3, and the references given therein.

C.3 Filtration, adaptation, and the “history of BROWNIAN motion”

Given a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ and a random variable $X : \Omega \rightarrow \mathbb{R}$, the σ -algebra \mathcal{A} codes the events that can be detected by the probability measure \mathbb{P} . What the single σ -algebra \mathcal{A} is for a single random variable X , is a whole family of σ -algebras for a stochastic process X_t .

C.1 Definition (Filtration, adapted process, simple process)

Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space, and $X : [0, \infty) \times \Omega \rightarrow \mathbb{R}^n$ be a stochastic process.

1. A *filtration* is a family $\{\mathcal{M}_t\}_{t \geq 0}$ of increasing sub- σ -algebras $\mathcal{M}_t \subset \mathcal{A}$, such that $0 \leq s < t \implies \mathcal{M}_s \subset \mathcal{M}_t$.
As a convention, we assume that \mathcal{M}_0 always contains all \mathbb{P} -null sets. Further, we frequently write $\{\mathcal{M}_t\}$ or just \mathcal{M}_t for $\{\mathcal{M}_t\}_{t \geq 0}$.
2. For a given filtration $\{\mathcal{M}_t\}$, the process X_t is called *\mathcal{M}_t -adapted*, if $X_t(\omega)$ is \mathcal{M}_t -measurable for each $t \geq 0$.

3. For a filtration $\{\mathcal{M}_t\}$, the quadruple $(\Omega, \mathcal{A}, \{\mathcal{M}_t\}, \mathbb{P})$ is called a *filtered probability space* or *stochastic basis* (again, we assume here that the filtered probability space is complete, i.e. \mathcal{M}_0 contains all \mathbb{P} -null sets).
4. The filtration $\mathcal{N}_t := \sigma(\{X_s\}_{s \leq t})$, i.e. the σ -algebra generated by the random variables $X_s(\omega); s \leq t$, is called the *natural filtration* of the process X_t . It describes the setting, that for every time point t the whole information about the history of X_t is known up to time t .

\mathcal{F}_t denotes the *natural filtration* of an n -dimensional BROWNIAN motion W_t . It is the smallest σ -algebra containing the sets $\{\omega, W_{t_1}(\omega) \in B_1, \dots, W_{t_k}(\omega) \in B_k\}$ with $t_j \leq t$ and $B_j \in \mathbb{B}^n$ being BOREL sets for all $j \leq k = 1, 2, 3, \dots$ □

For a WIENER process W_t and its filtration \mathcal{F}_t , a possible interpretation of an \mathcal{F}_t -adapted process $\{X_t\}$ is that although X_s is a random variable for every time point s , its value is completely defined by the path of the BROWNIAN motion $\{W_t\}_{0 \leq t \leq s}$ up to that time point s . The filtration \mathcal{F}_t can thus be seen as the *history* of the WIENER process W_t up to time t .

C.4 The ITÔ integral

We have not said anything yet about the class of functions to integrate. Similarly to common constructions of the LEBESGUE or RIEMANN-STIELTJES integral, we start with a definition of the ITÔ integral for a class of simple (step) functions, and extend this definition to more complicated functions that can be suitably approximated by simple functions.

C.2 Definition (class \mathcal{V})

Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a complete probability space and $\{\mathcal{F}_t\}$ the natural filtration of a WIENER process W_t . By $\mathcal{V} = \mathcal{V}([0, T])$ we denote the class of real-valued functions $f: [0, \infty) \times \Omega \rightarrow \mathbb{R}$ that fulfill the following conditions:

- i) f is $\mathbb{B} \times \mathcal{A}$ -measurable. Here, \mathbb{B} denotes the BOREL σ -algebra on $[0, \infty)$.
- ii) f is \mathcal{F}_t -adapted.
- iii) $\mathbb{E} \left[\int_0^T f(t, \omega)^2 dt \right] < \infty$. □

C.3 Definition (Simple processes)

An \mathcal{F}_t -adapted process $X_t \in \mathcal{V}$ is called *simple* or *elementary*, if it is a random step function, i.e. X_t can be written as

$$X_t(\omega) = \sum_{j=0}^{m-1} \xi_j(\omega) \cdot \mathbb{1}_{[\tau_j, \tau_{j+1})}(t)$$

for a time grid $0 = \tau_0 < \tau_1 < \tau_2 < \dots < \tau_m$, and random variables ξ_j that are \mathcal{F}_{t_j} -measurable ($j = 0, \dots, m - 1$).

Note that, when we talk about a simple process, we always implicitly assume that it is associated with a suitable time grid $0 = \tau_0 < \tau_1 < \tau_2 < \dots < \tau_m$. □

In this survey, we restrict ourselves to integrals on the domain $[0, T]$, as we can easily define integrals over intervals $[a, b]$ (equivalently $(a, b], [a, b), (a, b)$) with nonrandom $0 \leq a \leq b \leq T$ by

$$\int_a^b X_t dW_t := \int_0^T X_t \cdot \mathbb{1}_{[a,b]}(t) dW_t$$

C.4.1 The ITÔ integral in 1D

C.4 Definition (ITÔ integral for simple processes)

The ITÔ integral for a simple process $X_t(\omega) = \sum_{j=0}^{m-1} \xi_j(\omega) \cdot \mathbb{1}_{[\tau_j, \tau_{j+1})}(t)$ is defined as

$$\int_0^T X_t(\omega) dW_t(\omega) := \sum_{j=0}^{m-1} \xi_j(\omega) \cdot \Delta W_j(\omega) \tag{C.3}$$

with $\Delta W_j(\omega) := W_{\tau_{j+1}}(\omega) - W_{\tau_j}(\omega)$. □

Now that we have the ITÔ integral for simple processes, we extend the definition to arbitrary functions in \mathcal{V} . Following ØKSENDAL, we do this in three steps. First, we show that bounded continuous functions in \mathcal{V} can be suitably approximated by elementary functions. Second, we show that we can drop the continuity requirement, as we can approximate every bounded function in \mathcal{V} by continuous (and bounded) functions. Finally, we extend the approximability to arbitrary functions in \mathcal{V} . We first introduce the following isometry that will be needed in the more general definition C.7 of the ITÔ integral.

C.5 Lemma (ITÔ isometry for simple processes)

For any bounded and simple process $X_t \in \mathcal{V}$, i.e. $X_t(\omega) = \sum_{j=0}^{m-1} \xi_j(\omega) \cdot \mathbb{1}_{[\tau_j, \tau_{j+1})}(t)$, it holds:

$$\mathbb{E} \left[\left(\int_0^T X_t(\omega) dW_t(\omega) \right)^2 \right] = \mathbb{E} \left[\int_0^T X_t^2(\omega) dt \right] \tag{C.4}$$

Proof: Setting $\Delta W_j := W_{\tau_{j+1}} - W_{\tau_j}$ and using the linearity of expectation, we get:

$$\begin{aligned} \mathbb{E} \left[\left(\int_0^T X_t(\omega) dW_t(\omega) \right)^2 \right] &= \sum_{j=0}^{m-1} \sum_{i=0}^{m-1} \mathbb{E}[\xi_i(\omega) \xi_j(\omega) \Delta W_i \Delta W_j] \stackrel{(*)}{=} \sum_{j=0}^{m-1} \mathbb{E}[\xi_j^2(\omega)] \cdot (\tau_{j+1} - \tau_j) \\ &= \mathbb{E} \left[\sum_{j=0}^{m-1} \xi_j^2(\omega) \cdot (\tau_{j+1} - \tau_j) \right] = \mathbb{E} \left[\int_0^T \sum_{j=0}^{m-1} \xi_j^2(\omega) \cdot \mathbb{1}_{[\tau_j, \tau_{j+1})}(t) dt \right] = \mathbb{E} \left[\int_0^T X_t^2(\omega) dt \right] \end{aligned}$$

where we have used the independence of $\xi_i \xi_j \Delta W_i$ and ΔW_j for $i < j$ in (*). □

C.6 Lemma (Approximability of $X_t \in \mathcal{V}$ by simple functions)

For every process $X_t \in \mathcal{V}$, there is a sequence of simple processes $X_t^{(n)} \in \mathcal{V}$, such that

$$\mathbb{E} \left[\int_0^T (X_t - X_t^{(n)})^2 dt \right] \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

Proof (sketch):

1. If $X_t(\omega)$ is bounded and continuous in t for each ω , then, for $\tau_j^{(n)} := j \cdot \frac{T}{n}$, the processes $X_t^{(n)}(\omega) := \sum_{j=0}^n X_{\tau_j^{(n)}}(\omega) \cdot \mathbb{1}_{[\tau_j^{(n)}, \tau_{j+1}^{(n)})}(t)$ are simple processes in \mathcal{V} and continuous in t for each ω . Thus, $X_t^{(n)}(\omega) \rightarrow X_t(\omega)$ ($n \rightarrow \infty$) pointwise for each ω , and, since $X_t^{(n)}$ is also bounded, we have $\mathbb{E} \left[\int_0^T (X_t - X_t^{(n)})^2 dt \right] \rightarrow 0$ ($n \rightarrow \infty$) by the bounded convergence theorem.

2. Let $X_t \in \mathcal{V}$ be bounded, such that $|X_t(\omega)| \leq K \forall (t, \omega)$. We approximate X_t using a series of suitable convolution integrals. For that, we first construct a series h_n that converges to an identity. Let h_n be a non-negative and continuous function on \mathbb{R} , with $\int_{-\infty}^{\infty} h_n(x) dx = 1$ and $h_n(x) = 0$ for $x \leq -\frac{1}{n}$ and $x \geq 0$, such that $\text{supp}(h_n) = [-\frac{1}{n}, 0]$.

Then, we set $X_t^{(n)}(\omega) := \int_0^t h_n(s-t) X_s(\omega) ds$. It can be shown that $X_t^{(n)}(\cdot)$ is \mathcal{F}_t -measurable, and $\int_0^T (X_s(\omega) - X_s^{(n)}(\omega))^2 ds \rightarrow 0$ ($n \rightarrow \infty$) for each ω . Further, again using the bounded convergence theorem, one gets $\mathbb{E} \left[\int_0^T (X_t - X_t^{(n)})^2 dt \right] \rightarrow 0$ ($n \rightarrow \infty$).

3. Now, let $X_t \in \mathcal{V}$ be an arbitrary process in \mathcal{V} . Then, for every $n \in \mathbb{N}$,

$$X_t^{(n)}(\omega) := \begin{cases} -n & \text{if } X_t(\omega) < -n \\ X_t(\omega) & \text{if } -n \leq X_t(\omega) < n \\ n & \text{if } X_t(\omega) > n \end{cases}$$

is bounded, $X_t^{(n)} \in \mathcal{V}$, and $X_t^{(n)} \rightarrow X_t$ ($n \rightarrow \infty$). Using the dominated convergence theorem, we finally get $\mathbb{E} \left[\int_0^T (X_t - X_t^{(n)})^2 dt \right] \rightarrow 0$ ($n \rightarrow \infty$), i.e. the desired result for arbitrary processes $X_t \in \mathcal{V}$.

□

C.7 Definition (ITÔ integral in 1D)

Let $X_t \in \mathcal{V}$, and $X_t^{(n)} \in \mathcal{V}$ be a sequence of elementary processes as in lemma C.6, such that

$$\lim_{n \rightarrow \infty} \mathbb{E} \left[\int_0^T (X_t(\omega) - X_t^{(n)}(\omega))^2 dt \right] = 0.$$

Then, the ITÔ integral of X_t is defined as

$$\int_0^T X_t(\omega) dW_t(\omega) := \lim_{n \rightarrow \infty} \int_0^T X_t^{(n)}(\omega) dW_t(\omega) \tag{C.5}$$

where the limit is taken in $L^2(\mathbb{P})$, i.e. mean square convergence. □

The ITÔ isometry for simple processes (Lemma C.5) ensures, that the ITÔ integral is well-defined, i.e. that the limit in eq. (C.5) exists and is independent of the choice of the approximating processes $X_t^{(n)}$.

C.4.2 Properties of the ITÔ-Integral

C.8 Corollary (ITÔ isometry)

For $X_t \in \mathcal{V}$, it holds:

$$\mathbb{E} \left[\left(\int_0^T X_t(\omega) dW_t(\omega) \right)^2 \right] = \mathbb{E} \left[\int_0^T X_t^2(\omega) dt \right]$$

□

C.9 Lemma (Properties of the ITÔ integral)

Let $X_t, Y_t \in \mathcal{V}$.

1. The ITÔ integral is linear, i.e. for $\alpha, \beta \in \mathbb{R}$, it holds for almost all $\omega \in \Omega$:

$$\int_0^T (\alpha X_t + \beta Y_t) dW_t = \alpha \cdot \int_0^T X_t dW_t + \beta \cdot \int_0^T Y_t dW_t$$

2. For $0 \leq S < U < T$, it holds for almost all $\omega \in \Omega$:

$$\int_S^T X_t dW_t = \int_S^U X_t dW_t + \int_U^T X_t dW_t$$

3. The expectation of an ITÔ integral is zero:

$$\mathbb{E} \left[\int_0^T X_t(\omega) dW_t(\omega) \right] = 0$$

4. The ITÔ integral $\int_0^T X_t dW_t$ is \mathcal{F}_T -measurable.

Proof: By lemma C.6, it is sufficient to prove the propositions for simple processes. We take the union of the partitions of $[0, T]$ for two simple processes X_t and Y_t , i.e. we use a combined time grid $\{\tau_j\}_{j=0, \dots, m-1}$, and set

$$X_t = \sum_{j=0}^{m-1} \xi_j \cdot \mathbb{1}_{[\tau_j, \tau_{j+1})}, \quad Y_t = \sum_{j=0}^{m-1} \zeta_j \cdot \mathbb{1}_{[\tau_j, \tau_{j+1})}$$

$$\begin{aligned} 1. \text{ Linearity: } \int_0^T (\alpha X_t + \beta Y_t) dW_t &= \sum_{j=0}^{m-1} (\alpha \xi_j + \beta \zeta_j) \Delta W_j \\ &= \alpha \sum_{j=0}^{m-1} \xi_j \Delta W_j + \beta \sum_{j=0}^{m-1} \zeta_j \Delta W_j = \alpha \int_0^T X_t dW_t + \beta \int_0^T Y_t dW_t \end{aligned}$$

2. Splitting the time grid at U , such that $S = \tau_0 < \tau_1 < \dots < \tau_k = U < \tau_{k+1} < \dots < \tau_m = T$, and using the linearity of the integral, we immediately get the result.

3. Since $\mathbb{E}[\Delta W_j] = 0$, using the independence of X_{τ_j} and ΔW_j together with the linearity of expectation, we get: $\mathbb{E} \left[\int_0^T X_t(\omega) dW_t(\omega) \right] = \sum_{j=0}^{m-1} \mathbb{E}[\xi_j \Delta W_j] = \sum_{j=0}^{m-1} \mathbb{E}[\xi_j] \mathbb{E}[\Delta W_j] = 0$.

4. Since ξ_j and ΔW_j are measurable w.r.t. T , also $\int_0^T X_t dW_t = \sum_{j=0}^{m-1} \xi_j \Delta W_j$ is \mathcal{F}_t measurable. □

To show the continuity of ITÔ's stochastic integral, we have to introduce the concept of *martingales*.

C.10 Definition (Martingale)

Let $(\Omega, \mathcal{A}, \{\mathcal{M}_t\}, \mathbb{P})$ be a filtered probability space, and $X_t: [0, \infty) \times \Omega \rightarrow \mathbb{R}^n$ be an n -dimensional stochastic process. X_t is called a *martingale* with respect to the filtration $\{\mathcal{M}_t\}$, if all of the following hold:

- i) X_t is \mathcal{M}_t -adapted, i.e. X_t is \mathcal{M}_t -measurable for each $t \geq 0$,
 - ii) $\mathbb{E}[|X_t(\omega)|] < \infty \quad \forall t \geq 0$,
 - iii) $\mathbb{E}[X_s | \mathcal{M}_t] = X_t$ for all $s \geq t$.
-

C.11 Lemma (DOOB's inequalities for continuous martingales)

For a continuous martingale $\{X_t\}_{t \geq 0}$ and $p > 1$, it holds for any $t \geq 0$:

$$\mathbb{E} \left[\sup_{0 \leq s \leq t} |X_s|^p \right] \leq \left(\frac{p}{p-1} \right)^p \cdot \mathbb{E} \left[|X_t|^p \right] \quad (\text{C.6})$$

and especially for $p = 2$:

$$\mathbb{E} \left[\sup_{0 \leq s \leq t} |X_s|^2 \right] \leq 4 \mathbb{E} \left[|X_t|^2 \right]. \quad (\text{C.7})$$

Further, for all $p \geq 0$, $t \geq 0$, and $\lambda > 0$, it holds:

$$\mathbb{P} \left[\sup_{0 \leq s \leq t} |X_s| \geq \lambda \right] \leq \frac{\mathbb{E} \left[|X_t|^p \right]}{\lambda^p}. \quad (\text{C.8})$$

The first result in eq. (C.6) is frequently called *DOOB's maximal inequality for continuous martingales*, whereas the second eq. (C.8) is known as *DOOB's martingale inequality*.

