Humboldt-Universität zu Berlin
Fakultät für Lebenswissenschaften
Institut für Biologie

Information Processing in Cellular Signaling

Dissertation

zur Erlangung des akademischen Grades

doctor rerum naturalium (Dr. rer. nat.)

im Promotionsfach Biophysik

eingereicht von: Friedemann Uschner
geboren am:
geboren in:
Präsident der Humboldt-Universität zu Berlin:
Prof. Dr. Jan-Hendrik Olbertz

Dekan der Humboldt-Universität zu Berlin:
Prof. Dr. Richard Lucius

Gutachter/innen: Prof. Dr. Edda Klipp
Prof. Dr. Hanspeter Herzel
Prof. Dr. Nils Blüthgen

eingereicht am: 04.05.2015 verteidigt am: 22.10.2015

Abstract

Information plays a ubiquitous role in nature. It provides the basis for structure and development, as it is inherent part of the genetic code. But it also enables organisms to make sense of their environments and react accordingly. For this, a cellular interpretation and measure of information is needed. Cells have developed sophisticated signaling mechanisms to fulfill this task and integrate many different external cues with the help of such.

Here we focus on signaling processes that sense osmotic stress (the High Osmolarity Glycerol (HOG) pathway) as well as α-factor stimulation (the pheromone pathway) in the model organism *S.cerevisiae*. We employ stochastic modeling that simulates the inherent noisy nature of biological processes to assess how signaling systems process the information they receive. This information transmission is evaluated with an information theoretic approach by interpreting signal transduction as an information transmission channel in the sense of Shannon.

We use channel capacity to both constrain as well as quantify the fidelity in the phosphorelay system of the HOG pathway. In this model, simulated with the Stochastic Simulation Algorithm by D.T. Gillespie, the analysis of signaling behavior allows us to constrain the possible parameter sets for the system severely. A further approach to signal processing and integration is concerned with the mechanisms that conduct crosstalk between the HOG and the pheromone pathway. We find that the control for signal specificity lies especially with the scaffold proteins that tether signaling components and facilitate signaling by trans-location to the membrane and shielding against miss-activation. As conserved motifs of cellular signal transmission, these scaffold proteins show a particularly well suited structure for accurate information transmission. In the last part of this thesis, we examine the potential reasons for an evolutionary selection of the scaffolding structure. We show that solely due to its structural mechanisms, scaffolds are increasing information transmission fidelity and outperform a distributed signal in this regard as well as their robustness.

The presented analyses provide a structural view on signal processing in cellular systems and combine several mathematical methods with one another.

Keywords: *S.cerevisiae*, cellular signaling, HOG pathway, pheromone pathway, scaffolding, systems biology, information theory, stochastic modeling, chemical master equation, moment closure

Zusammenfassung

"Information" spielt in der Natur eine zentrale Rolle. Als intrinsischer Teil des genetischen Codes ist sie das Grundgerüst jeder Struktur und ihrer Entwicklung. Aber im Speziellen dient sie auch Organismen, ihre Umgebung wahrzunehmen und sich daran anzupassen. Die Grundvoraussetzung dafür ist, dass sie Information ihrer Umgebung sowohl messen als auch interpretieren können, wozu Zellen komplexe Signaltransduktionswege entwickelt haben, mit denen sie diese Aufgabe erfüllen können.

In dieser Arbeit konzentrieren wir uns auf Signalprozesse in *S.cerevisiae* die von osmotischem Stress (über den High Osmolarity Glycerol (HOG) Signalweg) und der Stimulation mit α-Faktor (Pheromon Signalweg) angesprochen werden. Wir wenden stochastische Modelle an, die das natürlich innewohnende Rauschen biologischer Prozesse darstellen können, um verstehen zu können wie Signalwege die ihnen zur Verfügung stehende Information umsetzen. Informationsübertragung wird dabei mit einem Ansatz aus Shannons Informationstheorie gemessen, indem wir Signaltransduktion als einen Kanal in diesem Sinne auffassen.
Wir verwenden das Maßder Kanalkapazität, um die Genauigkeit des Phosphorelays innerhalb des HOG Signalweges messen und einschränken zu können. In diesem Modell, welches mit dem Gillespie Algorithmus simuliert wurde, können wir durch die Analyse des Signalverhaltens den Parameterraum zusätzlich bereits stark einschränken. Eine weitere Herangehensweise der Signalverarbeitung und -integration beschäftigt sich mit dem Mechanismus des "Crosstalks" zwischen HOG und Pheromon Signalweg. Wir zeigen, dass die Kontrolle der Signalspezifizität vor allem bei Scaffold-Proteinen liegt, die Komponenten der Signalkaskade binden und bündeln und über Lokalisierung an der Zellmembran die Signalübertragung ermöglichen sowie vor falscher Aktivierung schützen. Diese konservierten Motive zellulärer Signaltransduktion besitzen eine geeignete Struktur, um Information getreu übertragen zu können. Im letzten Teil der Arbeit untersuchen wir potentielle Gründe für die evolutionäre Selektion der Scaffold Proteine. Wir zeigen, dass ihnen bereits durch die Struktur des Mechanismus möglich ist, Informationsübertragungsgenauigkeit zu verbessern und einer verteilten Informationsweiterleitung sowohl dadurch als auch durch ihre Robustheit überlegen sind.

Die hier vorgestellten Analysen bieten eine strukturelle Sicht auf Signalprozesse in zellulären Systemen und kombinieren verschiedene mathematische Methoden dafür miteinander.

Stichwörter: *S.cerevisiae*, zelluläre Signaltransduktion, HOG Signalweg, Pheromon Signalweg, Scaffolding, Systembiologie, Informationstheorie, stochastische Modellierung, chemical master equation, moment closure

Contents

1. Introduction

1.1. Motivation

Stochasticity in organisms and their environments has in the recent past more and more become a focus of scientific research as a key to heterogeneity and diversity in biology. Stochasticity is considered a driving force in the evolutionary "design" of biological systems, as handling it is an important functionality that such designs have to fulfill in order to be conserved. No organism can survive if it cannot deal with the natural fluctuations in its environments and in addition, every cell is "born" into its own intrinsic stochasticity of varying molecule numbers and inherently stochastic biological processes. Yet, even though this imperative noisy nature of biological mechanisms is all-abundant and well established, it has been largely avoided in previous research. To a certain degree, this can be attributed to the experimental techniques and computational power available in the past, as only recently we see more and more studies aiming their focus on observing single cells and molecules instead of population means (Spiller *et al.*, 2010). This development not only reflects our fast technical advancement and growing understanding of fundamental processes, but also opened up a variety of topics for deeper research in both experimental as well as theoretical work. Studying noise has been an urgent undertaking especially when low copy numbers and stochastic events are under research, like for example in the case of gene expression (McAdams and Arkin, 1997; Swain *et al.*, 2002; Elowitz *et al.*, 2002), cellular signaling (Samoilov *et al.*, 2005), but also for considerations of evolutionary developments (Rivoire and Leibler, 2011, 2014). Since we have been provided with huge datasets containing whole genomes and individual cell behaviors within populations, understanding this stochasticity has proven to be of great importance and fostered many new insights and implications. This development can be seen in many fields of molecular biology nowadays and is closely coupled to theoretical advances.[1]

[1] A very interesting testament to that is the citation metrics of the paper Gillespie (1977), whose popularity has risen only in the last decade in a very impressive fashion due to the higher attention of stochastical theories and simulations in biology that came especially with the data produced by experimental research, but also to some extend with the availability of increased computational power and theoretical developments for model development (e.g. Daigle *et al.*, 2012).

Fluctuations and variations span over a large diversity of scales and types and additionally their source, leverage point as well as contribution, is not always apparent in systems as complex as biological organisms. Neither is the discrimination between pure (biological) noise and vital information that fluctuations can potentially carry. Yet, this distinction plays a particularly important role when it comes to cellular signaling. External conditions vary naturally and information about those environmental cues, stimuli and stresses is relayed via sophisticated networks to evoke appropriate adaptations in the organism. Signaling networks have evolved to function reliably despite their fluctuating environments and to overcome their stochastic encoding and transmission properties, thus achieving a natural separation of noise and vital information. In a fascinating fashion, they manage to regulate a flow of information that is both specific as well as sufficiently accurate and allows the decision centers within the cell to infer what was sensed in the first place. In recent years, theoretical research has taken this matter of separation up (i.e. Bowsher and Swain, 2012) and is incorporating it into a better fundamental understanding of information flow and function. As cell signaling and the failing thereof is closely connected to cancer development and also its therapeutic treatment (Pfizer *et al.*, 2006; Bianco *et al.*, 2006; Hanahan and Weinberg, 2011), many scientific efforts have been aimed at understanding its basic principles, building blocks and functions. Creating knowledge on how this information flow is directed and what roles stochasticity plays within it adds an important dimension to this research. Our grasp of biology and also our capacities for creating and analyzing biological data has come to a point where it is possible and advisable to take this additional dimension into account as well. Understanding, step by step, how systems as complicated as cellular organisms (including ourselves) achieve robustness and function, will let us take many more steps forward in scientific research. As can be learned from economic and financial sciences, in large systems exhibiting a high degree of complexity it is important to regard more than just mean approximations, as some variations can have critical and unpredictable effects on the whole system.

In the work at hand, we are exploring the notion of “information” in biolog-

ical systems using the example of signaling in *Saccharomyces cerevisiae*. We aim at studying effects of structural properties of cellular mechanisms on the processing of stochasticity and the flow of information.

1.2. Signaling in Systems Biology - Information processing in biological systems

As with many applied fields of science, the development of biological research has over the centuries been closely coupled with the technical advances in physics and engineering. Although biology was mainly on the receiving end, there has been a strong back and forth between these fields. Technical advances delivered new methods for the experimental research of biological organisms and in turn, the high demand for this fed back into technical refine- and advancement as well as the development of theoretical analysis. This let to a point where it became necessary (and possible) to consider not only parts of molecular machinery, but also connections and interactions between them, creating a "systems view" of biological processes. Thus, the idea of "systems biology" was developed in order to focus on such a level of understanding by incorporating knowledge as well as data from many different experimental fields of biology (Klipp *et al.*, 2013; Alon, 2006; Kitano, 2002). One particular aim of this is to use mathematical frameworks and bio-informatic analysis in order to find and test hypotheses on structures as well as functions and create a loop back again to experimental studies to validate, refine and extend these hypotheses. In that way, systems biology brings together and connects the scientific fields of biology, mathematics, physics, chemistry and informatics. And although this involves tremendous amounts of communication between and joint understanding of the different sciences, this circle has on many occasions shown to be fruitful in developing new insights into biological systems.

A key point in this research has been to elucidate the connections between different machineries within cell systems. It has been shown that there exist designs and motifs that occur frequently and fulfill many functions in different contexts as well as organisms (Milo *et al.*, 2002; Yeger-Lotem *et al.*, 2004; Alon, 2007; Legewie *et al.*, 2008). Identifying them and exploring their potential cre-

ates a modular view on biology that places detailed research in a much broader context (Hartwell *et al.*, 1999). Much in analogy to the proof of "Fermat's last theorem" by Andrew Wiles (Wiles, 1995), the emphasize is placed on ways to exploit connections between the different fields in order to reach a higher goal. This also means that concepts and hypotheses along this way will have to be refined and understood in greater detail. Recent advances (Karr *et al.*, 2012; Sanghvi *et al.*, 2013; Macaulay *et al.*, 2015; Azeloglu and Iyengar, 2015) show that our growing understanding of how to do this integration enables further research, making a system more than the sum of its parts. And finally, this approach presents us with integral benchmarks on what we know and what we cannot comprehend yet, thus suggesting where to look next. This feeds back to experimental research and closes the mentioned "cycle of systems biology".[2]

One particularly important part of such a modular view on biology is the understanding of cellular signaling networks. They play a crucial role in the inter-connections between modules and incorporate many recurring motifs themselves, making them a highly integrated network. Their main function, however, is to link the cell's behavior to the outside environment. Many receptors act as sensors to encode and transmit information to cell decision centers, making them an integral part for survival and proliferation. They enable cells to adapt to stresses and other environmental cues, like growth factors, nutrient availability or signals on a population level. At the same time, their vital role also puts them at a high risk of causing malfunction of the organism if not working properly. Mutations causing cancer development (Hanahan and Weinberg, 2011; Berg *et al.*, 2002, chap. 15) have been predominantly associated with signaling pathways and attempts for treatments thereof are thus often aimed at restoring their functionality (Levitzki and Klein, 2010). Of course, this requires a deep understanding of how the motifs in question actually function. Regulation patterns like feedbacks (Klinger and Bluethgen, 2014), feed-forward loops (Mangan and Alon, 2003), crosstalk between pathways (Natarajan *et al.*,

[2]Exemplary for this kind of cycle is the joint work by the labs of Jens Timmer and Ursula Klingmüller, who engaged in many fruitful collaborations showing that with proper communication and mutual understanding of theory and experiment, this field of research provides progress that otherwise would not be possible in a more isolated approach (Faller *et al.*, 2003; Schilling *et al.*, 2005; Raue *et al.*, 2009).

2006) and many further mechanisms make this a non-trivial, yet important task.

Adding to this complexity is the fact that the environmental cues to be relayed are stochastic. Thus, the gathered information is noisy by nature and since biochemical signal transduction depends on diffusion and reactions, also the transmission of this information will be distorted. While structural research on molecular signaling modules has been conducted already for more than a decade, incorporating stochastical considerations into them is a more recent and very exciting development and is especially driven by the already mentioned advancements in experimental biology. As with a modular view of biology, the scope of such research varies immensely and can employ a zoom-out approach, going from simple modeling and the understanding of basic structural implications to more complex systems integrating such smaller modules (Powell, 2004).

Cell signaling is in need of processing capabilities that regulate and integrate the available information, yet how and to what extend this is done by biological systems is unclear. Applying a stochastic systems biology approach to this information processing of cell signaling networks grants access to a more comprehensive and integrated view of how fluctuations and variations are handled and additionally allows to apply evolutionary arguments to the selection of such structures Rivoire and Leibler (2014).

Yet, as we try to understand biological function in more and more detail, we have to restrict ourselves to basic principles and manageable systems upon which we build knowledge. Like mathematic proofs building on one another, biological sciences have advanced step by step to learn more (not without stumbling, yet always with progress). With the advent of molecular biology we have seen different model organisms arise and while we are venturing into human applications, those biological models are still teaching us fundamental principles upon which we can build.

1.3. *Saccharomyces cerevisiae* - A model organism for systems level science

Biological organisms are of immense complexity. Comparing theirs to the complexity of even the largest of human designed technical systems would still not do them justice. Especially since in that case, we ourselves are the designers and set the rules that elude our view in a biological "design by evolution". And even though within this enormous diversity and heterogeneity we find principles and recurring mechanisms that govern functionality, most of the time our observations are merely glimpses of the real biological truths and hidden behind convoluted causes and effects. So how can we learn and advance in the face of such overwhelming complexity? In mathematics, a well-built structure advances from a established statements and assumptions further to extended implications. Biologists share this approach to some extend, by building knowledge on a well-known, controlled and manageable environment, at least within our possibilities. Such environments in the age of molecular biology are model organisms like for example *Drosophila melanogaster*, *Caenorhabditis elegans* or (most notably) *Saccharomyces cerevisiae*, each used for different purposes in experimental science. While we are already venturing into human research and applications, those biological models are still teaching us fundamental principles upon which we can build. In the case of systems biology, *S.cerevisiae* ("bakers yeast", see Fig. 1) has proven to be one of the most important model organisms (Botstein and Fink, 2011).

The knowledge obtained from experimental as well as theoretical research of the eukaryotic *S.cerevisiae* spans a vast amount of biological science. More than 80% of its genome is annotated, generating large amounts of available knockouts and especially bridging the connection between genes and their function that is needed to understand broader contexts. In addition, it shares many regulatory features and mechanisms with more complex eukaryotes[3]. Moreover, the organism can be handled easier than others due to the relatively low doubling time around 100 minutes and is robust to many perturbations. All this enables the basis for a systems approach that is unparalleled by other

[3] About 1000 genes were shown to have ortholog families of genes mammalian cells, that are connected to human diseases (Heinicke *et al.*, 2007).

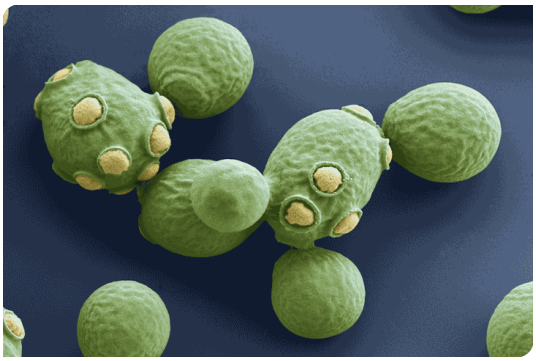


Fig. 1: The model organism *S. cerevisiae*, (Eye Of Science / Science Photo Library, accessed at `http://thebeerdiaries.tv/full-genome-sequencing-yeast/`)

organisms. Even though it is immensely important to eventually take the obtained knowledge to another level and research the mammalian counterparts with all their similarities and differences, the basic knowledge we obtain by studying yeast is still vital to scientific understanding of cellular functions and also will be for some time to come.

Within this work, we will be focusing on signal transmission in *S.cerevisiae* to utilize gathered knowledge in order to apply stochastic approaches and regulatory elements of signaling and test their implications.

1.4. Information in Biology

The notion of "information" can be ambiguous and generally depends on the broad context of its interpretation. This makes it hard to infer from an outside view what the fundamental semantics in a certain setting are. This is particularly true for biology. In a designed technical system it is possible to pre-define what alphabets one is working with and how valid messages are composed. In analogy to a key exchange protocol in cryptography, those agreements are subject to a larger context of a communicated agreement between sender and receiver. Unfortunately, it is impossible to know what these terms are exactly in a biological system and how this information is perceived and utilized by cells. There are many ways, in which an organism can encode, store and transmit information. The most obvious way is presented in its genetic

code in form of nucleotide sequences on DNA. It carries the information vital for everything: On how to facilitate development, growth, decision making, proliferation and actually, with the latter it also carries the information for itself[4]. Yet, this is regulated (in fact with the regulators being part of the code themselves as well) by incorporating information transduced by a complex system of protein signaling pathways. Information is stored and encoded in concentrations, gradients thereof, molecule conformations, activation states, action potentials and a plethora of further means.[5] This simply illustrates that "information" is all abundant and the building block for life.

To us, the meaning of information in a cell and the weighting of its importance for the biological functionality is hidden. This cellular interpretation and especially the extend to which information is also neglected by an organism (for example as a way of optimizing a balance of energy or robustness, see also the discussion in Voliotis *et al.* (2014)) is part of evolutionary development and it is important to keep in mind that this selection process concerns both signal transduction (the message) as well as responses (Smith, 2000). We, as the third party trying to understand this inherent biological information processing, can only make observations on the responses to such a communication and integration of signals and interpret them with what we already know about the internal processes. Experimental setups can control the environment to a certain extend and thus enable us to modulate the external informational cue. The experiment further leaves the interpretation to the organism and evaluates only the changes in its behavior of the chosen observables. We might not be able to grasp the full extend of the blackbox in between this signal and output, yet there is a way of inferring how information was processes in it. Here, the work of C.E. Shannon is particularly appealing. His "information theory" provides the tools to quantify the extend and boundaries of such a processing. Because of this, it was employed in many interesting applications in biological signaling mechanisms. One important aspect of the theory is that it disregards semantics. This means that any encoding or decoding can

[4] Yet, not in any paradox way that Bertrand Russell could use to cause chaos to the card house of biological research.

[5] As shown in Selimkhanov *et al.* (2014) and Tostevin and Ten Wolde (2009), many encodings make use of temporal profiles of such means.

be considered, the only prerequisite is that our approaches provide sufficient statistics. Mere processing of the data does not change information theoretical measures (they are invariant to transformations) and thus presents us with an objective way to measure and constraints information transmission. Yet when applying the theory, we have to be aware that we see upper boundaries and constrains on signaling. This in itself can elucidate many questions, but going into semantics is sometimes still needed as we cannot tell how much information is merely disregarded by the cell. Nevertheless, the application to many biological studies has shown that the numbers for such constraints are sometimes lower than would be expected (Voliotis *et al.*, 2014), which has two implications: Firstly, information processing can be adjusted to incorporate a certain amount of randomness. This could for example be less energy consuming than more sophisticated signaling mechanisms and thus be chosen even with the risk of faulty sensing. This consideration explains low capacities and argues that this boundaries is exhausted and optimal. A second implication can be, that we were looking for the wrong features. The observables we choose for studying biological information have to be chosen carefully. As mentioned earlier, an encoding can for example map stress strength to temporal profile of a downstream species. Evaluating such a structure only by using one point in time will not represent the full information that can be transmitted. At this point it is important to obtain enough knowledge on structure, behavior and basic functionality of the system in question and thus answer the question of the underlying semantics. Information theory provides a benchmark for what we observe and an important tool for refinement. Applying it to information in biology can be rewarding as it gives a valid measure for it, but should be done with careful consideration.

As in noisy signal transmission, uncertainty always remains in what we know about biological systems, but as we progress in our understanding of the underlying principles, we also push the boundaries of producing applicable knowledge. We need to reach a point, where the information contained within a cell can first be entangled, then interpreted, and ultimately also manipulated to be used to our advantage, for example in clinical applications.

1.5. Scope and aim of the work

Within this work, we explore the capabilities of cellular signaling mechanisms to conduct and process information from a structural point of view. We employ both stochastic and deterministic modeling to study their functionality with regard to how they enable fidelity of signaling despite the noise that is inherent in their biochemical processes and how they adapted evolutionary to still preserve function and can conduct information reliably to decision centers - even in the case of designs with multiple inputs and multiple outputs. The merging of structural research and stochastic concepts has only recently been gaining increased attention. Here we aim to study design principles that integrate both ideas to gain further insight into cellular signaling and model construction thereof.

The interpretation of signaling is closely connected to information theoretical concepts. The framework plays a key role in our work, as it enables us to quantify information processing, put constraints on what a system is capable of transmitting reliably and presents us with an optimality criterion that can be used to discriminate. We introduce information theory as a field that has many appealing applications in biology and a large potential to facilitate more. We believe that while its application can be tricky at times, it still provides many new insights.

Chapter 3 is focusing on a conserved two-component signaling system, the phosphorelay in the Sln1 branch of the High Osmolarity Glycerol (HOG) signaling pathway. This module is employed in many prokaryotic organisms and shows a robust behavior, even if noise is introduced in the system. It is capable of transducing a graded response to changes in osmolarity already (and especially) at this first stage of signaling. As the response in Hog1 double phosphorylation shows distinct temporal time courses and features to different stress strengths, it is possible that this diversity has its origin not only in feedback mechanisms and adaptational programs, but instead can potentially already arise in the encoding of the input. We extend this modeling approach and interpretation as a transmitting channel in the sense of information theory by using the measure in a parameter space restricting fashion. Viable signaling

and information transmission capabilities define how our system can be modeled.

Chapter 4 embeds this signaling motif into a broader context and regards the whole network of pheromone and HOG signaling. We study the effects of crosstalk as observed in a study by Vaga *et al.* (2014) with a deterministic structural model. We reexamine the experimental data and suggest a modeling approach to test a consensus view on the signaling pathway. The emphasize is put on how crosstalk can be prevented and find that the control in this specificity lies with scaffolding proteins. We reach the conclusion that the knowledge obtained in earlier studies is not sufficient to explain the novel findings of the dataset and further investigations are required for a more comprehensive understanding.

The last chapter 5 is concerned with reinterpreting the evolutionary development of a central role that scaffold proteins play in cellular signaling. This motif tethers signaling molecules to itself and has been shown to fulfill a plethora of different functions (as also seen in chapter 4). We discuss the idea that in order to develop a structure as abundantly used as scaffolding, yet with sometimes very low similarities between particular scaffold proteins, the incentive has to be based on a structural advantage to other motifs (for example with higher binding affinities between signaling species instead of a tethering). Many of the functions observed nowadays seem to be of secondary nature. As a strong potential candidate for these evolutionary selected primary features, we suggest an optimal and increased fidelity compared to non-scaffolded signaling. Furthermore we present the use of multiple channels and the integration of molecule concentrations downstream as a redundant coding for achieving higher fidelity in signal inference. With this, cells are able to reliably judge their environment and can adapt themselves accordingly.

Although increasingly focused on, single cell data of signaling systems providing a sufficiently good statistics to approximate probability distributions is still sparse. But since especially experimental techniques improve at a fast rates nowadays, we believe that this kind of data will increase over the next decades and potentially become a standard approach as it covers both single cell

dynamics or properties as well as the integrated population level by averaging over these. The availability of such data opens up applications of stochastic modeling and information theoretical analysis. With the work at hand we hope to provide a fundamental understanding of the application of these concepts to probability distributions of biological species and the assessment of signal processing for considerations in future research.

2. Mathematical background

Within this section we will provide the reader with the mathematical fundamentals and details to enable a better comprehension of the frameworks used in the following chapters. We will introduce the mathematical descriptions that are applied to biological settings, in our case information processing mechanisms in *S.cerevisiae*. As a center piece for analyzing information transmission we will familiarize the reader with C.E. Shannon's "Information Theory" and some of its applications in section 2.1. In order to make use of this framework, mathematics for stochastic modeling are needed and will be established subsequently in section 2.2. This includes the "chemical master equation" as a way to describe the development of probability distributions of states in a system as well as two important practical ways to realize this development: the "Stochastic Simulation Algorithm" (SSA) by Gillespie (1977) and the method of "moment closure". Both allow us to build models that can be analyzed using information theoretic measures.

2.1. Information Theory - A framework to quantify information processing

Over the last century, two scientific fields have showed a particularly impressive development. Starting out with the landmark paper of J.Watson and F.Crick (Watson *et al.*, 1953) that described the double-helix structure of DNA, molecular biology took a very impressive and fast-paced path within this century. In a similarly short time window after the second world war, computer sciences were formed[6] and, built upon mathematics and physics, matured with the advances in engineering into one of the most important scientific as well as applied fields. As in molecular biology, one single contribution played a particularly key role in defining and driving the progression of the field: Claude E. Shannon's fundamental paper (Shannon, 1948), that described the complete framework of "Information Theory". It was developed by Shannon in 1948 while working at Bell Laboratories - the very famous research and development department

[6] Of course, as with molecular biology, not without building on predecessors.

of a telephone company - to describe and quantify information transmission as well as its boundaries. At that time it was of importance for telegraph and telephone communication, but later it became an important basis of computer science and our modern information society, by providing applications in data-compression, -encoding, -transmission, -processing or -correction ultimately enabling inventions like the world wide web, wireless communication, data encryption and countless more. Also, biological sciences benefited from the theory with several applications. With its numerous implications Shannon's paper is one of the most important and influential papers of the 20th century[7] and one of the most impressive legacies passed on by a single man.[8]

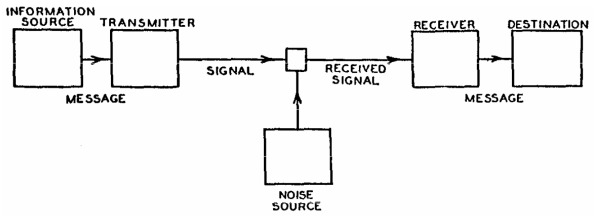


Fig. 2: "Schematic diagram of a general communication system" - Fig.1 from Shannon (1948)

While computer science thrived and built a sizable part of its advancement upon the invention of information theory, biology has only recently noticed the potential of the framework for its own applications.[9] Before going into the theory itself, we want to give a brief account of some of these interesting uses to motivate the upcoming derivations and guide the interested reader towards further information.

[7] One of the most cited as well.

[8] **Remark**: It also is interesting to know that Shannon's PhD-thesis was concerned with developing a framework for a biological setting (Shannon, 1940), namely a formal "algebra for theoretical genetics". The download is available from the MIT under `http://hdl.handle.net/1721.1/11174`. Even before that, he developed the basis for digital computing in his master thesis (Shannon, 1938). In the 60s, he invented one of the first portable computers and with it applied his information theory to gambling, even going through the risks of testing it in a casino in Las Vegas at that time (Thorp, 1998). Unfortunately, Shannon probably never fully experienced this legacy, as he later suffered from Alzheimer's disease.

[9] In fact, the early advances even considered the applications very skeptical (Johnson, 1970).

A very important use of information theoretical measures developed in the early 90s when T.D. Schneider defined "sequence logos" (Schneider and Stephens, 1990). Since then, they have become standard in displaying the information encoded in sequence alignments. This representation interprets the DNA as a code that contains information, where each position can encode 2 *bit*, as the possible characters at each position are A,G,C and T. Aligning such sequences (for example around the translation initiation site "START") gives probability distributions for each position of the binding site, depending on the conservation of the sequence. Thus, Schneider found a way of describing how much information e.g. a transcription factor would gain (and need, respectively) through binding. This has revolutionized the earlier used notion of "consensus sequences", where only the most likely sequence was used. Instead, Schneider's logos (based on the measured probabilities) show a much more detailed and correct way of possible bindings. Nowadays, this is employed in many applications besides the pure representation of binding sites (e.g. Schneider, 2001; Lyakhov *et al.*, 2008).

An important use of information theory is without doubt its measure of mutual information that quantifies how much one random variable can tell about the other and thus, naturally provides a measure of (even non-linear) correlation. The applications for that are manifold. In biology, especially network inference and classification techniques (Butte and Kohane, 2000; Zheng and Kwoh, 2006; Liang *et al.*, 1998; Qiu *et al.*, 2009; Slonim *et al.*, 2005) as well as applications in neuro-science (Borst and Theunissen, 1999; Tkacik *et al.*, 2010; Dimitrov *et al.*, 2011), but also studies on signaling networks (Tostevin and Ten Wolde, 2009) have made extensive use of that.

Besides the many theoretical advances, the application of information theory still encounters restrictions in experimental work. Measuring probability distributions with single cell data of living organisms is no simple task and data sets only started to show the quality and especially quantity needed for a good approximation. Very important advances to that were provided by the work of Cheong *et al.* (2011a) and more recently Selimkhanov *et al.* (2014). Whereas data in genomics and neuro-science already provide a good basis for working with probabilities, studies in cell signaling and metabolism are only slowly catching up to that. Then again, this merely means that many opportunities

still lie ahead and it will be exciting to see the development in the near future.

Last but not least, we want to mention a very thorough and rigorous study by Rivoire and Leibler (2011), that provides perspectives and conceptual work on how cells perceive their environment and integrate this into decision-making and optimizations in terms of theoretical long-term growth rates. This work has far reaching implications for the further development of the topic, even though it only creates unspecific concepts in many aspects. Together with the recent more technical follow-up (Rivoire and Leibler, 2014), it provides a comprehensive insight into how the value of information in varying environments is processed and integrated.

This is only a small and incomplete account of research employing the use of information theory. For further reviews and interesting applications of the theory, we refer to the diverse literature on the topic: Waltermann and Klipp (2011); Adami (2004, 2012); Battail (2005); Rhee *et al.* (2012); Schneider (2005); Vinga (2013).

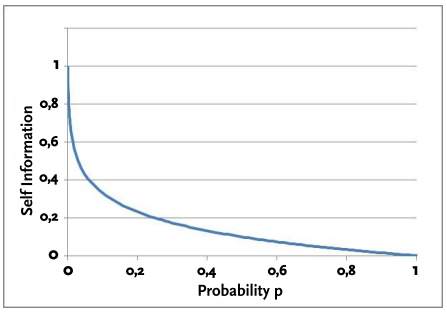
2.1.1. Uncertainty and Mutual Information - Building blocks of signal processing

To quantify the transmission of information within biological cell signaling, it is possible to consider biochemical signaling pathways as noisy channels in the sense of Information Theory as developed by Shannon (1948). This probabilistic mathematical framework provides "channel capacity" as a measure that can be used to evaluate how well different input signals are still distinguishable after the signal has been transduced. We aim to quantify and evaluate the system's capabilities of transmitting information by observing its ability to respond to certain inputs in the presence of noise. In a technical setting, this is the limit to which messages can be transmitted reliably. It is important to keep in mind that with capacity we can set an upper bound on information transmission. The biological implications however can be very complex and possibly even include the neglect of information. Nevertheless, it is (not without reason) assumed that biological systems typically evolve by optimizing efficiencies of certain biological functions. This could be an objective function like for

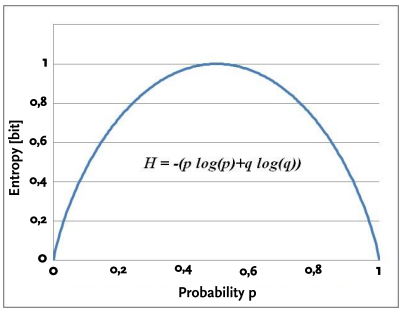
example a growth rate or cell proliferation. Bacterial organisms even combine the two with a "bet-hedging" strategy of phenotype switching that finds a balance between growth and pure survival of the population (Balaban *et al.*, 2004). Organisms thus often work in near-optimal regimes, leading to the idea of using this optimization for the study of biological principles (Rosen, 1986; Parker *et al.*, 1990). We use this argument in the way that we regard signaling pathways as evolutionary optimized (and thus fixed as a channel). Here we give an introduction to the main concepts of the framework, embedding them into biological settings. For more detailed information we refer to Cover and Thomas (2012) and Brillouin (2013).

