

Optimal Design of Signaling Modules: Key Drivers, Trade-Offs and Sustainability

THÈSE N° 5419 (2012)

PRÉSENTÉE LE 17 AOÛT 2012

À LA FACULTÉ DES SCIENCES ET TECHNIQUES DE L'INGÉNIEUR

LABORATOIRE D'AUTOMATIQUE

PROGRAMME DOCTORAL EN INFORMATIQUE, COMMUNICATIONS ET INFORMATION

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

POUR L'OBTENTION DU GRADE DE DOCTEUR ÈS SCIENCES

PAR

Andrijana RADIVOJEVIC

acceptée sur proposition du jury:

Dr D. Gillet, président du jury
Prof. D. Bonvin, Prof. V. Hatzimanikatis, directeurs de thèse
Prof. M. Dal Peraro, rapporteur
Prof. H. Köppl, rapporteur
Dr B. Schoeberl, rapporteur



ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

Suisse
2012

ACKNOWLEDGMENTS

The successful realization of this thesis would not been possible without tremendous support of many people, professors and colleagues, my friends and family.

First and foremost, I owe immense gratitude to my thesis advisors, Professor Dominique Bonvin and Professor Vassily Hatzimanikatis, for their support, insightful guidance and mentorship during the course of this work. Prof. Bonvin, je vous remercie d'avoir toujours eu le temps pour moi, pour l'encouragement et pour de précieux commentaires sur mon travail. Βασίλη, σε ευχαριστώ για το θετικό σου πνεύμα, που μου έδωσες την δυνατότητα να συμμετέχω σε συνεδριάσεις και να έχω πρακτικές εμπειρίες, που ενθάρρυνες τις συμβουλευτικές μου ικανότητες και για όλες τις γλυκο-δραστηριότητες που είχαμε.

I would like to express my recognition to Dr. Ljubiša Mišković for his precious advice from the first until the last days of my PhD journey, both from professional and personal aspect. Мишко, ти си ме увео у ”Српску мафију на EPFLy”, бринуо о балансу између два Лаба и био пријатељ на кога сам увек могла да се ослоним. На свему томе сам ти неизмерно захвална.

I also thank to Dr. Benoît Chacuat for helpful guidance and interesting discussions. Benoît, ton enthousiasme et ton dévouement à la science a eu un grand impact sur moi.

It was an honor and pleasure to have, as the members of my Thesis committee, Dr. Denis Gilles, Dr. Birgit Schoeberl, Professor Heinz Koepl and Professor Matteo dal Peraro. I am highly appreciative to them for accepting to read and evaluate the contents of this dissertation.

Many thanks go to all the LA and LCSB members for creating friendly and nice working atmosphere. Je vais chérir tous les moments amusants pendant nombreux petits-déjeuners, les orgies, les fêtes de Noël et les Journées de Recherche au Laboratoire d'Automatique, and photo-shooting, conferences, Xmas gatherings, Lab Retreats and strange-coffee tasting at Laboratory of Computational Systems Biotechnology. Sandy, inté kinté mitil ikhit ilé; 7abbait ktir al ra2ess ma3ik, salon el vernis, el latlaté, el tabekh wel fruit breaks. I thank Evgeny за розжиг моего интереса к конькам и за замечательного танцевального партнера, David for being the administrative executive of LA party club and Sandy&Andi's Nail Salon, Sandra που εισουν σαν μελος της οικογενειας μου, Paman, Christophe, Philippe, Noushin دوست و قشنگ های پایشنهاد و صمیمانه های گفتگو همه برای داشتنی, Liz for all the pouf-talks, past and future common project, Katerina που είσουν γιατρός μου, Marianne pour les différents problèmes techniques, and my buddies Emre, Jan, Luis, Keng and Ho Ki. I extend my gratitude to the friendly administrative staff Ruth, Francine, Sara, Christine and Francis pour leur soutien sincère à de nombreuses reprises.

I wish to thank my students Alen Brušnjak, Johannes Pfleging and Vincent Zimmern for choosing to work with me on different projects that are related to this thesis. Guys, thank you for keeping up with me as your supervisor. Your help is gratefully acknowledged.

The six months internship in Merrimack Pharmaceuticals (Cambridge, MA, USA) was an indispensable experience that made me mature in a scientific sense. Thanks Sasha for being such a great boss, Birgit for your amazing energy and the whole MM team who made my stay in Boston so enjoyable.

My stay in Lausanne would not be as nearly as pleasant and fun without my dear friends. My profound gratitude goes to my “bros” who have always been taking goof care of me: Vane за сва на нивоу акнем путовања, Хер Жикину дар-мар школу, пасуље, савијаче, бризле и урмашице, наступе и редовну аутобуску линију, Ivan за инспиративне дискусије, shopping consulting и семалирање пуфни, Мика за приступ великој масонској ложи, ритуално спаљивање карираних стољњака и акција исхране из заједничког казана иза музичке школе, Сојба за неисцрпне пик-ап стратегије, калибрацију иберлауфа и релажу ламбда-сонде, Аса за све инсомниа изласке и ђуске, Deki за све миле и рођене стил-royal активности. Thank you Nevena, Ivana, Tanja, Jelena, Nenad, Cane. Maria, thanks innik seZadtiné ta ifham 7alé aktar.

Thanks a lot to my friends in Serbia for all those looong skype hang-outs over these years and great time during vacations. Special thanks go to my girls: Sandra, Каћа, Mima, Sanja, and my guys: Boris for huge congratulations, кардан-торте, чешљање прстима, интерконтинентално бањање и месне седнице уз гвиски, Marko for escort и уредну пашошку контролу, Boki, Grujo, Luka, Aco, Dragan, Srđan, Stevo, Branko.

I am grateful to my boyfriend Matija, for his endless love, support and constant encouragement, especially during the most challenging moments. Љмлдимвтп!

Last but not least, my deepest gratitude goes to my sister Tijana, my mom Nada and my dad Cojo, for all their love, sacrifice and understanding. This thesis is dedicated to them. Хвала вам на свему – ово је мој поклон вама!

ABSTRACT

One of the basic characteristics of every living system is the ability to respond to extracellular signals. This is carried out through a limited number of protein-based signaling networks, whose function is not based only on simple transmission of the received signals, but incorporates the processing, encoding and integration of both external and internal signals. The results then lead to different changes in gene expression and regulate cell growth, mitogenesis, differentiation, embryo development, and stress responses in mammalian cells, whereas the malfunction is in correlation with diseases such as cancer, asthma and diabetes.

In signaling networks, the basic units are covalent modification cycles, which comprise the activation and deactivation of proteins by other proteins. Protein modification in cell signaling – typically a phosphorylation and dephosphorylation – is a general mechanism responsible for the transfer of a wide variety of chemical signals in biological systems. Although the concept does not seem to be complex from a biochemical point of view, these simple systems can nevertheless provide a large diapason of dynamical responses and are therefore ubiquitous building blocks of signaling pathways. These cycles are often linked, forming multiple layers of cycles, the so-called cascades.

Commonly observed instance of signal transduction through a series of protein kinase reactions are the kinases of the mitogen-activated protein kinase (MAPK) cascades. These pathways, which are found in almost all eukaryotes, play an important role in controlling different cellular processes, including fundamental functions. The activation of the cellular response by MAPK pathways typically involves at least three phosphorylation steps.

In order to better understand the nature of this regulation and to gain greater insight into the mechanisms that determine the function of cells, signaling modules have been intensively studied using mathematical modeling and computational simulations, through the fast growing field of systems biology and its disciplines. The primary aim is to faithfully describe the system and to be able to predict the system behavior. Synergistically with experimental analysis, the reported observations have allowed one to identify properties of these pathways, such as fast signal propagation, large amplification, short signal duration and noise resistance.

Since biochemical parameters in signaling pathways are not easily accessible experimentally, it is necessary to use advanced mathematical tools for their correct estimation. Using the paradigm of man-made optimal signal transduction systems, we chose to take the research path for discovering optimal design of cellular signaling modules.

To approach the main thesis objective, we first identified the key system parameters through global sensitivity analysis. Comparative analysis of differences and similarities within different system architectures revealed some insights for initial parameter classification and starting point for optimal system design.

In order to be able to interpret a broader range of phenotypes, we take into account both steady-state and dynamic properties simultaneously. Furthermore, we investigated the trade-offs between optimal characteristics. As a result, we found the biochemical and biophysical parameters that determine these trade-offs and we analyzed if there exist conditions under which we can simultaneously achieve optimal steady-state and dynamic performance. We first analyze what are the design principles that lead the system to have the minimal signaling times, subject to a certain level of amplification gain. In this setup, we bring out our main research question: are there any trade-offs and interplay between different steady-state and dynamic properties? Furthermore, we include the property of ultrasensitivity and eventually solve multi-objective optimization problems. A particularly insightful finding of this work is that, upon judicious selection of the kinetic parameters, a simple covalent modification cycle is able to meet multiple objectives simultaneously. In particular, this analysis may help explain why signaling cycles are so ubiquitous in cell signaling. The enhancement of ultrasensitivity and faster signal propagation in the multicyclic systems clearly show the advantages of the natural choice of designing signaling pathways in the form of signaling cascades.

The thesis concludes with the potential research steps that could be taken along the same path, and that would gather more quantitative knowledge about signaling pathways.

Keywords: Systems Biology, signal transduction, mathematical modeling, parametric analysis, optimization, sensitivity analysis, stochastic simulations.

RÉSUMÉ

Une des caractéristiques de base de tout système vivant, c'est la capacité à réagir aux signaux extracellulaires. Ceci est réalisé par l'intermédiaire d'un nombre limité de protéines à la base de réseaux de signaux, dont la fonction ne se limite pas à la simple transmission des signaux reçus, mais incorpore aussi le traitement, le codage et l'intégration des signaux externes et des signaux internes. Ceci engendre différents changements dans l'expression des gènes, et permet la régulation de la croissance cellulaire, la mitogenèse, la différenciation, le développement embryonnaire, ainsi que les réponses au stress dans les cellules de mammifères, alors que le dysfonctionnement est en corrélation avec les maladies, comme le cancer, l'asthme et le diabète.

Dans les réseaux de signalisation, les unités de base sont les cycles de modification covalente, qui comprennent l'activation et la désactivation des protéines par d'autres protéines. La modification des protéines dans la signalisation cellulaire – typiquement une phosphorylation et déphosphorylation – est un mécanisme général responsable du transfert d'une grande variété de signaux chimiques dans les systèmes biologiques. Bien que le concept ne semble pas être complexe d'un point de vue biochimique, ces systèmes simples peuvent néanmoins fournir une large gamme de réponses dynamiques et forment donc des blocs de construction importants pour les voies de signalisation. Ces cycles sont souvent liés, formant des couches de cycles multiples, appelées « cascades ».

Mitogen-activated protein kinase (MAPK) sont des instances communes observées dans la transduction des signaux à travers une série de réactions protéine-kinases. Ces voies se retrouvent dans presque tous les eucaryotes et jouent un rôle important dans le contrôle de différents processus cellulaires, y compris les fonctions fondamentales. En général, l'activation de la réponse cellulaire par la voie des MAPK implique au moins trois étapes de phosphorylation.

Afin de mieux comprendre la nature de cette régulation, ainsi que les mécanismes déterminant la fonction des cellules, des modules de signalisation ont été intensivement étudiés en se servant de modèles mathématique et de simulations informatiques, dans le domaine de la biologie des systèmes et de ses disciplines. L'objectif principal est de fournir une description authentique du système dans la mesure du possible et d'être capable de prédire son comportement. En accord avec les analyses expérimentales, les observations ont permis d'identifier les propriétés de les voies du système, telles la propagation rapide du signal, de fortes amplifications, un temps de signal court, et une résistance au bruit.

Etant donné que les paramètres biochimiques dans les voies de signalisation ne sont pas facilement accessibles expérimentalement, il est nécessaire d'utiliser des outils avancés de mathématiques pour une estimation correcte. Nous utilisons le paradigme de l'optimisation de systèmes artificiels pour la transduction du signal pour découvrir les conditions optimales de modules de signaux cellulaire.

Pour atteindre l'objectif principal de cette thèse, nous avons d'abord identifié les paramètres

clés du système à travers des analyses de sensibilité globale. L'analyse comparative des différences et des similitudes au sein des architectures de différents systèmes, a fourni les points de départ pour la classification des paramètres et la conception optimale du système.

Afin de pouvoir interpréter un large éventail de phénotypes, nous considérons à la fois les propriétés statiques et dynamiques. En outre, nous avons étudié des compromis entre les différentes caractéristiques optimales. Ainsi, nous avons trouvé les paramètres biochimiques et biophysiques qui permettent d'obtenir ces compromis et analysons s'il existait des conditions dans lesquelles nous pouvons simultanément réaliser une performance optimale, en état stationnaire et en régime dynamique. Nous analysons les principes de conception qui conduisent le système à avoir un temps minimal de signalisation, et également un certain niveau de gain d'amplification. Dans cette configuration, nous abordons problématique de recherche principale: y a-t-il une relation entre l'état d'équilibre et des propriétés dynamiques? En outre, nous tenons en compte les propriétés d'ultra-sensibilité déjà analysées et résolvons les problèmes d'optimisation multi-objectifs. Grâce à ce travail, nous avons pu constater que lors de la sélection judicieuse des paramètres cinétiques, un simple cycle de modification covalente est en mesure de répondre à plusieurs objectifs simultanément. En particulier, cette analyse pourrait aider à expliquer pourquoi les cycles de signalisation sont si omniprésents dans la signalisation cellulaire. L'amélioration de l'ultra-sensibilité et la propagation rapide du signal dans les systèmes multi-cycliques montre clairement les avantages du choix naturel de la conception des voies de signalisation en cascade.

La thèse se conclue par une discussion sur les prochaines étapes potentielles de recherche, et rassemble des connaissances plus quantitatives sur les voies de signalisation.

Mots-clés: biologie des systèmes, la transduction du signal, modélisation mathématique, analyse paramétrique, optimisation, analyse de sensibilité, simulations stochastiques.

CONTENTS

Acknowledgments	i
Abstract	iii
Résumé	v
Contents	vii
List of Figures	xi
List of Tables	xiii

Chapter 1: Exposition

Background and Scope.....	1
1.1 Systems theory + Biology = Systems Biology	1
1.1.1 Computational systems biology	2
1.1.2 Cellular signaling and its place in systems biology	2
1.2 Parameter estimation and research motivation	5
1.3 Thesis outline	7

Chapter 2: Meeting with Mathematics

Mathematical Modeling in Signaling	9
2.1 Model building	9
2.2 Models in systems biology	9
2.2.1 Quantitative vs. qualitative	10
2.2.2 Deterministic vs. stochastic	11
2.2.2.1 The Gillespie algorithm	14
2.2.3 Continuous vs. discrete	15
2.2.4 Static vs. dynamic.....	19
2.2.5 Modularity in cellular signaling.....	19
2.3 Parametric sensitivity analysis.....	20
2.3.1 Direct approach.....	20
2.3.1.1 The Sobol method.....	21
2.3.2 Inverse approach	24

Chapter 3: Ensembles in Signaling Symphony

Covalent Modification Cycle and Prototypical MAP Kinase Cascade.....	27
3.1 Covalent modification cycle – monocyclic system	27
3.1.1 Mechanistic modeling framework.....	28
3.1.2 Mathematical model of monocyclic system.....	28
3.2 Prototypical MAP kinase cascade – multicyclic cascade system.....	32
3.2.1 Mechanistic modeling framework.....	34
3.2.2 Mathematical model of multicyclic system	34
3.3 Design parameters	36

3.4	Design criteria.....	38
3.4.1	Ultrasensitivity.....	38
3.4.2	Amplification.....	40
3.4.3	Signaling times.....	40

Chapter 4: Pianoforte Sonata

Optimal Design for Ultrasensitivity.....	43
4.1. Materials and methods.....	43
4.2. Optimal design for ultrasensitive monocyclic system.....	45
4.3. The longer, the better: Optimal design for ultrasensitive tricyclic systems.....	48
4.4. Step back: Optimal design for ultrasensitive bicyclic system.....	50
4.5. Modular differences and similarities.....	51
4.6. Optimal design as entry point for sensitivity analysis.....	51

Chapter 5: Identifying The Violins in Signaling Orchestra

Sensitivity Analysis of Signaling Modules.....	55
5.1 Materials and methods.....	55
5.2 Sensitivity analysis of monocyclic system.....	56
5.2.1 Influence on ultrasensitivity.....	57
5.2.2 Influence on amplification.....	58
5.2.3 Influence on rise time.....	59
5.2.4 Influence on decay time.....	60
5.3 Sensitivity analysis of tricyclic system.....	61
5.3.1 Influence on ultrasensitivity.....	62
5.3.2 Influence on amplification.....	63
5.3.3 Influence on rise time.....	65
5.3.4 Influence on decay time.....	65
5.4 Sensitivity analysis as <i>prelude</i> for dynamic optimal design.....	66

Chapter 6: The Perfect Choreography

Optimal Design for Dynamic Properties.....	69
6.1 Optimal recipes for signaling times.....	69
6.1.1 Optimal design for signaling times under amplification constraints in monocyclic systems.....	69
6.1.2 Optimal design for signaling times under amplification constraints in multicyclic systems.....	73
6.2 The trade-offs and interplay.....	76
6.2.1 Optimal design for signaling times under amplification and ultrasensitivity constraints in monocyclic system.....	76
6.2.2 Optimal design for signaling times under amplification and ultrasensitivity constraints in multicyclic systems.....	78
6.3 Lessons learned from direct and inverse approach for parametric sensitivities.....	81

Chapter 7: *Dénouement and Encore*

Conclusions and Outlook	83
7.1 Thesis summary	83
7.2 Follow-up work and perspectives.....	84
7.2.1 Timing matters for sensitivity analysis	84
7.2.2 Avant-garde motif: noise.....	85
7.2.3 Drug delivery as optimal control problem	89
7.3 Closing intermezzo	91

Appendix: *The Backup Trio*

Supplementary Results	93
A.1 Global sensitivity analysis in optimal regime	93
A.2 Avant-garde motif: noise.....	97
A.3 Optimal control as a drug delivery problem.....	107

<i>Bibliography</i>	<i>111</i>
---------------------------	------------

<i>Curriculum Vitae</i>	<i>119</i>
-------------------------------	------------

List of Figures

Figure 1.1: The role of systems biology in drug discovery.....	3
Figure 1.2: Simplified input/output paradigm applied on signal transduction.....	4
Figure 2.1: Systems biology research cycle	10
Figure 2.2: Stochastic methods for chemically reacting systems	13
Figure 2.3: The Gillespie algorithm	15
Figure 2.4: Michaelis-Menten kinetics	18
Figure 2.5: The Sobol method	23
Figure 2.6: Numerical methods for nonlinear programming problems	25
Figure 3.1: Schematic representation of a covalent modification cycle	29
Figure 3.2: Schematic representation of tricyclic cascade	34
Figure 3.3: Design criteria.....	39
Figure 4.1: Computational framework for optimal design	44
Figure 4.2: Landscape of Hill coefficient values versus dimensionless Michaelis-Menten constants, in monocyclic system.....	46
Figure 4.3: Design for maximal ultrasensitivity in monocyclic system.....	48
Figure 4.4: Maximum Hill coefficient of tricyclic system	49
Figure 4.5: Maximum Hill coefficient of bicyclic system	50
Figure 4.6: Ultrasensitive responses of monocyclic system.....	52
Figure 4.7: Maximum Hill coefficient of tricyclic system, with the corresponding sets of kinase affinities.....	53
Figure 5.1: Computational framework for sensitivity analysis.....	56
Figure 5.2: Total sensitivity indices and scatter plots for each parameter of monocyclic system, with respect to ultrasensitivity	57
Figure 5.3: Total sensitivity indices and scatter plots for each parameter of monocyclic system, with respect to amplification	59
Figure 5.4: Total sensitivity indices and scatter plots for each parameter of monocyclic system, with respect to rise time	60
Figure 5.5: Total sensitivity indices and scatter plots for each parameter of monocyclic system, with respect to decay time.....	61
Figure 5.6: Total sensitivity indices for tricyclic system on ultrasensitivity, amplification gain, signaling rise and decay times	62
Figure 5.7: Scatter plots of the most influential parameters on ultrasensitivity of tricyclic system	63
Figure 5.8: Scatter plots of the most influential parameters on amplification of tricyclic system	64
Figure 5.9: Scatter plots of the most influential parameters on signaling rise time of tricyclic system	65
Figure 5.10: Scatter plots of the most influential parameters on signaling decay time of tricyclic system.....	66

Figure 6.1: Dynamic responses of monocyclic system for different levels of amplification.....	70
Figure 6.2: Design for minimal signaling rise time and decay time of monocyclic system, subject to given amplification.....	71
Figure 6.3: Design for minimal signaling rise time of multicyclic systems, subject to given amplification	74
Figure 6.4: Optimal design for ultrasensitivity in monocyclic system.....	77
Figure 6.5: Design for minimal signaling rise time and decay time of monocyclic system, subject to given amplification and ultrasensitivity	78
Figure 6.6: Design for minimal signaling rise time and decay time, subject to given amplification and ultrasensitivity	80
Figure A.1: Total sensitivity indices and scatter plots for each parameter of monocyclic system, with respect to ultrasensitivity in optimal regime.....	93
Figure A.2: Total sensitivity indices and scatter plots for each parameter of monocyclic system, with respect to amplification in optimal regime	94
Figure A.3: Total sensitivity indices and scatter plots for each parameter of monocyclic system, with respect to signaling rise time in optimal regime	95
Figure A.4: Total sensitivity indices and scatter plots for each parameter of monocyclic system, with respect to signaling decay time in optimal regime.....	96
Figure A.5: Design criteria for stochastic modeling.....	97
Figure A.6: Stochastic influence on steady-state responses in monocyclic system.....	98
Figure A.7: Stochastic influence on Hill coefficient values in monocyclic system.....	99
Figure A.8: Stochastic influence on effects of sequestration in monocyclic system.....	100
Figure A.9: Stochastic influence on maximal activation in monocyclic system.....	101
Figure A.10: Probability density functions of activation in monocyclic system	102
Figure A.11: Stochastic influence on standard deviation in monocyclic system.....	103
Figure A.12: Coefficient of variation in monocyclic system.....	104
Figure A.13: Stochastic influence on signaling rise time in monocyclic system.....	105
Figure A.14: Stochastic influence on signaling decay time in monocyclic system	106
Figure A.15: Example of therapeutical action in prototypical MAPK cascade.....	107
Figure A.16: Effect of different inhibitors on the dynamic responses of signaling network	108
Figure A.17: Drug delivery as optimal control problem.....	109

List of Tables

Table 3.1: Reactions in monocyclic system	30
Table 3.2: Reactions in monocyclic system and multicyclic cascade systems.....	35
Table 3.3: Kinetic equations comprising the mechanistic models	36
Table 3.4: Definition of dimensionless variables and parameters.....	37
Table 3.5: Dimensionless mechanistic models	38
Table 4.1: Maximal Hill coefficient of monocyclic and tricyclic systems, from semi-quantitative analysis.....	53
Table 5.1: Sensitivity analysis as <i>prelude</i> for dynamic optimal design.....	67
Table 6.1: Modular differences and similarities: optimal design for signaling rise time of monocyclic, bicyclic and tricyclic systems.....	75
Table 6.2: Modular differences and similarities: optimal design for minimal signaling rise and decay times and maximal ultrasensitivity of monocyclic, bicyclic and tricyclic systems.....	81
Table 7.1: Concentrations of different protagonists of MAPK signaling pathways, obtained experimentally in various cell types	86
Table 7.2: Relation of number of molecules and concentration in our study	87

Exposition

Background and Scope

1.1 Systems theory + Biology = Systems Biology

Whether in prokaryotic or multicellular arrangement, the cell has been the subject of intense investigation and prolific scientific undertaking. While imaging techniques continue to be developed for scrutinizing cellular mechanisms at deeper levels both *in vivo* and *in vitro*, computer simulations have summoned forth a new tool that has allowed for the interrogation of the cell's inner-workings.

Despite its infancy to be widely noted by the biological community, simulation begins to find its place among the leading methods of inquiry. While results from experiments remain the empirical standard, there seems to be an increase in the number of research groups using simulation to formulate hypotheses and make predictions about cellular mechanisms. Putting aside the rising cost of wet-lab materials and the comparatively inexpensive cost of software and hardware, one would expect to find more and more labs moving in this direction since computer simulations are an excellent way to gain quantitative insight into rather complicated phenomena.

Emerging from this collaboration between experiment and the theory is a field generally known as systems biology, which has already made substantial breakthroughs in understanding some of the major problems in biology and chemistry.

Field of systems biology aims at investigating the behavior and relationships of all elements in a particular biological system (1). The central task is to comprehensively gather information from each of the distinct levels of organization for individual biological systems and to integrate these data to generate predictive mathematical models of the system. System level understanding of biological system can be derived from insights into system structure, system dynamics and control and design methods (2). Progress in any of these components requires collaborative efforts from computational sciences, genomics and measurement technologies, and integration of discoveries with collected knowledge.

Some of the questions that fall into the scope of systems biology are:

- How does life behave as a system?
- How does each cellular constituent act within its environment?
- How do cellular constituents interact with each other?
- Do cellular constituents exhibit any particular features?
- Is there any predictability of the behavior of different cellular constituents?

1.1.1 Computational systems biology

One sub-discipline of systems biology is computational systems biology. It is focused on the development of efficient algorithms, customized data structures, visualization and communication tools and much more, all with the goal of improved use of computer simulations of biological systems (3). Improvements in both algorithm efficiency and computer power are likely to increase our capabilities in tackling massive computation tasks in the future. The daunting task for analysis is the increasing complexity of models approaching complexity of the biological systems, which these models are used to approximate.

Computational systems biology has two distinct directions: knowledge discovery, which extracts the patterns from huge quantities of experimental data, forming hypotheses as a result; and simulation-based analysis, which tests hypotheses with *in silico* experiments, providing predictions to be tested by *in vitro* and *in vivo* studies (4).

In order to engage systems theory and biology in a long-lasting relation, both involved sides need to be ready to compromise and accept the advantages of the partnership. Systems specialists need to realize that biology is not just another application area of standard systems theory and that the transfer of the concepts from physical to living systems is not so straightforward (5). On the other hand, the biologist should gain more confidence and accept the view that theoretical models and computer simulations can be useful to address the dynamic behavior of complex regulatory networks in biological systems (6). Only this way the overall objective of collecting the knowledge in a symbiotic way and closing the loop of a collaborative efforts as a foundation of systems biology approach can be achieved.

The systems biology approach, with its combination of computational, experimental and observational enquiry, is also highly relevant to drug discovery and the optimization of medical treatment regimes for individual patients (4). Indeed, a more widespread collaboration between mathematicians, computer scientists, physicians and experimental scientists is constantly improving drug discovery (7). Furthermore, this will allow pharmaceutical companies to both increase their lead compound and clinical candidate portfolios and satisfy market demands for continued innovation, leading to revenue growth (8). Fig. 1.1 illustrates the simplified cycle of systems biology steps in drug discovery.

Therefore, it is easy to understand why Systems Biology has become such an intensive field of research, enabling to understand the key functions of gene transcription, metabolic pathways and signal transduction. The latter will be the focus of this thesis.

1.1.2 Cellular signaling and its place in systems biology

Cells, like all living organisms, must be able to receive, transmit, and emit information through signals. Whether it is an isolated prokaryote or a part of a tissue, all cells must be able to accept signals from their external environment, process this signal and make decisions based on it. Over the years, biologists have shown that these processing mechanisms oftentimes take place in the form of signaling pathways (9). The activity of these protein-based networks further influence gene expression, regulate cell growth, mitogenesis, differentiation, embryo development, and stress responses in mammalian cells, whereas the malfunction is in correlation with diseases, such as cancer, asthma

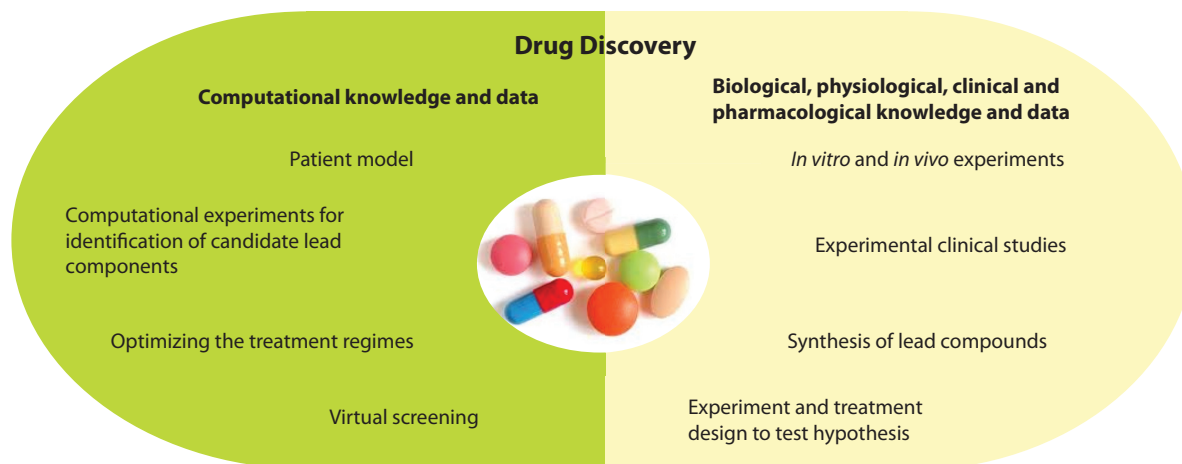


Figure 1.1: The role of systems biology in drug discovery. The application of systems biology in drug discovery and the design of multiple drug therapies and therapeutic gene circuits is believed to be the future of the medicine (4). Imposing the system-oriented view, this iterative cycle of hypothesis and simulation-driven processes is changing the way treatments are conducted.

and diabetes (10).

Within the cohort of subfields that now comprise systems biology, one finds the subject of cellular signaling. A stand-alone subject within cellular biology, signaling is an excellent candidate for computational analysis, for, at its core, cell signaling deals with the cellular mechanisms of information processing. The latter has been the privileged study of electrical engineers and computer scientists for nearly half of a century, and all the techniques and methods accumulated under those banners are set to make – and to a large extent already have made – an enormous contribution to our knowledge about the cell functioning.

Cellular signal processing generates a vast amount of interesting properties and it can be highly complex. Until recently, cellular signaling was only analyzed through linear pathways, connecting the cellular periphery to the metabolic and genetic machinery. Nevertheless, some of these pathways have been shown to interact with each other. This phenomenon is called signal *crosstalking*. As increasingly large numbers of signaling components are being identified, it has become clear that these pathways are not isolated and *crosstalking* is nothing more but a synonym for signaling network (11). Modeling these systems is not always straightforward; there is a high level of uncertainty, which cannot be neglected when building assumptions for a mathematical model.

From a system point of view, one can envision cell signaling as simplified input/output constellation (Fig. 1.2). Different external stimuli, such as signals from other cells, different hormones, survival factors, growth factors or cytokines, are sensed by receptors that are situated on cell surface. The complex network of signaling pathways do not only transmit the received signals, but process, encode and integrate both external and internal signals. Depending on their specificity, the output signals affect different cellular processes. The three elements receptors, signaling pathways and gene transcription are described subsequently.

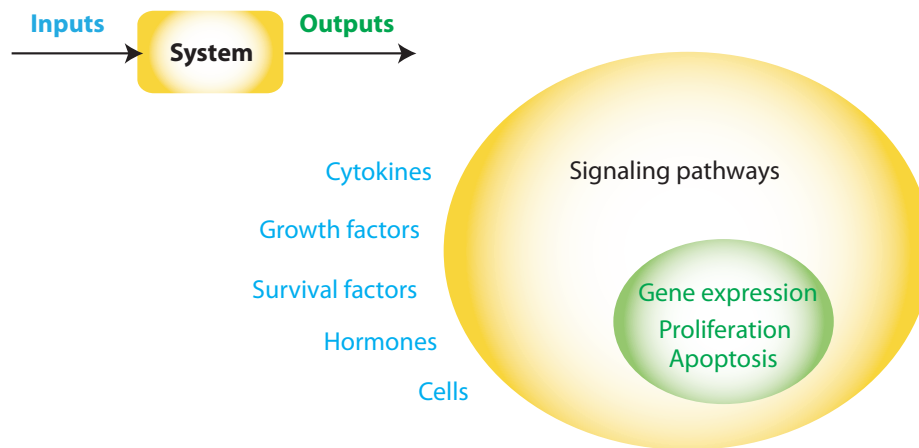


Figure 1.2: Simplified input/output paradigm applied on signal transduction. Activation of a signaling pathway usually involves binding of a ligand – hormones, mitogens, growth factors or neurotransmitters – to its cognate receptor, followed by phosphorylation or dephosphorylation of a target protein on tyrosine, serine or threonine residues by the appropriate downstream protein kinases or phosphatases. Modules of protein kinases further control cellular processes, such as gene expression, cell growth, mitogenesis, differentiation, embryo development, and stress responses.

Receptors. The receptors are the entry points of the signal to the cell. There are several families of receptors, each triggering different functions (12):

- Ion channels are transmembrane proteins that enable small molecules, usually ions, to cross the membrane. Their main function is to regulate the electric potential across the membrane, which is used to carry fast information mainly in neurons.
- G-protein coupled receptors (GPCR) are coupled with G-proteins, as their name indicates. The G-protein is activated with the binding of a ligand to the GPCR and can in turn activate or inhibit an intracellular enzyme.
- Receptors tyrosine kinases (RTK) are activated by cytokines, interferons and human growth factors. They lack intrinsic activity, but the ligand binding stimulates the formation of a dimeric receptor, which can then activate several protein kinases.
- Receptors with intrinsic catalytic activity are transmembrane proteins, which gain intracellular catalytic properties upon a binding of a ligand. They can act as phosphatase, kinase, or convert GTP to cGMP. They are common receptors for growth factors and insulin.

Intracellular signaling pathways. The internal signal carries the information to the genes. It can be described as a huge network of different pathways. A pathway can be activated by different receptors and the same pathway can activate different functions. There are large numbers of pathways and a selected set of them is presented here:

- The Akt signaling pathway can be activated by multiple receptors. The Akt is also called protein kinase B, which can bind to phospholipids, enabling anchoring to the cell membrane. The complex then triggers the phosphorylation of many downstream kinases. The Akt pathway is involved in processes regulating the glucose metabolism, growth and apoptosis. It has become an important target for the treatment of cancer, diabetes, stroke and neurodegenerative disease (13, 14).
- The NF- κ B signaling pathway plays a key role in regulating the immune response to infection. It is composed of NF- κ B proteins and, when the cell is not stimulated, they

remain inhibited by the protein I κ B. The signal can be triggered from many receptors, usually by cytokines and growth factors. The receptor activates and IKK protein, which is an I κ B kinase. The phosphorylation of the I κ B releases the NF- κ B from its inhibited state and enters the nucleus where it promotes the expression of specific genes. This pathway is involved in adaptive immunity, inflammations, stresses and B cell development (15, 16).

- The Jak/Stat signaling pathway is an important part of the cytokine signaling, but many other molecules can trigger it, as well. Like the NF- κ B signaling pathway, it is also an important signaling pathway of the immune system. Jaks (just another kinase or Janus kinase) are attached to the membrane receptors and can phosphorylate tyrosines. The binding of ligand provokes the dimerization of the receptor and the activation of Jak proteins. These latter phosphorylate themselves and the receptor, creating docking sites for Stats proteins, which will be also phosphorylated by Jak proteins. Stats then dimerize and migrate to the nucleus, regulating the gene expression (17, 18).

- MAPK signaling pathway is a highly studied and very important pathway in most cells. The mitogen-activated protein kinase (MAPK) pathway, also known as the MAPK protein cascade since the pathway looks like a cascade in which key enzymes flow down the waterfall to yield a final output quantity. Alongside a whole host of other very significant pathways, MAPK has been found to regulate cell death (also known as cell apoptosis), cell differentiation, and cell division. All three of these cellular processes are crucial to the development of cancers, which explains why the MAPK cascade has found itself at the cross-hairs of nearly every drug company that is interested in cancer medication (19, 20). The focus of this thesis will be wrapped around investigation of design principles in prototypical MAPK pathways and its constitutive elements. More details about these pathways will be given in subsequent chapters.

Gene transcription. The ability to adapt to environmental constrains is a matter of survival and the expression of genes has to be adjustable. The nucleus is, therefore, the destination of most of the signal transduction pathways. The regulation of genes is a complex field and possesses a large variety of mechanisms. The signal transduction is mainly controlling the expression of the genes at the initiation step and it happens in different ways. For example, to replace the initiation complex by the elongation complex, covalent modification, such as phosphorylation, are needed. This is often performed by a kinase, which is activated by signal transduction. Regulatory proteins described for the initiation phase are sometimes called transcription factors. They can bind to specific sequences of DNA, and promote or enhance the transcription of a particular gene, and they are often the downstream element of a signaling pathway.

1.2 Parameter estimation, optimization and research motivation

*“If one way be better than another, that you may be sure is nature’s way.”
Aristotle*

Systems biology and related disciplines opened a whole new avenue of possibilities for using different mathematical tools in order to facilitate further understanding of cellular processes and drug development. Since Systems Biology as a field was born, lots of attention was placed on simulating the results of experimental measurements or output variables for a

given system with some given set of parameters. However, the development of the systems biology predictive models requires the information about real kinetic parameters and species that exists in signaling pathways. Unfortunately, it is very difficult or sometimes even impossible to gather information about values of kinetic parameters or concentration profiles of species in an experimental setting. Furthermore, some of the mentioned entities can be prone to large variations, depending on how the experimental setting was performed. Therefore, there emerges a need of utilization of powerful mathematical methods for successful parameter estimation. Up to date, little effort has been devoted in this direction.

A common routine applied in systems biology consists in using available experimental data (so called “dose-time matrix” data), in order to calibrate mathematical model so that it mimics these data in the best possible way. This parameter identification task is often formulated as optimization problem with the objective of minimizing the difference between predicted and experimental values. To compare model results with experimental data, one first has to simulate the mathematical model to produce these results – the *forward* problem; and then to estimate the parameters – the *inverse* problem (21). Unfortunately, there are a number of difficulties involved, mostly arising from the fact that models are quite complex. Thus, there exists a need for superior time integrator in order to simulate the model properly. This integrator should be able to fulfill its task fast, since the model will be evaluated many times. Furthermore, it should be robust, giving the fact that a large parameter and state space will be explored and there is a high probability of different numbers and ranges of time scales involved. Nevertheless, these difficulties are just the starting point along the path of successful parameter estimation. Identifiability analysis aims of providing important information whether the parameters for the mathematical model can be determined at all, after obtaining an infinite number of observations from it (22). The data fitting process typically starts with an arbitrary guess about parameter values (nominal set) and then changes those values to minimize the mentioned discrepancy between model and data. Minimization of this difference is set as optimality criterion. The criterion selection will depend on the assumptions about the data disturbance and on the amount of information provided by the user. One of the main pitfalls occurring in inverse part of the problem is that, in general, models in systems biology have multiple sets of parameters that could satisfy the desired behavior and some of those sets are pinpointing to *local* solutions in contrast to *global* (unique) solutions that would be of interest. This comes as a consequence of the nonlinear and constrained nature of the systems dynamics. In mathematical jargon, problems of such a nature are said to be multimodal (non-convex). Thus, lot of methods may fail to identify the global solution and may converge to a local minimum when a better solution is just a small distance away. Moreover, in the presence of a bad fit, there is no guarantee whether this is the aftereffect of a wrong model formulation or simply due to local convergence.

Described process of model calibration *a priori* assumes at least some existence of experimental data. The high disbalance between available measures and unknown entities makes the whole problem of parameter estimation even more complex. For instance, some of the signaling networks consist of around 100 species and parameters, whereas the experimental data set is composed of dose-time matrices for 2–3 quantities. But, what to do when experimental data is not available at all? What is the right way to, at first, formulate such a problem, and then to solve it? These and much more other questions and challenges, both from the perspective of biology and systems theory, were intriguing enough to conceive the idea of this thesis.

It was stated many years ago that any approach towards the explanation of the kinetic parameters and the structural design of biological systems are the outcome of evolution

(23). An enormous literature exists on the natural selection and evolution of biological systems ranging from the reconstruction of evolutionary mutation trees up to the study of forces acting on evolution of whole organisms or populations. In many of these studies evolution is considered as an **optimization** process (24-26). Minimizing investment cost, maximizing profit, minimizing risks, maximizing effectiveness, minimizing energy, maximizing yield and much more are the phrases that are frequently used in everyday vocabulary. People are constantly trying to find the best strategy, given the specific constraints, in order to reach a goal. Optimization – the framework for defining this whole process – becomes omnipresent in different areas of human activity and interest. The base premise of this thesis is that the nature, as an ultimate designer, constructed cellular processes so that they also tend to behave optimally.

If we were about to assemble the electronic device that transduces signals, we would have good idea of its functionality. Signal processing, as an area of systems engineering, electrical engineering and applied mathematics, would give theoretical and practical guidance for implementation of such a device. Observing different kind of signals that we are encountering in our everyday lives, such as sounds, images, videos, telecommunication signals etc., we could define how a good or a perfect signal should look like. The signal should be clear (noise-free), it should be amplified enough so that we can be aware of it, it should arrive at its destination in reasonable time, etc. Even though the biochemical signals are occurring through chemical reactions and conformational change of species is considered as information that the signal is carrying, the same paradigm could be applied. Thus, our research path is leading to the discovery of optimal design in cellular signaling modules. We perform the full parameter estimation, which means that we did not use any available experimental data in order to guide the parameter search. Instead, we use engineering archetype for signals and we define desired systems behavior accordingly. Furthermore, we aim at exploring design principles that determine the dynamics of signal transduction, which adds additional complexity toward achieving the goal.

State-of-the-art approaches and efforts made with different tools from systems biology arsenal will be commented throughout subsequent chapters.

1.3 Thesis outline

The thesis is organized in two main parts. The first part (chapters 1–3) serves as extended description of different ways of defining the models in systems biology and introduces the signaling modules that are addressed during our study. The second part (chapters 4–6) presents the results and findings that emerge from our analysis.

The structure is as follows:

- In chapter 2, we describe main modeling concepts for signaling networks, which are extensively used through the scope of Systems Biology. In particular, we focus on approaches and methods that will be further used to answer our different research questions.
- Chapter 3 gives a general introduction to the covalent modification cycle and prototypical MAPK cascade – the two signaling patterns that will be explored throughout the thesis. The derivation of the models for these instances is presented, as well as the properties that we are interested in.

- In chapter 4, we start utilizing optimization framework to identify the design principles for optimal ultrasensitive signaling modules. Optimization applied here considers analysis of steady state, so the overall problem complexity remains within acceptable limits. The difficulties arise when we try to gain knowledge about the design principles of the signaling modules that are set to satisfy specific optimality criteria in dynamic regimes.
- In order to achieve the main thesis objective, we first proceed with identification of the key system parameters, done through global sensitivity analysis in chapter 5. Comparative analysis of differences and similarities within different system architectures reveals some insight for initial parameter classification.
- In chapter 6 we return back to the optimization framework. We first analyze what are the design principles that lead the system to have the minimal signaling times, subject to a certain level of amplification gain. In this setup, we bring out our main research question: are there any trade-offs and interplay between different steady-state and dynamic properties? Furthermore, we include already analyzed property ultrasensitivity and eventually solve multi-objective optimization problems.
- Discussion and outlook are given in closing chapter 7.

Meeting with Mathematics

Modeling in Signal Transduction

2.1 Model building

Being able to provide predictions of emergent network properties and to uncover the principles of cellular networks by merging prior knowledge with experimental data and model simulation, brought systems biology to become a powerful tool for studying the dynamics of the biological systems. Mathematical formalism and its ability to describe nature in a comprehensive way to humans allowed Systems Biology to persist in its role of helping us better understand complex network of metabolic and signal transduction pathways.

The collaborative process of model building (Fig. 2.1) consists of repetitive communication steps between experimental scientists and modelers. Once it is clear what the scientific question should answer, the best practice for the first step in model building considers *pen-and-paper* model representation. All the involved scientific sides should be able to articulate their perspective on the biological phenomena, so that biological and physiological data and knowledge can be translated into universally understandable version of the model. For instance, if we are considering a mechanistic model then ultimately the model should describe all biochemical reactions that determine the observed biological process. Next, the hypotheses that should be tested need to be integrated into the model. Comparison of model outputs with available experimental data is the iterative procedure that is employed in different steps of model building. Another important phase includes exposing the model to different tests. This should mature the infantile stage of the model and improve the model relevance.

One more benefit of this ping-pong communication could lead to improved experimental design and eventually decrease the overall cost and time invested in experiments.