Proof: See, e.g., [MoertersPeres2010], Proposition 2.43 and Theorem 12.30, or [Kallenberg2002], Proposition 7.16, and [KallenbergSztencel1991] for vector-valued martingales. \square

C.12 Theorem (Continuity of the ITÔ integral)

Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space and $X_t \in \mathcal{V}$ a stochastic process, and define

$$Y_t := \int_0^t X_s(\omega) dW_s(\omega) \quad (0 \leq t \leq T).$$

Then:

1. Y_t is an \mathcal{F}_t -martingale.
2. Y_t has a continuous version, i.e. there exists a process Z_t on $(\Omega, \mathcal{A}, \mathbb{P})$ that is continuous in t and $\mathbb{P}(Z_t = Y_t) = 1$ ($0 \leq t \leq T$).

\square

In the proof of theorem C.12, one shows for a series of simple processes $X_t^{(n)}$ approximating X_t that the ITÔ integral $I_n(t, \omega)$ of each $X_t^{(n)}$ is continuous and a martingale w.r.t. \mathcal{F}_t . Using *DOOB's martingale inequality*, one may choose a subsequence $I_{n_k}(t, \omega)$ of integrals, uniformly convergent on $[0, T]$ for almost all $\omega \in \Omega$, whose limit is the t -continuous integral Z_t . For details, we refer again to [Oksendal1998].

C.4.3 Extension of the ITÔ integral to n dimensions

With definition C.7, we have established the 1-dimensional version of the ITÔ integral for integrands from the class \mathcal{V} , in whose definition C.2 we now relax the measurability condition

- ii) $X_t(\omega)$ is \mathcal{F}_t -adapted

to

- ii)' There exists an increasing family of σ -algebras $\{\mathcal{H}_t\}_{t \geq 0}$, such that
- W_t is a martingale with respect to \mathcal{H}_t
 - X_t is \mathcal{H}_t -adapted.

This relaxation is (especially) important if the integrand and/or the differential of the ITÔ integral contain coordinate processes of a multi-dimensional WIENER process. The relaxation implies $\mathcal{F}_t \subset \mathcal{H}_t$ and $\mathbb{E}[W_s - W_t | \mathcal{H}_t] = 0$, which is sufficient to construct the ITÔ integral analogously for the extended class of integrands. For a more detailed description we refer, e.g., to the textbook of ØKSENDAL [Øksendal1998].

Before defining the multi-dimensional ITÔ integral, we introduce a notation for n -dimensional BROWNIAN motion and a class collecting the integrands described above:

C.13 Definition (Notation for n -dimensional BROWNIAN motion, class $\mathcal{V}_{\mathcal{H}}^{m \times n}$)

- The set of processes fulfilling the relaxed formulation above, i.e. conditions i) and iii) of definition C.2 and condition ii)' above, for a given filtration $\mathcal{H} = \{\mathcal{H}_t\}_{t \geq 0}$, is denoted as $\mathcal{V}'_{\mathcal{H}}$.
- Let $W : [0, \infty) \times \Omega \rightarrow \mathbb{R}^n$ be an n -dimensional BROWNIAN motion. Then, we denote its k -th coordinate function by $W_k(t, \omega)$, and write $W = (W_1, W_2, \dots, W_n)^T$, omitting the arguments t and ω .
- By $\mathcal{F}_t^{(n)}$ we define the σ -algebra generated by $W_1(s_1, \cdot), W_2(s_2, \cdot), \dots, W_n(s_n, \cdot)$ with $s_k \leq t$.
- Let $v_{ij} \in \mathcal{V}'_{\mathcal{H}}$ ($i = 1, \dots, m; j = 1, \dots, n$). Then, we define $\mathcal{V}_{\mathcal{H}}^{m \times n}(0, T)$ as the set of $m \times n$ matrices $V = [v_{ij}]_{\substack{i=1, \dots, m \\ j=1, \dots, n}}$. □

C.14 Definition (The multi-dimensional ITÔ integral)

Let W be an n -dimensional BROWNIAN motion. For $v \in \mathcal{V}_{\mathcal{H}}^{m \times n}$, we define the multi-dimensional ITÔ integral as the m column vector

$$\int_0^T v dW = \int_0^T \begin{bmatrix} v_{11} & \cdots & v_{1n} \\ \vdots & \ddots & \vdots \\ v_{m1} & \cdots & v_{mn} \end{bmatrix} \begin{bmatrix} dW_1 \\ \vdots \\ dW_n \end{bmatrix} = \begin{bmatrix} \sum_{j=1}^n \int_0^T v_{1j}(s, \omega) dW_j(s, \omega) \\ \vdots \\ \sum_{j=1}^n \int_0^T v_{mj}(s, \omega) dW_j(s, \omega) \end{bmatrix}$$

in which each component is a sum of 1-dimensional ITÔ integrals. □

S Software Package :sfit

In this section, we give some basic information about the implemented software. We first present the GAUSS-NEWTON solver :gnsolve that is based on the methods of chapter 1, and give a short description of the ODE solver suite :clradau.

The final part of this section describes the software package :sfit, a framework that allows the user to formulate the constrained nonlinear stochastic parameter estimation problems from chapter 5 in a convenient way. Basic usage is demonstrated using the example of the calcium ion oscillator in section 6.2. Further, we give the complete code for the WIENR exponential example in section 5.4.4 that easily fits on a single page.

S.1 GAUSS-NEWTON solver :gnsolve

The GAUSS-NEWTON method presented in chapter 1 has been implemented in the COMMON LISP package :gnsolve. Inequality constraints are treated with an active-set method; simple bounds are treated more effectively with a projection based strategy.

The class gnsolve::lsq-prob forms the basis class for least squares problems, with specializations for nonlinear and linear least squares problems, unconstrained, equality constrained, inequality constrained, or both. The class gnsolve::lsq-sol abstracts a solution to a least square problem.

As an example, the linear constrained least squares problem

$$\begin{aligned} \min_x \quad & \frac{1}{2} \|Cx - d\|^2 \\ \text{s.t.} \quad & Ex - b = 0 \end{aligned} \quad \text{with} \quad C = \begin{bmatrix} 1 & 2 & 0 \\ 4 & 0 & 6 \\ 7 & 0 & 0 \\ 1 & 1 & 1 \\ 0 & 4 & 5 \end{bmatrix}, d = \begin{pmatrix} 5.11 \\ 5.22 \\ 5.33 \\ 5.44 \\ 5.55 \end{pmatrix}, E = \begin{bmatrix} 0 & 2 & 0 \\ 3 & 4 & 0 \end{bmatrix}, b = \begin{pmatrix} 21.11 \\ 22.22 \end{pmatrix},$$

may be instantiated, solved (using the default solver), and printed by

```
(let* ((C #2a((1 2 0) (4 0 6) (7 0 0) (1 1 1) (0 4 5))) ;; least squares matrix
      (d #(5.11d0 5.22d0 5.33d0 5.44d0 5.55d0)) ;; least squares rhs
      (E #2a((0 2 0) (3 4 0))) ;; constraint matrix
      (b #(21.11d0 22.22d0)) ;; constraint rhs
      (lsqprob (make-instance 'lin-lsq-ec-prob :C C :d d :E E :b b :name "Txp"))
      (lsqsol (solve-lin-lsq lsqprob))) ;; solve
      (present-solution lsqsol))) ;; display solution
```

giving the output

```
Solution Txp-SOLUTION, generated at 2016-08-01-12-30-20-T4056436
of problem Txp, generated at 2016-08-01-12-30-20-T4056434:
Solved by solver #<FUNCTION SOLVE-LIN-LSQ-EC-DGGLSE>
Contents of optionset: [Type: NULL]
Solution vector: (first 3 of 3 components)
x( 0) = -6.666666666666666d0
x( 1) = 10.5550000000000001d0
x( 2) = 0.1535752688172103d0
Residuals: (first 5 of 5 components)
r( 0) = -9.3333333333333336d0
r( 1) = 30.9652150537634d0
r( 2) = 51.996666666666666d0
r( 3) = 1.3980913978494547d0
r( 4) = -37.437876344086064d0
Multipliers: (first 2 of 2 components)
μ( 0) = 403.44489695340496d0
μ( 1) = -159.9674283154121d0
Contents of optionset: [Type: NULL]
```

Linear least squares

For linear least squares problems, different solvers are provided. Inequalities are treated by an active-set strategy. For the common case with full-rank system matrix, a nullspace method based on QR decompositions has been implemented. For (possibly) rank-deficient problems, the generalized inverse can be computed (see section 1.3.3). Also, solvers based on the LAPACK routines `dgelss` and `dggls` computing minimum norm solutions are provided.

Sparse problems

For sparse equality constrained problems, a binding to the sparse direct solver SUPERLU⁽¹⁾ exists. The solution to the problem $\min \frac{1}{2} \|Cx - d\|^2$ s.t. $Ex - b = 0$ is thereby computed by solving the sparse linear system

$$\begin{bmatrix} 0 & 0 & E \\ 0 & I & C \\ -E^T & -C^T & 0 \end{bmatrix} \begin{pmatrix} \lambda \\ r \\ x \end{pmatrix} = \begin{pmatrix} b \\ d \\ 0 \end{pmatrix}$$

delivering the solution vector x , the Lagrangian multipliers λ and the residual vector $r = d - Cx$ simultaneously.

Nonlinear least squares

As presented in chapter 3, nonlinear least squares problems are solved in an iterative procedure, in which an initial guess x_0 is successively updated by the solutions Δx_k (search directions) of a series of linearized problems. This applies directly to unconstrained and equality constrained problems. Problems with inequality constraints are reduced to equality constrained problems using an active-set strategy.

Globalization

To ensure global convergence, both a backtracking line search as well as the *restrictive monotonicity test* (RMT, see section 1.3.4) are available.

S.2 Integrator :clradau

The integrator `:clradau` was developed and implemented in close collaboration with Mario S. MOMMER⁽²⁾ at the Interdisciplinary Center for Scientific Computing (IWR), Heidelberg University.

`:clradau` is a solver suite for stiff and non-stiff ODE and index-1 DAE systems, providing RUNGE-KUTTA based RADAU-IIa integration schemes⁽³⁾ up to order 17 with variable (error-controlled) step sizes.

Exact derivatives up to (in principle) arbitrary order are efficiently computed by algorithmic differentiation⁽⁴⁾ (AD) and freezing the adaptive parts (*internal numerical differentiation*, see section 2.2.3.2).

⁽¹⁾James W. Demmel, Stanley C. Eisenstat, John R. Gilbert, Xiaoye S. Li, and Joseph W. H. Liu: “A supernodal approach to sparse partial pivoting”. *SIAM J. Matrix Analysis and Applications* 20.3 (1999), pp. 720–755

⁽²⁾now: Modellierung und Systemoptimierung Mommer GmbH, Stettiner Straße 48, 69502 Hemsbach
<http://www.msmommer.de>

⁽³⁾In principle, every method that can be represented in a BUTCHER array is supported; besides the mentioned RADAU schemes, also a set of order 4 LOBATTO IIIa/b/c integration schemes is directly available and others may be upgraded easily.

⁽⁴⁾Based on a suggestion of FATEMAN [Fateman2006], we implemented automatic differentiation in COMMON LISP and extended it for arbitrarily high and mixed derivatives.

S.3 Framework :sfit

The software package `:sfit` implements the methods for parameter estimation in stochastic systems presented in chapter 5. It offers easy symbolic problem formulation to the user, from which it generates stochastic parameter estimation problems as described in problem 5.18 and problem 5.22 (the latter with per-interval parameters). Bindings to the GAUSS-NEWTON package `:gnsolve` allow efficient solution of these problems. Further, for simulation and theoretical studies, it offers a framework for artificial data generation using an EULER-MARUYAMA stochastic integrator.

S.3.1 Formulating an sfit-problem: Calcium oscillator example

We here provide some presentation on how to formulate a stochastic parameter estimation problem in `:sfit`, on the example of the calcium oscillator system studied in section 6.2.

The ODE system for the calcium oscillator was already given in eq. (6.5) on page 175 and is repeated here for convenience:

$$\begin{aligned}\frac{dG_\alpha}{dt} &= p_1 + p_2 \cdot G_\alpha - \frac{p_3 \cdot G_\alpha \cdot PLC}{p_4 + G_\alpha} - \frac{p_5 \cdot G_\alpha \cdot Ca}{p_6 + G_\alpha} \\ \frac{dPLC}{dt} &= p_7 \cdot G_\alpha - \frac{p_8 \cdot PLC}{p_9 + PLC} \\ \frac{dCa}{dt} &= p_{10} \cdot G_\alpha - \frac{p_{11} \cdot Ca}{p_{12} + Ca}\end{aligned}\tag{6.5}$$

S.3.1.1 Symbolic representation: states, parameters, and estimables

The macro `sfit::make-rhs-code` generates code representing the right hand side function. Its calling follows the rule

```
(sfit::make-rhs-code :states states-list
                    :parameters parameters-list
                    :rhs rhs-declaration
                    :result result-vector)
```

The lists `states-list` and `parameters-list` contain pairs of symbols and their default values used in the model; the former collecting the state variables, the latter the model parameters. Their form is `'((symbol1 value1) (symbol2 value2) ...)`, see the example below, and also the notes on *estimables* below.

The `rhs-declaration` is a list of intermediate quantities used in the model's right hand side function, each following the scheme `(quantity = (s-expression))`.

These intermediate quantities may be accessed in the `results-vector`, which returns the actual right-hand-side vector. The order of the result vector's components must correspond to the order of the states in `state-list`.

As a convention, the user is encouraged to store the `states-list` and `parameters-list`, and also the *estimables* discussed below, in the special variables⁽⁵⁾ `*states*`, `*parameters*`, and `*estimables*`.

We note that this symbolic formulation has no negative impacts on speed, as the generated code is compiled to machine code with execution speed comparative to native C code.

⁽⁵⁾ In COMMON LISP, a *special variable* defined with the `defparameter` macro may be seen as a global variable. We note that `:sfit` does not access these special variables; we use them here for convenient notation.

Of course, the user is free to formulate his own right-hand-side code without using the interface provided by `sfit::make-rhs-code`, as long as his code generates a function with signature

$$(\text{lambda } (\text{time states parameters}) \dots (\text{values rhs-vector})) \quad (\text{S.1})$$

taking the current time and the vectors of current state and parameters, and returning the vector of the ODE's right hand side⁽⁶⁾. This calling convention is generic in `:sfit` and `:clradau` and used frequently.

We note that `:sfit` is prepared to support different models on the individual shooting intervals, though that will not be discussed here.

Selecting the set of parameters to estimate: *estimables*

We shortly introduce here the concept behind *estimables*. Frequently, during parameter estimation, one does not know a priori about the identifiability of model parameters from the available measurement data (see also section 2.1.4). Also, if one has gotten additional information about parameter values, one might be interested in estimating only a subset of the model parameters, while the other shall stay at some fixed (known) value. If, in a later stage, the user would like to estimate a different set of parameters, he only has to adjust the set of *estimables* without any further modification of the model.

Calcium oscillator example: state and parameter symbols

We first store the symbols for the states G_α , PLC, and Ca and the symbols for the parameters p_1 to p_{12} in two lists bound to the special variables `*states*` and `*parameters*`, consisting of pairs of a symbol that is used in the right hand side and an associated value that is used as default.

The position in the list corresponds to the component index in the respective vectors. The user, however, may access the components simply by their name. Any admissible LISP symbol may be used as name.

```
(defparameter *states*          ;; 3 states with initial concentrations
  '((G 0.01d0) (PLC 0.01d0) (CA 0.01d0)))
(defparameter *parameters*    ;; 12 kinetic parameters with default values
  '((P1 0.106d0) (P2 1.475d0) (P3 0.76d0) (P4 0.095d0) (P5 2.44d0)
    (P6 0.59d0) (P7 0.62d0) (P8 16.12d0) (P9 14.545d0) (P10 6.79d0)
    (P11 76.5d0) (P12 0.08d0)))
```

In `:sfit`, the user explicitly specifies the (sub)set of parameters to be estimated. In this calcium oscillator example, say, we are only interested in estimating the five parameters p_7, p_9, p_{10}, p_{11} , and p_{12} of eq. (6.5), and thus store their symbols in the special variable `*estimables*`:

```
(defparameter *estimables*    ;; model parameters that shall be estimated
  '(P7 P9 P10 P11 P12))      ;; others remain on their default values
```

⁽⁶⁾ Technically, the user must provide *code* for the right hand side function, as the symbols will be *interned* in the generic arithmetic package `:GA`, and subsequently used in the algorithmic differentiation package `:DFC` for automatic derivative generation. The mentioned packages are part of the integrator suite `:clradau`.