Let X be a discrete random variable associated with the corresponding probabilities $P(x)$ of the events $x \in X$. Each of these events will be associated with a function called "surprisal" or "self-information" (see Fig. 3a), which is the negative logarithm of its probability: $S(x \in X) = -\log(P(x))$. This is based on the following intuition: The occurrence of rare event would be both surprising as well as a potent carrier for information. Imagine letters in an alphabet and the words composed from them: Rare letters (e.g. z, j, q, x for the English language) in a word narrows down the number of possible words immensely even without knowing all other letters, whereas the single occurrence of frequent letters (e.g. e, t, a, o) still leaves many choices and thus contains less self-information. A certain event ($P(x) = 1$) is never "surprising" and thus, its occurrence carries no information at all. As for the usage of the logarithm, Shannon noted that (with referring to Hartley, 1928) it is the natural choice for a measure of information in states, since most importantly it scales linearly with many processes in engineering as well as nature, not to mention its handiness in mathematical calculations that would otherwise need more complicated statements.

As a suitable representation for measuring "information, choice and uncertainty", Shannon deduced the so called *(Shannon) entropy* H, as visualized in Fig. 3b. For this, he defined important properties of our (intuitive) understanding of information and identified H as the only function satisfying these:



(a) The measure of "surprisal" or "self-information" of an event $x \in X$, dependent on the probability $P(x)$.



(b) Entropy of a random variable $X = \{x_1, x_2\}$ with the associated probability distribution $P(x_1) = p$ and $P(x_2) = q = 1 - p$.

Fig. 3: Fundamental measures for quantifying "information" and "uncertainty".

$$H(X) = -K \sum_{x \in X} P(x) \cdot \log P(x) \tag{1}$$

where $P(\cdot)$ is the associated probability distribution of the random variable X (Shannon, 1948, Appendix II). The constant K represents a choice of unit for the entropy. Following the general convention, we measure entropies in *bits*, referring to a base of 2 for the logarithm ($K = 1/\log(2)$). Intuitively 1 *bit* could be visualized by the toss of a fair coin: The equal probability of "heads" and "tails" as the outcome of the toss gives us two equally good choices for predicting the toss, reflecting the uncertainty about the random variable measured by the entropy. Manipulating the coin to favor one outcome will lower our uncertainty for the prediction (and thus the choice we are likely to make), but also the information that could be gathered by tossing the coin provided we know the probability distribution. Applying this to a set of possible external state variables for a cell (such as nutrients, temperature or in our case osmotic conditions), we get a measure of how uncertain our environment is and how informative measuring it will be.

As discussed before, "information" can be a very elusive term since from an outside point of view, interpreting the actual meaning of messages can

be riddling. For example, imagine an encrypted message being interrupted and read by a person who doesn't know the corresponding deciphering code. Yet, in the design of systems it can be defined and classified what the term will mean. This requires to agree on basic assumptions. For one, we need a common alphabet. This not only means characters or signs, but also their usage in the messages. This comes down to their probability of occurrence, stemming from the rules for building messages (words, sentences, codes, etc.). A typical example would be a language. Shannon gave a vivid illustration for his theory by calculating the measure of entropy (and thus redundancy) for different languages.

The notion of entropy as a prerequisite for describing information can be extended in a natural way to *conditional entropy* by using conditional probability distributions. The following concept then connects two random variables to one another and states the setup for successive definitions that conceptualize communication from a sender to a receiver. Consider two (not necessarily independent) random variables X and Y. We can measure our average "uncertainty" about Y when knowing X as:

$$H(Y|X) = -\sum_{x\in X, y\in Y} P(x,y)\cdot \log_2 P(y|x) \tag{2}$$

where $P(X,Y)$ denotes the joint distribution and $P(Y|X)$ the conditional distribution of Y given X. Eq. (2) measures the entropy of the output when the input is known. This can be used to define *mutual information* (MI), a measure commonly used to quantify how much information one random variable carries about the other.[10]

$$I(X;Y) = H(X) - H(X|Y) \tag{3}$$
$$= H(Y) - H(Y|X) \tag{4}$$

For our purposes, another (although equivalent) interpretation of mutual information is more convenient. Using Eq. (4), we can describe it as "the amount of information received less the uncertainty that still remains due to the noise

[10]Note that this is symmetric by definition. One can be transformed into the other by applying Bayes theorem in the definition of (conditional) entropy.

in the system". This scenario can be directly applied to the situation that we have in an experimental setup. One would evaluate a noisy output (e.g. the fluorescence of a tagged protein) as the response to a certain stimulus (e.g. stress level, input dose). Once the probability distribution for the (natural) input has been defined or inferred (experimentally a nearly impossible task), we can evaluate how much information our output still contains about the input, despite the inherent noise. The conditional distribution $P(Y|X)$ for such a scenario defines rules for the transmission through this biochemical channel.

The most descriptive use of mutual information is that of a measure for correlation between two random variables. Assuming a connection between the two (i.e. our channel), one can tell to which degree they are entwined and since we are working with probability distributions, mutual information even offers the advantage of covering non-linear correlations (in contrast to other standard correlation measures[11]). Besides other advantages like the invariance to transformations (see for example the representation and processing of biological data), this is an important reason for the application of MI as the metric of interest in the inference of gene networks (Butte and Kohane, 2000; Liang *et al.*, 1998; Qiu *et al.*, 2009) or even in clustering methods (Slonim *et al.*, 2005). Even though mutual information does not solve the problem of causality, it is a very useful and versatile tool due to its unique properties.

The next natural step is to maximize this mutual information. If our (biological) system is given, its fixed transmitting properties (conveyed by $P(Y|X)$) might have evolved to be optimally adapted to a certain input distribution and perhaps a weighting on how important the reaction to a certain level of stress is. This amounts to the formal definition of "channel capacity", namely

$$C = \max_{P(X)} I(X;Y).$$

This optimization is performed with respect to the input distribution. Nevertheless, it is important to note that this could also be done by optimizing the transmission probabilities themselves. But as already mentioned we assume these to be evolutionary optimized and thus as a fixed quantity characterizing the channel. Shannon defined in his paper how a general communication system

[11] A nice description of that can be found in Tkačik (2010).

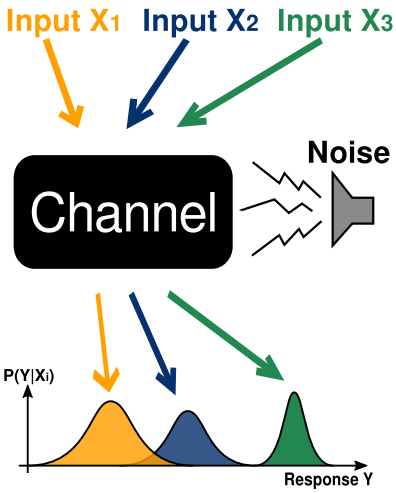


Fig. 4: Schematic diagram of the biochemical channel as it would be used in an experimental setup. Distinct inputs (e.g. environmental cues) are encoded and transmitted through the channel. Due to noise added during the process of transmission, we can only observe conditional probability distributions as the received signal.

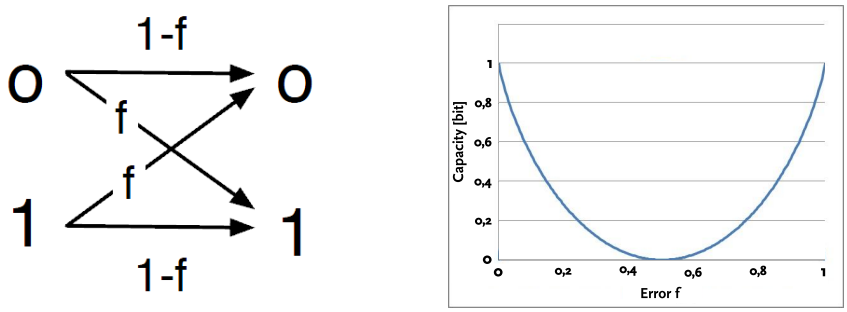
could look like (see Fig. 2). A biological experiment can be seen as an adapted representation of this concept of "channel", as visualized in Fig. 4.

In the following, we give two simple examples visualizing the approach to that.

Example 2.1. Binary Symmetric Channel (BSC, see also Cover and Thomas (2012))

The BSC (Fig.5) consists of two inputs $\{0, 1\}$ mapping over a (noisy) channel to two outputs $\{0, 1\}$. The error-probability is defined as f for the faulty transmission of both inputs, hence the term "symmetric". The matrix comprised of the conditional transmission probabilities $Q_{Y|X} = P(Y = y_j|X = x_i)$ that define our channel looks as follows:

$$\mathbf{Q}(\text{Output}|\text{Input}) = \begin{pmatrix} P(\text{off}|\text{off}) & P(\text{off}|\text{on}) \\ P(\text{on}|\text{off}) & P(\text{on}|\text{on}) \end{pmatrix} = \begin{pmatrix} 1-f & f \\ f & 1-f \end{pmatrix}$$



(a) Conditional transmission probabilities in the BSC.

(b) Capacity of the BSC as a function of the error f.

Fig. 5: Schematic representing the Binary Symmetric Channel. The channel capacity is a function of the error f.

Thus the capacity of the BSC can be computed as

$$\begin{aligned} C &= \max_{P(X)} I(X;Y) = \max_{P(X)} [H(Y) - H(Y \mid X)] \\ &= H(Y) - \max_{P(X)} \left[\sum_j P(X = x_j) H(Y \mid X = x_j) \right] \\ &= H(Y) - [(1-f)\log_2(1-f) + f\log_2(f)] \\ &= H(Y) - H(f) = 1 - H(f), \end{aligned}$$

giving a function of f as visualized in Fig.5. Intuitively, the capacity is maximal if the error probability is either 1 or 0, allowing for perfect inference. It is vanishing, when the correct input is sent in exactly half of the cases. No error correction (e.g. with a repetition code) can then achieve any information transmission. The signal is completely lost.

Example 2.2. The Noisy Typewriter (Cover and Thomas, 2012)

The "Noisy Typewriter" (see Fig. 6) can be described as a keyboard being so small that its keys are hard to hit accurately. Typing a symbol (for convenience we chose 27 different ones) will send either the selected or one of the adjacent characters with a probability of 1/3, meaning that the message will be distorted severely. The capacity can be calculated as

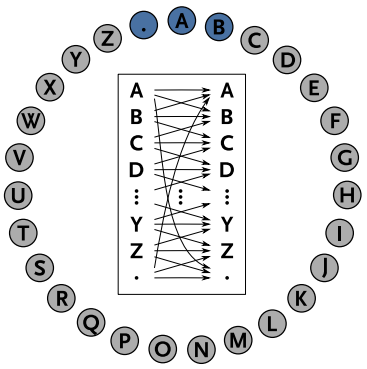


Fig. 6: The "Noisy Typewriter" types the correct letter with a probability of 1/3 and one of the neighboring characters with the same probability.

follows:

$$\begin{aligned} C &= \max_{P(X)} I(X;Y) \\ &= \max_{P(X)} H(X) - H(X \mid Y) \\ &= \max_{P(X)} H(X) - \log(3) \\ &= \log(27) - \log(3) \;=\; \log(9). \end{aligned}$$

Even if the noise in the channel for transmitting a message in such a way is preventing an accurate inference of the input, it is interesting to note that we can still select a subset of our input that can be transmitted reliably. Such a subset could be if we chose to only use every third key, leaving out the intermediate ones. A message can then be properly decoded after the transmission, provided the decoder is aware of this encoding decision. In a way, most channels look like the noisy typewriter providing a mapping from a distinct subset of inputs to distinct subsets of outputs (see also Shannon, 1948, Fig.10).

Within this work we apply these measures to the biological setting of signal transduction. What serves as an "encoder" and "decoder" has to be defined by

the respective system and is subject to the scope of the investigated mechanism. For example, an obvious way would be to regard the cellular receptors as encoders for external concentrations of ligands, transforming this information into a message that can be transmitted over the pathway. As in the case of mitogen-activated protein kinases (MAPK) cascades, this could mean a subsequent activation via phosphorylation of the next tier of proteins. The output level is then definable as being any of the downstream protein species. A sensible choice would be a species that provides good experimental access and a meaningful interpretation, like for example a downstream transcription factor that "interprets" the activation message by binding to the DNA and initiating the transcriptional responses.

Of course many such choices are possible. Instead of an output, a transcription factor could also be seen as a hub and a converter that transforms the message into another alphabet (base-pairs of the DNA) and relays the information further, potentially like a broadcast channel to multiple receivers.[12] Interpreting DNA and RNAs as further messages and the cell's ribosomes as decoders would then be viable as well. Thus, when dealing with this we have to keep the scope of our system in mind and find appropriate definitions.[13] As intuition would suggest, a longer channel-"blackbox" like the one we just stated is not beneficial *per se*. After all, signal transduction will also always be governed by the data (signal) processing inequality, stating that information can only be lost: In a Markov-chain $X \to Y \to Z$, it is intrinsic that $I(X;Z) \leq I(X;Y)$. This strict bound can be found in many every-day scenarios (be it the sophisticated post-processing of a distorted message or simply a game of "Chinese whispers") and cannot be overcome by any signal en- or decoding. In an experimental application, this means in particular that the post-processing of experimental data can only reveal things that were in the data beforehand (in whatever hidden way). Bio-informatics and other theoretical analysis cannot (re-) add information about the original biological processes that was lost or not transmitted to data in the experiment, it can only sort and transform it in an intelligent way to a human readable format,

[12] After all, transcription factors in general target the expression of many genes.

[13] Sometimes, the experimental procedures will determine what measures can be used as an output level beforehand.

still leaving the task of interpreting the results.

One major advantage of Information Theory is that in order to characterize and evaluate a (biological) system, one does not need to consider all the details within it when interpreting it as a transmitting channel. This very general "blackbox"-approach has numerous implications. In experiments, where the input can be well-defined in addition to a proper output-statistic of the "channel", one can draw conclusions on function, structure and boundaries by using this theory (see e.g. Cheong *et al.* (2011b); Rhee *et al.* (2012)). In a biological setting, this can be interpreted as the pathway repeatedly sensing environmental conditions and stresses, which could for example be provided in experimental setups. Measuring this channel capacity then means answering the question "How much can the receiver of a noise-distorted message tell about what was originally sent by the encoder, given this particular transmitting channel". To a biological system this would translate to "How accurate can our cells respond to external cues, given a noisy signal transmission over the respective pathways". With the appropriate assumptions we can obtain knowledge of our blackbox in this way. Not only does this make sense in experimental work, but the same approach can be fruitful for a (stochastic) modeling approach.

2.1.2. Optimization of information transmission - Arimoto and Blahut at work

For finite inputs and the given transmission probabilities of the channel (represented by a fixed matrix), we can use a numeric optimization algorithm developed concurrently by Arimoto (1972) and Blahut (1972) to calculate the channel capacity (see also Cover and Thomas, 2012) and the input distribution that achieves this capacity. Again this quantity is given in *bits*, as in this way transmitting 1 *bit* corresponds to measuring an on/off response representing two states of the environment (switch-like behavior). Capacities above 1 *bit* enable a channel to reliably distinguish more than just "on" and "off" and react appropriately, leading to an input-specific response.

The optimization method uses an iterative approach to find the optimizing

input-distribution that achieves the capacity given fixed transition probabilities (i.e. the channel), as can be derived from Eq. (3). Going into the details of the proofs necessary for developing the theory behind the algorithm is out of scope for the work at hand. Nevertheless, to understand and discuss certain properties of our outcomes, it is helpful to understand the idea behind the Arimoto-Blahut optimization algorithm. In the following, we only aim to sketch the approach. For detailed mathematical proof of the statements, we refer to Blahut (1972) and Arimoto (1972).

For setting up the optimization approach of Arimoto and Blahut, a Lagrange Multiplier restricting the solution $P(x)$ to a valid set of input distributions is introduced. This means satisfying the normalizing conditions $\forall i : P(x_i) \geq 0$ and $\sum_i P(x_i) = 1$, given that we have a finite discrete number of inputs and outputs.[14] Thus, we set up the variational problem as

$$\mathcal{L}\left[P(x)\right] = \sum_{x,y} P(y|x)P(x)\log_2 \frac{P(y|x)}{P(y)} - \lambda \sum_x P(x), \tag{5}$$

where the channel $P(y|x)$ is fixed and the optimization is achieved by varying $P(x)$.[15] In the paper, Blahut developed the idea to reformulate the equation for mutual information and execute the maximization by adding an additional variational object (the backwards transmission channel $P(x|y)$):

$$C = \max_{P(x)} \mathcal{L}\left[P(x)\right] \sim \max_{P(x)} \max_{P(x|y)} \mathcal{L}'\left[P(x)\right], P(x|y)] \tag{6}$$

$$= \max_{P(x)} \max_{P(x|y)} \sum_{x,y} P(y|x)P(x)\log_2 \frac{P(x|y)}{P(y)} - \lambda \sum_x P(x), \tag{7}$$

giving us the achievable maximum mutual information, the capacity of the channel.

The integral part is to show how variational objects that maximize the term look like when the respective other variational object is held fixed. First, let

[14] In a continuous case, this would correspond to using an integral instead of the sum.

[15] Notice, that $P(y)$ in the denominator is a function of $P(x)$ as well, as can be seen by applying Bayes theorem.

$P(X)$ be fixed. R.Blahut came to the conclusion that

$$P(x|y) = \frac{P(x)P(y|x)}{\sum_x P(x)P(y|x)} \tag{8}$$

achieves this maximization merely by restating the (conditional) probabilities and showing, that the inserted backward channel leads to the same expression of mutual information.

As the next step, we fix this channel $P(x|y)$. Following the Lagrange multiplier approach, we take the normalizing constraint and set the partial derivatives of $\mathcal{L}'$ with respect to $P(x)$ equal to zero (thus finding an optimum):

$$\frac{\partial}{\partial P(x)} \left[\sum_x \sum_y P(x)P(y|x) \log(\frac{P(x|y)}{P(x)}) + \lambda \left(\sum_x P(x) - 1 \right) \right] = 0 \tag{9}$$

$$- \log P(x) - 1 + \sum_y P(y|x) \cdot \log P(x|y) + \lambda = 0 \tag{10}$$

$$\Rightarrow P(x) = \frac{\exp\left(\sum_y P(y|x) \log(P(x|y))\right)}{\sum_x \exp\left(\sum_y P(y|x) \log(P(x|y))\right)}, \tag{11}$$

with λ fulfilling the normalization. Now the approach is to iteratively use these two maximization steps to create a stepwise increase in mutual information until the increase falls below an error-threshold. This leaves to show that the obtained sequence is strictly monotone increasing and approaches the channel capacity. We omit that step here and again refer to the original literature (Blahut, 1972). The pseudo-code for the algorithm is visualized in Fig. 7.

Besides the channel capacity, the algorithm produces a second output: An optimal input distribution that achieves this capacity, given the channel.[16] This probability distribution often looks spike-like and seem unnatural from a biological perspective. A typical outcome can be seen in Fig. 8. This behavior is nothing unexpected though. It conveys the fact that for a given noisy channel, one can select a subset of inputs (i.e. the spikes) that achieves the capacity. This can easily be seen in example 2.2, where restricting the input to only every third character (the subset $A, D, G, ...$) would result in a perfect inference of the input after a transmission.[17] In a natural setting, it makes sense to assume that

[16] Note that there might be more than one that does so.

[17] This subject is also briefly discussed in Tkačik *et al.* (2008a) placed in a setting of gene

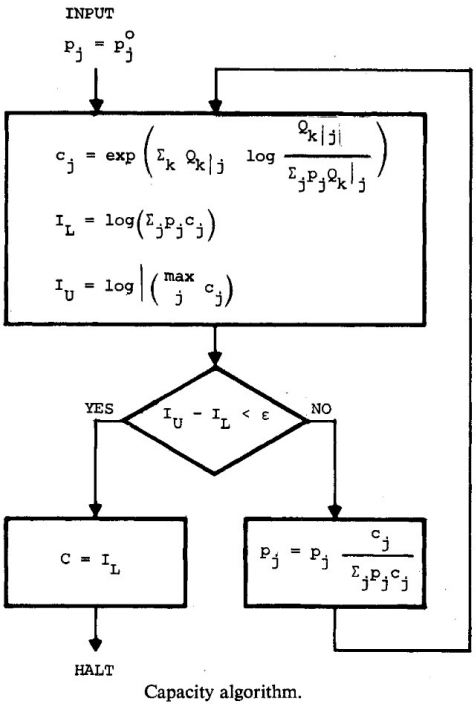


Fig. 7: Pseudo-code for the optimization Arimoto-Blahut-algorithm, taken from Blahut (1972).

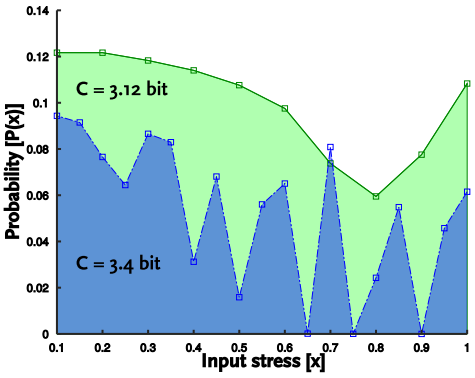


Fig. 8: Two optimal input distributions achieving capacity (3.12 *bit* (*green*) and 3.4 *bit* (*blue*), respectively) in an identical communication channel. The difference is the resolution of the input: The *green* channel bins the input into 10 distinct stresses, the *blue* channel employs a resolution of 19 stresses over the same range. We observe that both can distinguish roughly the same number of inputs after transmission, restricting the possible resolution for the input. *green* is thus a smoothed version of *blue*, able to transmit similar information.

these distributions would occur in a "smoothed" way, thus deviating slightly from achieving full capacity. This "smoothing" is visualized in Fig. 8 and shows how noise induces a maximal possible resolution that cannot be surpassed. At this point, it is interesting to note that to a large degree the capacity is determined by the channel transmission, whereas the input distribution adds less variation in capacity to that. This helps also in choosing an initial input probability distribution for the algorithm, as assigning it uniformly is usually a feasible choice and a good first approximation. Nevertheless, many biological processes show very specific and distinguished distributions, often in form of an on/off-signal or in a manner that allows a binning into characteristic states.

2.2. Stochastic Modeling

Key for information theoretical analyses of signal processing is a stochastical description of the processes in question. As can be seen from the equations of entropy and capacity, when thinking in terms of a channel transmitting informa-

expression noise.

tion we think in probabilities of its usage and noisy transmission probabilities, respectively. First, sender and receiver have knowledge about a (common) alphabet or an encoding thereof. For languages, this means characters (letters) and subsequently words and sentences composed of them, for computers this could mean numbers and analogously sequences of them depicting messages. Biological systems also employ various alphabets: As an apparent example, nucleotides composed in sequences on the DNA are used as an alphabet to code for messages in the form of mRNAs and subsequently proteins. Other examples are action potentials, concentrations and gradients thereof as well as many further means of providing, storing and transmitting the information of signals. This flow of information is omnipresent in nature and organisms cannot survive without a proper way to encode, conduct, decode and interpret it. This implies that, secondly, "knowing" the probability distributions of the signal's occurrence as well as how it is being transmitted is of importance as well. Biologically, this is subject to evolutionary adaptation, whereas technical systems are designed using this knowledge *a priori*. Yet, both have in common that they are subject to noise and variations. By describing cellular processes in stochastic terms, it is possible to capture this natural behavior more accurately, study the implications of variation and investigate how a separation between noise and information can be achieved. Ultimately, it enables us to apply an information theoretic framework for describing this cellular flow of information and its processing properties.

Although having been an important part of many sciences (especially in financial research) already for a very long time, the development of stochastic methods for modeling in biology has only seen an impressive uprise recently. This development is not surprising, it reflects the huge advancements that both experimental biology as well as our general knowledge of biological processes have achieved. Nowadays, there are many ways of using stochastic frameworks for biological settings. Here we first introduce the chemical master equation as a general way of describing the stochastic behavior of reaction systems. This equation is always applicable for a defined set of species and reactions, yet solving it is only possible for few biochemical systems and limited settings. Still, we believe that knowing the theoretical basis behind this representation gives

many good insights into how systems will behave as well as on how to construct them. Furthermore, it also creates awareness for the limitations of the approach. In a next step, the basic thoughts underlying the chemical master equation lead in a natural way to another stochastic description for the same biochemical reaction system. The Stochastic Simulation Algorithm developed by D.T. Gillespie will be introduced and later on applied in chapter 3, giving likely trajectories for the vector of species over time and subsequently a sampling of the probabilities in the master equation. As with the Gillespie-algorithm, the last section describes an important framework, namely the moment closure, with which we are able to approximate the distributions of the master equation by neglecting higher moments. This will be applied in chapter 5.

2.2.1. The chemical master-equation - A deterministic description of stochasticity

Conventional mathematical descriptions of chemical reaction systems are often put in terms of a deterministic framework describing the time-evolution of mean concentration values. This approach can answer many questions, yet in certain cases it fails to capture essential parts of the natural system behavior. This happens for example when species are sparse, the system exhibits unstable behavior (and thus critical points that are very susceptible to variation), when rare stochastic events pregnant with consequences are occurring (e.g. extinction of a species) but also, and especially, when both intrinsic and extrinsic noise plays a role in the system in question or its design. In particular, fluctuations of components and successively their covariances are neglected in such a framework. Due to the inherent noise that is the nature of biological processes, it is important to resort to other mathematical descriptions. Despite the fact that, as mentioned, the fundamental mathematics of stochastic approaches was developed quite some time ago, their application to chemical reaction systems and biological processes has only started much later in the middle of the 19^{th} century (see McQuarrie, 1967). One of the most important frameworks for such a description, embedded into the concept of Markov chains (Gillespie, 1991), is called the “chemical master equation” (CME). This equation can be interpreted as a “gain-loss” equation for the probabilities of the distinct states our system can occupy (Van Kampen, 2011). We will present the concept and structure

of the equation here. For detailed derivations and proofs, we refer the reader to standard literature on the topic: Gardiner (2009) and Van Kampen (2011) provide excellent introductions, deriving the master equation as a special case of the differential Chapman-Kolmogorov equation. We will follow the rigorous derivation as presented in Gillespie (1992), since the paper in based on the essential ideas that connect the master equation to the Stochastic Simulation Algorithm (SSA) described in the next section (Gillespie, 1977).

In the following, we use the nomenclature of bold symbols as the notation for vectors. We start off with a system of M chemical reactions R_m ($m = 1, \ldots, M$) involving a vector of N species $\boldsymbol{X}(t) = \{X_1(t), \ldots, X_N(t)\}$:

$$\begin{aligned}
\underline{\nu}_{11}X_1 + \underline{\nu}_{12}X_2 + \ldots + \underline{\nu}_{1N}X_N \quad &\xrightarrow{c_1} \quad \overline{\nu}_{11}X_1 + \overline{\nu}_{12}X_2 + \ldots + \overline{\nu}_{1N}X_N \\
\underline{\nu}_{21}X_1 + \underline{\nu}_{22}X_2 + \ldots + \underline{\nu}_{2N}X_N \quad &\xrightarrow{c_2} \quad \overline{\nu}_{21}X_1 + \overline{\nu}_{22}X_2 + \ldots + \overline{\nu}_{2N}X_N \\
&\vdots \\
\underline{\nu}_{m1}X_1 + \underline{\nu}_{m2}X_2 + \ldots + \underline{\nu}_{mN}X_N \quad &\xrightarrow{c_m} \quad \overline{\nu}_{m1}X_1 + \overline{\nu}_{m2}X_2 + \ldots + \overline{\nu}_{mN}X_N \\
&\vdots \\
\underline{\nu}_{M1}X_1 + \underline{\nu}_{M2}X_2 + \ldots + \underline{\nu}_{MN}X_N \quad &\xrightarrow{c_M} \quad \overline{\nu}_{M1}X_1 + \overline{\nu}_{M2}X_2 + \ldots + \overline{\nu}_{MN}X_N.
\end{aligned} \tag{12}$$

The coefficients $\underline{\nu}_{mn}$ and $\overline{\nu}_{mn}$ can be combined in the stoichiometric matrix $\nu_{mn} = \overline{\nu}_{mn} - \underline{\nu}_{mn}$ (McNaught *et al.*, 1997), describing the change in species X_n when reaction R_m occurs. The vector $\boldsymbol{x}$ denotes a certain state for the species vector to be in: $\boldsymbol{X}(t) = \boldsymbol{x} = (x_1, \ldots, x_N)$, where each x_n is an integer molecule number. Thus, simply adding one row of the stoichiometric matrix $\boldsymbol{\nu}_m = (\nu_{m1}, \ldots, \nu_{mN})$ to the state vector $\boldsymbol{x}$ tells us how the state changed after the reaction R_m occurred: $\boldsymbol{x} + \boldsymbol{\nu}_m$. For each of these reactions, there exists a distinct number of possible molecule combinations that could undergo this reaction. This number only depends on the state of the system and can be calculated as a product of binomial coefficients: $h_m(\boldsymbol{x}) = \binom{x_1}{\underline{\nu}_{m1}} \cdot \binom{x_2}{\underline{\nu}_{m2}} \cdot \ldots \cdot \binom{x_N}{\underline{\nu}_{mN}}$. The derivation of the master equation is concerned only with bimolecular reactions and handles reactions involving one or more than two molecules as special cases separately. Thus, probability rates c_m for reactions are found by examining collisions of two random molecules taking part in reaction R_m.

The stochastic rates c_m are derived rigorously in Gillespie (1992) from the physical basis of the molecule movement within a *well stirred* and *constant* volume at *thermal equilibrium*. Those prerequisites give rise to two basic assumptions: For one, the molecule positions are distributed uniformly in the volume. And on the other hand, the movement of each molecule will follow a Maxwell-Boltzmann velocity distribution.[18] Without going into the detailed derivation[19], this gives rise to the probability of a collision of two randomly selected R_m-reactants as well as the subsequent (conditional) probability for the reaction R_m to occur. The product of these two probabilities gives a (probability) rate, $c_m = P(\text{collision}) \cdot P(\text{reaction}|\text{collision})$, that is independent of dt and the probability for reaction R_m to occur in the time interval $[t, t + dt)$ can be written as $c_m \cdot dt$, which is of fundamental importance for the master equation.

As $\boldsymbol{X}(t)$ is the time evolution of system states and thus comprised of integer numbers, it is impossible for us to find a deterministic expression.[20] Nevertheless, the evolution of the probability to be in a certain state $\boldsymbol{x}$ at time t can be described, given some initial conditions: $P(\boldsymbol{x}, t|\boldsymbol{X}(t_0) = \boldsymbol{x}_0, t_0)$. Now, the key idea is to take apart how this probability distribution would change in a (infinitesimal) time interval dt, depending on the probability of certain events. The paper splits these events into three observations on numbers of reactions to occur in the time interval $[t, t + dt)$ (see Gillespie, 1992, Theorem 1-3), that we only state here without giving the proofs. Given $\boldsymbol{X}(t) = \boldsymbol{x}$, the probabilities are:

1. $P(\text{“}No \text{ reaction”}) = 1 - \sum_{m=1}^{M} c_m h_m(\boldsymbol{x})dt + o(dt)$[21]
2. $P(\text{“Exactly } one \text{ reaction } R_m\text{”}) = c_m h_m(\boldsymbol{x})dt + o(dt)$
3. $P(\text{“}More \text{ than } one \text{ reaction”}) = o(dt)$

[18] The arguments are actually concerned with infinitesimal notions of this, as the movement as well as the position of the molecules are regarded in an infinitesimal time interval dt. Very unfortunately, going into the details of the arguments will be out of the scope of this introduction and work.

[19] The interested reader may find these steps thoroughly explained in a very appealing way within the publication (Gillespie, 1992).

[20] The best we can do is to find sample paths in the state space with the Gillespie-algorithm (see 2.2.2).

[21] Note on the nomenclature: $o(dt)$ refers to the standard Landau-“little-o”-notation of terms that vanish much faster than dt.

Notice, that summing up these events for all reactions R_m will (in the limit of $dt \to 0$) equate to 1 and thus cover all possible events. Combining those statements gives the probability to be in state $\boldsymbol{x}$ at time $t + dt$, given the initial conditions.

$$P(\boldsymbol{x}, t + dt|\boldsymbol{x}_0, t_0) = \left(1 - \sum_{m=1}^{M} c_m h_m(\boldsymbol{x})dt + o(dt)\right) \cdot P(\boldsymbol{x}, t|\boldsymbol{x}_0, t_0) \tag{13}$$

$$+ \sum_{m=1}^{M} (c_m h_m(\boldsymbol{x} - \boldsymbol{\nu}_m)dt + o(dt)) \cdot P(\boldsymbol{x} - \boldsymbol{\nu}_m, t|\boldsymbol{x}_0, t_0) \tag{14}$$

$$+ o(dt) \tag{15}$$

The term (13) depicts the joint probability of "*No* reaction in $[t, t + dt)$" and "being in state $\boldsymbol{x}$ at t". As those events are independent, it can be stated as the product of the two probabilities. The same is true for every reaction R_m that could have taken place in $[t, t + dt)$, but in this case we need to consider the distribution of the state vector *before* the reaction. As already mentioned, a change in molecule numbers is described by $\boldsymbol{x} + \boldsymbol{\nu}_m$ for each reaction. Analogously, going backwards by subtracting $\boldsymbol{\nu}_m$ gives us the state the system was in before. The probability of "exactly *one* reaction R_m" taking place times the probability that the system was in the state $\boldsymbol{x} - \boldsymbol{\nu}_m$, summed up over $m = 1, \ldots, M$, now tells us the probability of "inflow" from other states. This is the meaning of the term (14). The last term is exactly statement 3. and tells us that the probability of more than one reaction happening in $[t, t + dt)$ is vanishing faster than dt.