2.2 Models in systems biology

Most models in systems biology can easily be located within the space that is spanned by four dimensions of modeling, namely, qualitative vs. quantitative; continuous vs. discrete; stochastic vs. deterministic, and static vs. dynamic (27). These dimensions are not entirely independent nor are they exclusive. Many modeling approaches are hybrid

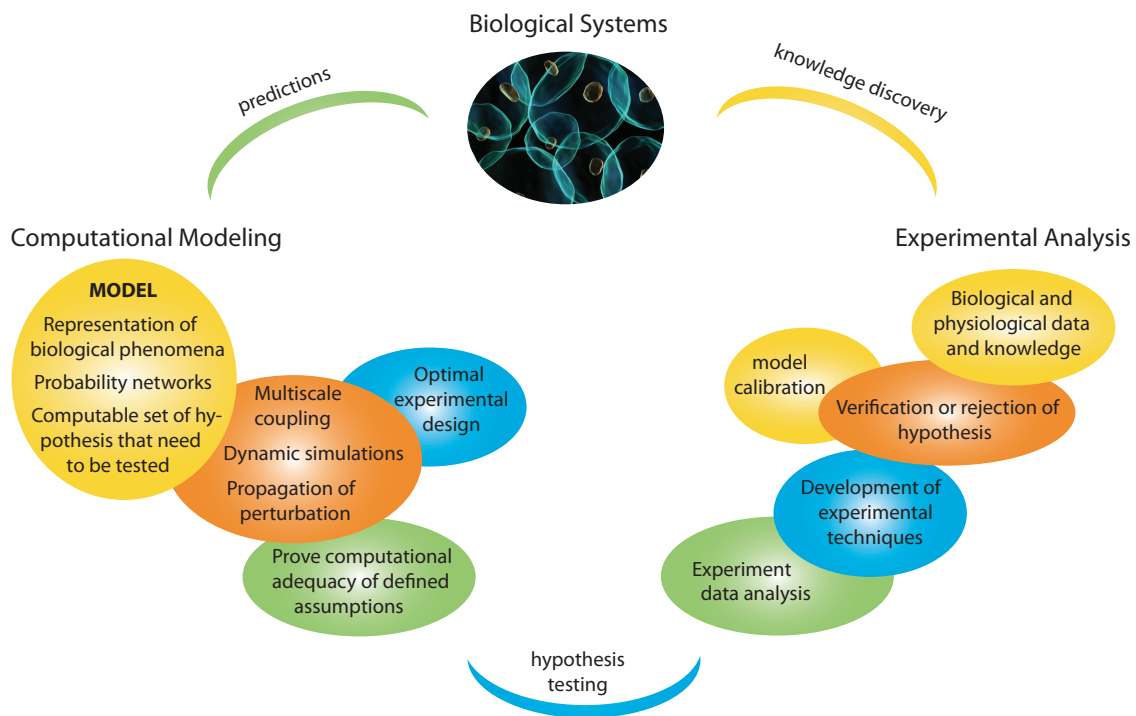


Figure 2.1: Systems biology research cycle: the closed-loop interaction between model building and experimental verification.

as they combine these different modeling aspects.

Literature offers large spectra of models of signaling pathways (28). Research groups of Peter Sorger and Douglas Lauffenburger have high reputation in scientific circles dealing with signaling. Over the years, they have shown quite a number of studies. Among other approaches, they proposed using Bayesian network analysis (29). The network model is a probabilistic graphical model that represents a set of random variables and their conditional dependencies via a directed acyclic graph. Also, they have been using Boolean logic to convert literature-derived signed protein signaling networks into a computable model (30). Furthermore, this group of scientist developed fuzzy logic modeling tailored to biologic networks (31). Modeling efforts done by other researchers will be noted in consecutive sections.

2.2.1 Qualitative vs. quantitative

Often the lack of quantitative data motivates the use of qualitative methods. Qualitative methods are usually built as a first step to develop a quantitative model. This way, one gains useful structural information by determining what variables play a role for certain kinetics and whether there exists any correlation between variables. Another motivation for the application of qualitative methods is that they are aimed at answering different kinds of questions than quantitative methods and offer different possibilities for analysis.

Models of signal transduction were addressed qualitatively (32, 33) but in a rather inferior frequency than quantitatively (i.e. (34)). Our analysis will be based on simulations of quantitative models.

2.2.2 Deterministic vs. stochastic

The most realistic way of simulating a chemically-reacting system is to use molecular dynamics, which involves keeping track of the spatial position and velocity of each particle and allowing for chemical reactions whenever two chemical species are within the same binding radius of each other (35, 36). The reactions involved may be unimolecular (meaning that a single species transforms into another) or bimolecular (meaning that two species combine to form a third, or a third and forth), while any reaction of higher order (trimolecular, for example) can usually be reduced to sets of bimolecular and unimolecular reactions (37).

The reason that molecular dynamics simulations are not used in every situation is that they are computationally quite intensive. Even with the world's best super-computers, simulations can barely exceed the mark of several thousand particles over a simulation time that exceeds a few microseconds (depending on the type of system being simulated) (38).

As a substitute for molecular dynamics, many chemists and biologists have turned to the so-called *deterministic* or classical approach. Our study will be dealing with deterministic models. A brief demonstration of stochastic modeling in the same domain will be given as a part of future research.

The classical formalism for chemically-reacting systems is to treat the number of molecules of a set of species S_i , for $i=1,2,\dots,N$, as a continuously-varying $X_i(t)$ that satisfies a set of coupled ordinary differential equations (ODEs) of the form:

$$\frac{dX_i(t)}{dt} = f_i(X_1(t), X_2(t), \dots, X_N(t)), \quad \forall i \in \{1, 2, \dots, N\}$$

where the specific f_i are taken from the system. This set of differential expressions is called a reaction-rate equation (RRE) (36, 37). Astonishingly enough, simulations based on the RRE work quite well.

However, there are at least three good reasons to withhold from using the RRE:

- Quantum mechanics play a not-altogether-trivial role in unimolecular and bimolecular reactions, in so much as quantum effects can alter the isomeric form of the resulting product. This effect results in some noisiness in the output data (36).
- Chemical systems are in thermodynamic equilibrium within some kind of heat bath. By contributing thermal fluctuations, the heat bath makes the overall species populations look less like smooth curves and a lot more like the index of a stock exchange (36).
- As often happens in biology, the number of molecules of a certain species may be no more than a few tens or hundreds. In these cases, the former expression is inaccurate since it does not take into account the stochasticity of the molecular species (39).

Stochastic chemical kinetics emerged nearly thirty years ago as a series of methods for considering these reactions in a way that is analogous to the RRE but somehow accounts for the inherent noise. Beginning with the same species S_i , $i=1,2,\dots,N$, and M chemical reactions R_1, R_2, \dots, R_M it considers the number of molecules $X_i(t)$ for a specific i in a volume V and analyzes the state vector $\mathbf{X}(t) = [X_1(t), X_2(t), \dots, X_N(t)]$, given the initial conditions $\mathbf{X}(t_0) = \mathbf{x}_0$ (37).

Instead of treating each molecule as an independent unit with a unique position and velocity at a certain time, the stochastic approach only deals with the total number of

molecules of a given species. It can disregard the position and velocity computations since thermal equilibrium guarantees that positions are uniformly distributed while velocities stabilize around the Maxwell-Boltzmann distribution. This assumption of thermal equilibrium is what makes stochastic chemical kinetics much faster than the usual molecular dynamics approach and still scientifically accurate: molecular dynamics simulations spend most of their computation time on non-product-yielding molecular collisions while the stochastic approach only focuses on events that can change the overall molecular population.

Each reaction R_j , $j = 1, 2, \dots, M$, is going to depend on two quantities:

1. The state-change vector: A state-change vector $v_j = [v_{1,j}, v_{2,j}, \dots, v_{N,j}]$ indicates how a state should change when a reaction R_j takes place. In particular, $v_{i,j}$ is the number of molecules that species S_i obtains from reaction R_j (this number may be negative). So, if the system is in state \mathbf{x} and reaction R_j occurs, then the system changes its state to $\mathbf{x} + v_j$.
2. The propensity function: A propensity function a_j is defined so that $a_j(\mathbf{x})dt$ is the probability that one R_j reaction will take place in a volume V and in the next infinitesimal interval $[t, t + dt]$, given $\mathbf{X}(t) = \mathbf{x}$. For unimolecular reactions, the quantum mechanics of molecular collisions gives the probability for a reaction $S_1 \rightarrow g$ to be $c_j x_1 dt$, for some constant c_j that depends on the reaction. The result for bimolecular reactions is similar: $c_j x_i x_k dt$. The propensities for unimolecular and bimolecular reactions are therefore $c_j x_i$ and $c_j x_i x_k$, respectively.

The formula for propensity is probabilistic, so it is natural to ask if we can find a formula for $P(\mathbf{x}, t | \mathbf{x}_0, t_0)$ (37). The time evolution of that probability is called the Chemical Master Equation (CME):

$$\frac{\partial P(\mathbf{x}, t | \mathbf{x}_0, t_0)}{\partial t} = \sum_{j=1}^M [a_j(\mathbf{x} - v_j)P(\mathbf{x} - v_j, t | \mathbf{x}_0, t_0) - a_j(\mathbf{x})P(\mathbf{x}, t | \mathbf{x}_0, t_0)] \quad (2.1)$$

A priori, the CME gives the whole time evolution of the conditional probability $P(\mathbf{x}, t | \mathbf{x}_0, t_0)$. Since it is a set of coupled ODEs, it is not tractable in most cases except for a few analytically well-defined cases. Interestingly, in the case of a completely noiseless process, the CME becomes the RRE, which shows that the stochastic kinetic approach is really the more general case of the deterministic approach. But since the CME is quite difficult to solve, it is better to look for numerical schemes capable of approximating the state vector $\mathbf{X}(t)$. To that end, it is valuable to build a slightly different conditional probability $p(\tau, j | \mathbf{x}, t)d\tau$ which can be defined as the probability that a reaction will take place in an infinitesimal time interval $[t + \tau, t + \tau + d\tau]$ and that the reaction will be R_j . Formally, this function is the joint probability density function of the two random variables time to the next reaction (τ) and index of the next reaction (j), given that the system is currently in state \mathbf{x} . A more explicit version of this same formula can be found with basic probability (37):

$$p(\tau, j | \mathbf{x}, t)d\tau = a_j(\mathbf{x})e^{-a_0(\mathbf{x})\tau} \quad (2.2)$$

$$a_0(\mathbf{x}) = \sum_{j=1}^M a_j(\mathbf{x}) \quad (2.3)$$

Many different algorithms can be built from this stochastic formulation, but there are three basic families, namely, fully stochastic, stochastic differential equations (SDE), and spatial-stochastic algorithms (40). Some of the representative methods of these classes are given in Fig 2.2. More details about fully stochastic methods will be given subsequently.

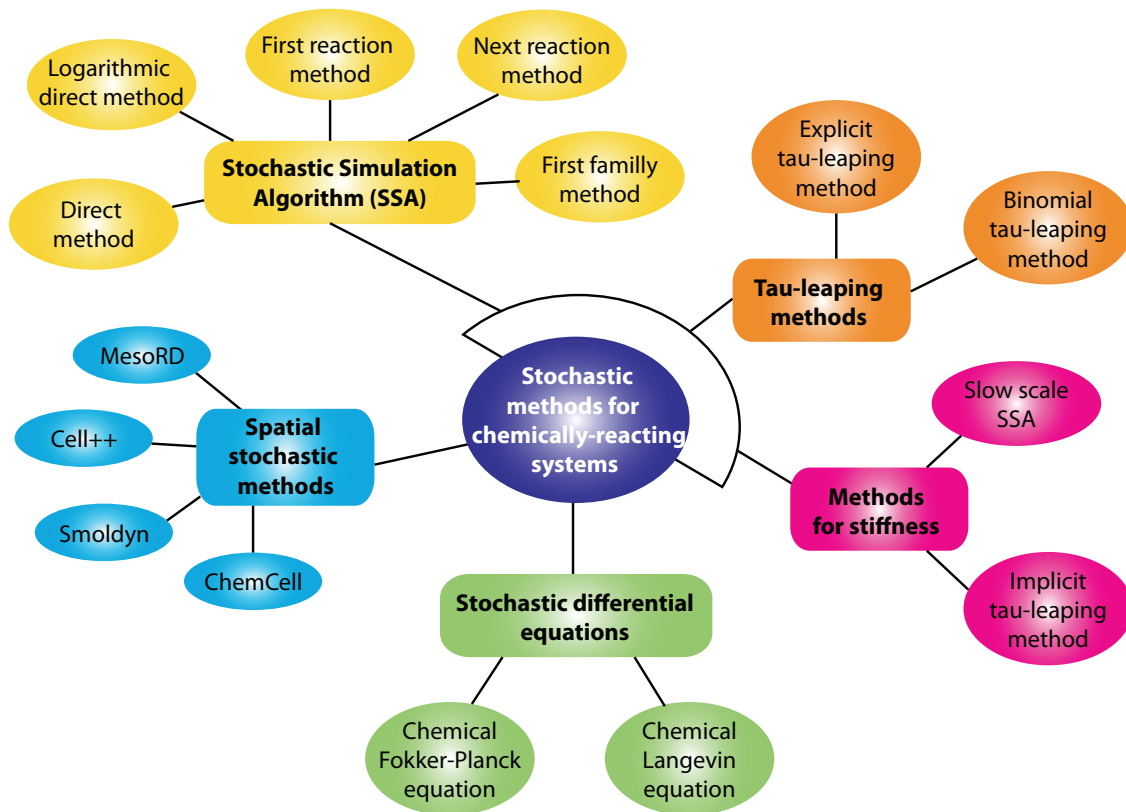


Figure 2.2: Stochastic methods for chemically reacting systems.

Stochastic differential equations (SDE) are often used when the formulation with ordinary or even partial differential equations does not reflect the nature of the simulated system so well (40). SDE maintain the appearance of an ODE or PDE but add a noise term in order to achieve a higher level of realism. The noise term may either be intrinsic (meaning that it emerges from the system) or extrinsic (meaning that it comes from external perturbations). The two popular implementations for biochemical reaction networks are the Chemical Langevin Equation (41) and the Chemical Fokker-Planck Equation (CFPE) (40).

Spatial stochastic methods are defined by their use of space as an analytical tool. As a general rule, these methods have been developed by biologists interested in simulating networks within the cell, where macromolecular crowding and other spatial effects are very important (42). Most of the available software that implement the techniques come from academia and are open-source. The advantages of this type of simulation are that the dynamics of the system are realistic and that larger systems can be simulated than can be with pure molecular dynamics. On the flip side, the spatial-stochastic simulations are among the most computationally intensive because even though they do not compute the position and velocity, they still keep track of the position and velocity for later statistical analyses. However, these methods seem to be the future of biological computing in the cell, since

spatial effects seem far too important to be neglected (42).

In past years, researchers have been exploring influence of stochastic noise in signaling pathways. In year of 2000, Paulsson *et al.* presented the concept of stochastic focusing in order to see how the signal noise influences the sensitivity amplification of threshold mechanism (43). Following the study of Goldbeter and Koshland (44), later same year this group of authors showed that zero-order ultrasensitivity is invariably coupled to large number of fluctuations in systems with low number of molecules (45). Thattai and Oudenaarden showed that these fluctuations could be attenuated if the signaling architecture includes cascades (46). Bhalla combined stochastic and diffusion effect on individual pathways and synaptic network properties (47, 48). Morishita and coworkers identified the optimal number of molecules for signal amplification and discrimination and they have been studying transient behavior of linear signaling cascade (49). Recently, Koepl and colleagues proposed unified framework for estimation of stochastic rate constants, accounting for extrinsic and intrinsic noise (50).

2.2.2.1 The Gillespie algorithm

The fully stochastic methods are all derivatives of the same algorithm, popularly known as the Gillespie algorithm after its inventor, but also widely referred as the SSA for Stochastic Simulation Algorithm or the *direct method* (37). Relying on central ideas of Monte Carlo theory, it draws two random numbers r_1 and r_2 from the uniform distribution on the unit interval, and uses them to compute t and j as follows:

$$\tau = \frac{1}{a_0(x)} \ln\left(\frac{1}{r_1}\right) \quad (2.4)$$

$$j = j \mid \sum_{i=1}^j a_i(x) > r_2 a_0(x) \quad (2.5)$$

Then the full-blown SSA follows along these lines (Fig. 2.3):

1. Start the system simulation at time $t = t_0$ and $\mathbf{x} = \mathbf{x}_0$
2. For a state \mathbf{x} at time t , evaluate all the propensities $a_j(\mathbf{x})$ and their sum $a_0(\mathbf{x})$
3. Compute t and j using the aforementioned methods
4. Produce the next reaction by allowing the time ($t \rightarrow t + \tau$) and the state-vector ($x \rightarrow x + v_j$) to change
5. Record $(\mathbf{x}; t)$ if needed, and start again at step 1 until the simulation time T is reached.

In his landmark 1977 article, Gillespie himself points out the strengths and weaknesses of his algorithm:

Strengths

- The SSA is exact, meaning that it is not based on averaging or other heuristics.
- There are no approximations of infinitesimal time steps. Instead, the time steps are as small as is necessary to allow for the next reaction to take place.
- Relatively easy to code
- Information about the individual ensemble behavior is not hard to extract.

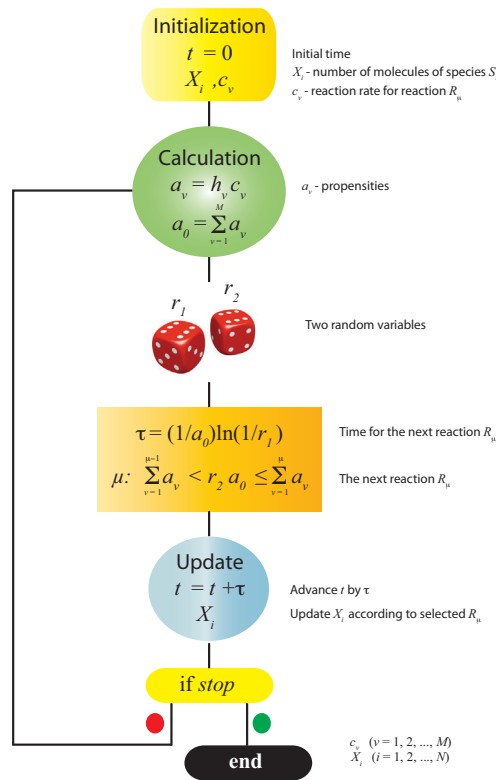


Figure 2.3: The Gillespie algorithm.

Weaknesses

- It requires long simulation time (not memory).
- The number of particles that one can simulate is limited. Beyond a certain number, the simulation may never converge.
- The SSA requires a good random-number generator. It may even require multiple number generators for crosschecking purposes.
- For statistics, multiple runs of a single SSA execution may be needed, which adds to the overall simulation time.

In an attempt to keep the strengths and remove some of the weaknesses – especially the long simulation time – many researchers have introduced modifications changes to the SSA, but at the expense of accuracy. Fig. 2.2 summarizes the main classes of stochastic methods.

Our clear requirement that all stochastic simulations should be exact made the choice of stochastic methods easier. Out of these methods, the Gillespie approach was by far the simplest to code and the best established of all the SSA-based algorithms, even though it is computationally inefficient. Though it was originally feared that the intracellular medium does not satisfy the assumptions of the Gillespie algorithm (namely, thermal equilibrium and mass-action kinetics), it was decided that these assumptions could be made for the sake of simplicity.

2.2.3 Continuous vs. discrete

The most common way of defining the entities of the model accounts for the representation

of the state of a molecular component as its concentration in some cellular compartment (cytoplasm, plasma membrane, etc.), which is treated as a function of time. The temporal dynamics are then described by an ordinary differential equation (ODE) for the net rate of production of the species. The model will comprise differential algebraic equations (DAEs) as well, but only in case when conservation relations are explicitly involved in its description. The model designed under this fashion is called a *continuous model*. Inspiration for this kind of modeling was taken from how the biochemistry of enzymes has been modeled (51).

Mass action kinetics. For example (52), the reversible reaction with three species involved:



will, in the most general notation, result in three differential equations:

$$\frac{dS_1}{dt} = -v, \quad \frac{dS_2}{dt} = -v \quad \text{and} \quad \frac{dP}{dt} = 2v. \quad (2.7)$$

The change of concentrations of each substance, expressed by the derivative over time t is on the left hand side and the net rate is on the right hand side. This net rate accounts for the sum of forward and backward reaction rates

$$v = v_+ - v_- = k_+ S_1 S_2 - k_- P^2 \quad (2.8)$$

k_+ and k_- are proportional factors, called *rate constants*. The power of each substance in the reaction depends on its molecularity. The following equation can be used to generate the net rate for reversible reactions in general:

$$v = k_+ \prod_i S_i^{m_i} - k_- \prod_j P_j^{m_j} \quad (2.9)$$

where S_i are substrate concentrations and P_j product concentrations and m_i and m_j denote the respective molecularities of S_i and P_j (53). Solving the system results into a time dependent trajectory for each substance.

Enzyme kinetics. Enzymes are the proteins that catalyze chemical reactions and they are involved in metabolism and signal transduction. A typical enzyme reaction can be written as



It reflects the conversion of an initial species, the substrate, into a resulting species, the product. The species dynamic read:

$$\frac{dS}{dt} = -k_1 E \cdot S + k_{-1} ES \quad (2.11)$$

$$\frac{dES}{dt} = k_1 E \cdot S - (k_{-1} + k_2) ES \quad (2.12)$$

$$\frac{dE}{dt} = -k_1 E \cdot S + (k_{-1} + k_2) ES \quad (2.13)$$

$$\frac{dP}{dt} = k_2 ES. \quad (2.14)$$

First, Michaelis and Menten assumed that the conversion of E and S to ES and vice versa

is much faster than the decomposition of ES into E and P , which leads to $k_1, k_{-1} \gg k_2$ (52). Second, Briggs and Haldane assumed that during the course of reaction a state is reached where the concentration of the ES complex remains constant. This assumption is justified only if the initial concentration of the substrate is much larger than the concentration of the enzyme, so that $dES/dt = 0$. With these simplifications, we can derive the reaction rate at steady state. Adding Eq. (2.12) and Eq. (2.13) lead to:

$$\frac{dES}{dt} + \frac{dE}{dt} = 0. \quad (2.15)$$

This also means that the total amount of enzyme stays constant, $E_{tot} = E + ES$.

Combining this conservation, Eq. (2.12) and steady-state assumption gives

$$ES = \frac{E_{tot}S}{S + \frac{k_{-1} + k_2}{k_1}} \quad (2.16)$$

and for the reaction rate

$$v = \frac{k_2 E_{tot} S}{S + \frac{k_{-1} + k_2}{k_1}}. \quad (2.17)$$

This reaction rate is typical for kinetics of many enzymes (Fig. 2.4). It is called Michaelis-Menten equation. Usually, it is written in the following form

$$v = V_{max} \frac{S}{S + K_m} \quad (2.18)$$

where $V_{max} = k_2 E_{tot}$ is the maximum reaction rate that is reached for very high substrate concentrations (saturation) and

$$K_m = \frac{k_{-1} + k_2}{k_1}. \quad (2.19)$$

is the Michaelis-Menten constant.

Ligand binding and Hill coefficient. Every molecule that binds to a protein is a ligand. The ligands bind to the subunits of the protein called binding sites, with the rule: one per subunit. Proteins are called monomers if they consist of only one subunit and oligomers if they consist of several subunits. The binding reactions of ligands under specific assumptions show different rate equations than simple mass action kinetics of Eq. (2.9) (52).

The binding reaction of a ligand S to a monomer E can be written as



with binding constant

$$K_B = \frac{ES}{E \cdot S}. \quad (2.21)$$

Since the enzymes are proteins, the assumption that the complex ES converts S to a product and releases the product afterwards can be used. Also, we assume that initial concentration of the ligand is much higher than initial concentration of the protein.

These assumptions lead to the conclusion that the reaction rate is proportional to ES (ES is converted to product E) and it holds that

$$\frac{v}{V_{max}} = \frac{ES}{E_{tot}} = Y, \quad (2.22)$$

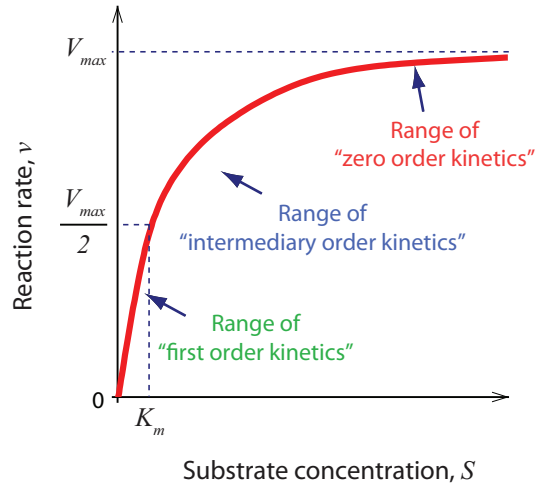


Figure 2.4: Michaelis-Menten kinetics. For very low substrate concentrations v follows first-order kinetics because of approximately linear change. For high S , reaction rate follows zero-order kinetics because there is almost no change of v in that region (52).

where v is the reaction rate and Y the fraction of protein-ligand complex. Using Eq. (2.21), we can also write that

$$Y = \frac{ES}{E_{tot}} = \frac{ES}{ES + E} = \frac{K_B \cdot S}{1 + K_B \cdot S}, \quad (2.23)$$

which leads to

$$v = V_{max} \frac{K_B \cdot S}{1 + K_B \cdot S}. \quad (2.24)$$

There are oligomers that show unexpected behavior in ligand binding. So-called allosteric proteins change their affinity to binding of further ligands depending on the number of ligands that are already bound to the protein. Hill discovered this behavior for hemoglobin. This protein, responsible for transport of oxygen in the blood, has got four binding sites. Depending on the oxygen partial pressure, hemoglobin binds either on all sites to oxygen if the pressure is high (like in the lungs) or on none of the sites if the pressure is low (like in the blood). States where there are still free binding sites do only rarely occur. This behavior can be represented in a sigmoid function, which is approximated by the Hill equation

$$v = V_{max} \frac{K_B \cdot S^{n_H}}{1 + K_B \cdot S^{n_H}}, \quad (2.25)$$

where n_H is the Hill coefficient which corresponds to the number of binding sites of the protein.

We distinguish between three cases:

1. If $n_H = 1$, affinity to further ligand binding is independent of ligands bound to the protein. The Hill equation is equivalent to Michaelis-Menten Eq. (2.18) and there is no cooperativity.
2. If $n_H > 1$, affinity to further ligand binding increases if more ligands are bound to the protein. v behaves like the sigmoid function with respect to S . The bigger n_H , the steeper is the curve and this is the case of positive cooperativity.

3. If $n_H < 1$, affinity to further ligand binding decreases if more ligands are bound to the protein. Since v behaves less sensitive with respect to S than Michaelis-Menten, this is the case of negative cooperativity.

Hemoglobin shows positive cooperativity with Hill coefficient of $n_H = 2.9$. The discrepancy to the number of binding sites of hemoglobin is based on the fact that the Hill equation is only valid for complete cooperativity, which means that every protein is either empty or fully bound to ligands.

Another way of defining the states of the integral components of the system could be using the discrete description, in a Boolean fashion as on or off, low or high, or with multiple discrete levels. This brings us to *discrete models*, which are mostly used to describe gene expression (54). One should be convinced that discrete representations of the biological phenomena is indeed relevant or otherwise lots of assumptions would be integrated into the model, which as such might not be able to answer its predefined questions.

2.2.4 Static vs. dynamic

Dynamic models are simplified representation of some real-world entity, in equations or computer code, and they describe how system properties change over time (55). On the contrary, static models represent isolated moment in time (snapshot) and therefore carry limited (but also useful) information about the system. The most common time point for system exploration is, for sure, the point where the system is in steady state.

Dynamic models in Systems biology are applied to areas such as cellular physiology, disease prevalence and extinction of endangered species (55).

In the context of cellular signaling, both static and dynamic models are able to reveal insights about system design and regulation. Depending of the properties of interest (steady-state or dynamic), one can use either of the two models or even both of them simultaneously for extended analysis.

Throughout this thesis, we have used steady-state representation of the model for part of the studies, but the main focus is placed on dynamic model and appropriate techniques tailored for its efficient exploration.

2.2.5 Modularity in cellular signaling

A functional unit that is integral part of more complex system and is devoted to perform a specific task is called a module. No matter what the overall nature of the whole system is, modular analysis can be successfully applied. Nevertheless, it is not always so straightforward to isolate these functionally independent subunits.

Coming from this systems theory concept, modularity is also widely utilized in signal transduction (56, 57). Typical examples for assignment of modules in cellular environment are the DNA-mRNA-enzyme-metabolism cascade and signal transduction cascades consisting of covalent modification cycles (52). In this thesis, we are following the idea of analyzing the simple module first and then proceeding to further analysis of more complex systems. In this context, the thesis consists of two parts: the study about covalent

modification cycle and the study about signaling cascades. The draw back of this kind of approach is that in every step of defining the modules at least some level of connectivity is neglected. Knowing that signaling networks possess a large number of already identified, as well as not identified, positive or negative feedbacks and feed-forward regulatory loops, one needs to be conscious about the significant information loss.

2.3 Parametric sensitivity analysis

Many studies have been based on parametric sensitivity analysis, which provides a powerful framework to relate the complex network structure to functions. There exist two types of parametric sensitivity: the direct sensitivity analysis, often referred simply as *sensitivity analysis* (58) and the inverse sensitivity analysis that relies on *numerical optimization* (59). The two approaches have been used for the analysis of both metabolic and signal transduction pathways. The direct sensitivity analysis – the consideration of changes in the system due to a variation in the model parameters – is widely applied (60-65). The inverse sensitivity analysis approach – the identification of the corresponding parameter values needed to achieve a desired functional behavior – is occasionally used for studies of metabolic networks to identify the relationship between model parameters and functions (23, 66-68). The later approach involves solving constrained optimization problems and it is well adapted for studying biochemical networks, as it makes it possible to deal with large-scale models (69-71). Moreover, the inverse approach leads to efficient parametric analysis and identification, contrary to an exhaustive parameter search.

2.3.1 Direct approach

Sensitivity analysis provides valuable insights about robustness of the observed model outputs with respect to the changes of model parameters. Even more, this tool can classify model parameters according to their influence on model outputs. It can also facilitate model development, model reduction, and it can guide parameter estimation and experimental design. As such, this powerful tool has a significant role in the arsenal of methods for System Biology (72).

Gutenkunst *et al.* (73) investigated the sensitivities of 17 published systems biology models, including few for signaling pathways, and studied the model output variations to the parameter changes. They found that systems biology models exhibit sloppy sensitivity spectra. This was one more proof of a common opinion that biological models seem to be quite robust. These findings again indicate the difficulty of uniquely determining the model parameters by fitting to a few experimental data.

Generally speaking, parametric sensitivity is usually performed as a series (of significantly high number) of trials for different parameter values, with observation on how the change in parameters causes a change in model outputs (74). Such procedure allows determining what level of accuracy is necessary for a parameter to make the model sufficiently useful and valid. If results reveal that the model is insensitive, then it may be possible to use an estimate rather than a value with greater precision for a concrete parameter.

There are two types of sensitivity analysis approaches: local and global sensitivity analysis. Local sensitivity analysis is a common approach, where the sensitivity of a model

output is performed by computing the first-order partial derivatives of the system output with respect to the parameters, which can be viewed as the gradients around the multidimensional reference parameter space (72). The second type of methods is global sensitivity analysis, which is used to quantify the overall effects of the parameter changes on the model output by perturbing the parameters within large ranges (58).

Since most of the models in systems biology are nonlinear, non-additive and non-monotonic, and in addition have numerous parameters, local sensitivity analysis might not be the most relevant approach for addressing the sensitivity questions. Accordingly, we performed global sensitivity analysis investigating the sensitivity over the entire parameter space.

There exist quite a number of different sensitivity analysis methods (72, 75) and some of them have been applied to the analysis of signaling networks. For example, using Monte Carlo method, Cho *et al.* performed multi-parametric global sensitivity analysis on the TNFa-mediated NF-kB pathway for experimental design (76). Using the same approach, Zi and coworkers studied IFN-g induced JAK-STAT signaling pathway (65). Chu and colleagues used four different techniques to identify key steps in the mathematical model of IL-6 signaling pathway (77). Coming from the same idea of using few methods and comparing their results, Zhang and Rundell addressed similar questions in TCR-activated Erk-MAPK signaling pathway (78).

2.3.1.1 The Sobol method

The Sobol method is a variance-based global sensitivity analysis approach that makes no assumptions on the relationship between the model inputs and outputs (79).

Let system be described as

$$Y = f(x) , \quad (2.26)$$

where $x = (x_1, x_2, \dots, x_p)$ is a vector of p input variables that influence the behavior of the system, Y is the model output and f the structure of the model. In this case, model inputs are equivalent to the parameters under study.

The main idea behind Sobol method is the decomposition of the variance of the model output function $f(x)$ into summands of variances in combinations of input parameters of increasing dimensionality:

$$f(x) = f_0 + \sum_{i=1}^p f_i(x_i) + \sum_{i=1}^p \sum_{j=i+1}^p f_{ij}(x_i, x_j) + \dots + f_{1\dots p}(x_1, x_2, \dots, x_p) . \quad (2.27)$$

Total variance D is defined as

$$D = \int_{\Omega^p} f^2(x) dx - f_0^2 = \int_{\Omega^p} f^2(x) dx - \left(\int_{\Omega^p} f(x) dx \right)^2 . \quad (2.28)$$

The partial variances are computed from each item in Eq. (2.27) and defined as

$$D_{i_1, i_2, \dots, i_s} = \int \dots \int f^2(x_{i_1}, x_{i_2}, \dots, x_{i_s}) dx_{i_1} dx_{i_2} \dots dx_{i_s} . \quad (2.29)$$

Furthermore, these variances, D and D_{i_1, i_2, \dots, i_s} can be approximated by Monte Carlo integrations.

The partial variance divided by the total variance is called first-order sensitivity index:

$$S_i = \frac{D_i}{D} \quad (2.30)$$

and it quantifies the contribution of parameter x_i to the output variance.

The total sensitivity indices are defined as

$$S_{i_1, i_2, \dots, i_S} = \frac{D_{i_1, i_2, \dots, i_S}}{D} . \quad (2.31)$$

In order to reduce computational effort, the following algorithm proposed by Saltelli *et al.* (58) offers an efficient way to approximate sensitivity indices, and as such can be easily implemented (Fig 2.5):

1. Generate two $(N \times p)$ matrices of random numbers, A and B , in the range that is valid for the variables (or parameters), where p is the number of input variables and N is the number of samples (experiments) per input variable. N should be carefully chosen, usually from hundreds to thousands, in order to achieve statistically relevant results.

$$A = \begin{bmatrix} x_1^{(1)} & x_2^{(1)} & \dots & x_i^{(1)} & \dots & x_p^{(1)} \\ x_1^{(2)} & x_2^{(2)} & \dots & x_i^{(2)} & \dots & x_p^{(2)} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ x_1^{(N-1)} & x_2^{(N-1)} & \dots & x_i^{(N-1)} & \dots & x_p^{(N-1)} \\ x_1^{(N)} & x_2^{(N)} & \dots & x_i^{(N)} & \dots & x_p^{(N)} \end{bmatrix}$$

$$B = \begin{bmatrix} x_{p+1}^{(1)} & x_{p+2}^{(1)} & \dots & x_{p+i}^{(1)} & \dots & x_{2p}^{(1)} \\ x_{p+1}^{(2)} & x_{p+2}^{(2)} & \dots & x_{p+i}^{(2)} & \dots & x_{2p}^{(2)} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ x_{p+1}^{(N-1)} & x_{p+2}^{(N-1)} & \dots & x_{p+i}^{(N-1)} & \dots & x_{2p}^{(N-1)} \\ x_{p+1}^{(N)} & x_{p+2}^{(N)} & \dots & x_{p+i}^{(N)} & \dots & x_{2p}^{(N)} \end{bmatrix}$$

2. Define matrix C_i that is identical to B , except for the i -th column, which is taken from A . i is the variable under study.

$$C_i = \begin{bmatrix} x_{k+1}^{(1)} & x_{k+2}^{(1)} & \dots & x_i^{(1)} & \dots & x_{2k}^{(1)} \\ x_{k+1}^{(2)} & x_{k+2}^{(2)} & \dots & x_i^{(2)} & \dots & x_{2k}^{(2)} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ x_{k+1}^{(N-1)} & x_{k+2}^{(N-1)} & \dots & x_i^{(N-1)} & \dots & x_{2k}^{(N-1)} \\ x_{k+1}^{(N)} & x_{k+2}^{(N)} & \dots & x_i^{(N)} & \dots & x_{2k}^{(N)} \end{bmatrix}$$

3. Compute model output for each row of input variables in A , B , C_i which gives three vectors of model outputs of dimension $(N \times 1)$:

$$y_A = f(A), \quad y_B = f(B), \quad y_{C_i} = f(C_i).$$

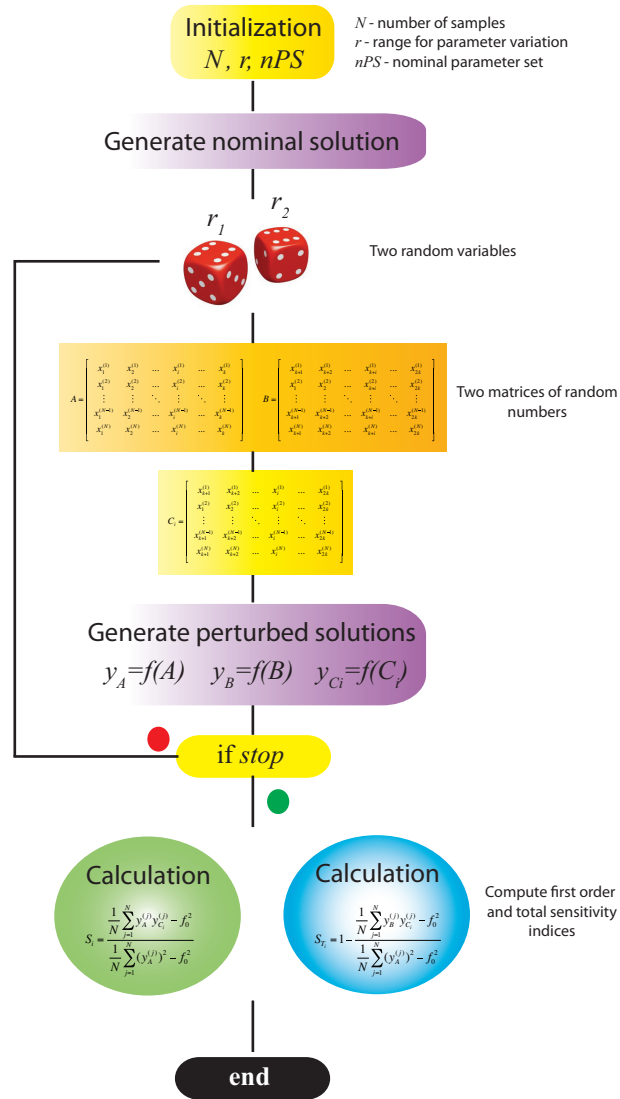


Figure 2.5: The Sobol method.

4. The first-order sensitivity indices can be estimated by

$$S_i = \frac{\frac{1}{N} \sum_{j=1}^N y_A^{(j)} y_{C_i}^{(j)} - f_0^2}{\frac{1}{N} \sum_{j=1}^N (y_A^{(j)})^2 - f_0^2} \quad \text{with} \quad f_0^2 = \left(\frac{1}{N} \sum_{j=1}^N y_A^{(j)} \right)^2.$$

5. Total-effect estimates can be computed by

$$S_{T_i} = 1 - \frac{\frac{1}{N} \sum_{j=1}^N y_B^{(j)} y_{C_i}^{(j)} - f_0^2}{\frac{1}{N} \sum_{j=1}^N (y_A^{(j)})^2 - f_0^2}$$

Saltelli *et al.* (58) suggested using the variance-based sensitivity analysis approaches when such application is possible. The main strengths of Sobol method lie around the facts that it

accounts for full sampling of parameter space, without requirement neither of any assumption about the model nor of parameter relations, allowing the interaction effects among parameters. It provides probability distribution of the output, given probability distributions of the input. The main drawback of all variance-based approaches is their high computational cost because they require more model evaluations than other types of methods. In case of complex models with significant number of parameters and state variables, this can become an issue, which will lead the choice of global sensitivity analysis method in some other direction.

Our systems were within the limits of not having extensively high computational cost, so we defined our sensitivity analysis framework around Sobol method.

2.3.2 Inverse approach

As introduced before, using optimization framework – the “*per aspera ad astra*” with all its challenges – is the core thread of the thesis. Furthermore, the motivation to explore trade-offs between different properties of cellular signals pointed to optimization as indispensable tool to relate system structure and functions.

In simple words, the setting of the optimization problem includes the minimization or maximization of an *objective function*:

$$\min_{x \in X} f(x) \quad (2.32)$$

where X is a subset of \mathbb{R}^{n_x} and $f: X \rightarrow \mathbb{R}$.

This can be any quantified measure of the performance of the system under study. The objective depends on the different input parameters, which are called *decision variables*:

$$x = (x_1, x_2, \dots, x_{n_x}). \quad (2.33)$$

If x_i , $i = 1, 2, \dots, n$ are independent of time, than the problem is called static optimization problem. Otherwise, these problems fall into the scope of dynamic optimization. In any case, the goal is to identify these variables in order to optimize the objective. Very often, the variables or even system performance are restricted, or *constrained* (i.e. positive values of molecular concentration) with inequality or equality constraints:

$$g_i(x) \leq 0, \quad i = 1, 2, \dots, n_g \quad (2.34)$$

$$h_i(x) = 0, \quad i = 1, 2, \dots, n_h, \quad (2.35)$$

where $g: X \rightarrow \mathbb{R}^{n_g}$ and $h: X \rightarrow \mathbb{R}^{n_h}$.

A vector $x \in X$ satisfying all the constraints is called a *feasible solution* of the problem and the collection of all such points forms the feasible region. Once the whole problem is set, the optimization algorithm aims at finding such a feasible point x^* so that:

$$f(x) \geq f(x^*). \quad (2.36)$$

x^* is then called an *optimal solution*. Depending on whether the Eq. (2.36) is valid only for the feasible points x that are in the neighborhood of x^* , or for all the points x in the feasible set, we can distinguish between *local* and *global optimum*, respectively.

Nowadays, strong and efficient mathematical programming techniques are available for solving a great variety of optimization problems, which are based on solid theoretical results and extensive numerical studies (80). Some of them are summarized in Fig. 2.6. The term “programming” does not have any correlation with

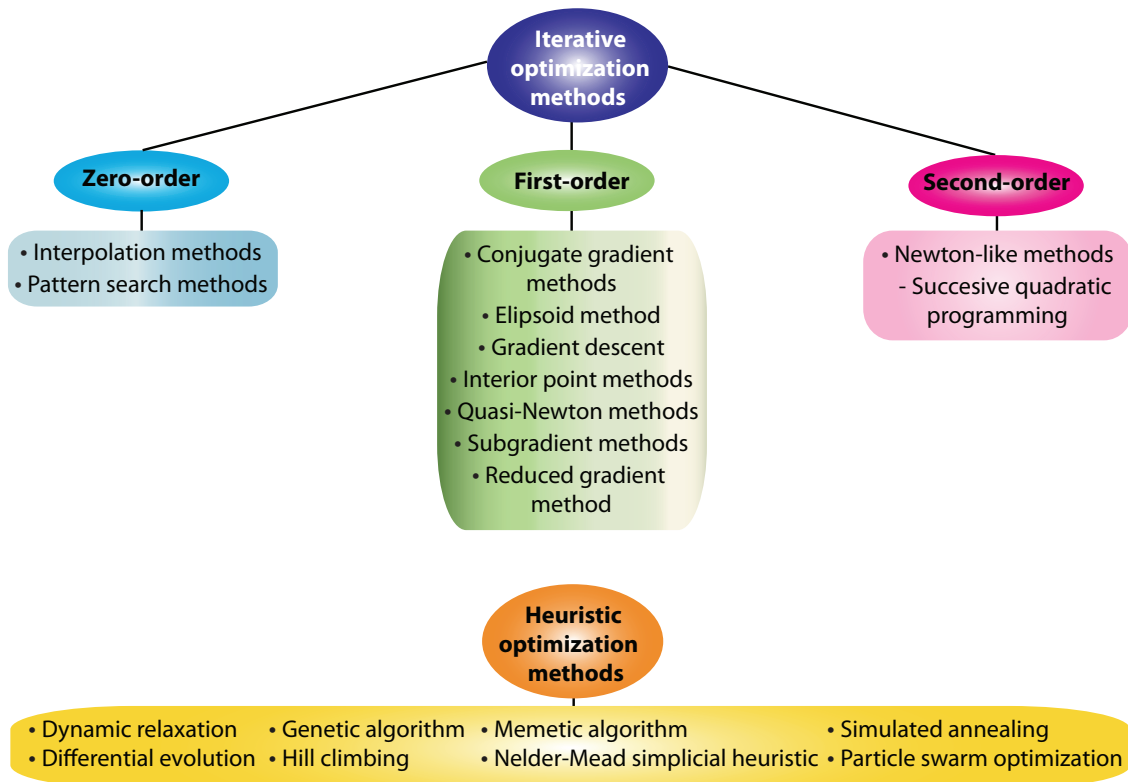


Figure 2.6: Numerical methods for nonlinear programming problems.

computer programming, but is rather associated with the history of optimization as mathematical discipline.