Calcium oscillator example: Generating r.h.s. code with `sfit::make-rhs-code`

Having defined the sets of state variables and parameters, we can now formulate the ODE system given in eq. (6.5) using the macro `sfit::make-rhs-code`

```
(defparameter *ode-code*
  (sfit::make-rhs-code
   :states *states*
   :parameters *parameters*
   :rhs ((dG   = (- (+ p1 (* p2 G))
                    (/ (* p3 G PLC) (+ G p4))
                    (/ (* p5 G Ca) (+ G p6))))
        (dPLC = (- (* p7 G)
                    (/ (* p8 PLC) (+ p9 PLC))))
        (dCa  = (- (* p10 G)
                    (/ (* p11 Ca) (+ p12 Ca))))
   :result (vector dG dPLC dCa)))
```

The above call to `sfit::make-rhs-code` generates the following code:

```
(LAMBDA (TIME STATE PARAMETERS)
  (DECLARE (IGNORABLE TIME STATE PARAMETERS))
  (LET ((P1 (ELT PARAMETERS 0))
        (P2 (ELT PARAMETERS 1))
        (P3 (ELT PARAMETERS 2))
        (P4 (ELT PARAMETERS 3))
        (P5 (ELT PARAMETERS 4))
        (P6 (ELT PARAMETERS 5))
        (P7 (ELT PARAMETERS 6))
        (P8 (ELT PARAMETERS 7))
        (P9 (ELT PARAMETERS 8))
        (P10 (ELT PARAMETERS 9))
        (P11 (ELT PARAMETERS 10))
        (P12 (ELT PARAMETERS 11))
        (G (ELT STATE 0))
        (PLC (ELT STATE 1))
        (CA (ELT STATE 2)))
    (DECLARE (IGNORABLE P1 P2 P3 P4 P5 P6 P7 P8 P9 P10 P11 P12 G PLC CA))
    (LET* ((DG (- (+ P1 (* P2 G)) (/ (* P3 G PLC) (+ G P4))
                  (/ (* P5 G CA) (+ G P6))))
           (DPLC (- (* P7 G) (/ (* P8 PLC) (+ P9 PLC))))
           (DCA (- (* P10 G) (/ (* P11 CA) (+ P12 CA))))
           (VECTOR DG DPLC DCA))))
```

S.3.1.2 Evaluation functions in package `:evalfuncs`

The class `evalfuncs::evalfunc` provides the basic function evaluation mechanism in `:sfit`. An `evalfunc` object is instantiated by

```
(make-instance 'sfit::evalfunc
  :funccode evalf-code
  :times    evalf-times
  :type     evalf-type)
```

In `evalf-code`, the `code` of the evaluation function has to be specified as ordinary LISP code, The code is compiled and the resulting function is bound to the slot `'func`, for which the reader function `evalfuncs::func` is available. The compiled function must return a *vector*; scalar values must be wrapped.

By *evalf-times*, a list of time points when to evaluate the function is specified. Evaluation functions are called at the specified time points during integration, and they are given the current time, current state vector, and the vector of parameters, with the same calling convention as in eq. (S.1).

The *evalf-times* argument specifies the kind of evaluation function. If the evaluation function shall be used as a measurement function in :sfit, the type :measurement is appropriate. Predefined types are:

- :MEASUREMENT - Represents *measurement functions* as described in section 2.1.2.
- :EQUALITY - Evaluation function type for equality constraints.
- :INEQUALITY - Evaluation function type for inequality constraints.
- :NODEEVAL - This evaluation function type is used for function evaluation at the shooting nodes at both the right and left limits, and is internally used to calculate the stochastic jumps α_k (see definition 5.17).
- :SIMULATION - Evaluation functions of this type may be used for visualization purposes.
- :PARAMEVAL - Reserved type, internally used.

The user is free to define and use additional types for special purposes; they are evaluated *en passant* during integration, and do not interfere with the parameter estimation.

Ready-to-use evaluation function objects may be **automatically created** by the :sfit package, if the package :measurement-handler is used for measurement management (discussed below).

Calcium oscillator example: evaluation function

In the calcium oscillator example, *full state measurements* of the system are taken at time points $\{0, 0.25, 0.5, \dots, 60\}$, i.e. one sample every 0.25 time units over the whole interval. An adequate evaluation function object may be created by

```
(defparameter *evalf*
  (make-instance 'sfit::evalfunc
    :funcrcode '(lambda (time states parameters)
                 "Full state evaluator" states)
    :type      :MEASUREMENT
    :times     (loop for ti from 0.0d0 upto 60.0d0 by 0.25d0
                     collecting ti)))
```

The string "Full state evaluator" is a user-chosen description of the evaluation function.

S.3.1.3 Convenient handling of measurements with the :measurement-handler

The management of measurements (of any kind) is simplified by using the interface provided by the :measurement-handler package (or :mhandler for short). A measurement handler object may be instantiated by

```
(defparameter *mhandler*
  (make-instance 'mhandler::measurement-handler
    :name "Calcium Oscillator measurements"
    :measfuncrcode (list '(lambda (time states parameters)
                          "Full state evaluator" states))))
```

Here, as for the evaluation functions, the keyword `:name` is used to specify the name of the measurement handler, and the argument to the `:measfuncscodes` keyword is a *list* of codes describing the individual measurement functions. From this list, evaluation functions will be automatically created by `:sfit`.

In the calcium oscillator example, there is only a single measurement function returning the full state vector.

Compiled functions are, after instantiation, available in the `measfuncs` slot of the measurement handler object, for which a reader function `mhandler::measfuncs` is available. Thus, the first measurement function may be accessed as `(first (measfuncs *mhandler*))`, the second one by `(second (measfuncs *mhandler*))`, and so on. The position in the `measfuncs` list corresponds to the order of the code definitions in the `measfuncscodes` slot.

Using the `mhandler::add-measurement` method, measurement data may be added to a measurement handler object according to the following call:

```
(add-measurement *mhandler* time measf data variance type)
```

After all data has been added to the measurement-handler, an update of the internal structures must be triggered by invoking the method `mhandler::consolidate-mhandler`.

Calcium oscillator example: adding data to a measurement handler

If, e.g., in an actual or simulated experiment the vector `#(8.981 16.55 0.6349)` was measured using the first measurement function at time 7.5, with associated variances `#(0.1 0.1 0.01)` describing the accuracy of these measurements, the following line of code would be appropriate:

```
(add-measurement *mhandler*
 7.5 (first (measfuncs *mhandler*))
      #(8.981 16.55 0.6349) #(0.1 0.1 1.0) :measurement)
(consolidate-mhandler *mhandler*)
```

The last line of code must be done after all data has been added to update internal structures.

We note that the user is not required to use a `measurement-handler`. `:sfit` itself uses a single vector with all measurement data, and a second vector with the associated variances, stored in the `measurements` and `variances` slots of an `sfit-problem` object, that can be automatically built from a measurement handler, together with the corresponding evaluation functions.

The measurement handler is a convenient tool for managing measurements; it offers way more data selection and processing methods than described here.

S.3.1.4 Simple bounds on estimable parameters and shooting variables

`:sfit` offers an easy way to formulate simple bounds on the estimables and on the state variables at the shooting nodes on a *per interval* basis. This is done by a collection (list) of bound-specifiers of type

```
(symbol interval-index lower-bound upper-bound)
```

where (1) `symbol` specifies the respective parameter or state symbol, i.e. a symbol present in `*states*` or `*estimables*`, or T for all, (2) `interval-index` is the index (number) of a shooting interval or T for all, (3) `lower-bound` is the lower bound value and (4) `upper-bound` the upper bound value for the specified symbol.

More special bound definitions overwrite more general ones. If no node initializer function is specified, the default values (set up in the `*states*` variable above) are used for initialization. As an example: `'(X1 3 0.0d0 10.0d0)` represents the requirement that the amount of the species X1 at the fourth shooting node (zero-based indexing!) must be between zero and ten.

Calcium oscillator example: specifying simple bounds

For our example here, we give a common bound for all state and parameter values on all intervals, and specify some individual bounds for the estimable parameters.

```
(defparameter *bounds* ;; simple bounds
  '( (t t 0.0d0 50.0d0) ;; common bounds
    (p7 t 0.1d0 10.0d0) ;; true value: 1.24
    (p9 t 1.0d0 100.0d0) ;; true value: 29.09
    (p10 t 1.0d0 50.0d0) ;; true value: 13.58
    (p11 t 10.0d0 500.0d0) ;; true value: 153.0
    (p12 t 0.01d0 10.0d0))) ;; true value: 0.16
```

S.3.1.5 Shooting node variables initialization

By a *node initializer function*, the initial values of the variables s_i at the shooting nodes can be set. It has the form

```
(lambda (interval time state-symbol sfitp) ... (values inival))
```

where `interval` and `time` denote the number of the shooting interval and the associated time-point for which an initial value is requested, the variable `state-symbol` contains the symbol of the state variable that is to be initialized, and `sfitp` is the `sfit-problem` object. The function must return a valid initialization value `inival`.

Calcium oscillator example: shooting node initialization

Assuming the measurement data is available in a measurement-handler object bound to the slot `mhandler` of the `sfit-problem`, a typical node initializer function that initializes the shooting node variables with the most proximate measurement data looks like

```
(defparameter *node-initializer*
  (lambda (intvl time statesymb sfitp)
    "Determine a value for the specified state symbol on specified time"
    (let* ((mhandler (sfit::mhandler sfitp)) ;; use sfitp's mhandler
          (measstruct (mhandler::get-measurement-nearby mhandler time
                                                         :restrict-to-type :measurement))
          (data (mhandler::get-data measstruct)))
      (ecase statesymb
        (G (elt data 0))
        (PLC (elt data 1))
        (CA (elt data 2))))))
```

S.3.1.6 Setting up the shooting grid

Instead of a set of grid points, a list of individual *integration intervals*

```
((start1 end1) (start2 end2) ... (startn endn))
```

is given to :sfit. This allows non-contiguous integration domains, e.g. for problems with longer unobserved time ranges.

Calcium oscillator example: shooting grid initialization

In the calcium oscillator example, the shooting nodes are equidistantly distributed over the whole time domain [0, 60], placed every ten time units, such that the list of intervals may be specified as

```
(defparameter *msintervals*
  (loop for i from 0.0d0 below 60.0d0 by 10.0d0
        collecting (list i (+ i 10.0d0))))
```

giving the nested list of intervals

```
((0.0d0 10.0d0) (10.0d0 20.0d0) (20.0d0 30.0d0) (30.0d0 40.0d0) (40.0d0 50.0d0) (50.0d0 60.0d0))
```

as result.

S.3.1.7 Instantiating the sfit-problem

The class `sfit::sfitp` represents the stochastic parameter estimation problem 5.18, and its variant problem 5.22 with local parameters.

The instantiation and primary configuration is done using the LISP's `make-instance` with the following calling convention

```
(make-instance 'sfit-problem
  :name problem-name
  :model-code code-list
  :model-parameters parameters-list
  :model-estimables estimables-list
  :model-states states-list
  :node-initializer initializer-function
  :msintervals intervals-list
  :bounds bounds-spec
  :mhandler mhandler
  :use-local-parameters lp-flag
  :evalfuncs evalfuncs-list
  :measurements measurements-vector
  :variances variances-vector
  :measweights measweights-vector
  :jumpweights jumpweights-vector
  :userdata userdata
  :msgrid shooting-grid
  :solver-options solver-options)
```

The initialization arguments `problem-name`, `code-list`, `parameters-list`, `estimables-list`, `states-list`, `initializer-function`, `intervals-list`, `bounds-spec`, and `mhandler` correspond to the objects described in the previous sections S.3.1.1 to S.3.1.6 and are sufficient to fully describe an `sfit-problem`; see section S.3.2 for a complete example.

When set to `T`, the `lp-flag` indicates to automatically set up local parameters as in problem 5.22; setting it to `NIL` (default) leads to a unique set of parameters for all shooting intervals as in problem 5.18.

If no `measurement-handler` is used, the `evalfuncs-list` is a list of evaluation functions as described in section S.3.1.2. Then, the `measurements-vector` contains all measurement data as a single vector, as the evaluation of the `evalfuncs` deliver, with the associated variances specified by the initialization keyword argument `variances-vector`, from which the measurement weights are automatically calculated as squared reciprocals. Alternatively, they can be directly specified in `measweights-vector`, in the same order as the measurements. The argument `jumpweights-vector` specifies the jump regularization weights $\omega_{k,l}^2$ of the stochastic jumps $\alpha_{k,l}$ (see definition 5.17). They can be given as a vector of the same length as the full jump vector α , usually $n_x \cdot (n_{ms} - 1)$, or as a scalar value; the default is a jump regularization weight of 1.0 for all jumps.

The user can specify and store arbitrary data via the `userdata` keyword argument.

An alternative way to specifying the shooting intervals by the `:msintervals` is to deliver the `shooting-grid` as a list of shooting nodes, from which the integration intervals will be internally created.

Lastly, by specifying `solver-options` the underlying solver can be configured; the possible options depend on the used solver. The default set of options is available in the special variable `sfit::*sfit-standard-solver-options*`.

The resulting `sfit-problem` can be solved by invoking `sfit::solve-with-gnsolve` on it:

```
(solve-with-gnsolve *sfitp*)
```

Calcium oscillator example: instantiating the `sfit-problem`

With the above definitions from the previous sections, we have all the ingredients together and preparations done to instantiate an `sfit-problem` by

```
(defparameter *sfitp*
  (make-instance 'sfit-problem
    :name "Calcium oscillator"
    :model-code *ode-code*
    :model-parameters *parameters*
    :model-estimables *estimables*
    :model-states *states*
    :node-initializer *node-initializer*
    :msintervals *msintervals*
    :bounds *bounds*
    :mhandler *mhandler*
    :use-local-parameters NIL))
```

The value to the keyword `:use-local-parameter` specifies whether or not to use per-interval parameters (see section 5.5.5). Here, its value is `NIL`, indicating that no local parameters shall be used.

Finally, to start the solution process of the `sfit-problem` using the `:gnsolve` package, we invoke the method `sfit::solve-with-gnsolve` on `*sfitp*`

```
(solve-with-gnsolve *sfitp*)
```

S.3.1.8 Printing information about an `sfit-problem` using `describe-sfitp`

Using the method `sfit::describe-sfitp` on an `sfit-problem` displays information about that problem. The level of details may be chosen as argument to the keyword parameter `:verbosity`. Valid are `:minimum`, `:default`, or `:maximum`.

An example output is given for the WIENER exponential example in the adjacent section.