This equation is nearly the result we were looking for. The only question remaining is how to calculate and solve such an equation, as the contribution of the terms $o(dt)$ is unclear as long as dt is unspecified. One can immediately see that dt has to be "sufficiently" small in order to make our statements valid and that the equation invites us to form the differential quotient in the following

way:

$$\begin{aligned}
&\frac{P(\boldsymbol{x}, t+dt|\boldsymbol{x}_0, t_0) - P(\boldsymbol{x}, t|\boldsymbol{x}_0, t_0)}{dt} \overset{(dt\to 0)}{=} \frac{\partial}{\partial t} P(\boldsymbol{x}, t|\boldsymbol{x}_0, t_0) \\
&= \sum_{m=1}^{M} c_m h_m(\boldsymbol{x} - \boldsymbol{\nu}_m) P(\boldsymbol{x} - \boldsymbol{\nu}_m, t|\boldsymbol{x}_0, t_0) \\
&\quad - c_m h_m(\boldsymbol{x}) P(\boldsymbol{x}, t|\boldsymbol{x}_0, t_0),
\end{aligned} \tag{16}$$

with the initial conditions set by

$$P(\boldsymbol{x}, t = t_0|\boldsymbol{x}_0, t_0) = \begin{cases} 1, & \text{if } \boldsymbol{x} = \boldsymbol{x}_0, \\ 0, & \text{if } \boldsymbol{x} \neq \boldsymbol{x}_0. \end{cases} \tag{17}$$

This approach provides us with a system of partial differential equations, called the "master equation"[22], that can be solved to give a continuous description for the evolution of the probability distribution of chemical species $\boldsymbol{X}(t)$ in the model. As mentioned at the beginning, interpreting it as balancing the "gain" and "loss" in probability to be in a certain state gives an immediately tractable intuition on what the equation actually describes.

The master equation can be written down for every reaction system stated like above. Yet, solving the system is in general immensely hard. A very interesting and elegant study on the special case of mono-molecular reactions was performed by Jahnke and Huisinga (2007), where the authors were able to deduce a way of solving the master-equation analytically. This special case is of great importance, yet in a general case the master-equation is intractable to analytical solutions as its complexity grows exponentially with the number of species and reactions in the system. To circumvent this dilemma, the classic approach is to use numerical simulations or dynamic approximations. We will present the approaches that are employed in this work in the following.

[22] As an interesting side note: The name actually originates from a paper by (Nordsieck *et al.*, 1940, paragraph 3) where it was used for energy distributions as a master template "from which all other equations [...of the paper...] can be derived". The name got stuck with this type of equation and, in our opinion, deserves it in many ways.

2.2.2. Gillespie's algorithm - Probabilities at work

As mentioned in the previous section, solving the master-equation for bigger systems is usually not feasible. One way of circumventing the need to do so is by obtaining paths (single trajectories) as possible solutions. Sampling such exact solution paths then gives rise to estimating the probability distributions that are solutions of the master-equation. D.T.Gillespie described such an algorithm in 1977 (Gillespie, 1977), called the "Stochastic Simulation Algorithm" (SSA) that uses a Monte-Carlo approach to simulate exact trajectories for the chemical reaction system. The derivation of the algorithm can be established using the exact same premises as for the master-equation as shown in Gillespie (1992), which makes the two approaches equivalent to one another despite their distinct outputs.

The technique basically splits the problem in two questions per step: "*1. When will the next reaction occur?*" and "*2. Which reaction R_m will be occurring?*". This is the joint probability distribution $P(\tau, m|\boldsymbol{x}, t)$ of the time τ and reaction R_m given the state $\boldsymbol{x}$ at time t. The time τ until the next reaction will occur is an exponentially distributed random number as was shown in the original paper. This fact can be proven from the premises of (I) a well-mixed, spatially homogeneous system in which molecules can be regarded as uniformly distributed and (II) the molecules following a Maxwell-Boltzmann velocity distribution. These were exactly the prerequisites of the derivation for the master equation and in fact, both use similar derivations and are equivalent to one another. Making use of the previously stated probabilities, this can be expressed in the following manner: We see the interval $[t, t+\tau]$ divided into k evenly spaced subintervals of length $\epsilon = \frac{\tau}{k}$ in which the event "*no* reaction" occurs with the probability described in term (13). These are k independent successive events and the probabilities can thus be multiplied to give the overall probability of "*no* reaction" in the whole interval $[t, t+\tau]$. For a shorter notation, we define the sum over the reactions in term (13) to be

$$a(\boldsymbol{x}) = \sum_{m=1}^{M} c_m h_m(\boldsymbol{x}).$$

Now, in the next infinitesimal interval $[t+\tau, t+\tau+d\tau]$ after that, the reaction

R_m is supposed to occur. And, again referring to the previous section for this probability is the term (14). Thus with $d\tau \to 0$, we can write

$$P(\tau, m|\boldsymbol{x}, t) = (1 - a(\boldsymbol{x})\epsilon + o(\epsilon))^k \cdot c_m h_m(\boldsymbol{x}) \tag{18}$$

and by making the subintervals $\epsilon = \frac{\tau}{k}$ infinitely small by $k \to \infty$ we obtain the central formula for the SSA:

$$P(\tau, m|\boldsymbol{x}, t) = \lim_{k\to\infty} (1 - a(\boldsymbol{x})\epsilon + o(\epsilon))^k \cdot c_m h_m(\boldsymbol{x}) \tag{19}$$

$$= \lim_{k\to\infty} \left(1 - \frac{a(\boldsymbol{x})\tau}{k}\right)^k \cdot c_m h_m(\boldsymbol{x}) \tag{20}$$

$$= a(\boldsymbol{x}) e^{-a(\boldsymbol{x})\tau} \cdot \frac{c_m h_m(\boldsymbol{x})}{a(\boldsymbol{x})}. \tag{21}$$

The two terms of the product (21) are what is now to be simulated in Monte Carlo fashion.

The first term in (21), $a(\boldsymbol{x})e^{-a(\boldsymbol{x})\tau}$, is the exponential distribution from which we sample the time τ. We calculate the quantile function for this distribution (by taking the inverse of its cumulative distribution function) and thus can map a uniform random number r_1 of the interval $[0, 1]$ back to such a τ:

$$\tau = \frac{1}{a(\boldsymbol{x})} \cdot \ln\left(\frac{1}{r_1}\right). \tag{22}$$

Next, the reaction to be chosen will be weighted with its occurrence: We divide the unit interval into M fractions, each of a width proportional to the amount that the corresponding reaction R_m contributes to the sum $a(\boldsymbol{x})$, namely the fraction in the second term of (21). The algorithm does this by defining m as "the smallest integer satisfying"

$$\sum_{i=1}^{m} c_i h_i(\boldsymbol{x}) > r_2 a(\boldsymbol{x}), \tag{23}$$

with r_2 being again a random number from the uniform distribution on $[0, 1]$.

As can be seen, the specific reaction that will then be chosen to occur at

time τ is dependent on both the state $\boldsymbol{X}(t) = \boldsymbol{x}$ of the system (i.e. the number of molecules for each species) as well as the stochastic reaction constants c_m associated with the reactions R_m. Again, the probability to choose a certain reaction is proportional to the product of possible distinct substrate-molecule-combinations $h_m(\boldsymbol{x})$ and the stochastic rate of the reaction. The rate itself also depends on the volume in which the reaction will take place, since within a smaller space the probability of molecules colliding and reacting will increase.[23]

After selecting the reaction R_m that is occurring, the molecule numbers $\boldsymbol{x}$ are updated according to the reactants and products of the reaction in question by adding the corresponding row $\boldsymbol{\nu}_m$ of the stoichiometric matrix. The time is increased by τ and unless the desired simulation time is reached, these steps are repeated. The algorithm can be summarized as follows:

1. Determine the time τ of next reaction event
2. Determine the reaction R_m that will occur next
3. Update time $t = t + \tau$ and the systems state-vector $\boldsymbol{x} = \boldsymbol{x} + \boldsymbol{\nu}_m$
4. If $t < t_{\text{final}}$, go to step 1

A typical example of a trajectory for a SSA model for gene expression can be found in Fig. 9. The result of one SSA-run is a possible trajectory out of the solution for the master-equation of the process, discrete in the system-state $\boldsymbol{x}$ (molecule numbers) and continuous in time. By simulating sufficiently enough trajectories, we can then sample and estimate the underlying probability distribution of molecule numbers at times t, i.e. the species population vector $\boldsymbol{X}(t)$.

The SSA is computationally expensive if the rates and/or the numbers of molecules are high - corresponding to the huge number of reactions that are occurring in that case. There are techniques to speed up this process by making simplifications and assumptions that lower the complexity and thus accuracy of the simulations. One such technique is the τ-Leaping version of the SSA (Gillespie, 2001; Cao *et al.*, 2006, 2007), where the species vector

[23] Of course, by definition reactions of first order are unaffected by the volume since no collisions have to take place.

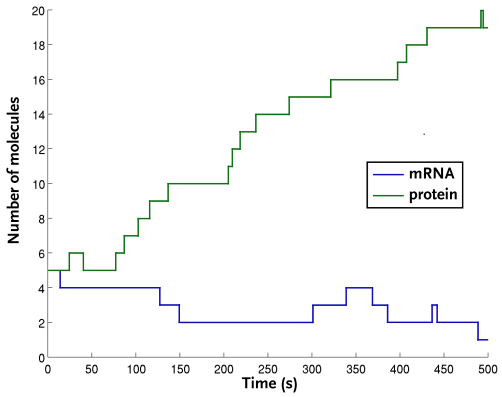


Fig. 9: Trajectories in one run of a SSA simulation - mRNA and protein in a model for stochastic gene expression (see Shahrezaei and Swain (2008)).

is regarded as constant over a time window larger than the "time to next reaction". Thus the firing of several successive reactions will be computed by drawing random numbers (from appropriate distributions derived from the equations in this section) telling which and how many occur in this time frame. The approximation of constant species introduces a controllable error. Another technique would be to employ continuous versions like the chemical Langevin-equation (Gillespie, 2000) and others.

The original detailed description of the algorithm can be found in Gillespie (1977). Tutorials and applications of the Stochastic Simulation Algorithm were subject or tool in numerous publications. As a review, we recommend Gillespie (2007) written by D.T.Gillespie himself.

2.2.3. Moment closure - On how to make the CME tractable

As this sampling with Monte Carlo techniques like the Gillespie-algorithm becomes quite complex for large systems as well as in situations where events occur too frequently (e.g. high abundant molecules or fast reaction rates), often other approximations are employed to obtain the probability distributions needed for our information theoretic analysis. Here we introduce one of the most prominent of such approximations: The "moment closure".[24]

[24] **Remark:** Usually, such simplifications run into problems when the system exhibits strong non-linearities or abundances of some of the species involved become sparse. In cellular

As a direct consequence from the master-equation, we can state the time development of the moment generating function $M(t)$[25] of a stochastic variable X consisting of the states $x \in \{0, 1, 2, \ldots\}$. Each of the states could for example denote a number of molecules with the associated probability $P(x, t)$ of being in that particular state at time t. In order to familiarize the reader with the approach, we limit the derivation of the moment closure to the accessible one-dimensional case. The multivariate extension for multiple species $X_1, X_2, \ldots$ follows the same steps in a straightforward way (for a complete derivation we refer to Gillespie (2009)). The moment generating function[26] of a probability distribution is defined as:

$$M(t) := \sum_{x=0}^{\infty} P(x, t) \cdot e^{x\theta}. \tag{24}$$

Remark: Notice that the lower and upper boundary of the sum do not necessarily have to be zero and infinity, respectively. Rolling a dice would, for example, require them to be 1 and 6. We, however, are applying this to biochemical entities and thus our states are molecule numbers greater than or equal to zero. The upper boundary might be definable by space limitations for molecules, yet a particular number is in no way unique and thus of no importance for us. Enclosing it in the infinite sum with the probabilities tending to zero states the general case and also raises no difficulties with the further derivations.

signaling, the latter is not too much of a concern. Signaling molecules generally occur in abundances that are low enough for variations and noise to matter, yet large enough that stochastic "shocks" are of controllable magnitudes. Here, one has to factor in spatial considerations as well. Mammalian cells are typically larger than for example yeast cells. Concentrations of signaling molecules still need to be able to functionally cover the volume sufficiently, thus making higher molecule numbers necessary in bigger cells. Since our models do not mix abundant with sparse species (like mRNA production in gene expression would for example), the scales can be handled with said approximations. Different splitting techniques would be required, if one broadens the scope of the models to allow for such a mix. The non-linearities occurring in our modeling structures were seen to be in tolerable boundaries for the approximations to still be viable.

[25] Remark on nomenclature: As in this work the employed systems and models reflect dynamic time-courses of molecule numbers, the parameter t will denote "time" here. In some nomenclatures this character is used for the definition of the moment generating function. We use θ for that matter.

[26] See also classical literature for statistics, e.g. Kenney and Keeping (1951); Bailey (1990); Ross *et al.* (1996).

In a next step, expanding the exponential function $f(x\theta) = e^{x\theta}$ as a Taylor series around the origin results in a natural representation in terms of the moments:

$$\sum_{x=0}^{\infty} P(x,t) \cdot e^{x\theta} = \sum_{x=0}^{\infty} P(x,t) \cdot \left[\sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} \cdot (x\theta - 0)^n \right] \tag{25}$$

$$= \sum_{x=0}^{\infty} P(x,t) \cdot \left[\sum_{n=0}^{\infty} \frac{1}{n!} \cdot x^n \theta^n \right] \tag{26}$$

$$= \sum_{n=0}^{\infty} \frac{\theta^n}{n!} \cdot \sum_{x=0}^{\infty} P(x,t) x^n \tag{27}$$

$$= \sum_{n=0}^{\infty} \frac{\theta^n}{n!} \cdot E\left[X^n(t)\right] \tag{28}$$

$$= \sum_{n=0}^{\infty} \frac{\theta^n}{n!} \cdot \mu_n(t) \tag{29}$$

$$= \mu_0(t) + \frac{\mu_1(t) \cdot \theta^1}{1!} + \frac{\mu_2(t) \cdot \theta^2}{2!} + \ldots =: M(t), \tag{30}$$

where $\mu_n(t)$ denotes the n^{th} moment.[27] As can already be seen in this equation, the terms in the sum are separated both in the moments as well as the distinct powers of θ. This will be taken up again later. For now, we go back to the definition of the master-equation as stated in Eq. (16) and follow the idea of the moment generating function by multiplying this expression (16) with $e^{x\theta}$ and subsequently sum over the states x.[28] Omitting the conditioning on the

[27] Note that the 0^{th} moment is solely the normalization $\sum P(x,t) = 1$ in the case that P is a probability distribution.

[28] Note that, whereas we had in Eq. (16) a vector $\boldsymbol{x}$ for all possible combinations of states, we now only regard one species and thus only x, e.g. all possible numbers of molecules of species X in the system!

initial state (x_0, t_0) for better readability, we thus obtain

$$\frac{\partial M(t)}{\partial t} = \sum_{x=0}^{\infty} \underbrace{\frac{\partial P(x,t)}{\partial t}}_{(16)} \cdot e^{x\theta} \tag{31}$$

$$= \sum_{x=0}^{\infty} \sum_{m=1}^{M} \left(c_m h_m(x-\nu_m) P(x-\nu_m, t) - c_m h_m(x) P(x,t)\right) \cdot e^{x\theta} \tag{32}$$

$$= \sum_{m=1}^{M} \sum_{i=0}^{\infty} c_m h_{m,i} \frac{\partial^i M(t)}{\partial \theta^i} \left[e^{\nu_m \theta} - 1\right] \tag{33}$$

Remark: The step from (32) to (33) is not straightforward to see, yet can be grasped as a convoluted consequence of the following statement used in (Bailey, 1990, chapter 7.4):

$$M(\theta, t+\Delta t) = M(\theta, t) \Delta M(\theta, t) \tag{34}$$

and

$$\Delta M(\theta, t) = E\left[e^{\theta \Delta X(t)}\right]. \tag{35}$$

$X(t)$ and $\Delta X(t)$ are assumed to be independently distributed for the product in Eq. (34) (as they are in our systems). For a detailed derivation, we refer to the aforementioned book chapter in Bailey (1990). In our case, $X(t)$ is identified by ν_m, giving the result. A use of this for cumulant neglects can be found in Matis and Guardiola (2010)

We now write the partial derivative as established in Eq. (29) leaving out the vanishing terms and develop the exponential in a Taylor series:

$$\frac{\partial M(t)}{\partial t} = \sum_{m=1}^{M} \sum_{i=0}^{\infty} c_m h_{m,i} \sum_{j=i}^{\infty} \frac{\theta^{j-i}}{(j-i)!} \mu_j(t) \cdot \left(\sum_{k=0}^{\infty} \frac{(\nu_m \theta)^k}{k!} - 1 \right) \tag{36}$$

and finally sort the equation to extract powers of θ, giving the final equation and result of this section:

$$\frac{\partial M(t)}{\partial t} = \sum_{n=0}^{\infty} \theta^n \sum_{m=1}^{M} \sum_{i=0}^{\infty} c_m h_{m,i} \sum_{k=0}^{\infty} \nu_m^k \binom{n}{k} \mu_{n-k+i}(t) \tag{37}$$

Now, comparing this final Eq. (37) with Eq. (29) and equating the coefficients (powers of θ) gives us one expression for each moment μ_i. Each of these expressions is a differential equation, allowing us to solve the system of equations for a dynamic solution of all moments. Due to the fact that we only used the stoichiometric matrix and the propensity functions for our reaction system, we were able to translate the chemical reactions to an approximate stochastic representation. Finally, setting moments starting from a certain order to zero will "close" the (by definition infinite) number of equations and give a closed unique ODE system. In our work, we will use a closure of second order as this will give us Gaussian distributions of the involved species, which is sufficiently exact for our systems.[29] This closure is also termed "Gaussian approximation" in the literature.

Of course, the same procedure can be done with neglecting orders of cumulants instead of moments, as they are equivalent formulations of the concept (e.g. Matis and Guardiola, 2010). Other long established procedures of probability distribution truncations are used in the literature as well. Since we decided for the application of the moment closure, we only want to mention some of them here and refer to the corresponding literature for the interested reader. A very popular way to approximate fluctuations in biological models is the "linear noise approximation", as used by Elf and Ehrenberg (2003); Hayot and Jayaprakash (2004). This approach is also known as the "van Kampen system size expansion" (Kampen, 1961; Van Kampen, 2011)[30], named after the physicist who first developed the idea. The "Kramers-Moyal expansion" (Kramers, 1940; Moyal, 1949) (already employed in a variant by Einstein (1905)) is a close relative to that. Both methods are also introduced in Gardiner (2009). A study on the validity of such approximations can be found in Wallace *et al.* (2012).

The procedure of moment closure as reviewed here is a powerful approach for biological processes. With its direct way of going from reaction systems to a stochastic approximation via Eq. (37), it makes fluctuations within the system comprehensible and enables studying their implications on top of the

[29] As, in fact, many processes in biology follow normal distributions.

[30] A comparison to the second order moment closure (called "gaussian approximation") can be found in Lafuerza and Toral (2010).

conventional description of means. In this work, we apply the moment closure in chapter 5 in order to study noise in signal transduction that incorporates scaffolding proteins.

2.3. Parameter estimation with "Data2Dynamics"

Parameterization is a crucial step in model development. Wide ranges and combinations of parameter choices are thinkable to serve the task at hand, yet validating the "right" parameters that occur in nature is not possible. At the same time, it is also not the central purpose of the modeling in the first place. Of course, using parameters that are based on actual findings and knowledge of the system is necessary and important, but considering all the potential unknowns that are hidden to our view for now and the complexity of the biological "truth", we can only go so far as to develop and test hypotheses about how a structural mechanism could achieve what we observe. Simplifications and approximations are thus tailored to explain how dynamic system behavior can arise and parameterizations are fitted to this desired output. The ways of how to do that are well-established and are used extensively in other fields, most importantly economics. Fitting parameters encounters many caveats. The most recognized is "unidentifiability" of parameters, meaning that within our system, degrees of freedom exist that lead to sets of parameters that allow infinite value combinations that produce the same behavior. Recent research has produced software that provides many functionalities for such fitting problems and the overall task to combine data and modeling. In this regard, especially the group of Jens Timmer from the Institute of Physics in Freiburg, Germany, has provided many useful concepts and studies in both theoretical as well as applied work (e.g. Maiwald and Timmer, 2008; Kreutz and Timmer, 2009; Raue *et al.*, 2009, 2010). These studies are exemplary for the approach of systems biology, enforcing the cycle of model and experiment.

For implementing the model of chapter 4 and estimating its parameters, we utilized the software "Data2Dynamics"[31] from Raue *et al.* (2013). This environment enabled us to connect our theoretical model of the crosstalk between HOG- and pheromone-pathway with the data from Vaga *et al.* (2014). By

[31] Available at `http://www.data2dynamics.org/`.

using latin hypercube sampling of the parameter space and the provided optimization methods for data fitting, we obtained potential parameter sets for our model that fit the structure reported in literature. As we are looking for both agreements and disagreements between data and model output, the metrics on the goodness of fit that "Data2Dynamics" provides will be used as a benchmark.

Many advances have also been made for parameter estimation in biological stochastic models (e.g. Reinker *et al.*, 2006; Daigle *et al.*, 2012) more recently, but this field is still in need of further development and practical application, particularly for models of larger complexity. Yet, "Data2Dynamics" can potentially and in a straightforward way be used with the aforementioned approximations. Since a moment closure approach is resulting in nothing else than a system of ODEs for the moments of species, this tool could basically also be applied to fit such models to experimental data that include sufficiently good statistics on variations of the species. Yet, as our approach in chapter 5 aims to understand noise in signaling from a structural point of view, and thus is not connected to particular datasets, we did not utilize "Data2Dynamics" in this case. Going into details for a particular scaffolding system could use this as a simulating and fitting environment with great potential.

3. Information processing in stress-adaptation: An analysis of the Sln1-Phosphorelay

Within this chapter, we want to revisit a well studied pathway, the High Osmolarity Glycerol (HOG) pathway, and approach it from a new angle. We investigate the Sln1-Ypd1-Ssk1-phosphorelay as a modular input to the HOG pathway and propose a stochastic view together with the idea of evaluating signal processing capabilities in this system. Within this framework, we can imitate the noisy perception of the cell and interpret the phosphorelay as an information transmitting channel in the sense of C.E. Shannon's Information Theory (see section 2.1). We refer to the channel capacity as a measure to quantify and investigate the transmission properties of this system, enabling us to draw conclusions on viable parameter sets for modeling the system. This view extends the existing approaches in a way that it incorporates not only the signal and its transmission, but also the noise that inherently arises in the system and limits its potential. The pathway has been described and researched for about two decades in great detail both experimentally as well as from a modeling perspective. This gives us the opportunity to build upon existing knowledge and a sound foundation to test our hypotheses on this prototypical signaling pathway.

3.1. The HOG-pathway in yeast osmotic stress response

Stress responses in *S.cerevisiae* are mediated by a complex network of signaling pathways that coordinate adaptational programs. Prevalently, conserved three-kinase cascades of mitogen-activated protein (MAP) kinases are employed for this task (Seger and Krebs, 1995; Klipp *et al.*, 2005). In the case of increased osmolarity (e.g. high salt concentrations in the environment), one of the most intensively studied signal transduction pathways, the High-Osmolarity Glycerol (HOG) pathway, is employed. It is a close relative to the p38 signaling pathways in mammalian cells and has been under research both from an experimental (e.g. Posas *et al.*, 1996b; Hohmann, 2002; Macia *et al.*, 2009; Patterson *et al.*, 2010a) as well as a computational side (e.g. Klipp *et al.*, 2005; Muzzey *et al.*,

2009; Petelenz-Kurdziel *et al.*, 2013; Patel *et al.*, 2013). The pathway provides the cell with answers to an increased concentration of osmolytes in the environment. This concentration can be lethal for the cell as it severely changes the pressure on the cell wall, water exchange and with it the cell volume as well as chemical reactions within the cell. For the cell to survive, a complex system of adaptation processes is initiated. An immediate answer is the closure of Fps1 channels, preventing the further outflow of glycerol, the osmolyte being accumulated within the cell. The further adaptation is regulated over the HOG pathway facilitating a transcriptional answer, ultimately leading to glycerol production and near perfect longterm adaptation (Muzzey *et al.*, 2009). The connections of the pathway are manifold and there is a strong interplay with the cell cycle, as it is important that an adaptation to this stress is prioritized and proliferation is stopped immediately to avoid flawed replication. For a review of the pathway we refer to Hohmann (2009).

When facing different stress levels, the cell exhibits very distinct profiles in Hog1 activation and cellular response. This can in part be attributed to feedbacks in the system and the complexity of interacting mechanisms. Several signal processing techniques have been applied to characterize the response of the HOG pathway (e.g. Mettetal *et al.*, 2008; Hersen *et al.*, 2008). But even though those studies have common aims towards a better understanding of how signaling in complex systems works and enables adaptation, the stochastic nature of biochemical signaling itself has been largely disregarded. Here we explore the idea of this stress answer being part of a fidelity problem to the cell. Ultimately, this would require that by encoding and transmitting the extracellular signal appropriately, the cell can distinguish an optimal number of input levels for further processing. For this we focus on the Sln1-Ypd1-Ssk1-phosphorelay, an extended two-component signaling system that forms the first module of the HOG pathway (Maeda *et al.*, 1994; Stock *et al.*, 2000). Observing the cell's ability to discriminate profiles already in this first signaling instance will add a further "stochastical layer" to the study on input-output relations (Shinar *et al.*, 2007) of these systems.

3.2. A stochastic model of the Yeast Sln1-phosphorelay

Here we want to investigate the role of the so called "Sln-branch".[32] For evaluating the capabilities of *S.cerevisiae* to react to osmotic pressure, we are zooming in on this first part of the HOG stress response pathway: As visualized in Fig. 10, the phosphorelay consists of three proteins of interest: Sln1, Ypd1 and Ssk1. They form a biochemical signal transduction chain that belongs to the family of "two-component regulators" that are a common feature in prokaryotic signaling, but also found in eukaryotes (West and Stock, 2001). The signaling chain can be described by the following coupled system of species and reactions:

$$\text{Sln1} \xrightarrow{k_1} \text{Sln1-P} \tag{38}$$

$$\text{Sln1-P} + \text{Ypd1} \underset{k_{2,\text{off}}}{\overset{k_{2,\text{on}}}{\rightleftarrows}} \text{Sln1-P-Ypd1} \tag{39}$$

$$\text{Sln1-P-Ypd1} \underset{k_{2,\text{on}}}{\overset{k_{2,\text{off}}}{\rightleftarrows}} \text{Sln1} + \text{Ypd1-P} \tag{40}$$

$$\text{Ypd1-P} + \text{Ssk1} \xrightarrow{k_3} \text{Ypd1} + \text{Ssk1-P} \tag{41}$$

$$\text{Ssk1-P} \xrightarrow{k_4} \text{Ssk1} \tag{42}$$

Sln1 is a trans-membrane protein that reacts on the turgor pressure put on the cell wall. In the non-stressed situation it constantly auto-phosphorylates His-576 under the consumption of ATP. In our model this rate (k_1) of auto-phosphorylation plays a fundamental role as it is considered the encoder of the (osmotic) input. Depending on the stress level this rate will be decreased by a certain factor, mimicking a reduction in the auto-phosphorylation rate.[33] The phosphate group is subsequently transfered to the response regulator domain Asp-1144 of the protein[34], from where it can be further relayed to His-64 of

[32] The other branch (named "Sho-branch" after the membrane bound protein Sho1, that takes part in the activation of the pathway) will be introduced in chapter 4.

[33] Since we are thinking in stochastical terms and rates, this could also be interpreted as a decreasing probability of each Sln1 molecule in the whole set of receptors to be (auto-)activated, capturing a conformational change. The link to a deterministic description then appears by averaging over the ensemble, thus giving us a ratio of phosphorylated to dephosphorylated molecules and reflecting the stress level.

[34] This is believed to happen between dimerized Sln1 molecules, instead of intra-molecular, as kind of an exchange (Qin *et al.*, 2000).

the intermediate signaling protein Ypd1. This is mediated by the (reversible) formation of a complex Sln1-P-Ypd1. Because the phosphate transfer within the complex happens at a very fast rate (Janiak-Spens *et al.*, 2005; Kaserer *et al.*, 2009), this step will not be modeled explicitly.

The intermediate protein Ypd1 is a more abundant although smaller molecule. It comes in a copy number of roughly 6300 molecules per cell[35] that can enter the nucleus freely. Compared to the other species of our model (Sln1: ~ 650 molecules, Ssk1: ~ 1200 molecules) this is a relatively high copy number.[36] This might be due to the fact that the cell needs to circumvent a bottleneck in information shuttling.[37] Ypd1 is able to interact with both Ssk1 in the cytosol as well as nuclear Skn7 (Li *et al.*, 1998; Lu *et al.*, 2003) to transfer the phosphate to the respective response regulator domain, but a transfer to Ssk1 is strongly favored as demonstrated in Janiak-Spens *et al.* (2005). The phosphoryl group was not observed to be transported back to Ypd1.

Ssk1 is the protein that is used in our model as an output in its unphosphorylated form by catalyzing the phosphorylation reactions of the downstream MAPK cascade leading towards the double phosphorylation and thus activation of Hog1.

In an unstressed environment, Ssk1 will constantly be phosphorylated and its activating function thereby inhibited. Upon osmotic shock, Sln1 acts on the variation of turgor pressure by a change in its conformation (Tao *et al.*, 2002). Its auto-phosphorylation rate will be decreased and thus successively also the inhibition of Ssk1. Ssk1 becomes free to catalyze the downstream reactions and activates a chain of amplification, signaling the presence of stress. The model uses the probability distribution of this species as the relevant observable for the system. Its fidelity defines how detailed the response of the cell will be.

Because of the way this phospho-transfer-system is designed, mutating the phosphorylation sites of Sln1, Ypd1 or Ssk1 will have lethal consequences (Maeda *et al.*, 1994; Fassler and West, 2010; Posas *et al.*, 1996a): Since Ssk1 is

[35] Numbers taken from "http://www.yeastgenome.org/".

[36] In our model, this enables us to choose the shuttling rate k_3 in a non restrictive but computationally more efficient manner.

[37] A similar consideration can be found in the subsection 5.2.3.

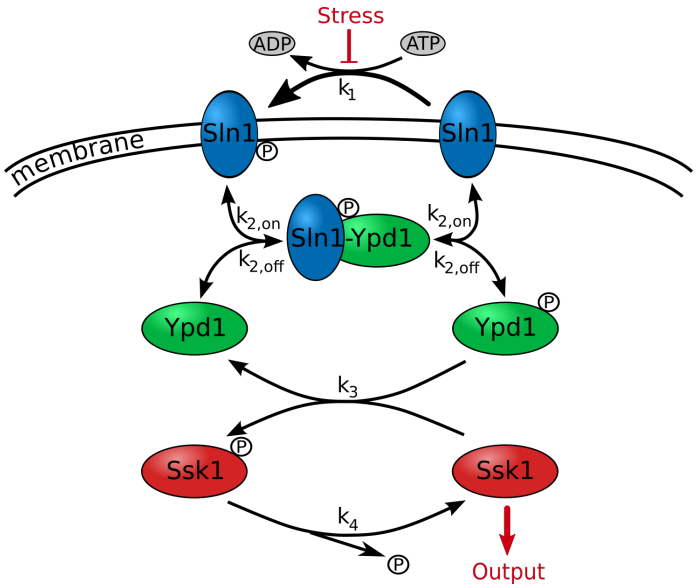


Fig. 10: Schematic overview of the Sln-branch phosphorelay in the Hog-pathway.

never inactivated and constitutively signals a stress situation, the cell constantly produces an excessive amount of osmolytes that increase the pressure from within the cell, ultimately bursting it. This also reflects the importance of the phosphorelay system, as it was designed by nature in a way that does not allow for instability.

The value for the association rate of the complex Sln1-P-Ypd1 was chosen to be varied with the values $k_{2,on} = \{1,\ 5,\ 10\}\ \cdot 10^6 (Ms)^{-1}$ and the dissociation constant is kept at $K_d = \frac{k_{2,off}}{k_{2,on}} = 300nM$, thus defining $k_{2,off}$. In our model this kinetic change will vary the response of the system and through an decreased/increased number of forward as well as backward reactions forming the complex also the observed variability.

As was experimentally investigated by Janiak-Spens *et al.* (2005); Kaserer *et al.* (2009); Janiak-Spens *et al.* (1999) the corresponding reaction

$$\text{Ypd1-P} + \text{Ssk1} \rightarrow \text{Ypd1} + \text{Ssk1-P}$$

was not observed to be reversible. In addition, this phosphotransfer from Ypd1 to Sln1 was very rapid. Because of the lack of reversibility we reduce this

reaction chain in our model, thus simplifying the formation of a Ypd1_P-Ssk1-complex, the subsequent transfer of the phosphate and the dissociation of the complex to just one single reaction with a sufficiently high reaction rate.

The volume in which the reactions take place is important for the probability of molecules reacting in our model. Since we only look at reactions happening in the cytosol of the cell, we take this number to be $vol = 30\text{fL}$.[38] This volume will have an effect only on the simulations for reactions of second order, since a higher density of reactants will cause a higher probability for a reaction. Projecting the number of molecules to concentrations illustrates the same argument in another way as it means dividing by the volume, resulting in higher concentrations for smaller values. Thus, adjusting this number is crucial for comparing the variation of $k_{2,\text{on}}$ in the model.