There are different ways to classify main subfields of optimization:

- Depending on whether the variables of the system can take any value or their values can be chosen only from the predefined set, we distinguish between continuous and discrete optimization, respectively. Widely used approach is integer programming, where the variables are constrained to have integer values. A hybrid case, where only some of the variables need to be integers, is called mixed-integer programming (59).
- The next classification criteria take into account the nature of constraints and objective function. Thus, there exist unconstrained and constrained optimization problems.
- Furthermore, there is a distinction between linear programming, where one studies the case in which the objective function is linear and the set of constraints is specified using only linear equalities and inequalities, and nonlinear programming, which is the general case in which the objective function or the constraints or both contain nonlinear parts. Quadratic programming allows the objective function to have quadratic terms, while the feasible set must be specified with linear equalities and inequalities.
- A subclass of optimization methods called convex optimization, which studies only convex functions over convex sets, has a very complete theory and arises in a variety in applications (81). The fact that any local optimum in a convex set is at the same time global optimum makes this class of methods very efficient and not so expensive

computationally. Unfortunately, models of signal transduction in general do not belong to this class.

- Depending on the way the model is defined, the optimization can be deterministic or stochastic, which is the case in which some of the constraints or parameters depend on random variables. Robust programming attempt to capture uncertainty in the data underlying the optimization problem, and, unlike in the case of stochastic programming, the problem is solved taking into account inaccuracies in the input data rather than dealing with random variables.
- Heuristics and metaheuristics methods make few or no assumptions about the problem being optimized. Usually, heuristic approach does not guarantee that any optimal solution need be found. On the other hand, heuristics are used to find approximate solutions for many complicated optimization problems.
- A very important classification of optimization methods is local versus global. The fastest algorithms seek only a local solution, but these methods are not always able to find the best point of the whole search space, the global solution. Even though global solutions are difficult to be identified, they are highly desirable.

Numerous state-of-the-art algorithms for numerical optimization of nonlinear problems are used in Systems biology. Among them, quite a number of methods are gathered in the class of *iterative methods*. Their main characteristic is that after a given initial point x^0 , a sequence of points $\{x^k\}$ is obtained by repeated application of an algorithmic rule, with the objective to make this sequence converge to the point \bar{x} in a pre-specified solution set. The algorithm is said to be *globally convergent* if the previous statement holds for any initial point x^0 . Global convergence is a very desirable property of an optimization algorithm.

Depending whether a method requires only the evaluation of the function values, gradients or Hessian, we can distinguish between zero-, first- and second-order optimization methods. The advantage of calculating the first and second derivatives improves the rate of convergence, but increase the computational cost, that can sometimes be the unsolved issue.

Such a vast number of methods do not make the initial step of choosing the appropriate algorithm so straightforward. The wrong choice could lead to increase of computational cost and time, or even to not being able to find a solution at all. Since one of the main goals of this work is to gain the knowledge of inherent dynamics of signaling modules, the dynamic optimization methods (versus static optimization methods) would suit better. Dynamic programming, together with calculus of variations and optimal control, is one of the main subfields of the optimization approaches that are designed primarily for optimization in dynamic contexts, where the decision making process is occurring over the time (82).

The choice of the appropriate method in our studies was already established through some previous collaboration. Similar approach was proven to be adequate in other studies of signaling networks (69) and that was assuring enough for us not to focus on comparing of different methods and developing new ones. More details are given in the section Materials and methods of chapter 4.

Ensembles in Signaling Symphony

Covalent Modification Cycle and Prototypical MAP Kinase Cascade

3.1 Covalent modification cycle – monocyclic system

The control of cellular processes is a consequence of the evolution of an extremely complex system of regulatory mechanisms. One of the main aims of Systems biology is to elucidate the interplay of specific components and their characterization, which will in turn improve general understanding of cell regulation. There are many processes of cell regulation whereby a protein is reversibly and covalently modified by the enzyme-catalyzed transfer of a group from a donor to a specific amino acid residue usually located at the active site of the acceptor (83-85). Specific converter enzymes, often protein kinases and phosphoprotein phosphatases, catalyze these covalent modification and demodification reactions. The modification of the protein by a converter enzyme and the opposite reaction, in which the modified protein is “demodified” by another converter enzyme, forms a *monocyclic enzyme cascade system*. The converter enzymes for both the modification and demodification reactions also undergo a modification process (activation or inactivation) induced by an allosteric effector. Both allosteric effectors may be the same substance.

Monocyclic enzyme cascades are ubiquitous in biological systems. They play an important role in the regulation of many physiological processes involving the cyclic interconversion of enzymes between phosphorylated and unphosphorylated forms, e.g., the repair of lesions, differentiation, growth, evolution metabolism and motility (86, 87) the regulation of neurotransmitter receptor function and the efficiency of synaptic transmission (88). Almost all proteomic signaling networks in prokaryotes and eukaryotes are based on phosphorylation–dephosphorylation monocyclic cascades (89). Phosphorylation of the enzyme is catalyzed by a specific kinase, whereas dephosphorylation is catalyzed by a specific phosphatase, the activities of the converter enzymes being controlled by numerous allosteric effectors (84).

The special significance of enzyme cascades is their ability to impose upper and lower boundaries on the rates of a biological process and to act as physiological switches (90). Furthermore, the abundance of design features in enzyme cascades provides many possibilities for response and adaptation to environmental cues and challenges. Such cascades are therefore essential for the success of evolutionary systems. Accordingly, our study will place equal attention to monocyclic cascade systems, as it will be done for the more complex systems later.

3.1.1 Mechanistic modeling framework

The prototypical monocyclic system under consideration consists of three elements: i) a kinase, ii) an activating enzyme, which can be a receptor stimulated by a ligand or an activated kinase, and iii) a deactivating enzyme, usually a phosphatase. The active receptor initiates the internal signaling cascade, including a series of protein phosphorylation state changes, which, as mentioned, represents the basic mechanism in signal transduction networks.

The complicated structure of many enzyme cascades renders their kinetic analysis difficult (91). However, such knowledge is a prerequisite for understanding biological regulation at a high level. In order to facilitate the kinetic analysis of monocyclic enzyme cascades a considerable number of simplifying assumptions must be made. Obviously, the more assumptions are made, the simpler the results obtained. The gain in simplicity means a loss of accuracy in the results.

The development of our model relies on the following assumptions:

- the activation step involves the reversible binding of the activating enzyme to inactive kinase and the complex is irreversibly released;
- the inactivation step involves the reversible binding of the deactivating enzyme to active kinase and the complex is irreversibly released;
- the total amount of kinase is taken to be constant for the signaling time scale considered;
- the rates of the various processes follow mass-action kinetics;
- the dynamics can be described by ordinary differential equations.

A commonly used simplification considers that the concentrations of intermediate enzyme-substrate complexes are small compared to the substrate concentrations and can therefore be neglected (44, 92, 93), leading to simpler mathematical models. This simplification has been used mainly to reduce the computational effort or to obtain an exact analytic solution to the problem. However, it has been shown and we found in our studies that if we neglect the so-called substrate sequestration in the form of intermediate enzyme-substrate complexes, the analysis leads to incorrect model predictions in some cases (94-96). Besides its effect on model prediction accuracy, substrate sequestration has been shown to induce both positive and negative feedback mechanisms in signaling cascades (97, 98). One of the advantages of our computational approach is that it does not require such simplifications and, accordingly, substrate sequestration will be considered in this work.

3.1.2 Mathematical model of monocyclic system

The covalent modification cycle represented in Fig. 3.1 comprises 7 species S , S^* , X , X^* , P_X , $X:S^*$ and $X^*:P_X$. The receptor changes its own state from susceptible S to active S^* . Since we will focus our interest on the dynamics of the internal module, this event of tuning on/off the receptor is assumed to occur under very fast kinetics. The kinetic mechanism and the initial conditions for the system will be defined later. The variables X and X^* represent the two interconvertible forms of one protein, e.g. the phosphorylated and dephosphorylated forms of a kinase. Proteins S^* and P_X catalyze the activation and deactivation reactions, respectively. In the following, they will be called enzymes. The activation steps proceed through the formation of intermediate enzyme-substrate

complexes $X:S^*$ and $X^*:P_X$. The enzyme S^* can be either the activating kinase of an upstream cycle or the receptor responding to an external input signal (e.g. growth factor or hormone level), while the enzyme P_X corresponds to a phosphatase.

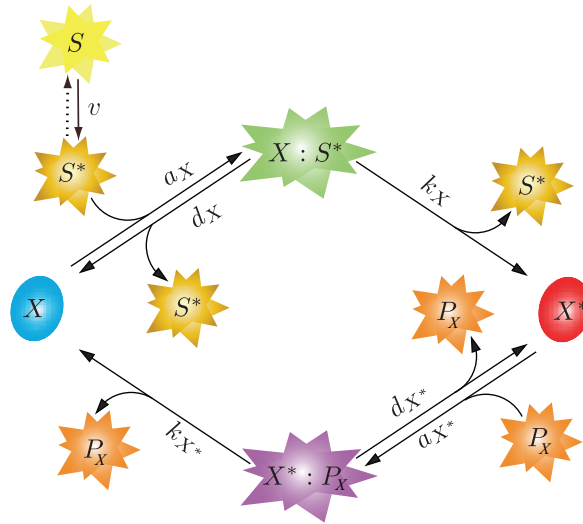


Figure 3.1: Schematic representation of a covalent modification cycle. A cycle is composed of two states of the same protein, namely the inactive state X and the active state X^* . The activation and deactivation reactions are catalyzed by the enzymes S^* and P_X , respectively.

The system consists of 2 reversible and 3 irreversible reactions, which gives 7 reactions in total (Table 3.1).

The total amount of interconvertible protein X_T , as well as the total amounts of the activation/deactivation proteins S_T and P_{XT} , are considered being constant in the time scale considered. The corresponding conservation equations relating the concentrations of the seven species read

$$[X]_T = [X] + [X^*] + [X:S^*] + [X^*:P_X], \text{ with } \frac{d[X]_T}{dt} = 0 \quad (3.1)$$

$$[S]_T = [S] + [S^*] + [X:S^*], \text{ with } \frac{d[S]_T}{dt} = 0 \quad (3.2)$$

$$[P_X]_T = [P_X] + [X^*:P_X], \text{ with } \frac{d[P_X]_T}{dt} = 0. \quad (3.3)$$

The conservation laws Eq. (3.1–3.3) relate the concentrations of the species in the system.

No.	Reaction	Reaction Rate	Notes
1	$S \xrightarrow{v} S^*$	$v \cdot S$	Activation of signaling enzyme
(1)	$S^* \xrightarrow{v} S$	$v \cdot S^*$	Deactivation of signaling enzyme
2	$X + S^* \xrightarrow{a_X} X : S^*$	$a_X \cdot X \cdot S^*$	Formation of complex
3	$X : S^* \xrightarrow{d_X} X + S^*$	$d_X \cdot X : S^*$	Degradation of complex
4	$X : S^* \xrightarrow{k_X} X^* + S^*$	$k_X \cdot X : S^*$	Degradation of complex
5	$X^* + P_X \xrightarrow{a_{X^*}} X^* : P_X$	$a_{X^*} \cdot X^* \cdot P_X$	Formation of complex
6	$X^* : P_X \xrightarrow{d_{X^*}} X^* + P_X$	$d_{X^*} \cdot X^* : P_X$	Degradation of complex
7	$X^* : P_X \xrightarrow{k_{X^*}} X + P_X$	$k_{X^*} \cdot X^* : P_X$	Degradation of complex

Table 3.1: Reactions in monocyclic system.

Four additional (independent) relations can be obtained by writing mass-balance equations for X^* , $X : S^*$ and $X^* : P_X$ as

$$\frac{d[X^*]}{dt} = -a_{X^*}[X^*][P_X] + k_X[X : S^*] + d_{X^*}[X^* : P_X] \quad (3.4)$$

$$\frac{d[X : S^*]}{dt} = a_X[X][S^*] - (d_X + k_X)[X : S^*] \quad (3.5)$$

$$\frac{d[X^* : P_X]}{dt} = a_{X^*}[X^*][P_X] - (d_{X^*} + k_{X^*})[X^* : P_X] \quad (3.6)$$

$$\frac{d[S]}{dt} = -v[S] \quad (1) \quad (3.7)$$

where $a_X, d_X, k_X, a_{X^*}, d_{X^*}, k_{X^*}$ and v are the parameters of the mass-action kinetic laws, as indicated in Table 3.1.

The basic structure of signaling modules of interest is well conserved in cells, even though it can generate a high variety of biological responses. In order to facilitate the discovery of more general features of the system, it was proposed that model assumes dimensionless parameters, rather than to be related to a particular parameter set (44). For that purpose, we introduce the set of dimensionless variables and parameters used in all of our system analysis (later in Table 3.4). In addition, the time is scaled as $\tau := \tilde{k}_X t$.

¹ Depending on which property is the focus of our analysis, the reaction describing the steps between active and inactive state of signaling enzyme and corresponding differential equation will be defined accordingly. More details will follow by the end of this chapter.

The dynamic model Eq. (3.1–3.7) becomes the following differential-algebraic equation (DAE) system:

$$\frac{dx^*}{d\tau} = \rho^{S/X} \tilde{k}_X \{x : s^*\} + \rho^{P_X/X} (\tilde{d}_{X^*} \{x^* : p_X\} - \tilde{a}_{X^*} x^* p_X) \quad (3.8)$$

$$\frac{d\{x : s^*\}}{d\tau} = \tilde{a}_X x s^* - (\tilde{d}_X + \tilde{k}_X) \{x : s^*\} \quad (3.9)$$

$$\frac{d\{x^* : p_X\}}{d\tau} = \tilde{a}_{X^*} x^* p_X - (\tilde{d}_{X^*} + 1) \{x^* : p_X\} \quad (3.10)$$

$$\frac{ds}{d\tau} = -\nu_S \quad (1) \quad (3.11)$$

$$1 = x + x^* + \rho^{S/X} \{x : s^*\} + \rho^{P_X/X} \{x^* : p_X\} \quad (3.12)$$

$$1 = s + s^* + \{x : s^*\} \quad (3.13)$$

$$1 = p_X + \{x^* : p_X\}. \quad (3.14)$$

The compact notation of the system is defined as:

$$\mathbf{F}(\xi(\tau), \dot{\xi}(\tau), \mathbf{p}, \mathbf{r}) = \mathbf{0}, \quad (3.15)$$

where ξ is the vector of state variables: $\xi := (x^* \ s \ \{x : s^*\} \ \{x^* : p_X\})^T$, \mathbf{p} is the vector of kinetic parameters: $\mathbf{p} := (\tilde{a}_X \ \tilde{a}_{X^*} \ \tilde{d}_X \ \tilde{d}_{X^*} \ \tilde{k}_X)^T$, and \mathbf{r} is the vector of concentration ratios: $\mathbf{r} := (\rho^{S/X} \ \rho^{P_X/X})^T$.

If we assume that the entire signaling enzyme is already in its active form at steady state (i.e. $\bar{s} = 0$), then the model can be shown to reduce to the set of algebraic equations:

$$0 = \overline{\{x : s^*\}} - \frac{\bar{x}}{\tilde{K}_X + \bar{x}} \quad (3.16)$$

$$0 = \overline{\{x^* : p_X\}} - \frac{x^*}{\tilde{K}_{X^*} + x^*} \quad (3.17)$$

$$0 = \overline{\{x^* : p_X\}} - \alpha_X \overline{\{x : s^*\}} \quad (3.18)$$

$$1 = \bar{x} + \bar{x}^* + \rho^{S/X} \overline{\{x : s^*\}} + \rho^{P_X/X} \overline{\{x^* : p_X\}} \quad (3.19)$$

$$1 = \bar{s}^* + \overline{\{x : s^*\}} \quad (3.20)$$

$$1 = \bar{p}_X + \overline{\{x^* : p_X\}}, \quad (3.21)$$

where $(\bar{\cdot})$ indicates a steady-state value and

$$\tilde{K}_X = \frac{\tilde{d}_X + \tilde{k}_X}{\tilde{a}_X} \quad (3.22)$$

$$\tilde{K}_{X^*} = \frac{\tilde{d}_{X^*} + 1}{\tilde{a}_{X^*}} \quad (3.23)$$

$$\alpha_X = \tilde{k}_X \frac{\rho^{S/X}}{\rho^{P_X/X}}. \quad (3.24)$$

3.2 Prototypical MAP kinase cascade – multicyclic cascade system

Monocyclic patterns form the backbone of most signaling systems and they are often linked, forming multiple layers of cycles, the so-called cascades (referred as *bi-*, *tri-*, *multicyclic cascade systems*). Commonly observed instance of signal transduction through a series of protein kinase reactions are the kinases of the mitogen-activated protein kinase (MAPK) cascades (99).

MAP kinase is the collective term for the serine/threonine (Ser/Thr)-specific protein kinases generating the output signals of the MAP kinase modules. They are by no means activated only by mitogenic signals, but also are involved in cell differentiation and cell death (100).

MAP kinase modules are complexes of three protein kinases that are interconnected in series, known as MAP kinase, MAPKK kinase, and MAPKKK kinase (9). Each of these kinases exists in various isoforms, which are integrated in individual modules. The core of each module is the pair of MAP kinase and MAPKK kinase, and they interact with each other in a highly specific manner. In contrast, the interaction between MAPKK kinases and MAPKKK kinases is much less specific: many protein kinases may function as MAPKKK kinases, thus transmitting a wide variety of input signals to the module. The same holds for MAP kinases, which are firmly wired with the MAPKK kinases, but are not choosy as far as their substrate proteins are concerned. Thus, MAP kinase modules transform a wide variety of input signals into a comparatively large variety of output signals. As such, they play a key role in cellular data processing ranging from yeast to humans (101). While the input signals are mainly derived from signal receptors, the output signals address metabolic reactions, the architecture and mobility of cells, and gene transcription (102), which may account for its evolutionary success throughout the tree of life.

All MAPK pathways share a common structure, namely, the sequential activation of three protein kinases in a kinase cascade (19). At the very top of the cascade, the first protein kinase is activated by a molecular signal received from an upstream activator, while at the very bottom of the cascade, the output protein kinase serves as a signal to transcription factors or to other protein kinases. Getting the cascade started may be the result either of a signal relayed to the cascade by a member of the family of the Ras proteins or of a response to some form of physical or chemical stress. When the last kinase is activated, it binds to a substrate that will either go on to activate a transcription factor within the nucleus or activate another protein kinase. Coordinating and organizing this process is the protein scaffold, a multiprotein complex whose importance will be highlighted later.

MAPKKKs: The starting points of the protein cascades are the MAPKKKs (also called MEKKs). These kinases are usually activated by small GTPases, like many of the members of the Ras superfamily of proteins. Without getting into many details, the activation of MAPKKKs is a complex procedure that requires membrane translocation, phosphorylation, oligomerization, and binding to a scaffold protein - the actual details of which are not well known.

Individual MAPK modules are associated with distinct MAPKKK type:

- RAF kinases of ERK1,2 modules and the related mixed-linear kinases (MLKs) of JNK and p38 modules
- MEK kinases (MEKKs), mainly JNK and p38 modules
- A heterogeneous collection of other kinases named ASK1, TAK1, TPL2, NIK and

TAO.

MAPKKs: The MEKs that phosphorylated MAPK are known as dual-specificity protein kinases because they can accept two different kinds of amino acids at the phosphorylation site of the MAPK/ERK they bind. As the intermediate step of a tricyclic cascade, the MEKs are phosphorylated by the MEKKs or MAPKKKs at two serine residues that make up part of a conserved sequence in all known MEK proteins - perhaps an indication of the long evolutionary history of this pathway.

MAPKs: The activation of MAPK/ERK proteins by MEK occurs via a phosphorylation reaction at the tyrosine and threonine residues of the activation loop. In addition to relieving a steric hindrance to substrate binding, the phosphorylation makes it possible for homodimers to form. Without these homodimers, MAPK would be unable to penetrate inside the nuclear envelope where gene expression takes place.

The mammalian (human) genome encodes 14 MAPK's and 8 MAPKK's. The MAP kinases have been arranged in several subfamilies, each of which is named according to the protein kinase that emerges at the end of the cascade:

- ERK1,2 family (Extracellular signal-Regulated Kinases), including isoforms ERK1 and ERK2
- JNK family (cJun N-terminal Kinases, also called SAPK, Stress-Activated Protein Kinases), including isoforms JNK1-3
- p38 family (molecular mass of 38kDa), including isoforms p38a, -b, -g, -d
- ERK3,4 family and ERK5 family (known as big MAP kinases).

Each family, even each isoform, organizes the module of its own.

In mammals, the ERK1,2, JNK, and p38 kinases pathways happen to be the most important ones (103).

Phosphatases: There are many enzymatic proteins that intervene in the cascade to regulate the amount of signal output. Of these, the most researched are the phosphatases, which preferentially dephosphorylate MAPK, MAPKK, or MAPKKK depending on where they bind in the protein scaffold. It is interesting to note that the MAPK phosphatases also constitute a regulatory network that may either amplify or degrade the output signal via activation or negative feedback, respectively.

Given the limited number of MAPK constellations and impressively large number of functions they are involved, there arises the question of how individual module can distinguish its functions. A scaffold protein that interacts with each of the protein kinases may mediate the organization of the module. As such, scaffolds play an important role in producing specificity in these signaling mechanisms (104-106).

The attributes of the protein scaffold, as pertain to MAPK signaling, include:

- contribution to selectivity by assembling distinct kinases into distinct modules;
- allosteric activation these distinct kinases, which enhances signaling efficiency;
- reduction the cross-talk between various cycles;
- recruitment of proteins phosphatases for feedback mechanisms;
- contact between the protein scaffold and the signal effectors maintains the flow through the cascade.

3.2.1 Mechanistic modeling framework

Accumulating evidence of role of MAPK modules in cancer development and growth (107) was yet one more reason why these modules are focus of extensive research in recent years (108). Systems biology also tries to decode their complexity through mathematical modeling and prediction in order to gain a deeper insight into the inner works of signaling networks (109). There are quite a number of computational models that, in some way, incorporate MAPK module (110). The three most cited models that focus on investigation of the properties and behavior of the core cascade itself are: Huang and Ferrell's model from 1996 (111), Kholodenko's model from 2000 (93) and Heinrich *et al.* model from 2002 (112). The choice of prototypical MAPK module architecture in this study lies somewhere in between modeling frameworks of these three models. Since the activation of the cellular response by MAPK pathways typically involves three phosphorylation steps (113), we will limit our analysis to the systems that have up to three joined cycles. Furthermore, we assumed the simple module concatenation with single phosphorylation events (unlike in (111)), but with assumed intermediate complex formation (unlike in (112)), and without any feedback loops (unlike in (93)).

3.2.2 Mathematical model of multicyclic system

Tricyclic cascade (Fig. 3.2) repeats the fundamental network motif of one-step enzymatic modification and reverse modification reaction. Similar to the single interconvertible cascade, the input signal S activates the cascade and each kinase then activates the next level kinase. At each level, a separate phosphatase deactivates the active kinase.

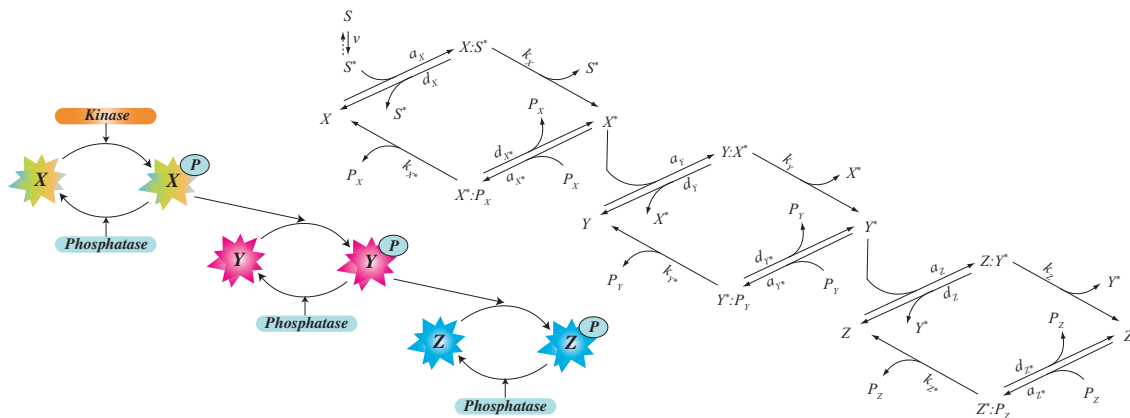


Figure 3.2: Schematic representation of tricyclic cascade. X , X^* , $X:S^*$ and $X^*:P_X$ denote the species/complexes relative to the first level of cascade, with S^* and P_X being the upstream kinase and phosphatase, respectively; Y , Y^* , $Y:X^*$ and $Y^*:P_Y$ denote the species/complexes relative to the second level of cascade, with X^* and P_Y being the upstream kinase and phosphatase, respectively; and Z , Z^* , $Z:Y^*$ and $Z^*:P_Z$ denote the species/complexes relative to the third level of cascade, with Y^* and P_Z being the upstream kinase and phosphatase, respectively.

The system reactions are given in Table 3.2.

Model	No.	Reaction	Reaction Rate	Notes		
Tricyclic system	Bicyclic system	Monocyclic system	1	$S \xrightarrow{v} S^*$	$v \cdot S$	Activation of signaling enzyme
			(1)	$S^* \xrightarrow{v} S$	$v \cdot S^*$	Deactivation of signaling enzyme
			2	$X + S^* \xrightarrow{a_X} X : S^*$	$a_X \cdot X \cdot S^*$	Formation of complex
			3	$X : S^* \xrightarrow{d_X} X + S^*$	$d_X \cdot X : S^*$	Degradation of complex
			4	$X : S^* \xrightarrow{k_X} X^* + S^*$	$k_X \cdot X : S^*$	Degradation of complex
			5	$X^* + P_X \xrightarrow{a_{X^*}} X^* : P_X$	$a_{X^*} \cdot X^* \cdot P_X$	Formation of complex
			6	$X^* : P_X \xrightarrow{d_{X^*}} X^* + P_X$	$d_{X^*} \cdot X^* : P_X$	Degradation of complex
		7	$X^* : P_X \xrightarrow{k_{X^*}} X + P_X$	$k_{X^*} \cdot X^* : P_X$	Degradation of complex	
		8	$Y + X^* \xrightarrow{a_Y} Y : X^*$	$a_Y \cdot Y \cdot X^*$	Formation of complex	
		9	$Y : X^* \xrightarrow{d_Y} Y + X^*$	$d_Y \cdot Y : X^*$	Degradation of complex	
		10	$Y : X^* \xrightarrow{k_Y} Y^* + X^*$	$k_Y \cdot Y : X^*$	Degradation of complex	
		11	$Y^* + P_Y \xrightarrow{a_{Y^*}} Y^* : P_Y$	$a_{Y^*} \cdot Y^* \cdot P_Y$	Formation of complex	
	12	$Y^* : P_Y \xrightarrow{d_{Y^*}} Y^* + P_Y$	$d_{Y^*} \cdot Y^* : P_Y$	Degradation of complex		
	13	$Y^* : P_Y \xrightarrow{k_{Y^*}} Y + P_Y$	$k_{Y^*} \cdot Y^* : P_Y$	Degradation of complex		
	14	$Z + Y^* \xrightarrow{a_Z} Z : Y^*$	$a_Z \cdot Z \cdot Y^*$	Formation of complex		
	15	$Z : Y^* \xrightarrow{d_Z} Z + Y^*$	$d_Z \cdot Z : Y^*$	Degradation of complex		
	16	$Z : Y^* \xrightarrow{k_Z} Z^* + Y^*$	$k_Z \cdot Z : Y^*$	Degradation of complex		
	17	$Z^* + P_Z \xrightarrow{a_{Z^*}} Z^* : P_Z$	$a_{Z^*} \cdot Z^* \cdot P_Z$	Formation of complex		
	18	$Z^* : P_Z \xrightarrow{d_{Z^*}} Z^* + P_Z$	$d_{Z^*} \cdot Z^* : P_Z$	Degradation of complex		
19	$Z^* : P_Z \xrightarrow{k_{Z^*}} Z + P_Z$	$k_{Z^*} \cdot Z^* : P_Z$	Degradation of complex			

Table 3.2: Reactions in monocyclic system and multicyclic cascade systems.

The time dependent behavior described by a set of differential equations and additional conservation relations derived from stoichiometry, build a mechanistic model of the prototypical MAPK cascade (Table 3.3). As mentioned before, Table 3.4 defines dimensionless variables and parameters. Following the same steps as in the case of monocyclic system, dimensional model is further transformed into dimensionless model (Table 3.5).

$$\begin{aligned}
d[X^*]/dt &= -a_{X^*}[X^*][P_X] + d_{X^*}[X^* : P_X] + k_X[X : S^*] - \\
&\quad -a_Y[Y][X^*] + (d_Y + k_Y)[Y : X^*] \\
d[X : S^*]/dt &= a_X[X][S^*] - (d_X + k_X)[X : S^*] \\
d[X^* : P_X]/dt &= a_{X^*}[X^*][P_X] - (d_{X^*} + k_{X^*})[X^* : P_X] \\
d[Y^*]/dt &= -a_{Y^*}[Y^*][P_Y] + d_{Y^*}[Y^* : P_Y] + k_Y[Y : X^*] - \\
&\quad -a_Z[Z][Y^*] + (d_Z + k_Z)[Z : Y^*] \\
d[Y : X^*]/dt &= a_Y[Y][X^*] - (d_Y + k_Y)[Y : X^*] \\
d[Y^* : P_Y]/dt &= a_{Y^*}[Y^*][P_Y] - (d_{Y^*} + k_{Y^*})[Y^* : P_Y] \\
d[Z^*]/dt &= -a_{Z^*}[Z^*][P_Z] + d_{Z^*}[Z^* : P_Z] + k_Z[Z : Y^*] \\
d[Z : Y^*]/dt &= a_Z[Z][Y^*] - (d_Z + k_Z)[Z : Y^*] \\
d[Z^* : P_Z]/dt &= a_{Z^*}[Z^*][P_Z] - (d_{Z^*} + k_{Z^*})[Z^* : P_Z]
\end{aligned}$$

Conservation relations:

$$\begin{aligned}
[X]_T &= [X] + [X^*] + [X : S^*] + [X^* : P_X] + [Y : X^*] \\
[S]_T &= [S] + [S^*] + [X : S^*] \\
[P_X]_T &= [P_X] + [X^* : P_X] \\
[Y]_T &= [Y] + [Y^*] + [Y : X^*] + [Y^* : P_Y] + [Z : Y^*] \\
[P_Y]_T &= [P_Y] + [Y^* : P_Y] \\
[Z]_T &= [Z] + [Z^*] + [Z : Y^*] + [Z^* : P_Z] \\
[P_Z]_T &= [P_Z] + [Z^* : P_Z]
\end{aligned}$$

Table 3.3: Kinetic equations comprising the mechanistic models: monocyclic system (green), addition for bicyclic system (magenta), addition for tricyclic system (blue).

3.3 Design parameters

Steady-state analysis identifies dimensionless Michaelis-Menten constants, which have been shown to determine the most important features of steady-state responses. They are in turn nonlinear combinations of the dimensionless kinetic rates and as such they constraint the allowable ranges of parameters for desired dynamic performance. This relationship is the important link between dynamic and the steady-state characteristics.

Furthermore, the parameter α_X represents the *activation potential* of the first cycle and describes the balance between activation and deactivation of the first kinase. It has been shown that in special case, where all the corresponding parameters of the upper and lower branch in the covalent modification cycle are the same and $\alpha_X = 1$, the system is at or very close to its inflection point (44).

Concentration ratios emerge from transformation of the system into dimensionless form, and thus will be analyzed as parameters of the system.

Model	Concentrations	Kinetic constants	Concentration ratios	Michaelis-Menten constants	Activation potentials		
Tricyclic system	Bicyclic system	Monocyclic system	$x := \frac{[X]}{[X]_T}$	$\tilde{a}_x := \frac{a_x}{k_{x^*}} [X]_T$	$\rho^{S/X} := \frac{[S]_T}{[X]_T}$	$\tilde{K}_x = \frac{\tilde{d}_x + \tilde{k}_x}{\tilde{a}_x}$	$\alpha_x = \tilde{k}_x \frac{\rho^{S/X}}{\rho^{P_x/X}}$
			$x^* := \frac{[X^*]}{[X]_T}$	$\tilde{a}_{x^*} := \frac{a_{x^*}}{k_{x^*}} [X]_T$	$\rho^{P_x/X} := \frac{[P_x]_T}{[X]_T}$	$\tilde{K}_{x^*} = \frac{\tilde{d}_{x^*} + 1}{\tilde{a}_{x^*}}$	
			$s := \frac{[S]}{[S]_T}$	$\tilde{d}_x := \frac{d_x}{k_{x^*}}$			
			$s^* := \frac{[S^*]}{[S]_T}$	$\tilde{d}_{x^*} := \frac{d_{x^*}}{k_{x^*}}$			
			$x : s^* := \frac{[X : S^*]}{[S]_T}$	$\tilde{k}_x := \frac{k_x}{k_{x^*}}$			
			$p_x := \frac{[P_x]}{[P_x]_T}$				
	$x^* : p_x := \frac{[X^* : P_x]}{[P_x]_T}$						
	Bicyclic system	$y := \frac{[Y]}{[Y]_T}$	$\tilde{a}_y := \frac{a_y}{k_{y^*}} [Y]_T$	$\rho^{X/Y} := \frac{[X]_T}{[Y]_T}$	$\tilde{K}_y = \frac{\tilde{d}_y + \tilde{k}_y}{\tilde{a}_y}$	$\alpha_y = \frac{\tilde{k}_y}{\tilde{k}_{y^*}} \frac{\rho^{X/Y}}{\rho^{P_y/Y}}$	
		$y^* := \frac{[Y^*]}{[Y]_T}$	$\tilde{a}_{y^*} := \frac{a_{y^*}}{k_{y^*}} [Y]_T$	$\rho^{P_y/Y} := \frac{[P_y]_T}{[Y]_T}$	$\tilde{K}_{y^*} = \frac{\tilde{d}_{y^*} + \tilde{k}_{y^*}}{\tilde{a}_{y^*}}$		
		$y : x^* := \frac{[Y : X^*]}{[X]_T}$	$\tilde{d}_y := \frac{d_y}{k_{x^*}}$				
		$p_y := \frac{[P_y]}{[P_y]_T}$	$\tilde{d}_{y^*} := \frac{d_{y^*}}{k_{x^*}}$				
		$y^* : p_y := \frac{[Y^* : P_y]}{[P_y]_T}$	$\tilde{k}_y := \frac{k_y}{k_{x^*}}$				
			$\tilde{k}_{y^*} := \frac{k_{y^*}}{k_{x^*}}$				
	Bicyclic system	$z := \frac{[Z]}{[Z]_T}$	$\tilde{a}_z := \frac{a_z}{k_{z^*}} [Z]_T$	$\rho^{Y/Z} := \frac{[Y]_T}{[Z]_T}$	$\tilde{K}_z = \frac{\tilde{d}_z + \tilde{k}_z}{\tilde{a}_z}$	$\alpha_z = \frac{\tilde{k}_z}{\tilde{k}_{z^*}} \frac{\rho^{Y/Z}}{\rho^{P_z/Z}}$	
		$z^* := \frac{[Z^*]}{[Z]_T}$	$\tilde{a}_{z^*} := \frac{a_{z^*}}{k_{z^*}} [Z]_T$	$\rho^{P_z/Z} := \frac{[P_z]_T}{[Z]_T}$	$\tilde{K}_{z^*} = \frac{\tilde{d}_{z^*} + \tilde{k}_{z^*}}{\tilde{a}_{z^*}}$		
		$z : y^* := \frac{[Z : Y^*]}{[Y]_T}$	$\tilde{d}_z := \frac{d_z}{k_{x^*}}$				
		$p_z := \frac{[P_z]}{[P_z]_T}$	$\tilde{d}_{z^*} := \frac{d_{z^*}}{k_{x^*}}$				
		$z^* : p_z := \frac{[Z^* : P_z]}{[P_z]_T}$	$\tilde{k}_z := \frac{k_z}{k_{x^*}}$				
		$\tilde{k}_{z^*} := \frac{k_{z^*}}{k_{x^*}}$					

Table 3.4: Definition of dimensionless variables and parameters.

Transient model:	Steady-state model:
$\frac{dx^*}{d\tau} = \rho^{S/X} \tilde{k}_X \{x : s^*\} + \rho^{PX/X} (\tilde{d}_{X^*} \{x^* : p_X\} - \tilde{a}_{X^*} x^* p_X) + (\tilde{d}_Y + \tilde{k}_Y) \{y : x^*\} - \tilde{a}_Y x^* y$	$0 = \{x : s^*\} - \frac{x}{\tilde{K}_X + x}$
$\frac{d\{x : s^*\}}{d\tau} = \tilde{a}_X x s^* - (\tilde{d}_X + \tilde{k}_X) \{x : s^*\}$	$0 = \{x^* : p_X\} - \frac{x^*}{\tilde{K}_{X^*} + x^*}$
$\frac{d\{x^* : p_X\}}{d\tau} = \tilde{a}_{X^*} x^* p_X - (\tilde{d}_{X^*} + 1) \{x^* : p_X\}$	$0 = \alpha_X \{x : s^*\} - \{x^* : p_X\}$
$\frac{dy^*}{d\tau} = \rho^{X/Y} \tilde{k}_Y \{y : x^*\} + \rho^{PY/Y} (\tilde{d}_{Y^*} \{y^* : p_Y\} - \tilde{a}_{Y^*} y^* p_Y) + (\tilde{d}_Z + \tilde{k}_Z) \{z : y^*\} - \tilde{a}_Z y^* z$	$0 = y x^* - \tilde{K}_Y \{y : x^*\}$
$\frac{d\{y : x^*\}}{d\tau} = \tilde{a}_Y y x^* - (\tilde{d}_Y + \tilde{k}_Y) \{y : x^*\}$	$0 = \{y^* : p_Y\} - \frac{y^*}{\tilde{K}_{Y^*} + y^*}$
$\frac{d\{y^* : p_Y\}}{d\tau} = \tilde{a}_{Y^*} y^* p_Y - (\tilde{d}_{Y^*} + \tilde{k}_{Y^*}) \{y^* : p_Y\}$	$0 = \alpha_Y \{y : x^*\} - \{y^* : p_Y\}$
$\frac{dz^*}{d\tau} = \rho^{Y/Z} \tilde{k}_Z \{z : y^*\} + \rho^{PZ/Z} (\tilde{d}_{Z^*} \{z^* : p_Z\} - \tilde{a}_{Z^*} z^* p_Z)$	$0 = z y^* - \tilde{K}_Z \{z : y^*\}$
$\frac{d\{z : y^*\}}{d\tau} = \tilde{a}_Z z y^* - (\tilde{d}_Z + \tilde{k}_Z) \{z : y^*\}$	$0 = \{y^* : p_Z\} - \frac{z^*}{\tilde{K}_{Z^*} + z^*}$
$\frac{d\{z^* : p_Z\}}{d\tau} = \tilde{a}_{Z^*} z^* p_Z - (\tilde{k}_{Z^*} + \tilde{d}_{Z^*}) \{z^* : p_Z\}$	$0 = \alpha_Z \{z : y^*\} - \{y^* : p_Z\}$
Conservation relations:	
$0 = 1 - x - x^* - \rho^{S/X} \{x : s^*\} - \rho^{PX/X} \{x^* : p_X\} - \{y : x^*\}$	
$0 = 1 - s - s^* - \{x : s^*\}$	
$0 = 1 - p_X - \{x^* : p_X\}$	
$0 = 1 - y - y^* - \rho^{X/Y} \{y : x^*\} - \rho^{PY/Y} \{y^* : p_Y\} - \{z : y^*\}$	
$0 = 1 - p_Y - \{y^* : p_Y\}$	
$0 = 1 - z - z^* - \rho^{Y/Z} \{z : y^*\} - \rho^{PZ/Z} \{z^* : p_Z\}$	
$0 = 1 - p_Z - \{z^* : p_Z\}$	

Table 3.5: Dimensionless mechanistic models: monocyclic system (green), addition for bicyclic system (magenta) and addition for tricyclic system (blue).

3.4 Design criteria

Several criteria have been considered to assess signal transduction properties. Ultrasensitivity (Fig. 3.3A) and gain amplification (Fig. 3.3B) are steady-state criteria, whereas rise time (Fig. 3.3C) and decay time (Fig. 3.3D) are transient criteria.

3.4.1 Ultrasensitivity

Signaling pathways with ultrasensitive input-output characteristics convert gradual changes in all-or-nothing type decisions. This property is defined as the response of a system that is more sensitive to the input concentration than a normal hyperbolic Michaelis-Menten response. For instance, to increase the reaction rate 9-fold from 10% to 90% of the maximum activation, a typical Michaelis-Menten response requires 81-fold increase in input concentration. An ultrasensitive response should need less ligand concentration change to accomplish the same. The degree of ultrasensitivity depends on the size of the input window within which the response changes from *nothing* to *all*.

The ability of signaling modules to produce ultrasensitive response has been observed experimentally, and many physiological phenomena such as cancer progression and morphogenesis are associated with it (see, e.g. (111)). Several mechanisms can lead to switch-like stimulus-response curves: cooperative allosteric effect of multisite protein (114), saturation of enzymes (44) and positive feedback (115). When combined with negative feedback loop, ultrasensitive modules can lead to oscillations (93), but these modules make system more robust against stochastic fluctuations (46).

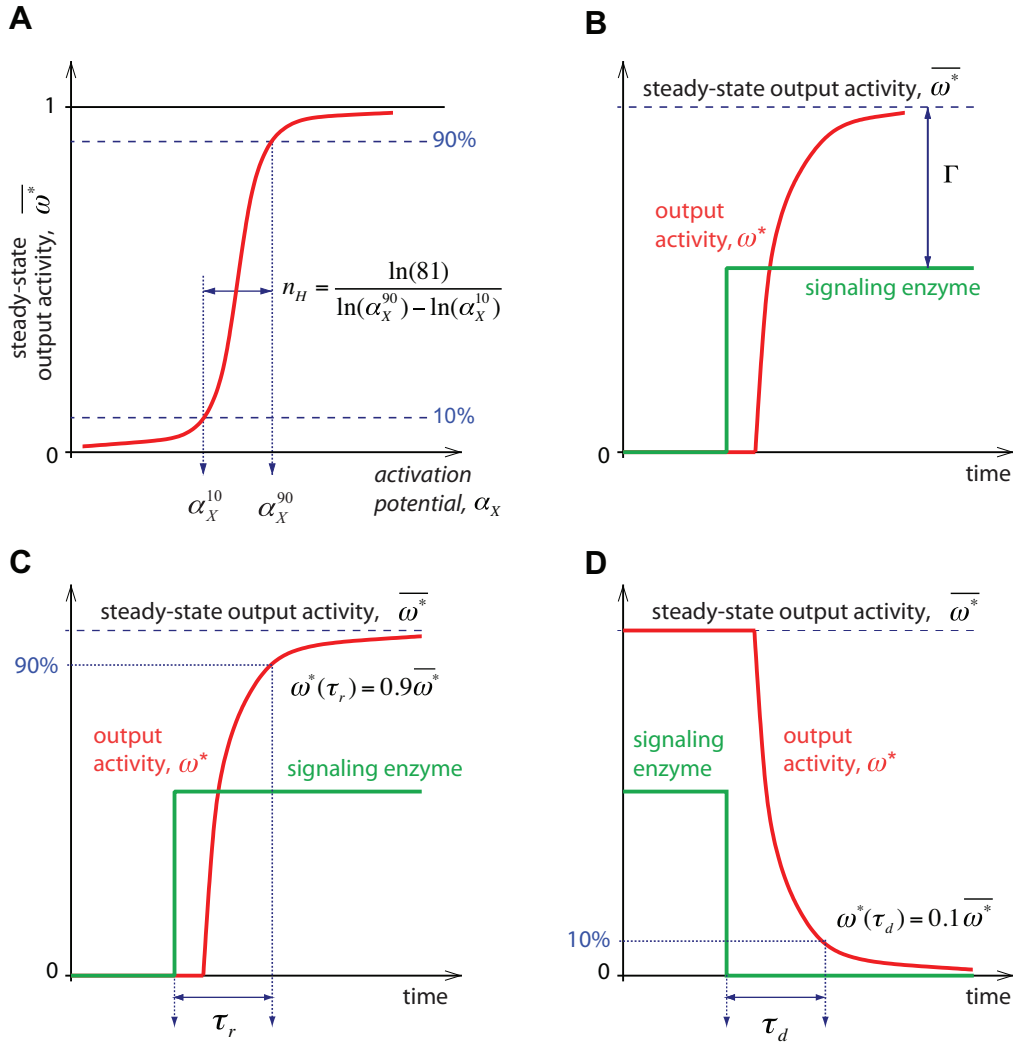


Figure 3.3: Design criteria. (A) The ultrasensitivity coefficient n_H provides a measure of the sensitivity of output activity to changes in input signal activity, at steady state. (B) The amplification gain Γ is a measure of the strength of the response to a given input level, at steady state. (C) The rise time τ_r is a measure of the speed at which stimulation is transduced. (D) The decay time τ_d provides a measure of the speed at which an activated state vanishes upon canceling the stimulation.