S.3.2 A complete example: WIENER exponential

Here, we show how to formulate and solve the stochastic parameter estimation problem for the WIENER exponential discussed in section 5.4.4 on page 145. We present the whole code that is necessary, beginning with listing the measurement data. As this is a stand-alone example, we give explicit specifications here (e.g. bounds are not necessary).

```
(in-package :sfit)

;; set the measurement times and data
(defparameter *measurement-times*
  #(0.5d0 0.55d0 1.5d0 1.55d0 2.5d0 2.55d0 3.5d0 3.55d0 4.5d0 4.55d0
    5.5d0 5.55d0 6.5d0 6.55d0 7.5d0 7.55d0 8.5d0 8.55d0 9.5d0 9.55d0))
(defparameter *measurement-data*
  #(0.79d0 1.12d0 4.94d0 5.37d0 -2.09d0 -1.5d0 -1.31d0 -1.82d0 -5.24d0 -5.42d0
    -8.08d0 -7.94d0 -7.4d0 -7.7d0 -12.82d0 -13.6d0 -16.43d0 -16.91d0 -21.32d0 -21.58d0))

;; instantiate the measurement-handler with measurement function code
(defparameter *mhandler*
  (make-instance 'mhandler::measurement-handler
    :name "Wiener Exponential Data"
    :measfuncscode '(lambda (tt ss pp) ; single measurement function
                     (declare (ignore tt pp)) ; no need for time and parameter vector
                     "State Evaluator" ss))) ; return full state vector ss

;; copy data into measurement-handler
(loop
  for ti across *measurement-times*
  for xi across *measurement-data*
  with evalf = (first (mhandler::measfuncs *mhandler*))
  do
    (mhandler::add-measurement *mhandler* ti evalf xi 1.0d0 :measurement)
  finally
    (mhandler::consolidate-mhandler *mhandler*))

;; set up ODE system, estimables and bounds
(defparameter *states* '(X 0.01d0)) ; true: 1.00 ; single state X, initial guess 0.01
(defparameter *parameters* '(p 0.01d0)) ; true: 0.25 ; single parameter p, initial guess 0.01
(defparameter *estimables* '(p)) ; select p to be estimated
(defparameter *ode-code* (sfit::make-rhs-code :states *states*
                                             :parameters *parameters*
                                             :rhs ((dX = (* p X)))
                                             :result (vector dX)))

(defparameter *bounds* '(T T -1.0d10 1.0d10) ; common bounds
  (p t -100.0d0 100.0d0)) ; bounds on parameter p

;; initialization function for shooting nodes
(defparameter *node-initializer*
  (lambda (intvl time statesymb sfitp)
    (declare (ignore intvl))
    "Determine a value for the specified state symbol on specified time/interval"
    (let* ((mhandler (sfit::mhandler sfitp)) ; get the sfitp's mhandler
           (measstruct (mhandler::get-measurement-nearby mhandler time :restrict-to-type :measurement))
           (data (mhandler::get-data measstruct)))
      (ecase statesymb
        (X (elt data 0)))))

;; set the shooting intervals
(defparameter *msintervals*
  (loop for i from 0.0d0 below 10.0d0 by 1.0d0
        collecting (list i (+ i 1.0d0))))

;; initialize the sfit-problem
(defparameter *sfitp*
  (make-instance 'sfit-problem
    :name "Wiener exponential example"
    :model-code *ode-code*
    :model-parameters *parameters*
    :model-estimables *estimables*
    :model-states *states*
    :node-initializer *node-initializer*
    :msintervals *msintervals*
    :bounds *bounds*
    :mhandler *mhandler*
    :allow-other-keys T
    :add-node-evaluators T
    :use-local-parameters NIL))

;; Solve the sfit-problem
(solve-with-gnsolve *sfitp*)
```

The last call to `sfit::solve-with-gnsolve` then produces the following output:

```

INFO: Saving instantiation data
INFO: Processing mhandler...
INFO: ... using mhandler "Wiener Exponential Data" to initialize sfitp "Wiener exponential example"
INFO: ... ensuring no shooting nodes are on measurements
INFO: ... initializing evalfuncs
INFO: Using DFC:COMPILE-DIFFERENTIABLE-LAMBDA to compile evalfunc "State Evaluator"
INFO: ... initializing measurements and variances
INFO: Processing of mhandler "Wiener Exponential Data" done.
INFO: No msyslist given. Building it.
INFO: Initializing shooting nodes using function
      #<FUNCTION (LAMBDA (INTVL TIME STATESYMB SFITP) :IN "wienerexponentialexample.lisp") {1012FB94DB}>
INFO: ... updating msys's
INFO: Adding node evaluators...
INFO: Using DFC:COMPILE-DIFFERENTIABLE-LAMBDA to compile evalfunc "Node eval for shooting nodes"
INFO: Initializing measurement weights from variances
INFO: Initializing VARIABLES vector...
INFO: Initializing variables
INFO: Initializing VARIABLES-SCALING vector...
INFO: ... initializing scaling (*use-scaling* = NIL)
INFO: Processing evalfuncs...
INFO: Initialization of SFIT-problem "Wiener exponential example" done.
INFO: Solving with :GNSOLVE
INFO: ITERATION #000 --- OBJECTIVE: 210.7487      [RESIDUAL-PART: 146.2050      JUMP-PART: 64.54376      ]
INFO: ITERATION #001 --- OBJECTIVE: 25.65407     [RESIDUAL-PART: 8.111939     JUMP-PART: 17.54213     ]
INFO: ITERATION #002 --- OBJECTIVE: 20.73017     [RESIDUAL-PART: 10.21355     JUMP-PART: 10.51662     ]
INFO: ITERATION #003 --- OBJECTIVE: 20.72096     [RESIDUAL-PART: 10.19394     JUMP-PART: 10.52702     ]
INFO: ITERATION #004 --- OBJECTIVE: 20.72092     [RESIDUAL-PART: 10.18772     JUMP-PART: 10.53319     ]
INFO: ITERATION #005 --- OBJECTIVE: 20.72092     [RESIDUAL-PART: 10.18730     JUMP-PART: 10.53361     ]
INFO: ITERATION #006 --- OBJECTIVE: 20.72092     [RESIDUAL-PART: 10.18727     JUMP-PART: 10.53364     ]
INFO: :GNSOLVE finished with status: SUCCESS
Contents of optionset [Type: CONS]:
KEY = ITERATIONS          VAL = 7
KEY = LAST-PK-2NORM      VAL = 1.5835941834291773d-6
KEY = LAST-PK-1NFNORM    VAL = 7.728591393232885d-7
KEY = MESSAGE             VAL = Apparantly converged!
KEY = OBJ-VALUE          VAL = 20.720917052698397d0
KEY = RESIDUAL-2NORM     VAL = 6.437533231401357d0
KEY = STATUS              VAL = SUCCESS
KEY = SUCCESS             VAL = SUCCESS
KEY = TERMINATION-TESTS  VAL = (MINIMUM-INCREMENT-NORM)
KEY = TIMING-REALTIME    VAL = 0.181d0
KEY = TIMING-RUNTIME     VAL = 0.364d0
Solution vector::
vector, length: 11, element-type: DOUBLE-FLOAT, lisp-type: (SIMPLE-ARRAY DOUBLE-FLOAT (11))
  1.183    2.316   -0.6453   -1.845   -4.215   -6.203   -7.342   -10.84   -14.21   -18.58    0.2896

```

The solution vector, printed on the last line, consists of the 10 initial state values s_i at the shooting nodes and the last value gives the estimate for the exponential parameter. The true initial value is $x^*(0) = 1.0$ (estimated as 1.183, initial guess was 0.01); the true parameter value is $p^* = 0.25$ (estimated as 0.2896, initial guess was 0.01). See table 5.1 on page 146 in section 5.4.4 for a detailed discussion.

We shortly mention that much more informational output, especially during the solution process, is available and may be triggered by setting the respective output flags (documented in the code).

Printing information about the sfitt-problem

Invoking the method `sfitt::describe-sfitt` on an `sfitt-problem` displays information about it, as mentioned in section S.3.1.8.

We reprint here the information given for the above WIENER exponential example when maximum detail level is requested:

```
(describe-sfitt *sfitt* :verbosity :maximum)
```

This information is highly beneficial during problem set-up and for debugging purposes.

```
Displaying stats of sfitt-problem "Wiener exponential example" - created 2016-08-09-10-45-11-T256123
# SHOOTING-INTERVALS : 10
SHOOTING-INTERVALS : ((0.0d0 1.0d0) (1.0d0 2.0d0) (2.0d0 3.0d0) (3.0d0 4.0d0) (4.0d0 5.0d0) (5.0d0 6.0d0)
(6.0d0 7.0d0) (7.0d0 8.0d0) (8.0d0 9.0d0) (9.0d0 10.0d0))
# SHOOTING-NODES : 11
SHOOTING-GRID : (0.0d0 1.0d0 2.0d0 3.0d0 4.0d0 5.0d0 6.0d0 7.0d0 8.0d0 9.0d0 10.0d0)
VARIABLES-INIT : #(0.79d0 1.12d0 5.37d0 -1.5d0 -1.82d0 -5.42d0 -7.94d0 -7.7d0 -13.6d0 -16.91d0 0.01d0)
MHANDLER : #<MEASUREMENT-HANDLER {10143BEC73}>
# MEASUREMENTS : 20
# MEAS-VARIANCES : 20
NODE-Initializer : #<FUNCTION (LAMBDA (INTVL TIME STATESYMB SFITP)
:IN "wienerexponentialexample.lisp") {10143C021E}>
# VARIABLES : 11
--> # STATE-VARS : 10
--> # PARAMETERS : 1
LOCAL PARAMETERS : NIL
MODEL-CODE given : T
# EVALFUNCS : 2
# EVALFUNCTYPES : 2
SOLVER-OPTIONS : PRESENT

Information about the model (on availability):
MODEL-CODE : (LAMBDA (TIME STATE PARAMETERS)
(DECLARE (IGNORABLE TIME STATE PARAMETERS))
(LET ((P (ELT PARAMETERS 0)) (X (ELT STATE 0)))
(DECLARE (IGNORABLE P X))
(LET* ((DX (* P X)))
(VECTOR DX))))
STATE SYMBS : (X)
PARAMETER SYMBS : (P)
ESTIMABLE SYMBS : (P)
NON-ESTIM.PARAM.VALS : NIL
```

```
Description of the VARIABLES vector (value --> list-of-symbols-and-intervals)
Entry # 0: :STATE val= 1.1834 --> ((0 X))
Entry # 1: :STATE val= 2.3156 --> ((1 X))
Entry # 2: :STATE val= -0.64532 --> ((2 X))
Entry # 3: :STATE val= -1.8453 --> ((3 X))
Entry # 4: :STATE val= -4.2150 --> ((4 X))
Entry # 5: :STATE val= -6.2029 --> ((5 X))
Entry # 6: :STATE val= -7.3418 --> ((6 X))
Entry # 7: :STATE val=-10.838 --> ((7 X))
Entry # 8: :STATE val=-14.214 --> ((8 X))
Entry # 9: :STATE val=-18.576 --> ((9 X))
Entry # 10: :PARAMETER val= 0.28957 --> ((0 P) (1 P) (2 P) (3 P) (4 P) (5 P) (6 P) (7 P) (8 P) (9 P))
```

Details about the underlying SHOOTING NODES INITIALIZATION:

```
Total number of systems: 10
System #0 on interval [0.0d0, 1.0d0]: (X 0.79d0)
System #1 on interval [1.0d0, 2.0d0]: (X 1.12d0)
System #2 on interval [2.0d0, 3.0d0]: (X 5.37d0)
System #3 on interval [3.0d0, 4.0d0]: (X -1.5d0)
System #4 on interval [4.0d0, 5.0d0]: (X -1.82d0)
System #5 on interval [5.0d0, 6.0d0]: (X -5.42d0)
System #6 on interval [6.0d0, 7.0d0]: (X -7.94d0)
System #7 on interval [7.0d0, 8.0d0]: (X -7.7d0)
System #8 on interval [8.0d0, 9.0d0]: (X -13.6d0)
System #9 on interval [9.0d0, 10.0d0]: (X -16.91d0)
```

Information about the evalfuncs:

Evalfuncs: 2

Evalfunc #0

ID: #<EVALFUNC {101443B5E3}>

Name: State Evaluator

Func: #<FUNCTION {101444F09B}>

Type: MEASUREMENT

Dim: NIL

Evals: 20

Times: #(0.5d0 0.55d0 1.5d0 1.55d0 2.5d0 2.55d0 3.5d0 3.55d0 4.5d0 4.55d0
5.5d0 5.55d0 6.5d0 6.55d0 7.5d0 7.55d0 8.5d0 8.55d0 9.5d0 9.55d0)

Evalfunc #1

ID: #<EVALFUNC {101443B603}>

Name: Node eval for shooting nodes

Func: #<FUNCTION {101444F11B}>

Type: NODEEVAL

Dim: NIL

Evals: 9

Times: #(1.0d0 2.0d0 3.0d0 4.0d0 5.0d0 6.0d0 7.0d0 8.0d0 9.0d0)

List of Tables

2.1	Example on automatic differentiation.	34
3.1	Individual and shared components in JAK-STAT signalling.	48
3.2	Abbreviations of experimental settings.	55
3.3	Time course data of pSTAT-3 upon four different stimulation settings.	57
3.4	Measurements of recombinant STAT-3	66
3.5	Statistical analysis of pSTAT-3 peak levels.	78
3.6	List of model species	82
3.7	Reactions in the crosstalk model.	83
3.8	ODE model equations.	84
3.9	Initial values for the crosstalk model.	85
3.10	Kinetic parameters used for SOCS-3 prediction.	87
3.11	Time course data of cytosolic SOCS-3 upon IL-6 stimulation.	88
3.12	Final estimates of kinetic parameters.	90
3.13	Phospho-STAT-3 peak levels in measurements, preliminary, and final calibrated model.	91
5.1	Estimation results for the WIENER exponential with and without jump regularization	146
6.1	Estimation test series: FITZHUGH-NAGUMO oscillator.	171
6.2	Estimation test series: Calcium oscillator (full observations).	181
6.3	Estimation test series: Calcium oscillator (partial observations).	182
6.4	Estimation results with intermittent observations.	185
6.5	Estimation test series: BISTABAER allosteric regulation model.	191
6.6	Improved estimates by adjusting grid and jump regularization weights.	194
6.7	Estimation test series: ORNSTEIN-UHLENBECK process.	202
A.1	Time course data upon four different stimulation settings.	207
A.2	Final estimates of parameters in the crosstalk model – non-rounded values.	208

List of Figures

2.1	Distribution of female body size in Germany	23
2.2	Solving boundary value problems by multiple shooting.	30
3.1	Biological assembly image of human IL-6 and GM-CSF	43
3.2	3D organotypic co-culture model as stromal equivalent.	44
3.3	IL-6 facilitates tumour progression.	45
3.4	Simplified IL-6-induced JAK-STAT signalling pathway.	46
3.5	Proliferation of HaCaT in different stimulation scenarios	47
3.6	Sketch of gel eletrophoresis.	53
3.7	Randomization breaks time correlation in errors.	54
3.8	BLU data of cytosolic pSTAT-3 measurements in four experimental settings.	58
3.9	Curve separating between uni- and bimodal density of $T = (a + X)/(b + Y)$	61
3.10	Densities of $T = (a + X)/(b + Y)$	62
3.11	Possible destructive effects of normalization.	67
3.12	Distribution of additive error ϵ_a	70
3.13	Distribution of multiplicative bias ϵ_m	72
3.14	Distribution of simulated amplification factors.	73
3.15	Raw and time-transformed artificial signal.	74
3.16	Normalization effects on distribution of maximum amplification factors and their time point	76
3.17	Measurements of cytosolic pSTAT-3 upon four experimental settings.	78
3.18	JAK-STAT model by STEVEN et al.	80
3.19	The crosstalk model.	81
3.20	Preliminary model calibration using solely pSTAT-3 measurements.	86
3.21	Prediction and measurements of cytosolic SOCS-3 upon IL-6 stimulation.	87
3.22	Model fit of cytosolic SOCS-3 upon IL-6 stimulation.	89
3.23	Final model calibration using pSTAT-3 and SOCS-3 measurements from four experimental settings.	89
3.24	Prediction of time course in co-stimulation experiments.	93
3.25	Predicted time course of activated IL-6/gp130 ligand receptor complex.	93
3.26	Dose dependency on GM-CSF in IL-6-induced phospho-STAT-3 signalling.	96
4.1	Five example realizations of standard WIENER processes	105
4.2	Comparing EULER-MARUYAMA approximations to a strong solution of an S-IVP.	115
4.3	Sample paths of (compound) POISSON processes.	120
5.1	Stochasticity in the BISTABAER model completely alters its behaviour.	122
5.2	Short term similarity of ODE and SDE interpretation of the BISTABAER model.	123
5.3	Illustration of stochastic jumps \mathcal{A}_k	144
5.4	Trajectories of a WIENER exponential.	146

5.5	Fitted exponential trajectory with and without jump regularization	147
5.6	Illustration of interval solutions	150
5.7	Sparsity patterns of the combined residual vector's Jacobian	153
5.8	Examples of sparsity patterns of Jacobian J and decomposition factor R	158
6.1	Time course and phase portrait of the FHN oscillator upon excitation	165
6.2	Steady state sensitivities of the FITZHUGH-NAGUMO oscillator	167
6.3	Simulation of the FITZHUGH-NAGUMO oscillator with a driving WIENER process effective on the first component.	168
6.4	Sparsity pattern of the combined residual vector's Jacobian - FITZHUGH-NAGUMO oscillator	170
6.5	FITZHUGH-NAGUMO example fit	172
6.6	Simulation of the calcium oscillator with a driving WIENER process effective on the PLC component.	176
6.7	Initial trajectory in parameter estimation scenario (G) for a certain realization.	179
6.8	Sparsity pattern of the combined residual vector's Jacobian - Calcium oscillator	180
6.9	Calcium oscillator example fit.	184
6.10	Calcium oscillator fit to intermittent data.	186
6.11	Stochasticity in the BISTABAER model completely alters its behaviour.	187
6.12	Sparsity pattern of the combined residual vector's Jacobian - BISTABAER model	190
6.13	Example fit of the BISTABAER model with exact and noisy measurements	192
6.14	Detail enlargement of a BISTABAER fit	194
6.15	Realisation of an ORNSTEIN-UHLENBECK process with a driving compound POISSON process.	196
6.16	Example of fitted trajectories of an ORNSTEIN-UHLENBECK process.	199
6.17	Sparsity pattern of the combined residual vector's Jacobian - OU model	201

Picture Credits

Figure 3.1a on page 43: *IL-6 (PDB-1ALU)*

RCSB Protein Data Base: <http://www.rcsb.org/pdb/explore.do?structureId=1ALU>

Primary citation: Somers W., Stahl M., Seehra J. S.: *A crystal structure of interleukin 6: implications for a novel mode of receptor dimerization and signaling*. EMBO J. 1997 Mar 3;16(5):989-97.