To summarize the settings and parameters used for the simulations, we give an overview over all parameter values in Tab. 1. The stoichiometric matrix of the system depicted in the reactions (38)-(42) as used in the SSA can be defined as

$$\nu_{MN} = \begin{pmatrix} -1 & 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 1 \\ 1 & -1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & -1 \\ 0 & 1 & -1 & -1 & 1 & 0 & 0 \end{pmatrix}^T, \; X = \begin{pmatrix} \text{Sln1} \\ \text{Ypd1} \\ \text{Ssk1} \\ \text{Sln1-P} \\ \text{Ypd1-P} \\ \text{Ssk1-P} \\ \text{Sln1-Ypd1-P} \end{pmatrix},$$

where X is the vector of species and the matrix ν_{mn} has been transposed for better readability in correspondence to the species. The number of species is equal to the number of reactions used $M = N = 7$. The rates used for the reactions follow the order of the reaction system (38)-(42).

[38] Volume for yeast cells varies a lot and the number chosen is just a first assumption. With regard to the biological experiments this can be adjusted. Zi *et al.* (2010) i.e. uses a cell volume of 34.8fL for computations and accounting for about 50% of the cell's volume belonging to the cytosol, this would result in a volume of 17.4fL in our model.

Name	Values	Description
vol	$30fl$	Volume of cytosol in the cell.
X	$[0\%, ..., 90\%]$	Stress level, 39 equally spaced values.
k_1	$[0.01, \ldots, 7]\ s^{-1}$	Phosphorylation rate of Sln1.
		(Input function of the system: $k_1 * (1 - X)$.)
$k_{2,\text{on}}$	$[1, 5, 10] \cdot 10^6 (Ms)^{-1}$	Association rate between Sln1 and Ypd1.
$k_{2,\text{off}}$	$K_d/k_{2,\text{on}}$	Association rate of Sln1-Ypd1 complex.
k_3	$10\ \cdot 10^6 (Ms)^{-1}$	Rate of phospho-transfer from Ypd1 to Ssk1.
k_4	$[0.01, \ldots, 7]\ s^{-1}$	Dephosphorylation rate of Ssk1.
K_d	$300nM$	Dissociation constant for Sln1-Ypd1-complex.
t_{ststst}	$150s$	Simulation time for steady-state.
t_{stress}	$40s$	Simulation time for stress response.
Sln1	656	Total Sln1 # (Sln1 + Sln1-P).*
Ypd1	6330	Total Sln1 # (Ypd1 + Ypd1-P).*
Ssk1	1200	Total Sln1 # (Ssk1 + Ssk1-P).*
*(For the initial state, the system is set to be in the fully phosphorylated form, with Sln1-Ypd1-P set to 0.)		

Table 1: Overview over the parameters used for the simulations.

Here, we implement the proposed phosphorelay model with the Gillespie SSA (Gillespie, 1977) in order to simulate a sufficient number of trajectories. Fig. 11 illustrates one typical simulation run. As expected, we observe characteristic dynamics for each species depending on the chosen parameter set. By using the stochastic framework, we introduce noise into the system as well, enabling us to examine its properties of signal fidelity.[39] To observe the output, we sample its probability distributions as a function of time depending on a defined input[40], simulated with an adequate number of runs. We vary the two crucial

[39] "Fidelity" in this sense refers to a measure on how accurate the signaling can reproduce the input signal.

[40] This input being a percentage of the auto-phosphorylation rate k_1 of Sln1, depending on the stress level.

parameters for input (k_1) and output (k_4) within the system to observe the dependence of information transmission on them.

Remark: It is interesting to note that, provided a sufficient signal transmission within the system without bottlenecks occurring, the control lies with those two nodes. This is a characteristic that can be observed in general: The system will balance and switches for the input will in a viable system produce the desired output. For the phosphorelay, it is already apparent at this stage that the dephosphorylation rate k_4 will have to be low enough as to not produce a strong basal output, yet high enough to let the system be susceptible to stress in a timely manner. The signaling in HOG stress response usually works on a timescale below one minute, requiring a fast reaction time. Likewise, the input signal will have to be strong enough to excite the system constitutively, yet also enable a switching off during a stressed situation.

For analyzing the phosphorelay, first a functional system needs to be ensured by simulating its behavior in the non-stressed environment and thus its steady state. Subsequently, the system's capacities can be evaluated by our chosen framework. The results of these two steps are then combined.

3.3. Implications of information transmission on parameter spaces

When working with mathematical models, considering the parametrization is always a very important step and an art in itself. The question of what can be achieved using different sets of parameters has to be posed and the answer should ideally be congruent with what the aims of your model will be. As there are more answers than there are possible combinations of parameter sets, it is quite a difficult task to choose which variant will support a claim to the best possible extend. After all, compared to the overwhelming complexity of biological systems in nature, we are only working with very limited theoretical projections of them, in a best case scenario solving barely a few riddles at a time. Yet this doesn't mean that modeling is helpless in the face of this complexity. Parameters can be adjusted to fulfill certain requirements. The most prominent one is certainly a fit to data. For our phosphorelay model, we are able to find additional constraints on signaling by making two assumptions that our signaling mechanism has to fulfill. The first is a functionality constraint for the

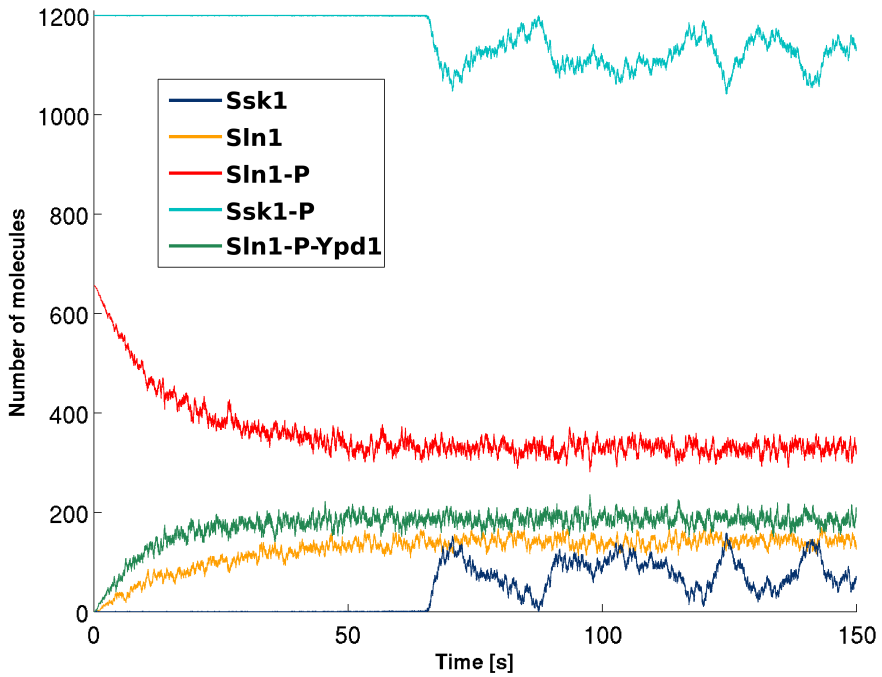


Fig. 11: Detailed view of a single exemplary simulation run of the model with the Gillespie algorithm. The varied parameters (see Tab. 1) are set to $k_1 = 2s^{-1}$, $k_4 = 0.25s^{-1}$ and $k_{2,on} = 5 \cdot 10^6 (Ms)^{-1}$. The higher abundant protein Ypd1 has been omitted. The unstressed steady state can be observed after about $70s$. Starting from this state the systems' signal propagation properties can be examined by applying stress. Here we focus on the analysis of the noise emerging in the system. Matching the timing in transduction to biological behavior could provide further insight as it places further constraints on the dynamical kinetic parameters.

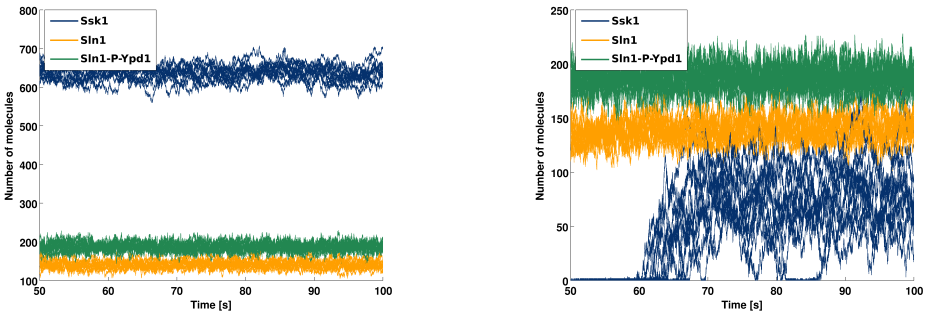
phosphorelay and the second uses an optimality criterion to give a lower and upper boundary on information transmission (which for a signaling module is a sensible choice) and thus restricts the potential parameter set. The lower boundary is a prerequisite for a reliable on/off signal and thus we use it as a sharp requirement. The upper boundary provides optimal parameter choices for the fidelity in the system, but can also be disregarded by the cell. Thus, emphasizing the particular numbers of capacity should be handled with care. In the following, we explain the two restrictions to our parameter space.

3.3.1. Saturated responses - Phosphorylation of Ssk1 in unstressed steady state

Signaling molecules in yeast stress responses show a characteristically high saturation in a stressed environment. For one, they are only abundant at a relatively low number of copies, emphasizing the need for reliable signaling. Additionally, the volume of yeast cells is small enough for even low concentrations to sufficiently cover a viable signaling response.[41] A high saturation also shows an efficient use of the molecule numbers, as it implies a low energy loss due to protein maintenance.

Aiming at an operational signal transduction in the cell we first require that, if our environment exhibits no stress, also no signal (or only a basal level) is transmitted. Thus, we demand that a high percentage ($> 80\%$, or arguably even a higher threshold) of our signaling output Ssk1 remains phosphorylated in unstressed environmental conditions. This ensures that the relay is not constantly activating the downstream signaling MAP kinase and thus doesn't keep the cell stressed without an immediate need for it. This is crucial to be taken into account since a permanent activation of the pathway can be lethal to the cell (Maeda *et al.*, 1994). The threshold we are setting selects for feasible rate combinations between phosphorylating Sln1 (k_1) and dephosphorylating Ssk1 (k_4). Fig. 12 illustrates two possible scenarios for the simulations, one activating the downstream signal constitutively and the other exhibiting a functional non-stressed behavior.

[41] In comparison, mammalian cells utilize far higher molecule numbers, also owing to the fact that their volume is larger and regulation exhibits a higher complexity.



(a) $k_4 = 0.5s^{-1}$ (b) $k_4 = 0.25s^{-1}$

Fig. 12: Trajectories of steady state simulations for the output (Ssk1, dark blue) with $k_1 = 2s^{-1}$ and $k_{2,on} = 5 \cdot 10^6 (Ms)^{-1}$, further parameters see Tab. 1. Varying k_4 yields two characteristic behaviors: (a) Shows a combination that activates the downstream pathway constitutively due to a high level of dephosphorylated Ssk1 and is hence not fit for transmitting the signal. (b) Shows a "feasible parameter combination", exhibiting only a basal level of dephosphorylated Ssk1 in steady state. This allows for a functional pathway and can be further investigated.

It is imperative that if k_4 is set to a low enough value, phosphorylation will always exceed the dephosphorylation and prevent Ssk1 from signaling further downstream in steady state as expected. So there is no lower boundary on k_4 that violates the functionality, only one on information transmission as explained in section 3.3.2. For the time being it will be considered the upper limit on the dephosphorylation rate that is of interest to us. After simulating the model for a sufficiently long time to reach the steady state, the feasibility of the parameters can be evaluated using said threshold.

As can be seen in Fig. 13, we observe that the steady state behavior already sets a very strict upper boundary on the rate combinations of k_1 and k_4. Within the simulated range, k_4 could only be chosen at a magnitude lower than the input rate. In addition to that it only took a marginal difference in k_4 to shift the equilibrium towards a non-functional system. This small margin of possible combinations gives an idea of how the system is constructed: While the input rate k_1 can take on a range of values that excite the system in an unstressed

state, the dephosphorylation rate k_4 has sharp restrictions. Setting it too high will signal a basal level of activation downstream and thus potentially activate Hog1. An interesting perspective would be to test a phosphatase whose activity is dependent in some way on the stress level or a downstream target as well. Signaling could then function in a low energy regime that uses slow rate combinations and still would be able to react fast on stress induction. For this hypothesis to be tested, a kinetic study on the activation of Ssk1 would have to be performed.

Analyzing the parameter space with regard to the steady state behavior in non-stressed environments forms the first part of restricting the possible parametrization aiming towards a functional model. Alternatively, this analysis could be done by observing the mean value of the output, corresponding to simulations of differential equations for this system. This would enable an analytical view on the problem, but it also would deny the possibility to exploit the noise that we observe. We use this steady state simulations as part of the subsequent analysis of channel capacity.

3.3.2. Information as a lower bound

One tempting way to think about objectives for biological organisms is in terms of the optimizations of certain aspects. An organism could for example optimize its growth rate with respect to a particular environment as suggested in Rivoire and Leibler (2011). On the other hand, fast responses might be evolutionary selected in the case of severe and urgent stresses that need to be dealt with in a timely manner. Depending on the environmental variables that biological systems are subjected to, this aim will change and pronounce certain features more and others less.[42] Yet, as to what really is important for a cell, we can only speculate and make assumptions for sensible choices. In a natural setting it will most definitely be a mixture of many optimizations to be considered and balanced. Our interpretation of the phosphorelay as a channel gives us an optimization with respect to the fidelity in the system. Developing a sensitive and accurate information transmission necessary for biological systems and thus we argue that such an optimization is a good choice

[42] After all, both the channel as well as the response are subject to evolutionary development.

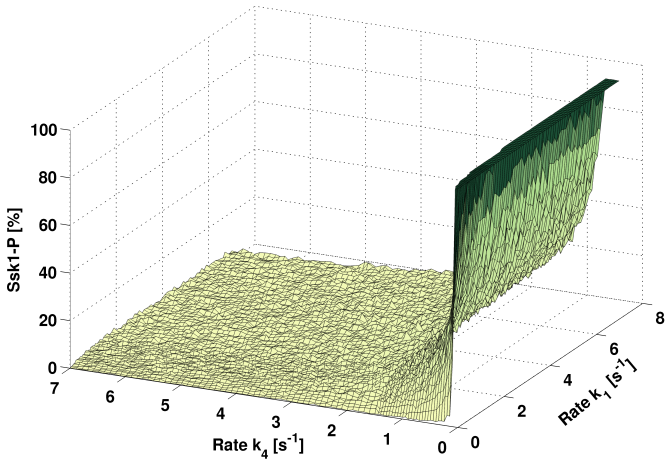


Fig. 13: Restrictions on the output rate k_4 (with fixed $k_{2,on} = 5 \cdot 10^6 (Ms)^{-1}$, see Tab.1 for other parameters). In non-stressed steady state, only the values exhibiting more than 80% of phosphorylated Ssk1 are selected as feasible behavior of the system. This provides only a narrow range for choosing k_4 as illustrated by the figure.

for cell signaling. Of course, the neglect of information has to be considered as well. Because of that, we rather see this interpretation as presenting us with a lower instead of an upper bound.

As described in section 2.1, we can measure the channel capacity of the phosphorelay using the proposed setup (see Fig. 4). For this, we simulate the system with increasing stress levels. In the model, this means that depending on the level of stress, the initial phosphorylation rate k_1 is linearly downscaled until a basal level of activation is reached. This influence of the turgor pressure on the ensemble of Sln1 molecules is an important assumption that will be discussed later. We sample the probability distribution of the output species Ssk1 over time depending on the parameter sets used for simulation. By doing this, we simulate the time courses of the transmission probabilities $P(Y|X)$ that define the channel (visualized in Fig. 14). We now could directly compute the channel capacity if we considered a certain distribution $P(X)$ for the input. Yet, this natural distribution is not known to us and it is the incentive that

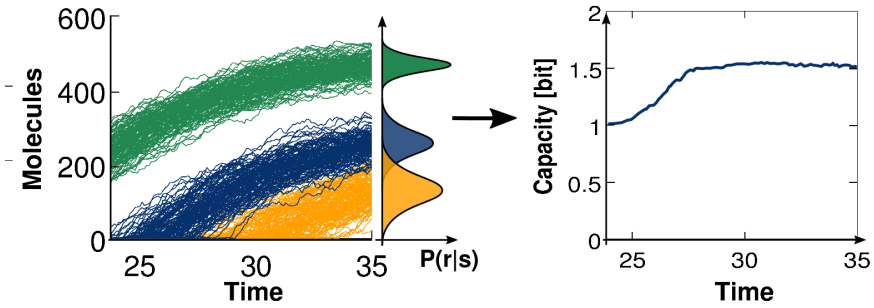


Fig. 14: The time course simulations give rise to the conditional probabilities $P(\text{response } r|\text{stress } s)$ at time t. Subsequently, the corresponding capacities can be calculated and give a dynamic view on signaling to answer the question: "How much can the system infer reliably at time t?"

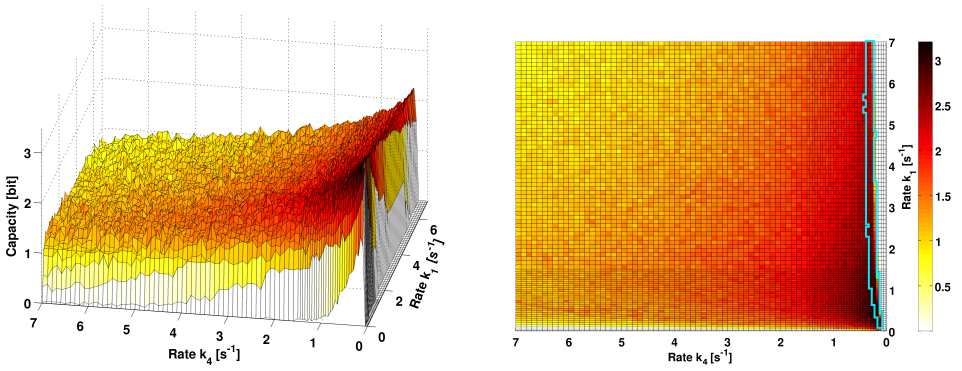
the channel has adapted to. So, in a further step we will have to find another way of computing capacity. As an aside, employing a moment closure of this reaction system would be an alternative to the Monte Carlo approximation. When using this approach, it is important to keep in mind that the derivatives of the second moments (as they are products of our random variables) starting at a sharp initial distribution can have very steep slopes and thus make solving the ODEs numerically problematic, as a grid for solving below a certain error threshold must be very fine grained. For our purposes, the simulation results using the SSA are sufficient to approximate the probability distributions in a manageable time window and are thus preferred for our use.

This leaves us with the optimization problem of finding the input distribution that achieves the capacity, i.e. that fits the channel. We do this by employing the Arimoto-Blahut-Algorithm (see section 2.1.2) that finds the maximum capacity as well as the achieving input distribution numerically. Interpreting these optimal input distributions would be interesting but is not part of the analysis here. Generally, it will look sharper than it could be the case in a natural setting, which is also due to the binning process that we get by choosing a number of inputs (for numerical reasons) instead of a continuous range of concentrations. The number of inputs that we subject the system to is determining an upper boundary to the capacity as can be seen by Eq. (3). However, capacity will usually saturate at a much lower level because of the

noise that the system exhibits. This has been visualized in Fig. 15 and is part of the discussion in section 2.1.2.

The maximum value for information transmission is found at a low activating rate k_1. Increasing this parameter introduces a higher variability and thus more noise in the system as can be seen in Fig. 15. Although the absolute amount of capacity has to be debated (see section 3.5), a capacity of 3 *bit* means that the cell could potentially identify a number of 8 distinct signals. Looking at the landscape of capacities, we observe in the system a sharp transition from non-informational signaling to full capacity in the lower regimes of k_4. This suggests a sensitivity of the system that the cell will have to either overcome or use to its advantage.

By connecting the analyses for steady state phosphorylation and the channel capacity, we observe a narrow margin (Fig. 15) that is viable for simulating the phosphorelay model.



(a) We observe a steep gradient for the capacity in the regime of a low dephosphorylation rate k_4, implying a strong sensitivity in this parameter.

(b) The combination with the results of Fig. 13 (see contoured area), the analysis restricts our parameter space strongly.

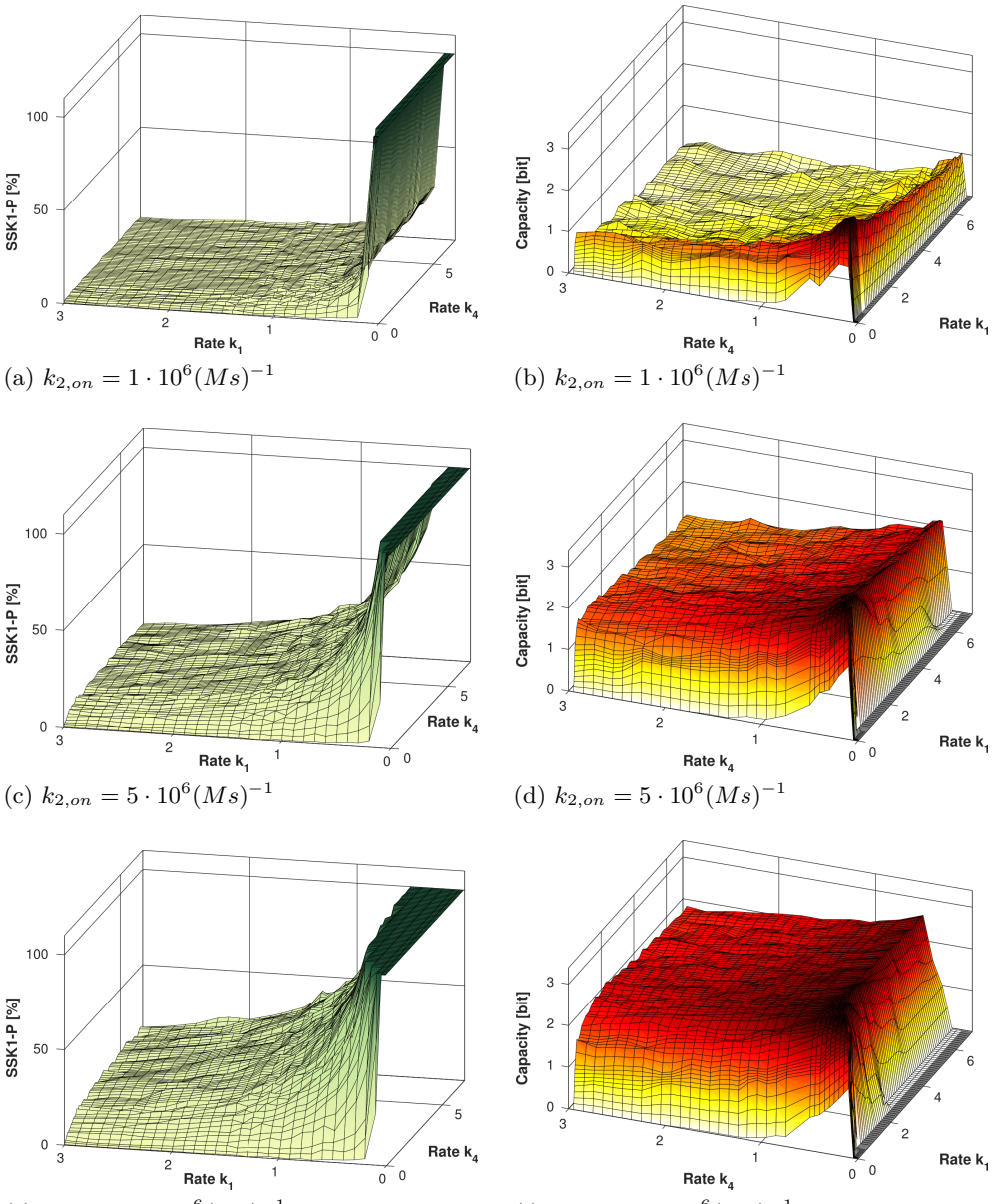
Fig. 15: Capacity as a function of the input rate k_1 and the output rate k_4 (with fixed $k_{2,on} = 5 \cdot 10^6 (Ms)^{-1}$, see Tab.1 for other parameters).

3.4. Improving information processing in yeast osmotic stress response - "The story of on and off"

In collaboration with C. Kiel and the lab of Luis Serrano at the Centre for Genomic Regulation in Barcelona we apply this modeling approach to examine the role of kinetic constants within this signaling motif. For this, we focus on the association and dissociation of the complex between Sln1 and Ypd1. As already described, we are aiming at increasing $k_{2,on}$ and $k_{2,off}$ in a way that their ratio, the affinity for the complex building, is constant. In an experimental setup, this was realized through electrostatic engineering in the case of $k_{2,on}$ and a mutation of hot-spot interface residues that increases the dissociation rate $k_{2,off}$. As described in section 3.2, we mimic this alterations by varying the association rate $k_{2,on} = \{1,\ 5,\ 10\}\ \cdot 10^6 (Ms)^{-1}$ and kept the dissociation constant at a value $K_d = \frac{k_{2,off}}{k_{2,on}} = 300nM$. These values resemble suitable rates and affinities as occurring in natural settings, yet have not been adjusted to the experimental data for the model at hand (in preparation). Nevertheless, we can already observe interesting features.

What is most prominent in the performed simulations (see Fig. 16) is the bottleneck arising through low association and dissociation for the complex between Sln1 and Ypd1. We can see that for most parameter combinations the pathway exhibits a constitutively active signaling. Ssk1 cannot be phosphorylated sufficiently to attenuate the downstream signal. This is the result of an accumulation of the complex between Ypd1 and Sln1. The upstream signal is "buffered" in this accumulation and further transmission of the signal is too low to allow for a constantly phosphorylated Ssk1 to attenuate the signal in the pathway. This bottleneck removes the ability of the input to exert control over the pathway. A second important observation concerns the transmission of information. While for a small set of (very low) parameters we still see functional information transmission > 1 *bit*, we observe a strong decrease in overall fidelity with low association rates. This is a critical point as it means that this modulation impairs severely the systems ability to signal and thus fulfill its functions. It would be interesting to compare this to the behavior as observed in an experimental setup.

These observations are part of a larger extension to the model that is able



(a) $k_{2,on} = 1 \cdot 10^6 (Ms)^{-1}$

(b) $k_{2,on} = 1 \cdot 10^6 (Ms)^{-1}$

(c) $k_{2,on} = 5 \cdot 10^6 (Ms)^{-1}$

(d) $k_{2,on} = 5 \cdot 10^6 (Ms)^{-1}$

(e) $k_{2,on} = 10 \cdot 10^6 (Ms)^{-1}$

(f) $k_{2,on} = 10 \cdot 10^6 (Ms)^{-1}$

Fig. 16: Simulations with varying association and dissociation rates $k_{2,on}$ ($k_{2,off}$ is scaled accordingly to keep the ratio K_d constant, see Tab.1 for other parameters) as discussed in section 3.2 and 3.4. Low rates lead in a larger degree to constitutive signaling and at the same time exhibit lower information transmission capability.

to predict and explain *in silico* the behavior as seen *in vitro*. Our findings substantiate the hypotheses as obtained from the experimental observations. The work is still in progress and will be published in the near future.

3.5. Discussion

Within this chapter, we performed an analysis of the phosphorelay module in the HOG pathway by interpreting the system in an information theoretical way as a channel that transduces environmental cues to inner-cellular decision centers. Our aim was to identify how, given a stochastic biochemical nature, the phosphorelay system translates an input (decreasing phosphorylation of Sln1) into its output (dephosphorylated Ssk1). We focused on the question of fidelity that this system can achieve despite inherent noise.

Enabling fidelity

We observe that the fidelity and thus diverse response patterns of the HOG pathway (see e.g. Macia *et al.*, 2009, Fig.1) might already have its origin in the first step of osmo-sensing studied by us. The capacities that potentially could be achieved with the phosphorelay, provided a suitable input function, exceed an on/off response that would correspond to a capacity of 1 *bit*. We consider this with several implications:

With our analysis we gain a method of estimation on how to restrict the parameter space of the model. Here we regard 1 *bit* as a lower bound on information capacity. If this would not be achievable by the system under the configuration in question, a functional adaptation of the cell to the osmo-stress would be error prone and probably insufficient.[43] Thus, we can conclude to disregard parameter sets with capacity below 1. The capacity that exceeds this boundary nevertheless is not necessarily used by the cell. As discussed below,

[43] It is important to note that an error prone system can still survive, since even a random choice can by chance be the right one and cellular systems could make trade-offs between risk and investment. A very interesting discussion of that can be found in Voliotis *et al.* (2014). Nevertheless, we argue that in the case of life-threatening stresses, this risk will have to be minimized over evolution or shared (similar to a bet-hedging strategy) over a population (see e.g. Kussell and Leibler, 2005). In our case, reliable sensing should be preferred to stochastic switching.

the mechanism of the input plays an important role in this regard. But we also need to consider the distribution of the external variables as well. Bowsher and Swain (2012) used a concept related to that of mutual information, called "informational fraction". Similar to our output of the Arimoto-Blahut algorithm, they draw conclusions on how the pathway could potentially have evolved to adapt the cell to a certain scenario of environmental state distributions.

The structure of the phosphorelay as observed in experimental settings (Janiak-Spens *et al.*, 2005) as well as the molecule numbers that the system works with are crucial to enable this signaling branch to work and seem to follow a fine-tuned optimization. For one, we showed that the region of parameters in which the system is neither constitutively active nor unreliable is a narrow margin and needs to follow a tightly defined range to be viable. This is in particular the case for the de-phosphorylation rate k_4. Another structural feature of the phosphorelay are the ratios of signaling molecules and transfer rates in the cascade. These numbers fulfill two important points: a) Bottlenecks within the cascade are avoided and b) the resolution of the noisy output can be matched to input levels in a noisy-typewriter fashion. The first point is going hand in hand with our choice to set a threshold on basal signaling. A bottleneck (e.g. through low Ypd1 expression or insufficient transmission strength between the layers) would lead to a constitutive activation of the pathway and like in knock-out mutants for Sln1 or Ypd1, to a lethal phenotype. Changing the association and dissociation of the Ypd1-Sln1 complex has shown such a bottleneck. Lower association rates lead to higher leakage of the pathway as can be seen in Fig. 16 and must be avoided for the functionality of this system. In addition, this impairs the transmission of information severely. The second important point b) is concerned with the overall achievable resolution of the pathway. The molecule numbers as observed in the yeast Sln1-phosphorelay represent a good example of how a signal transduction mechanism can cope with inherent biochemical noise. If we interpret the number N of receptors as a resolution (in analogy to pixel on a screen for example), we can tell that there is a maximum of 2^N states (on/off) that can be transmitted. Even if the intermediate species follows the rules of a) and allows for the efficient transmission of the on/off signal for each receptor protein, the noisy nature of molecular interactions and particle movement will distort the signal and map

to a (more or less broad) conditional probability distribution, depending on the channel transmission properties. Thus, as visualized in Fig. 8, the resolution determines a subset of distinguishable inputs like a noisy typewriter. In the case of the output species Ssk1, this resolution roughly doubles the resolution of the receptor level and facilitates a reasonable level of fidelity.

An important point that can be addressed, when extending the signaling pathway to include the downstream Hog1 and its implications, is the encoding of signals in other features. As with constructed codes in computer sciences, information can be stored in different characteristics, like absolute activation, signal duration, time until full activation, area under the curve, frequency responses and many more (e.g. Cai *et al.*, 2008; Mettetal *et al.*, 2008; Locke *et al.*, 2011; Hao and O'Shea, 2012; Batchelor *et al.*, 2011). Information theory can address such features as well, since it is merely a matter of definition of the input or output events. Choosing any of the potential functions and measuring the corresponding probability distributions then gives the opportunity to incorporate more knowledge into estimating responses and inferring inputs from observables. This approach has for example been used with time courses (Tostevin and Ten Wolde, 2009) and vectors of timepoints in single cell data (Selimkhanov *et al.*, 2014) to increase the observed capacities. In a way, this seems trivial (after all, using a single timepoint for a complex dynamic process will naturally have lower information content than using the full description) but these approaches show an important direction for future research and take many features of dynamic cell responses into consideration. In this way we can learn about what cells actually use to encode and transmit information.

Capturing efficiency

Furthermore, we observed a pattern of information capacity that prefers low reaction rates over faster ones. Although slowly, capacity decreases towards a higher auto-phosphorylation rate k_1. This can be explained by the increase of variability (and thus a lower signal to noise ratio) at higher rates. This observation underlines the intuitive notion that cells also try to optimize their energy consumption. Since the first reaction of Sln1 auto-phosphorylation is constantly consuming ATP in order to keep Ssk1 phosphorylated downstream,

lower rates could be preferred for efficiency. As the results of our study have shown, a distinct signal transduction in a reasonable time window is still manageable by the cell. Further study of the rates could validate this finding, as it would also be very interesting to observe such an optimization in an experimental setup.

In our study, we observed a steep slope of capacity of the channel (see Fig.15). This means that within that small region, we change very quickly from no information transmission to a good signal transduction for the system. This sensitivity allowed us to put very sharp boundaries on the parameter space, thereby explaining the regimes of functionality in our model without the need of fitting it to data.

Interpreting the amounts of capacity that we observe in our system is a debatable topic. It is impossible for us to know if that information can and will be used. What a biological system neglects and what it actually infers through different channels can't be evaluated. As a matter of fact, many studies observe relatively low capacities (e.g. Cheong *et al.*, 2011a, Fig. S1), even below 1 *bit* suggesting that not even a reliable inference of an on/off signal would be achievable. Yet, Voliotis *et al.* (2014) note that this is not necessarily what nature needs. It is possible that an unreliable inference is sufficient enough on a population level. A cell that does not adapt properly would potentially not survive the stress, yet investing into better sensing mechanisms could be a costly task and thus, the optimal strategy would be to live with uncertainty and run a (possibly only small) risk of not being able to cope with the environment. On the other hand, as was noticed earlier, Selimkhanov *et al.* (2014) show an often employed feature, where the information about a response is not only stored in single timepoints or particular features, but in multiple characteristics. In this case, it is the encoding of information in certain temporal profiles. Yet, other encodings are certainly possible and probable. An important example would be to regard cell populations that communicate and thus share and increase information. In our case, the initial response of the phosphorelay suggests that it is able to convey even more than just an on/off switch. The distinct temporal profiles downstream of Ssk1 hint into the same direction, yet the use of feedbacks and adaptation is integrated in those profiles as well. So

for now, it is not possible to disentangle this complex response and tell how much the phosphorelay contributes. Integrating this approach into a larger setting could present answers to these remaining questions.