Goldbeter and Koshland (44) introduced a quantification of ultrasensitivity in signaling cycles based on the similarity of the ultrasensitive response with the sigmoidal (Hill) kinetics of allosteric proteins. They quantified ultrasensitivity by the apparent Hill coefficient n_H of the output response as

$$n_H = \frac{\ln(81)}{\ln(\alpha_X^{90}) - \ln(\alpha_X^{10})}, \quad (3.25)$$

where the variables α_X^{10} , α_X^{90} are the *activation potentials* of the signaling kinase required to achieve 10% and 90%, respectively, of the maximal activation of the output (Fig. 3.3A):

$$\alpha_X^{10} : \bar{\omega}^*(\alpha_X^{10}) = 0.1 \times \lim_{\alpha_X \rightarrow \infty} \bar{\omega}^*(\alpha_X) \quad (3.26)$$

$$\alpha_X^{90} : \overline{\omega^*}(\alpha_X^{90}) = 0.9 \times \lim_{\alpha_X \rightarrow \infty} \overline{\omega^*}(\alpha_X) \quad (3.27)$$

where ω^* represents the last activated kinase in the system, $\omega^* \in \{x^*, y^*, z^*\}$.

The larger the apparent Hill coefficient in Eq. (3.25), the more ultrasensitive the cascade.

3.4.2 Amplification

The amplification gain in signaling modules defines a measure of response strength. It is a relevant biological characteristic because the response produced by a signaling pathway must exceed a certain magnitude in order to trigger downstream reactions. However, a very large amplification may not always be warranted for a signal to be transmitted to its final target.

The amplification gain is defined as the ratio of activated substrate concentration to input concentration at steady state. We consider that the relevant input concentration accounts for the active form of the receptor, either free or in the complex with the kinase (Fig. 3.3B).

3.4.3 Signaling times

Signaling time is the time needed for an output to reach a certain threshold with respect to a reference state after input activation. This is an important characteristic of signal transduction pathways, and short signaling time has been proposed as a desirable biological characteristic in these systems (112, 116). The signaling time also depends on the input activation levels. We define this characteristic time as the time from the initial state to 90% of the corresponding steady state for a step change in the input signal (Fig. 3.3, C and D).

Specifically, we consider the signaling time in two different cases:

- The *rise time* τ_r – a measure of how fast an activation signal propagates through a cycle. In this study, we specifically define it as the time needed to reach 90% of the steady-state substrate activation in response to sustained step activation of the input, starting from the ground state (Fig. 3.3C):

$$\tau_r : \omega^*(\tau_r) = 0.9 \times \overline{\omega^*}(\tau_r), \quad \text{with } \omega^*(0) = 0, \quad (3.28)$$

where the coefficient 0.9 denotes that the system needs to reach 90% of its maximal activation $\overline{\omega^*}$ and the ground state corresponds to the no-signal case. Note that this definition of signaling time differs from the one in (112), which corresponds to the average time needed to activate the substrate. The latter definition cannot be considered here since it grows to infinity in the case of a permanently activated pathway.

- The *decay time* τ_d – the time needed for the substrate activity of interconvertible kinase, starting from a stimulated state $\omega^*(0)$, to decrease to within 10% of this initial activity after the stimulus has been removed (Fig. 3.3D):

$$\tau_d : \omega^*(\tau_d) = 0.1 \times \omega^*(0), \quad \text{with } \omega^*(0) = \overline{\omega^*}. \quad (3.29)$$

The above definitions assume step inputs. Alternatively, one could consider exponential, impulse or rectangular inputs, which would require redefining Eq. (3.28) and Eq. (3.29).

According to our definition of inputs, we can use two different formulations of the fraction of inactive receptor, namely:

$$s := \begin{cases} 1, & \tau \leq 0 \\ 0, & \tau > 0 \end{cases} \quad (3.30)$$

and

$$s := \begin{cases} 0, & \tau \leq 0 \\ 1, & \tau > 0 \end{cases} \quad (3.31)$$

The event of the tuning on/off the receptor is assumed to occur under very fast kinetics (i.e. $\nu = 1000$) and in the two cases of different inputs it can be defined as:

$$\frac{ds}{d\tau} = -\nu s \quad (3.32)$$

and

$$\frac{ds}{d\tau} = \nu s^* \quad (3.33)$$

We incorporate Eq. (3.32), when analyzing signaling rise time and Eq. (3.33), when analyzing signaling decay time, into transient model.

Pianoforte Sonata

Optimal Design for Ultrasensitivity



Our computational framework for optimal design is wrapped into Simulation and Optimization Collection – SOC (117), a collection of C++ classes for definition and solution of simulation and optimization problems. DAEOC is the main exploited class and it serves for the definition and numerical solution of optimal control problem in differential-algebraic equations (DAEs). In DAEOC, the control profiles are approximated by Lagrange polynomials (control parameterization) to yield a finite-dimensional optimization problem with parametric DAEs in the constraints.

The class DAEOC allows the definition and solution of optimal control problems with index-1, multistage, nonlinear DAEs embedded that conform to the following general formulation:

$$\min_{\substack{u^{(j)}(t) \in [u^{(j),L}(t), u^{(j),U}(t)], \\ \Delta t^{(j)} \in [\Delta t^{(j),L}, \Delta t^{(j),U}], \\ p \in [p^L, p^U]}} \phi_0[x^{(j)}(t_k^{(j)}), x^{(j)}(t_k^{(j)}), y^{(j)}(t_k^{(j)}), u^{(j)}(t_k^{(j)}), p, t_k^{(j)}]_{\substack{j=1, \dots, n_m \\ k=1, \dots, n_s^{(j)}}}$$

subject to:

$$\phi_i[x^{(j)}(t_k^{(j)}), x^{(j)}(t_k^{(j)}), y^{(j)}(t_k^{(j)}), u^{(j)}(t_k^{(j)}), p, t_k^{(j)}]_{\substack{j=1, \dots, n_m \\ k=1, \dots, n_s^{(j)}}} =, \leq, \geq 0, i = 1, \dots, n_c$$

with the following initial value problem in DAEs in the constraints:

$$F^{(j)}[x^{(j)}(t), x^{(j)}(t), y^{(j)}(t), u^{(j)}(t), p, t] = 0, \forall t \in (t_0^{(j)}, t_{n_s^{(j)}}^{(j)}], j = 1, \dots, n_m$$

$$G^{(1)}[x^{(1)}(t_0^{(1)}), x^{(1)}(t_0^{(1)}), y^{(1)}(t_0^{(1)}), u^{(1)}(t_0^{(1)}), p] = 0$$

$$G^{(j)}[x^{(j-1)}(t_{n_s^{(j-1)}}^{(j-1)}), x^{(j-1)}(t_{n_s^{(j-1)}}^{(j-1)}), y^{(j-1)}(t_{n_s^{(j-1)}}^{(j-1)}), x^{(j)}(t_0^{(j)}), x^{(j)}(t_0^{(j)}), y^{(j)}(t_0^{(j)}), u^{(j)}(t_0^{(j)}), p] = 0, j = 2, \dots, n_m$$

where n_m denotes the number of modes, $n_s^{(j)}$ the number of stages in the j^{th} mode, and n_c the number of (equality or inequality) constraints. In this formulation, $x^{(j)}$ and $y^{(j)}$ stand for the differential and algebraic state variables, respectively, in the mode j^{th} ; $u^{(j)}$, for the (infinite dimensional) control variables in the j^{th} mode; p , for the time-invariant parameters; and $t_k^{(j)}$, for the k^{th} stage time in the j^{th} mode. The numbers of state variables $n_x^{(j)}$, algebraic variables $n_y^{(j)}$, control variables $n_u^{(j)}$ and control stages $n_s^{(j)}$ are allowed to change from one mode to the other. Continuity of the differential state variables is automatically enforced from one time stage to the next in a given mode.

The solution of such optimal control problems relies on the parameterization of the infinite-dimensional control profiles $u(t)$ using Lagrange polynomials. For flexibility, the order of the Lagrange polynomial used to approximate each control variable can be set separately, and different orders can be used from one mode to the other. Continuity of the control variables at the time stages of a given mode can also be enforced during the solution process.

The organization of the framework, considering top-level user functions, is demonstrated in the Fig. 4.1.

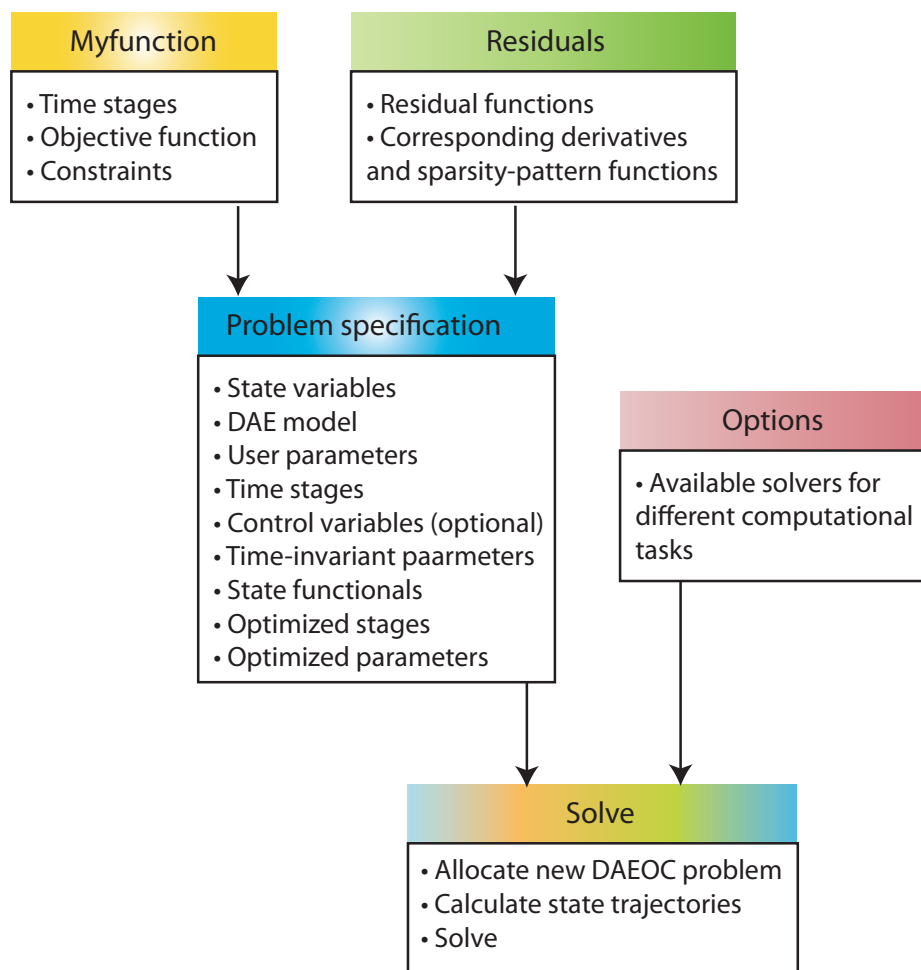


Figure 4.1: Computational framework for optimal design.

Systems of DAEs that build the models of observed signaling modules are both sparse and

stiff. The sparsity pattern comes from the fact that time derivatives of state variables depend only on a small number of other state variables. Considering the parameter ranges of six order of magnitude again brings the stiffness issues, so appropriate solver should be chosen.

Ideally, the optimization problem with embedded dynamics should be solved with a deterministic global optimization method. Local optimization of variables was carried out from the large number of random starting points (multi start). The local optimization was carried out with the nonlinear programming (NLP) solver SNOPT (118), which implements a sequential quadratic programming (SQP) algorithm (80). SQP is one of the most effective methods for nonlinearly constrained optimization and it generates its steps by solving quadratic subproblems. SNOPT terminates each local optimization when a Karush-Kuhn-Tucker point is reached, to within a numerical tolerance 10^{-9} . Local optimization from a large set of random initial parameter values was used and the best optimum was taken as the global optimum. As a check of convergence, we required this global optimum to occur multiple times starting from different initial conditions.

We have used the solver DSL48S (119), which is part of DAEPACK (120), for solving the initial value problems in DAEs. This package is particularly suited for large-scale problems. DAEPACK is also used for consistent initialization and for calculating first-order parametric sensitivities and for all the necessary differentiations (of the model objective and constraint functions).

4.2 Optimal design for ultrasensitive monocyclic system

The ultrasensitivity of covalent modification cycles has been studied extensively, since it was first discussed by Goldbeter and Koshland (44). Their classical results suggest that a monocyclic cascade can display ultrasensitive responses even when the interconversion steps follow Michaelis-Menten kinetics. Here, we perform inverse sensitivity analysis in order to identify the subdomains of the parameter space that lead to ultrasensitive responses.

The calculation of apparent Hill coefficients in complex signaling cycles can be computationally expensive, especially considering that the study of design criteria involves exhaustive parameter search. To automate the computations, the problem can be reformulated and rearranged as a set of DAEs in α_X , using numerical continuation (121).

Then, steady-state model Eq. (3.16–3.21) is differentiated with respect to α_X :

$$0 = \frac{d\{\bar{x} : s^*\}}{d\alpha_X} + \frac{\tilde{K}_X}{(\tilde{K}_X + \bar{x})^2} \left(\frac{d\bar{x}^*}{d\alpha_X} + \rho^{S/X} \frac{d\{x : s^*\}}{d\alpha_X} + \rho^{P_X/X} \frac{d\{x^* : p_X\}}{d\alpha_X} \right) \quad (4.1)$$

$$0 = \frac{d\{x^* : p_X\}}{d\alpha_X} - \frac{\tilde{K}_{X^*}}{(\tilde{K}_{X^*} + \bar{x}^*)^2} \frac{d\bar{x}^*}{d\alpha_X} \quad (4.2)$$

$$0 = \frac{d\{x^* : p_X\}}{d\alpha_X} - \alpha_X \frac{d\{x : s^*\}}{d\alpha_X} - \{x : s^*\} \quad (4.3)$$

$$1 = \bar{x} + \bar{x}^* + \rho^{S/X} \{x : s^*\} + \rho^{P_X/X} \{x^* : p_X\} \quad (4.4)$$

$$1 = \bar{s} + \bar{s}^* + \{x : s^*\} \quad (4.5)$$

$$1 = \bar{p}_X + \{x^* : p_X\}. \quad (4.6)$$

The values of α_X^{10} and α_X^{90} are identified via Eq. (4.1–4.6), e.g. using state event detection techniques (122).

The compact notation of the steady-state model is represented as:

$$\mathbf{G}(\xi, \alpha_X, \mathbf{q}, \mathbf{p}) = \mathbf{0}, \quad (4.7)$$

with the parameters that explicitly appear in this formulation and which are linear combination of model parameters: $\mathbf{q} := (\tilde{K}_X \ \tilde{K}_{X^*})^T$.

We measure how the input signal s^* affects the relative concentration of protein in active state x^* . We identify the dimensionless Michaelis-Menten constants \tilde{K}_X , \tilde{K}_{X^*} and the concentration ratios $\rho^{S/X}$, $\rho^{P_x/X}$ as parameters responsible for the coupling of the network. First, we observe how ultrasensitivity is determined by the values of \tilde{K}_X and \tilde{K}_{X^*} for fixed values of $\rho^{S/X}$ and $\rho^{P_x/X}$ (Fig. 4.2).

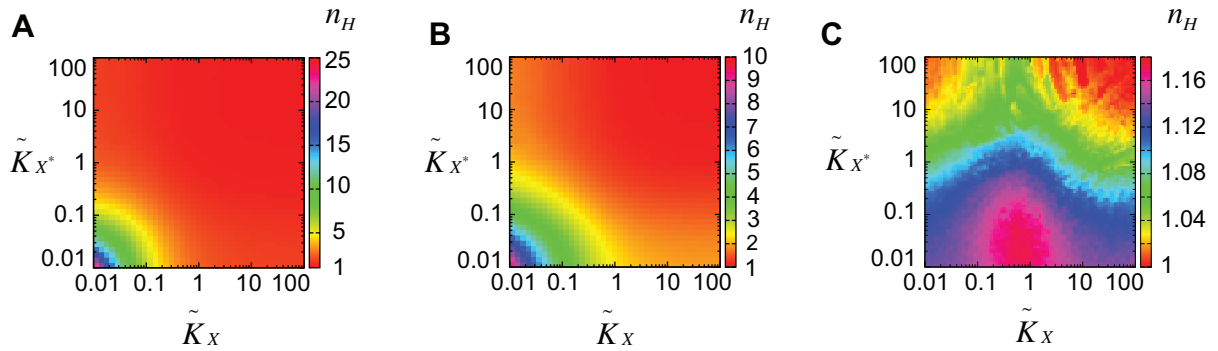


Figure 4.2: Landscape of Hill coefficient values n_H versus dimensionless Michaelis-Menten constants \tilde{K}_X and \tilde{K}_{X^*} for two different values of concentration ratios: (A) $\rho^{S/X} = \rho^{P_x/X} = 10^{-2}$, (B) $\rho^{S/X} = \rho^{P_x/X} = 10^{-1}$, (C) $\rho^{S/X} = \rho^{P_x/X} = 1$.

This analysis is in good agreement with main conclusions of Goldbeter and Koshland, namely:

- Ultrasensitivity is promoted by low values of dimensionless Michaelis-Menten constants \tilde{K}_X and \tilde{K}_{X^*} .
- Ultrasensitivity is only feasible when the total concentration of activating enzyme $[S]_T$ and the total concentration of deactivating enzyme $[P_x]_T$ are small compared to the total concentration of the interconvertible protein $[X]_T$, i.e. $\rho^{S/X}$ and $\rho^{P_x/X} \ll 1$.
- The system response is symmetric if: $\tilde{K}_X = \tilde{K}_{X^*}$ and $\rho^{S/X}, \rho^{P_x/X} \rightarrow 0$.

A number of studies have since elaborated on these results (123, 124). We propose to study this property using the inverse sensitivity approach. To determine the optimal values of dimensionless Michaelis-Menten constants $\{\tilde{K}_X, \tilde{K}_{X^*}\}$, the following optimization problem is considered:

$$\begin{aligned} &\text{find} && \tilde{K}_X, \tilde{K}_{X^*} \text{ that} && (P_1) \\ &\text{maximize} && n_H, \text{ Eq. (3.25)} \end{aligned}$$

subject to steady-state model, Eq. (4.1–4.6).

The parameters \tilde{K}_X and \tilde{K}_{X^*} are taken in the range $[10^{-2}, 10^2]$, which was found to be wide enough from comparisons with larger parameter ranges. Moreover, following the observations in Fig. 4.2, the concentration ratio space is set to $\{\rho^{S/X}, \rho^{P_X/X}\} \in [10^{-2}, 1]$.

The mathematical formulation reads:

$$\begin{aligned} \max_{\mathbf{q}, \alpha_X^{10}, \alpha_X^{90}} \quad & n_H = \frac{\ln(81)}{\ln(\alpha_X^{90}) - (\alpha_X^{10})} \quad (P_1) \\ \text{subject to} \quad & \frac{\partial \mathbf{G}}{\partial \xi} \frac{d\xi}{d\alpha_X} + \frac{\partial \mathbf{G}}{\partial \alpha_X} = \mathbf{0}, \quad 0 \leq \alpha_X \leq \alpha_X^\infty, \quad \mathbf{G}(\xi(0), 0, \mathbf{q}, \mathbf{r}) = \mathbf{0}, \\ & x^*(\alpha_X^{10}) = 0.1x^*(\alpha_X^\infty), \quad x^*(\alpha_X^{90}) = 0.9x^*(\alpha_X^\infty), \\ & 0 \leq \alpha_X^{10} \leq \alpha_X^{90} \leq \alpha_X^\infty, \quad 10^{-2} \leq \mathbf{q} \leq 10^2. \end{aligned}$$

As noted, initial conditions $\xi(0)$ are determined from $\mathbf{G}(\xi(0), 0, \mathbf{q}, \mathbf{r}) = \mathbf{0}$. The two latter equations give $x^*(0) = \{x^* : p_X\}(0) = 0$ and $p_X(0) = 1$; the remaining initial concentrations $x(0)$, $s^*(0)$ and $\{x : s^*\}(0)$ depend on the kinetic constant \tilde{K}_X and concentration ratio $\rho^{S/X}$ only, and are given by:

$$0 = x(0)^2 + (\tilde{K}_X + \rho^{S/X} - 1)x(0) - \tilde{K}_X \quad (4.8)$$

$$\{x : s^*\}(0) = \frac{1 - x(0)}{\rho^{S/X}} \quad (4.9)$$

$$s^*(0) = 1 - \frac{1 - x(0)}{\rho^{S/X}}. \quad (4.10)$$

The results, which are given in graphical form in Fig. 4.3, lead to the following observations:

- Significant ultrasensitivity ($n_H > 6$) can only be achieved i) for values of dimensionless Michaelis-Menten constants \tilde{K}_X and \tilde{K}_{X^*} lower than 10^{-1} and ii) for small values of the two concentration ratios $\rho^{S/X}, \rho^{P_X/X} \ll 1$.
- Ultrasensitivity can accommodate higher fluctuations in phosphatase level ($\rho^{P/X} < 0.5$) than in activating enzyme level ($\rho^{S/X} < 0.1$).

The first observations is not a new result for it was suggested as a design criterion for ultrasensitivity already by (44). On the other hand, the interplay and relative effect of the activation and deactivation steps constitutes a new observation that follows from the proposed optimization methodology. In particular, this result also suggests that overexpression in activating proteins, such as receptors relative to target kinases, might lead to (pathological) loss of ultrasensitivity.

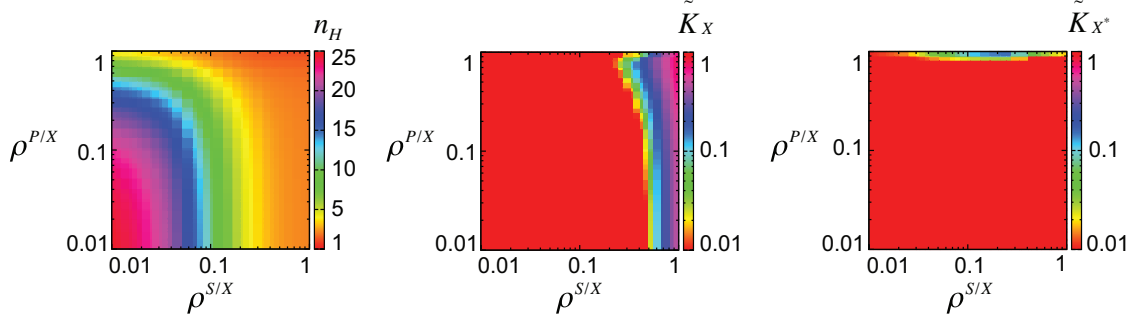


Figure 4.3: Design for maximal ultrasensitivity in monocyclic system. The optimal Hill coefficient and Michaelis-Menten constants are calculated by solving the optimization problem (P₁), for values of the concentration ratios $\rho^{S/X}$ and $\rho^{P_X/X}$ in the range $[10^{-2}, 1]$. The largest possible Hill coefficient n_H (left plot), with corresponding dimensionless Michaelis-Menten constant \tilde{K}_x (middle plot) and corresponding dimensionless Michaelis-Menten constant \tilde{K}_{x^*} (right plot).

4.3 The longer, the better: Optimal design for ultrasensitive tricyclic system

Early work on modeling these cascades revealed that they are capable of displaying switch-like activation of the third level kinase (MAPK) in response to activation of the top-level kinase (MAPKKK), without including any cooperative kinetic mechanism (III). The ability of an interconvertible enzyme cascade to generate an ultrasensitive response is important for the regulation of the various processes under control of the MAPK pathways. Ultrasensitive systems will be insensitive to small fluctuations in stimuli, with no response elicited until the threshold is crossed. With a hyperbolic response, the cell partially responds to any change in stimulus level. Thus, all-or-none responses of ultrasensitive cascades are advantageous for regulating processes such as cell differentiation or division.

Coupling in the tricyclic signaling networks is occurred through the formation of kinase-kinase complexes. The concentration of these complexes, and consequently the strength of interactions, are determined by two classes of dimensionless parameters: i) the ratios of concentrations of the first to the second stage kinases $\rho^{X/Y}$ and the second to the third stage kinases of the cascade $\rho^{Y/Z}$, and ii) the affinities of complex formation, i.e. the dimensionless Michaelis-Menten constants of the three kinases $\tilde{K}_X, \tilde{K}_Y, \tilde{K}_Z$.

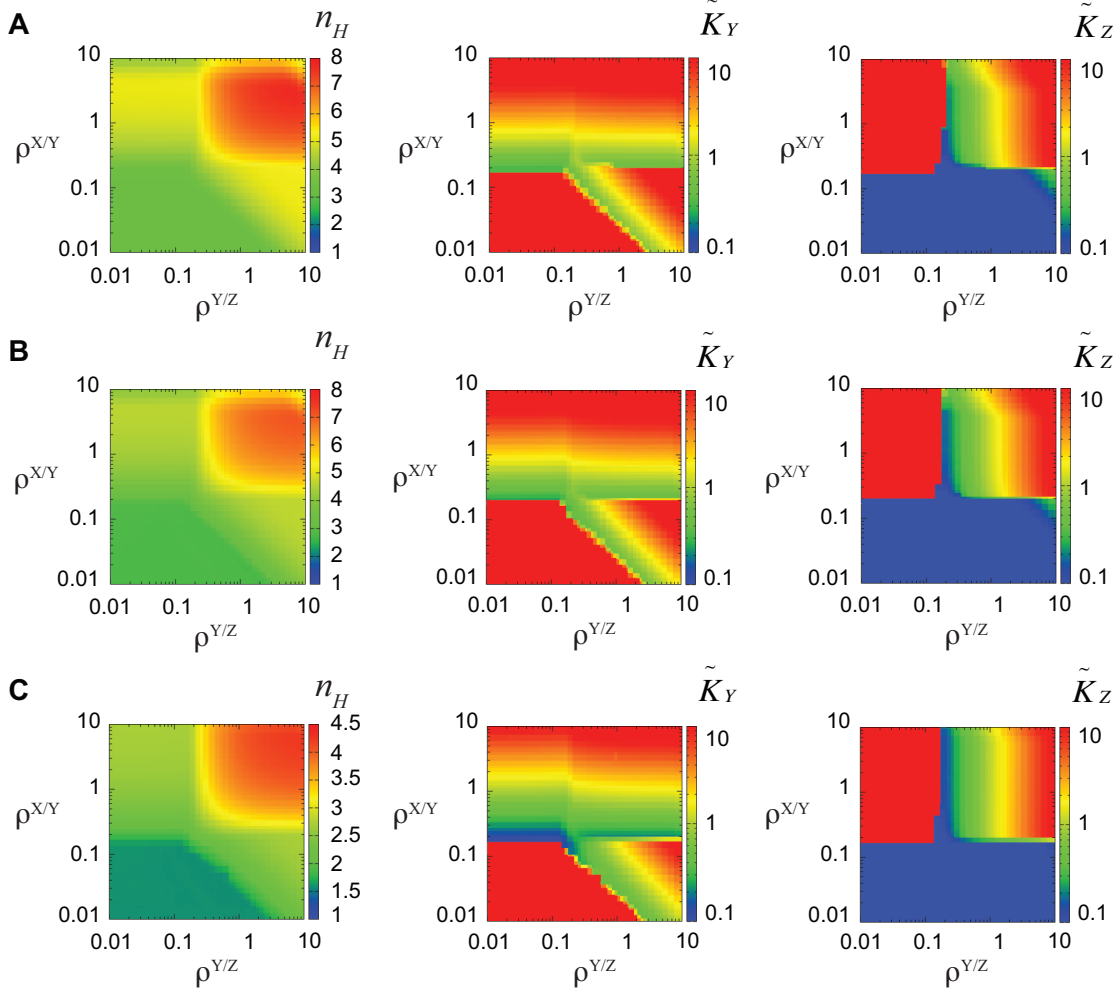


Figure 4.4: Maximum Hill coefficient n_H of tricyclic cascade system, for specified concentration ratios $\rho^{X/Y}, \rho^{Y/Z}$ and corresponding Michaelis-Menten dimensionless constants \tilde{K}_Y and \tilde{K}_Z . (A) $\rho^{S/X} = 10^{-2}$, (B) $\rho^{S/X} = 10^{-1}$, (C) $\rho^{S/X} = 1$.

The ratios of phosphatases to their substrate kinase concentrations were fixed at $\{\rho^{P_x/X}, \rho^{P_y/Y}, \rho^{P_z/Z}\} = \{0.1, 0.1, 0.1\}$ and their dimensionless Michaelis-Menten constants at $\{\tilde{K}_X^*, \tilde{K}_Y^*, \tilde{K}_Z^*\} = \{0.1, 0.1, 0.1\}$.

As in previous subsection, we perform inverse sensitivity analysis in order to identify parameter landscape that leads to ultrasensitive response. The problem is structured as:

$$\begin{array}{ll}
 \text{find} & \tilde{K}_X, \tilde{K}_Y, \tilde{K}_Z \text{ that} \\
 \text{maximize} & \text{ultrasensitivity } n_H, \text{ Eq. (3.25)} \\
 \text{subject to} & \text{steady-state model,}
 \end{array} \tag{P_2}$$

for given values of $\{\rho^{X/Y}, \rho^{Y/Z}\} \in [0.01, 10]$. The focus is to determine how the input signal s^* affects the relative concentration of the activated protein of the last cycle z^* .

We analyze the cases with $\rho^{S/X}$ fixed to the three different values: 0.01, 0.1 and 1 (Fig. 4.4). The graphs for the Michaelis-Menten constants which will bring the system to

maximal ultrasensitivity, display almost the same trends in Fig. 4.4A and Fig. 4.4B. The only effect that different values of concentration ratio have is reflected on values of n_H , which is slightly higher if $\rho^{S/X}$ decreases. In both cases, the global optimum is placed approximately around the point $\{\rho^{X/Y}, \rho^{Y/Z}\} = \{3, 6\}$. Fig. 4.4C though clearly shows the significant decrease in the value of n_H , still preserving the similar trends for Michaelis-Menten constants.

In order to achieve values of $n_H > 6$, the system should operate in the region of $0.5 < \rho^{X/Y} < 10$ and $0.5 < \rho^{Y/Z} < 10$. Furthermore, the best possible combination of kinase affinities can be achieved when the first kinase is saturated and the other two are unsaturated by their target kinases. Interestingly, the Michaelis-Menten constant \tilde{K}_X keeps the minimal possible value along the whole concentration ratio space ($\tilde{K}_X = 0.1$), again suggesting the obvious sensitivity of the pathway to the upstream activator (data not shown).

4.4 Step back: Optimal design for ultrasensitive bicyclic system

Using the same ideas as in the previous sections, we further explored the optimal design for the ultrasensitive bicyclic cascade. The purpose is to complete the analysis and to make an optimal design bridge between monocyclic and tricyclic systems.

Following the same patterns as before, we fixed the ratios of phosphatases to their substrate kinase concentrations at $\{\rho^{P_x/X}, \rho^{P_y/Y}\} = \{0.1, 0.1\}$ and their dimensionless Michaelis-Menten constants at $\{\tilde{K}_X^*, \tilde{K}_Y^*\} = \{0.1, 0.1\}$. We formulated the optimization problem in order to get the best possible combination of kinase affinities, for the whole concentration ratio space. The results obtained by varying $\rho^{S/X}$ and $\rho^{X/Y}$ in the range $[0.01, 10]$, while optimizing the rate constants \tilde{K}_X and \tilde{K}_Y in the range $[0.1, 10]$, are shown in Fig. 4.5.

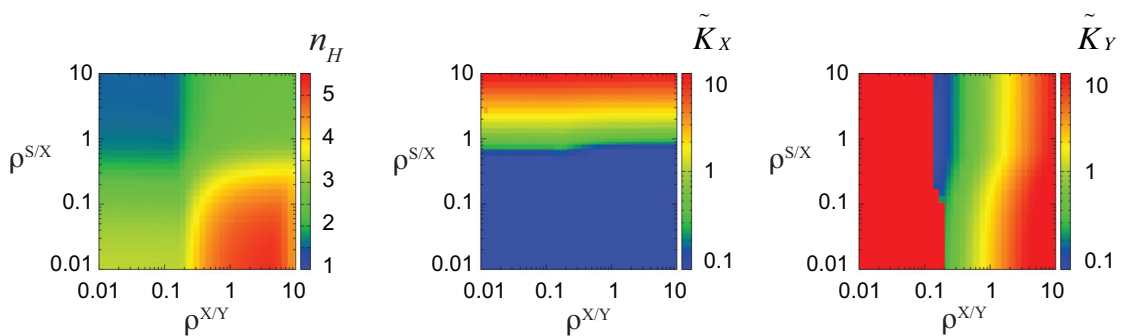


Figure 4.5: Maximum Hill coefficient n_H of bicyclic cascade system, for specified concentration ratios $\rho^{S/X}$, $\rho^{X/Y}$ and corresponding values of Michaelis-Menten dimensionless constants \tilde{K}_X and \tilde{K}_Y .

This time, region of the global maximum is in between $0.01 < \rho^{S/X} < 0.1$ and $1 < \rho^{X/Y} < 10$, and as before the first kinase should be saturated and the second should be unsaturated with its target kinase. Unlike it was the case in the tricyclic system, first-level kinase affinity does not keep the lowest value in the whole concentration ratio space. The values of n_H produced that way are not superior if compared with the corresponding monocyclic cascade. This suggests that, even if the optimal design dictates the certain arrangement, it does not necessarily mean that the overall improvement of ultrasensitivity is satisfied.

4.5 Modular differences and similarities

The ability of signaling pathways and other enzymatic cascades to generate a switch-like response through ultrasensitivity has been widely debated in recent years (111, 114, 123-126). However, the idea remains attractive as a mean of regulating cellular responses to stimuli, and genetic circuits capable of displaying Hill coefficients up to 7.5 have been constructed (127). Therefore, there should be increasing interest in the future in the rational design of genetic circuits and further quantitative analyses of existing signaling pathways with increasingly sophisticated experimental tools. However, the rational design or analysis of an ultrasensitive network requires an understanding of how the different network components couple together and interact. Here, we have systematically explored the parameter space for simple ultrasensitive cascade networks and studied their emerging properties when all the interactions are explicitly included. The conditions shown for global maxima in the Hill coefficients for tricyclic cascade models are not similar to conditions in single cascades. The minimum ratio of kinase concentrations and most saturated kinase affinities resulted in Hill coefficients of less than three. This demonstrated the importance of accounting for the formation of intermediate complexes and considering the actual network, rather than just the core module of the network.

4.6 Optimal design as entry point for sensitivity analysis

As we have concluded from this section, optimization techniques gave precise answers on how to design signaling modules in order to achieve maximal ultrasensitivity. Problem solving in steady state was not computationally very expensive, providing the parameter space was relatively small.

Proceeding toward more challenging tasks, such as optimal design in dynamic regime, we encounter a larger number of parameters. The questions that arise are:

- Are all the parameters equally important?
- What are the key parameters that drive certain behaviors?
- Are there any parameters that we can neglect when doing parameter estimation?
- What are the relations between different parameters?

Also, very important: are there some clear patterns for parameter values that drive the system to have optimal behavior. All this and much more can be addressed using global sensitivity analysis techniques.

We follow the idea from (96) and attribute semi-quantitative description of Michaelis-Menten constants as one of the three forms: saturated ($\tilde{K} = 0.1$), mildly saturated ($\tilde{K} = 1$), or unsaturated ($\tilde{K} = 10$). Their simulation results are aligned with our results obtained with the optimization approach, therefore the derived conclusions still hold.

For the demonstration purposes, the simulations of our monocyclic and tricyclic models with described values of Michaelis-Menten constants are presented. Setting the parameters of monocyclic system into predefined arrangement give the notable ultrasensitivity only when the system operates close to saturation (Fig. 4.6)

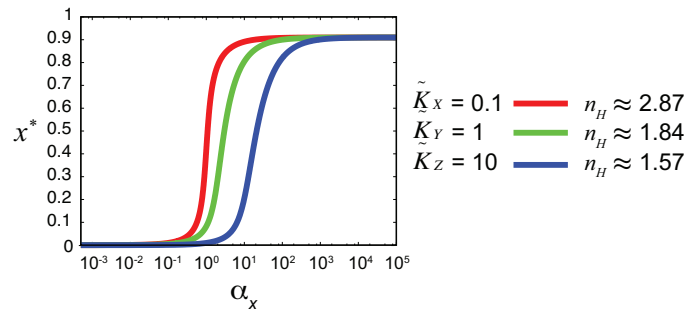


Figure 4.6: Ultrasensitive responses of monocyclic system, with different values of dimensionless Michaelis-Menten constant \tilde{K}_x , and fixed $\tilde{K}_{x^*} = 0.1$, $\rho^{S/X} = \rho^{P_x/X} = 0.1$.

Considering all possible combination of the defined Michaelis-Menten constants (27 combinations in total) led to similar findings about the significant enhancement of ultrasensitivity in tricyclic model, as well as different design rules. Fig. 4.7 summarizes all $n_H > 5$ with the sets of kinase affinities that generate these maximal Hill coefficients.

Having in mind the conclusions that monocyclic system needs to be saturated and the first kinase in multicyclic system should be saturated, whereas the others should be unsaturated by their target kinases, we proceed to our next investigation which aims at identifying the important kinetic parameters that can ensure satisfying ultrasensitivity. Furthermore, Table 4.1, which summarizes maximal values of Hill coefficient for each of 27 combinations, provides initial idea about the critical parts for ultrasensitivity in tricyclic system. Without any dilemma, Michaelis-Menten constant attributed to the first level of the signaling cascade \tilde{K}_x , influences the value of Hill coefficient the most. Michaelis-Menten constant attributed to the second cascade level \tilde{K}_y follows in the ranking. The only exception from this sequential rule is that the combinations with $\{\tilde{K}_y, \tilde{K}_z\} = \{10, 0.1\}$ seem to be less favorable than combinations with $\{\tilde{K}_y, \tilde{K}_z\} = \{1, 1\}$. Interestingly, the whole pattern for ranking of these two parameters is preserved independently of the value of \tilde{K}_x .

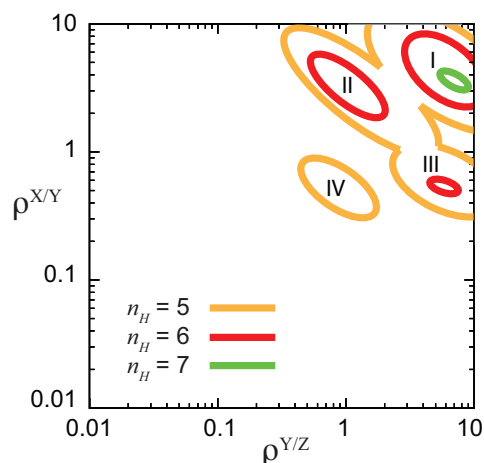


Figure 4.7: Maximum Hill coefficient of tricyclic system ($n_H \in (6, 9)$), with the corresponding sets of kinase affinities. Contours represent the maximum possible n_H for each $\{\rho^{X/Y}, \rho^{Y/Z}\}$ pair and ordinal numbers indicate the particular n_H of the contour. The set of kinase affinities that produce each n_H region are indicated by Roman Numerals: I – $\{\tilde{K}_X, \tilde{K}_Y, \tilde{K}_Z\} = \{0.1, 10, 10\}$; II – $\{\tilde{K}_X, \tilde{K}_Y, \tilde{K}_Z\} = \{0.1, 10, 1\}$; III – $\{\tilde{K}_X, \tilde{K}_Y, \tilde{K}_Z\} = \{0.1, 1, 10\}$; IV – $\{\tilde{K}_X, \tilde{K}_Y, \tilde{K}_Z\} = \{0.1, 1, 1\}$.

\tilde{K}_Z										
0.1	1	10	0.1	1	10	0.1	1	10		
2.46	3.38	3.65	2.12	2.75	2.92	1.97	2.55	2.71	0.1	\tilde{K}_Y
4.09	5.61	6.1	2.96	3.92	4.2	2.64	3.42	3.66	1	
4.84	6.59	7.15	3.34	4.38	4.68	2.91	3.79	4.01	10	
0.1			1			10				
\tilde{K}_X										
2.87			1.84			1.57			Monocyclic	

Table 4.1: Maximal Hill coefficient from semi-quantitative analysis. 27 sets of Michaelis-Menten constants for tricyclic system and comparison with monocyclic system.

Identifying the Violins in Signaling Orchestra

Sensitivity Analysis of Signaling Modules



In order to gain insights about how robust the biological responses are with respect to changes in parameters and which parameters are the key factors that affect the model outputs, we have used software package `sbtoolbox2` (128) as a base of our computational framework for sensitivity analysis. `Sbtoolbox2` operates on MATLAB and it is specialized for the creation and analysis of biological models. Models can also be imported and exported in Systems Biology Markup Language (SBML) (129), which increases the portability and generality of all studies. In its wide range of analysis tools, `sbtoolbox2` offers different algorithms for parameter sensitivity analysis, including the implementation of the Sobol method (79).

A schematic of the framework is presented in Fig. 5.1. The main file prepares the model and gives simulation specification (time, number of runs, etc.). The model is simulated with each given parameter set and each property of interest is observed. The file containing the implemented procedure for sensitivity analysis generates $N = 10\,000$ pairs of random parameter sets, chosen in the range $[10^{-3}, 10^3]$. The matrices of output values for each property are finally used to calculate first-order (Eq. (2.30)) and total sensitivity indices (Eq. (2.31)).

We observed steady-state properties (amplification and ultrasensitivity), as well as dynamic properties (rise and decay times), using dynamic representations of monocyclic and tricyclic models (Table 3.5). The models were simulated for the time up to 10^7 and the final output values were considered to be at steady state. In case no steady state had been reached after 10^7 , the corresponding parameter set was discarded, assuming that such a scenario is not of interest. Numerical integration of all models was done with MATLAB function `ode15s`, particularly designed to solve stiff differential equations and DAEs, with the relative and absolute tolerances of 10^{-6} . Computational time for the study of monocyclic system was approximately 13 hours, whereas for tricyclic system it could be as long as 10 days. The time cost accounts for the simulations done on a single desktop machine.

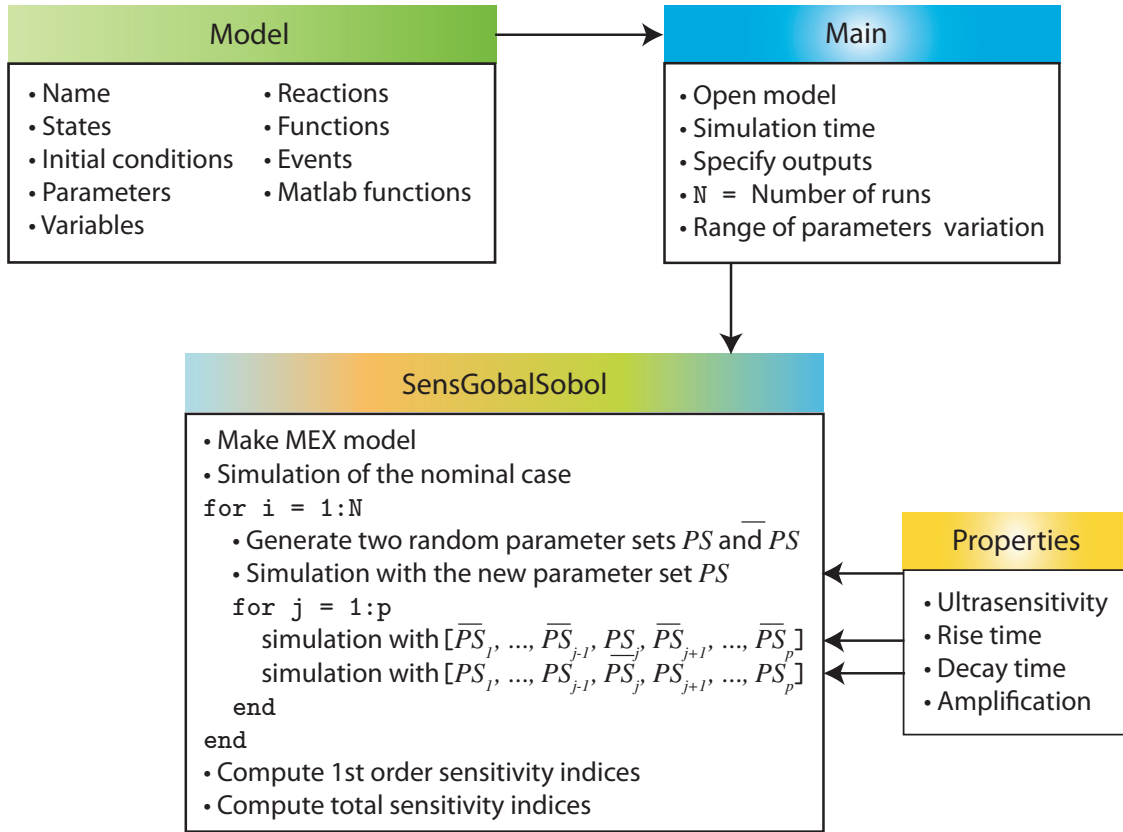


Figure 5.1: Computational framework for sensitivity analysis.