Figure 3.1b on page 43: *GM-CSF (PDB-2GMF)*

RCSB Protein Data Base: <http://www.rcsb.org/pdb/explore.do?structureId=2GMF>

Primary citation: Rozwarski D. A., Diederichs K., Hecht R., Boone T., Karplus P. A.: *Refined crystal structure and mutagenesis of human granulocyte-macrophage colony-stimulating factor*. Proteins. 1996 Nov;26(3):304-13.

Figure 3.3 on page 45: *IL-6 facilitates tumour progression*

Assembled with H&E stained OTC pictures kindly provided by Margareta M. MÜLLER and co-workers of the group “Tumour and Microenvironment” at the German Cancer Research Center (DKFZ) in Heidelberg (now: Faculty of Mechanical and Medical Engineering, Hochschule Furtwangen University).

Figure 3.5 on page 47: *Proliferation of HaCaT in different stimulation scenarios*

Kindly provided by Margareta M. MÜLLER and co-workers of the group “Tumour and Microenvironment” at the German Cancer Research Center (DKFZ) in Heidelberg (now: Faculty of Mechanical and Medical Engineering, Hochschule Furtwangen University).

Figure 3.9 on page 61: *Curve separating between uni- and bimodal density of $T = (a + X)/(b + Y)$*

Adapted from George Marsaglia: “Ratios of Normal Variables”. *Journal of Statistical Software* 16.4 (May 2006), pp. 1–10. DOI: 10.18637/jss.v016.i04. Figure 4.

Figure 3.10 on page 62: *Densities of $T = (a + X)/(b + Y)$*

Adapted from George Marsaglia: “Ratios of Normal Variables and Ratios of Sums of Uniform Variables”. *Journal of the American Statistical Association* 60.309 (1965), pp. 193–204. DOI: 10.2307/2283145. Figure 1.

Figure 3.18 on page 80: *JAK-STAT model by STEVEN et al.*

Fnu Steven et al.: *Mathematical model of IL6 signal transduction pathway*. 2009. URL: http://homepages.rpi.edu/~hahnj/Models/Updated_IL_6_Model.pdf (visited on 09/23/2014)

Bibliography

- [Albersmeyer2010] Jan Albersmeyer: “Adjoint based algorithms and numerical methods for sensitivity generation and optimization of large scale dynamic systems”. PhD thesis. Universität Heidelberg, 2010. eprint: <http://www.ub.uni-heidelberg.de/archiv/11651>.
- [AlbersmeyerDiehl2010] Jan Albersmeyer and Moritz Diehl: “The Lifted Newton Method and Its Application in Optimization”. *SIAM Journal on Optimization* 20.3 (2010), pp. 1655–1684. DOI: 10.1137/080724885.
- [AsaiKloeden2016] Yusuke Asai and Peter E. Kloeden: “Multi-step methods for random ODEs driven by Itô diffusions”. *Journal of Computational and Applied Mathematics* 294 (2016), pp. 210–224. DOI: 10.1016/j.cam.2015.08.019.
- [Ascher1995] Uri M. Ascher, Robert M. M. Mattheij, and Robert D. Russell: *Numerical solution of boundary value problems for ordinary differential equations*. Classics in applied mathematics 13. SIAM: Philadelphia, 1995.
- [BandiPhillips2003] Federico M. Bandi and Peter C. B. Phillips: “Fully Nonparametric Estimation of Scalar Diffusion Models”. *Econometrica* 71.1 (2003), pp. 241–283. DOI: 10.1111/1468-0262.00395.
- [Bard1974] Yonathan Bard: *Nonlinear parameter estimation*. Academic Press, 1974.
- [Barndorff2001] Ole E. Barndorff-Nielsen, Thomas Mikosch, and Sidney I. Resnick, eds.: *Lévy processes. theory and applications*. Birkhäuser: Berlin, 2001.
- [Beigel2012] Dörte Beigel: “Efficient goal-oriented global error estimation for BDF-type methods using discrete adjoints”. PhD thesis. Universität Heidelberg, 2012. eprint: <http://www.ub.uni-heidelberg.de/archiv/14317>.
- [Bertsekas1995] Dimitri P. Bertsekas: *Nonlinear programming*. Optimization and neural computation series 2. Athena Scientific: Belmont, Mass., 1995.
- [BianchiCleur1996] Carlo Bianchi and Eugene M. Cleur: “Indirect estimation of stochastic differential equation models: some computational experiments”. *Computational Economics* 9.3 (1996), pp. 257–274. DOI: 10.1007/BF00121638.
- [Bichteler2002] Klaus Bichteler: *Stochastic Integration with Jumps*. Encyclopedia of Mathematics and its Applications Bd. 89. Cambridge University Press, 2002.
- [Bjorck1996] Åke Björck: *Numerical Methods for Least Squares Problems*. Society for Industrial and Applied Mathematics, 1996.
- [Bock1981] H. Georg Bock: “Numerical treatment of inverse problems in chemical reaction kinetics”. In: *Modelling of Chemical Reaction Systems*. Ed. by Klaus H. Ebert, Peter Deuffhard, and Willi Jäger. Vol. 18. Springer Series in Chemical Physics. Springer: Heidelberg, 1981, pp. 102–125. DOI: 10.1007/978-3-642-68220-9_8. eprint: <http://www.iwr.uni-heidelberg.de/groups/agbock/FILES/Bock1981.pdf>.
- [Bock1983] H. Georg Bock: “Recent advances in parameter identification techniques for ODE”. In: *Numerical Treatment of Inverse Problems in Differential and Integral Equations*. Ed. by Peter Deuffhard and Ernst Hairer. Birkhäuser: Boston, 1983, pp. 95–121. eprint: <http://www.iwr.uni-heidelberg.de/groups/agbock/FILES/Bock1983.pdf>.
- [Bock1987] H. Georg Bock: *Randwertproblemmethoden zur Parameteridentifizierung in Systemen nichtlinearer Differentialgleichungen*. Vol. 183. Bonner Mathematische Schriften. Universität Bonn, 1987. eprint: <http://www.iwr.uni-heidelberg.de/groups/agbock/FILES/Bock1987.pdf>.

- [Bock2000b] H. Georg Bock, Ekaterina A. Kostina, and Johannes P. Schlöder: “On the Role of Natural Level Functions to Achieve Global Convergence for Damped Newton Methods.” In: *System Modelling and Optimization. Methods, Theory and Applications*. Ed. by Michael J. D. Powell and Stefan Scholtes. Kluwer, 2000, pp. 51–74.
- [Bohl2009] Sebastian Bohl: “Dynamic modeling of signal processing for IL-6-induced STAT3 signal transduction in primary mouse hepatocytes”. PhD thesis. Universität Heidelberg, 2009.
- [Boukamp1988] Petra Boukamp, Rule T. Petrussevska, Dirk Breitkreutz, Jürgen Hornung, Alex Markham, and Norbert E. Fusenig: “Normal keratinization in a spontaneously immortalized aneuploid human keratinocyte cell line”. *J. Cell Biol.* 106.3 (Mar. 1988), pp. 761–771.
- [Bouniakowsky1859] Victor Bouniakowsky: “Sur quelques inégalités concernant les intégrales ordinaires et les intégrales aux différences finies”. *Mémoires de l'Académie Impériale des Sciences de St.-Pétersbourg. 7e série. Tome 1. No. 9* (1859).
- [Brown1828] Robert Brown: “A Brief Account on Microscopical Observations on the Particles Contained in the Pollen of Plants and on the General Existence of Active Molecules in Organic and Inorganic Bodies”. In: *The miscellaneous botanical works of Robert Brown: Volume 1*. Ed. by John J. Bennett. Robert Hardwicke, London, 1828.
- [Buettner2002] R. Buettner, Linda B. Mora, and Richard Jove: “Activated STAT signaling in human tumors provides novel molecular targets for therapeutic intervention”. *Clin. Cancer Res.* 8.4 (Apr. 2002), pp. 945–954.
- [Bulirsch1971] Roland Bulirsch: *Die Mehrzielmethode zur numerischen Lösung von nicht-linearen Randwertproblemen und Aufgaben der optimalen Steuerung*. Tech. rep. Oberpfaffenhofen: Carl-Cranz-Gesellschaft, 1971.
- [Chiba1990] Shigeru Chiba, Arinobu Tojo, Toshio Kitamura, Akio Urabe, Kohei Miyazono, and Fumimaro Takaku: “Characterization and molecular features of the cell surface receptor for human granulocyte-macrophage colony-stimulating factor”. *Leukemia* 4 (Jan. 1990), pp. 29–36.
- [Chung2000] Kai Lai Chung: *A Course in Probability Theory, Revised Edition*. Elsevier Science, 2000.
- [CobelliDiStefano1980] Claudio Cobelli and Joseph J. DiStefano: “Parameter and structural identifiability concepts and ambiguities: a critical review and analysis”. *Am. J. Physiol.* 239.1 (July 1980), pp. 7–24.
- [Cysne2004] Rubens Penha Cysne: “On the statistical estimation of diffusion processes: a partial survey”. *Brazilian review of econometrics* 24.2 (2004), pp. 273–301.
- [Davidenko1953] D. F. Davidenko: “On a new method of numerical solution of systems of nonlinear equations”. *Doklady Akad. Nauk SSSR (N.S.)* 88 (1953), pp. 601–602.
- [deGroot1998] Rolf P. de Groot, Paul J. Coffey, and Leo Koenderman: “Regulation of proliferation, differentiation and survival by the IL-3/IL-5/GM-CSF receptor family”. *Cell. Signal.* 10 (Oct. 1998), pp. 619–628.
- [Demmel1999] James W. Demmel, Stanley C. Eisenstat, John R. Gilbert, Xiaoye S. Li, and Joseph W. H. Liu: “A supernodal approach to sparse partial pivoting”. *SIAM J. Matrix Analysis and Applications* 20.3 (1999), pp. 720–755.
- [Depner2014] Sofia Depner, Wiltrud Lederle, Claudia Gutschalk, Nina Linde, Alexandra Zajonz, and Margareta M. Müller: “Cell type specific interleukin-6 induced responses in tumor keratinocytes and stromal fibroblasts are essential for invasive growth”. *International Journal of Cancer* (2014). DOI: 10.1002/ijc.27951.

- [DereichHeidenreich2011] Steffen Dereich and Felix Heidenreich: “A multilevel Monte Carlo algorithm for Lévy-driven stochastic differential equations”. *Stochastic Processes and their Applications* 121.7 (2011), pp. 1565–1587. DOI: 10.1016/j.spa.2011.03.015.
- [Deuffhard1974] Peter Deuffhard: “A Modified Newton Method for the Solution of Ill-conditioned Systems of Nonlinear Equations with Applications to Multiple Shooting”. *Numerische Mathematik* 22 (1974), pp. 289–311.
- [Deuffhard1975] Peter Deuffhard: “A Relaxation Strategy for the Modified Newton Method”. In: *Optimization and Optimal Control, Lecture Notes in Mathematics*. Ed. by Josef Stoer, Roland Bulirsch, and W. Oettli. Vol. 477. Springer, 1975.
- [DiGianantonioEdalat2013] Pietro Di Gianantonio and Abbas Edalat: “A Language for Differentiable Functions”. English. In: *Foundations of Software Science and Computation Structures*. Ed. by Frank Pfenning. Vol. 7794. Lecture Notes in Computer Science. Springer: Berlin, Heidelberg, 2013, pp. 337–352. DOI: 10.1007/978-3-642-37075-5_22.
- [Doss1977] Halim Doss: “Liens entre équations différentielles stochastiques et ordinaires”. *Annales de l’institut Henri Poincaré (B) Probabilités et Statistiques* 13.2 (1977), pp. 99–125. eprint: <http://eudml.org/doc/77063>.
- [Elerian2001] Ola Elerian, Siddhartha Chib, and Neil Shephard: “Likelihood Inference for Discretely Observed Nonlinear Diffusions”. *Econometrica* 69.4 (2001), pp. 959–993. DOI: 10.1111/1468-0262.00226.
- [Ernst2008] Matthias Ernst, Meri Najdovska, Dianne Grail, Therese Lundgren-May, Michael Buchert, Hazel Tye, Vance B. Matthews, Jane Armes, Prithi S. Bhathal, Norman R. Hughes, Eric G. Marcusson, James G. Karras, Songqing Na, Jonathon D. Sedgwick, Paul J. Hertzog, and Brendan J. Jenkins: “STAT3 and STAT1 mediate IL-11-dependent and inflammation-associated gastric tumorigenesis in gp130 receptor mutant mice”. *J. Clin. Invest.* 118.5 (May 2008), pp. 1727–1738.
- [Fateman2006] Richard Fateman: *Building Algebra Systems by Overloading Lisp: Automatic Differentiation*. 2006. eprint: <http://www.cs.berkeley.edu/~fateman/papers/overload-small.pdf>.
- [FitzHugh1961] Richard FitzHugh: “Impulses and Physiological States in Theoretical Models of Nerve Membrane”. *Biophys. J.* 1.6 (July 1961), pp. 445–466.
- [FreidlinWentzell1998] Mark I. Freidlin and Alexander D. Wentzell: *Random Perturbations of Dynamical Systems*. Grundlehren der mathematischen Wissenschaften. Springer, 1998.
- [Fusenig1998] Norbert E. Fusenig and Petra Boukamp: “Multiple stages and genetic alterations in immortalization, malignant transformation, and tumor progression of human skin keratinocytes”. *Mol. Carcinog.* 23 (Nov. 1998), pp. 144–158.
- [GallantTauchen1996] A. Ronald Gallant and George Tauchen: “Which Moments to Match?”. *Econometric Theory* 12.4 (1996), pp. 657–681.
- [Ghosh2011] Suvojit Ghosh, Subbiah Elankumaran, and Ishwar K. Puri: “Mathematical model of the role of intercellular signalling in intercellular cooperation during tumorigenesis”. *Cell Prolif.* 44 (Apr. 2011), pp. 192–203.
- [Gillespie1976] Daniel T. Gillespie: “A general method for numerically simulating the stochastic time evolution of coupled chemical reactions”. *Journal of Computational Physics* 22.4 (1976), pp. 403–434. DOI: 10.1016/0021-9991(76)90041-3.
- [Gillespie2000] Daniel T. Gillespie: “The chemical Langevin equation”. *The Journal of Chemical Physics* 113.1 (2000), pp. 297–306. DOI: 10.1063/1.481811.

- [GolubVanLoan1996] Gene H. Golub and Charles F. Van Loan: *Matrix Computations*. Johns Hopkins Studies in the Mathematical Sciences. Johns Hopkins University Press, 1996.
- [Gourieroux1993] Christian Gourieroux, Alain Monfort, and Eric Renault: “Indirect inference”. *Journal of Applied Econometrics* 8.S1 (1993), pp. 85–118. DOI: 10.1002/jae.3950080507.
- [GriewankWalther2008] Andreas Griewank and Andrea Walther: *Evaluating Derivatives: Principles and Techniques of Algorithmic Differentiation*. SIAM e-books. Society for Industrial and Applied Mathematics, SIAM, Philadelphia, 2008.
- [Groetzinger1999] Joachim Grötzinger, Thomas Kernebeck, Karl-Josef Kallen, and Stefan Rose-John: “IL-6 type cytokine receptor complexes: hexamer, tetramer or both?” *Biol. Chem.* 380.7-8 (1999), pp. 803–813.
- [Guerriero2009] Maria L. Guerriero, Anna Dudka, Nicholas Underhill-Day, John K. Heath, and Corrado Priami: “Narrative-based computational modelling of the Gp130/JAK/STAT signalling pathway”. *BMC Syst Biol* 3 (2009).
- [Hairer2000] Ernst Hairer, Syvert P. Nørsett, and Gerhard Wanner: *Solving ordinary differential equations I*. 2. rev. ed., 2. corr. print. 2000. Springer series in computational mathematics 8. Springer: Berlin, Heidelberg, 1993.
- [Hansen1982] Lars Peter Hansen: “Large Sample Properties of Generalized Method of Moments Estimators”. *Econometrica* 50.4 (1982), pp. 1029–1054.
- [HansenScheinkman1995] Lars Peter Hansen and José Alexandre Scheinkman: “Back to the Future: Generating Moment Implications for Continuous-Time Markov Processes”. *Econometrica* 63.4 (1995), pp. 767–804.
- [Heinrich1998] Peter C. Heinrich, Iris Behrmann, Gerhard Müller-Newen, Fred Schaper, and Lutz Graeve: “Interleukin-6-type cytokine signalling through the gp130/Jak/STAT pathway”. *Biochem. J.* 334 (Pt 2) (Sept. 1998), pp. 297–314.
- [HodgkinHuxley1952] Alan Lloyd Hodgkin and Andrew Fielding Huxley: “A quantitative description of membrane current and its application to conduction and excitation in nerve”. *J. Physiol. (Lond.)* 117.4 (Aug. 1952), pp. 500–544.
- [Humpherys2012] Jeffrey Humpherys, Preston Redd, and Jeremy West: “A Fresh Look at the Kalman Filter”. *SIAM Review* 54.4 (2012), pp. 801–823. DOI: 10.1137/100799666.
- [Iacus2008] Stefano M. Iacus: *Simulation and inference for stochastic differential equations. With R examples*. Springer series in statistics. Springer: New York, 2008. DOI: 10.1007/978-0-387-75839-8.
- [IkedaWatanabe1981] Nobuyuki Ikeda and Shinzo Watanabe: *Stochastic differential equations and diffusion processes*. Vol. 24. North-Holland Mathematical Library. North-Holland Publishing Co.: Amsterdam, 1981.
- [ImkellerSchmalfuss2001] Peter Imkeller and Björn Schmalfuß: “The Conjugacy of Stochastic and Random Differential Equations and the Existence of Global Attractors”. *Journal of Dynamics and Differential Equations* 13.2 (2001), pp. 215–249. DOI: 10.1023/A:1016673307045.
- [IngenHousz1784] Johann Ingen-Housz: “Bemerkungen über den Gebrauch des Vergrößerungsglases”. In: *Verm. Schriften physisch-medicinischen Inhalts*. Christian Friderich Wappler: Wien, 1784.
- [InmanBradley1989] Henry F. Inman and Edwin L. Bradley: “The overlapping coefficient as a measure of agreement between probability distributions and point estimation of the overlap of two normal densities”. *Communications in Statistics - Theory and Methods* 18.10 (1989), pp. 3851–3874. DOI: 10.1080/03610928908830127.