Choosing the input

When we performed our simulations, we assumed a linear input that turgor pressure has on the phosphorylation of Sln1, namely the linear decrease in k_1. Although there is still ongoing research on the topic (Tanigawa *et al.*, 2012), the mechanism itself has not yet been characterized comprehensively. Neither has the stochastic influence of the whole ensemble of Sln1 sensing the external signal been studied. So the question remains: Is the linearization of the input function a valid assumption? This question will have to be answered experimentally. Our analysis shows that many different choices for this input can result in a similar behavior. In the extreme case, the input to Sln1 would be an "on/off" for the phosphorylation rate. As can be seen from Eq. (3), this would limit the absolute value of capacities with an upper bound[44], but neither the observations on functionality nor a information transmission (even if minimal) are impaired by this.

We regard the channel as an evolutionary evolved and thus fixed property. This is an important prerequisite, but under natural circumstances this is hard to prove. How does the channel change when stress is applied? Will there be an adaptation that the cell performs *ad hoc*? Whether our findings on information transmission will hold in a living cell remains to be seen in an experimental setup, as this is the only way to observe what a cell's behavior will be. Yet, the beauty of the applied methods is that they are not restricted to analyzing a mathematically modeled system, but can instead also be used to evaluate mechanisms and motifs solely based on observing the noisy input-output-relation as an information transmission problem. Provided it is possible to sufficiently capture the stochastical nature experimentally[45], we believe that this is a powerful tool to find functions and characterize biological systems and ultimately connect theoretical and experimental work.

[44] In the extreme case with 1 *bit*.

[45] This requires a sufficiently large set of single cell data to approximate the conditional probabilities needed for the analysis.

4. Crosstalk in Yeast signaling - "Conducting Information" or "To talk cross or not, that is here the question"

Signaling not only in biological systems, but in a broader general sense consists of myriads junctions and turns. The best example for this is our world wide web, a network patched together by millions of nodes and connections, able to transmit reliably countless numbers and messages over huge distances. In biological terms, these networks are comprised of interacting protein species that employ many different mechanisms to signal the cell a biochemical visualization of its environment. In yeast, the highly conserved mechanism of MAPK cascades (Widmann *et al.*, 1999) facilitate many of the functions needed to not only visualize this environment, but also embed the cell's behavior into these cues. An interesting feature that can be found in those signaling motifs is the crosstalk by using identical species for information transmission. In this chapter, we extend the description of the HOG pathway and add the network of α-factor signaling, the pheromone (mating) pathway, to that. We present and evaluate a modeling approach based on data of temporal phosphorylation profiles of the species involved (Vaga *et al.*, 2014).[46]

4.1. The full HOG pathway

In the previous chapter, we described a very important part of the HOG pathway, the Sln1-branch. Yet, as mentioned, this is only one part of how yeast cells are activating their adaptive responses. Alternatively the so-called "Sho"-branch can facilitate the same activation by other means of sensing. Both branches converge at the phosphorylation of the MAPKK Pbs2 which in turn activates Hog1. While for a weak stress response each of the branches is sufficient, for the adaptation to be perfect yeast is in need of both of them (Tanaka *et al.*, 2014).[47] The mechanisms that activate those branches have

[46] The data are available at `http://www.cellnopt.org/data/yeast/`.

[47] **Remark:** Amongst others, this recent research also suggests that the Sho-branch actually splits in even more redundant ways. Since only one input in this branch is important for

been focused by many studies over the last two decades and still are providing insights into how sensing mechanisms in biological systems is facilitated (Maeda *et al.*, 1994; Janiak-Spens *et al.*, 1999; Tanigawa *et al.*, 2012). In the following, we give a short introduction into the functionality of both signaling branches and the full HOG pathway.

For the description of the phosphorelay in the Sln1-branch, we refer to chapter 3. In this model of the phosphorelay, we regarded Ssk1 in its unphosphorylated form as the output. Yet, the pathway continues, as Ssk1 is catalyzing an activation of the (redundant) MAPKK kinases Ssk2 and Ssk22 downstream of it (Posas and Saito, 1998). Their phosphorylated state is able to promote a phosphorylation in the MAPKK of the signaling cascade, the scaffold protein Pbs2. Two phosphosites that have been shown to be of vital importance are located at the Ser514 and Thr518 (Maeda *et al.*, 1995; Wurgler-Murphy *et al.*, 1997; Soufi *et al.*, 2009). Due to the same reason as mutations in the phosphorelay, mutating those sites to an aspartic acid will lead to a constitutive activation of Pbs2 which is lethal due to the excessive production of osmolytes. This also shows the important activation characteristics of those two phosphosites, that catalyze the phosphorylation of MAPK Hog1 at Thr174 and Tyr182 (Murakami *et al.*, 2008). This double phosphorylation has been identified as the main activating factor for the subsequent adaptation. Hog1 is regulating many downstream responses Petelenz-Kurdziel *et al.* (2013), most importantly the enzyme Gpd1 which is facilitating the subsequent synthesis of glycerol to counter the osmotic pressure put on the cell from the environment.

The second activating branch for this pathway is regulated by Sho1. We regard this species as the input for the pathway, yet its activity has been shown to depend on the transmembrane osmo-sensors Hkr1 and Msb2 (O'Rourke and Herskowitz, 2002; Tatebayashi *et al.*, 2007). Sho1 assembles the downstream MAPKK Pbs2 to the cell membrane upon osmotic stress (Raitt *et al.*, 2000), facilitating the interaction with Ste50 and the MAPKKK Ste11 (activated through Cdc42-bound Ste20), which in turn activates Pbs2 by phosphorylating

our study on crosstalk, we merely mention these results for now and further lump those mechanisms together in one single activation of Sho1 by osmotic stress.

it at exactly the same residues as in the case of Ssk2. Thus it conducts the flow of information through the HOG pathway, if stimulated. The interaction with Ste20, Ste50 and Ste11 is an unusual feature, as these species are also observed to be involved in the signaling MAPK cascades for filamentous growth and pheromone sensing. The pheromone pathway be introduced in the next section.

4.2. The yeast pheromone pathway

A critical checkpoint within the cell cycle is the transition from Growth Phase 1 (G1) to Synthesis (S) Phase, denoting the commitment to either asexual reproduction via mitosis or the mating with a partner Herskowitz (1988). The latter enables yeast cells to form diploid cell types that can undergo sporulation under challenging environmental conditions to survive. Yet for mating to be possible, haploid yeast cells need to communicate with one another. This is done by producing either a- or α-factor, pheromones that indicate the presence of potential mating partners nearby and can be sensed by the corresponding opposite cell type. Depending on a sufficiently strong dose of a- or α-factor (signaling a mating partner in the vicinity) (see also Dohlman and Thorner, 2001; Moore *et al.*, 2008; Hao *et al.*, 2008), this facilitates the "shmoo-ing" and ultimately a fusion between the cells, creating a single diploid cell that then can undergo sporulation and meiosis. However, before that another crucial reaction must be triggered: The cell cycle in both mating partners needs to be stopped first to prevent them from entering S phase. This synchronization is a prerequisite for mating. Yet, not only is this pheromone induced cell cycle arrest important for yeast populations in their natural habitats, but it plays a significant role in experimental setups as well (e.g. see Fantes and Brooks (1993) and Breeden (1997)). Using haploid MATa yeast strains that are unable to switch their mating type is one of the main techniques to achieve a homogeneous and synchronized cell culture that can be studied after release from the pheromone treatment. Here we want to introduce the structure and function of the pathway enabling this synchronization: The yeast pheromone pathway.

The input to this pathway is a G Protein-Coupled Receptor (GPCR), consisting of the transmembrane proteins Ste2 or Ste3 (depending on the mating type)

that accept binding of a- and α-factor, the G_α unit Gpa1 and the $G_{\beta\gamma}$-subunit comprised of Ste4 and Ste18. Upon osmotic shock, the scaffold protein Ste5 translocates to Ste4 at the membrane, tethering the MAPK cascade including Ste11 (MAPKKK), Ste7 (MAPKK) and Fus3 (MAPK). Again, as in the HOG pathway, Cdc42-bound Ste20 is able to phosphorylate Ste11 with the help of Ste50, thus activating the first step of the kinase cascade tethered by Ste5. In successive activation over Ste7, the MAPK Fus3 is phosphorylated at Thr180 and Tyr182 (Gruhler *et al.*, 2005). This kinase is considered the output of the pathway. It induces for example the activity of the transcription factor Ste12 and phosphorylates Far1, the factor arresting the cell cycle and involved in cell polarization.

Because of its key role in experimental techniques and its well established functionality, scientific research has focused on this interesting pathway on many occasions over the last two decades. Nevertheless, as with the HOG pathway, it still produces surprising results and nurtures important insights into the functionalities of signaling systems.

4.3. Overlap in Signaling networks - Mass-spec-data from Crosstalk in yeast

A very intriguing feature of these two pathways and also one of the reasons for them to be studied this well, is the degree of crosstalk within them. As both are sharing components (Ste20, Ste50 and Ste11) in their activating branches, the two pathways could under certain circumstances activate the wrong protocol for one of the two stresses. Yet, the cells react in quite a distinctive and accurate manner: Using one "channel" for multiple signals doesn't disturb their function. This has also been observed experimentally (McClean *et al.*, 2007). How do they maintain their signal specificity? This question has drawn a lot of attention and is still fueling research (e.g. Schwartz and Madhani, 2004; Patterson *et al.*, 2010b; Schaber *et al.*, 2012; Baltanás *et al.*, 2013). Vaga *et al.* (2014) have been performing extensive phospho-proteomic studies on yeast cells in order to determine the extend to which the phospho-proteome is affected by the two stresses. This includes not only each stress applied separately, but measurements of consecutive excitation with both pathways for different time

periods sampled at 0, 1, 5, 10, 20 and 45 minutes after stress induction. Overall, this results in 36 experimental conditions, where the aforementioned time courses are included. The study reveals that more residues of involved signaling proteins seem to be differentially regulated by the pathways and by crosstalk. We were interested in using this data to model potential regulation patterns and test the data against the knowledge obtained in previous studies on the insulation mechanisms of the two pathways. The data seemed to provide a solid base for fitting a reasonable model to it and extending it to more functionalities.

The work of Vaga *et al.* (2014)[48] shows a complexity in crosstalk regulation between the two pathways, that is higher than previously thought. The data comprises a number of 2536 detected phosphopeptides, including 57 that are associated with either the HOG or the pheromone pathway. A number of known phosphosites in the two pathways was detected, foremost to be mentioned the critical Hog1 phosphorylations at Thr174 and Tyr182 as well as Fus3 phospho-residues Thr180 and Tyr182. The behavior of these outputs was used as a benchmark for the validation of the data.

Yet, the relevance of the regulatory interactions found as well as their implications cannot be covered by the analysis. As a matter of fact, several critical nodes in the known structure of the two pathways were not detected by the experimental setup, leaving fundamental gaps for evaluating relations and causalities. We reviewed the literature extensively in order to validate the mechanisms that were previously shown to function in the two pathways and can be considered as vital. Tab. 2 lists important phosphosites reported in the literature, that are not contained in the data. The reasons for those sites being missing can be various. For one, the experimental techniques are not able to detect phosphorylations on aspartic acid or histidin residues. This concerns for example the phosphorelay in the Sln1-branch of the HOG pathway, which consists of a phosphate transfer only involving phosphate transfer via Asp and His residues. Another reason can be that proteins are not of sufficient length to be detected, as in the case of Ypd1 which has only a length of 167 amino acids and thus is below a detection threshold.

[48] (See also Vaga, 2013).

Protein	Phosphosite	References for the evidence (PMID)
'FAR1'	'T306'	8334305, 9148934, 9632750
'GPA1'	'S200'	21521692
'KSS1'	'T183'	1628831, 8668180
'KSS1'	'Y185'	1628831, 8668180
'PBS2'	'S250'	19823750
'PBS2'	'S514'	9032256, 9180081, 7624781
'PBS2'	'T518'	9032256, 9180081, 7624781
'SSK2'	'T1420'	9482735
'SSK2'	'T1460'	9482735
'STE11'	'S302'	10837245
'STE11'	'S306'	10837245
'STE11'	'T307'	10837245
'STE7'	'S359'	8131746
'STE7'	'T363'	8131746
'STE50'	'S155'	20932477,1885432
'STE50'	'S196'	20932477,1885432
'STE50'	'S248'	20932477,1885432
'SLN1'	'H576'	2957298
'SLN1'	'D1144'	2957298
'YPD1'	'H64'	2957298
'SSK1'	'D544'	2957298

Table 2: Summary of important phospho-sites reported in literature, that were not detectable in the mass-spec approach of Vaga *et al.* (2014). The His and Asp within the phosphorelay could not be detected due to the experimental procedure aiming for Ser, Thr and Tyr. The information is obtained from `http://www.phosphogrid.org`.

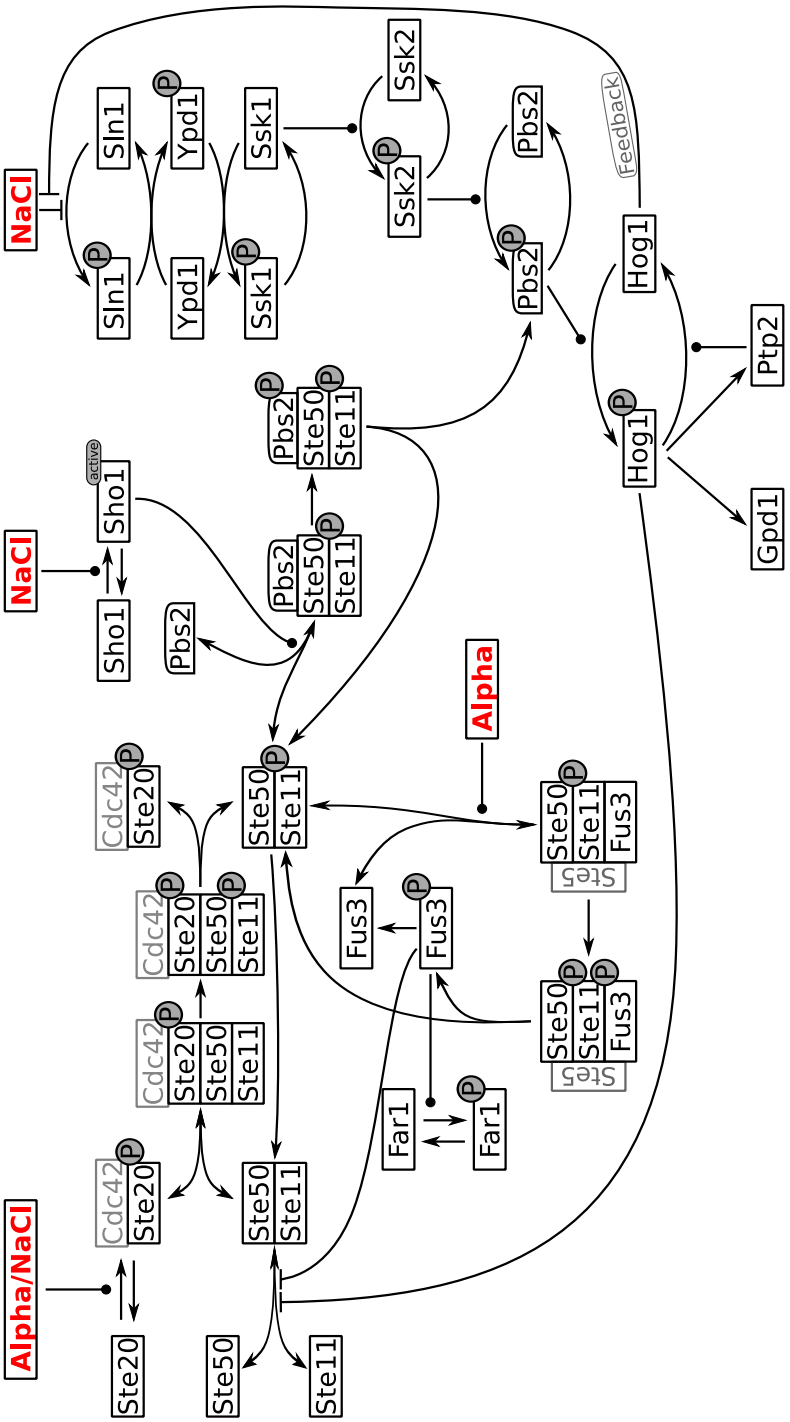


Fig. 17: Schematic of the crosstalk model between HOG- and pheromone-pathway, incorporating the main mechanisms as reported in literature.

4.4. The model

We were interested in the modeling of potential regulation patterns that could provide a mechanism for the temporal profiles seen in the data and to study, which regulation patterns of the data could potentially be described by a mechanistic model. In the context of chapter 5, the aim was to examine the data set for features of insulation and how further mechanisms could influence the crosstalk in the pathways. For this, we built a model of ordinary differential equations that combines and harmonizes insights from a broad range of past research to a comprehensive view of the signaling mechanisms in the two pathways. The model describes the events as introduced in section 4.1 and 4.2. A schematic structure of the model can be seen in Fig. 17. We carefully reviewed the known interactions and incorporated them to a combined, but in parts simplified view. In order to create a parametrization for the model, we turned to the data and looked for reasonable overlaps. The phosphopeptides HOG1_174_Y176 and FUS3_T180_Y182 were included in the data an thus used for fitting procedures in our modeling approach. The implementation was done using "Data2Dynamics"[49] introduced in Raue *et al.* (2013). The ODE equations as well as parameters and initial values for the model are documented in Appendix A.

Further, species showed a large variability in phosphorylation patterns compared to the anticipated behavior. For example Ste20 facilitates many interactions and thus showed also a large number of regulated phosphopeptides upon both stresses. Yet, the regulation was unspecific to a certain stress. Thus, we decided to disregard the data for now.

As a result of the modeling approach, we obtained a mechanistic view on the two signal transduction pathways that captures the anticipated behavior sufficiently, as can be seen in Fig. 18. In particular, the strong influence of the scaffold proteins Ste5 and Pbs2 could be observed. We see the regulating power of this biochemical functionality in the maintenance of specificity between the two signals. In the model, Ste5 activation of the MAPK signaling cascade Ste50-Ste7-Fus3 depends on the recruitment to the membrane at Ste4 and

[49] Available at `http://www.data2dynamics.org/`. The environment was simulated on MATLAB R2015a.

initiates as well as insulates the signal. A similar behavior can be seen by the membrane located signal initiation process of the Sho1-branch in HOG signaling. This behavior already arises with a very simple (and non-biological) parametrization of the model, owed to the strong control that these tethering functions exert in the pathways. This mechanical way of insulating signals thus arises from structure already, emphasizing its importance for biological signal transmission. Even though the modeling process was based on the findings of phospho-proteomic data, we observed a large disparity between our simulations and the data. This proposes that with the knowledge available at the moment, it is not possible to explain the full extend of the data set and further investigations are needed.

As discussed, the data obtained by Vaga et.al. shows that many previously undiscovered regulation patterns of phospho-sites are existent in a crosstalk-like manner between the HOG- and the pheromone-pathway. Further experimental studies are required to either show the significance of the individual regulations or whether certain sites are only byproducts of the crucial signaling. Nevertheless the data depicts a profound basis to assess further aspects of the complex network and its inter-dependencies.

Remark: The publication of (Vaga *et al.*, 2014, Fig. 8) includes a parallel developed modeling approach employing the boolean ODE simulation framework "CellNOpt" published in Terfve *et al.* (2012). This modeling approach is based on a similar consensus network of the pathways (yet models the separate phosphosites that were detected) and allows for model variant construction by adding or subtracting interactions in the network. This model was used to validate certain proposed mechanisms in the paper by introducing edges in the model graph between different phosphosites. This approach shows that the model then gains a better quality of fit, yet the extend of this is unclear. Objectively seen, the qualitative behavior of most of the peptides was also imperfect.

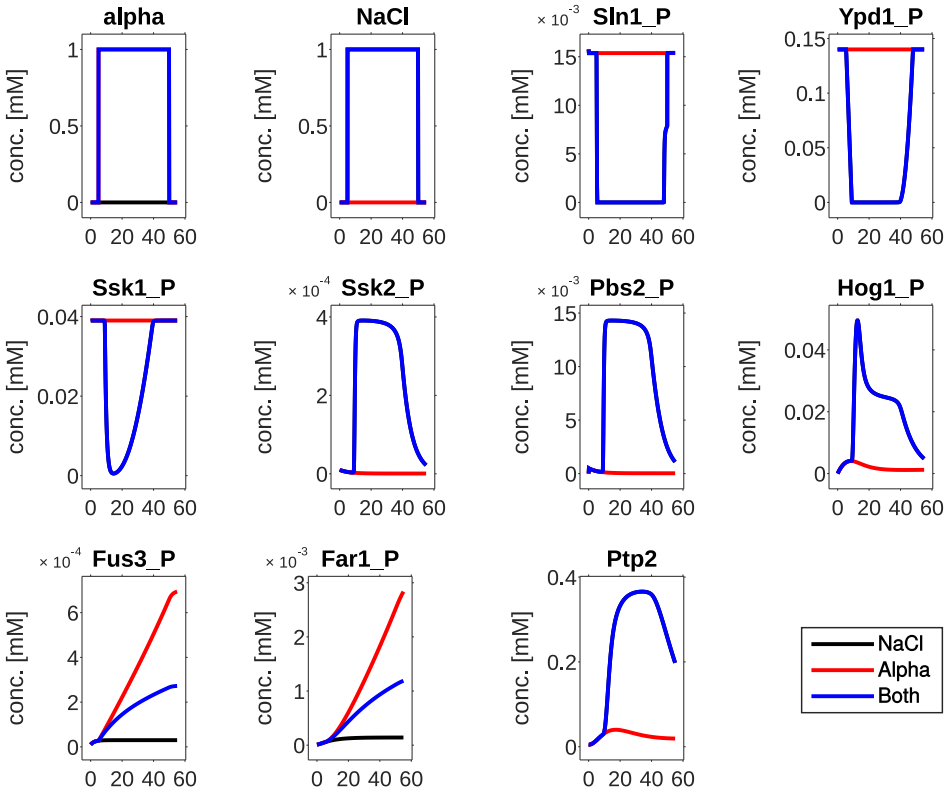


Fig. 18: Simulated time courses for an important subset of the species in the model. The expected activation patters can be shown with the model. Especially appealing is the influence of scaffolding by Ste5 as well as the assembly of Ste20, Ste50, Ste11 and Pbs2 via the activation of Sho1. Both mechanisms insulate efficiently as was observed in previous research. For parametrization, see Appendix A, Tab. 4 and 5.

Consecutive activation patterns The mechanisms that have been identified so far for the functionality of both pheromone and HOG signaling are to a large extend consecutive activations in chains of signaling. There exists an interesting overlap between species and transient as well as longterm feedbacks factor into the system as well. Despite those non-linearities, the chain of activation will (until a certain point) follow an order of activation. In the face of the phospho-proteomic data, a reexamination of this has been suggested. Our modeling approach has shown as well that the behavior of most of the detected phosphosites does not follow the picture painted by scientific studies until now. And while the data exhibits unfortunate gaps, that would (if filled) presumably show exactly the anticipated consecutive activation pattern, it also suggests that possibly more mechanisms facilitate such a chain. This could be phosphorylations (or the missing thereof) that unlock certain conformation changes and subsequently allow further activations in the cascade. Yet as to what these mechanisms are, it will be inevitable to go into molecular details that are not in the scope of a modeling approach. Since the significance of the detected phosphosites has not been shown and the observed patterns could merely be by-standing phosphorylations that occur more or less without any consequences, it is vital to follow up this research in order to understand the nature of these patterns. The data provides a critical starting point for that, yet so far is limited in unveiling mechanistic regulations by itself. The model contained in the publication (Vaga *et al.*, 2014, Fig.8) validates proposed mechanisms to a certain degree, yet we believe that more detailed experimental research on single phosphopeptides is necessary to substantiate proofs for their significance and functional mechanisms. The data by Vaga et.al. can thus be primarily viewed as a comprehensive, although non-specific basis for further investigations.

4.5. Summary & Discussion

In this chapter we have re-evaluated data obtained by Vaga *et al.* (2014), containing phosphorylation profiles with high coverage of the phospho-proteome in yeast cells that are subjected to osmotic as well as pheromone stress at different time points. Using a high-throughput mass spectrometry approach, the data shows a level of interplay between the two pathways under research,

namely the HOG and the pheromone pathway. It includes previously shown temporal phosphorylation profiles, yet unfortunately the data was not able to cover all critical phosphopeptides, thus leaving large gaps for modeling approaches. We built a consensus model from the literature and compared the outcome of the structure to the observed phosphorylation time courses of the data. We were not able to show the required overlap and even with intensive search of the models parameter space could not find a fit to most phosphorylation behavior. This is not too surprising, as most of the observations were not reported previously. Nevertheless, a functioning system was established in our work and can further be used to include potential mechanisms.

Vaga et.al. employed a crosstalk measure as introduced in Schaber *et al.* (2006)[50]. It would be interesting to see, whether this measure can be reproduced in a modeling approach and be compared to what was observed in the data. Yet, with the low overlap between measured phosphosites and the model species, the comparison for now would be ambiguous until more of the underlying mechanisms that give rise to the observed temporal profiles in the data are elucidated.

As a further extension for our model, including the species Kss1 as a competitor to Fus3 in the binding of Ste7 would be an interesting study. The binding of Fus3 is preferred due to the scaffolding function of Ste5, yet Kss1 has been observed to be transiently activated with a linear dose-response to the input α-factor (Hao *et al.*, 2008). As this interaction facilitates a transient activity and diverts the phosphorylating activation of Ste11 via the MAPKK Ste7 from other potential interactions (Sabbagh *et al.*, 2001; Hao *et al.*, 2008), we believe that including this species could explain one finding of the data that has been disregarded by the modeling part of the paper: No matter how long the cells have been subjected to osmotic stress, an α-factor shock always leads to a strong but transient decrease in Hog1 double phosphorylation. The cause for this is unclear as of yet and since it is maybe the most surprising result of Vaga et.al., elucidating it should be focused on. Kss1 could provide a candidate for such a function.

[50](See Vaga *et al.*, 2014, Fig.4)

5. Scaffolding improves information transmission in cell signaling - "On how to play the right tone accurately"

Gathering information is important for organisms. Gathering the *right* information even more so. Environmental signals themselves are subject to (sometimes severe) fluctuations and adding to that is the noisy nature of the biological processes that make up cellular machinery, including its signal transmission. So how can cells develop a reliable discrimination between "good" or "bad" information at their decision centers and make sure that their responses are appropriately matched to the particular input? The idea of evolutionary development in nature is providing a simple, yet powerful answer to this question.[51] Handling external cues is an adaptational process that - by a procedure of variation, selection and elimination over large time scales - evolves structures that are (optimally) fit to map an external input to a distinct behavior, thus ensuring survival and proliferation. This is thought to be the most important way of biological design, providing robust and reliable processes embedded in enormous complexity.

Technical systems invented by humans did not go through this long-time optimization process, but rather were designed *a priori* in a way that their structure fulfills similar requirements. The need for accuracy and fidelity is a basic requisite in most engineering applications, especially when it comes to the transmission of information. After all, reliable input-output mapping is the basis of controlling nearly any program or machine. To fulfill this, sophisticated ways of signal transmission and processing as well as error correction were invented. As introduced earlier, an important basis for this development were the theories C.E.Shannon formalized in the middle of the 20^{th} century.[52]

This need for a certain precision and a correct interpretation of external

[51] Very interesting takes on evolution can be found in Chen and Nowak (2012) and Rivoire and Leibler (2014).

[52] Both his Master thesis (Shannon, 1938) as well as his monumental paper on information theory (Shannon, 1948).

cues connects the fields of signaling in cellular and engineered systems closely. And even though in biology we face unknown settings with an immense complexity, we can identify motifs that occur frequently and allow us to draw analogies to human designed systems. Thus, we can investigate evolutionary "designed" modules in biological organisms with our means of human technical understanding aiming towards those similarities. In this chapter we revisit such a highly conserved module in cell signaling, the *scaffolding* of proteins, and show that if interpreted in technical terms as a signal transducing channel, this module inherently comprises beneficial properties in its structure already. We argue that this creates a strong potential for selection and thus adds to the conservation of scaffolding in evolutionary development.

5.1. An evolutionary role of Scaffolds - From Accuracy to Crosstalk

Again serving as a prototypical model organism for the discovery of biological concepts, scaffolding structures of signaling proteins have first been described in *S.cerevisiae* (concurrently, yet independently researched and reported in close succession by Chol *et al.* (1994) and Marcus *et al.* (1994)). Ste5 was observed to tether several components of the pheromone signaling pathway, namely the MAPK cascade comprised of Ste11 (MAPKKK), Ste7 (MAPKK) and Fus3 (MAPK). Since then, the motif has been shown to exist in many different pathways and organisms, respectively (Posas and Saito, 1997; Schaeffer *et al.*, 1998; Dickens *et al.*, 1997; Witzel *et al.*, 2012; Dhanasekaran *et al.*, 2007). Not only does this scaffolding motif appear often in cellular signaling, but it also has been observed to fulfill many different roles and functions. The ever growing number of identified scaffolding proteins and their usages has made our perception of cellular signaling far more complex than previously thought, much like the discovery of microRNAs has impacted the view on gene expression regulation (see e.g. Bartel, 2004, 2009, for reviews).[53] This has made the scaffolding motif a very interesting source for fundamental knowledge of signaling systems as well as a promising subject of investigation, especially

[53]Remark: We already noted that the analogy to an information theoretic channel has been applied to gene expression before (Tkačik *et al.*, 2008b, 2009; Tkačik and Walczak, 2011, e.g.) and recently, the role of microRNAs has been included in such an approach as well (Finn and Searles, 2013; Zheng and Kwoh, 2006).

for the field of synthetic biology. Here we will briefly review some of the central roles that scaffolds have been associated with.

The most basic property of scaffold proteins (and their name-giving function as it was the first to be observed) is the assembly and co-localization of signaling molecules within a pathway. This already allows for several important implications: As reviewed in Good *et al.* (2011), it leads to a spatial regulation and coordination of the pathway components (see also Mahanty *et al.*, 1999). It enables them to "steer" the assembled complex to certain locations in the cell (e.g. the cell membrane or other organelles) where signaling is then to be initiated. While this can be used as well as a "pre-assembly" to provide quicker reactions, scaffolds also regulate the concentrations that are experienced locally by the signaling components, as it brings them into close proximity of one another. This effectively increases reaction rates and efficiency of the pathways in question. Scaffolds in this way help to catalyze reaction cascades. Yet, this could also be achieved by increasing the affinity of binding between adjacent signaling layers.

The concentration of scaffolding proteins has another interesting effect on the regulation of signaling. As was investigated both theoretically and experimentally (Levchenko *et al.*, 2000; Witzel *et al.*, 2012; Chapman and Asthagiri, 2009), there exist optimal ranges of scaffold expression due to either a saturation (low concentrations lead to a bottleneck of signal transmission) or a combinatorial effect (high concentrations allow the assembly of incomplete and thus non-responsive cascades). Thus, again in analogy to microRNAs, regulating the expression of scaffolds allows the organism to regulate and fine-tune its responsiveness to certain stresses.

While the aforementioned functions are of a passive nature, scaffolds have also been shown to actively take part in the signaling process in addition to that. Through their tethering function, they increase the specificity and also sequester the activation of signaling molecules (Good *et al.*, 2009). In that particular case, the pathway would only be activating the MAPK Fus3 weakly without the scaffold protein Ste5, making it essential for the transmission process itself. This is similar to the enhancement of reactions as described

in the previous paragraph, yet going even one step further (Sabbagh *et al.*, 2001): Kss1, a MAPK involved in pheromone signaling and invasive growth, is activated by the same MAPKK as Fus3, namely Ste7. It can take the place of Fus3 in the scaffolding complex of Ste5 (for example when Fus3 is knocked out), but does not need the scaffold itself for its activating function. This separates the two scenarios and makes the scaffold not only an enhancer and catalyst, but an integral part of the signaling structure itself. Other active functions, facilitating allosteric activation of Fus3 (Bhattacharyya *et al.*, 2006b) and the use of conformational changes to enable signaling (Sette *et al.*, 2000; Zalatan *et al.*, 2012), have been shown for the scaffold Ste5.

While this is already an impressive number of roles that scaffolds have been associated with, the primary one is thought to be the prevention of crosstalk between pathways that make use of the same signaling proteins (Garrington and Johnson, 1999; Whitmarsh and Davis, 1998; Elion, 1998). As with many other prototypical motifs, this function has first been identified in *S.cerevisiae*. We introduced the mechanism in chapter 4, where the yeast pheromone pathway and the HOG pathway both make use of Ste11, Ste20 and Ste50, yet still are capable of signaling distinct inputs to distinct outputs. Specificity is achieved by the use of the scaffold proteins Ste5 and Pbs2[54] that control the flow of information. This management of crosstalk is one of the most important and also most exciting functions of scaffolds as it hands interesting possibilities to synthetic biology (Lai *et al.*, 2015). Re-wiring pathways and thus re-routing information is an interesting concept and possibly evolved many further functionalities (Bhattacharyya *et al.*, 2006a; Good *et al.*, 2011), yet this would only make sense as a secondary development for an already established system. Additionally, as with the enhancement of signaling efficiency, such specificity could also be achieved through developing separate pathway species with high recognition affinity patterns.