5.2 Sensitivity analysis of monocyclic system

The dynamic model of monocyclic system (Eq. (3.8–3.14)) comprises five kinetic parameters $\tilde{a}_X, \tilde{a}_{X^*}, \tilde{d}_X, \tilde{d}_{X^*}, \tilde{k}_X$ and two concentration ratios $\rho^{S/X}, \rho^{P/X}$. Since we do not use any prior knowledge about parameter values, we chose to fix them all to 1. Defining the nominal parameter set in such a way, we aim to perform unbiased identification of key parameters (reactions) in the system. The variation of three orders of magnitude on both sides around the nominal parameter value, resulting in the variation range of six orders of magnitude, is assumed to be representative enough to display parametric sensitivities. Both concentration ratios are set to be equal to 0.1. This will be the base setting for our first parameter scan and classification. Simulation of the system with chosen nominal set gives values for amplification gain of 4.8, signaling rise time of 35.8, signaling decay time of 52.2 and Hill coefficient of 1.1. As seen in previous chapter, this value of Hill coefficient is situated far in suboptimal region. Beside the purpose of the sensitivity analysis to identify the most influential parameters, we will focus on identifying the parameter ranges that will push the system from suboptimal to optimal regime. Namely, our objective is to further explore the design that could allow system to achieve simultaneously high ultrasensitivity, high amplification and short signaling times (fast response).

5.2.1 Influence on ultrasensitivity

Fig. 5.2 shows total sensitivity indices (A) of all five parameters and scatter plots (B) give direct relations between parameters and ultrasensitivity. Results suggest that the kinetic rate \tilde{a}_X has a significant influence on the change of Hill coefficient, closely followed by \tilde{a}_{X^*} . This is expected, taking into account the knowledge from optimal design of this property. These two kinetic rates are inversely proportional to Michaelis-Menten constants \tilde{K}_X (Eq. (3.20)) and \tilde{K}_{X^*} (Eq. (3.21)), which are the key parameters that drive the ultrasensitivity in the system. Scatter plots for \tilde{a}_X and \tilde{a}_{X^*} clearly show how ultrasensitivity becomes highly promoted when these two parameters are increased and how their low values restrict Hill coefficient to be smaller than two. On the other hand, kinetic rates \tilde{d}_X and \tilde{d}_{X^*} insignificantly influence ultrasensitivity. Parameter \tilde{k}_X is not taken into account, since it is in direct relation with α_X , which is used to determine value of n_H (Eq. (3.23)).

Interestingly, random sampling of the parameter space performed during all simulations statistically favors small Hill coefficient values between 1 and 2.

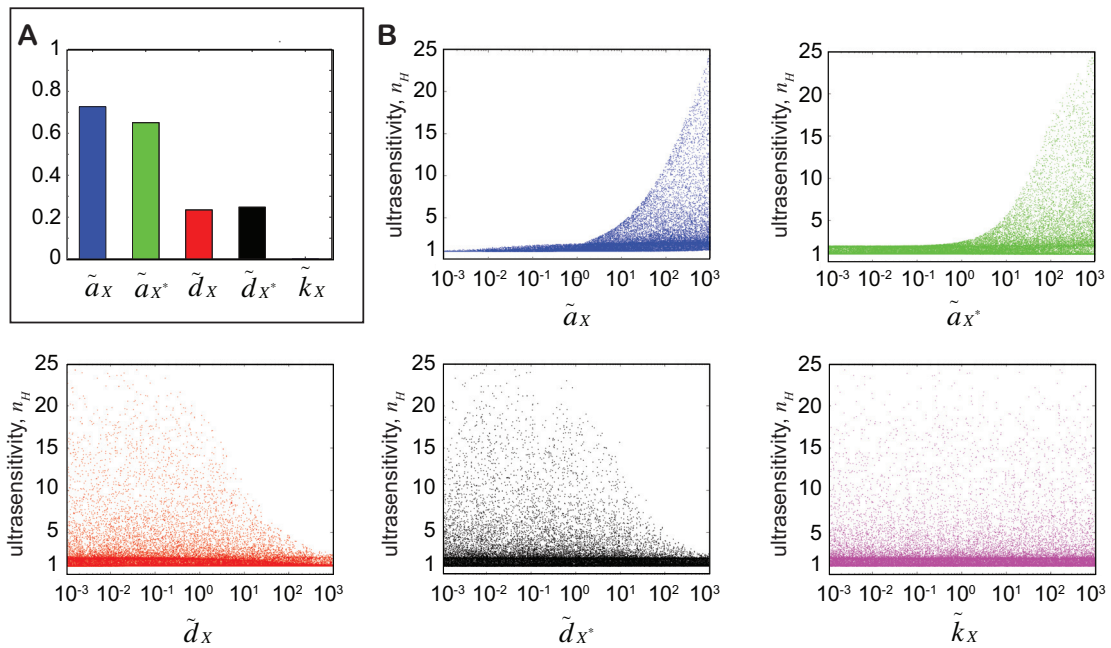


Figure 5.2: (A) Total sensitivity indices and (B) scatter plots for each parameter of monocyclic system, with respect to ultrasensitivity. The concentration ratios are set to $\{\rho^{S/X}, \rho^{P/X}\} = \{0.1, 0.1\}$. The nominal parameter set $\{\tilde{a}_X, \tilde{a}_{X^*}, \tilde{d}_X, \tilde{d}_{X^*}, \tilde{k}_X\} = \{1, 1, 1, 1, 1\}$ determines the Hill coefficient $n_H = 1.1$.

5.2.2 Influence on amplification

As mentioned before, we define the amplification gain as the ratio of activated kinase concentration to input concentration at steady state. If the relevant input concentration accounts for the active form of the receptor (either free or in the complex with the kinase), then the amplification gain of monocyclic system can be written in a dimensionless form as:

$$\Gamma := \frac{\bar{x}^*}{\rho^{S/X}} \quad (5.1)$$

where \bar{x}^* stands for the steady-state kinase activity corresponding to a given level of signaling enzyme (Fig. 3.3B). From this definition, it is clear that high amplification is favored by a small signaling enzyme-to-substrate ratio, $\rho^{S/X} \ll 1$. This also suggests that the models that ignore intermediate complex formation tend to overestimate amplification (44). Since all the concentration ratios, including $\rho^{S/X}$, are fixed to the value of 0.1 throughout this chapter, the maximal possible gain that can occur in the system would be 10.

Given the diversity of signals in cells, one cannot just simply describe what would be the expected signal amplification. Therefore, the scatter plots on Fig. 5.3B show interesting natural system design: intermediate levels of amplification gain are less likely to occur. Independently of the parameter values, the density is very high near the bottom and top lines, but low in the middle of range of possible amplification outcomes. This means that, in most experiments, the kinase gets either fully phosphorylated or stays fully unphosphorylated after activation of the pathway – half phosphorylated kinases rarely occur. Furthermore, this coincides with the ultrasensitive behavior of covalent modification cycles.

The scatter plot for parameter \tilde{a}_X itself even reflects the shape of this none-or-all type of steady-state response.

Parameters \tilde{a}_X and \tilde{a}_{X^*} are again the parameters that have the highest influence on amplification. For values of $\tilde{a}_X > 1$, the activation of the kinase is more likely to be a dominant process. This will in turn accumulate more of the active kinase form x^* , resulting in higher amplification gain. The same applies to kinetic rate \tilde{k}_X , which directly leads to x^* production. As expected, trend is the opposite for values of $\tilde{a}_X < 1$ and in the case of parameter \tilde{a}_{X^*} .

As in the case of ultrasensitivity, parameters \tilde{d}_X and \tilde{d}_{X^*} have lower sensitivity indices.

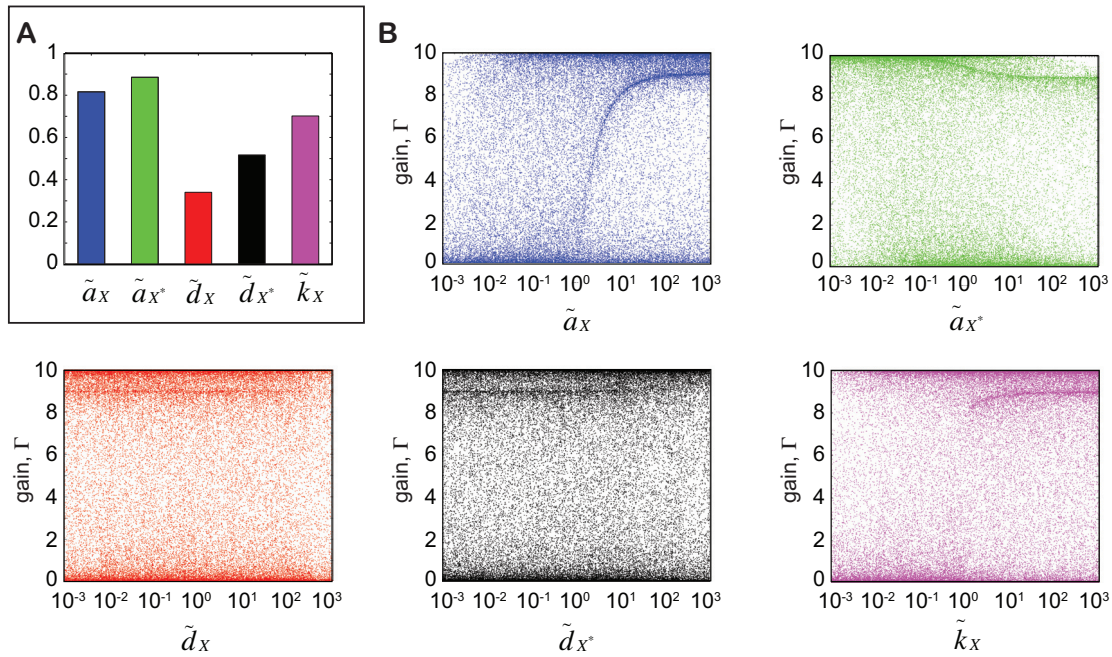


Figure 5.3: (A) Total sensitivity indices and (B) scatter plots for each parameter of monocyclic system, with respect to amplification. The concentration ratios are set to $\{\rho^{S/X}, \rho^{P_x/X}\} = \{0.1, 0.1\}$. The nominal parameter set $\{\tilde{a}_x, \tilde{a}_{x^*}, \tilde{d}_x, \tilde{d}_{x^*}, \tilde{k}_x\} = \{1, 1, 1, 1, 1\}$ determines the amplification gain $\Gamma = 4.8$.

5.2.3 Influence on rise time

Random sampling of parameter space revealed high variation in signaling rise time values, from 10^{-1} s up to 10^5 s (Fig 5.4B).

Parameters \tilde{a}_{x^*} and \tilde{d}_{x^*} are the most sensitive parameters with almost equal values followed by parameters \tilde{a}_x , \tilde{d}_x and \tilde{k}_x , which also have about the same values of total sensitivity indices. This kind of classification of parameters into two groups underlines that the kinetic rates involved in deactivating reactions are the ones that perturb the values of the rise time the most. On the other hand, the direction of parameter change that increase/decrease rise time can only be assessed based on statistical observation. According to that perspective, high values of \tilde{a}_x , \tilde{a}_{x^*} and \tilde{k}_x lead to a significant decrease in rise time.

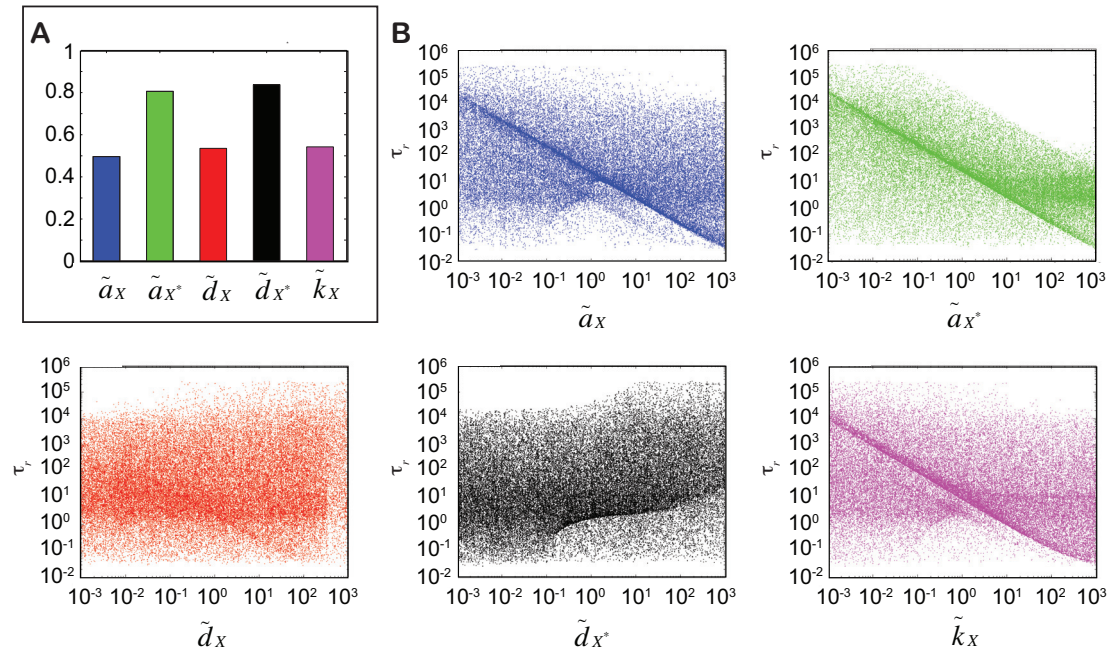


Figure 5.4: (A) Total sensitivity indices and (B) scatter plots for each parameter of monocyclic system, with respect to rise time. The concentration ratios are set to $\{\rho^{S/X}, \rho^{P_x/X}\} = \{0.1, 0.1\}$. The nominal parameter set $\{\tilde{a}_x, \tilde{a}_{x^*}, \tilde{d}_x, \tilde{d}_{x^*}, \tilde{k}_x\} = \{1, 1, 1, 1, 1\}$ determines the signaling rise time $\tau_r = 35.8$.

5.2.4 Influence on decay time

Fig. 5.5A clearly show that only parameters \tilde{a}_{x^*} and \tilde{d}_x influence the signaling decay time, with scatter plots on Fig. 5.5B confirming this observation.

Since the decay time is measured after signal withdrawal, it is not surprising that only the deactivating branch of the cycle plays a role in bringing the output of the system to the zero-state. In other words, the simulation for signaling decay time is equivalent to the simulation of the system that consist of only one set of Michaelis-Menten reactions. The above mentioned is still to be proven and compared for all the cases.

In the same way as for rise time, high values of \tilde{a}_{x^*} and low values of \tilde{d}_{x^*} restrict decay times to $< 10^0$.

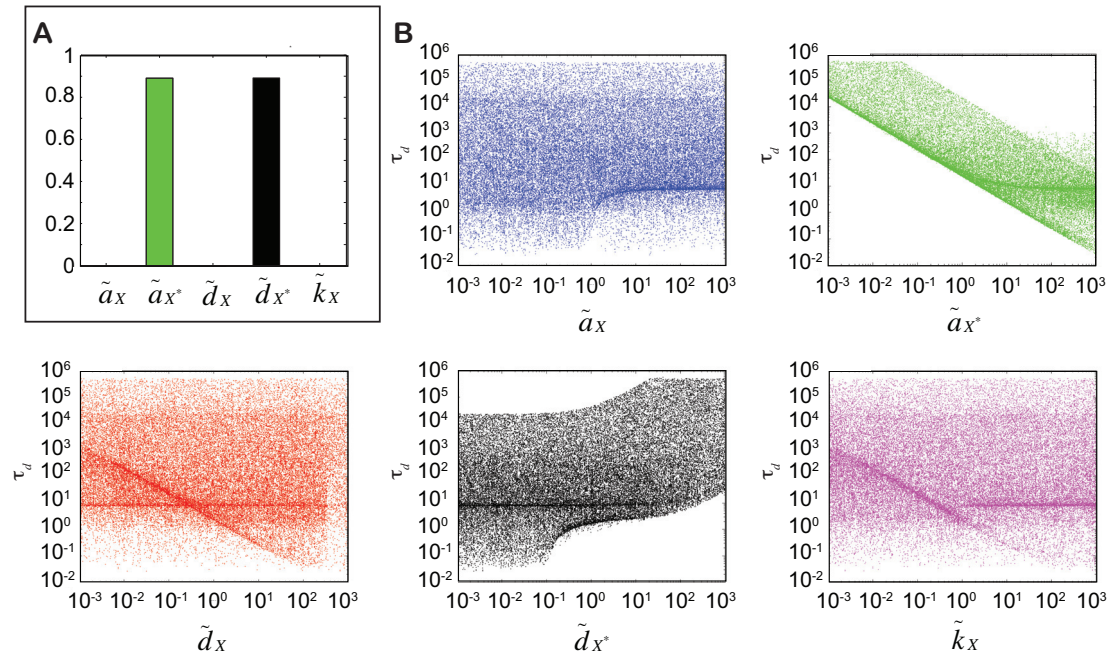


Figure 5.5: (A) Total sensitivity indices and (B) scatter plots for each parameter of monocyclic system, with respect to decay time. The concentration ratios are set to $\{\rho^{S/X}, \rho^{P_X/X}\} = \{0.1, 0.1\}$. The nominal parameter set $\{\tilde{a}_X, \tilde{a}_{X^*}, \tilde{d}_X, \tilde{d}_{X^*}, \tilde{k}_X\} = \{1, 1, 1, 1, 1\}$ determines the signaling decay time $\tau_d = 52.2$.

5.3 Sensitivity analysis of tricyclic system

The role of sensitivity analysis in the exploration of the models of signal transduction was quite recognized in recent years and MAPK cascade was part of these studies. Birtwistle and coworkers were analyzing ligand-dependent responses of the ErbB signaling network and their effects on ERK regulation (130). They performed linear sensitivity analysis by making a 1% change in each model parameter and looking at the fractional effect on each observable. Their analysis was analogous to calculating the control coefficients and response coefficients in metabolic control analysis. Zhang and Rundell were interested in the sensitivity to parameter variation in T-cell receptor-activated Erk-MAPK signaling pathway model (78). Here we perform the global sensitivity analysis on the isolated module of prototypical MAPK cascade. Total sensitivity indices for all the parameters and all four properties are illustrated in Fig. 5.6.

Defining that in nominal set of parameters all the values are fixed to 1 (as we have done in the previous subsection as well), leads to the values of Hill coefficient 1.1, amplification gain of 8.64, signaling rise time of 21.2, and signaling decay time of 232.13.

Computational cost of our analysis was approximately 10 times more expensive than it was the case for the global sensitivity analysis of monocyclic system. Except for the higher system complexity and larger number of parameters, this is mainly due to the fact that random choice of parameters would not lead to successful system simulation; either the system would not reach the steady state during the predefined simulation time, or the numerical issues arising from the stiffness of the system would be more promoted. More

than 10% of all simulations (combinations of parameters) needed to be discarded, as they were not contributing to the calculation of the sensitivity indices. Nevertheless, this “failed” set of parameters can potentially be used for further analysis, such as mapping the conditions under which the system becomes “broken” – an analogy with the diseased state of the cell. Since we aim to identify the conditions under which the system works “perfectly”, we are closing this subsection.

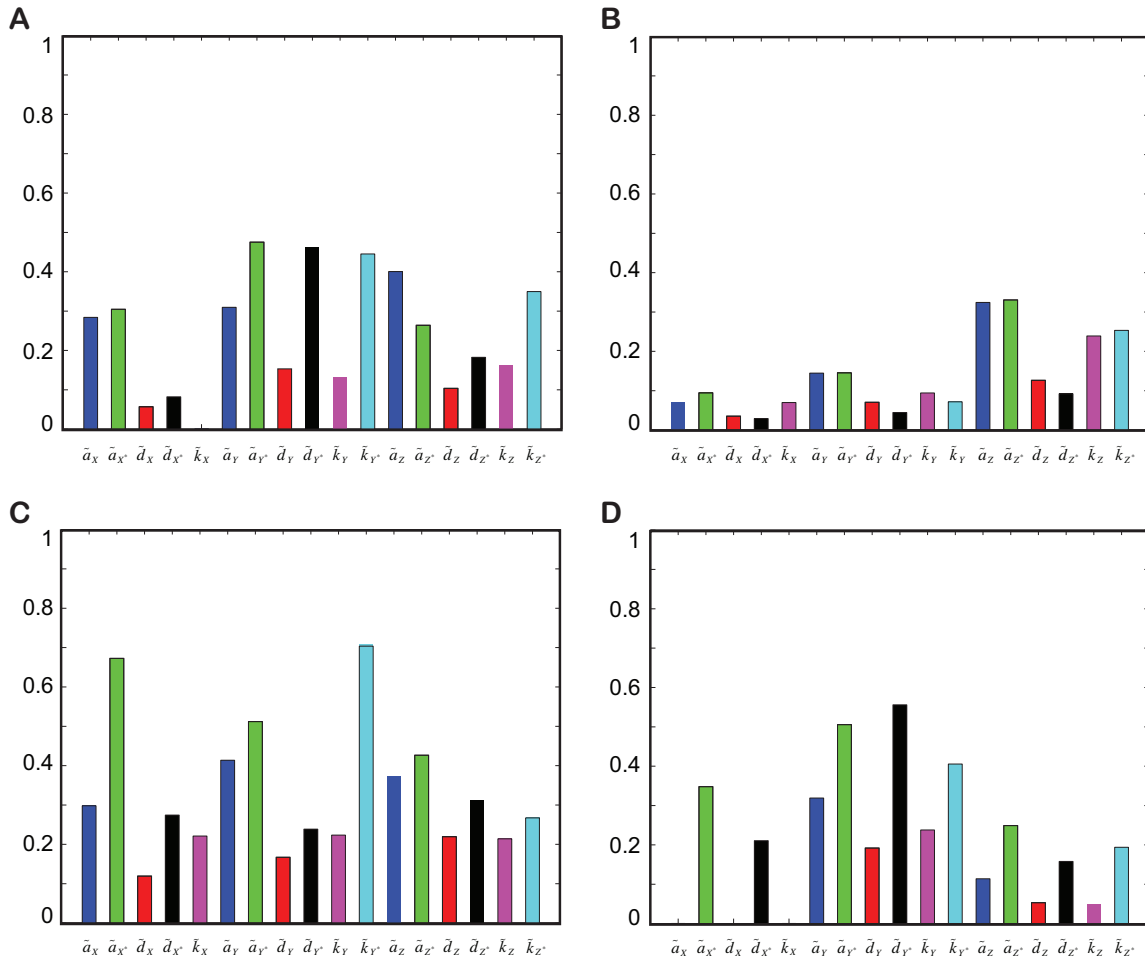


Figure 5.6: Total sensitivity indices for tricyclic system on (A) ultrasensitivity, (B) amplification gain, signaling (C) rise and (D) decay times.

5.3.1 Influence on ultrasensitivity

Having in mind the observation from the optimal design (Table 4.1), one would expect that the most sensitive part of the tricyclic model with respect to ultrasensitivity is the first level of the cascade, precisely the parameters that are lambded into the Michaelis-Menten constant \tilde{K}_X . Global sensitivity analysis does not underline this fact very clearly (Fig. 5.6A), but the scatter plots of the most influential parameters (Fig. 5.7) confirm the design rules for parameter values identified in the study of optimal ultrasensitive tricyclic systems. Indeed, higher values of parameters \tilde{a}_X and \tilde{a}_{X^*} (and lower values of \tilde{K}_X and \tilde{K}_{X^*}) lead the first level to the kinase saturation, and lower values of \tilde{a}_Y and \tilde{a}_Z (and higher values of \tilde{K}_Y and

\tilde{K}_Z) directly dictate that the second and the third level of the cascade are unsaturated with the signaling kinase. Another important observation that arises from global sensitivity analysis for ultrasensitivity is that the group of parameters directly related to Michaelis-Menten constant \tilde{K}_Y^* has the overall most important ranking. In chapter 4, this parameter was fixed to one value, which prevented the determination of its influence on the Hill coefficient.

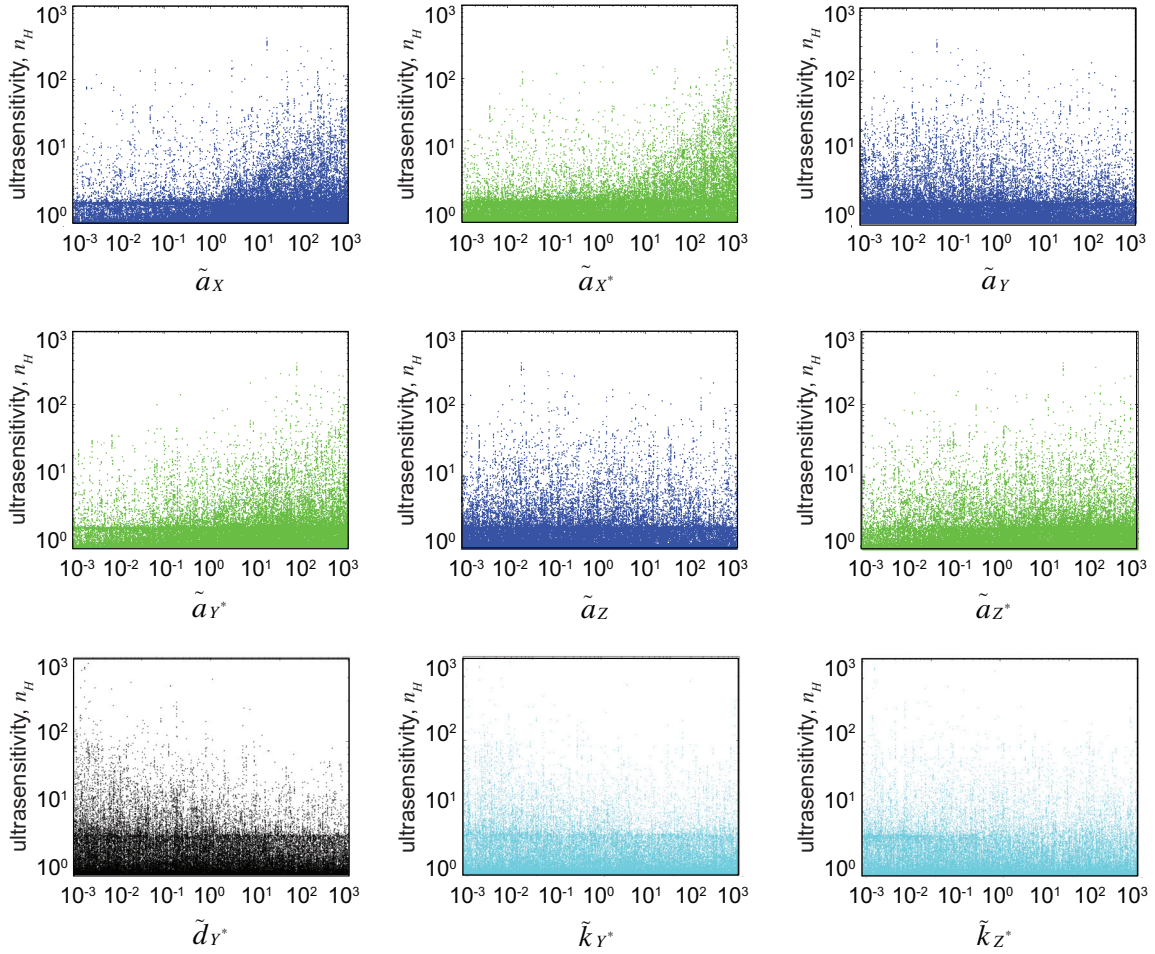


Figure 5.7: Scatter plots of the most influential parameters on ultrasensitivity of tricyclic system. The concentration ratios are set to $\{\rho^{S/X}, \rho^{X/Y}, \rho^{Y/Z}\} = \{0.1, 1, 1\}$ and $\{\rho^{P_X/X}, \rho^{P_Y/Y}, \rho^{P_Z/Z}\} = \{0.1, 0.1, 0.1\}$. The nominal values for all the parameters are set to 1, and such a set determines the Hill coefficient $n_H = 1.1$.

5.3.2 Influence on amplification

As introduced before, the amplification gain is the ratio of activated kinase concentration to input concentration at steady state. In the case of the tricyclic system, output of interest is the last activated kinase in cascade, z^* . Following the definition for the dimensionless form of amplification in monocyclic systems, we arrive to the one valid for tricyclic systems:

$$\Gamma := \frac{\bar{z}^*}{\rho^{S/X} \cdot \rho^{X/Y} \cdot \rho^{Y/Z}} \quad (5.2)$$

where \bar{z}^* stands for the steady-state kinase activity corresponding to a given level of signaling enzyme (Fig. 3.3B). In order to keep amplification gain comparable in both mono- and tricyclic systems, we fixed concentration ratios to the values: $\rho^{S/X} = 0.1$ and $\rho^{X/Y} = \rho^{Y/Z} = 1$. Thus, the highest level of amplification gain is kept to the value 10.

If we gather parameters into 3 groups according to the level of the cascade they belong, we see that amplification of the signaling kinase z^* is mostly influenced by the parameters directly related to its synthesis and degradation (Fig. 5.6B). The single sensitivity ranking of each parameter in all the groups remains fairly similar within each level. These rankings correspond to the ones revealed by global sensitivity analysis of monocyclic system (Fig. 5.3A). Furthermore, the sensitivity indices gradually increase in each subsequent level of the cascade. Important finding is that overall sensitivities are quite decreased, comparing tricyclic with monocyclic systems. This might be yet one more incentive for having multiple cycles versus one cycle as signaling mechanism, since the overall robustness of the system is largely improved.

Intermediate amplification gains are again less likely to happen, and parameters \tilde{a}_z and \tilde{a}_{z^*} determine the amplification level of z^* the most, followed by \tilde{k}_z and \tilde{k}_{z^*} .

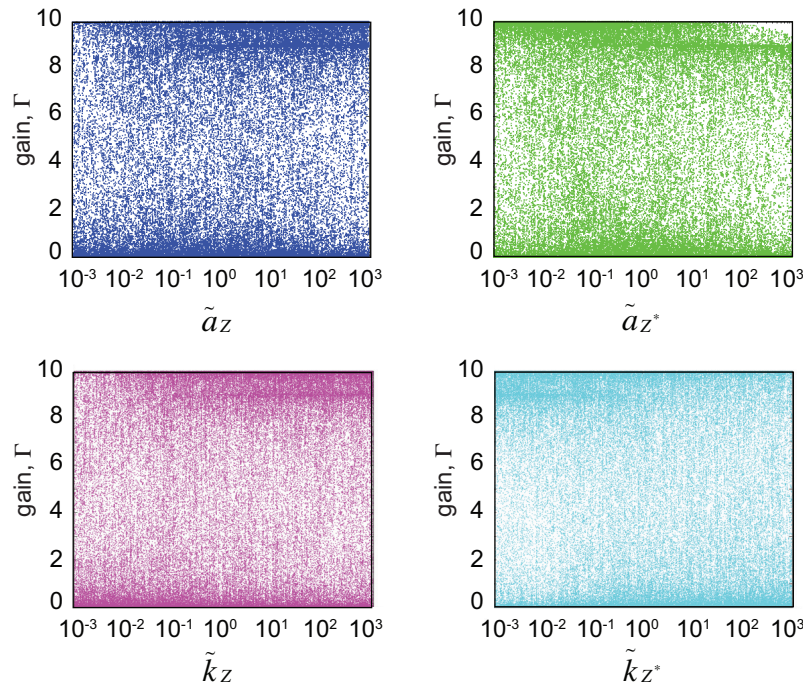


Figure 5.8: Scatter plots of the most influential parameters on amplification of tricyclic system. The concentration ratios are set to $\{\rho^{S/X}, \rho^{X/Y}, \rho^{Y/Z}\} = \{0.1, 1, 1\}$ and $\{\rho^{P_x/X}, \rho^{P_y/Y}, \rho^{P_z/Z}\} = \{0.1, 0.1, 0.1\}$. The nominal values for all the parameters are set to 1, and such a set determines the amplification gain $\Gamma = 8.64$.

5.3.3 Influence on rise time

The classification of the key parameters influencing the dynamics of the tricyclic system is not straightforward. Even though global sensitivity analysis pinpoints on parameters \tilde{a}_Y and \tilde{k}_Y as the most influential ones, there could be also other groups of equally important parameters for determining signaling rise time (Fig. 5.6C). We are underlining the six parameters that are most likely to contribute to the low values of signaling rise time in Fig. 5.9. Those are three pairs of corresponding \tilde{a} and \tilde{k} kinetic rates at each level of the cascade. Their high value is in a direct relation with fast signal propagation through the cascade, also suggesting that information flow is downstream.

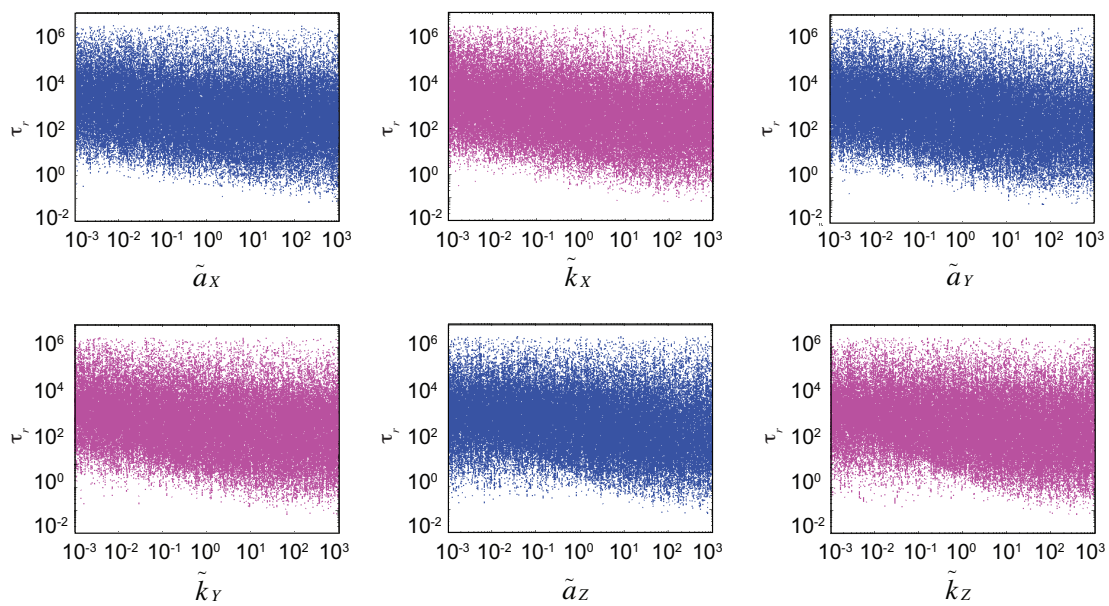


Figure 5.9: Scatter plots of the most influential parameters on signaling rise time of tricyclic system. The concentration ratios are set to $\{\rho^{S/X}, \rho^{X/Y}, \rho^{Y/Z}\} = \{0.1, 1, 1\}$ and $\{\rho^{P_X/X}, \rho^{P_Y/Y}, \rho^{P_Z/Z}\} = \{0.1, 0.1, 0.1\}$. The nominal values for all the parameters are set to 1, and such a set determines the signaling rise time $\tau_r = 21.2$.

5.3.4 Influence on decay time

It is striking that \tilde{a}_X , \tilde{d}_X and \tilde{k}_X have sensitivities almost equal to zero, as it is the case for monocyclic system (Fig. 5.6D). It is also obvious that parameters of the second cycle have much more influence on decay time of z^* than parameters of the third cycle, indicating on the gradual decrease of sensitivity by going down the cascade. We can observe that the group of kinetic rates from the lower branches of each cycle (deactivation of the kinase) is promoted with the highest sensitivity indices.

Scatter plots in Fig. 5.10 display the direct dependence of short signaling decay time

with respect to high values of parameters \tilde{a}_* of each level in the cascade. Introducing more levels in the system architecture, parameters \tilde{k}_{Y^*} and \tilde{k}_{Z^*} come forward as the influential ones and their high value promote fast signal decay. Since these two parameters are directly leading to accumulation of inactive form of the apparent kinase, the question whether the same pattern would occur also in the first level of the cascade or isolated monocyclic system if the parameter \tilde{k}_{X^*} were the part of the study, still remains open.

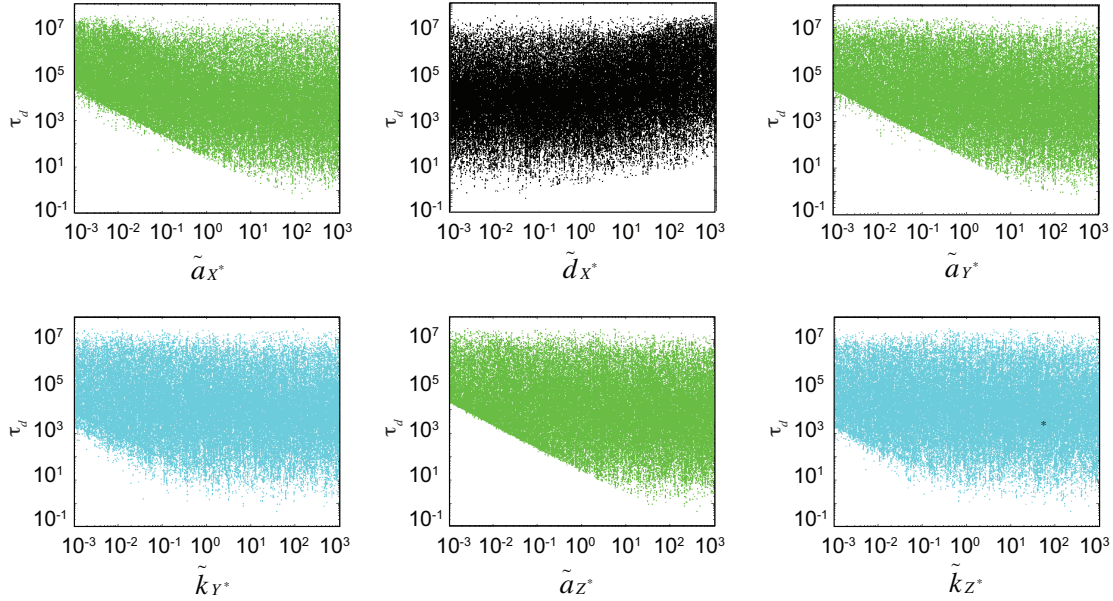


Figure 5.10: Scatter plots of the most influential parameters on signaling decay time of tricyclic system. The concentration ratios are set to $\{\rho^{S/X}, \rho^{X/Y}, \rho^{Y/Z}\} = \{0.1, 1, 1\}$ and $\{\rho^{P_x/X}, \rho^{P_y/Y}, \rho^{P_z/Z}\} = \{0.1, 0.1, 0.1\}$. The nominal values for all the parameters are set to 1, and such a set determines the signaling decay time $\tau_d = 232.13$.

5.4 Sensitivity analysis as prelude for dynamic optimal design

Table 5.1 summarizes the main results for global sensitivity analysis of monocyclic and tricyclic system. The influence of at least two key parameters on each property is indicated with the arrow, showing the direction of the change of parameter value that would lead to optimal value of the property.

Red arrow in increasing direction suggests high values of parameters and blue arrow in decreasing direction suggests low values of parameters in order to achieve desired system performance: high ultrasensitivity, high amplification gain and low signaling rise and decay times. With these results in hand, we are ready to proceed to more challenging tasks – identifying design principles and parameter values that ensure the optimal performance.

		n_H	Γ	τ_r	τ_d
Monocyclic system	\tilde{a}_X	↗	↗	↗	
	\tilde{a}_{X^*}	↗	↙	↗	↗
	\tilde{d}_X	↙			
	\tilde{d}_{X^*}			↙	↙
	\tilde{k}_X		↗	↗	
Tricyclic system	\tilde{a}_X	↗		↗	
	\tilde{a}_{X^*}	↗			↗
	\tilde{d}_X				
	\tilde{d}_{X^*}			↗	↙
	\tilde{k}_X			↗	
	\tilde{a}_Y	↙		↗	
	\tilde{a}_{Y^*}	↗			↗
	\tilde{d}_Y				
	\tilde{d}_{Y^*}				
	\tilde{k}_Y			↗	
	\tilde{k}_{Y^*}				↗
	\tilde{a}_Z	↙	↗		
	\tilde{a}_{Z^*}	↗	↙		↗
	\tilde{d}_Z				
	\tilde{d}_{Z^*}				
	\tilde{k}_Z		↗	↗	
	\tilde{k}_{Z^*}		↙		↗

Table 5.1: Modular differences and similarities: global sensitivity analysis of monocyclic and tricyclic systems. Red arrows represent the value on the upper bound and blue arrows represent the value on the lower bound of variation range.

The Perfect Choreography

Optimal Design for Dynamic Properties

6.1 Optimal recipes for signaling times

The optimal performance of a signaling network or its integral part mostly depends on a particular biological context. Typically, it is determined by a relationship between the size of the network, the amplitude of the signal and its duration. Heinrich *et al.* posed an interesting question of whether signaling cascades can respond with sharp signals, i.e., signals with short duration and high amplitude (112). Their model of interest included simple linear signaling cascade and neglected the formation of intermediate complexes. Chaves and coworkers further expanded the seminal study and demonstrated mathematical conditions for sharp signals (67).

In this chapter, we first show that a simple system, consisting only of covalent modification cycle, can simultaneously achieve the mentioned desired performance. Furthermore, we identify the rules that lead a tricyclic cascade towards the system that provides sharp signals.

6.1.1 Optimal design for signaling times under amplification constraints in monocyclic system

In this subsection, we follow the inverse approach and formulate an optimization problem to identify the kinetic parameter values that minimize the response time, while at the same time satisfying a given amplification level Γ . Separate formulations are considered for the rise time τ_r (P₂) and for the decay time τ_d (P₃),

These optimization problems are formulated as follows:

$$\begin{array}{ll}
 \text{find} & \tilde{a}_X, \tilde{a}_{X^*}, \tilde{d}_X, \tilde{d}_{X^*}, \tilde{k}_X \quad \text{that} \\
 \text{minimize} & \text{rise time } \tau_r, \text{ Eq. (3.28)} \\
 \text{subject to} & \text{amplification gain } \Gamma, \text{ Eq. (5.1)} \\
 & \text{transient model, Eq. (3.8–3.14)}
 \end{array} \tag{P_3}$$

and

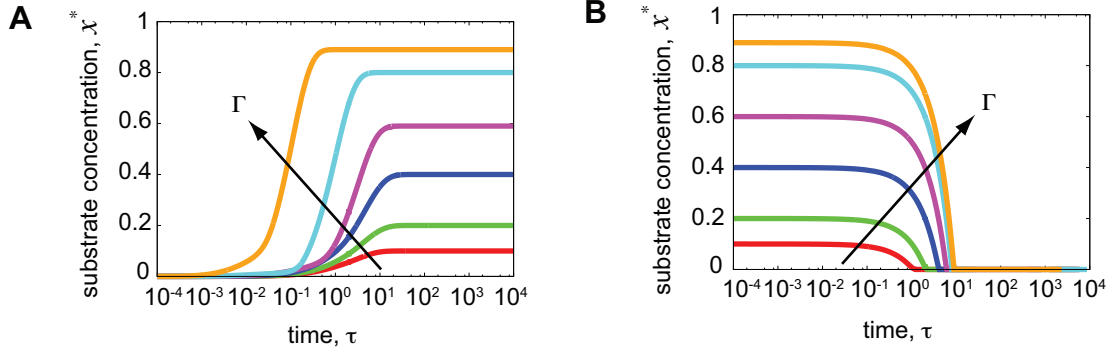


Figure 6.1: Dynamic responses of monocyclic system for different levels of amplification. (A) Positive and (B) negative step function as the input. Fixed amplification gain takes values from $\Gamma \in \{1, 2, 4, 6, 8, 9\}$ respectively (direction of the arrows).

$$\begin{aligned}
 &\text{find} && \tilde{a}_X, \tilde{a}_{X^*}, \tilde{d}_X, \tilde{d}_{X^*}, \tilde{k}_X && \text{that} && (P_4) \\
 &\text{minimize} && \text{decay time } \tau_d, \text{ Eq. (3.29)} \\
 &\text{subject to} && \text{amplification gain } \Gamma, \text{ Eq. (5.1)} \\
 &&& \text{transient model, Eq. (3.8–3.10, 3.33, 3.12–3.14).}
 \end{aligned}$$

The model parameters $\tilde{a}_X, \tilde{a}_{X^*}, \tilde{d}_X, \tilde{d}_{X^*}, \tilde{k}_X$ are all taken in the range $[10^{-3}, 10^3]$ subsequently. Recall that the transient model was made dimensionless, and so a kinetic parameter whose value is either at its lower bound 10^{-3} or at its upper bound 10^3 reflects a very small or very large value *relative to* other parameters. To confirm that this parameter range is large enough to reveal the actual set of behaviors, we performed similar computations with wider parameter ranges as $[10^{-5}, 10^5]$ (data not shown). It was found that the general trend is conserved in that the parameters that were at their lower/upper bounds remained at their bounds, and those taking intermediate values too remained intermediate; moreover, optimization with wider parameter ranges only leads to marginal improvements in signaling times. We also performed additional computations with more narrow parameter ranges as $[10^{-2}, 10^2]$, which again led to essentially the same trade-offs and minor differences in signaling times.