- [IsogenLifeScience2014] Isogen Life Science: *Important Factors in Performing Quantitative Western Blots*. Nov. 2013. URL: <http://www.isogen-lifescience.com/blotting/western-blotting-documentation/important-factors-in-performing-quantitative-western-blots> (visited on 08/08/2016).
- [Jacod2005] Jean Jacod, Thomas G. Kurtz, Sylvie Méléard, and Philip Protter: “The approximate Euler method for Lévy driven stochastic differential equations”. *Annales de l’Institut Henri Poincaré (B) Probability and Statistics* 41.3 (2005). En hommage a Paul André Meyer, pp. 523–558. DOI: 10.1016/j.anihpb.2004.01.007.
- [Jazwinski1970] Andrew H. Jazwinski: *Stochastic Processes and Filtering Theory*. Mathematics in Science and Engineering. Elsevier Science, 1970.
- [JohnstonGrandis2011] Paul A. Johnston and Jennifer R. Grandis: “STAT3 signaling: anticancer strategies and challenges”. *Mol. Interv.* 11.1 (Feb. 2011), pp. 18–26.
- [Kahrimanian1953] Harry G. Kahrimanian: “Analytical Differentiation by a Digital Computer”. MA thesis. Temple University, May 1953.
- [Kallenberg2002] Olav Kallenberg: *Foundations of Modern Probability*. Probability and Its Applications. Springer New York, 2002.
- [KallenbergSztencel1991] Olav Kallenberg and Rafal Sztencel: “Some dimension-free features of vector-valued martingales”. *Probability Theory and Related Fields* 88.2 (1991), pp. 215–247. DOI: 10.1007/BF01212560.
- [KalmanBucy1961] Rudolf E. Kálmán and Richard S. Bucy: “New results in linear filtering and prediction theory”. *Trans. ASME, Ser. D, J. Basic Eng* (1961), p. 109.
- [Kanerva2013] Anna Kanerva, Petri Nokisalmi, Iulia Diaconu, Anniina Koski, Vincenzo Cerullo, Ilkka Liikanen, Siri Tähtinen, Minna Oksanen, Raita Heiskanen, Saira Pesonen, Timo Joensuu, Tuomo Alanko, Kaarina Partanen, Leena Laasonen, Kalevi Kairemo, Sari Pesonen, Lotta Kangasniemi, and Akseli Hemminki: “Antiviral and antitumor T-cell immunity in patients treated with GM-CSF-coding oncolytic adenovirus”. *Clin. Cancer Res.* 19.10 (May 2013), pp. 2734–2744.
- [KesslerSorensen1999] Mathieu Kessler and Michael Sørensen: “Estimating equations based on eigenfunctions for a discretely observed diffusion process”. *Bernoulli* 5.2 (Apr. 1999), pp. 299–314.
- [Kikuchi1991] K. Kikuchi, M. Yoshida, T. Maekawa, and H. Watanabe: “Metropolis Monte Carlo method as a numerical technique to solve the Fokker-Planck equation”. *Chemical Physics Letters* 185.3 (1991), pp. 335–338. DOI: 10.1016/S0009-2614(91)85070-D.
- [Klingmueller1995] Ursula Klingmüller, Ulrike Lorenz, Lewis C. Cantley, Benjamin G. Neel, and Harvey F. Lodish: “Specific recruitment of SH-PTP1 to the erythropoietin receptor causes inactivation of JAK2 and termination of proliferative signals”. *Cell* 80 (Mar. 1995), pp. 729–738.
- [KloedenPlaten1995] Peter E. Kloeden and Eckhard Platen: *Numerical solution of stochastic differential equations*. 2. corr. print. Applications of mathematics 23. Springer: Berlin, Heidelberg, 1995.
- [Krylov2002] Nicolai V. Krylov: *Introduction to the Theory of Random Processes*. American Mathematical Society, 2002.
- [Kummer2000] Ursula Kummer, Lars F. Olsen, C. Jane Dixon, Anne K. Green, Erich Bornberg-Bauer, and Gerold Baier: “Switching from simple to complex oscillations in calcium signaling”. *Biophys. J.* 79.3 (Sept. 2000), pp. 1188–1195.
- [Kummer2005] Ursula Kummer, Borut Krajnc, Jürgen Pahle, Anne K. Green, C. Jane Dixon, and Marko Marhl: “Transition from stochastic to deterministic behavior in calcium oscillations”. *Biophys. J.* 89.3 (Sept. 2005), pp. 1603–1611.

- [Lauritzen1981] Steffen L. Lauritzen: “Time Series Analysis in 1880: A Discussion of Contributions Made by T.N. Thiele”. *International Statistical Review / Revue Internationale de Statistique* 49.3 (1981), pp. 319–331.
- [Lederle2011] Wiltrud Lederle, Sofia Depner, Sabine Schnur, Eva Obermüller, Nicola Catone, Alexandra Just, Norbert E. Fusenig, and Margareta M. Müller: “IL-6 promotes malignant growth of skin SCCs by regulating a network of autocrine and paracrine cytokines”. *Int. J. Cancer* 128 (June 2011), pp. 2803–2814.
- [Lenz2014] Simon M. Lenz: “Impulsive Hybrid Discrete-Continuous Delay Differential Equations”. PhD thesis. Universität Heidelberg, 2014.
- [Ljung2002] Lennart Ljung: “Recursive identification algorithms”. *Circuits, Systems and Signal Processing* 21.1 (2002), pp. 57–68. DOI: 10.1007/BF01211651.
- [LjungSoderstrom1985] Lennart Ljung and Torsten Söderström: *Theory and practice of recursive identification*. MIT Press series in signal processing, optimization, and control. MIT Press, 1985.
- [Lo1998] Andrew W. Lo: “Maximum Likelihood Estimation of Generalized Itô Processes with Discretely Sampled Data”. *Econometric Theory* 4.2 (1988), pp. 231–247.
- [LuschgyPages2006] Harald Luschgy and Gilles Pagès: “Functional quantization of a class of Brownian diffusions: A constructive approach”. *Stochastic Processes and their Applications* 116.2 (2006), pp. 310–336. DOI: 10.1016/j.spa.2005.09.003.
- [Lyness1967] James N. Lyness: “Numerical Algorithms Based on the Theory of Complex Variable”. In: *Proceedings of the 1967 22nd National Conference in Washington, D.C., USA*. ACM ’67. Association for Computing Machinery: New York, NY, USA, 1967, pp. 125–133. DOI: 10.1145/800196.805983.
- [LynessMoler1967] James N. Lyness and Cleve B. Moler: “Numerical Differentiation of Analytic Functions”. *SIAM Journal on Numerical Analysis* 4.2 (June 1967), pp. 202–210.
- [Marsaglia1965] George Marsaglia: “Ratios of Normal Variables and Ratios of Sums of Uniform Variables”. *Journal of the American Statistical Association* 60.309 (1965), pp. 193–204. DOI: 10.2307/2283145.
- [Marsaglia2006] George Marsaglia: “Ratios of Normal Variables”. *Journal of Statistical Software* 16.4 (May 2006), pp. 1–10. DOI: 10.18637/jss.v016.i04.
- [Martin1994] Chris S. Martin and Irena Bronstein: “Imaging of chemiluminescent signals with cooled CCD camera systems”. *J. Biolumin. Chemilumin.* 9 (1994), pp. 145–153.
- [MartinezMoczygamba2003] Margarita Martinez-Moczygamba and David P. Huston: “Biology of common beta receptor-signaling cytokines: IL-3, IL-5, and GM-CSF”. *J. Allergy Clin. Immunol.* 112 (Oct. 2003), pp. 653–665.
- [Martins2003] Joaquim R. R. A. Martins, Peter Sturdza, and Juan J. Alonso: “The complex-step derivative approximation”. *ACM Trans. Math. Softw.* 29.3 (2003), pp. 245–262. DOI: 10.1145/838250.838251.
- [Miao2011] Hongyu Miao, Xiaohua Xia, Alan S. Perelson, and Hulin Wu: “On Identifiability of Nonlinear ODE Models and Applications in Viral Dynamics”. *SIAM Rev Soc Ind Appl Math* 53.1 (Jan. 2011), pp. 3–39.
- [Milstein1975] Grigori N. Mil’shtejn: “Approximate integration of stochastic differential equations”. English translation by K. Durr, original in Russian. *Theory Probab. Appl.* 19 (1975), pp. 557–562. DOI: 10.1137/1119062.
- [MoertersPeres2010] Peter Mörters and Yuval Peres: *Brownian Motion*. Cambridge University Press, 2010. DOI: 10.1017/CB09780511750489.

- [MollerMadsen2010] Jan Kloppenborg Møller and Henrik Madsen: *From State Dependent Diffusion to Constant Diffusion in Stochastic Differential Equations by the Lamperti Transform*. Tech. rep. Technical University of Denmark, DTU Informatics, 2010.
- [Mueller2001] Margareta M. Müller, Wolfgang Peter, Marion Mappes, Andrea Hülsen, Heinrich Steinbauer, Petra Boukamp, Michael Vaccariello, Jonathan Garglick, and Norbert E. Fusenig: “Tumor progression of skin carcinoma cells in vivo promoted by clonal selection, mutagenesis, and autocrine growth regulation by granulocyte colony-stimulating factor and granulocyte-macrophage colony-stimulating factor”. *Am. J. Pathol.* 159 (Oct. 2001), pp. 1567–1579.
- [Nagumo1962] Jinichi Nagumo, Suguru Arimoto, and Shuji Yoshizawa: “An Active Pulse Transmission Line Simulating Nerve Axon”. *Proceedings of the IRE* 50.10 (Oct. 1962), pp. 2061–2070. DOI: 10.1109/jrproc.1962.288235.
- [Nici2014] Marco Nici: “IL-6 induzierte Signaltransduktion in HaCaT-ras A5 und Fibroblasten”. PhD thesis. Universität Heidelberg, 2014.
- [Ninck2003] Simon Ninck, Christoph Reisser, Gerhard Dyckhoff, Burkhard Helmke, Harald Bauer, and Christel Herold-Mende: “Expression profiles of angiogenic growth factors in squamous cell carcinomas of the head and neck”. *Int. J. Cancer* 106 (Aug. 2003), pp. 34–44.
- [NocedalWright2006] Jorge Nocedal and Stephen J. Wright: *Numerical Optimization*. Springer series in operations research and financial engineering. Springer, 2006.
- [Nolan1953] John F. Nolan: “Analytical Differentiation on a Digital Computer”. MA thesis. Massachusetts Institute of Technology, May 1953.
- [Obermueller2004] Eva Obermüller, Silvia Vosseler, Norbert E. Fusenig, and Margareta M. Müller: “Cooperative autocrine and paracrine functions of granulocyte colony-stimulating factor and granulocyte-macrophage colony-stimulating factor in the progression of skin carcinoma cells”. *Cancer Res.* 64 (Nov. 2004), pp. 7801–7812.
- [Oksendal1998] Bernt Øksendal: *Stochastic Differential Equations: An Introduction with Applications*. Fifth edition. Springer, 1998.
- [Pachpatte1998] Baburao G. Pachpatte. Vol. 197. Mathematics in Science and Engineering. Elsevier, 1998, pp. 9–97. DOI: 10.1016/S0076-5392(98)80003-9.
- [Pearson2001] Gray Pearson, Fred Robinson, Tara Beers Gibson, Bing e Xu, Mahesh Karandikar, Kevin Berman, and Melanie H. Cobb: “Mitogen-Activated Protein (MAP) Kinase Pathways: Regulation and Physiological Functions”. *Endocrine Reviews* 22.2 (2001). PMID: 11294822, pp. 153–183. DOI: 10.1210/edrv.22.2.0428.
- [Pedersen1995] Asger Roer Pedersen: “A New Approach to Maximum Likelihood Estimation for Stochastic Differential Equations Based on Discrete Observations”. *Scandinavian Journal of Statistics* 22.1 (1995), pp. 55–71.
- [Pichler2013] Lukas Pichler, Arif Masud, and Lawrence A. Bergman: “Numerical Solution of the Fokker-Planck Equation by Finite Difference and Finite Element Methods - A Comparative Study”. In: *Computational Methods in Stochastic Dynamics*. Ed. by M. Papadrakakis, G. Stefanou, and V. Papadopoulos. Computational Methods in Applied Sciences. Springer Netherlands, 2013, pp. 69–85.
- [Protter2004] Philip E. Protter: *Stochastic integration and differential equations*. 2nd ed. Applications of mathematics 21. Springer: Berlin, Heidelberg, 2004.
- [ProtterTalay1997] Philip E. Protter and Denis Talay: “The Euler scheme for Lévy driven stochastic differential equations”. *Ann. Probab.* 25.1 (Jan. 1997), pp. 393–423. DOI: 10.1214/aop/1024404293.

- [Putoczki2010] Tracy Putoczki and Matthias Ernst: “More than a sidekick: the IL-6 family cytokine IL-11 links inflammation to cancer”. *J. Leukoc. Biol.* 88.6 (Dec. 2010), pp. 1109–1117.
- [RamisConde2009] Ignacio Ramis-Conde, Mark A. J. Chaplain, Alexander R. A. Anderson, and Dirk Drasdo: “Multi-scale modelling of cancer cell intravasation: the role of cadherins in metastasis”. *Physical Biology* 6.1 (2009).
- [Rawlings2004] Jason S. Rawlings, Kristin M. Rosler, and Douglas A. Harrison: “The JAK/STAT signaling pathway”. *J. Cell. Sci.* 117 (Mar. 2004), pp. 1281–1283.
- [Riskén1996] Hannes Riskén and Till Frank: *The Fokker-Planck Equation: Methods of Solution and Applications*. Springer Series in Synergetics. Springer Berlin Heidelberg, 1996.
- [Schilling2005] Marcel Schilling, Tim Maiwald, Sebastian Bohl, Markus Kollmann, Clemens Kreutz, Jens Timmer, and Ursula Klingmüller: “Quantitative data generation for systems biology: the impact of randomisation, calibrators and normalisers”. *Syst Biol (Stevenage)* 152 (Dec. 2005), pp. 193–200.
- [Schilling2005a] Marcel Schilling, Tim Maiwald, Sebastian Bohl, Markus Kollmann, Clemens Kreutz, Jens Timmer, and Ursula Klingmüller: “Computational processing and error reduction strategies for standardized quantitative data in biological networks”. *FEBS J.* 272 (Dec. 2005), pp. 6400–6411.
- [Schloeder1983] Johannes P. Schlöder and H. Georg Bock: “Identification of Rate Constants in Bistable Chemical Reactions”. In: *Numerical Treatment of Inverse Problems in Differential and Integral Equations, Progress in Scientific Computing*. Ed. by Deuffhard and Hairer. Birkhäuser: Basel, Boston, Berlin, 1983, pp. 27–47.
- [Schloeder1987] Johannes P. Schlöder: *Numerische Methoden zur Behandlung hochdimensionaler Aufgaben der Parameteridentifizierung*. Vol. 187. Bonner Mathematische Schriften. Universität Bonn, 1987.
- [Schoeberl2002] Birgit Schoeberl, Claudia Eichler-Jonsson, Ernst D. Gilles, and Gertraud Müller: “Computational modeling of the dynamics of the MAP kinase cascade activated by surface and internalized EGF receptors”. *Nat. Biotechnol.* 20 (Apr. 2002), pp. 370–375.
- [Schroers2005] Andreas Schroers, Oliver Hecht, Karl-Josef Kallen, Michael Pachta, Stefan Rose-John, and Joachim Grotzinger: “Dynamics of the gp130 cytokine complex: a model for assembly on the cellular membrane”. *Protein Sci.* 14 (Mar. 2005), pp. 783–790.
- [Seruga2008] Bostjan Seruga, Haibo Zhang, Lori J. Bernstein, and Ian F. Tannock: “Cytokines and their relationship to the symptoms and outcome of cancer”. *Nat. Rev. Cancer* 8.11 (Nov. 2008), pp. 887–899.
- [Shoji2002] Isao Shoji: “Nonparametric state estimation of diffusion processes”. *Biometrika* 89.2 (2002), pp. 451–456.
- [Shoji2003] Isao Shoji: “Estimation of Diffusion Parameters by a Nonparametric Drift Function Model”. In: *Intelligent Data Engineering and Automated Learning: 4th International Conference, IDEAL 2003, Hong Kong, China. Revised Papers*. Ed. by Jiming Liu, Yiu-ming Cheung, and Hujun Yin. Springer: Berlin, Heidelberg, 2003, pp. 211–217. DOI: 10.1007/978-3-540-45080-1_29.
- [ShojiOzaki1997] Isao Shoji and Tohru Ozaki: “Comparative study of estimation methods for continuous time stochastic processes”. *Journal of Time Series Analysis* 18.5 (1997), pp. 485–506. DOI: 10.1111/1467-9892.00064.
- [ShojiOzaki1998] Isao Shoji and Tohru Ozaki: “A Statistical Method of Estimation and Simulation for Systems of Stochastic Differential Equations”. *Biometrika* 85.1 (1998), pp. 240–243.