With the many functions and the high conservation of the scaffolding motif in many organisms, the question arises as to what made the re-utilization of

[54] With both fulfilling additional roles (Pbs2 as a MAPKK (Posas and Saito, 1997) and Ste5 as described in this section).

this structure favorable. As far as their usage might be spread, the (sequence) similarity between different scaffolds is very limited and thus, it seems unlikely that their origin is of particularly close relation. We argue that the benefits of scaffolds are especially based on the structure and its properties, making it appear in many different contexts. Functions like inter-pathway insulation, the branching of information transmission routes and higher regulation patterns are likely to be developed as improvements and extensions of an already established and favorable design. We found that a potential reason for evolutionary selection can be found by investigating the signal processing properties of prototypical scaffolded signal transmission. In the following sections, we show that scaffolding limits inherent noise and benefits the fidelity that signaling is capable of. In addition, it shapes dose-response alignment in a near-linear way.

An important observation is that, in contrast to many other functionalities reviewed, the here proposed enhancements of insulated signaling by scaffold proteins are not contradictory to other roles that scaffolds have been shown to play. On the contrary, they propose a strong incentive for reusing and refining this structure and thus evolving into the mechanisms that we find nowadays in nature.

5.2. A model comparison for "mixed" and "insulated" information transmission

Our hypothesis that beneficial signaling properties exhibited by a scaffolded pathway are leading to a preferential evolutionary selection is naturally not verifiable, since evolution is not repeatable for us. Nevertheless, the properties we focus on describe very fundamental levels of scaffolded signal transmission and can set a solid basis for further development of various functionalities. In the following sections we introduce a modeling approach to examine whether we can achieve and recreate such a biological incentive theoretically. This enables us to propose and test potential reasons behind the strong selection of this motif in nature that have not been examined before. We introduce two structural models comprised of the key features for signaling with and without the participation of a scaffold protein and present the results of an information

theoretic analysis. This new view provides insights into the characteristics of potential information processing optimality of insulated signal transmission.

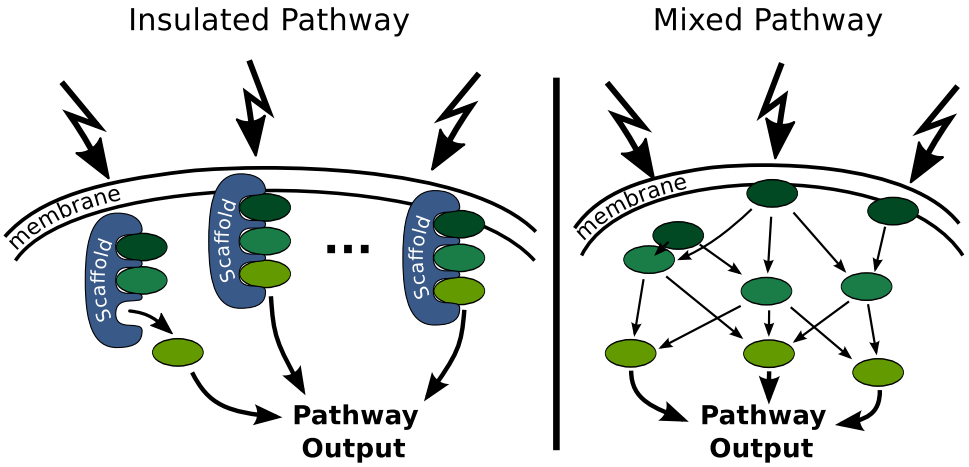


Fig. 19: Schematic representation of the main characteristics of *insulated* (scaffolded) and *mixed* (non-scaffolded) signal transmission. Scaffolds tether signaling molecules together and insulate the transmission process, whereas a free diffusing mixed model enables all proteins of adjacent tiers to interact with one another.

5.2.1. Setting up the models

As a prototypical model structure for signal transduction we constructed a signaling cascade that assumes a free diffusion for the species, similar to standard MAPK cascade models (Goldbeter and Koshland, 1981; Heinrich *et al.*, 2002; Blüthgen and Herzel, 2003; Klipp *et al.*, 2005; Muñoz-García *et al.*, 2009, 2010). Because of the free diffusing molecules, we will denote this model with *mixed channel* in the following. A number of molecular signaling tiers $\{P_1, P_2, P_3, ...\}$ forms a chain of successive activation, as every molecule of a tier can facilitate the activation (e.g. phosphorylation) of all molecules of the next downstream layer. In the model, active protein forms are denoted as P_i^A, the forward and reverse rates as $k_{i,forw}$ and $k_{i,rev}$, respectively.[55] The reverse reaction is counteracting this activation process and ensures that the system

[55] The parameters and their selection will be described in more detail in 5.2.2 as well as the subsequent sections.

can be switched off. Usually this is facilitated by a phosphatase, yet this will not be explicitly modeled here. A reaction scheme for this module can be described as follows (also see Fig.20):

$$\begin{aligned} P_1 &\xrightarrow{k_{1,forw}} P_1^A; & P_1^A &\xrightarrow{k_{1,rev}} P_1 \\ P_1^A + P_2 &\xrightarrow{k_{2,forw}} P_2^A; & P_2^A &\xrightarrow{k_{2,rev}} P_2 \\ P_2^A + P_3 &\xrightarrow{k_{3,forw}} P_3^A; & P_3^A &\xrightarrow{k_{3,rev}} P_3 \\ &\vdots & &\vdots \end{aligned}$$

For each species in the signaling chain, the number of molecules usually shows different expression levels. These numbers can play an important role, since for example low concentrations of intermediate species can create bottlenecks (see Yu *et al.*, 2007). Especially the output species for the cascade is key to determine the resolution with which levels of stress can be differentiated after the signal transmission as well as the total saturation for the pathway. The key feature for the mixed signal transmission is the large degree of inter-connectivity between molecules of successive tiers. Even a small number of activated receptors can induce a "snow-balling" effect on the downstream layers, leading to a switch-like ultrasensitive behavior (Huang and Ferrell, 1996). With our information theoretical analysis we will investigate this from a new perspective.

As an alternative view on the same signal transmission process, we introduce a second structural model that restricts the interaction of adjacent layers, thus capturing the basic property defining a scaffolded cascade. As described in section 5.1, a scaffold protein co-localizes several pathway components within a close proximity, enabling and shaping the transmission of the signal (e.g. a phosphorylation cascade) as well as shielding it from external interference. This model will be referred to as the *insulated channel*. From a modeling perspective, a single (fully assembled) scaffold could be viewed as the state-space of a Markov chain $X(t) = \{0, 1\}^n$ involving the active and in-active forms of the joined protein layers, where n is the number of successive signaling tiers. The state-space is a hypercube that scales with the number of tiers, as we have 2^n combinations of scaffold states. In nature, many systems employ three such stages (Raman *et al.*, 2007; Bardwell, 2005; McKay *et al.*, 2009). In scaffolded

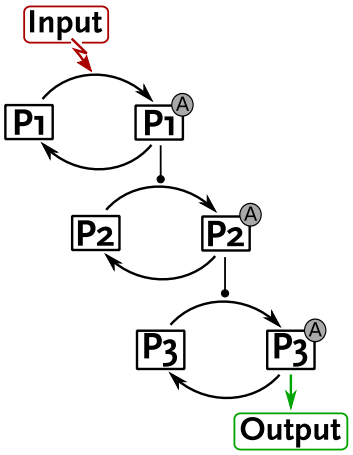


Fig. 20: Schematic representation of the mixed channel consisting of three signaling tiers. The cascade involves inactive (P_i) and active (P_i^A) states of the species, with the active form catalyzing the downstream activation of the next signaling tier.

as well as non-scaffolded signaling motifs, also multiple phosphorylations that can be interpreted as successive tiers (Schüller *et al.*, 1994; Posas and Saito, 1997)) and in addition cascade chains of six tiers (Seger and Krebs, 1995) are observed. The transitions between states follow the activation cascade from the external signal downstream towards the output, directing the flow of information through our scaffold as shown in Fig. 21 for three tiers. The rules for state transitions are fairly simple: "backwards" transitions (corresponding to a de-activation, e.g. through a phosphatase) are allowed from every state, whereas an activating transition is only possible if the adjacent layer upstream of the activated species is active itself. Only one transition at a time is regarded.

In contrast to the mixed channel, the scaffolded signaling is comprised out of m separate copies of this hypercube, each describing one instance of a scaffold protein. The active input level encloses the sum of states, where the first species is equal to 1. If compared to the mixed model, this sum corresponds one to one to the P_1^A layer and is identical with that. The output of the model is the number of activated molecules in the last cascade layer, which is the sum of states with the n^{th} coordinate being 1 (see Fig. 21 green box). So on the receptor as well as output level the two models will have the same interpretations. Yet, the internal information transmission will differ heavily

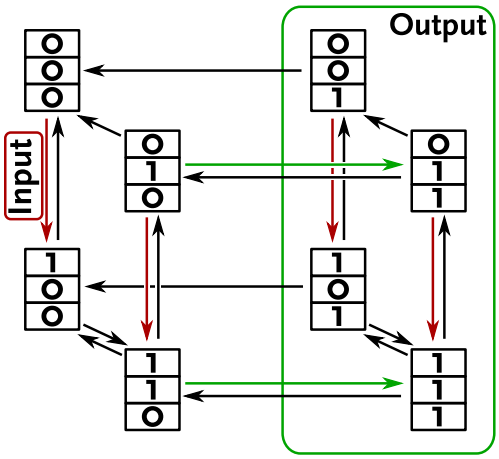


Fig. 21: State-space of one scaffold channel with transitions that can be interpreted as a Markov chain, restricted to three species. States within the green box share an activated third species, thus depicting the output of our model. Red transitions are subject to the input function (stress signal), activating the first layer of the cascade. Each activation from a signaling layer to the adjacent downstream layer requires the activated state of the former, thus giving a natural description of the transitions between states. De-activating transitions are possible in any of the states.

due to the different wirings and thus will produce a distinct input-output relation as well as noise behavior that uniquely characterizes each of the structures.

Remark: In our work, we omit the process of assembling and disassembling the scaffold, as its implications have been researched previously (e.g. Levchenko *et al.*, 2000; Witzel *et al.*, 2012; Chapman and Asthagiri, 2009, as mentioned earlier). Our modeling approach is focusing on the transmission process within the scaffold itself, helping us understand what structural properties a scaffold protein offers as a single entity for signal processing. This gives us a characteristic that is less specific to particular pathways, but more general and comprehensive for this large class of proteins.

In order to investigate and compare the two models using an information theoretical approach, we apply a moment closure of order two[56] to the chemical

[56] Wallace *et al.* (2012) show the applicability of the Linear Noise Approximation to systems like ours. Analogous reasonings can be made for applying the moment closure as well.

master equation that describes these reaction systems.[57] We obtain systems of $\frac{N(N+3)}{2}$ Ordinary Differential Equations (ODEs), where N is the number of species. For every layer i in the mixed channel, active and inactive forms of the molecules are considered a single species each, giving $N = 2 \cdot n$, where n is the index of the final layer. For the insulated channel, the number of species corresponds to the number of states in the hypercube, i.e. $N = 2^n$.

The full set of equations for the moment closure of order 2 for the systems can be found in Appendix B, Fig. 31 (mixed channel) and Fig. 32 (insulated channel). Parameter values for the standard models are given in Tab. 3, subsequent variations for the analysis can be found in the respective sections. The equation systems describe the time-courses for the moments up to 2^{nd} order. This comprises the means $E[X_i]$ for all species as well as the second moments $E[X_iX_j]$ from which we can calculate the variances $Var[X_i] = E[X_i^2] - E[X_i]^2$ and covariances $Cov[X_i, X_j] = E[X_iX_j] - E[X_i] \cdot E[X_j]$, respectively. Solving this ODE system enables us to analyze the dynamic responses to environmental stress and thus capture the stochastic signaling behavior.

5.2.2. Analysis of information transmission accuracy - the Fidelity of signaling

We designed the two models to compare simple but fundamental structural properties and how they influence signaling. In the following sections we want to state our prerequisites and modeling assumptions for that and present the obtained results.

First, we define a standard variant for each model and compare both to one another. In this reference version we assume three signaling layers with molecule numbers of $m_i = 1000$ for each involved layer $i = 1, 2, 3$. This corresponds to the approximate order of magnitude for most of the MAPK signaling molecules in *S.cerevisiae*. These numbers can vary immensely in different organisms and pathways, from cell to cell or (and especially) from experiment

[57] The equations for this can be derived as described in section 2.2.3, yet to circumvent the lengthy calculations we chose to employ the python script provided in Gillespie (2009), which exports the moment ODEs from an SBML description of the reaction system automatically by applying equation (37). The script can be downloaded from `http://pysbml.googlecode.com`.

to experiment (for comparisons, see Thomson *et al.*, 2011; Fujioka *et al.*, 2006; Maeder *et al.*, 2007; Slaughter *et al.*, 2007).[58] This is in part due to the large differences in cell sizes and can also be attributed to functionality pathways. Nevertheless, our findings result from the structural nature of signaling in the two systems and are applicable to a wide range of possible setups. This robustness is examined in section 5.2.3.

In order to make the two models comparable, it is important that we consider the corresponding transmission strengths in a way that adjusts for the spatial effects that scaffolds exhibit. By bringing molecules into a close proximity, scaffold proteins enhance the strength of signal transduction immensely as the molecules experience very high local concentrations. In nature, this enables scaffolded signaling to e.g. react already to very low ligand concentrations and explains the signal enhancing effect.[59] This is achieved by the scaffolding structure, which by assembling the pathway components rids the signal transmission process of its spatial dimension and sequesters stronger activation instead. The forward reactions (and corresponding rates) in our insulated model are thus of first order, whereas the mixed signaling employs second order forward reactions. We correct for this by defining the forward reactions of the mixed channel with pseudo-first-order reaction rates in analogy to experimental techniques. This means dividing the second-order forward rate by the number m of signaling molecules of the corresponding layer.[60] The "signal transmission ratio" or "transmission strength" is defined as the ratio between forward ($k_{i,forw}$) and backward ($k_{i,rev}$, "reverse") reaction, corresponding to the activation and deactivation of the single signaling layers (e.g. through (de-) phosphorylation),

[58] See also `http://yeastgfp.yeastgenome.org/` for specific numbers from literature in *S.cerevisiae*. `http://bionumbers.hms.harvard.edu/` reports $\approx 10\text{nM} - 1\mu\text{M}$ as the "characteristic concentration for a signaling protein" (Milo *et al.*, 2010).

[59] For example, Chapman and Asthagiri (2009) find that Ste5 is expressed at a sub-optimal level and attribute this to their observation that a higher Ste5 concentration sequesters a higher basal activation of the pathway. This shows a high potential for signal activation at low input-levels due to strongly amplified signal transduction rate.

[60] From another perspective it also means that the forward rates of the insulated channel are m-times stronger than those of the mixed channel, as expected from the closer proximity of reaction partners. An equivalent way to achieve this would be this exact notion of enhancing the rate constants of the scaffolded channel by m instead of the downscaling of the mixed channel. The result would be the same, yet from a numerical perspective the lower rates are easier to handle and were thus chosen for this work.

and set to

$$\frac{k_{i,forw}}{k_{i,rev}} = \frac{5s^{-1}}{1s^{-1}} = 5, \ \forall i > 1$$

for the two reference models.[61] This ratio determines the absolute level of activation as well as the dynamic properties (time to reach steady state, temporal variance before a steady state is reached, etc.). As with molecule numbers, these parameters will vary to a large degree for different scenarios and thus the influence of these variations on signal transduction will be investigated later (see page 105).

As set up in section 5.2.1, the outputs of both models are the probability distributions of the activated last species in the signaling cascade. We can compare them qualitatively, as they are equivalent and differ only in the upstream transmission process. Yet, to quantify the information processing capability of our stochastic channel models, we employ an information theoretic analysis by defining the measure of "Fidelity F" as:

$$F(X;Y) = \frac{\max_{P(X)} I(X;Y)}{H(X)}, \tag{43}$$

giving us a measure with $F \in [0,1]$ that enables a comparison between the two models for how accurately the input is inferred at the output Y with respect to the number of inputs $X = k_{1,forw}$. This measure is merely a normalization of the channel capacity by one of its upper boundaries, as $I(X;Y) \leq \max(H(X), H(Y)$. It implies that $F = 1$ stands for a loss-less information transmission (and thus the potential for perfect inference of the input) and $F = 0$ denotes that barely noise is transmitted. Choosing $H(X)$ as a normalization factor is sensible for a low number of inputs and a high number of output states. This is a matter of the binning employed: Since for cell signaling, the resolution of the output (in that case defined by the molecule numbers) will usually be larger than the number of inputs, we refer to the input-entropy. Yet, as can be seen in Fig. 7, the channel itself sets the achievable resolution and thus, both selections are plausible. Using the "Fidelity" F is only viable for a discrete computational approach with finite

[61] $k_{1,forw}$ plays a special role as it encodes the input strength to the system (stress level X), as explained later in this section.

Name	Values	Description
$X = k_{1,forw}$	$[0.001, \ldots, 10]s^{-1}$	Input strength (stress level).
i	$i \in \mathbb{N}$	Index of signaling layer (tier).
m_i	1000	Molecule # per signaling layer i.*
$k_{\text{i,forw}}$	$5s^{-1},\ \forall i > 1$	"Forward" rate, activation of next layer.
$k_{i,rev}$	$1s^{-1},\ \forall i$	"Reverse", de-activation of active species.
t	$150s$	Simulation time used for steady state.

* (This number is the sum of activated and inactive proteins of one cascade layer.)

Table 3: Overview of the parameters used in the standard model in section 5.2.2. The parameters are subsequently varied for analysis as indicated in the sections. For the exact values of the input strength X, see Appendix B.

inputs, as $H(X) \to \infty$ with the number of input states X tending to infinity. Yet, it resembles the biological nature of finite number of molecules in a cell that act as an encoder[62] and thus can be used without restrictions for our application.

The fidelity measure conveys a simple way to interpret the performance of a system without putting emphasize on the meaning of the number of transmitted bits. After all, it is impossible to know what a particular number could mean to a cell without going into more details of the system in question[63] - this is a matter of semantics and not part of this theoretical analysis. Nevertheless, it proofs to be a simple and objective way to compare our parsimonious modeling approaches.

Signal amplification or graded response - The scaffold makes the difference

[62] Or, as a matter of fact, at a decoding level, should we choose to use $H(Y)$ for normalization.

[63] Nevertheless, it is interesting to know that interpretations for such bit-numbers are in fact sometimes possible: Dubuis *et al.* (2013) measure gradients of morphogens and the corresponding expression levels of gap genes in *Drosophila* embryos. With this they are able to show that the biologically encoded and information transmitted (over a "gene-expression channel") corresponds exactly to the information needed to encode each of the 100 unique positions for cells along the anteroposterior axis. This is also an extra-ordinary example for a biological system working at an optimized level of maximized information.

In order to get a first grasp on how signaling through the two motifs differs from one another, we compare the behavior of the reference models. Given distinct stress levels (i.e. the probability rate for the first activation step), the two system structures exhibit a strong influence on the input-output relation as shown in Fig. 22 (see the mapping of color coded arrows). As expected, the first species in the cascade (being the receptor level) naturally shows the same conditional distributions of activation. This can be interpreted as the encoding of the external stimuli into molecule numbers and since this encoding is conducted by the same mechanism for both models, the outputs on this level are identical by design. This changes once the different modes of transduction enter the equation. Multiple layers of successive signaling molecules produce a significantly different output, diverging with the number of tiers: While scaffolded signaling exhibits a near linear dose-response relation, the mixed relation shows a strong amplification effect that entails a switch-like saturation behavior.

Both these behaviors can be observed in natural settings and can be associated with different functions. A switch-like response enables thresholded signaling and would make sense in cases, where signals and/or cellular decisions are both rare and severe. This would enable a cell to commit to adaptations to an external stress if need be, while still reserving normal function when a threshold is not surpassed. Apoptosis would for example be such a fatal cellular decision, if committed to prematurely. Stresses that pose a strong threat to the cell need to be handled within a very fast manner and could potentially require a strong commitment and thus an amplified signal transmission as well. MAPK cascades (Seger and Krebs, 1995) have been under investigation for many years and are largely considered to act as such amplifiers (Goldbeter and Koshland, 1981; Ferrell and Machleder, 1998; Kholodenko, 2000). On the other hand, graded signaling allows for fine-tuning and distinct cellular behavior. As an example, while the decision to undergo mating is certainly of binary nature, the pathway output of the pheromone response shows a linear alignment to the α-factor concentration sensed in the environment (Poritz *et al.*, 2001; Yu *et al.*, 2008, Fig.2). Moreover, yeast has several distinct modes of cellular behavior (budding, elongated growth and shmooing) depending on

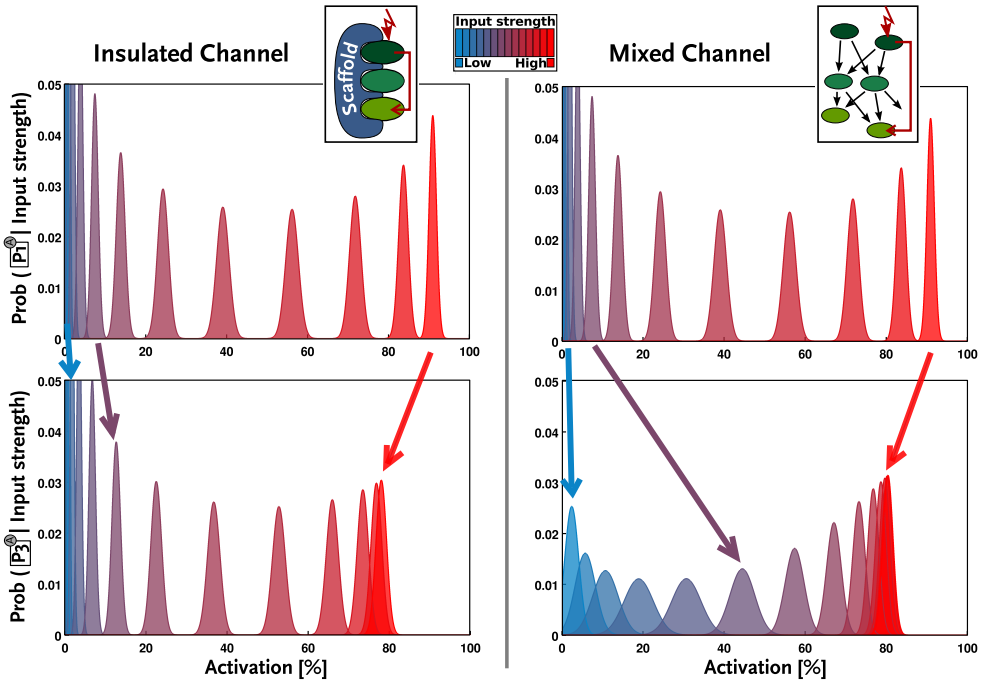


Fig. 22: Simulation results of the input-output relations for the two reference signaling models. We display the conditional probability for the distribution of molecules (normalized to the total protein abundance) with respect to the input strength (color-coded), i.e. distinct stochastic rates of the first activation step (see Tab. 3). The relation is shown for the first and the third species in the cascades. Whereas the first layer of signaling (the receptor level) in each model experiences identical distributions for matching stresses, the third layer already exhibits strong differences both in transmission as well as intrinsic noise level resulting from the structure.

the strength of α-factor stimulation (Dohlman and Thorner, 2001; Moore *et al.*, 2008; Hao *et al.*, 2008). Other studies have shown an integration of both graded and switch-like behaviors due to the use of several mechanisms in pheromone responses (Paliwal *et al.*, 2007) and the adaptation to high osmolarity (English *et al.*, 2015), respectively. An important way to facilitate both modes of responses can be to use feedbacks (Ninfa and Mayo, 2004; Bhalla and Iyengar, 1999; Legewie *et al.*, 2008), yet we observe that scaffolding can achieve the same behavior in a controlled way. A combination of many such mechanisms is most likely in a natural setting.

In our analysis, the amplification effect strongly influences fidelity. As a fast

saturation can be observed upon crossing a certain threshold, signals below and above that cannot be distinguished reliably. This produces channel capacity at lower numbers, slightly above an on/off switch. This behavior results from the structural connection between the successive signaling layers: The received signals spread in a snowballing fashion starting at the receptor level. The strong inter-connection between the signaling layers is the key mechanism behind this behavior of mixed signaling. In contrast to this amplification, graded responses allow for higher achievable channel capacities and thus a more precise inference, depending on the noise that is introduced in the system. This is the case for the scaffolded model. Its linear structure of signal transmission leads to a sizable enhancement of fidelity compared to the mixed channel. With restricting each activation transmission to one molecule per tier is insulating the process internally, achieving a high level of coordination and control.

Yet, fidelity is not defined by the input-output relation alone. When comparing the conditional output distributions, we can draw a second important conclusion:

Scaffolding limits the propagation of noise

With each layer of signaling molecules the propagation of noise is diverging for the two reference models, as can be observed in Fig. 22. Scaffolds seem to severely limit the amount of noise that is intrinsically introduced on each layer of the transmission process, whereas the mixed model shows (together with signal amplification) a large spread of the conditional output distributions. This observation depicts the potential of the scaffold structure to shield the transmission from outside perturbations, like it has been shown for example in Perlson *et al.* (2006). While this might be true for particular scaffold systems in nature, what we observe in our modeling approach is not a consequence of this active shielding due to conformational changes. A close look into the design of our model reveals that no external perturbations (except for the input) are modeled explicitly. This would certainly be straightforward to do, as it would only require to extend the backward reactions to model e.g. the influence of different phosphatase activities. Yet, what we actually observe is a passive

"intra-pathway-insulation" effect that is implemented naturally by the scaffold structure. The co-localization restricts the transmission of a signal in a one-to-one relation between signaling molecules and thus noise occurring in one scaffold will not spread to other molecules of the subsequent downstream tier. This is a very interesting feature, facilitated solely by the structure of the scaffold protein.

The mixed model, on the other hand, not only amplifies the signal as described previously, but also amplifies noise on every tier. The introduced inherent fluctuations of reactions have larger consequences in this model. As the connection between layers is stronger due to every signaling molecule on one tier potentially being able to activate every molecule of the next, a stochastic variation of this particular molecule will increase covariances in many downstream molecules and thus propagates its noise further than a scaffolded molecule would. The amplification effect of distributed signaling is the reason for lower fidelity. Yet, when the system aims at a switch-like behavior, this might not necessarily be a restriction. Although it also affects the sharpness of the threshold, one can assume that such a system might have adapted to a stress that typically only occurs in a correspondingly switch-like "low/high" fashion.

Comparison of Fidelity under fluctuating components

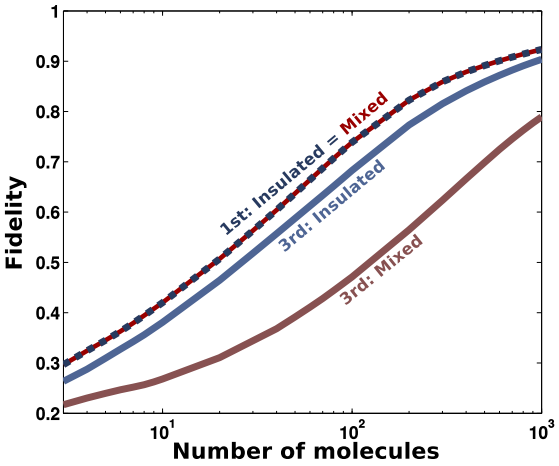
Up to this point, we merely regarded the reference models in their basic configuration. We were able to state important implications that scaffolding has on signal transmission. Yet, to rigorously compare the performance of the two, it is necessary to vary the models systematically and observe the changes in signal processing and the dependence of fidelity on our parameters. The two features described in the previous paragraphs act in unison and are integrated in the measure for fidelity. The results of the comparison under variation of molecule numbers and transmission strengths are visualized in Fig. 23a and 23b, respectively. As a surprising outcome, we find that our statements of signaling behavior hold for a wide range of variations. This is very important, as in natural settings, those parameters will fluctuate to a large degree and also occur in

many different configurations (e.g. de Godoy *et al.*, 2008; Thomson *et al.*, 2011).

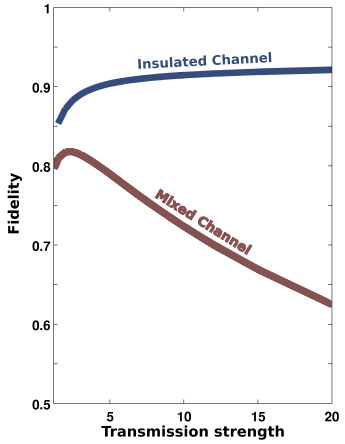
When molecule numbers are varied, the most important feature that changes linearly with it is the possible resolution at the output level: The more molecules we employ, the more input signals can potentially be resolved. Of course, again as with the noisy type writer (see example 2.2), this resolution is under a strong restriction by the noise induced in the channel, leading to a saturation of information capacity and fidelity. Once this saturation is reached, higher molecule numbers might be futile and lower the efficiency. In fact, for a mixed channel, fidelity might decrease again if molecule numbers are too high (not shown), since the amplification effect acts in a non-linear way and will saturate the response. Fig. 23a also exemplifies the natural application of the data (signal) processing inequality: While the receptor levels equal one another, the signaling layers further downstream can only decrease in fidelity. The important question, however, is to what degree this happens. We observe that a scaffolded channel always outperforms the mixed channel with respect to fidelity.

The same is true for the variation in signal transmission strength. Fig. 23b visualizes this dependency and shows a disadvantage for the mixed channel: Increasing the transmission strength will amplify the snowball effect of mixed signaling in a non-linear way and lead to both a faster saturation profile (Fig. 24) and a stronger noise propagation. As these parameters determine the speed of the signaling process and the absolute activation of the pathway (see also Fig. 23c), it implies that scaffolding can be tuned very efficiently to fulfill certain requirements without loosing its profile for dose-response. This is a feature that makes scaffold signaling a very interesting target for synthetic biological research.

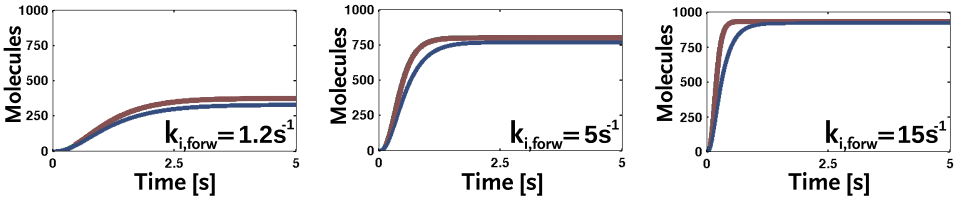
In a similar argumentation, the profiles in Fig. 24 show a very surprising feature for the comparison between the two models: It is possible to tune the mixed channel in a way that it behaves like the insulated channel. Yet doing so requires a very low signal transmission ratio. This low ratio in turn leads to a very low total activation of the pathway, making it inefficient, as well as a slower reaction time. This behavior is visualized in figure 23c. While



(a) Molecule numbers m (per signaling layer, $m = m_i$) are setting the resolution of signaling and severely influence the amplification effect. With increasing molecule numbers, higher resolutions are able to increase the fidelity up to a boundary set by the channel transmission probabilities. With exception of the first signaling layer, scaffolding fidelity (blue lines) always exceeds mixed fidelity (red lines) due to both amplification as well as noise propagation effects. This result also visualizes the data processing inequality (see section 2.1.1).



(b) Fidelities corresponding to the variation of transmission strength ($k_{i,forw}/k_{i,rev}$, with $k_{i,rev} = 1s^{-1}$). This figure complements Fig. 24 with the calculated output of transmission fidelity. Compared to the insulated channel (blue), higher rates severely limit the fidelity of the mixed channel (red) by increasing the non-linear amplification effect.



(c) Exemplary time courses for the mean activation of scaffolded (blue) and mixed (red) signaling depending on the transmission strength. Transmission rates are responsible for the dynamic response of the systems and also influence the saturation level of the output.

Fig. 23: Fidelity as a function of varied system parameters m and $k_{i,forw}$. The measure visualizes how scaffolding outperforms a distributed mixed signaling. Parameters follow the standard set from Tab. 3, unless varied as indicated in the figures.