The optimization routine was repeated for a range of amplification gain Γ and for two sets of concentration ratios: $\rho^{S/X} = \rho^{P/X} = 10^{-2}$ and $\rho^{S/X} = \rho^{P/X} = 10^{-1}$.

Mathematically, these two optimization problems read:

$$\begin{aligned}
 &\min_{\mathbf{p}, \tau_r} && \tau_r && (P_2) \\
 &\text{subject to} && \mathbf{F}_r(\xi(\tau), \dot{\xi}(\tau), \mathbf{p}, \mathbf{r}) = \mathbf{0}, \quad 0 \leq \tau \leq \tau^\infty, \quad \xi(0) = \xi_{r_0}, \\
 &&& x^*(\tau_r) - 0.9x^*(\tau^\infty) = 0, \\
 &&& \Gamma \rho^{S/X} - x^*(\tau^\infty) = 0, \\
 &&& 10^{-3} \leq \mathbf{p} \leq 10^3,
 \end{aligned}$$

and

$$\begin{aligned}
 &\min_{\mathbf{p}, \tau_d} && \tau_d && (P_3) \\
 &\text{subject to} && \mathbf{F}_d(\xi(\tau), \dot{\xi}(\tau), \mathbf{p}, \mathbf{r}) = \mathbf{0}, \quad 0 \leq \tau \leq \tau^\infty, \quad \xi(0) = \xi_{d_0}, \\
 &&& x^*(\tau_d) - 0.1x^*(\tau^\infty) = 0,
 \end{aligned}$$

$$\Gamma \rho^{S/X} - x^*(0) = 0,$$

$$10^{-3} \leq \mathbf{p} \leq 10^3.$$

Fig. 6.2 displays the optimal signaling times as a function of the amplification gain Γ , together with the optimal kinetic parameters $\tilde{a}_x, \tilde{a}_{x^*}, \tilde{d}_x, \tilde{d}_{x^*}, \tilde{k}_x$. Figs. 6.2A and 6.2C correspond to operating regime $\rho^{S/X} = \rho^{P/X} = 10^{-2}$ and Figs. 6.2B and 6.2D to operating regime $\rho^{S/X} = \rho^{P/X} = 10^{-1}$.

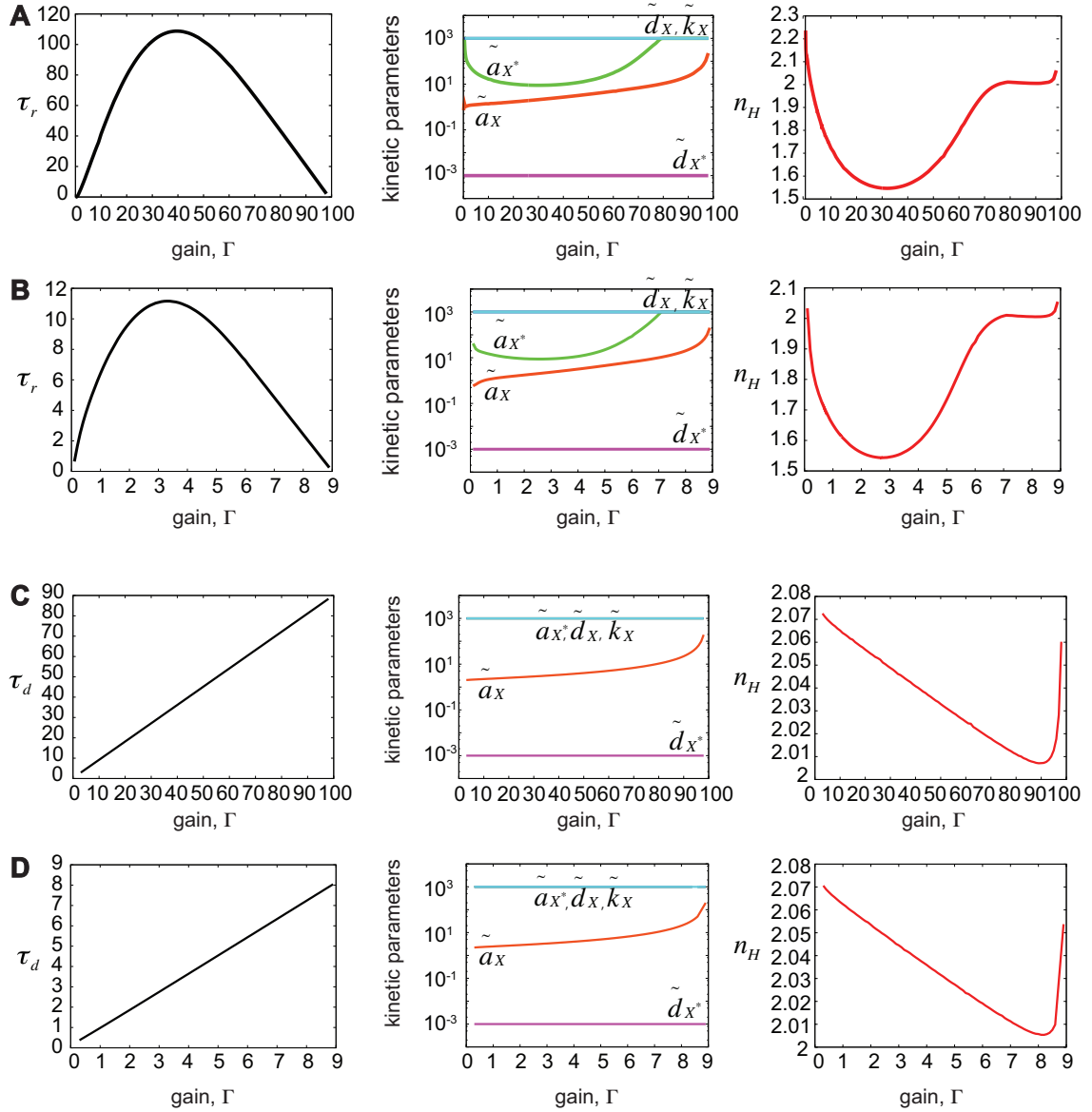


Figure 6.2: Design for minimal signaling rise time τ_r (A and B) and decay time τ_d of monocyclic system (C and D), subject to given amplification. Optimal signal propagation times (*left plots*), corresponding kinetic parameters $\tilde{a}_x, \tilde{a}_{x^*}, \tilde{d}_x, \tilde{d}_{x^*}, \tilde{k}_x$ (*middle plots*) and calculated Hill coefficient n_H (*right plots*) versus amplification gain Γ , for the concentration ratio regime $\rho^{S/X} = \rho^{P/X} = 10^{-2}$ (A and C) and concentration ratio regime $\rho^{S/X} = \rho^{P/X} = 10^{-1}$ (B and D).

Qualitatively, the optimization results seem to be very similar in both regimes. The feasible range of amplification gains is much wider with the lower values of concentration ratios – as expected from the theoretical analysis, Eq. (5.1). The rise time τ_r (Figs. 6.2A and 6.2B) first increases monotonically with Γ for low amplification. Beyond a certain amplification threshold, the rise time decreases monotonically with Γ , thus indicating that the system can respond faster while achieving higher levels of amplification. This amplification threshold depends on the concentration ratios $\rho^{S/X}$ and $\rho^{P/X}$; it is close to $\Gamma = 40$ in the operating regime $\rho^{S/X} = \rho^{P/X} = 10^{-2}$ (Fig. 6.2A) and around $\Gamma \approx 3.3$ in the regime $\rho^{S/X} = \rho^{P/X} = 10^{-1}$ (Fig. 6.2B). On the other hand, the decay time τ_d (Figs. 6.2C and 6.2D) increases linearly with the amplification gain Γ . This is due to the fact that higher amplification leads to higher concentrations in activated protein form, which in turn takes more time to return to inactive state. Moreover, higher concentration ratios $\rho^{S/X}$ and $\rho^{P/X}$ lead to lower maximal possible levels of free active protein and, consequently, decay times are almost 10 times shorter in the regime $\rho^{S/X} = \rho^{P/X} = 10^{-1}$ compared to the regime $\rho^{S/X} = \rho^{P/X} = 10^{-2}$ (Figs. 6.2B and 6.2D vs. Figs. 6.2A and 6.2C).

Non-monotonic correlations between various objectives demonstrate the usefulness of optimization methods to study signaling pathways in a systematic way. In order to better understand these relationships, we investigate the combination of parameters that lead to the minimum signaling times using optimization methodology. The results indicate that the optimal values for certain parameters are on the extreme bounds of the variation intervals for all values of the amplification gain, while the optimal values vary with Γ for others.

Both dissociation rate constants \tilde{d}_X and \tilde{k}_X of the kinase complex $X:S^*$ into x and x^* , respectively, stay at their upper bounds and the rate constant \tilde{d}_{X^*} of the phosphatase complex $X^*:P_X$ into x^* remains at its lower bound in all cases and regardless of the amplification level Γ . Particularly counter-intuitive is the finding that minimum response times are *not* achieved when the rate of formation \tilde{a}_X of the kinase complex $X:S^*$ is maximum, which suggests that faster signal propagation with amplification is promoted by a more unstable complex $X:S^*$. These findings are in good agreement with the computational results for a covalent modification cycle obtained in (69). It is also observed that higher levels of amplification together with shorter rise and decay times are achieved for an increasing value of the complex formation rate constant \tilde{a}_X , which suggests that \tilde{a}_X is the primary determining parameter for minimal signaling times under amplification constraints.

For sure, the choice of parameters would be determined by the relative importance of the different objectives, as they have evolved in the system or designed in the context of synthetic biology.

If we now compare our results from previous chapter, we see that sensitivity analysis nicely predicted some of the parameter regimes in order to have minimal signaling rise and decay times (Figs. 5.4 and 5.5). Even though we now combine signaling times with the certain amplification level, there seems to be a correlation of design rules for amplification on its own, with the cases when also other properties are taken into account. Namely, the kinetic rates \tilde{a}_X and \tilde{a}_{X^*} are identified as the ones that have the biggest influence on amplification gain (Fig. 5.3). We hypothesize that the change of these two parameters in optimal design for signaling times is due to the change of amplification gain level, whereas the other parameters make sure that system transmits input signal very fast. This is done by

having the minimal values of kinetic rate \tilde{d}_{X^*} and maximal values of kinetic rates \tilde{d}_X and \tilde{k}_X . The same emerged from Fig. 5.4 and Fig. 5.5.

Nevertheless, monocyclic cascades designed to achieve minimal signaling times do not promote high ultrasensitivity (Fig. 6.2 *right plots*). In these studies, the constraints imposed on the kinetic parameters correspond to profiles of \tilde{K}_X and \tilde{K}_{X^*} that can lead to maximal $n_H \approx 2$. In the case of rise time (Figs. 6.2A and 6.2B, *right plots*), any loss of ultrasensitivity occurs in the range of amplification where τ_r is increasing. Hence, it is important to take the signaling times and ultrasensitivity together into account when investigating the design of signaling modules in cells.

6.1.2 Optimal design for signaling times under amplification constraints in multicyclic systems

Using the same framework, we further analyze the optimal design for signaling times in multicyclic systems. The optimization problem that aims at minimizing the rise and the decay times in tricyclic (but also in bicyclic) systems, significantly grows in complexity and brings many numerical issues. In this case, the decision variables set consists of 17 members, which are presumably varied in the large range $[10^{-3}, 10^3]$. Multi-modality becomes much more obvious and it increases computational cost dramatically. In order to overcome some of these issues, we first solved the problems within the ranges $[10^{-1}, 10^1]$ and $[10^{-2}, 10^2]$. Such an approach allowed to classify the parameters into three groups: i) the ones that stay at the lower bound of the interval, ii) the ones that take intermediate values and iii) the ones that stay at the upper bound of the varying interval. Interestingly, the patterns were consistent regardless of the range. As expected, the optimal values of signaling times were smaller when considering larger range. In the case of monocyclic system, going from the range $[10^{-2}, 10^2]$ to $[10^{-3}, 10^3]$ did not improve the overall quality of signaling times and in the case of bicyclic and tricyclic systems the improvement was for an order of magnitude (data not shown).

Solving the problem (P₃) for the minimization of rise time, subject to an appropriate dynamic system representation (bicyclic and tricyclic systems) and amplification gain, resulted into the design rules presented in Fig. 6.3. As it was the case in chapter 5, the concentration ratios are fixed to the values: $\rho^{S/X} = 0.1$ and $\rho^{X/Y} = \rho^{Y/Z} = 1$. The first striking observations are the quantitative and qualitative improvements of the optimal values of signaling rise time in tricyclic system, with respect to the different levels of amplification gain. Parabolic (concave) dependence is flattened out and the rise times are more than 300 times lower, compared to the rise time of monocyclic system. This is yet one more advantage of the naturally occurring

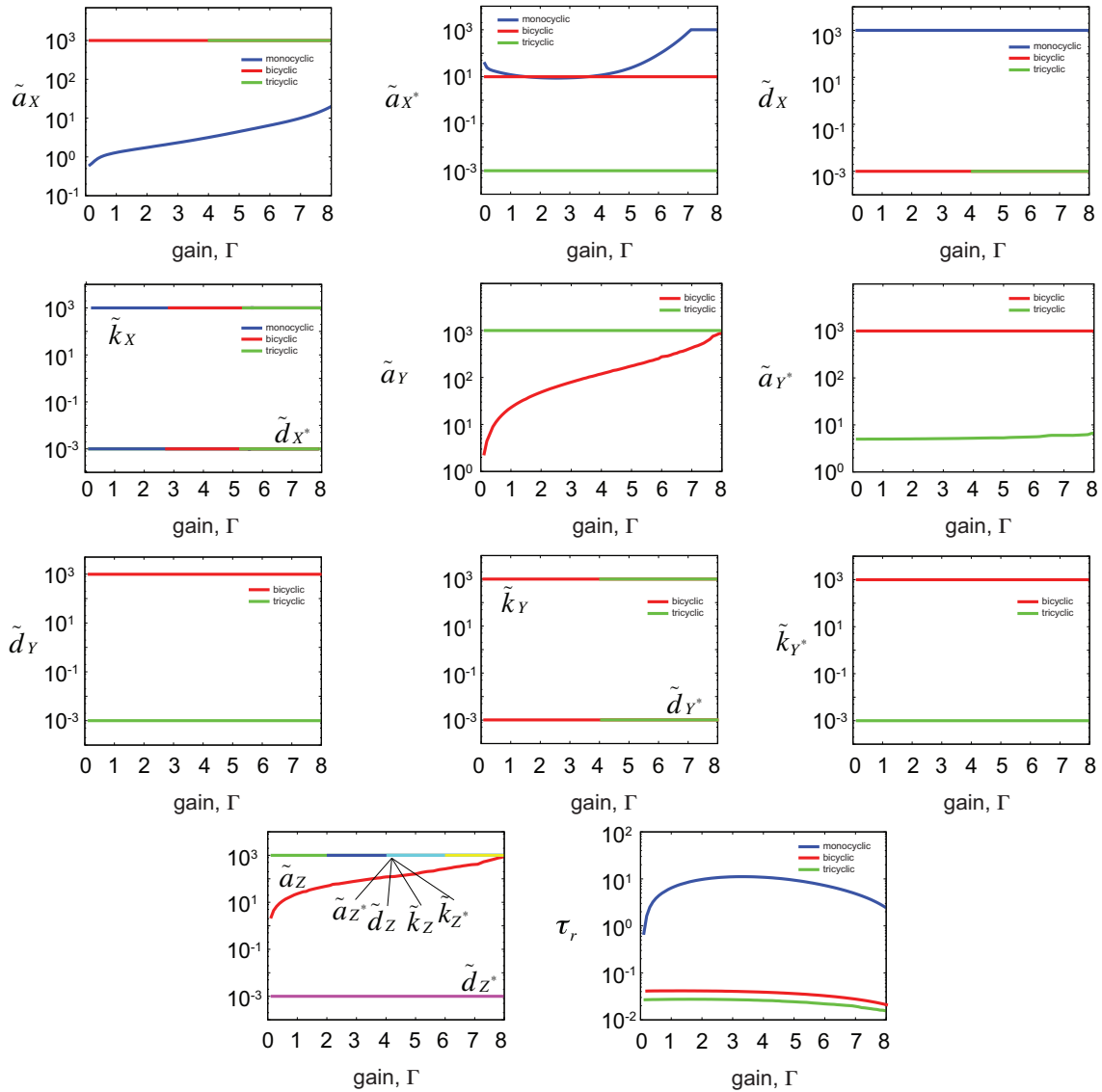


Figure 6.3: Design for minimal signaling rise time τ_r of monocyclic, bicyclic and tricyclic systems subject to given amplification.

architecture through the cascades. Therefore, not only that the response will be faster in multicyclic systems, but also all amplification levels could be reached without losing the transduction velocity.

The discussion about the kinetic rates, comparing differences and similarities for parameter values across different system architectures, is summarized in Table 6.1. This comparison is based on a qualitative analysis, whereas Fig. 6.3 summarizes the results of a quantitative analysis. It is interesting to notice how the same patterns are preserved regardless of whether the kinetic rates are compared horizontally (across the three systems) or vertically (in each system separately, but across their different levels). For example, the kinetic rates associated with the formation of the complex between the last inactive kinase in the cascade and active kinase from the previous cascade level (or in the case of monocyclic systems, the complex between the inactive kinase and signaling enzyme), namely \tilde{a}_x in monocyclic, \tilde{a}_y in bicyclic and \tilde{a}_z in tricyclic system,

	Monocyclic system	Bicyclic system	Tricyclic system
\tilde{a}_X	...	↗	↗
\tilde{a}_X^*	↙
\tilde{d}_X	↗	↙	↙
\tilde{d}_X^*	↙	↙	↙
\tilde{k}_X	↗	↗	↗
\tilde{a}_Y		...	↗
\tilde{a}_Y^*		↗	...
\tilde{d}_Y		↗	↙
\tilde{d}_Y^*		↙	↙
\tilde{k}_Y		↗	↗
\tilde{k}_Y^*		↗	↙
\tilde{a}_Z			...
\tilde{a}_Z^*			↗
\tilde{d}_Z			↗
\tilde{d}_Z^*			↙
\tilde{k}_Z			↗
\tilde{k}_Z^*			↗

Table 6.1: Modular differences and similarities: optimal design for signaling rise time of monocyclic, bicyclic and tricyclic systems. Red arrow represents the value on the upper bound of variation range, the blue arrow represents the value on the lower bound of variation range and three dots represent intermediate values.

always have the intermediate values. Beside these parameters, the kinetic rates associated with the degradation of the corresponding complexes in the second to last cascade level, \tilde{d}_X in bicyclic and \tilde{d}_Y in tricyclic system are at the upper bound. This again suggests the instability of these complexes. However, the corresponding kinetic rates that degrade the active form of the kinase in the same levels, \tilde{a}_X^* in bicyclic and \tilde{a}_Y^* in tricyclic system are somewhere in between the bounds of variation interval. Consequently, we can assume that the two groups of parameters, that are on neither bound, are responsible to determine different values of the amplification gain and the rest are ensuring fast signal propagation, as we have hypothesized in the previous subsection for monocyclic system. Fast signal propagation is occurring upon maximizing the forward reaction rates of the upper branch of each cycle – the ones that convert kinases to their active forms – and minimizing all the

other rates. This finding also suggests the one-directional information flow through the cascade system.

6.2 The trade-offs and interplay

We finally arrive to the central question of this thesis: are there any trade-offs between different system properties? Can signaling modules respond fast to the input signals, have high amplification and high ultrasensitivity at the same time? Answers to these questions will arise after solving optimization and multi-objective optimization problems in the following subsections.

6.2.1 Optimal design for signaling times under amplification and ultrasensitivity constraints in monocyclic system

The ability of monocyclic cascades to achieve a high Hill coefficient for small values of the Michaelis-Menten constants is one of the basic findings of (44). However, under these same conditions the system might exhibit excessively long response times, as well as poor amplification capabilities. We investigate next whether a simple covalent modification cycle can achieve fast signaling, high amplification and high ultrasensitivity simultaneously.

Previous analysis has underlined that ultrasensitivity strongly depends on the two parameters \tilde{K}_X and \tilde{K}_{X^*} , which themselves are functions of five parameters of the dynamic model (Eq. (3.22–3.23)).

The incorporation of ultrasensitivity objective together with response time and amplification objectives can be done in either one of the two ways:

1. Optimize the steady-state and transient kinetic parameters jointly and enforce the constraints Eq. (3.22) and Eq. (3.23) directly. This requires accounting for both the transient model Eq. (3.8–3.14) and the reformulated steady-state model Eq. (4.1–4.6) in the optimization problem;
2. Optimize the transient kinetic parameters only, while enforcing an ultrasensitivity target indirectly via constraining the values of Michaelis-Menten constants to fixed values $\tilde{\kappa}_X, \tilde{\kappa}_{X^*}$ as:

$$\tilde{d}_X + \tilde{k}_X - \tilde{\kappa}_X \tilde{a}_X = 0 \quad (6.1)$$

$$\tilde{d}_{X^*} + 1 - \tilde{\kappa}_{X^*} \tilde{a}_{X^*} = 0. \quad (6.2)$$

The latter approach is considered next. It does not require the reformulated steady-state model Eq. (4.1–4.6) in the optimization problem, and the constraints Eq. (6.1–6.2) are linear. Note also that there remains flexibility in the choice of the kinetic parameters despite these two constraints (3 remaining degrees of freedom out of 5). In particular, the same dynamic optimization problems as before can be used, with the additional two constraints:

$$\begin{array}{ll} \text{find} & \tilde{a}_X, \tilde{a}_{X^*}, \tilde{d}_X, \tilde{d}_{X^*}, \tilde{k}_X \quad \text{that} \\ \text{minimize} & \text{rise time } \tau_r, \text{ Eq. (3.28)} \\ \text{subject to} & \text{amplification gain } \Gamma, \text{ Eq. (5.1)} \end{array} \quad (P_5)$$

target n_H , Eq. (6.1–6.2)
 transient model, Eq. (3.8–3.14)

and

find $\tilde{a}_X, \tilde{a}_{X^*}, \tilde{d}_X, \tilde{d}_{X^*}, \tilde{k}_X$ (P₆)
 minimize decay time τ_d , Eq. (3.29)
 subject to amplification gain Γ , Eq. (5.1)
 target n_H , Eq. (6.1–6.2)
 transient model. Eq. (3.8–3.10, 3.33, 3.12–3.14).

The Michaelis-Menten constant $\tilde{\kappa}_{X^*}$ is set to 10^{-3} and the value of $\tilde{\kappa}_X$ is chosen in order to meet a desired ultrasensitivity target $n_H \in \{3, 4, 6, 8, 10, 12, 16\}$ (Fig. 6.4). The influence of the amplification and ultrasensitivity constraints on the minimal signaling times is considered in the operating regime $\rho^{S/X} = \rho^{P/X} = 10^{-1}$ and for kinetic parameters $\tilde{a}_X, \tilde{a}_{X^*}, \tilde{d}_X, \tilde{d}_{X^*}, \tilde{k}_X$ varying in the range $[10^{-3}, 10^3]$ (Fig. 6.5). From the section 6.1.1, we have learned that operating in different regimes of concentration ratio space does not give additional information about design principles. Accordingly, the following studies will concern one isolated combination of concentration ratios, which will be representative enough to capture the systems behavior.

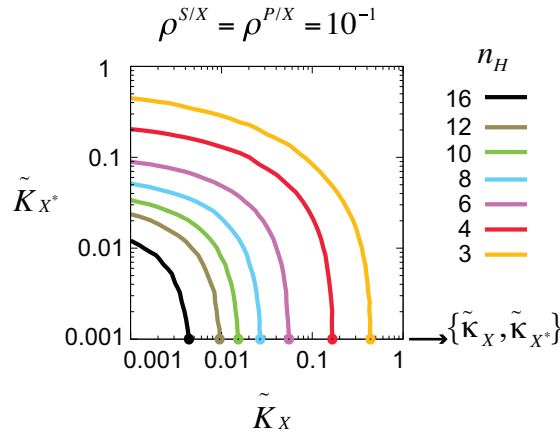


Figure 6.4: Optimal design for ultrasensitivity in monocyclic system. Contour graphs of Hill coefficient values n_H versus dimensionless Michaelis-Menten constants \tilde{K}_X and \tilde{K}_{X^*} .

The optimal rise time τ_r exhibits a non-monotonic relationship with respect to the amplification gain and it increases with the Hill coefficient target for a constant gain. In contrast, the ultrasensitivity requirement has a limited effect on the design for minimal decay time τ_d . The kinetic parameters \tilde{a}_X and \tilde{a}_{X^*} stay at their upper bounds and \tilde{d}_{X^*} at its lower bound in all cases (Fig. 6.5). On the other hand, the optimal values of the dissociation rate constant \tilde{d}_X of the kinase complex $X:S^*$ decrease significantly with increasing n_H , unlike those of the dissociation rate constant \tilde{k}_X .

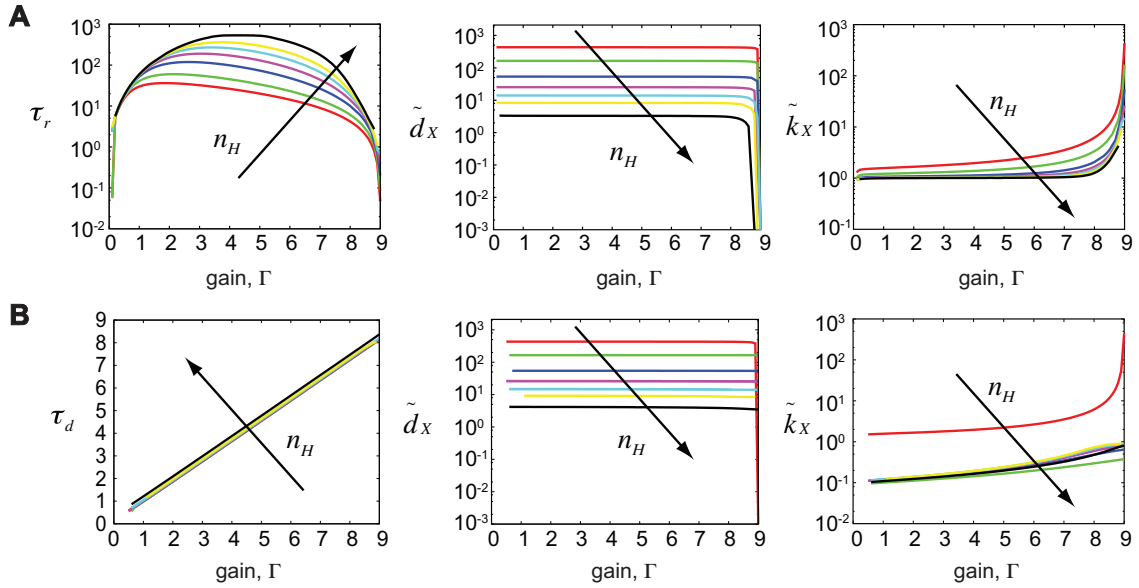


Figure 6.5: Design for minimal signaling rise time τ_r (upper plots) and decay time τ_d (lower plots) of monocyclic system, subject to given amplification and ultrasensitivity. The amplification gain Γ is varied in the range $[0, 9]$ and the Hill coefficient is chosen in the set $n_H \in \{3, 4, 6, 8, 10, 12, 16\}$. The concentration ratios are $\rho^{S/X} = \rho^{P/X} = 10^{-1}$.

This is attributed to the fact that high ultrasensitivity requires small values of the Michaelis-Menten constant \tilde{K}_X (since the value of Michaelis-Menten constant \tilde{K}_{X^*} is fixed to 10^{-3}), which in turn requires that $\tilde{d}_X \ll a_X - \tilde{k}_X$ (see Eq. (3.22)).

Perhaps the most striking finding from this inverse sensitivity analysis is that simple covalent modification cycles can be designed in such a way that they achieve high amplification and high ultrasensitivity, along with relatively short signaling times, on the order of 10 to 100 times the characteristic time $(\tilde{k}_X)^{-1}$ of the dissociation of the $X^* : P_X$ complex. It has often been postulated that multiple cascades are needed to achieve multiple objectives. Interestingly, our results show that even a single interconvertible cycle is in fact capable of meeting several goals simultaneously.

6.2.2 Optimal design for signaling times under amplification and ultrasensitivity constraints in multicyclic systems

Since there are no conceptual barriers for applying the same methodology to multicyclic systems, we investigated next how they satisfy multi-objective performance. As an illustrative example on how the optimization framework allows comfortable environment for setting different research questions, we formulate the problem in a slightly different manner:

find kinetic rates that (P₇)
 minimize sum of the rise time and decay time, Eqs. (3.28–3.29)
 subject to amplification gain Γ
 target n_H , Eqs. (6.1–6.2)
 transient model.

Here, we gather both signaling times and we minimize their sum. We have seen in the case of monocyclic system that different rules should be applied to minimize each of these two properties. In any case, the motivation to explore the design patterns of the systems that simultaneously meet optimal performance for different properties makes this formulation legitimate. Alternatively, one can assign different weight factors to each signaling time and thus underline the non-equivalent importance of each objective, if needed.

To be able to compare the performance and design principles across different system architectures, we use semi-quantitative description of Michaelis-Menten constants and incorporate them directly to Eqs. (6.1–6.2). We follow the rule: the first kinase is saturated and the rest of them are unsaturated by their signaling kinases. This arrangement gives the maximal Hill coefficient values of ~ 2.87 in monocyclic, ~ 5 in bicyclic and ~ 7.15 in tricyclic system.

For example, the optimization problem for tricyclic system in a mathematical form reads:

$$\begin{aligned}
& \min_{\mathbf{p}, \tau_r, \tau_d} && \tau_r + \tau_d && (P_7) \\
& \text{subject to} && \mathbf{F}_r(\xi(\tau), \dot{\xi}(\tau), \mathbf{p}, \mathbf{r}) = \mathbf{0}, \quad 0 \leq \tau \leq \tau^\infty, \quad \xi(0) = \xi_{r_0}, \\
& && z^*(\tau_r) - 0.9z^*(\tau^\infty) = 0, \\
& && \mathbf{F}_d(\xi(\tau), \dot{\xi}(\tau), \mathbf{p}, \mathbf{r}) = \mathbf{0}, \quad 0 \leq \tau \leq \tau^\infty, \quad \xi(0) = \xi_{d_0}, \\
& && z^*(\tau_r) - 0.9z^*(\tau^\infty) = 0, \\
& && \Gamma \rho^{S/X} \rho^{X/Y} \rho^{Y/Z} - z^*(\tau^\infty) = 0, \\
& && \tilde{d}_X + \tilde{k}_X - 0.1\tilde{a}_X = 0, \quad \tilde{d}_{X^*} + 1 - 0.1\tilde{a}_{X^*} = 0, \\
& && \tilde{d}_Y + \tilde{k}_Y - 10\tilde{a}_Y = 0, \quad \tilde{d}_{Y^*} + \tilde{k}_{Y^*} - 0.1\tilde{a}_{Y^*} = 0, \\
& && \tilde{d}_Z + \tilde{k}_Z - 10\tilde{a}_Z = 0, \quad \tilde{d}_{Z^*} + \tilde{k}_{Z^*} - 0.1\tilde{a}_{Z^*} = 0, \\
& && 10^{-3} \leq \mathbf{p} \leq 10^3.
\end{aligned}$$

Solution of this problem was computationally less expensive than problems (P₃) and (P₄) for multicyclic systems, given the fact that the parameter space is quite constrained. Fig. 6.6 displays the results. Once more, our results show the superiority of the multicyclic system structure, as opposed to the monocyclic one. Even though the multicyclic systems exhibit significantly higher ultrasensitivity, their times to response are still significantly lower than in the case of monocyclic system. The more apparent difference between monocyclic and bicyclic system is again valid, whereas the difference between bicyclic and tricyclic remains minimal. Therefore, our results prove the functional superiority of a cascade versus a single cycle composition. Given the fact that cascades found in prokaryotes tend to be much shorter than those found in eukaryotes (34), our results could be widely applied on different living entities.

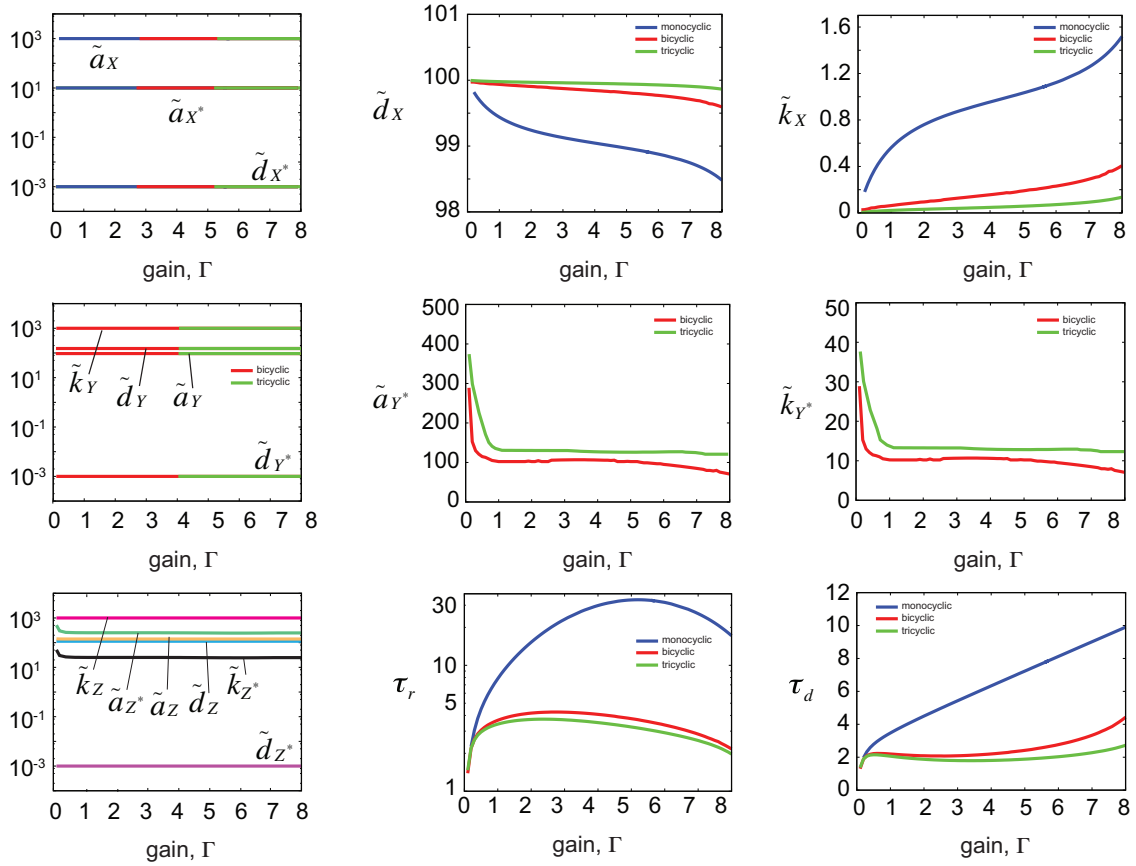


Figure 6.6: Design for minimal signaling rise time τ_r and decay time τ_d subject to given amplification and ultrasensitivity: comparison of monocyclic, bicyclic and tricyclic systems.

A qualitative comparison of the values of kinetic rates is given in Table 6.2. Unlike in the previous subsection, the horizontal scan (across different system architectures) reveals the same trends for every parameter. This is due to the fact that, clear requirements for ultrasensitivity, imposed through the Michaelis-Menten constants, limits the degrees of freedom during the optimization search. The way the problem is formulated predetermines the optimal design. As discussed previously, an alternative formulation could include taking into account both steady-state and dynamic representation of the system and solve them simultaneously, thus allowing more flexibility in choosing the appropriate combination of Michaelis-Menten constant that leads to a certain ultrasensitivity. In this case, it is important to clearly state the reference point when comparing the three different systems. Imposing one discrete level of ultrasensitivity as constraint on all three systems would not be so relevant, because of the inherent differences in the systems performances. For instance, as seen in chapter 4, tricyclic system always has a better performance than monocyclic system in terms of ultrasensitivity, even when we compare the worst desirable combinations of Michaelis-Menten constants.

	Monocyclic system	Bicyclic system	Tricyclic system
\tilde{a}_X	↗	↗	↗
\tilde{a}_X^*
\tilde{d}_X
\tilde{d}_X^*	↙	↙	↙
\tilde{k}_X
\tilde{a}_Y	
\tilde{a}_Y^*	
\tilde{d}_Y	
\tilde{d}_Y^*		↙	↙
\tilde{k}_Y		↗	↗
\tilde{k}_Y^*	
\tilde{a}_Z			...
\tilde{a}_Z^*
\tilde{d}_Z
\tilde{d}_Z^*	↙		↙
\tilde{k}_Z	↗		↗
\tilde{k}_Z^*

Table 6.2: Modular differences and similarities: optimal design for minimal signaling rise and decay times and maximal ultrasensitivity of monocyclic, bicyclic and tricyclic systems. Red arrows represent the value on the upper bound of variation range, blue arrows represent the value on the lower bound of variation range and three dots represent intermediate values.

6.3 Lessons learned from direct and inverse approach for parametric sensitivity analysis

All the results presented in chapters 4–6 can be wrapped into a big bundle of parametric sensitivity analysis. Approaching the same systems from different sides and addressing the different questions with the same goal, we completed the challenging task of parameter estimation for the optimal design. Beside the fact that our results contribute to quantitative knowledge on mathematical and computational analysis of signaling networks, we have also underlined the advantages of transducing the signals via cascades. A wide spectrum of

system functions that could be observed in tricyclic system are not just more superior than in the case of monocyclic system, but also the fact that the overall sensitivity indices for all properties of interest are lower in tricyclic than in monocyclic system, suggests that the system performance is more robust to parameter variations.

Dénouement and Encore **Discussion and Outlook**

7.1 Thesis summary

Mitogen-activated protein kinase cascade is one of the most studied cellular signaling pathways. It plays an important regulatory role in controlling cell proliferation, differentiation, cell survival and apoptosis, from yeast to mammals.

Systems biology literature accounts for a large number of studies of isolated MAPK module, as well as the studies that include numerous mechanism external to the cascade. Yet, a bundle of various questions about underlying principles of these networks still remain open for research. This thesis pursued towards identifying the conditions under which the MAPK module satisfies the demand for optimal system design. The optimality was defined through the analogy with man-made systems for signal transmission. Namely, the criterion accounts for fast signal propagation, large amplification and high sensitivity with respect to the input signal.

Prior to analysis of the prototypical MAPK cascade, we imposed all our research questions on the covalent modification cycle (referred to as monocyclic system), the integral motif found repeatedly (usually three times) throughout the structure of MAPK (referred as tricyclic system). This allowed for mapping of design patterns for the two signaling instances and their comparison.

Our study began with the optimal design for the steady-state property ultrasensitivity. Sigmoidal steady-state characteristic has been suggested as the preferred pattern over the hyperbolic response. The ability of the signaling module to operate in a truly all-or-none fashion has been explored both experimentally and theoretically. Setting the problem of identifying the rules for maximal ultrasensitivity into the optimization framework revealed new insights about the design of tricyclic systems that are not in line with the findings in monocyclic system.

Continuing the journey in the direction for optimal design of signaling modules in dynamic regime, we made a stopover and performed global sensitivity analysis. The aim was to identify the most influential parameters for a given system performance. Such analysis underlined the weak and the strong system nodes. Namely, we explored the way parameters contribute toward the transmission of sharp signals – the signals with high amplification, high ultrasensitivity and short propagation times. Modular analysis and comparison resulted in an important observation: a system composed of multiple cycles tends to be more robust to parametric uncertainty. Robustness of biological networks has often been postulated as one of the nature's design criteria. Our results quantitatively testify how the transition of robustness of signaling modules travels from one system architecture

to another. This work adds mechanistic insight to the contribution of key pathway components, and thus may support the identification of biomarkers for pharmaceutical drug discovery processes.

Finally, the last lap of the thesis elucidates the optimal design of signaling modules in the optimal regime, under the predefined criteria for steady-state behavior. We assembled the objective criteria as the mosaic of different system properties and investigated whether there are any compromises that are forced on the system, to simultaneously achieve minimal signal propagation times, maximal amplification and maximal ultrasensitivity. The arrangement of enzymatic cycles into multistep cascades again demonstrated admirable improvement in overall system behavior. Not just that the tricyclic system transduces the signal faster, but the same signal has higher amplification and higher ultrasensitivity. The benefit of assessing the research questions through optimization framework was a prerequisite to enhance the comprehensive knowledge about the complex dynamics of signaling modules.

7.2 Follow-up work and perspectives

Expanding the focus of analysis towards more complex signaling systems would result in reconsideration of the methods used to survey research questions presented in this thesis. Furthermore, congregation of all results produces a new starting point for further analysis. In the next subsections, the possible extensions and similar applications of this work are presented.

7.2.1 Timing matters for sensitivity

Among the many existing parameter sensitivity analysis approaches, global sensitivity analysis is one of the widely used methods to obtain parameter sensitivities within a large region about the nominal values. Sensitivity indices computed for persistent parameter changes, which means that parameters are constantly changed over time. However, as the impact of model parameters on the model output changes over time, time-dependent parameter sensitivity analysis has been proposed to study the effect of parameter variation on model output at different time (72). A parameter may have positive impact on the change of model output at early stage, but its effect can switch from positive to negative due to the complex interactions in the biological network. Therefore, one needs to know not only which parameters are critical for affecting a model output, but also at which time point they are important.

In chapter 5, we investigated the global sensitivities defining the arbitrary nominal parameter set. Such a setup was leading the system to operate in far suboptimal region. Chapter 6 brought forth the parameter values that accommodate optimal system behavior. The question that appears is whether the sensitivity indices of the parameters identified in chapter 5 still remain the same under the new conditions. More concretely, we are interested to see whether the same conclusions sustain along the wide spectra of possible scenarios.

To illustrate how this could be done and to obtain one more time-point snapshot for system analysis, we chose the parameter set that would bring the monocyclic system to have the amplification gain of 7, Hill coefficient of 6 and the sum of the two signaling times

would not exceed 35. The results in a graphic form are given in the *Supplementary Results* section. In the same manner as before, we observe the parametric sensitivities towards ultrasensitivity (Fig. A.1), amplification (Fig. A.2), signaling rise (Fig. A.3) and decay times (Fig. A.4.). Change of variation ranges of parameters significantly perturbed the influence of parameters on each property of interest. Random parameter sampling around the optimal point conveys the hypothesis that all-embracing robustness of the system is improved. In addition, the possible outcomes in this setting are restricted to more narrow stripes than it was the case in chapter 4.

All together, these new results and the results from chapter 4 imply that the more complex approach for sensitivity analysis needs to be chosen. Other researchers also recognized the need of using time-dependent sensitivity analysis. For instance, Schwacke and Voit presented a Taylor integration method for the efficient computation of time-dependent sensitivities for generalized mass-action systems (61). They investigated the effects of different initial species concentrations on the system dynamics. Yue and colleagues proposed the use of local dynamic sensitivity analysis and demonstrated how this framework could be used on NF- κ B signaling pathway model (63). Hu and Yuan used the same approach and revealed that the role of PI3K branch in the coupled pathways is to enhance the robustness of the MAPK pathway (131).

To close this subsection, we reiterate and emphasize the advantages of studying the systems dynamics with time-dependent sensitivity analysis. This framework can provide the information about multiple time scales existing in a complex signaling networks, the signs and strengths of responses to perturbations and more relevant identification of weak nodes in the pathway, as potential drug targets.

7.2.2 Avant-garde motif: noise

The question regarding which modeling framework, either differential equations or stochastic simulations, is more suitable for cellular signaling pathways has been the topic of numerous scientific discussions. Many authors would point out that signaling molecules exist in high number of copies and therefore proceed with deterministic modeling. Others would deny this fact and promote the stochastic simulation as the relevant one. Wolkenhauer and coworkers highlighted subtle differences and relationships of mathematical basis between generalized mass-action models and stochastic simulation (132). They concluded with the statement that one should not argue one way or the other regarding the numerically accurate representation, but whether a biological principle is reflected by the model or not. In any case, various assumptions are made about the physical context, including a constant volume, temperature, rapid diffusion, etc. Table 7.1 summarizes some of the experimentally observed concentration of different MAPK protagonist (48, 111, 133-137). For the comparison, the number of molecules is given beside the concentration values.

We were also interested to explore how the intrinsic fluctuations, coming from the low copies of signaling molecules, influence the optimal design identified in previous chapters. Does the system performance sustain despite the large stochastic influence? We formulated three concrete questions related to signaling cycles, aiming at demonstrating the utilization of stochastic modeling:

- How does the switching ability of the cycle depend on the number of molecules and the kinetic parameters?
- How can the cycle work as an intrinsic noise filter?