- [Singer2002] Hermann Singer: “Parameter Estimation of Nonlinear Stochastic Differential Equations: Simulated Maximum Likelihood versus Extended Kalman Filter and Itô-Taylor Expansion”. *Journal of Computational and Graphical Statistics* 11.4 (2002), pp. 972–995. DOI: 10.1198/106186002808.
- [Singer2004] Hermann Singer: “A survey of estimation methods for stochastic differential equations”. In: *Proceedings of 6th international conference on social science methodology, Amsterdam*. 2004.
- [Singh2006] Abhay Singh, Arul Jayaraman, and Jürgen Hahn: “Modeling regulatory mechanisms in IL-6 signal transduction in hepatocytes”. *Biotechnol. Bioeng.* 95 (Dec. 2006), pp. 850–862.
- [Soebiyanto2007] Radina P. Soebiyanto, Sree N. Sreenath, Cheng-Kui Qu, Kenneth A. Loparo, and Kevin D. Bunting: “Complex systems biology approach to understanding coordination of JAK-STAT signaling”. *BioSystems* 90 (2007), pp. 830–842.
- [Sommer2014] Andreas Sommer, Sofia Depner, Mario S. Mommer, Hans Georg Bock, Johannes P. Schlöder, and Margareta M. Müller: “Crosstalk effects in cytokine signaling: how important are they? - The case of IL-6 and GM-CSF”. *International Journal of Biomathematics and Biostatistics* (accepted).
- [Soong1973] T. T. Soong: *Random differential equations in science and engineering*. Mathematics in Science and Engineering. Elsevier Science, 1973.
- [Sorensen2004] Helle Sørensen: “Parametric Inference for Diffusion Processes Observed at Discrete Points in Time: A Survey”. *International Statistical Review* 72.3 (2004), pp. 337–354.
- [Southern1975] Edwin M. Southern: “Detection of specific sequences among DNA fragments separated by gel electrophoresis”. *J. Mol. Biol.* 98 (Nov. 1975), pp. 503–517.
- [SquireTrapp1998] William Squire and George Trapp: “Using Complex Variables to Estimate Derivatives of Real Functions”. *SIAM Review* 40.1 (1998), pp. 110–112. DOI: 10.1137/S003614459631241X.
- [Stacey1997] Alison Stacey and Alan Doyle: “Routine testing of cell cultures and their products for mycoplasma contamination”. *Methods Mol. Biol.* 75 (1997), pp. 305–311.
- [Steven2009] Fnu Steven, Zuyi Huang, and Jürgen Hahn: *Mathematical model of IL6 signal transduction pathway*. 2009. URL: http://homepages.rpi.edu/~hahnj/Models/Updated_IL_6_Model.pdf (visited on 09/23/2014).
- [Strand1970] John L. Strand: “Random ordinary differential equations”. *Journal of Differential Equations* 7.3 (1970), pp. 538–553. DOI: 10.1016/0022-0396(70)90100-2.
- [Stratonovich1960] Ruslan L. Stratonovich: “Conditional Markov Processes”. *Theory of Probability & Its Applications* 5.2 (1960), pp. 156–178. DOI: 10.1137/1105015.
- [Sun2008] Xiaodian Sun, Li Jin, and Momiao Xiong: “Extended Kalman filter for estimation of parameters in nonlinear state-space models of biochemical networks”. *PLoS ONE* 3 (2008), e3758.
- [Sussmann1978] Hector J. Sussmann: “On the Gap Between Deterministic and Stochastic Ordinary Differential Equations”. *Ann. Probab.* 6.1 (Feb. 1978), pp. 19–41. DOI: 10.1214/aop/1176995608.
- [Swameye2003] Ira Swameye, T. G. Müller, Jens Timmer, O. Sandra, and Ursula Klingmüller: “Identification of nucleocytoplasmic cycling as a remote sensor in cellular signaling by databased modeling”. *Proc. Natl. Acad. Sci. U.S.A.* 100 (Feb. 2003), pp. 1028–1033.
- [Swerling1959] Peter Swerling: *First Order Error Propagation in a Stagewise Smoothing Procedure for Satellite Observations*. Rand Corporation, 1959.

- [Thiele1880] Thorvald N. Thiele: *Om Anvendelse af mindste Kvadraters Methode i nogle Tilfælde, hvor en Komplikation af visse Slags uensartede tilfældige Fejlkilder giver Fejlene en 'systematisk' Karakter*. Vol. Raekke 5, Bd. 12. Kongelige Danske Videnskabernes Selskab Skrifter, Naturvidenskabelig og Mathematisk Afdeling: Kjøbenhavn: Munksgaard, 1880, pp. 381–408.
- [Thiele1880a] Thorvald N. Thiele: *Sur la compensation de quelques erreurs quasi-systématiques par la méthode des moindres carrés*. French translation of [Thiele1880]. Copenhagen: C. A. Reitzel, 1880.
- [Trikha2003] Mohit Trikha, Robert Corringham, Bernard Klein, and Jean-François Rossi: “Targeted anti-interleukin-6 monoclonal antibody therapy for cancer: a review of the rationale and clinical evidence”. *Clin. Cancer Res.* 9 (Oct. 2003), pp. 4653–4665.
- [UhlenbeckOrnstein1930] George E. Uhlenbeck and Leonard S. Ornstein: “On the Theory of the Brownian Motion”. *Physical Review* 36 (5 Sept. 1930), pp. 823–841. DOI: 10.1103/PhysRev.36.823.
- [Valdembri2002] Donatella Valdembri, Guido Serini, Angelo Vacca, Domenico Ribatti, and Federico Bussolino: “In vivo activation of JAK2/STAT-3 pathway during angiogenesis induced by GM-CSF”. *FASEB J.* 16.2 (2002), pp. 225–227.
- [vanderPol1926] Balthasar van der Pol: “On 'relaxation-oscillations'”. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* 2.11 (1926), pp. 978–992. DOI: 10.1080/14786442608564127.
- [vanSnick1990] Jacques van Snick: “Interleukin-6: An Overview”. *Annual Review of Immunology* 8.1 (1990). PMID: 2188664, pp. 253–278. DOI: 10.1146/annurev.iy.08.040190.001345.
- [Vasicek1977] Oldřich Alfons Vašíček: “An equilibrium characterization of the term structure”. *Journal of Financial Economics* 5.2 (1977), pp. 177–188. DOI: 10.1016/0304-405X(77)90016-2.
- [Welch1947] Bernard L. Welch: “The generalization of “Student’s” problem when several different population variances are involved”. *Biometrika* 34 (1-2) (1947), pp. 28–35.
- [Wengert1964] R. E. Wengert: “A simple automatic derivative evaluation program”. *Communications of the ACM* 7.8 (1964), pp. 463–464. DOI: 10.1145/355586.364791.
- [Wiener1863] Christian Wiener: “Erklärung des atomistischen Wesens des tropfbar flüssigen Körperzustandes und Bestätigung desselben durch die sogenannten Molekularbewegungen”. *Annalen der Physik* 118 (1863), pp. 79–94.
- [Wiener1923] Norbert Wiener: “Differential-Space”. *Journal of Mathematics and Physics* 2.1-4 (1923), pp. 131–174. DOI: 10.1002/sapm192321131.
- [Williams1988] Douglas E. Williams, David C. Bicknell, Linda S. Park, John E. Straneva, Scott Cooper, and Hal E. Broxmeyer: “Purified murine granulocyte/macrophage progenitor cells express a high-affinity receptor for recombinant murine granulocyte/macrophage colony-stimulating factor”. *Proc. Natl. Acad. Sci. U.S.A.* 85 (Jan. 1988), pp. 487–491.
- [WongHajek2012] Eugene Wong and Bruce Hajek: *Stochastic Processes in Engineering Systems*. Springer Texts in Electrical Engineering. Springer New York, 2012.
- [WongZakai1965] Eugene Wong and Moshe Zakai: “On the Convergence of Ordinary Integrals to Stochastic Integrals”. *Ann. Math. Statist.* 36.5 (Oct. 1965), pp. 1560–1564. DOI: 10.1214/aoms/1177699916.
- [Woods1986] N. M. Woods, K. S. Cuthbertson, and P. H. Cobbold: “Repetitive transient rises in cytoplasmic free calcium in hormone-stimulated hepatocytes”. *Nature* 319.6054 (1986), pp. 600–602.

- [Yamada2003] Satoshi Yamada, Satoru Shiono, Akiko Joo, and Akihiko Yoshimura: “Control mechanism of JAK/STAT signal transduction pathway”. *FEBS Lett.* 534 (Jan. 2003), pp. 190–196.
- [Yu2014] Hua Yu, Heehyoung Lee, Andreas Herrmann, Ralf Buettner, and Richard Jove: “Revisiting STAT3 signalling in cancer: new and unexpected biological functions”. *Nat. Rev. Cancer* 14.11 (Nov. 2014), pp. 736–746.
- [Zimmer2015] Christoph Zimmer: “Reconstructing the hidden states in time course data of stochastic models”. *Math Biosci* 269 (Nov. 2015), pp. 117–129.

Index

A

ACM SIGPLAN: History of Programming Languages . . . 34
activator (protein) *see* BISTABAER model
active set 3
AD (automatic differentiation) . . . *see* derivative generation
adapted process *see* stochastic process
additive noise *see* stochastic differential equation
adjoint sensitivities *see* derivative generation
algorithmic differentiation *see* derivative generation
aliasing *see* stochastic process
allosteric regulation *see* BISTABAER model
allosteric site *see* BISTABAER model
almost sure convergence 211
amplification factors method 41, 63, **68–72**
 additive error 69
 multiplicative bias 69
approximative EULER method for LÉVY-driven SDE . .
 *see* stochastic integrator
ARIMOTO 164
ARMIJO 20
automatic differentiation *see* derivative generation

B

BÉNARD 106
BACHELIER 106
background driving LÉVY process . . . *see* LÉVY process
backward mode (of algorithmic differentiation) . . .
 *see* derivative generation
ball (closed, open) 1
BDLP *see* LÉVY process
BELLMAN 133, 134, 137, 138
BESSEL 57, 66, 69, 77, 78
BISTABAER model **187–193**
 allosteric regulation 122, 187
 activator 122
 allosteric site 122
 effector protein 122
 inhibitor 122
 model 187
 reconstruction of hidden states 193
 simulation study 188
blotting *see* WESTERN blotting
BOCK xvii, 1, 17, 19, 20, 28
BONHÖFFER-VAN DER POL model 164

BOREL 106
BOREL set 99
BOREL σ -algebra *see* σ -algebra
boundary value problem *see* parameter estimation
BRADLEY 71
BROWN 106
BROWNIAN motion 106–109, 112, 127, 213, 215,
 216, 222
BUCY xix, 127
BULIRSCH 28
BUNYAKOVSKY 133, 136
BUTCHER 224
BVP (boundary value problem) . . . *see* parameter estimation
 see also BONHÖFFER-VAN DER POL model

C

càdlàg 104
calcium oscillator **174–187**
 calcium-induced calcium release 174
 G-protein 174
 identifiability analysis 177
 inositol 1,4,5-trisphosphate 174
 model 175
 phosphatidylinositol 4,5-bisphosphate 174
 phosphoinositid-phospholipase C 174
 recovering trajectories 183
 simulation study 178
cancer 42
CAUCHY 60, 61, 133, 136
central limit theorem 23
CHOLESKY 155
CLARKE 34
(CNLP) *see* constrained nonlinear problem
COBELLI 25
collocation *see* numerical integration
combined residual vector 121, 144, 151, 155
 decomposition factor R 155, 157, 158, 160,
 170, 180, 190, 200, 201, 239
common factor *see* parameter estimation
compatibility condition 20
complement (of a set) 97
complementarity condition 4
 strict 5
 strict condition (SC) 16
complete probability space *see* probability space

complex step derivative approximation
 *see* derivative generation
 compound POISSON process *see* POISSON process
 condensing *see* parameter estimation
 conditional expectation *see* expectation
 conditional probability density
 *see* parameter estimation in SDE
 confidence intervals *see* parameter estimation
 confidence region *see* parameter estimation
 constrained nonlinear least squares
 *see* parameter estimation
 constrained nonlinear problem (CNLP) 2
 global, local, strict solution 2
 constraint qualification
 condition (CQ) 16
 linear independence (LICQ) 3
 MANGASARIAN-FROMOVITZ (MFCQ) 4
 continuity conditions . *see* multiple shooting parametrization
 convergence of random sequences
 almost sure 211
 in distribution 211
 in L^2 211
 in law 211
 in probability 211
 mean square 211
 pointwise 211
 stochastic 211
 strong 211
 sure 211
 weak 211
 with probability 1 211
 convergence order *see* stochastic integrator
 correlated measurements *see* parameter estimation
 COTTLE 4
 counting process *see* stochastic process
 covariance matrix *see* parameter estimation
 (CQ) *see* constraint qualification
 crosstalk *see* JAK-STAT signalling pathway
 cytokines 43
 GM-CSF (granulocyte-macrophage colony-stimulating factor) 43
 IL-6 (interleukin 6) 43
 see also signalling pathways

D

DAVIDENKO 20
 decomposition factor R *see* combined residual vector
 derivative generation
 algorithmic differentiation 33
 backward mode 34
 forward mode 33
 algorithmic differentiation (AD)
 history of AD 34
 operator overloading 33
 source transformation 33
 complex step approximation 32

external numerical differentiation (END) 31
 finite differences (FD) 31
 rounding error 31
 truncation error 31
 i(maginary) trick 32
 internal numerical differentiation (IND) 32
 perturbed trajectories 31
 sensitivities 31, 33
 variational differential equations 33
 deterministic initial value problem
 *see* initial value problem
 difference process *see* stochastic process
 differential equation
 see ordinary differential equation
 see random ordinary differential equation
 see stochastic differential equation
 see also initial value problem
 diffusion
 of a WIENER process 106, 107
 of an SDE 108, 110, 139
 see also stochastic process
 DIRAC 127
 direct elimination method 10, 11
 discretization *see* multiple shooting parametrization
 dissimilarity index 71
 DISTEFANO 25
 distribution
 GAUSSian 102
 normal 102
 of a random variable 100
 POISSON 102
 tails 23
 D-IVP *see* initial value problem
 DOOB 131, 221
 DOOB's martingale inequalities . . *see* stochastic process
 dose dependency 92
 DOSS 112
 drift
 of a WIENER process 107
 of an SDE 108, 110, 139
 see also stochastic process

E

effector proteins *see* BISTABAER model
 efficient method of moments
 *see* parameter estimation in SDE
 EINSTEIN 106
 EKF (extended KÁLMÁN filter)
 *see* parameter estimation in SDE
 electrophoresis 50
 elementary process *see* stochastic process
 EMM (efficient method of moments)
 *see* parameter estimation in SDE
 END (external numerical differentiation)
 *see* derivative generation

error model *see* model
 EULER 113
 EULER-MARUYAMA method *see* stochastic integrator
 event 98
 expectation 101
 conditional 101
 extended KÁLMÁN filter (EKF)
 *see* parameter estimation in SDE
 external numerical differentiation (END)
 *see* derivative generation