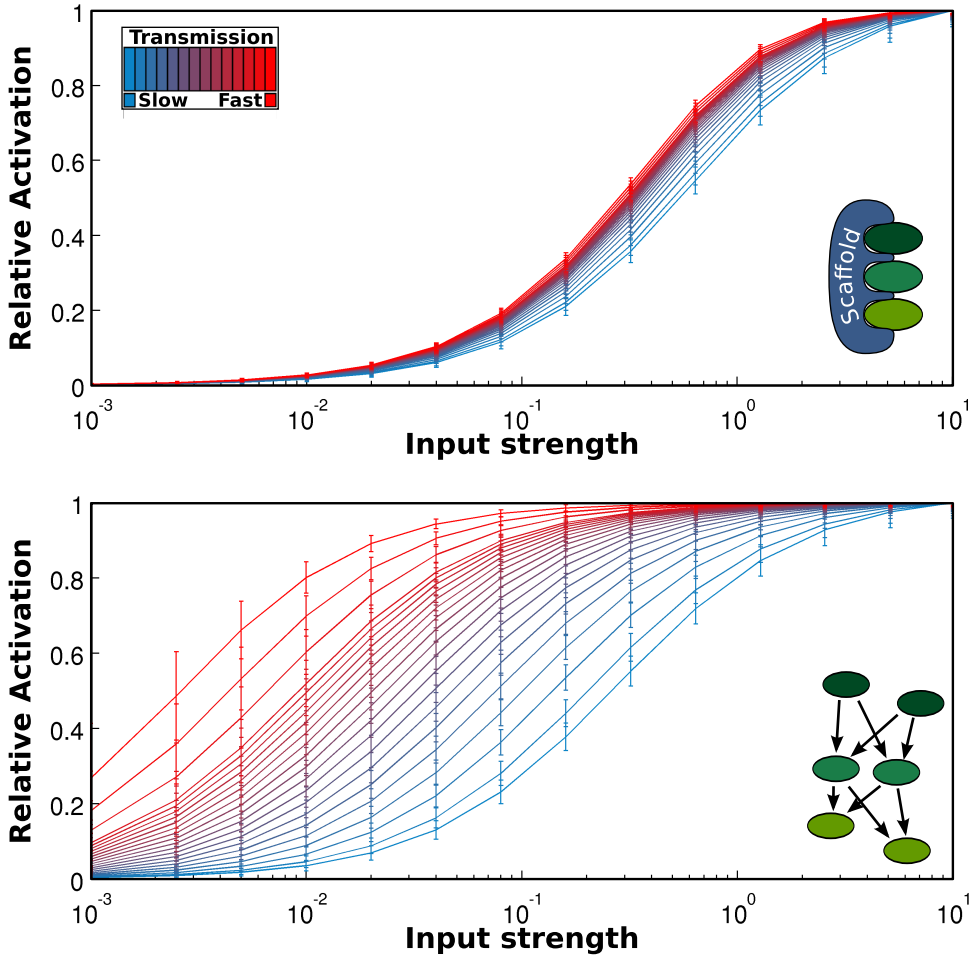


Fig. 24: Tuning the signal transmission strength (ratio of activation to deactivation) between the layers results in different profiles of relative activation for the two channel models. Especially interesting is the fact that the mixed channel can be tuned to show similar behavior, yet at the cost of total activation level. Transmission strength (color coded) has been varied from 1.2 (blue) to 20 (red).

scaffolded signaling is not prone to variation in this strength and thus shows a high robustness, mixed signaling needs to be tuned correctly to achieve the same behavior and while doing that looses in performance.

For now, we have only been looking at three consecutive species in the cascade. The question stands to reason what happens if the chain consists of more layers as is the case for several signaling pathways (especially, if double-phosphorylations are counted as two tiers). As can be expected from the analysis of three signaling layers already, such an extended chain would result in even more pronounced properties of the channels. First of all, the signal processing inequality (see also section 2.1.1) will by definition limit the information transmission even more severely, implying that information cannot be gained in a transmission chain. So adding more complexity inherently increases the noise in the system. At the same time, the amplification effect in the mixed channel is reinforced severely, leading to a fast saturation of the signal. These two effects are pronounced when increasing the length of the signaling chain.

Having reached the middle of our study, we already want to summarize the main finding of this chapter and subsequently substantiate it: Scaffolded signaling outperforms a distributed signal transmission with regard to signaling fidelity due to 1. a near-linear dose-response alignment as well as 2. the improved attenuation of noise propagation inherent to the system. A third point will extend the importance for evolutionary design and substantiate that scaffolding has indeed the potential to promote and sustain biological fitness.

5.2.3. Robustness of the Fidelity measure in model comparison

Using simple models of the basic processes, we were able to capture a desirable behavior for scaffolded signaling and showed the dependence of our statements on the parameters. In a next step we want to examine the effects that stochastic variations on the system parameters would entail. This corresponds to the unique expression of proteins and fluctuating transmission strength in each single cell, as in a natural setting they can vary to a large extend without breaking the systems function (see e.g. de Godoy *et al.*, 2008; Thomson

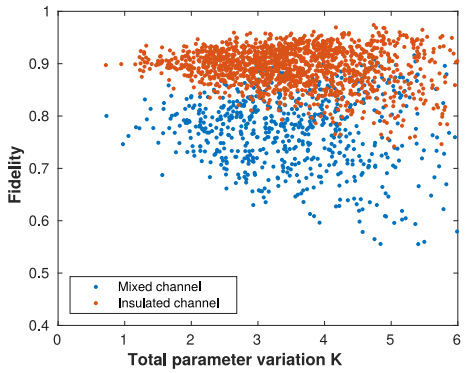
et al., 2011). For analyzing the effects of these natural variations, we follow the approach of randomizing parameter values as described in Barkai and Leibler (1997).[64] We create an ensemble of alternative models, each obtained by randomly varying the reference parameter set $\boldsymbol{k}_0$ (see section 5.2.2). The vector $\boldsymbol{k}_0$ consists of the values for molecule numbers and forward as well as backward reaction probability rates. Variation around the reference values is chosen to resemble a reasonable change in molecule numbers or reaction rates by multiplying them with a randomized factor of 2^x, with x drawn from a Gaussian random distribution $\mathcal{N}(\mu, \sigma)$. This results in log-normal distributed values in a range that roughly halves or doubles the population and strengths, with the weight distributed around the reference values. For the number of signaling molecules, we aim to create random parameter sets that show a coefficient of variation of slightly over 30% by using $\mathcal{N}(0, 0.5)$, as protein abundance variation of this magnitude is for example reported in Thomson *et al.* (2011). For comparing the two models we combine the variational changes in the measure of "total parameter variation" (adapted from Barkai and Leibler, 1997):

$$K = \sum_{i=1}^{size(\boldsymbol{k}_0)} |log_2(\frac{k_i}{k_{i,0}})| \tag{44}$$

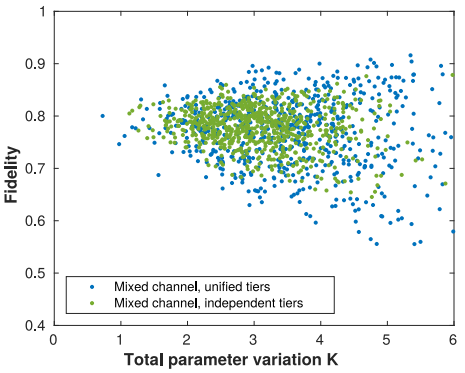
This approach allows us to evaluate the robustness of our models as perceived by the fidelity of the system. Fig.25 visualizes fidelity of varied model versions depending on how strong their parameters were varied.

With the analysis we are able to conclude that our previous observations hold also in the face of cell-to-cell variability in molecule numbers and kinetic transmission parameters. Over a wide range of K, the observed fidelity of insulated signal transmission exceeds the mixed fidelity, as can be seen in Fig. 25a. A visible overlap is only present at higher values of total parameter variation. By looking closer at the model variants that show this behavior, one realizes that this is due to larger differences in output resolutions between the variants (not shown here). Low output molecule numbers are unable

[64] The authors use this approach to evaluate robustness in bacterial chemotaxis. See also the use in Blüthgen and Herzel (2003) in an analysis of MAPK cascades.



(a) Comparison of scaffolded (orange) and mixed channel (blue) on the third signaling tier under the variation of all parameters (while $m_1 = m_2 = m_3$ for scaffolded signaling). The angle of the spanned cone measures the robustness of the model. Scaffolded signaling exceeds the mixed model in both fidelity as well as robustness for the majority of variations.



(b) Comparison between the variation of molecules in unison for all tiers (blue, in analogy to the way scaffold numbers would vary, see also (a)) and the variation on each single signaling tier (green) added to the varied parameters. Single tier variations mildly buffer the effects of variation in transmission strength and provide slightly more information capacity.

Fig. 25: Scatterplots comparing the robustness in fidelity to parameter variation (see Eq. (44)) between the two signaling models. Each dot represents the fidelity on the 3^{rd} signaling level for one model variant, altered in parametrization for m_i, $k_{i,forw}$ and $k_{i,rev}$, per signaling tier $i = 1, \ldots, 3$.

to achieve the same fidelity as the ones of higher magnitudes. The second important observation from the analysis is that the scaffolded signaling motif is more robust towards these variations, as the cone is spanning a lower angle.

One particular difference between the two models in this approach occurs on the level of molecule numbers: The mixed channel enables us to change molecule numbers on the level of each single tier. In nature, this can lead to bottlenecks in the transmission cascade and also determines the saturation of the downstream response. For example, with a high abundant intermediate species, the pathway can ensure a good diffusion coverage over the cell volume and thus a high activation of the pathway. The described amplification effect additionally makes it possible to activate a larger number of downstream molecules compared to a scaffolded setting. Effects like this have to be investigated systematically in separate studies. For the focus of this work we factor

them merely into the robustness analysis by allowing each tier of molecules to vary independently (yet with the same parameter variation model) of the other species. This can be seen in Fig. 25b. With a sufficient number of model variants[65], it can be seen that the variational cone is slightly less broad. When going into the details of the simulations, we observe that this is due to a buffering between molecule numbers and transmission rates: A lower transmission strength can be balanced by using a larger number of molecules and *vice versa*. This interesting effect allows for a tuning, yet the motif still lacks the regulatory possibilities of a scaffold.

Scaffolded signaling can only be realized with exactly the number of scaffolds that can be fully assembled and its saturation is thus restricted by the lowest abundant species. In addition, this leads to the well-known attenuation effect of signaling in the case of an over-expression of scaffold proteins due to combinatorial reasons. This effect has been studied from both experimental (Chapman and Asthagiri, 2009) and theoretical perspectives (Heinrich *et al.*, 2002; Levchenko *et al.*, 2000). Since we omit the assembly process in our modeling, the assumed number of scaffold proteins is set to the corresponding number of functionally assembled scaffolds, namely the lowest abundant species of the cascade. As discussed, this constrains our analysis to only consider the transmission process and the use of one single parameter for the number of channels. The mixed model is varied under the same constrains to enable a fair comparison. The results are visualized in Fig. 25a.

We can conclude that the main aspects of improving signal transmission with the use of scaffolds do not rely on a fine-tuned system, but rather are implemented in the structure of the motif itself. They hold for a large set of possible configurations in the parameter space and have shown to be more resilient to cell-to-cell variation. Our simple approach of parsimonious modeling shows that this robust structural behavior represents a strong candidate for selection in the early stage of evolutionary development. Such a simple, yet fundamental concept could provide a solid basis for an organism to subsequently build up complexity and add higher functionalities for a successful advancement

[65]For better visualization the figure only shows a subset of simulations.

of the species.

5.3. Information gain through teamwork - Channels working together

As we have seen in this chapter, introducing a scaffold protein into a signaling structure can have beneficial effects on the information transmission by shaping the input-output relation and the propagation of inherent noise. Yet seeing single assembled scaffolds as a separated instance for error-prone transmission opens up a related view of signal processing. In electronic systems, error correcting codes are employed in order to ensure that a message is transduced and decoded correctly. This can include adding redundant information, e.g. by repeatedly sending a message. To a certain extend this can be observed in this biological structure as well. In this chapter we already have established that scaffolded intra-pathway-insulation leads to enhanced signal fidelity and the temporal activation of pathways can easily be interpreted as a repetition of the message. Yet, instead of interpreting the activation distribution of the last species in the cascade, we can see this output species as an m-dimensional codeword. Interpreting each scaffold protein as one transmitting channel allows us in addition to regard the whole ensemble of channels in the cell as a multivariate aggregate. In a short detour to this project, we want to examine how this would feed into the correction of even severely distorted signal transmission.

Referring to our model of the insulated channel, we reinterpret this motif as a simple on-off-transmission in analogy to the Binary Symmetric Channel (BSC, see example 2.1). We established in the example, that the capacity of this channel depends directly on the transmission error-rate f. Now, going a small step further we can skew this transmission towards a different behavior in stressed and non-stressed situations by introducing an additional error-rate e for the transmission in one of the states, thus making it more adaptable to different scenarios. Using 0 and 1 (or “off” and “on”, respectively) as input states makes sense for example in situations where a conformation change keeps a protein in two distinct states, with one being active and the other not. We presented a case like this in section 3, where Sln1 played the role of an input by changing its conformation state upon stress. One such a Binary

Non-Symmetric Channel would have the transition matrix

$$\mathbf{Q}_1(\text{Output}|\text{Input}) = \begin{pmatrix} P(\text{off}|\text{off}) & P(\text{off}|\text{on}) \\ P(\text{on}|\text{off}) & P(\text{on}|\text{on}) \end{pmatrix} = \begin{pmatrix} 1-f & e \\ f & 1-e \end{pmatrix}. \quad (45)$$

Setting the error rates to $f = 0.05$ and $e = 0.2$ can then be interpreted as a channel firing occasionally even in an environment with low stress conditions (meaning a 0 is given as the input), yet under stress input the channel will be activated with a high probability. Such a conditional probability distribution would only give a channel capacity of $C = 0,48$ *bit* (for example computed with the Arimoto-Blahut algorithm using the transition matrix Q, see section 2.1.2). Assuming that the noise in transmission cannot be reduced, there still exist ways of improving the inference capabilities at the output level. We want to investigate the example of an ensemble of redundant channels as a way of doing that. Consider the previous setup. If instead of one binary channel we now employ two identical ones, observing them as a combined channel with four possible input states (the number of possible combinations of *on* and *off* equals 2^n) and aiming to infer a binary output. The transmitting channel has now the following structure:

$$\mathbf{Q}_2(\text{Out}|\text{In}) = \begin{pmatrix} P(\{\text{off},\text{off}\}|\text{off}) & P(\{\text{off},\text{off}\}|\text{on}) \\ P(\{\text{on},\text{off}\}|\text{off}) & P(\{\text{on},\text{off}\}|\text{on}) \\ P(\{\text{off},\text{on}\}|\text{off}) & P(\{\text{off},\text{on}\}|\text{on}) \\ P(\{\text{on},\text{on}\}|\text{off}) & P(\{\text{on},\text{on}\}|\text{on}) \end{pmatrix} = \begin{pmatrix} (1-f)^2 & e^2 \\ (1-f)f & e(1-e) \\ f(1-f) & (1-e)e \\ f^2 & (1-e)^2 \end{pmatrix}, \quad (46)$$

giving us an achievable capacity of $C = 0.71$ *bit* and thus a sizable increase in information. Adding a third in a similar manner, increases this number to $C = 0.83$ *bit*. The obvious next step is to enlarge the set of binary channels for improvement of inference. The channel will transmit a codeword of length n to a binary output, which represents a repetition code employing redundancy. The same scenario can be viewed as the reversed setting, corresponding more natural to a cell's inference mechanism: For this, we define the input as binary and the transmission an ensemble of binary channels, thus yielding a codeword of length n as the output which then can be used for our inference. The approach used to prove the Arimoto-Blahut algorithm (see subsection 2.1.2)

and the definition of mutual information tells us that this reversed channel will have the same capacity. Adding more channel instances increases the channel capacity until a saturation point (see Fig. 26). In fact, this behavior of redundant messages could be seen as a natural analogy of Shannon's well-known "Source Coding Theorem" from information theory (Shannon, 1948; Cover and Thomas, 2012). For our channel construction, this boundary is set by the maximum of input or output entropy, namely $C \leq H(Y) = 1\ bit$ of information transmission. We observe that, provided a sufficient amount of redundancy, even a poor single transmission channel can be used to transmit information. It would be interesting to experimentally study and measure this conditional transmission probabilities and compare the theoretical number of such channels needed to reach capacity with the observed abundance of output molecules within a cell. Obviously, extensions to our model would be needed for such a comparison.

In summary, the consequence of our observations is that by employing a sufficient amount of receptors and a parallel transduction of information, cells can recover poor performances of single signaling channels and thus are enabled to infer an external cue with arbitrary precision. This is an interesting example for a cellular error correcting code. Of course, it also is merely a theoretical consideration that omits parts of the embedding setting. After all, even if for a certain number this inference could in principle be perfect, a natural setting will always include a spatial component in addition. This, together with the here presented toy model, provides us with an idea of what can be interpreted by a cell: Codewords of certain lengths carrying a large amount of redundancy can simply be an analogy to molecular concentrations that encode a (more or less complex) information state of the cell depending on its environment. Increasing the length of codewords enhances the resolution and as with scaffolded channels, it can enable higher fidelities depending on our input states and the information processing. Diffusion within the cell or other means of transportation are then additional channels with other transmission properties that add upon this resolution. Taking them into account would be an interesting, yet arguably very complex extension to our setup.

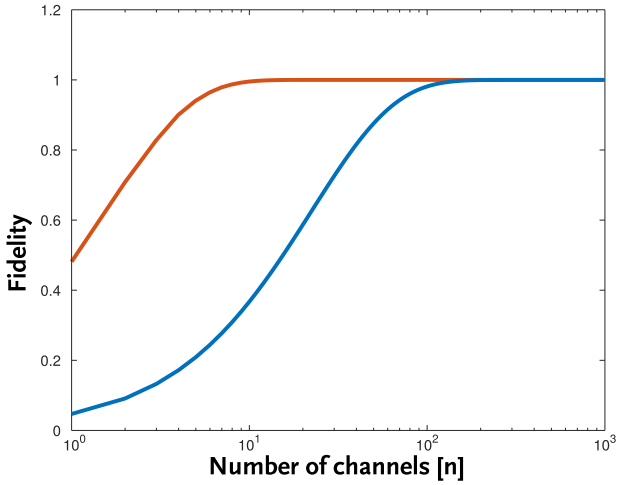


Fig. 26: Performance of combining multiple channels to a set producing codewords of length n. Error rates of the transition matrix $\mathbf{Q}_n$ are set to $f = 0.2$, $e = 0.05$ (red) and $f = 0.7$, $e = 0.1$ (blue), where f is the probability for a single channel to transmit the false state given an "off" input and e the respective false transmission given the "on" input. The maximum channel capacity of 1 bit can be achieved when enough channels are present for reliable inference. This can be thought of as error-correcting code through adding redundancy to a message. After transmission, the concentration of active molecules of the output species enables a perfect inference on the input signal, even in the case of a flawed (yet, not random) signal transmission process.

5.4. Summary & Discussion

Fidelity in the presence of noise in cellular systems is no miracle but a consequence of design. Even with the many sources for variations within the cell as well as the outside fluctuations of the natural environment, cells have managed to evolve ways to circumvent and cope with this ubiquitous variability. In our investigation, we focused on one of these mechanisms, namely the structural motif of scaffolding proteins. This motif has been evolutionary conserved in many different organisms and in addition fulfills just as many functional roles. Yet, since many scaffolds show low similarity, we argue that this conservation is the consequence of fundamental structural properties instead of a high specialization to certain scenarios. Our analysis examined benefits that a prototypical scaffold can exhibit on its own in a single pathway and allow the cell to react more reliable to its environmental settings. We found that scaffolding shapes the relation of dose to response and strongly improves information transmission fidelity.

In this chapter, we presented two models (see section 5.2.1) that were aimed at assessing and comparing particular features of signal transmission. One model employed scaffolding of signaling species, the other a mixed diffusive signaling. Both are identical on the receptor-level and were made comparable on the output by adjusting for spatial considerations. We found that the scaffold structure shows advantages in its signal processing capabilities in comparison to the reference model: Both a near-linear dose-response alignment as well as a severely limited propagation of noise due to intra-pathway insulation allows the scaffolded channel to outperform the mixed channel with respect to fidelity by a large margin. In addition, the scaffolded model was shown to be more robust to parameter variations, which can be considered one of the most important features for evolutionary design. As is inherent in evolutionary arguments, we are assuming that cells are optimizing features and are selected for beneficial capabilities in evolution. This prerequisite for our hypothesis is likely, yet impossible to prove. Nevertheless, scaffolding motifs show intriguing features that, in particular due to their basic nature, could explain the abundant and diverse use in nature. An additional important aspect for this is also that the observed features in no way contradict other functions scaffolds have been shown to fulfill.

As a last consideration, we studied a similar toy model to show a principle frequently used in computer science applications (see section 5.3). We showed that using redundant codes enhances a cell's inference behavior and ensures that the interpretation of extracellular cues can be sufficiently correct. This is an analogy that suggests further ways of error correction and presents us with interesting ideas for further studies. For a given channel, we can calculate the redundancy needed in order to infer a binary signal. An experimental setup measuring such transmission properties and signaling molecule abundances would allow for a comparison. Yet, for this to be meaningful, the modeling approach will have to be refined. For one, it does not take spatial organization of signaling and the biological meaning of molecular concentrations into account. Another caveat is that both input and output are hard to determine in a natural setting. Considering a synthetic approach and combining it with our theoretical consideration will be a very interesting and challenging task.

6. Discussion & Outlook

6.1. Summary of the work

In this work, we introduced the application of information theory to cellular signaling systems and built upon stochastical frameworks to enable the usage of this theory. Our main question motivating this exercise was to observe how cellular organisms can systematically cope with inherent fluctuations as well as noisy environments and how they could be designed to implement this coping ability already on a structural level. To test the theoretical considerations, we applied our modeling approaches to cellular signaling in *S.cerevisiae*:

First, we zoomed in on the Sln1-Ypd1-Ssk1 phosphorelay, a vital input module to the HOG pathway reacting on osmotic stress in the environment. We were able to show properties of signal transduction capacity for different setups of the system and employed signaling constraints to limit the set of possible parameter configurations. For this, we used stochastic simulation with the Gillespie algorithm to sample dynamic probability distributions. Comparing our results with an experimental setup predicted the role of association and dissociation constants for increasing fidelity as well as enabling a feasible signaling behavior.

In a second study, we reevaluated a phospho-proteomic dataset of temporal profiles for phosphosites (Vaga *et al.*, 2014) in the network for pheromone and HOG signaling. We built a deterministic model to combine and harmonize insights from a broad range of past research to a comprehensive view. Unfortunately, we were not able to bring the model behavior in accordance with and comprehensively explain the regulation patterns exhibited by the detected phosphosites. This was due to the small overlap between established mechanisms and the reported sites. Nevertheless, we were able to show that our model itself produces the desired output and especially emphasizes the importance of scaffolds in the inter-pathway insulation and the resulting signal specificity.

Although from a different perspective as in the previous exercise, this scaffolding motif was subject to another investigation. We connected stochastic modeling with information theory to show that scaffolds follow an optimal

design that exceeds other signaling motifs with respect to signaling fidelity. Three important points led to that conclusion: 1. We observed that due to the scaffolded structure facilitating an "intra-pathway" insulation for each protein, a near-linear dose-response alignment was enabled and 2. additionally, the motif displays a low amplification of the noise induced inherently in the signal transmission process. 3. We were able to show that scaffolded signaling is more prone to fluctuations in molecule numbers as well as rate conditions. These points allow us to propose that scaffolding could provide strong incentives for the selection in evolution and explain its abundance and myriad of further secondary functions.

6.2. Outlook

The three modeling approaches that we included in this work are all connected and take on a systems level share in important signaling mechanisms of *S.cerevisiae*: The sensing of pheromones as well as the osmotic stress response. An interesting way to continue this would be a more comprehensive integration of the approaches. On one hand, introducing noise into a larger system will be challenging, but it certainly has the potential elucidate many open questions. For example, our results on scaffold fidelity could be evaluated in a larger context and also be directly compared to mixed signaling in the same system. Nevertheless, it is advisable to only slowly build such a integrated view, as we have seen that many things are still unknown as of yet and the problems of parameter estimation and applicability of bigger system still can lead to non-conclusive results.

For the approach of extending this work further is of experimental nature. The collaborative work with the lab of Luis Serrano is still ongoing and has only been discussed briefly in the work at hand. Not all extension to our modeling approach were shown and evaluated. The phosphorelay system comprises for example further complex building between Ssk1 and Ypd1 than has been modeled in our approach of 3. Adding further complexity shows an interesting feature in the control of basal signaling of the pathway: The basal "leakage" of the phosphorelay is correlated with the strength of association between bot Ypd1 and Sln1 as well as Ypd and Ssk1. This correlation can be seen *in vitro*

at the output level of Hog1. As the setup of our phosphorelay model (see section3.2) was using artificial, yet sensible chosen parameters, we adjusted the model in a way that better represents the biological data and continued to study and capture the observed behavior *in silico*. The results will be published in the near future, yet are still preliminary and currently under preparation.

Another interesting way of extending considerations of crosstalk and bringing it together with the idea of multiple input multiple output (MIMO) design from communications engineering would be to interpret the whole network used in chapter 4 as such a channel[66] and measure experimentally the multi-input-output relations. This would be doable with the experimental setup employed by Vaga *et al.* (2014), but unfortunately the statistics provided by the approach are not sufficient to approximate the required probability distributions. A setup using single cell measurements of the combined output (e.g. two fluorescent reporters induced by the HOG and pheromone pathways as employed in Baltanás *et al.* (2013) or Patterson *et al.* (2010b)) would be the right scope for this kind of analysis. It would be interesting to combine this with our considerations of section 5.3, where we consider multiple instances of a pathway as a cellular error-correcting code, and see if achieving a functional capacity and thus a reliable signal inference is possible as would be expected.

Within this thesis, we reviewed how information theoretic frameworks have been applied to biological settings in many interesting ways and presented the usage on modeling approaches. Yet, we believe that the vast repertoire that has been developed build on Shannon's theories is by far not exhausted as of yet. In computer sciences, we saw the emergence of many sophisticated systems that without the theory would not have been possible. And even though the application to biological systems is even more complex and difficult due to them not being "designed" by us[67], it has shown to be of use for the understanding of further principles. This is especially true for signaling in cellular sensing. The analogy of interpreting these systems as channels, as presented in this work,

[66] Possibly, this could mean two binary channels with cross-over potential, resulting in an input of four values and the respective output.

[67] After all, with all our vast knowledge we still understand very little of biological complexity and the large networks of interplay in biological systems.

can be applied both experimentally as well as theoretically and presents us with an alternative way of viewing information transmission in biology. The most important aspect of the application of information theory is that its measures can be employed for many contexts without knowing the particular semantics. We obtain a measure that is objectively representing correlation and potential resolutions for information processing of such systems. This always will have to be considered critically, as semantics are unknown to an outside view, but it presents us with boundaries on what is achievable and could be optimal. Many more sophisticated concepts of information theory have not been applied in a biological research as of yet and current research only makes infrequent use of the theory, but on both experimental as well as theoretical fields are increasingly aware of its potential. Knowing the framework and being able to interpret and judge its meaning, which also includes being able to recognize both its benefits as well as boundaries, is important for the development in the field. In this work we provided a basis for such an understanding and hope that this will contribute to further studies. As biology seems to only have started grasping the full potential of information theory, we believe that many more scientific findings based on the framework can be expected in the near future.

Systems Biology has come a long way since its (still recent) introduction. It has matured to facilitate more and more successful integrative approaches and also develop the connection between different fields of science. Linking research and building a more comprehensive view on biological systems is an important key for the future advancement of biological sciences. With the increasing awareness of the importance for single cell behavior as well as the noisy nature of, well, nature, Systems Biology has drawn the attention of even another field, information theory, and included it in its merging process that creates interesting science.

Information is the fundamental property that is needed to create, sustain and evolve life. It is ubiquitous and thus hard to grasp as a concept. Nevertheless, its value can be observed and measured. Biological systems are the living proof of that: they integrate this value of information and with it create, sustain and evolve themselves.

References

Adami, C. (2004). Information theory in molecular biology. *Physics of Life Reviews*, **1**(1), 3–22.

Adami, C. (2012). The use of information theory in evolutionary biology. *Annals of the New York Academy of Sciences*, **1256**(1), 49–65.

Alon, U. (2006). *An introduction to systems biology: design principles of biological circuits*. CRC press.

Alon, U. (2007). Network motifs: theory and experimental approaches. *Nature Reviews Genetics*, **8**(6), 450–461.

Arimoto, S. (1972). An algorithm for computing the capacity of arbitrary discrete memoryless channels. *Information Theory, IEEE Transactions on*, **18**(1), 14–20.

Azeloglu, E. U. and Iyengar, R. (2015). Signaling Networks: Information Flow, Computation, and Decision Making. *Cold Spring Harbor perspectives in biology*, **7**(4), a005934.

Bailey, N. T. (1990). *The elements of stochastic processes with applications to the natural sciences*, volume 25. John Wiley & Sons.

Balaban, N. Q., Merrin, J., Chait, R., Kowalik, L., and Leibler, S. (2004). Bacterial persistence as a phenotypic switch. *Science (New York, N.Y.)*, **305**(5690), 1622–5.

Baltanás, R., Bush, A., Couto, A., Durrieu, L., Hohmann, S., and Colman-Lerner, A. (2013). Pheromone-induced morphogenesis improves osmoadaptation capacity by activating the HOG MAPK pathway. *Science signaling*, **6**(272), ra26.

Bardwell, L. (2005). A walk-through of the yeast mating pheromone response pathway. *Peptides*, **26**(2), 339–350.

Barkai, N. and Leibler, S. (1997). Robustness in simple biochemical networks. *Nature*, **387**(6636), 913–917.

Bartel, D. P. (2004). MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell*, **116**(2), 281–297.

Bartel, D. P. (2009). MicroRNAs: target recognition and regulatory functions. *Cell*, **136**(2), 215–233.

Batchelor, E., Loewer, A., Mock, C., and Lahav, G. (2011). Stimulus-dependent dynamics of p53 in single cells. *Molecular systems biology*, **7**(488), 488.

Battail, G. (2005). Should genetics get an information-theoretic education? *IEEE Engineering in Medicine and Biology Magazine*, **25**(1), 34–45.

Berg, J. M., Tymoczko, J. L., and Stryer, L. (2002). Biochemistry. *New York*.

Bhalla, U. S. and Iyengar, R. (1999). Emergent properties of networks of biological signaling pathways. *Science*, **283**(5400), 381–387.

Bhattacharyya, R. P., Reményi, A., Yeh, B. J., and Lim, W. A. (2006a). Domains, motifs, and scaffolds: the role of modular interactions in the evolution and wiring of cell signaling circuits. *Annu. Rev. Biochem.*, **75**, 655–680.

Bhattacharyya, R. P., Reményi, A., Good, M. C., Bashor, C. J., Falick, A. M., and Lim, W. A. (2006b). The Ste5 scaffold allosterically modulates signaling output of the yeast mating pathway. *Science*, **311**(5762), 822–826.

Bianco, R., Melisi, D., Ciardiello, F., and Tortora, G. (2006). Key cancer cell signal transduction pathways as therapeutic targets. *European journal of cancer*, **42**(3), 290–294.

Blahut, R. (1972). Computation of channel capacity and rate-distortion functions. *Information Theory, IEEE Transactions on*, **18**(4), 460–473.

Blüthgen, N. and Herzel, H. (2003). How robust are switches in intracellular signaling cascades? *Journal of theoretical biology*, **225**(3), 293–300.

Borst, a. and Theunissen, F. E. (1999). Information theory and neural coding. *Nature neuroscience*, **2**(11), 947–57.

Botstein, D. and Fink, G. R. (2011). Yeast: an experimental organism for 21st Century biology. *Genetics*, **189**(3), 695–704.

Bowsher, C. G. and Swain, P. S. (2012). Identifying sources of variation and the flow of information in biochemical networks. *Proceedings of the National Academy of Sciences of the United States of America*, **109**(20), E1320–8.

Breeden, L. L. (1997). alpha-Factor synchronization of budding yeast . In W. G. Dunphy, editor, *Cell Cycle Control*, volume 283 of *Methods in Enzymology*, pages 332 – 342. Academic Press.

Brillouin, L. (2013). *Science and information theory*. Courier Corporation.

Butte, A. J. and Kohane, I. S. (2000). Mutual information relevance networks: functional genomic clustering using pairwise entropy measurements. *Pacific Symposium on Biocomputing. Pacific Symposium on Biocomputing*, pages 418–429.

Cai, L., Dalal, C. K., and Elowitz, M. B. (2008). Frequency-modulated nuclear localization bursts coordinate gene regulation. *Nature*, **455**(7212), 485–490.

Cao, Y., Gillespie, D. T., and Petzold, L. R. (2006). Efficient step size selection for the tau-leaping simulation method. *The Journal of chemical physics*, **124**(4), 044109.

Cao, Y., Gillespie, D. T., and Petzold, L. R. (2007). Adaptive explicit-implicit tau-leaping method with automatic tau selection. *The Journal of chemical physics*, **126**(22), 224101.

Chapman, S. A. and Asthagiri, A. R. (2009). Quantitative effect of scaffold abundance on signal propagation. *Molecular systems biology*, **5**(1).

Chen, I. A. and Nowak, M. A. (2012). From prelife to life: how chemical kinetics become evolutionary dynamics. *Accounts of chemical research*, **45**(12), 2088–2096.

Cheong, R., Rhee, A., Wang, C. J., Nemenman, I., and Levchenko, A. (2011a). Information Transduction Capacity of Noisy Biochemical Signaling Networks. *Science*, **354**.

Cheong, R., Rhee, A., Wang, C. J., Nemenman, I., and Levchenko, A. (2011b). Information transduction capacity of noisy biochemical signaling networks. *Science (New York, N.Y.)*, **334**(6054), 354–8.

Chol, K.-Y., Satterberg, B., Lyons, D. M., and Elion, E. A. (1994). Ste5 tethers multiple protein kinases in the MAP kinase cascade required for mating in S. cerevisiae. *Cell*, **78**(3), 499–512.

Cover, T. M. and Thomas, J. A. (2012). *Elements of information theory*. Wiley-interscience.

Daigle, B. J., Roh, M. K., Petzold, L. R., and Niemi, J. (2012). Accelerated maximum likelihood parameter estimation for stochastic biochemical systems. *BMC bioinformatics*, **13**(1), 68.

de Godoy, L. M., Olsen, J. V., Cox, J., Nielsen, M. L., Hubner, N. C., Fröhlich, F., Walther, T. C., and Mann, M. (2008). Comprehensive mass-spectrometry-based proteome quantification of haploid versus diploid yeast. *Nature*, **455**(7217), 1251–1254.