	Fujioka et al. 2006				Schoeberl et al. 2002		Bhalla 2004	
	HeLa		COS7		HeLa		n.d.	
	1.2e-12		2.5e-12		1.2e-12		1e-16	
<i>volume</i>	<i>uM</i>	<i>molecules</i>	<i>uM</i>	<i>molecules</i>	<i>uM</i>	<i>molecules</i>	<i>uM</i>	<i>molecules</i>
(upstream)	0.4	2.88e+05	0.53	7.95e+05	1.6	9.6e+05	0.2	1.2e+01
MAPKKK	0.013	9.36e+03	0.054	8.1e+03	0.0057	3.62e+03	0.2	1.2e+01
MAPKK	1.4	1.01e+06	1.8	2.7e+06	3.1	1.86e+06	0.18	1.08e+01
MAPK	0.96	6.91e+05	0.81	1.22e+06	2.1	1.26e+06	0.36	2.16e+01

	Huang and Ferrell 1996				Xenopus oocyte		Hatakeyama et al. 2003	
	Budding yeast		CHO		5e-07		CHO	
	1e-13		1e-12		5e-07		1e-12	
<i>volume</i>	<i>uM</i>	<i>molecules</i>	<i>uM</i>	<i>molecules</i>	<i>uM</i>	<i>molecules</i>	<i>uM</i>	<i>molecules</i>
(upstream)					0.003	1e+09	0.12	7.2e+04
MAPKKK							0.1	6e+04
MAPKK	0.035	2e+03	1.3	8e+05	1.2	3.5e+10	0.12	7.2e+04
MAPK	0.1	5e+03	2.8	1.7e+06	0.33	1e+11	1	6e+05

	Brightman and Fell 2000		Bhalla and Iyengar 1999	
	PC12		PC12	
	1.2e-12		1e-12	
<i>volume</i>	<i>uM</i>	<i>molecules</i>	<i>uM</i>	<i>molecules</i>
(upstream)	0.033	1.98e+04	0.1	6e+04
MAPKKK	0.017	1.02e+04	0.5	3e+05
MAPKK	0.6	3.6e+05	0.68	4.08e+05
MAPK	1.25	7.5e+05	0.26	1.56e+05

Table 7.1: Concentrations of different protagonists of MAPK signaling pathways, obtained experimentally in various cell types.

- What is the influence of kinetic parameters and the number of molecules on the transient behavior of the system?

From our results, we have learned that the stochastic nature has an effective influence on systems behavior. Beyond this, we identified the approximate number of molecules that could serve as the inflection point between the stochastic and deterministic modeling.

A chemical reaction is occurring when a molecule happens to collide with another, carrying sufficient energy. Deterministic formulation of chemical kinetics considers chemical constants as “rates” and the concentration of species are represented by single-valued functions of time. When the number of molecules in a system becomes very small, the probability, and not the rate at which a molecule will meet another one and react, generates high fluctuations. Considering a volume V , which contains a pair of molecules, it is possible to compute the probability of the two molecules colliding and reacting, using statistical mechanics (37). Intuitively, the probability of reaction is inversely proportional to V . Changing the volume in a deterministic formulation would have the result of reducing or increasing the concentration of each of the reacting species. In our study of monocyclic system, the volume is held constant to 10^{-15} L, which is the approximate volume of bacteria or some yeast. The results will assume numbers of molecules, rather than concentrations.

Hence, for comparison proposes, Table 7.1 shows some useful conversions.

Number of Molecules	μM
10	0.02
100	0.17
500	0.83
1000	1.66

Table 7.2: Relation of number of molecules and concentrations in our study of monocyclic system.

In a system where R_1, R_2, \dots, R_M reactions can take place, involving S_1, S_2, \dots, S_N chemically active species, the chemical reaction R_μ ($\mu = 1, 2, \dots, M$) is associated with a stochastic rate constant c_μ . The relation between the stochastic and deterministic rate constant can be derived from statistical mechanics and depends on the order of the reaction:

- Zero order: $c = N_A \cdot V \cdot k$
- First order: $c = k$
- Second order: $c = \frac{k}{N_A \cdot V}$

Here, N_A is the Avogadro number and k is the classical deterministic kinetic rate constant. Higher orders of reaction are not considered, knowing that the probability of three molecules colliding simultaneously is very low. Those are not often used in stochastic models.

The aim of the stochastic formulation is to define the probability of a reaction occurring at a certain time. For that purpose, we define the reaction probability density function $P(\tau, \mu)$ as the probability at time t that the next reaction in V will occur in the differential time interval $[t, t + \tau]$, and that the reaction will be the R_μ reaction. This probability depends on the number of molecules of each chemical species at time t and on the way they can be combined. Therefore, a state variable h_μ is defined as the number of distinct molecular reactant combinations for a reaction R_μ at time t . This state variable has to take in account the order of the reaction as follows:

- Zero order: $h_\mu = 1$
- First order: $h_\mu = X_j$
- Second order: $h_\mu = X_j X_k$

Here, X_1, X_2, \dots, X_N represent the current numbers of molecules of the chemical species S_1, S_2, \dots, S_N in V .

All the previous definitions were prerequisites for successful implementation of the direct stochastic method (37), which is the methodological framework for our stochastic simulations. The choice of the algorithm was prioritized to drawing of realistic distribution functions, without caring too much about computational cost. The reaction scheme for monocyclic system (Table 3.1) is the basis that is incorporated into stochastic framework by converting the deterministic reaction rates into stochastic reaction rates and using the number of molecules instead of concentrations.

Design criteria. The degree of fluctuations in the systems is quantified with noise. In most of the cases, a simple standard deviation σ is enough to quantify fluctuations but it is more suited to see the percentage of the mean μ . Thus, we can introduce the coefficient of variation η :

$$\eta = \frac{\sigma}{\mu} \quad (7.1)$$

The larger the coefficient of variation, the noisier the response will be. This measure is used in signaling theory to evaluate how much noise is carried in the signal.

Noise is the only additional property that arises from stochastic analysis. We apply the same rules and definitions for quantification of ultrasensitivity and signaling rise and decay times (Fig. 7.1). The simulations are performed on the time interval $[t_0, t_f]$, which is divided into N subintervals Δt . Hence, 10^6 simulations were performed for each t_p , and the arithmetic mean was calculated over the simulation outcomes. In each simulation, the system is activated with a step function enabling to measure τ_r (Fig. A.5A). When the system reaches steady state, it is pushed back to inactivated state with an inverse step function, which allows the measurement of τ_d (Fig. A.5B). Steady-state values were collected in order to calculate ultrasensitivity (Fig. A.5C).

The total number of molecules X_T is fixed and the concentration is acquired by simple conversion:

$$[X_T] = \frac{X_T}{V \cdot N_A} \quad (7.2)$$

In order to limit the number of simulations (and without losing the generality), two concentration ratios, as well as Michaelis-Menten constants will be held equal and simply referred to as ρ and K , respectively. Using the correlation of deterministic kinetic constants from the definition of Michaelis-Menten constants, the require balance between the parameters is ensured by:

$$\text{if } a = \frac{d+k}{K \cdot [X_T]} \text{ then: } a_X = a \cdot b, d_X = d \cdot b, k_X = k \cdot b, a_{X^*} = a, d_{X^*} = d, k_{X^*} = k.$$

In this model, d and k are fixed to 1. From the definition of α_X (Eq. (3.24)), we have $\alpha_X = b$.

The measures were taken at $t > 10^5$ s in order to be sure that steady state has been reached. The mean μ_{X^*} and the standard deviation σ_{X^*} were calculated over a number of reactions that has reached a set of $N = 10^6$ as follows:

$$\mu_{X^*} = \frac{\sum_{i=1}^N X_i^* \cdot \Delta t_i}{\sum_{i=1}^N \Delta t_i} \quad (7.3)$$

$$\sigma_{X^*} = \left(\frac{\sum_{i=1}^N (X_i^*)^2 \cdot \Delta t_i}{\sum_{i=1}^N \Delta t_i} - \mu_{X^*}^2 \right)^{1/2} \quad (7.4)$$

Hereafter, the mean will simply be represented as x^* .

Our study of ultrasensitivity affirms that, if the system is held to work outside of the ultrasensitive regime ($K=1$), the mean response of the system does not depend on the numbers of molecules and all the responses overlap with the deterministic simulation (Fig.

A.6). However, as K is decreased, the responses diverge and systems with small numbers of molecules have a reduced steepness compared to deterministic ones, whereas a system with 1000 molecules is close to the deterministic simulation. In other words, the restriction in the number of molecules causes the sensitivity to a threshold stimulus to decrease, therefore lowering the Hill coefficient of the system (Fig. A.7). At the same time, the complex sequestration has a more promoted dampening effect on ultrasensitivity when the number of molecules is limited (Fig. A.8). However, taking only maximal activation into account, the signal remains clear (Fig. A.9). This comes as a result of the inherent noise filtering property of ultrasensitivity (Figs. A.11 and A.12). Our hypothesis that the maximally activated ultrasensitive cycles are less noisy than hyperbolic ones again captures the importance of switch-like steady-state responses. As seen in chapter 6, the Hill coefficient has significant effect on the signaling rise time. Fig. A.13 gives more light for a parabolic dependence of the rise time and amplification gain: the intermediate levels of activation can hardly be achieved due to long rise time around the switching point and large deviations (Fig. A.11). Analysis for decay time confirms the advantage of the system being in the ultrasensitive regime, with additional requirement for the system being fully activated (Fig. A.14). The parameters in Fig. 6.6B resembles this case and accordingly the moderate influence of Hill coefficients on decay time is a consequence of such a setup.

In summary, the stochastic modeling of signaling cycle resulted into additional observations that could not be supposed intuitively. The question whether stochastic modeling in the scope of cellular signaling is justified still remains open. The fact that pharmaceutical companies mostly use deterministic modeling as the essential framework for model development might give it the practical role (60, 138-140).

7.2.3 Optimal control for drug delivery

Cancer occurs when there is a disturbance in normal cell growth regulation. Such breakdown is often associated to accumulation of defects in one or more signaling pathways. Ras proteins, upstream of MAPK pathway, are among the first molecules that were identified as oncogenes. Studies show that about 30% of all human tumor growth is related to mutations of the Ras protein (141). This number is likely to be even higher, if one considers all effectors in the Ras-related signaling pathways.

Systems biology uses an integrated approach to study and understand the function of biological systems, but, at the same time, it inquires how perturbations of such systems, for example the administration of a therapeutic drug, affect their function. An increasing appreciation of the functional dysregulation of protein interactions as the underlying cause of cancer has promoted the development of therapeutic agents that target specific signaling pathways involved in tumor development. These agents are supposed to activate or inhibit specific disease targets (usually proteins), while having as few side effects as possible, as opposed to the conventional chemotherapeutic agents that perturb different cell functions. In particular, drug intervention can be seen as control of signaling in cellular networks. Identification of control parameters presents an extreme challenge due to the combinatorial explosion of control possibilities in combination therapy and to the incomplete knowledge of the systems biology of cells.

In this section, we discuss the application of optimization framework for drug delivery. Concretely, optimal control (80, 142), as the subdiscipline of dynamic optimization, seems well suited to address this kind of problems. A solution of optimal control problem, referred just as *optimal control*, is a set of differential equations describing the trajectories of the control variables that minimize the cost functional. The cost functional to be minimized

can be chosen to take into account different factors: the toxicity of some reactants, the drug action, the costs of the drug itself, etc.

As a very simple and general example, we define the drug delivery problem as follows:

- the system consists of the prototypical MAPK cascade (Fig. A.15);
- the system is supposed to directly influence cancer progression;
- protein X (i.e., Raf protein) in its inactive form is identified as target and the inhibitor (drug) I is introduced in the system to act on it;
- the initial distribution of the drug is supposed to linearly increase over time of 100 arbitrary units;
- the concentration profile of the last activated kinase Z^* is observed as the system output and its effect on, for instance, cell proliferation, needs to be turned off;
- the cost functional is the minimization of the drug concentration.

Intuitively, one would envision that the more drug is introduced in the system, the overall effect would be more prominent (Fig. A.16). In this kind of unconstrained setting, the trivial solution of using the maximal possible drug concentration would probably kill the patient and therefore the set of constraints should be applied. For instance, one should precisely define the upper and lower bounds for drug intake, and the time regimen for drug administration.

The optimal control problem is formulated as follows:

$$\begin{array}{ll}
 \text{find} & \text{drug administration regime} \quad \text{that} \\
 \text{minimize} & \text{the drug concentration} \\
 \text{subject to} & \text{transient model} \\
 & \text{the given time interval} \\
 & \text{turning-off the last activated kinase.}
 \end{array} \tag{P_8}$$

This problem sketch-up can be comfortably translated into mathematical form and Optimal Control could be used as a reliable mathematical tool for pharmacological pre-clinical research. The mathematical formulation reads:

$$\begin{array}{ll}
 \min_{u(t)} & \rho^{IX}(t_f) \\
 \text{subject to} & \mathbf{F}_r(\xi(\tau), \dot{\xi}(\tau), \mathbf{p}, \mathbf{r}) = \mathbf{0}, \quad 0 < t < t_f, \quad \xi(0) = \xi_0, \\
 & d\rho^{IX}(t)/dt = u(t), \quad 10^{-3} < u(t) < 10^3, \\
 & z^*(t_f) < 0.001.
 \end{array} \tag{P_8}$$

The effect of a drug can be simulated by varying one or more parameters simultaneously, and comparing the response of the system to a reference (“health”) dynamical state of the cell. Finally, the solution of optimal control problem is quantifying the optimal drug administration protocol (Fig. A.17).

The exploitation of optimal control techniques for simulating and tuning drug effects has just recently drawn some attention (143, 144). We believe that these techniques offer wide diapason of opportunities for a better drug development and disease curing initiatives.

7.3 Closing intermezzo

In the intracellular signaling networks that regulate important cell processes, the base pattern comprises the cycle of reversible phosphorylation of a protein, catalyzed by kinases and opposing phosphatases. Concatenation of these cycles into a cascade has been identified as the underlying structure of mitogen-activated protein kinase cascades – a very important signaling motif found in all organisms.

Mathematical modeling and analysis have been used for better understanding of the functions of signaling modules and to capture the rules governing systems behavior. However, since biochemical parameters in signaling pathways are not easily accessible experimentally, it is necessary to explore possibilities for either steady-state or dynamic responses in these systems. A number of studies have focused on analyzing these properties separately. In order to be able to interpret a broader range of phenotypes, it is necessary to take into account both of these responses simultaneously.

In this thesis, we investigated the trade-offs between optimal characteristics of both steady-state and dynamic responses. As a result, we found the biochemical and biophysical parameters that determine these trade-offs and we analyzed if there exist conditions under which we can simultaneously achieve optimal steady-state and dynamic performance. Remarkably, we discovered that even a single covalent modification cycle could simultaneously achieve high ultrasensitivity, high amplification and rapid signal transduction. Furthermore, the arrangement of cycles in a cascade significantly improves the systems performance, increases the robustness and provides multiple nodes for potential regulation.

The closing statement of this thesis goes along the often-cited words of Sutherland (145):

“Optimization: this beguilingly simple idea allows biologists not only to understand current adaptations, but also to predict new designs that may yet evolve”.

We believe that the presented study contributes to the quantitative knowledge about underlying principles of cellular signaling and the overall approach represents the indispensable framework to expand this knowledge further.

The Backup Trio Supplementary Results

A.1 Global sensitivity analysis in optimal regime

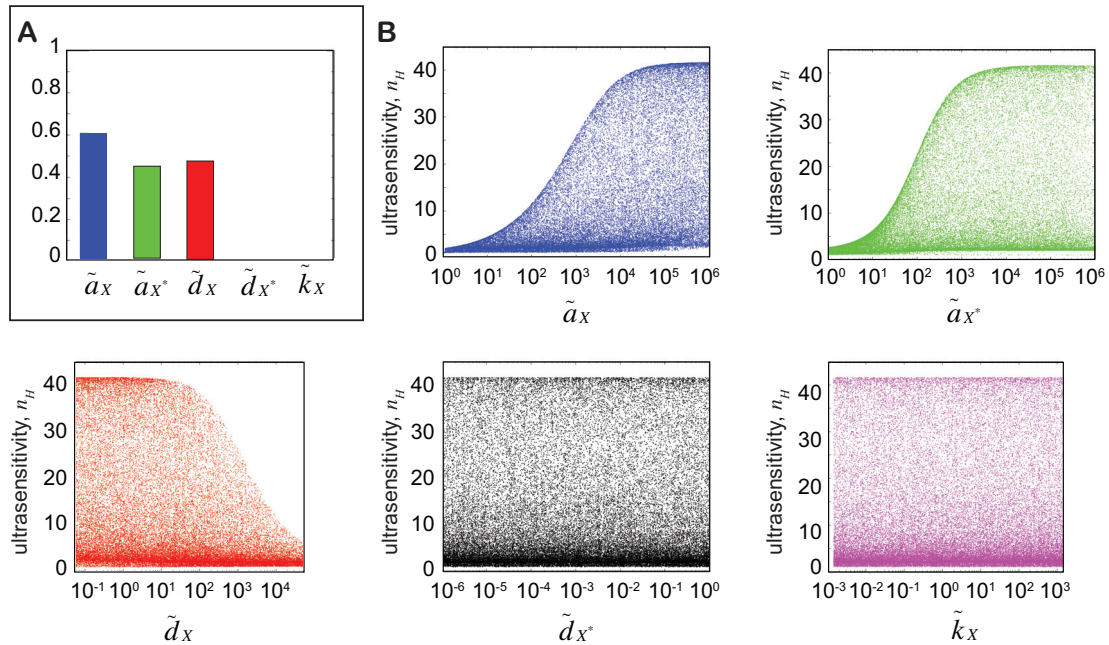


Figure A.1: (A) Total sensitivity indices and (B) scatter plots for each parameter of monocyclic system, with respect to ultrasensitivity in optimal regime. The concentration ratios are set to $\{\rho^{S/X}, \rho^{P/X}\} = \{0.1, 0.1\}$. The nominal parameter set $\{\tilde{a}_X, \tilde{a}_{X^*}, \tilde{d}_X, \tilde{d}_{X^*}, \tilde{k}_X\} = \{1000, 1000, 53, 0.001, 1.42\}$ determines the Hill coefficient $n_H = 6$. Higher values of parameters \tilde{a}_X and \tilde{a}_{X^*} still contribute to higher Hill coefficients. The shift to the right of variation range for parameter \tilde{d}_X and the shift to the left of variation range for parameter \tilde{d}_{X^*} underline the importance of low values of these parameters. This is resulting in the increase of sensitivity index for \tilde{d}_X and the decrease for \tilde{d}_{X^*} . Comparison with Fig. 5.2 suggests the threshold value of 10^0 for parameter \tilde{d}_{X^*} : if $\tilde{d}_{X^*} < 10^0$, this parameter does not have the influence on Hill coefficient.

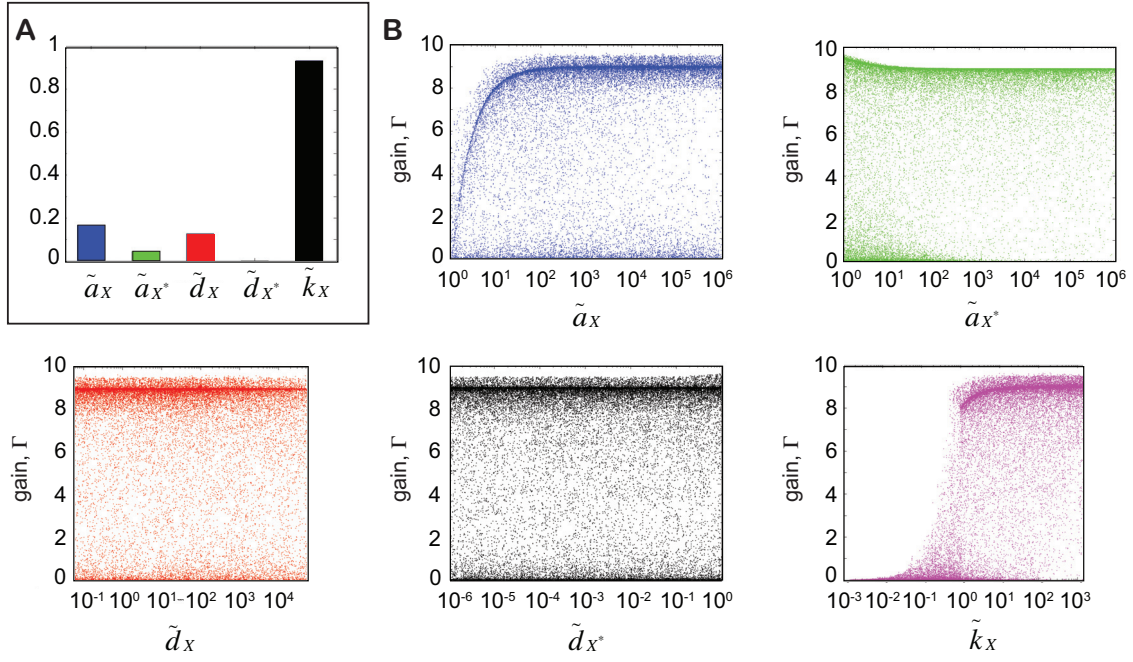


Figure A.2: (A) Total sensitivity indices and (B) scatter plots for each parameter of monocyclic system, with respect to amplification in optimal regime. The concentration ratios are set to $\{\rho^{S/X}, \rho^{P_x/X}\} = \{0.1, 0.1\}$. The nominal parameter set $\{\tilde{a}_x, \tilde{a}_{x^*}, \tilde{d}_x, \tilde{d}_{x^*}, \tilde{k}_x\} = \{1000, 1000, 53, 0.001, 1.42\}$ determines the amplification gain $\Gamma = 6$. Allowing the parameters \tilde{a}_x and \tilde{a}_{x^*} to have higher values, global sensitivity analysis highlights the parameter \tilde{k}_x as the most important. Higher values of this parameter will lead to faster accumulation of the active form of the kinase x^* , leading to higher amplification gain. Lower values of \tilde{k}_x allow the complex $\{x:s^*\}$ to stay more stable. Random parameter sampling in the optimal regime favors high amplification.

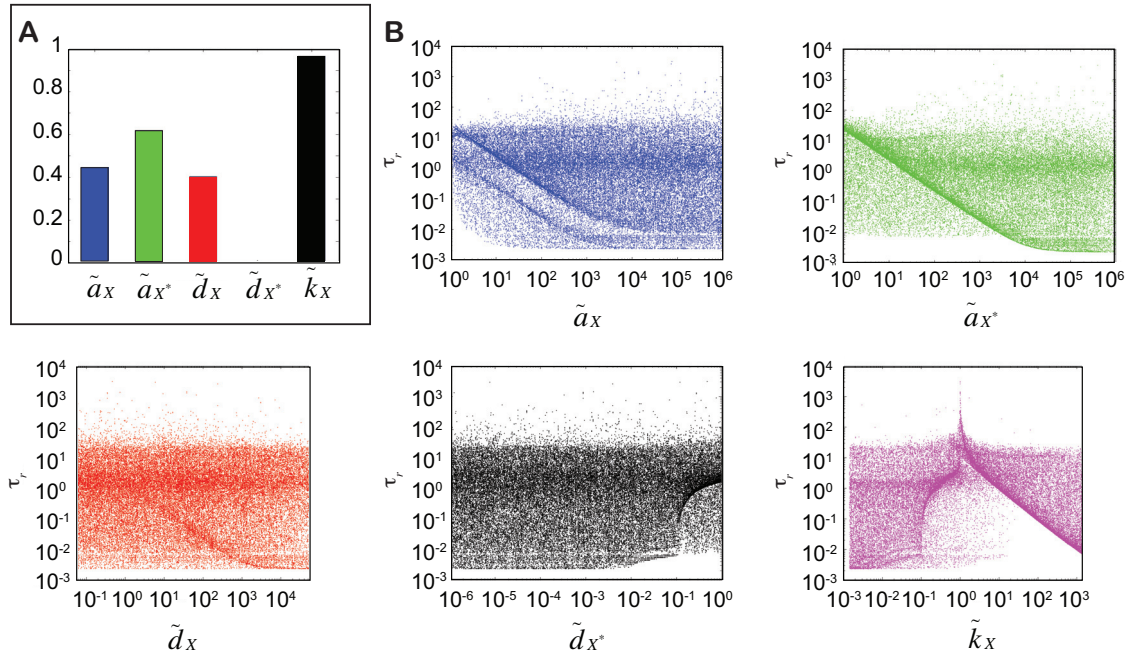


Figure A.3: (A) Total sensitivity indices (A) and scatter plots (B) for each parameter of monocyclic system, with respect to rise time in optimal regime. The concentration ratios are set to $\{\rho^{S/X}, \rho^{P/X}\} = \{0.1, 0.1\}$. The nominal parameter set $\{\tilde{a}_x, \tilde{a}_{x^*}, \tilde{d}_x, \tilde{d}_{x^*}, \tilde{k}_x\} = \{1000, 1000, 53, 0.001, 1.42\}$ determines the signaling rise time $\tau_r = 26.7$. As seen for amplification in optimal regime, parameter \tilde{k}_x strikes out as the most important to determine signaling rise time. Parabolic shape of rise time versus \tilde{k}_x is most probably the artifact of the interplay between this parameter and parameters \tilde{a}_x and \tilde{a}_{x^*} . Higher values of these two parameters allow for the fast response of monocyclic system, thus letting \tilde{k}_x to be the tuning factor.

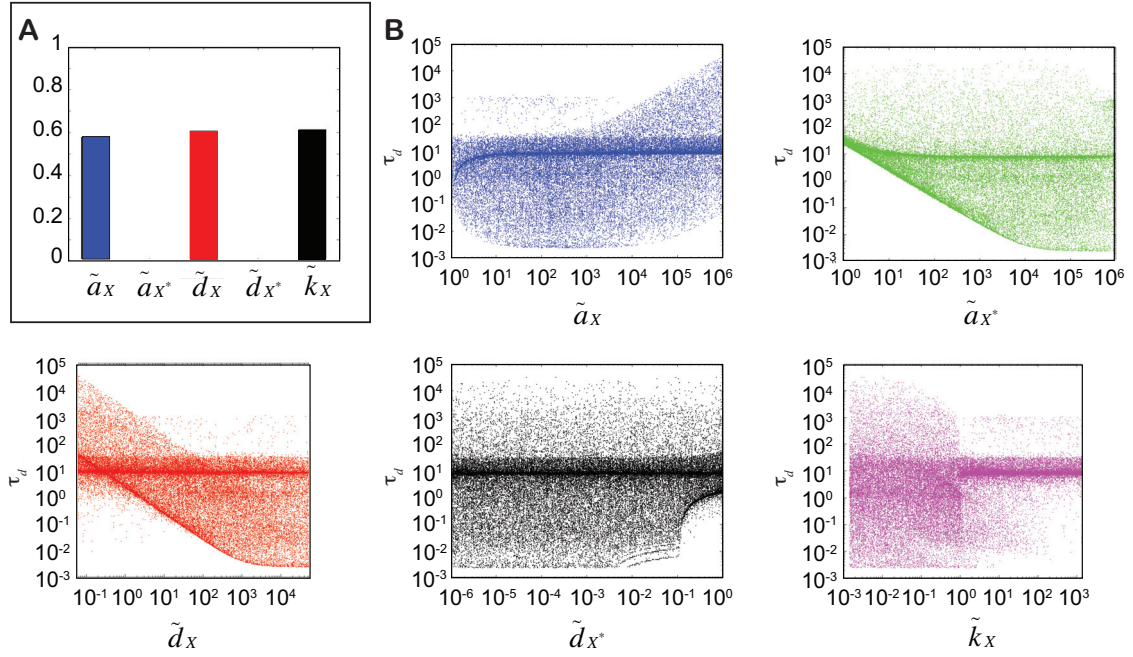


Figure A.4: (A) Total sensitivity indices and (B) scatter plots for each parameter of monocyclic system, with respect to decay time in optimal regime. The concentration ratios are set to $\{\rho^{S/X}, \rho^{P_X/X}\} = \{0.1, 0.1\}$. The nominal parameter set $\{\tilde{a}_X, \tilde{a}_{X^*}, \tilde{d}_X, \tilde{d}_{X^*}, \tilde{k}_X\} = \{1000, 1000, 53, 0.001, 1.42\}$ determines the signaling decay time $\tau_d = 6.4$. As in all previous cases, parameter \tilde{d}_{X^*} does not effect the signaling decay time. Low values of this parameter are ensuring high ultrasensitivity, high amplification and short signaling rise and decay times. An interesting point is that global sensitivity analysis in optimal regime points out on the parameters that had sensitivity indices almost zero in the analysis in chapter 5. Random parameter sampling favors decay time equals to 10.

A.2 Avant-garde motif: Noise

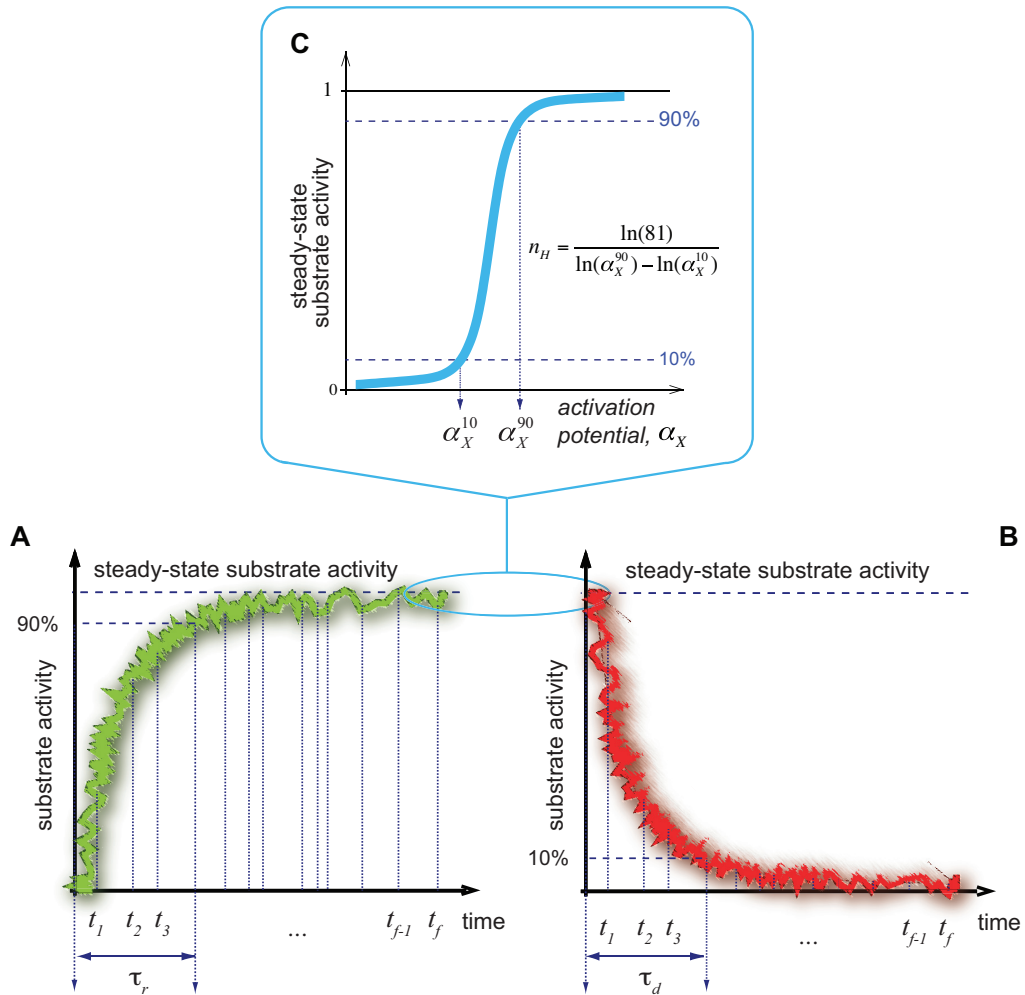


Figure A.5: Design criteria for stochastic modeling. Stochastic simulations and calculation of (A) signaling rise time, (B) signaling decay time and (C) ultrasensitivity.

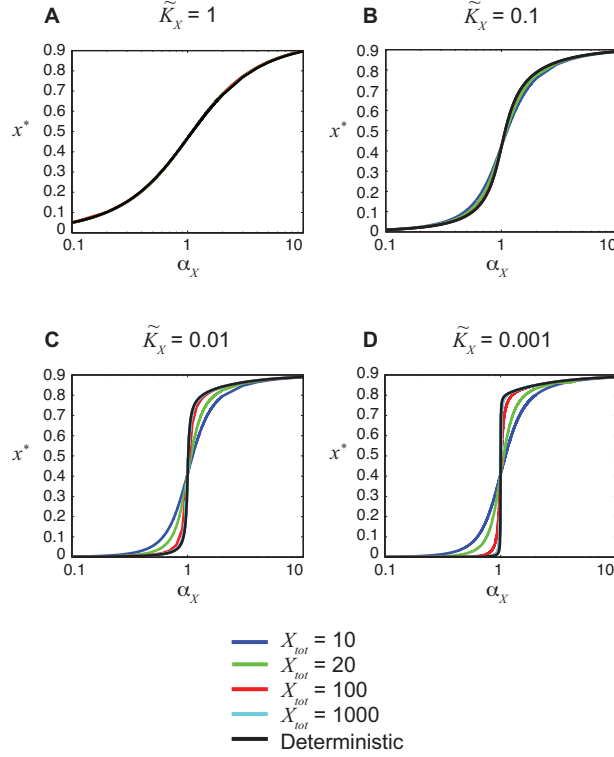


Figure A.6: Stochastic influence on steady-state responses in monocyclic system. The mean activation was plotted as a function of the activation potential α_X for different number of total molecules in the system $X_T = 10, 20, 100$ and 1000 . Deterministic simulation is also presented to enable comparison. The concentration ratios are both held to $\rho^{S/X} = \rho^{P_X/X} = \rho = 0.1$. The Michaelis-Menten constants are fixed as: (A) $\tilde{K}_X = \tilde{K}_{X^*} = K = 1$, (B) $\tilde{K}_X = \tilde{K}_{X^*} = K = 0.1$, (C) $\tilde{K}_X = \tilde{K}_{X^*} = K = 0.01$, (D) $\tilde{K}_X = \tilde{K}_{X^*} = K = 0.001$. When the system is held to work outside of the ultrasensitive regime ($K = 1$), the mean response of the system does not depend on the number of molecules and all the responses overlap with the deterministic simulation. However, as K is decreased, the responses diverge and systems with small number of molecules have a reduced steepness with respect to deterministic ones. The system with 1000 molecules is always close to the deterministic simulation, suggesting that stochastic fluctuations do not affect this system in steady state.

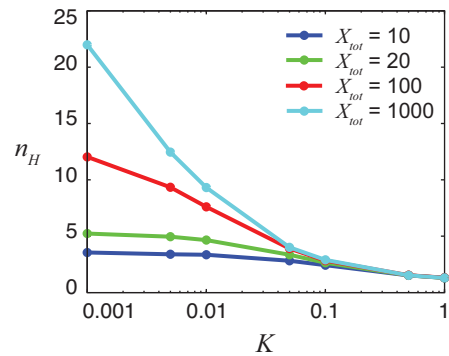


Figure A.7: Stochastic influence on Hill coefficient values in monocyclic system. When $X_T = 1000$ molecules, the Hill coefficient highly increases as Michaelis-Menten constants are reduced, whereas a system composed of 10 molecules has only a slow change in ultrasensitivity. The regime with values of Michaelis-Menten constants higher than unity is considered being hyperbolic and there is no stochastic dampening in this region. This steady-state behavior reveals a very interesting property of the system: if the system operates at enzyme saturation, the lower the number of molecules, the higher the loss of ultrasensitivity. The fact that ultrasensitivity is decreasing as the total number of molecules gets restricted is in line with studies done in (45).

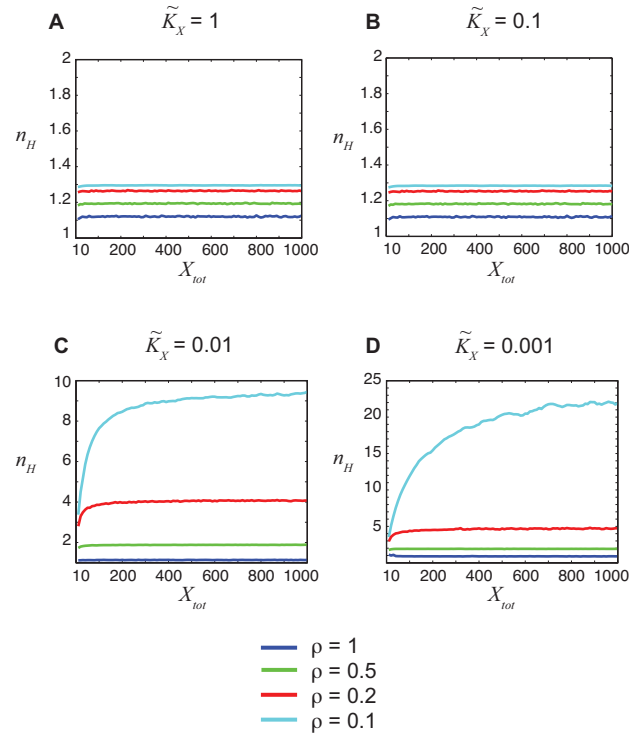


Figure A.8: Stochastic influence on effects of sequestration in monocyclic system. The stochastic dampening is measured with the Hill coefficient n_H as a function of the number of molecules X_T for different enzyme/kinase concentration ratios $\rho = 0.1, 0.2, 0.5, 1$. The Michaelis-Menten constants were fixed as: (A) $K = 1$, (B) $K = 0.1$, (C) $K = 0.01$, (D) $K = 0.001$. As we have seen in chapter 4, the accumulation of substrate protein in the complex has a very important dampening effect on the ultrasensitivity. This dampening depends on the ratio between protein and enzymes concentration ρ . Even a change from $\rho = 0.1$ to $\rho = 0.2$ introduces a five-fold decrease of the Hill coefficient when the number of molecules equals 1000. Nevertheless, when the number of molecules is reduces, this trend is conserved with diminished differences. If the number of protein kinases is equal to the number of enzyme molecules (i.e. $\rho = 1$), the ultrasensitivity is lost.

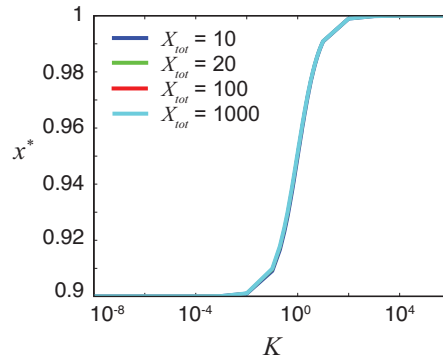


Figure A.9: Stochastic influence on maximal activation in monocyclic system. Concentration ratios are fixed to $\rho = 0.1$. As seen in Fig. A.7, the influence of the complex formation is quite important, since the increase of enzyme to kinase concentration ratios ρ reduces the ultrasensitivity dramatically. However, it also reduces the maximal activation achievable and, therefore the amplification of the signal. So far, it is not known if the stochasticity can modify this maximal activity. In deterministic model, changing the values of the Michaelis-Menten constants with respect to maximal activation makes the system switch from ultrasensitive to hyperbolic regime. The transition between these regimes for 10, 100 and 1000 molecules and $\rho = 0.1$ lead to overlapping responses. In the ultrasensitive regime (i.e. $K < 0.1$), the maximal achievable activation is around 0.9, since the kinetic rates directly related to the complex formation, a_X and a_{X^*} , are much faster than the other ones and this generates complex sequestration. These results indicate that maximal activation is completely independent on intrinsic noise and that it is only a function of kinetic parameters.

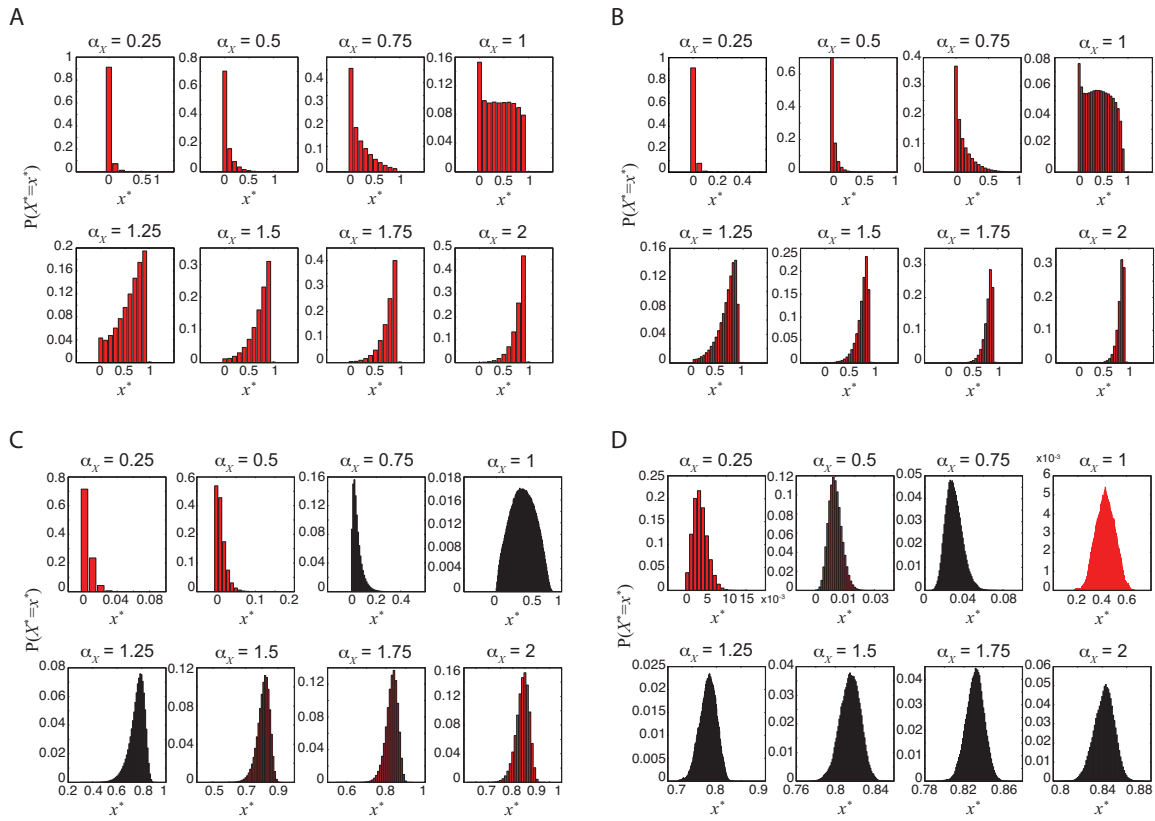


Figure A.10: Probability density functions of activation in monocyclic system. Measures were done for $\alpha_X = 0.25, 0.5, 0.75, 1, 1.25, 1.5, 1.75, 2$ for: (A) $X_T = 10$, (B) $X_T = 20$, (C) $X_T = 100$ and (D) $X_T = 1000$ molecules. The densities were generated with 10^6 samples and $K = 0.01$. By definition, stochastic simulation generates a set of outcomes following a certain distribution function. The probability distribution function for a small number of molecules $X_T = 10$ is naturally very discrete and displays heavy tails. The number of possible outcomes is increasing around $\alpha_X = 1$ and this is also the point where the variability is the largest. Moving towards $X_T = 1000$, the distribution is much closer to a normal distribution, displaying small or no tails. As in the previously mentioned case, the largest variation is occurring around $\alpha_X = 1$. However, this variation is much smaller than the one displayed by the system containing 10 molecules. The systems containing 20 and 100 molecules have intermediate distribution and the probability of the substrate being inactivated ($P(X^* = 0)$) is high for $\alpha_X < 1$. These differences in probability distribution function show the usefulness of using exact stochastic algorithm. Intuitively, one would assume a normal distribution for all the cases, which could lead to undesirable mistakes.

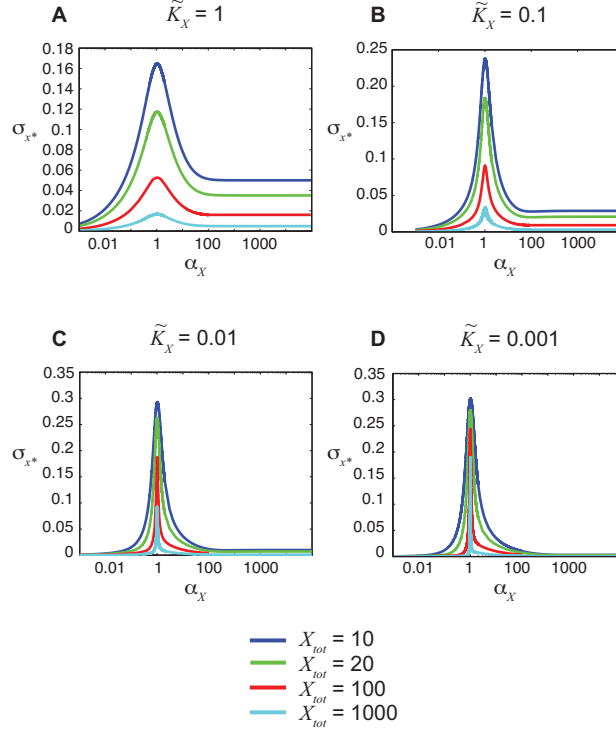


Figure A.11: Stochastic influence on standard deviation in monocyclic system. The standard deviation σ_{x^*} is measured as a function of the activation potential α_X for $X_T = 10, 20, 100, 1000$ molecules and fixed Michaelis-Menten constants to: (A) $K = 1$, (B) $K = 0.1$, (C) $K = 0.01$, (D) $K = 0.001$. So far, only the mean results of the stochastic simulations were presented, without performing an explicit computation of the noise and discussing its qualitative characteristics. The standard deviation σ_{x^*} could be recognized as a qualitative measure of fluctuations in the system. Although the results from Fig. A.10 indicate that an assumption about normal distribution was not relevant for modeling; it is sufficiently descriptive to use it in the context of qualitative visualization of standard deviation. For a full quantitative study, the calculation of quintiles should be carried out. The highest fluctuations are generated when $\alpha_X = 1$ (i.e. the summarized rate of substrate activation equals to the summarized rate of substrate deactivation) forming the top of a peak. Small numbers of molecules are generating more fluctuations and therefore a greater standard deviation. These fluctuations are increasing as the Michaelis-Menten constants are decreased. Therefore, ultrasensitive systems are generating a higher standard deviation values compared to hyperbolic systems at switching point, but the standard deviation is smaller for high activations.