F

FATEMAN 34, 224
 FD (finite differences) *see* derivative generation
 feasibility condition
 dual 4
 primal 4
 feasible point 2
 feasible set 2
 filtered probability space *see* probability space
 filtration 215
 natural 216
 finite differences *see* derivative generation
 FISHER information matrix *see* parameter estimation
 FITZHUGH-NAGUMO oscillator 164–174
 estimation test series 170
 model 166
 simulation study 167
 steady state 164
 FOKKER 126, 127
 FOKKER-PLANCK equation
 *see* parameter estimation in SDE
 forward mode (of algorithmic differentiation)
 *see* derivative generation
 FREIDLIN 132
 FUBINI 137, 138

G

GALLANT 128
 GAUSS-NEWTON method 1, 14–20, 37–40, 145,
 151, 155, 169, 179, 189, 200, 223
 search direction 14
 GAUSSian distribution *see* distribution
 gel electrophoresis 50
 generalized inverse 8, 17
 see also pseudo-inverse
 generalized method of moments (GMM)
 *see* parameter estimation in SDE
 genuine EULER method for LÉVY-driven SDE
 *see* stochastic integrator
 GGN *see* GAUSS-NEWTON method
 GHOSH 80

GILLESPIE 80, 174, 175
 GIVENS 209
 global solution *see* solution
 GM-CSF *see* cytokines
 GMM (generalized method of moments)
 *see* parameter estimation in SDE
 :gnsolve *see* software
 GOLDSTEIN 20
 GRAM 209
 granulocyte-macrophage colony-stimulating factor
 *see* cytokines
 GRIEWANK 33
 GRONWALL 133, 134, 137, 138
 GRÖTSCHEL 4
 GUERRIERO 79

H

HaCaT (cell line) 43
 HÄNGGI 106
 HANSEN 128
 history (of a WIENER process) *see* WIENER process
 HODGKIN 164
 HOOKE 195
 HOPPER 34
 house-keeping protein 64
 HOUSEHOLDER 121, 155, 157, 158, 209
 HUXLEY 164

I

i(maginary) trick *see* derivative generation
 identifiability *see* parameter estimation
 IL-6 *see* cytokines
 ImageJ *see* WESTERN blotting
 IMKELLER 112
 immunoblotting *see* WESTERN blotting
 immunolabelling *see* WESTERN blotting
 immunoprecipitation 50
 incompatibility constant 19
 increments
 of a stochastic process *see* stochastic process
 of a WIENER process *see* WIENER process
 IND (internal numerical differentiation)
 *see* derivative generation
 independent increments *see* stochastic process
 indicator function 100
 indifference region *see* parameter estimation
 indirect inference method *see* parameter estimation in SDE
 indistinguishable *see* stochastic process
 information matrix *see* parameter estimation
 INGEN-HOUSZ 106
 inhibitor (protein) *see* BISTABAER model

initial value problem
 deterministic (D-IVP) 129
 stochastic (S-IVP) 109, 110, 129
see also differential equation
 INMAN 71
 integration *see* numerical integration
 interleukin 6 *see* cytokines
 internal numerical differentiation (IND)
 *see* derivative generation
 interval solutions *see* parameter estimation in SDE
 ITÔ 97, 107, 109, 110, 112, 115, 117, 119, 124, 126,
 128, 130, 131, 215–217, 219–222

J

JACOD 120
 JAK-STAT signalling pathway 42, **45**
 crosstalk model 41, 42, 55, **79–88**
 illustration 46, **81**
 induction by GM-CSF 47
 induction by IL-6 45
 interactions in IL-6 and GM-CSF signalling48
 model 45
 JENSEN 130, 131
 JENSEN's inequality 130
 jump regularization *see* parameter estimation in SDE

K

KAHRIMANIAN 34
 KÁLMÁN xix, 127, 128
 KÁLMÁN-BUCY filter *see* parameter estimation in SDE
 KARUSH 4
 KARUSH-KUHN-TUCKER *see* KKT
 KESSLER 128
 KHINTCHINE 119, 120
 KKT
 conditions 4
 point 5
 tuple 5
 KLOEDEN 97
 KOLMOGOROV 105, 126
 KOLMOGOROV's forward equation
 *see* parameter estimation in SDE
 KUHN 4
 KUMMER 174, 177
 KURTZ 175
 KUTTA 29

L

Lagrangian
 function 3
 multipliers 3, 4

LAMPERTI transform *see* stochastic differential equation
 LANGEVIN equation 175
 least squares
 constrained nonlinear problem 14
 objective 25, **35**
 residual function 14
 sequential 11
 total least squares 37
 LEBESGUE 101, 106, 109, 216
 LEBESGUE-BOREL measure *see* measure
 level function 20
 LÉVY 23, 97, 119, 120, 195
 LÉVY process 110, **118**, 119, 120, 164, 195
 background driving (BDLP) 195
 LÉVY-driven SDE *see* stochastic differential equation
 LÉVY-KHINTCHINE representation
 *see* stochastic process
 LICQ (linear independence constraint qualification)
 *see* constraint qualification
 lifting *see* parameter estimation in SDE
 likelihood 24, 25, 126
 LINDBERG 23
 linear independence constraint qualification (LICQ)
 *see* constraint qualification
 LIPSCHITZ 18–21, 27, 111, 114, 119, 129, 134–138,
 150
 LO 127
 local linearization method
 *see* parameter estimation in SDE
 local solution *see* solution
 LumiImager *see* WESTERN blotting
 LYAPUNOV 117
 LYNESS 32
 lysis 49

M

MADSEN 125
 MANGASARIAN-FROMOVITZ constraint qualification
 *see* constraint qualification
 MARKOV 126–128
 Markov chain Monte Carlo (MCMC)
 *see* parameter estimation in SDE
 MARSAGLIA 41, 59, 60, 65, 70
 martingale *see* stochastic process
 martingale estimation functions
 *see* parameter estimation in SDE
 MARUYAMA 113
 matching conditions *see* multiple shooting parametrization
 maximum process *see* stochastic process
 MCMC (Markov chain Monte Carlo)
 *see* parameter estimation in SDE
 mean reversion level, rate 195
see also ORNSTEIN-UHLENBECK model
 mean square convergence 211
 measurable function 98

measurable set 98
 measurable space 98
 measure 98
 LEBESGUE-BOREL measure 99
 measure space 98
 measurement function *see* model
 measurements *see* model
 MEF (martingale estimation functions)
 *see* parameter estimation in SDE
 merit function 20
 MFCQ *see* constraint qualification
 MIAO 25
 MICHAELIS-MENTEN 68, 82
 MILSTEIN 116, 117, 119
 model xiii, 21
 error model 22
 measurement function 22, 140
 measurements 22, 140
 model response 22
 ODE model 21
 right hand side function 21
 see also combined residual vector
 modification *see* stochastic process
 MOLER 32
 MØLLER 125
 MOORE 8, 256
 MOORE-PENROSE pseudo-inverse . . *see* pseudo-inverse
 MÜLLER xv, 42, 45, 47, 240
 multi-experiment setting 197
 multiple shooting parametrization 21, **29**
 continuity conditions 29
 shooting grid 29, 139
 shooting nodes 29, 139
 shooting variables 29, 139
 multipliers *see* Lagrangian

N

NAGUMO x, 145, 164
 natural filtration *see* filtration
 see also WIENER process
 natural level function 20
 NEWTON 19, 20, 29, 30, 128
 NICHOLSON 60
 NOLAN 34
 nonparametric approach . *see* parameter estimation in SDE
 normal distribution *see* distribution
 nullspace 1
 nullspace method 9
 numerical integration
 collocation 21, 28
 single shooting 21, **27**, 28
 see also multiple shooting parametrization
 see also stochastic integrator

O

ODE model *see* model
 ØKSENDAL 97, 108, 213, 214, 217, 222
 operator overloading *see* derivative generation
 optimality conditions
 necessary, first order 5
 second order 5, 6
 ordinary differential equation model *see* model
 ORNSTEIN 112, 195
 ORNSTEIN-UHLENBECK model **195–201**
 estimating the diffusion 197
 heuristic 195
 mean reversion 122
 model 195
 simulation study 198
 ORNSTEIN-UHLENBECK process . . . 195, 196, 201,
 239
 outcomes 98
 overlapping coefficient (OVL) 71, 72
 OZAKI 128

P

parameter estimation
 boundary value problem 27
 common factor 23, 37, 40
 condensing 36
 confidence intervals 40
 confidence region 26
 constrained nonlinear least squares problem 38
 correlated measurements 36
 covariance matrix 26
 identifiability *see* 25
 direct testing 27
 local, global 26
 practical 25–27
 structural 25
 indifference region 40
 information matrix 26
 point constraints 35
 problem formulation **35, 36**
 statistical analysis 38
 see also combined residual vector
 see also likelihood
 see also model
 see also multiple shooting parametrization
 parameter estimation in SDE
 conditional probability density 126
 existing methods 125–129
 discretized maximum likelihood 127
 efficient method of moments 128
 extended KÁLMÁN filter (EKF) 128
 generalized method of moments 128
 indirect inference 128
 KÁLMÁN filtering 127

local linearization 128
 MARKOV chain Monte Carlo (MCMC) 127
 martingale estimation functions 128
 nonparametric approaches 128
 prediction error estimation method 128
 second-order nonlinear filter (SNF) 128
 homotopy for regularization 161
 idea description 123
 interval solutions 141
 jump regularization weight 141, 143
 lifting 159
 likelihood 126
 linear complexity growth 154
 maximum likelihood estimator 126
piecewise deterministic approach . . . 121, 123,
139
 point constraints 141
 selector functions 140
 shooting grid refinement 161
 software package `:sfit` 223
 sparsity structure 155–159
 stochastic jump 141, **143**
 trajectory 141
 transition probability density 126
 FOKKER-PLANCK equation 126
 see also model
 see also multiple shooting parametrization
 (PD) *see* positive definiteness condition
 PENROSE 8, 256
 PERRIN 106
 piecewise deterministic parameter estimation approach .
 *see* parameter estimation in SDE
 PLANCK 126, 127
 PLATEN 97
 point constraints
 see parameter estimation
 see parameter estimation in SDE
 pointwise convergence 211
 POISSON distribution *see* distribution
 POISSON process 118–120, 195, 197, 198
 compound 119, 120, 195, 196, 199, 239
 positive definiteness condition (PD) 16
 practical identifiability *see* parameter estimation
 prediction error estimation method
 *see* parameter estimation in SDE
 probability 98
 probability measure 98
 probability space 98
 complete 98
 filtered 216
 proliferation assay 47
 PROTTER 97
 pseudo-inverse 7, 8, 210
 see also generalized inverse

Q

quantitative immunoblotting . . . *see* WESTERN blotting

R

r.h.s. (right hand side) *see* model
 RADAU 224
 random ordinary differential equation 111, 112
 random variable 100
 independence *see* stochastic independence
 ratio of two random variables 59
 RADON-NIKODÝM derivative 101
 range 1
 ratio of normal random variables 59
 RAYLEIGH 106
 realization *see* stochastic process
 recombinant protein 64
 residual function *see* least squares
 residuals 24
 see also combined residual vector
 restrictive monotonicity test 20
 RIEMANN 213, 214, 216
 right hand side function *see* model
 RMT *see* restrictive monotonicity test
 RODE *see* random ordinary differential equation
 rounding error (of finite differences)
 *see* derivative generation
 RUNGE 29, 224

S

sample path *see* stochastic process
 sample space 98
 (SC) *see* complementarity condition
 SCHEINKMAN 128
 SCHLÖDER 36, 142, 197
 SCHMALFUSS 112
 SCHMIDT 209
 SCHOEBERL 80
 SCHWARTZ 108
 SCHWARZ 133, 136
 SDE *see* stochastic differential equation
 second-order nonlinear filter (SNF)
 *see* parameter estimation in SDE
 selector functions *see* parameter estimation in SDE
 sensitivities *see* derivative generation
 sequential least squares *see* least squares
`:sfit` *see* software
 SHOJI 128
 shooting grid, nodes, variables
 *see* multiple shooting parametrization

- σ -algebra 97
 BOREL 99
 generated by a mapping 98
 generated by a set 98
 signalling pathways 44
 MAPK/ERK 45
see also JAK-STAT signalling pathway
 simple process *see* stochastic process
 SINGH 80
 single shooting *see* numerical integration
 singular value decomposition (SVD) 210
 S-IVP *see* initial value problem
 SMOLUCHOWSKI 106
 SNF (second-order nonlinear filter)
 *see* parameter estimation in SDE
 SOCS-3 46
 SOEBIYANTO 80
 software 223–236
 package :clradau 224
 package :gnsolve 223
 package :sfit **225**
 SORENSEN 128
 source transformation *see* derivative generation
 SOUTHERN 49, 63
 SQUIRE 32
 stationarity condition 4
 stationary increments *see* stochastic process
 stationary point *see* KKT
 statistical analysis *see* parameter estimation
 STEVEN 80, 238, 240
 STIELTJES 213, 214, 216
 stochastic basis 216
 stochastic convergence 211
 stochastic differential equation 109
 additive noise 110, 139
 diffusion 110
 drift 110
 LAMPERTI transform 124–125
 LÉVY-driven 118
 multiplicative noise 110
 numerical integration 112
see also stochastic integrator
 strong solution 110
 weak solution 110
see also model
 stochastic EULER *see* stochastic integrator
 stochastic independence
 of random variables and events 100
 sub- σ -algebras 100
 stochastic initial value problem *see* initial value problem
 stochastic integral
 ITÔ 215, **216–222**
 STRATONOVICH 215
 stochastic integral equation *see* initial value problem
 stochastic integrator
 convergence order (strong, weak) 113
 EULER-MARUYAMA 97, 105, 113–116, 119,
 127, 225, 238
 algorithm 114
 convergence order 114
 explicit MILSTEIN 116
 for LÉVY-driven SDE 119, 120
 higher order schemes 117
 implicit EULER 116
 implicit MILSTEIN 116
 stochastic jump *see* parameter estimation in SDE
 stochastic process 103
 adapted 215
 counting process 118
 difference process 130
 diffusion 139
 drift 108, 139
 elementary process 216
 independent increments 104
 indistinguishable 103
 LÉVY-KHINTCHINE representation 119
 martingale 220
 DOOB’s inequalities 221
 maximum process 130
 modification 103
 natural filtration 216
 realization 103
 sample path 103
 simple process 216
 stationary increments 104
 version 103
 volatility 108
 white noise 108
see also LÉVY process
see also ORNSTEIN-UHLENBECK process
see also POISSON process
see also WIENER process
 STRATONOVICH xix, 117, 127, 215
 STRATONOVICH-KÁLMÁN-BUCY filter
 *see* parameter estimation in SDE
 strict complementarity condition (SC)
 *see* complementarity condition
 strict solution *see* solution
 stripping buffer solution *see* WESTERN blotting
 strong convergence 211
 strong solution *see* stochastic differential equation
 structural identifiability *see* parameter estimation
 SUN 80
 suppressor of cytokine signalling 46
 sure convergence 211
 SUSSMANN 112
 SVD *see* singular value decomposition
 SWAMEYE 80
 SWERLING 127

T

tails *see* distribution
 TAUCHEN 128
 TAYLOR 31, 32, 34, 116, 117, 119, 128
 tempered distribution 108
 THIELE 127
 topology, topological space 99
 total least squares *see* least squares
 trajectory *see* parameter estimation in SDE
 transitional probability density
 *see* parameter estimation in SDE
 TRAPP 32
 truncation error (of finite differences)
 *see* derivative generation
 TUCKER 4

U

UHLENBECK 112, 195
 UNIVAC 34

V

VAN DER POL 164
 variational differential equations
 *see* derivative generation
 VAŠÍČEK 164, 195
 version *see* stochastic process
 volatility 195

W

WALTHER 33
 weak convergence 211
 weak solution *see* stochastic differential equation
 weighted least squares *see* least squares
 WELCH 42, 77
 WENGERT 34
 WENTZELL 132
 WESTERN blotting 49, 52, 63
 data analysis 63
 ImageJ 53
 LumiImager 52, 206
 quantification
 amplification factors method 68
 normalization method 64
 simulation study 73
 stripping buffer solution 53
see also amplification factors method
 Whirlwind I 34
 white noise process *see* stochastic process
 WIENER, Christian 106

WIENER, Norbert 106
 WIENER process . 97, 105, 106, 107, 109, 110, 112,
 113, 115, 118, 119, 122, 123, 125, 129–
 131, 135, 136, 138, 139, 145, 146, 148,
 150, 163, 166–168, 170–172, 175, 176, 178,
 181–183, 185, 187–189, 191, 193, 195, 213,
 214, 216, 222
 continuous modification 105
 drift 106, 107
 history 216
 increments 105
 natural filtration 216
 some expectation results 131
 standard
 1D 105
 nD 107
 variance 107
 volatility 106, 107
 WONG 112
 WOODS 174

Y

YAMADA 80
 YOSHIZAWA 164
 YU 42

Z

ZAKAI 112
 ZIMMER 174