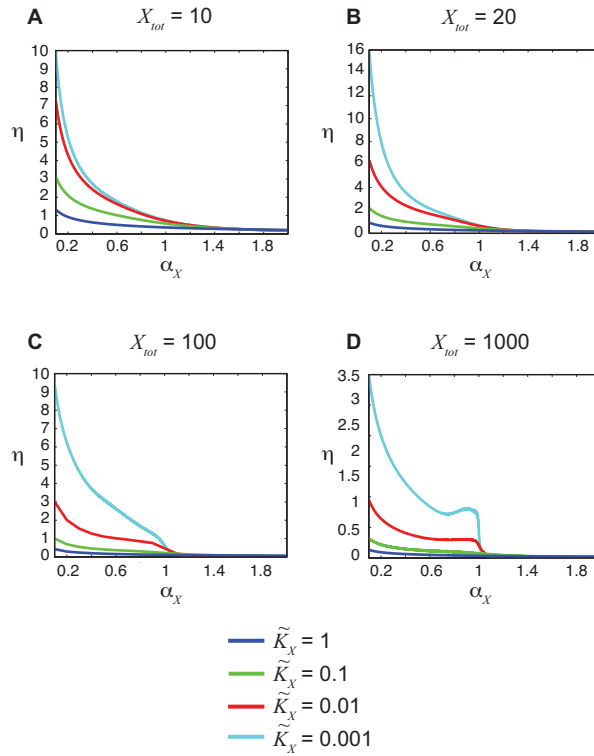


Figure A.12: Coefficient of variation in monocyclic system, as a function of the activation potential α_X , for different Michaelis-Menten constants $K = 0.001, 0.01, 0.1, 1$. Number of molecules are held to: (A) $X_T = 10$, (B) $X_T = 20$, (C) $X_T = 100$ and (D) $X_T = 1000$. The coefficient of variation η is a quantitative measure of the noise, as it gives the ratio of fluctuations regarding the mean value of activation. The first noticeable fact is that, as ultrasensitivity is increasing (i.e. by reducing K), the cycle becomes much noisier in the region of $\alpha_X < 1$. Intrinsic fluctuations, arising from stochastic chemical reactions, behave differently in four cases of total number of molecules. For 10 molecules, the noise does not display any abrupt change as we enhance the activity, and it is gradually decreasing. On the other hand, for 1000 molecules, there arise a steep attenuation of the noise after the switching point ($\alpha_X = 1$), leading to low noise for highly activated systems. Since the noise-free signal is preferred from signal transduction point of view, the results suggest that the signal has to reach high activation in order to be clear. Therefore, ultrasensitive cycles can truly work as noise filter devices. Thattai and van Oudenaarden came up with the similar conclusion in their study of generic stochastic cascade, where they used Langevin technique to model random fluctuations (46).

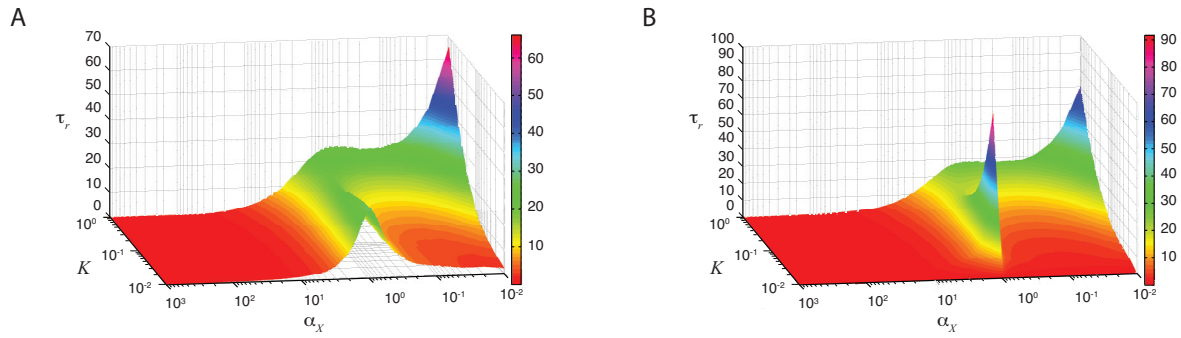


Figure A.13: Stochastic influence on signaling rise time. Rise time is presented as a function of activation potential α_x and Michaelis-Menten constants K . The measures are carried out for different number of molecules: (A) $X_T = 10$ and (B) $X_T = 1000$. From our deterministic analysis, we have learned that there is a significant trade-off between signaling rise time and ultrasensitivity. The first interesting observation is presence of the peak at switching point (around $\alpha_x = 1$). The size of the peak is decreased if the number of molecules is reduced. All of this suggests that the switching area brings higher fluctuations and instability, as seen in Fig. A.10. The effect of stochastic fluctuations in dynamic regime is aligned with the findings from steady-state analysis: number of molecules modifies the ultrasensitivity for $K < 1$, resulting in increasing values of signaling rise time. The signaling rise times are equivalent after the switching point, regardless of the number of molecules in the system, but also regardless of the value of Michaelis-Menten constants K . Even more, they are barely detectable for this region of high activation. In the unsaturated regime, a system promoting small activation of kinases will need more time to reach steady state, which is in a direct relation with higher values of signaling rise time. This latter result could show another benefit of operating in the ultrasensitive regime. Indeed, if the number of signaling molecules is not sufficient to make the system overpass the switching point, then the signal will take too much time and could be brought back to initial state. However, this is still to be proved.

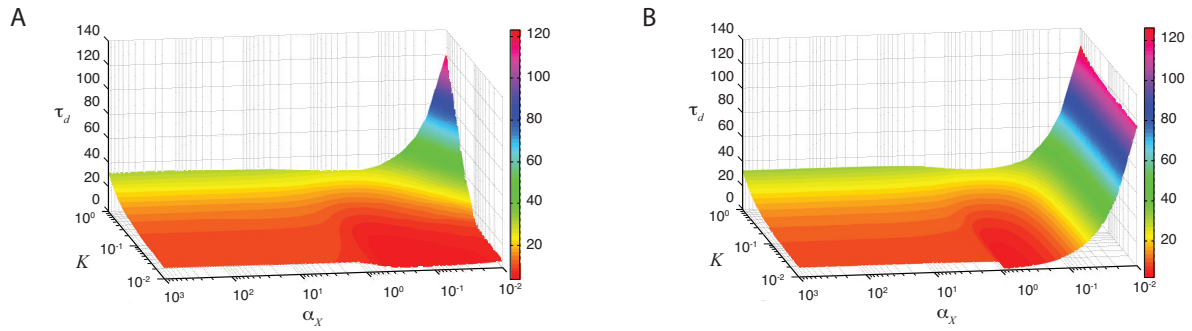


Figure A.14: Stochastic influence on signaling decay time. Decay time is presented as a function of activation potential α_X and Michaelis-Menten constants K . The measures were carried out for different number of molecules: (A) $X_T = 10$ and (B) $X_T = 1000$. First noticeable conclusion is that there are quite a number of dissimilarities between signaling rise and decay times, but the overall recommendation of operating under high activation and with high ultrasensitivity is still valid. Switching point of activation potential does not perturb signaling decay time, as it was the case with signaling rise time. Nevertheless, after the switching point, in the region of high activation, decay time increases when the system is less ultrasensitive. This is in a direct connection with the fact that the only way of degradation of X^* is through the kinetic rate a_{X^*} , which becomes higher how Michaelis-Menten constant K_{X^*} decreases. Region of low activation brings increase in decay time, especially with the higher copy of molecules. Analysis of intermediate number of molecules (i.e. $X_T = 50$ and $X_T = 100$) shows that this increase is consistent (data not shown).

A.3 Drug delivery as an optimal control problem

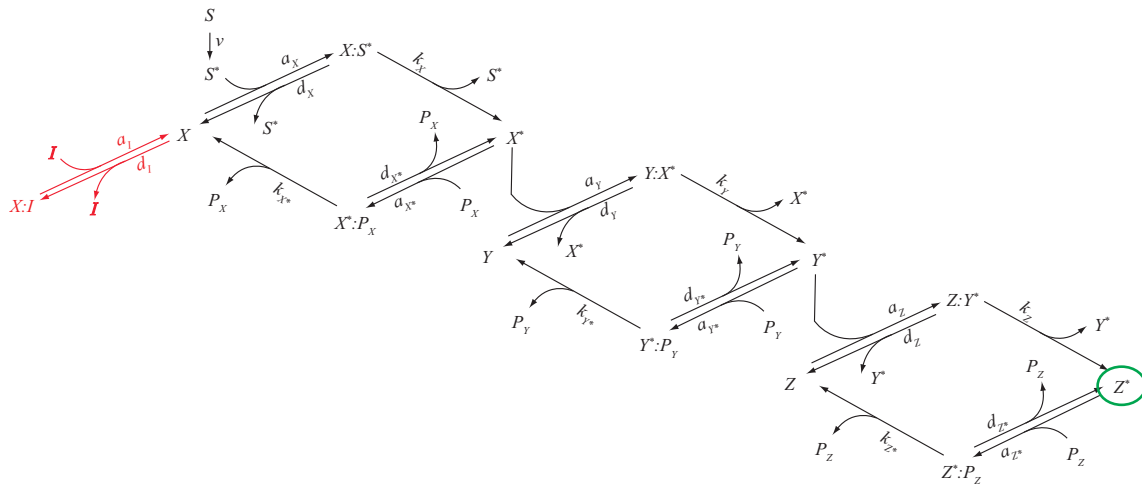


Figure A.15: Example of therapeutical action in prototypical MAPK cascade. The inhibitor (drug) I is introduced in the system and specified to react only on the inactive form of signaling protein X . The overall drug effect on the signaling pathway is observed through the concentration of the last activated kinase Z^* . The trajectory of concentration ratio between the total concentration of the drug and the total concentration of the signaling protein X , $\rho^{I/X}$ is defined as the control variable.

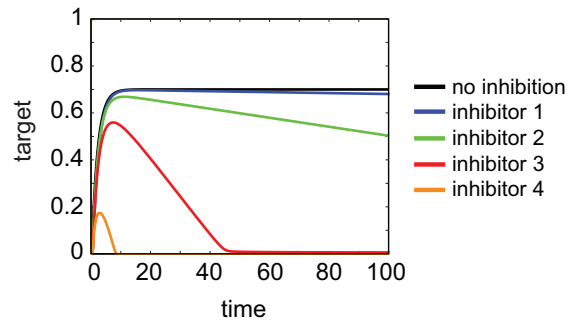


Figure A.16: Effect of different inhibitors on the dynamic responses of signaling network. The inhibitors (drugs) are distributed in linearly increasing manner over time. The difference between inhibitors 1–4 is the dynamics of their administration.

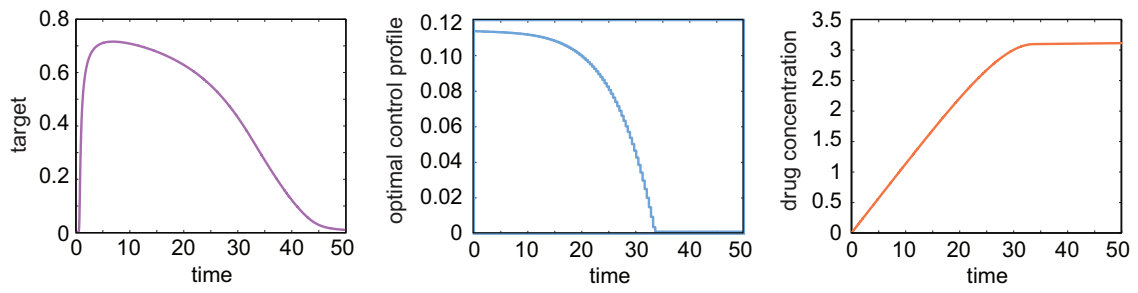


Figure A.17: Drug delivery as optimal control problem: therapeutic effect on the specified target (*left plot*), optimal control profile (*middle plot*) and the concentration profile of the inhibitor being added in the system (*right plot*).

BIBLIOGRAPHY

1. Ideker, T., T. Galitski, and L. Hood. 2001. A new approach to decoding life: Systems biology. *Annu. Rev. Genomics Hum. Genet.* 2:343-372.
2. Kitano, H. 2002. Systems biology: A brief overview. *Science* 295:1662-1664.
3. Levchenko, A. 2001. Computational cell biology in the post-genomic era. *Mol. Biol. Rep.* 28:83-89.
4. Kitano, H. 2002. Computational systems biology. *Nature* 420:206-210.
5. Wellstead, P., E. Bullinger, D. Kalarnatianos, O. Mason, and M. Verwoerd. 2007. The role of control and system theory in systems biology. In *10th IFAC Symposium on Computer Applications in Biotechnology/8th IFAC Symposium on Dynamics and Control of Process Systems*. Pergamon-Elsevier Science Ltd, Cancun, MEXICO. 33-47.
6. Goldbeter, A. 2004. Computational biology: A propagating wave of interest. *Current Biology* 14:R601-R602.
7. Cho, C. R., M. Labow, M. Reinhardt, J. van Oostrum, and M. C. Peitsch. 2006. The application of systems biology to drug discovery. *Current Opinion in Chemical Biology* 10:294-302.
8. Davidov, E. J., J. M. Holland, E. W. Marple, and S. Naylor. 2003. Advancing drug discovery through systems biology. *Drug Discov. Today* 8:175-183.
9. Marks, F., U. Klingmüller, and K. Müller-Decker. 2009. Cellular Signal Processing: An Introduction to the Molecular Mechanisms of Signal Transduction. Garland Science, New York.
10. Hornberg, J. J., F. J. Bruggeman, H. V. Westerhoff, and J. Lankelma. 2004. Cancer: A systems biology disease. In *5th International Conference on Systems Biology (ICSB 2004)*. Elsevier Sci Ltd, Heidelberg, GERMANY. 81-90.
11. Gomperts, B., I. M. Kramer, and P. E. R. Tatham. 2003. Signal Transduction. Elsevier/Academic Press, San Diego.
12. Lauffenburger, D., and J. J. Linderman. 1993. Receptors: Models for Binding, Trafficking, and Signaling. Oxford University Press, New York.
13. Nicholson, K. M., and N. G. Anderson. 2002. The protein kinase B/Akt signalling pathway in human malignancy. *Cellular Signalling* 14:381-395.
14. Vara, J. A. F., E. Casado, J. de Castro, P. Cejas, C. Belda-Iniesta, and M. Gonzalez-Baron. 2004. PI3K/Akt signalling pathway and cancer. *Cancer Treat. Rev.* 30:193-204.
15. Chen, Z. J. J. 2005. Ubiquitin signalling in the NF-kappa B pathway. *Nature Cell Biology* 7:758-U719.
16. Courtois, G., and T. D. Gilmore. 2006. Mutations in the NF-kappa B signaling pathway: implications for human disease. *Oncogene* 25:6831-6843.
17. Darnell, J. E., I. M. Kerr, and G. R. Stark. 1994. Jak-STAT pathways and transcriptional activation in response to IFNs and other extracellular signaling proteins. *Science* 264:1415-1421.
18. Seidel, H. M., P. Lamb, and J. Rosen. 2000. Pharmaceutical intervention in the JAK/STAT signaling pathway. *Oncogene* 19:2645-2656.
19. Cobb, M. H. 1999. MAP kinase pathways. *Prog. Biophys. Mol. Biol.* 71:479-500.
20. Wagner, E. F., and A. R. Nebreda. 2009. Signal integration by JNK and p38 MAPK pathways in cancer development. *Nat. Rev. Cancer* 9:537-549.

21. Ashyraliyev, M., Y. Fomekong-Nanfack, J. A. Kaandorp, and J. G. Blom. 2009. Systems biology: parameter estimation for biochemical models. *Febs J.* 276:886-902.
22. Miao, H., X. Xia, A. S. Perelson, and H. Wu. 2011. On Identifiability of Nonlinear ODE Models and Applications in Viral Dynamics. *SIAM Rev.* 53:3-39.
23. Heinrich, R., S. Schuster, and H. G. Holzhutter. 1991. Mathematical analysis of enzymatic reaction systems using optimization principles. *European Journal of Biochemistry* 201:1-21.
24. Allem, A. C. 2003. Optimization theory in plant evolution: An overview of long-term evolutionary prospects in the angiosperms. *Bot. Rev.* 69:225-251.
25. Goldsmith, T. H. 1990. Optimization, constraints, and history in the evolution of eyes. *Q. Rev. Biol.* 65:281-322.
26. Smith, J. M. 1978. Optimization theory in evolution. *Annu. Rev. Ecol. Syst.* 9:31-56.
27. Uhrmacher, A. M., D. Degenring, and B. Zeigler. 2005. Discrete event multi-level models for systems biology. *Lect Notes Comput Sc* 3380:66-89.
28. Vayttaden, S. J., S. M. Ajay, and U. S. Bhalla. 2004. A spectrum of models of signaling pathways. *Chembiochem* 5:1365-1374.
29. Sachs, K., D. Gifford, T. Jaakkola, P. Sorger, and D. A. Lauffenburger. 2002. Bayesian network approach to cell signaling pathway modeling. *Sci STKE* 2002:pe38.
30. Saez-Rodriguez, J., L. G. Alexopoulos, J. Epperlein, R. Samaga, D. A. Lauffenburger, S. Klamt, and P. K. Sorger. 2009. Discrete logic modelling as a means to link protein signalling networks with functional analysis of mammalian signal transduction. *Molecular Systems Biology* 5.
31. Morris, M. K., J. Saez-Rodriguez, D. C. Clarke, P. K. Sorger, and D. A. Lauffenburger. 2011. Training Signaling Pathway Maps to Biochemical Data with Constrained Fuzzy Logic: Quantitative Analysis of Liver Cell Responses to Inflammatory Stimuli. *PLoS Comput. Biol.* 7.
32. Ironi, L., and L. Panzeri. 2009. A computational framework for qualitative simulation of nonlinear dynamical models of gene-regulatory networks. *BMC Bioinformatics* 10:14.
33. Yi, M., K. L. Xia, and M. Zhan. 2010. Theoretical study for regulatory property of scaffold protein on MAPK cascade: A qualitative modeling. *Biophysical Chemistry* 147:130-139.
34. Sauro, H. M., and B. N. Kholodenko. 2004. Quantitative analysis of signaling networks. *Prog. Biophys. Mol. Biol.* 86:5-43.
35. Andrews, S. S., T. Dinh, and A. P. Arkin. 2009. Stochastic models of biological processes. In *Encyclopedia of Complexity and System Science*. R. Meyers, editor. Springer, New York. 8730-8749.
36. Gillespie, D. T. 2007. Stochastic simulation of chemical kinetics. In *Annual Review of Physical Chemistry*. Annual Reviews, Palo Alto. 35-55.
37. Gillespie, D. T. 1977. Exact stochastic simulation of coupled chemical reactions *Journal of Physical Chemistry* 81:2340-2361.
38. Boulianne, L., S. Al Assaad, M. Dumontier, and W. J. Gross. 2008. GridCell: a stochastic particle-based biological system simulator. *BMC Syst. Biol.* 2:9.
39. Bray, D., R. B. Bourret, and M. I. Simon. 1993. Computer-Simulation of the Phosphorylation Cascade Controlling Bacterial Chemotaxis. *Molecular Biology of the Cell* 4:469-482.
40. Gardiner, C. 2009. *Stochastic Methods: A Handbook for the Natural and Social Science*. Springer.
41. Gillespie, D. T. 2000. The chemical Langevin equation. *J. Chem. Phys.* 113:297-306.
42. Lemerle, C., B. Di Ventura, and L. Serrano. 2005. Space as the final frontier in

- stochastic simulations of biological systems. *Febs Letters* 579:1789-1794.
43. Paulsson, J., O. G. Berg, and M. Ehrenberg. 2000. Stochastic focusing: Fluctuation-enhanced sensitivity of intracellular regulation. *Proc. Natl. Acad. Sci. USA* 97:7148-7153.
 44. Goldbeter, A., and D. E. Koshland. 1981. An amplified sensitivity arising from covalent modification in biological systems. *Proc. Natl. Acad. Sci. USA* 78:6840-6844.
 45. Berg, O. G., J. Paulsson, and M. Ehrenberg. 2000. Fluctuations and quality of control in biological cells: Zero-order ultrasensitivity reinvestigated. *Biophysical Journal* 79:1228-1236.
 46. Thattai, M., and A. van Oudenaarden. 2002. Attenuation of noise in ultrasensitive signaling cascades. *Biophysical Journal* 82:2943-2950.
 47. Bhalla, U. S. 2004. Signaling in small subcellular volumes. II. Stochastic and diffusion effects on synaptic network properties. *Biophysical Journal* 87:745-753.
 48. Bhalla, U. S. 2004. Signaling in small subcellular volumes. I. Stochastic and diffusion effects on individual pathways. *Biophysical Journal* 87:733-744.
 49. Morishita, Y., T. J. Kobayashi, and K. Aihara. 2006. An optimal number of molecules for signal amplification and discrimination in a chemical cascade. *Biophysical Journal* 91:2072-2081.
 50. Koepl, H., C. Zechner, A. Ganguly, S. Pelet, and M. Peter. 2012. Accounting for extrinsic variability in the estimation of stochastic rate constants. *Int. J. Robust. Nonlinear Control*
 51. Cornish-Bowden, A. 1995. *Fundamentals of Enzyme Kinetics*. Portland Press, London.
 52. Klipp, E., R. Herwig, A. Kowald, C. Wierling, and H. Lehrach. 2005. *Systems Biology in Practice*. Wiley-VCH, Berlin.
 53. Heinrich, R., and S. Schuster. 1996. *The Regulation of Cellular Systems*. Chapman & Hall, New York.
 54. Gunawardena, J. 2008. Models in systems biology: the parameter problem and the meanings of robustness. In *Elements of Computational Systems Biology*. H. Lodhi, and S. Muggleton, editors. John Wiley and Sons, New York.
 55. Ellner, S. P., and J. Guckenheimer. 2006. *Dynamic Models in Biology*. Princeton University Press, Princeton.
 56. Saez-Rodriguez, J., A. Kremling, H. Conzelmann, K. Bettenbrock, and E. D. Gilles. 2004. Modular analysis of signal transduction networks. *IEEE Control Systems Magazine* 24:35-52.
 57. Saez-Rodriguez, J., A. Kremling, and E. D. Gilles. 2005. Dissecting the puzzle of life: modularization of signal transduction networks. *Comput. Chem. Eng.* 29:619-629.
 58. Saltelli, A., M. Ratto, T. Andres, F. Campolongo, J. Cariboni, D. Gatelli, M. Saisana, and S. Tarantola. 2008. *Global Sensitivity Analysis: The Primer*. John Wiley & Sons, Chichester.
 59. Nocedal, J., and S. J. Wright. 1999. *Numerical Optimization*. Springer-Verlag, New York.
 60. Chen, W. W., B. Schoeberl, P. J. Jasper, M. Niepel, U. B. Nielsen, D. A. Lauffenburger, and P. K. Sorger. 2009. Input-output behavior of ErbB signaling pathways as revealed by a mass action model trained against dynamic data. *Molecular Systems Biology* 5:19.
 61. Schwacke, J. H., and E. O. Voit. 2005. Computation and analysis of time-dependent sensitivities in generalized mass action systems. *Journal of Theoretical Biology* 236:21-38.
 62. van Riel, N. A. W. 2006. *Dynamic modelling and analysis of biochemical networks:*

- mechanism-based models and model-based experiments. *Briefings in Bioinformatics* 7:364-374.
63. Yue, H., M. Brown, J. Knowles, H. Wang, D. S. Broomhead, and D. B. Kell. 2006. Insights into the behaviour of systems biology models from dynamic sensitivity and identifiability analysis: a case study of an NF-kappa B signalling pathway. *Mol. Biosyst.* 2:640-649.
 64. Zak, D. E., J. Stelling, and F. J. Doyle. 2005. Sensitivity analysis of oscillatory (bio)chemical systems. *Comput. Chem. Eng.* 29:663-673.
 65. Zi, Z. K., K. H. Cho, M. H. Sung, X. F. Xia, J. S. Zheng, and Z. R. Sun. 2005. In silico identification of the key components and steps in IFN-gamma induced JAK-STAT signaling pathway. *Febs Letters* 579:1101-1108.
 66. Banga, J. R. 2008. Optimization in computational systems biology. *BMC Syst. Biol.* 2.
 67. Chaves, M., E. D. Sontag, and R. J. Dinerstein. 2004. Gains and optimal design in signaling pathways. In *43rd IEEE Conference on Decision and Control*. IEEE, San Diego, CA. 596-601.
 68. Kikuchi, S., D. Tominaga, M. Arita, K. Takahashi, and M. Tomita. 2003. Dynamic modeling of genetic networks using genetic algorithm and S-system. *Bioinformatics* 19:643-650.
 69. Adiwijaya, B. S., P. I. Barton, and B. Tidor. 2006. Biological network design strategies: discovery through dynamic optimization. *Mol. Biosyst.* 2:650-659.
 70. Biegler, L. T., A. M. Cervantes, and A. Wachter. 2002. Advances in simultaneous strategies for dynamic process optimization. *Chem. Eng. Sci.* 57:575-593.
 71. Goh, C. J., and K. L. Teo. 1988. Control parametrization: a unified approach to optimal control problem with general constraints. *Automatica* 24:3-18.
 72. Zi, Z. 2011. Sensitivity analysis approaches applied to systems biology models. *IET Syst. Biol.* 5:336-346.
 73. Gutenkunst, R. N., J. J. Waterfall, F. P. Casey, K. S. Brown, C. R. Myers, and J. P. Sethna. 2007. Universally sloppy parameter sensitivities in systems biology models. *PLoS Comput. Biol.* 3:1871-1878.
 74. Breierova, L., and M. Choudhari. 2001. An Introduction to Sensitivity Analysis. Massachusetts Institute of Technology, Cambridge.
 75. Turanyi, T. 1990. Sensitivity Analysis of Complex Kinetic Systems - Tools and Applications. *J. Math. Chem.* 5:203-248.
 76. Cho, K. H., S. Y. Shin, W. Kolch, and O. Wolkenhauer. 2003. Experimental design in systems biology, based on parameter sensitivity analysis using a Monte Carlo method: A case study for the TNF alpha-mediated NF-kappa B signal transduction pathway. *Simulation-Transactions of the Society for Modeling and Simulation International* 79:726-739.
 77. Chu, Y., A. Jayaraman, and J. Hahn. 2007. Parameter sensitivity analysis of IL-6 signalling pathways. *IET Syst. Biol.* 1:342-352.
 78. Zhang, Y., and A. Rundell. 2006. Comparative study of parameter sensitivity analyses of the TCR-activated Erk-MAPK signalling pathway. *Syst Biol (Stevenage)* 153:201-211.
 79. Sobol, I. M. 2001. Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. *Math. Comput. Simul.* 55:271-280.
 80. Chachuat, B. 2006. Nonlinear and dynamic optimization: From theory to practice. EPFL, Lausanne.
 81. Boyd, S. P., and L. Vandenberghe. 2004. Convex optimization. Cambridge University Press, New York.
 82. Kamien, M. I., and N. L. Schwartz. 1992. Dynamic Optimization: The Calculus of Variations and Optimal Control in Economics and Management. Elsevier North

- Holland Amsterdam and New York.
83. Cardenas, M. L., and A. Cornishbowden. 1989. Characteristics necessary for an interconvertible enzyme cascade to generate a highly sensitive response to an effector. *Biochemical Journal* 257:339-345.
 84. Krebs, E. G., and J. A. Beavo. 1979. Phosphorylation-dephosphorylation of enzymes. *Annu. Rev. Biochem.* 48:923-959.
 85. Stadtman, E. R., and P. B. Chock. 1977. Superiority of interconvertible enzyme cascades in metabolic regulation: analysis of monocyclic systems. *Proc. Natl. Acad. Sci. USA* 74:2761-2765.
 86. Goldbeter, A., and D. E. Koshland. 1987. Energy-expenditure in the control of biochemical systems by covalent modification. *Journal of Biological Chemistry* 262:4460-4471.
 87. Cardenas, M. L., and A. Goldbeter. 1999. The glucose-induced switch between glycogen phosphorylase and glycogen synthase in the liver: outlines of a theoretical approach. *Journal of Theoretical Biology* 198:143-143.
 88. Cohen, P. 1982. The role of protein phosphorylation in neuronal and hormonal control of cellular activity. *Nature* 296:613-620.
 89. Sauro, H. M. 2004. The computational versatility of proteomic signaling networks. *Current Proteomics* 1:67-81.
 90. Ortega, F., and L. Acerenza. 1998. Optimal metabolic control design. *Journal of Theoretical Biology* 191:439-449.
 91. Varon, R., E. Valero, M. Molina-Alarcon, F. Garcia-Canovas, F. Garcia-Molina, M. E. Fuentes, and M. Garcia-Moreno. 2006. Expressions for the fractional modification in different monocyclic enzyme cascade systems: Analysis of their validity tested by numerical integration. *Bulletin of Mathematical Biology* 68:1461-1493.
 92. Angeli, D., J. E. Ferrell, and E. D. Sontag. 2004. Detection of multistability, bifurcations, and hysteresis in a large class of biological positive-feed back systems. *Proc. Natl. Acad. Sci. USA* 101:1822-1827.
 93. Kholodenko, B. N. 2000. Negative feedback and ultrasensitivity can bring about oscillations in the mitogen-activated protein kinase cascades. *European Journal of Biochemistry* 267:1583-1588.
 94. Bluthgen, N., F. J. Bruggeman, S. Legewie, H. Herzel, H. V. Westerhoff, and B. N. Kholodenko. 2006. Effects of sequestration on signal transduction cascades. *Febs J.* 273:895-906.
 95. Millat, T., E. Bullinger, J. Rohwer, and O. Wolkenhauer. 2007. Approximations and their consequences for dynamic modelling of signal transduction pathways. *Math. Biosci.* 207:40-57.
 96. Rieger, T. R. 2005. Mathematical modeling of the eukaryotic heat shock response and associated ultrasensitive signaling cascades. Northwestern University, Evanston IL.
 97. Legewie, S., B. Schoeberl, N. Bluthgen, and H. Herzel. 2007. Competing docking interactions can bring about bistability in the MAPK cascade. *Biophysical Journal* 93:2279-2288.
 98. Ventura, A. C., J. A. Sepulchre, and S. D. Merajver. 2008. A hidden feedback in signaling cascades is revealed. *PLoS Comput. Biol.* 4:14.
 99. Pearson, G., F. Robinson, T. B. Gibson, B. E. Xu, M. Karandikar, K. Berman, and M. H. Cobb. 2001. Mitogen-activated protein (MAP) kinase pathways: Regulation and physiological functions. *Endocrine Reviews* 22:153-183.
 100. Neary, J. T. 1997. MAPK cascades in cell growth and death. *News in Physiological Sciences* 12:286-293.

101. Jonak, C., E. Heberle-Bors, and H. Hirt. 1994. MAP kinases: universal multi-purpose signaling tools. *Plant Mol.Biol.* 24:407-416.
102. Wilkinson, M. G., and J. B. A. Millar. 2000. Control of the eukaryotic cell cycle by MAP kinase signaling pathways. *FASEB Journal* 14:2147-2157.
103. Roux, P. P., and J. Blenis. 2004. ERK and p38 MAPK-activated protein kinases: a family of protein kinases with diverse biological functions. *Microbiology and Molecular Reviews* 68:320–344.
104. Garrington, T. P., and G. L. Johnson. 1999. Organization and regulation of mitogen-activated protein kinase signaling pathways. *Current Opinion in Cell Biology* 11:211-218.
105. Kolch, W. 2005. Coordinating ERK/MAPK signalling through scaffolds and inhibitors. *Nature Reviews Molecular Cell Biology* 6:827-837.
106. Morrison, D. K., and R. J. Davis. 2003. Regulation of map kinase signaling modules by scaffold proteins in mammals. *Annu. Rev. Cell Dev. Biol.* 19:91-118.
107. Sebolt-Leopold, J. S., and R. Herrera. 2004. Targeting the mitogen-activated protein kinase cascade to treat cancer. *Nat. Rev. Cancer* 4:937-947.
108. Shvartsman, S. Y., M. Coppey, and A. M. Berezhkovskii. 2009. MAPK signaling in equations and embryos. *Fly* 3:62-67.
109. Kolch, W., M. Calder, and D. Gilbert. 2005. When kinases meet mathematics: the systems biology of MAPK signalling. *Febs Letters* 579:1891-1895.
110. Orton, R. J., O. E. Sturm, V. Vyshemirsky, M. Calder, D. R. Gilbert, and W. Kolch. 2005. Computational modelling of the receptor-tyrosine-kinase-activated MAPK pathway. *Biochemical Journal* 392:249-261.
111. Huang, C. Y. F., and J. E. Ferrell. 1996. Ultrasensitivity in the mitogen-activated protein kinase cascade. *Proc. Natl. Acad. Sci. USA* 93:10078-10083.
112. Heinrich, R., B. G. Neel, and T. A. Rapoport. 2002. Mathematical models of protein kinase signal transduction. *Molecular Cell* 9:957-970.
113. Widmann, C., S. Gibson, M. B. Jarpe, and G. L. Johnson. 1999. Mitogen-activated protein kinase: Conservation of a three-kinase module from yeast to human. *Physiological Reviews* 79:143-180.
114. Liu, X. F., L. Bardwell, and Q. Nie. 2010. A combination of multisite phosphorylation and substrate sequestration produces switchlike responses. *Biophysical Journal* 98:1396-1407.
115. Ferrell, J. E. 2002. Self-perpetuating states in signal transduction: positive feedback, double-negative feedback and bistability. *Current Opinion in Cell Biology* 14:140-148.
116. Asthagiri, A. R., and D. A. Lauffenburger. 2001. A computational study of feedback effects on signal dynamics in a mitogen-activated protein kinase (MAPK) pathway model. *Biotechnology Progress* 17:227-239.
117. Chachuat, B. 2008. Simulation and optimization collection (SOC). *Personal communication*
118. Gill, P. E., W. Murray, and M. A. Saunders. 2005. SNOPT: An SQP algorithm for large-scale constrained optimization (Reprinted from SIAM Journal Optimization, vol 12, pg 979-1006, 2002). *SIAM Rev.* 47:99-131.
119. Feehery, W. F., J. E. Tolsma, and P. I. Barton. 1997. Efficient sensitivity analysis of large-scale differential-algebraic systems. *Applied Numerical Mathematics* 25:41-54.
120. Tolsma, J., and P. I. Barton. 2000. DAEPACK: An open modeling environment for legacy models. *Ind. Eng. Chem. Res.* 39:1826-1839.
121. Eissing, T., S. Waldherr, F. Allgower, P. Scheurich, and E. Bullinger. 2007. Steady state and (bi-) stability evaluation of simple protease signalling networks. *Biosystems* 90:591-601.
122. Park, T., and P. I. Barton. 1996. State event location in differential-algebraic models.

- ACM Transactions on Modeling and Computer Simulation* 6:137–165.
123. Ferrell, J. E. 1996. Tripping the switch fantastic: How a protein kinase cascade can convert graded inputs into switch-like outputs. *Trends in Biochemical Sciences* 21:460-466.
 124. Legewie, S., N. Bluthgen, and H. Herzl. 2005. Quantitative analysis of ultrasensitive responses. *Febs J.* 272:4071-4079.
 125. Lipshtat, A., G. Jayaraman, J. C. He, and R. Iyengar. 2010. Design of versatile biochemical switches that respond to amplitude, duration, and spatial cues. *Proc. Natl. Acad. Sci. USA* 107:1247-1252.
 126. Soyer, O. S., H. Kuwahara, and A. Csikasz-Nagy. 2009. Regulating the total level of a signaling protein can vary its dynamics in a range from switch like ultrasensitivity to adaptive responses. *Febs J.* 276:3290-3298.
 127. Hooshangi, S., S. Thiberge, and R. Weiss. 2005. Ultrasensitivity and noise propagation in a synthetic transcriptional cascade. *Proc. Natl. Acad. Sci. USA* 102:3581-3586.
 128. Schmidt, H., and M. Jirstrand. 2006. Systems Biology Toolbox for MATLAB: a computational platform for research in systems biology. *Bioinformatics* 22:514-515.
 129. Hucka, M., A. Finney, H. M. Sauro, H. Bolouri, J. C. Doyle, H. Kitano, A. P. Arkin, B. J. Bornstein, D. Bray, A. Cornish-Bowden, A. A. Cuellar, S. Dronov, E. D. Gilles, M. Ginkel, V. Gor, Goryanin, I., W. J. Hedley, T. C. Hodgman, J. H. Hofmeyr, P. J. Hunter, N. S. Juty, J. L. Kasberger, A. Kremling, U. Kummer, N. Le Novere, L. M. Loew, D. Lucio, P. Mendes, E. Minch, E. D. Mjolsness, Y. Nakayama, M. R. Nelson, P. F. Nielsen, T. Sakurada, J. C. Schaff, B. E. Shapiro, T. S. Shimizu, H. D. Spence, J. Stelling, K. Takahashi, M. Tomita, J. Wagner, J. Wang, and S. Forum. 2003. The systems biology markup language (SBML): a medium for representation and exchange of biochemical network models. *Bioinformatics* 19:524-531.
 130. Birtwistle, M. R., M. Hatakeyama, N. Yumoto, B. A. Ogunnaik, J. B. Hoek, and B. N. Kholodenko. 2007. Ligand-dependent responses of the ErbB signaling network: experimental and modeling analyses. *Molecular Systems Biology* 3:16.
 131. Hu, D. W., and J. M. Yuan. 2006. Time-dependent sensitivity analysis of biological networks: Coupled MAPK and PI3K signal transduction pathways. *Journal of Physical Chemistry A* 110:5361-5370.
 132. Wolkenhauer, O., M. Ullah, W. Kolch, and K. H. Cho. 2004. Modeling and simulation of intracellular dynamics: Choosing an appropriate framework. *IEEE Trans. Nanobiosci.* 3:200-207.
 133. Bhalla, U. S., and R. Iyengar. 1999. Emergent properties of networks of biological signaling pathways. *Science* 283:381-387.
 134. Brightman, F. A., and D. A. Fell. 2000. Differential feedback regulation of the MAPK cascade underlies the quantitative differences in EGF and NGF signalling in PC12 cells. *Febs Letters* 482:169-174.
 135. Fujioka, A., K. Terai, R. E. Itoh, K. Aoki, T. Nakamura, S. Kuroda, E. Nishida, and M. Matsuda. 2006. Dynamics of the Ras/ERK MAPK cascade as monitored by fluorescent probes. *Journal of Biological Chemistry* 281:8917-8926.
 136. Hatakeyama, M., S. Kimura, T. Naka, T. Kawasaki, N. Yumoto, M. Ichikawa, J. H. Kim, K. Saito, M. Saeki, M. Shirouzu, S. Yokoyama, and A. Konagaya. 2003. A computational model on the modulation of mitogen-activated protein kinase (MAPK) and Akt pathways in heregulin-induced ErbB signalling. *Biochemical Journal* 373:451-463.
 137. Schoeberl, B., C. Eichler-Jonsson, E. D. Gilles, and G. Muller. 2002. Computational modeling of the dynamics of the MAP kinase cascade activated by surface and

- internalized EGF receptors. *Nat. Biotechnol.* 20:370-375.
138. Fitzgerald, J. B., B. Schoeberl, U. B. Nielsen, and P. K. Sorger. 2006. Systems biology and combination therapy in the quest for clinical efficacy. *Nat. Chem. Biol.* 2:458-466.
 139. Schoeberl, B., E. A. Pace, J. B. Fitzgerald, B. D. Harms, L. H. Xu, L. Nie, B. Linggi, A. Kalra, V. Paragas, R. Bukhalid, V. Grantcharova, N. Kohli, K. A. West, M. Leszczyniecka, M. J. Feldhaus, A. J. Kudla, and U. B. Nielsen. 2009. Therapeutically targeting ErbB3: A key node in ligand-induced activation of the ErbB receptor-PI3K axis. *Sci. Signal.* 2:14.
 140. McDonagh, C. F., A. Huhlov, B. D. Harms, S. Adams, V. Paragas, S. Oyama, B. Zhang, L. Luus, R. Overland, S. Nguyen, J. M. Gu, N. Kohli, M. Wallace, M. J. Feldhaus, A. J. Kudla, B. Schoeberl, and U. B. Nielsen. 2012. Antitumor activity of a novel bispecific antibody that targets the ErbB2/ErbB3 oncogenic unit and inhibits heregulin-induced activation of ErbB3. *Mol. Cancer Ther.* 11:582-593.
 141. Bos, J. L. 1989. RAS oncogenes in human cancer - a review. *Cancer Res.* 49:4682-4689.
 142. Sargent, R. W. H. 2000. Optimal control. *J Comput Appl Math* 124:361-371.
 143. Bersani, A. M., E. Bersani, and L. Mastroeni. 2011. Modeling the action of drugs on cellular enzymes by means of optimal control techniques. *J. Math. Chem.* 49:776-795.
 144. Varigonda, S., T. T. Georgiou, R. A. Siegel, and P. Daoutidis. 2008. Optimal periodic control of a drug delivery system. *Comput. Chem. Eng.* 32:2256-2262.
 145. Sutherland, W. J. 2005. The best solution. *Nature* 435:569-569.



ANDRIJANA RADIVOJEVIC

✉ **anries@gmail.com**
 ☎ +41 (0) 78 925 95 97
 🏠 Chemin des Cèdres 1
 CH-1004 Lausanne, Switzerland

- PhD in Systems Biology, MSc in Electrical and Computer Engineering.
 - Strong analytical skills.
 - Dynamic person with excellent social skills suited for teamwork.

EDUCATION

- Jan 2008
 August 2012

 ●

PhD degree
 École Polytechnique Fédérale de Lausanne, Switzerland
 Automatic Control Laboratory & Laboratory for Computational Systems Biotechnology
 PhD Thesis: *Optimal design of signaling modules: key players, trade-offs and sustainability*
 Building mechanistic models of signal transduction pathways that play a very important role in major developmental changes in organisms and whose malfunction is associated with diseases, such as cancer, asthma and diabetes; identifying system design principles; using computational analysis as entry point for experimental design tailored to signaling and for parameter identification in real systems.
- Oct 2002
 Nov 2007

 ●

B.Sc. / M.Sc. degree
 Faculty of Technical Sciences, University of Novi Sad, Serbia
 School of Electrical and Computer Engineering, Department of Computing and Control
 Master Thesis: *Development of application software for controlling a home automation system based on Windows CE*

WORK EXPERIENCE

- Jan 2008
 July 2012

 ●

Research assistant
 École Polytechnique Fédérale de Lausanne, Switzerland
 Teaching: one Bachelor and one Master level course.
 Other: presentations at international conferences, publication in international peer-review journal, supervised Bachelor and Master level student thesis and projects.
- July 2011
 Dec 2011

 ●

intern
 Merrimack Pharmaceuticals, Cambridge, MA, USA
 Contributions: statistical mRNA data analysis, development of mechanistic mathematical models of large scale signaling pathways, model fitting, parameter estimation.
- July 2007
 Oct 2007

 ●

intern
 ION Solutions, doo., Novi Sad, Serbia
 Contributions: Master Thesis: design and implementation of configurable "Smart Home" system, implementation of X10 communication protocol.
- July 2006
 Oct 2006

 ●

intern
 Alcad, doo., Slovenska Bistrica, Slovenia
 Contributions: design and implementation of functional modules for information system for the client company, system internationalization.

SKILLS OVERVIEW

Programming C++, C, C#, MATLAB, Fortran, Java, Basic.
Applied mathematics Nonlinear dynamics, dynamic optimization, optimal control, parametric sensitivity analysis, stochastic modeling and simulations.

LANGUAGES

English (proficient), French (independent), German (basic), Serbian (mother tongue).

ADDITIONAL INFORMATION

• 2007 - One of 16 participants of Serbia in the one month long project "European Rails of Peace", sponsored by an Italian NGO. The project agenda included peace-building and conflict-resolution workshops related to the recent history of the Balkans.

Other Interest

- Dancing and fitness.
- Music: studied piano.
- Reading: psychology.

PERSONAL INFORMATION

28 years old, single, Serbian nationality.