Dhanasekaran, D., Kashef, K., Lee, C., Xu, H., and Reddy, E. (2007). Scaffold proteins of MAP-kinase modules. *Oncogene*, **26**(22), 3185–3202.

Dickens, M., Rogers, J. S., Cavanagh, J., Raitano, A., Xia, Z., Halpern, J. R., Greenberg, M. E., Sawyers, C. L., and Davis, R. J. (1997). A cytoplasmic inhibitor of the JNK signal transduction pathway. *Science*, **277**(5326), 693–696.

Dimitrov, A. G., Lazar, A. a., and Victor, J. D. (2011). Information theory in neuroscience. *Journal of computational neuroscience*, **30**(1), 1–5.

Dohlman, H. G. and Thorner, J. (2001). Regulation of G protein-initiated signal transduction in yeast: paradigms and principles. *Annual review of biochemistry*, **70**(1), 703–754.

Dubuis, J. O., Tkacik, G., Wieschaus, E. F., Gregor, T., and Bialek, W. (2013). Positional information, in bits. *Proceedings of the National Academy of Sciences of the United States of America*, **110**(41), 16301–8.

Einstein, A. (1905). Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen. *Annalen der Physik*, **322**(8), 549–560.

Elf, J. and Ehrenberg, M. (2003). Fast evaluation of fluctuations in biochemical networks with the linear noise approximation. *Genome research*, **13**(11), 2475–2484.

Elion, E. A. (1998). Routing MAP kinase cascades. *Science*, **281**(5383), 1625–1626.

Elowitz, M. B., Levine, A. J., Siggia, E. D., and Swain, P. S. (2002). Stochastic gene expression in a single cell. *Science (New York, N.Y.)*, **297**(5584), 1183–6.

English, J. G., Shellhammer, J. P., Malahe, M., McCarter, P. C., Elston, T. C., and Dohlman, H. G. (2015). MAPK feedback encodes a switch and timer for tunable stress adaptation in yeast. *Science signaling*, **8**(359), ra5–ra5.

Faller, D., Klingmüller, U., and Timmer, J. (2003). Simulation methods for optimal experimental design in systems biology. *Simulation*, **79**(12), 717–725.

Fantes, P. and Brooks, R. (1993). *The cell cycle: a practical approach*. IRL Press at Oxford University Press.

Fassler, J. S. and West, A. H. (2010). Genetic and biochemical analysis of the SLN1 pathway in Saccharomyces cerevisiae. *Methods in enzymology*, **471**, 291–317.

Ferrell, J. E. and Machleder, E. M. (1998). The biochemical basis of an all-or-none cell fate switch in Xenopus oocytes. *Science*, **280**(5365), 895–898.

Finn, N. A. and Searles, C. D. (2013). Using information theory to assess the communicative capacity of circulating microRNA. *Biochemical and biophysical research communications*, **440**(1), 1–7.

Fujioka, A., Terai, K., Itoh, R. E., Aoki, K., Nakamura, T., Kuroda, S., Nishida, E., and Matsuda, M. (2006). Dynamics of the Ras/ERK MAPK cascade as monitored by fluorescent probes. *Journal of biological chemistry*, **281**(13), 8917–8926.

Gardiner, C. (2009). Stochastic Methods: A Handbook for the Natural and Social Sciences Springer Series in Synergetics. *Springer, Berlin, Germany*.

Garrington, T. P. and Johnson, G. L. (1999). Organization and regulation of mitogen-activated protein kinase signaling pathways. *Current opinion in cell biology*, **11**(2), 211–218.

Gillespie, C. S. (2009). Moment-closure approximations for mass-action models. *IET systems biology*, **3**(1), 52–58.

Gillespie, D. T. (1977). Exact stochastic simulation of coupled chemical reactions. *The Journal of Physical Chemistry*, **81**(25), 2340–2361.

Gillespie, D. T. (1991). *Markov processes: an introduction for physical scientists*. Elsevier.

Gillespie, D. T. (1992). A rigorous derivation of the chemical master equation. *Physica A: Statistical Mechanics and its Applications*, **188**(1), 404–425.

Gillespie, D. T. (2000). The chemical Langevin equation. *The Journal of Chemical Physics*, **113**(1), 297.

Gillespie, D. T. (2001). Approximate accelerated stochastic simulation of chemically reacting systems. *The Journal of Chemical Physics*, **115**(4), 1716–1733.

Gillespie, D. T. (2007). Stochastic simulation of chemical kinetics. *Annual review of physical chemistry*, **58**, 35–55.

Goldbeter, A. and Koshland, D. E. (1981). An amplified sensitivity arising from covalent modification in biological systems. *Proceedings of the National Academy of Sciences*, **78**(11), 6840–6844.

Good, M., Tang, G., Singleton, J., Reményi, A., and Lim, W. A. (2009). The Ste5 scaffold directs mating signaling by catalytically unlocking the Fus3 MAP kinase for activation. *Cell*, **136**(6), 1085–1097.

Good, M. C., Zalatan, J. G., and Lim, W. a. (2011). Scaffold proteins: hubs for controlling the flow of cellular information. *Science (New York, N.Y.)*, **332**(6030), 680–6.

Gruhler, A., Olsen, J. V., Mohammed, S., Mortensen, P., Færgeman, N. J., Mann, M., and Jensen, O. N. (2005). Quantitative phosphoproteomics applied to the yeast pheromone signaling pathway. *Molecular & Cellular Proteomics*, **4**(3), 310–327.

Hanahan, D. and Weinberg, R. A. (2011). Hallmarks of cancer: the next generation. *Cell*, **144**(5), 646–674.

Hao, N. and O'Shea, E. K. (2012). Signal-dependent dynamics of transcription factor translocation controls gene expression. *Nature structural & molecular biology*, **19**(1), 31–9.

Hao, N., Nayak, S., Behar, M., Shanks, R. H., Nagiec, M. J., Errede, B., Hasty, J., Elston, T. C., and Dohlman, H. G. (2008). Regulation of cell signaling dynamics by the protein kinase-scaffold Ste5. *Molecular cell*, **30**(5), 649–656.

Hartley, R. V. (1928). Transmission of information. *Bell System technical journal*, **7**(3), 535–563.

Hartwell, L. H., Hopfield, J. J., Leibler, S., and Murray, A. W. (1999). From molecular to modular cell biology. *Nature*, **402**(6761 Suppl), C47–C52.

Hayot, F. and Jayaprakash, C. (2004). The linear noise approximation for molecular fluctuations within cells. *Physical biology*, **1**(4), 205.

Heinicke, S., Livstone, M. S., Lu, C., Oughtred, R., Kang, F., Angiuoli, S. V., White, O., Botstein, D., and Dolinski, K. (2007). The Princeton Protein Orthology Database (P-POD): a comparative genomics analysis tool for biologists. *PLoS One*, **2**(8), e766.

Heinrich, R., Neel, B. G., and Rapoport, T. A. (2002). Mathematical models of protein kinase signal transduction. *Molecular cell*, **9**(5), 957–970.

Hersen, P., McClean, M. N., Mahadevan, L., and Ramanathan, S. (2008). Signal processing by the HOG MAP kinase pathway. *Proceedings of the National Academy of Sciences of the United States of America*, **105**(20), 7165–70.

Herskowitz, I. (1988). Life cycle of the budding yeast Saccharomyces cerevisiae. *Microbiological Reviews*, **52**(4), 536–553.

Hohmann, S. (2002). Osmotic Stress Signaling and Osmoadaptation in Yeasts. *Microbiology and Molecular Biology Reviews*, **66**(2), 300–372.

Hohmann, S. (2009). Control of high osmolarity signalling in the yeast Saccharomyces cerevisiae. *FEBS letters*, **583**(24), 4025–9.

Huang, C.-Y. and Ferrell, J. E. (1996). Ultrasensitivity in the mitogen-activated protein kinase cascade. *Proceedings of the National Academy of Sciences*, **93**(19), 10078–10083.

Jahnke, T. and Huisinga, W. (2007). Solving the chemical master equation for monomolecular reaction systems analytically. *Journal of mathematical biology*, **54**(1), 1–26.

Janiak-Spens, F., Sparling, J. M., Gurfinkel, M., and West, A. H. (1999). Differential stabilities of phosphorylated response regulator domains reflect functional roles of the yeast osmoregulatory SLN1 and SSK1 proteins. *Journal of bacteriology*, **181**(2), 411–417.

Janiak-Spens, F., Cook, P. F., and West, A. H. (2005). Kinetic analysis of YPD1-dependent phosphotransfer reactions in the yeast osmoregulatory phosphorelay system. *Biochemistry*, **44**(1), 377–86.

Johnson, H. A. (1970). Information theory in biology after 18 years. *Science*, **168**(3939), 1545–1550.

Kampen, N. v. (1961). A power series expansion of the master equation. *Canadian Journal of Physics*, **39**(4), 551–567.

Karr, J. R., Sanghvi, J. C., MacKlin, D. N., Gutschow, M. V., Jacobs, J. M., Bolival, B., Assad-Garcia, N., Glass, J. I., and Covert, M. W. (2012). A whole-cell computational model predicts phenotype from genotype. *Cell*, **150**(2), 389–401.

Kaserer, A. O., Andi, B., Cook, P. F., and West, A. H. (2009). Effects of osmolytes on the SLN1-YPD1-SSK1 phosphorelay system from Saccharomyces cerevisiae. *Biochemistry*, **48**(33), 8044–50.

Kenney, F. and Keeping, E. (1951). Mathematics of statistics-part two.

Kholodenko, B. N. (2000). Negative feedback and ultrasensitivity can bring about oscillations in the mitogen-activated protein kinase cascades. *European Journal of Biochemistry*, **267**(6), 1583–1588.

Kitano, H. (2002). Systems biology: a brief overview. *Science*, **295**(5560), 1662–1664.

Klinger, B. and Bluethgen, N. (2014). Consequences of feedback in signal transduction for targeted therapies. *Biochemical Society transactions*, **42**(4), 770–775.

Klipp, E., Nordlander, B., Krüger, R., Gennemark, P., and Hohmann, S. (2005). Integrative model of the response of yeast to osmotic shock. *Nature biotechnology*, **23**(8), 975–82.

Klipp, E., Liebermeister, W., Wierling, C., Kowald, A., Lehrach, H., and Herwig, R. (2013). *Systems biology*. John Wiley & Sons.

Kramers, H. A. (1940). Brownian motion in a field of force and the diffusion model of chemical reactions. *Physica*, **7**(4), 284–304.

Kreutz, C. and Timmer, J. (2009). Systems biology: experimental design. *The FEBS journal*, **276**(4), 923–942.

Kussell, E. and Leibler, S. (2005). Phenotypic diversity, population growth, and information in fluctuating environments. *Science*, **309**(5743), 2075–8.

Lafuerza, L. F. and Toral, R. (2010). On the Gaussian approximation for master equations. *Journal of Statistical Physics*, **140**(5), 917–933.

Lai, A., Sato, P. M., and Peisajovich, S. G. (2015). Evolution of Synthetic Signaling Scaffolds by Recombination of Modular Protein Domains. *ACS synthetic biology*.

Legewie, S., Herzel, H., Westerhoff, H. V., and Blüthgen, N. (2008). Recurrent design patterns in the feedback regulation of the mammalian signalling network. *Molecular systems biology*, **4**(1).

Levchenko, A., Bruck, J., and Sternberg, P. W. (2000). Scaffold proteins may biphasically affect the levels of mitogen-activated protein kinase signaling and reduce its threshold properties. *Proceedings of the National Academy of Sciences*, **97**(11), 5818–5823.

Levitzki, A. and Klein, S. (2010). Signal transduction therapy of cancer.

Li, S., Ault, A., Malone, C. L., Raitt, D., Dean, S., Johnston, L. H., Deschenes, R. J., and Fassler, J. S. (1998). The yeast histidine protein kinase, Sln1p, mediates phosphotransfer to two response regulators, Ssk1p and Skn7p. *The EMBO journal*, **17**(23), 6952–62.

Liang, S., Fuhrman, S., and Somogyi, R. (1998). Reveal, a general reverse engineering algorithm for inference of genetic network architectures.

Locke, J. C. W., Young, J. W., Fontes, M., Hernández Jiménez, M. J., and Elowitz, M. B. (2011). Stochastic pulse regulation in bacterial stress response. *Science (New York, N.Y.)*, **334**(6054), 366–9.

Lu, J. M.-Y., Deschenes, R. J., and Fassler, J. S. (2003). Saccharomyces cerevisiae histidine phosphotransferase Ypd1p shuttles between the nucleus and cytoplasm for SLN1-dependent phosphorylation of Ssk1p and Skn7p. *Eukaryotic cell*, **2**(6), 1304–14.

Lyakhov, I. G., Krishnamachari, A., and Schneider, T. D. (2008). Discovery of novel tumor suppressor p53 response elements using information theory. *Nucleic acids research*, **36**(11), 3828–33.

Macaulay, I. C., Haerty, W., Kumar, P., Li, Y. I., Hu, T. X., Teng, M. J., Goolam, M., Saurat, N., Coupland, P., Shirley, L. M., Smith, M., Van der Aa, N., Banerjee, R., Ellis, P. D., Quail, M. a., Swerdlow, H. P., Zernicka-Goetz, M., Livesey, F. J., Ponting, C. P., and Voet, T. (2015). G&T-seq: parallel sequencing of single-cell genomes and transcriptomes. *Nature Methods*, (April), 1–7.

Macia, J., Regot, S., Peeters, T., Conde, N., Solé, R., and Posas, F. (2009). Dynamic signaling in the Hog1 MAPK pathway relies on high basal signal transduction. *Science signaling*, **2**(63), ra13.

Maeda, T., Wurgler-Murphy, S. M., and Saito, H. (1994). A two-component system that regulates an osmosensing MAP kinase cascade in yeast. *Nature*, **369**(6477), 242–5.

Maeda, T., Takekawa, M., and Saito, H. (1995). Activation of yeast PBS2 MAPKK by MAPKKKs or by binding of an SH3-containing osmosensor. *Science*, **269**(5223), 554–558.

Maeder, C. I., Hink, M. A., Kinkhabwala, A., Mayr, R., Bastiaens, P. I., and Knop, M. (2007). Spatial regulation of Fus3 MAP kinase activity through a reaction-diffusion mechanism in yeast pheromone signalling. *Nature Cell Biology*, **9**(11), 1319–1326.

Mahanty, S. K., Wang, Y., Farley, F. W., and Elion, E. A. (1999). Nuclear shuttling of yeast scaffold Ste5 is required for its recruitment to the plasma membrane and activation of the mating MAPK cascade. *Cell*, **98**(4), 501–512.

Maiwald, T. and Timmer, J. (2008). Dynamical modeling and multi-experiment fitting with PottersWheel. *Bioinformatics*, **24**(18), 2037–2043.

Mangan, S. and Alon, U. (2003). Structure and function of the feed-forward loop network motif. *Proceedings of the National Academy of Sciences*, **100**(21), 11980–11985.

Marcus, S., Polverino, A., Barr, M., and Wigler, M. (1994). Complexes between Ste5 and components of the pheromone-responsive mitogen-activated protein kinase module. *Proceedings of the National Academy of Sciences*, **91**(16), 7762–7766.

Matis, T. I. and Guardiola, I. G. (2010). Achieving moment closure through cumulant neglect. *The Mathematica Journal*, **12**.

McAdams, H. H. and Arkin, A. (1997). Stochastic mechanisms in gene expression. *Proceedings of the National Academy of Sciences*, **94**(3), 814–819.

McClean, M. N., Mody, A., Broach, J. R., and Ramanathan, S. (2007). Cross-talk and decision making in MAP kinase pathways. *Nature genetics*, **39**(3), 409–14.

McKay, M. M., Ritt, D. A., and Morrison, D. K. (2009). Signaling dynamics of the KSR1 scaffold complex. *Proceedings of the National Academy of Sciences*, **106**(27), 11022–11027.

McNaught, A. D., Wilkinson, A., *et al.* (1997). *Compendium of chemical terminology*, volume 1669. Blackwell Science Oxford.

McQuarrie, D. A. (1967). Stochastic approach to chemical kinetics. *Journal of applied probability*, **4**(3), 413–478.

Mettetal, J. T., Muzzey, D., Gómez-Uribe, C., and van Oudenaarden, A. (2008). The frequency dependence of osmo-adaptation in Saccharomyces cerevisiae. *Science (New York, N.Y.)*, **319**(5862), 482–4.

Milo, R., Shen-Orr, S., Itzkovitz, S., Kashtan, N., Chklovskii, D., and Alon, U. (2002). Network motifs: simple building blocks of complex networks. *Science*, **298**(5594), 824–827.

Milo, R., Jorgensen, P., Moran, U., Weber, G., and Springer, M. (2010). BioNumbers—the database of key numbers in molecular and cell biology. *Nucleic acids research*, **38**(suppl 1), D750–D753.

Moore, T. I., Chou, C.-S., Nie, Q., Jeon, N. L., and Yi, T.-M. (2008). Robust spatial sensing of mating pheromone gradients by yeast cells. *PloS one*, **3**(12), e3865.

Moyal, J. (1949). Stochastic processes and statistical physics. *Journal of the Royal Statistical Society. Series B (Methodological)*, **11**(2), 150–210.

Muñoz-García, J., Neufeld, Z., and Kholodenko, B. N. (2009). Positional information generated by spatially distributed signaling cascades. *PLoS computational biology*, **5**(3), e1000330.

Muñoz-García, J., Kholodenko, B. N., and Neufeld, Z. (2010). Formation of intracellular concentration landscapes by multisite protein modification. *Biophysical journal*, **99**(1), 59–66.

Murakami, Y., Tatebayashi, K., and Saito, H. (2008). Two adjacent docking sites in the yeast Hog1 mitogen-activated protein (MAP) kinase differentially interact with the Pbs2 MAP kinase kinase and the Ptp2 protein tyrosine phosphatase. *Molecular and cellular biology*, **28**(7), 2481–2494.

Muzzey, D., Gómez-Uribe, C. a., Mettetal, J. T., and van Oudenaarden, A. (2009). A systems-level analysis of perfect adaptation in yeast osmoregulation. *Cell*, **138**(1), 160–71.

Natarajan, M., Lin, K.-M., Hsueh, R. C., Sternweis, P. C., and Ranganathan, R. (2006). A global analysis of cross-talk in a mammalian cellular signalling network. *Nature Cell Biology*, **8**(6), 571–580.

Ninfa, A. J. and Mayo, A. E. (2004). Hysteresis vs. graded responses: the connections make all the difference. *Science Signaling*, **2004**(232), pe20.

Nordsieck, A., Lamb Jr, W., and Uhlenbeck, G. (1940). On the theory of cosmic-ray showers I The furry model and the fluctuation problem.

O'Rourke, S. M. and Herskowitz, I. (2002). A third osmosensing branch in Saccharomyces cerevisiae requires the Msb2 protein and functions in parallel with the Sho1 branch. *Molecular and cellular biology*, **22**(13), 4739–4749.

Paliwal, S., Iglesias, P. A., Campbell, K., Hilioti, Z., Groisman, A., and Levchenko, A. (2007). MAPK-mediated bimodal gene expression and adaptive gradient sensing in yeast. *Nature*, **446**(7131), 46–51.

Parker, G. A., Smith, J. M., *et al.* (1990). Optimality theory in evolutionary biology. *Nature*, **348**(6296), 27–33.

Patel, A. K., Bhartiya, S., and Venkatesh, K. V. (2013). Analysis of osmoadaptation system in budding yeast suggests that regulated degradation of glycerol synthesis enzyme is key to near-perfect adaptation. *Systems and Synthetic Biology*.

Patterson, J. C., Klimenko, E. S., and Thorner, J. (2010a). Single-cell analysis reveals that insulation maintains signaling specificity between two yeast MAPK pathways with common components. *Science signaling*, **3**(144), ra75.

Patterson, J. C., Klimenko, E. S., and Thorner, J. (2010b). Single-cell analysis reveals that insulation maintains signaling specificity between two yeast MAPK pathways with common components. *Science signaling*, **3**(144), ra75.

Perlson, E., Michaelevski, I., Kowalsman, N., Ben-Yaakov, K., Shaked, M., Seger, R., Eisenstein, M., and Fainzilber, M. (2006). Vimentin binding to phosphorylated Erk sterically hinders enzymatic dephosphorylation of the kinase. *Journal of molecular biology*, **364**(5), 938–944.

Petelenz-Kurdziel, E., Kuehn, C., Nordlander, B., Klein, D., Hong, K.-K., Jacobson, T., Dahl, P., Schaber, J., Nielsen, J., Hohmann, S., and Klipp, E. (2013). Quantitative analysis of glycerol accumulation, glycolysis and growth under hyper osmotic stress. *PLoS computational biology*, **9**(6), e1003084.

Pfizer, G., Eccleston, A., Dhand, R., Shaw, R. J., Cantley, L. C., Karin, M., Pouysségur, J., Dayan, F., Mazure, N. M., Christofori, G., *et al.* (2006). Insight: Signalling in cancer. *Nature*, **441**, 424–430.

Poritz, M. A., Malmstrom, S., Kim, M. K.-H., Rossmeissl, P. J., and Kamb, A. (2001). Graded mode of transcriptional induction in yeast pheromone signalling revealed by single-cell analysis. *Yeast*, **18**(14), 1331–1338.

Posas, F. and Saito, H. (1997). Osmotic activation of the HOG MAPK pathway via Ste11p MAPKKK: scaffold role of Pbs2p MAPKK. *Science*, **276**(5319), 1702–1705.

Posas, F. and Saito, H. (1998). Activation of the yeast SSK2 MAP kinase kinase kinase by the SSK1 two-component response regulator. *The EMBO journal*, **17**(5), 1385–94.

Posas, F., Wurgler-Murphy, S. M., Maeda, T., Witten, E. A., Thai, T. C., and Saito, H. (1996a). Yeast HOG1 MAP kinase cascade is regulated by a multistep phosphorelay mechanism in the SLN1–YPD1–SSK1 "two-component" osmosensor. *Cell*, **86**(6), 865–875.

Posas, F., Wurgler-Murphy, S. M., Maeda, T., Witten, E. a., Thai, T. C., and Saito, H. (1996b). Yeast HOG1 MAP kinase cascade is regulated by a multistep phosphorelay mechanism in the SLN1-YPD1-SSK1 "two-component" osmosensor. *Cell*, **86**(6), 865–75.

Powell, K. (2004). All systems go. *The Journal of cell biology*, **165**(3), 299–303.

Qin, L., Dutta, R., Kurokawa, H., Ikura, M., and Inouye, M. (2000). A monomeric histidine kinase derived from EnvZ, an Escherichia coli osmosensor. *Molecular microbiology*, **36**(1), 24–32.

Qiu, P., Gentles, A. J., and Plevritis, S. K. (2009). Fast calculation of pairwise mutual information for gene regulatory network reconstruction. *Computer methods and programs in biomedicine*, **94**(2), 177–180.

Raitt, D. C., Posas, F., and Saito, H. (2000). Yeast Cdc42 GTPase and Ste20 PAK-like kinase regulate Sho1-dependent activation of the Hog1 MAPK pathway. *The EMBO journal*, **19**(17), 4623–31.

Raman, M., Chen, W., and Cobb, M. (2007). Differential regulation and properties of MAPKs. *Oncogene*, **26**(22), 3100–3112.

Raue, a., Kreutz, C., Maiwald, T., Bachmann, J., Schilling, M., Klingmüller, U., and Timmer, J. (2009). Structural and practical identifiability analysis of partially observed dynamical models by exploiting the profile likelihood. *Bioinformatics (Oxford, England)*, **25**(15), 1923–9.

Raue, a., Becker, V., Klingmüller, U., and Timmer, J. (2010). Identifiability and observability analysis for experimental design in nonlinear dynamical models. *Chaos (Woodbury, N.Y.)*, **20**(4), 045105.

Raue, A., Schilling, M., Bachmann, J., Matteson, A., Schelke, M., Kaschek, D., Hug, S., Kreutz, C., Harms, B. D., Theis, F. J., Klingmüller, U., and Timmer, J. (2013). Lessons Learned from Quantitative Dynamical Modeling in Systems Biology. *PLoS ONE*, **8**(9), e74335.

Reinker, S., Altman, R., and Timmer, J. (2006). Parameter estimation in stochastic biochemical reactions. *IEE Proceedings-Systems Biology*, **153**(4), 168–178.

Rhee, A., Cheong, R., and Levchenko, A. (2012). The application of information theory to biochemical signaling systems. *Physical biology*, **9**(4), 045011.

Rivoire, O. and Leibler, S. (2011). The Value of Information for Populations in Varying Environments. *Journal of Statistical Physics*, pages 1124–1166.

Rivoire, O. and Leibler, S. (2014). A model for the generation and transmission of variations in evolution. *Proceedings of the National Academy of Sciences*, **111**(19), E1940–E1949.

Rosen, R. (1986). Optimality in biology and medicine. *Journal of mathematical analysis and applications*, **119**(1), 203–222.

Ross, S. M. *et al.* (1996). *Stochastic processes*, volume 2. John Wiley & Sons New York.

Sabbagh, W., Flatauer, L. J., Bardwell, A. J., and Bardwell, L. (2001). Specificity of MAP kinase signaling in yeast differentiation involves transient versus sustained MAPK activation. *Molecular cell*, **8**(3), 683–691.

Samoilov, M., Plyasunov, S., and Arkin, A. P. (2005). Stochastic amplification and signaling in enzymatic futile cycles through noise-induced bistability with oscillations. *Proceedings of the National Academy of Sciences of the United States of America*, **102**(7), 2310–2315.

Sanghvi, J. C., Regot, S., Carrasco, S., Karr, J. R., Gutschow, M. V., Bolival, B., and Covert, M. W. (2013). Accelerated discovery via a whole-cell model. *Nature methods*, **10**(12), 1192–5.

Schaber, J., Kofahl, B., Kowald, A., and Klipp, E. (2006). A modelling approach to quantify dynamic crosstalk between the pheromone and the starvation pathway in baker's yeast. *Febs Journal*, **273**(15), 3520–3533.

Schaber, J., Baltanas, R., Bush, A., Klipp, E., and Colman-Lerner, A. (2012). Modelling reveals novel roles of two parallel signalling pathways and homeostatic feedbacks in yeast. *Molecular systems biology*, **8**, 622.

Schaeffer, H. J., Catling, A. D., Eblen, S. T., Collier, L. S., Krauss, A., and Weber, M. J. (1998). MP1: a MEK binding partner that enhances enzymatic activation of the MAP kinase cascade. *Science*, **281**(5383), 1668–1671.

Schilling, M., Maiwald, T., Bohl, S., Kollmann, M., Kreutz, C., Timmer, J., and Klingmüller, U. (2005). Computational processing and error reduction strategies for standardized quantitative data in biological networks. *Febs Journal*, **272**(24), 6400–6411.

Schneider, T. D. (2001). Strong minor groove base conservation in sequence logos implies DNA distortion or base flipping during replication and transcription initiation. *Nucleic acids research*, **29**(23), 4881–91.

Schneider, T. D. (2005). Claude Shannon: Biologist. *IEEE Engineering in Medicine and Biology Magazine*, **25**(1), 30–33.

Schneider, T. D. and Stephens, R. M. (1990). Sequence logos: a new way to display consensus sequences. *Nucleic acids research*, **18**(20), 6097–100.

Schüller, C., Brewster, J., Alexander, M., Gustin, M., and Ruis, H. (1994). The HOG pathway controls osmotic regulation of transcription via the stress response element (STRE) of the Saccharomyces cerevisiae CTT1 gene. *The EMBO Journal*, **13**(18), 4382.

Schwartz, M. A. and Madhani, H. D. (2004). Principles of MAP kinase signaling specificity in Saccharomyces cerevisiae. *Annu. Rev. Genet.*, **38**, 725–748.

Seger, R. and Krebs, E. G. (1995). The MAPK signaling cascade. *The FASEB journal*, **9**(9), 726–735.

Selimkhanov, J., Taylor, B., Yao, J., Pilko, A., Albeck, J., Hoffmann, A., Tsimring, L., and Wollman, R. (2014). Accurate information transmission through dynamic biochemical signaling networks. *Science*, **346**(6215), 1370–1373.

Sette, C., Inouye, C. J., Stroschein, S. L., Iaquinta, P. J., and Thorner, J. (2000). Mutational analysis suggests that activation of the yeast pheromone response mitogen-activated protein kinase pathway involves conformational changes in the Ste5 scaffold protein. *Molecular biology of the cell*, **11**(11), 4033–4049.

Shahrezaei, V. and Swain, P. S. (2008). Analytical distributions for stochastic gene expression. *Proceedings of the National Academy of Sciences of the United States of America*, **105**(45), 17256–17261.

Shannon, C. E. (1938). A symbolic analysis of relay and switching circuits. *American Institute of Electrical Engineers, Transactions of the*, **57**(12), 713–723.

Shannon, C. E. (1940). *An algebra for theoretical genetics*. Ph.D. thesis, Massachusetts Institute of Technology.

Shannon, C. E. (1948). A mathematical theory of communication. *Bell system technical journal*, **27**(3), 379—423.

Shinar, G., Milo, R., Martínez, M. R., and Alon, U. (2007). Input output robustness in simple bacterial signaling systems. *Proceedings of the National Academy of Sciences of the United States of America*, **104**(50), 19931–5.

Slaughter, B. D., Schwartz, J. W., and Li, R. (2007). Mapping dynamic protein interactions in MAP kinase signaling using live-cell fluorescence fluctuation spectroscopy and imaging. *Proceedings of the National Academy of Sciences*, **104**(51), 20320–20325.

Slonim, N., Atwal, G. S., Tkačik, G., and Bialek, W. (2005). Information-based clustering. *Proceedings of the National Academy of Sciences of the United States of America*, **102**(51), 18297–18302.

Smith, J. M. (2000). The Concept of Information in Biology.

Soufi, B., Kelstrup, C. D., Stoehr, G., Fröhlich, F., Walther, T. C., and Olsen, J. V. (2009). Global analysis of the yeast osmotic stress response by quantitative proteomics. *Molecular bioSystems*, **5**(11), 1337–1346.

Spiller, D. G., Wood, C. D., Rand, D. A., and White, M. R. (2010). Measurement of single-cell dynamics. *Nature*, **465**(7299), 736–745.

Stock, A. M., Robinson, V. L., and Goudreau, P. N. (2000). Two-component signal transduction. *Annual review of biochemistry*, **69**(1), 183–215.

Swain, P. S., Elowitz, M. B., and Siggia, E. D. (2002). Intrinsic and extrinsic contributions to stochasticity in gene expression. *Proceedings of the National Academy of Sciences of the United States of America*, **99**(20), 12795–800.

Tanaka, K., Tatebayashi, K., Nishimura, A., Yamamoto, K., Yang, H.-Y., and Saito, H. (2014). Yeast osmosensors hkr1 and msb2 activate the hog1 MAPK cascade by different mechanisms. *Science signaling*, **7**(314), ra21–ra21.

Tanigawa, M., Kihara, A., Terashima, M., Takahara, T., and Maeda, T. (2012). Sphingolipids regulate the yeast high-osmolarity glycerol response pathway. *Molecular and cellular biology*, **32**(14), 2861–70.

Tao, W., Malone, C. L., Ault, A. D., Deschenes, R. J., and Fassler, J. S. (2002). A cytoplasmic coiled-coil domain is required for histidine kinase activity of the yeast osmosensor, SLN1. *Molecular microbiology*, **43**(2), 459–73.

Tatebayashi, K., Tanaka, K., Yang, H.-Y., Yamamoto, K., Matsushita, Y., Tomida, T., Imai, M., and Saito, H. (2007). Transmembrane mucins Hkr1 and Msb2 are putative osmosensors in the SHO1 branch of yeast HOG pathway. *The EMBO Journal*, **26**(15), 3521–3533.

Terfve, C., Cokelaer, T., Henriques, D., MacNamara, A., Goncalves, E., Morris, M. K., van Iersel, M., Lauffenburger, D. A., and Saez-Rodriguez, J. (2012). CellNOptR: a flexible toolkit to train protein signaling networks to data using multiple logic formalisms. *BMC systems biology*, **6**(1), 133.

Thomson, T. M., Benjamin, K. R., Bush, A., Love, T., Pincus, D., Resnekov, O., Yu, R. C., Gordon, A., Colman-Lerner, A., Endy, D., and Brent, R. (2011). Scaffold number in yeast signaling system sets tradeoff between system output and dynamic range. *Proceedings of the National Academy of Sciences of the United States of America*, **108**(50), 20265–70.

Thorp, E. O. (1998). The invention of the first wearable computer. In *Wearable Computers, 1998. Digest of Papers. Second International Symposium on*, pages 4–8. IEEE.

Tkačik, G. (2010). From statistical mechanics to information theory: understanding biophysical information-processing systems. *arXiv preprint arXiv:1006.4291*.

Tkacik, G., Prentice, J. S., Balasubramanian, V., and Schneidman, E. (2010). Optimal population coding by noisy spiking neurons. *Proceedings of the National Academy of Sciences of the United States of America*, **107**(32), 14419–24.

Tkačik, G. and Walczak, A. M. (2011). Information transmission in genetic regulatory networks: a review. *Journal of physics. Condensed matter : an Institute of Physics journal*, **23**(15), 153102.

Tkačik, G., Callan, C., and Bialek, W. (2008a). Information capacity of genetic regulatory elements. *Physical Review E*, **78**(1), 1–17.

Tkačik, G., Callan, C. G., and Bialek, W. (2008b). Information flow and optimization in transcriptional regulation. *Proceedings of the National Academy of Sciences of the United States of America*, **105**(34), 12265–70.

Tkačik, G., Walczak, A., and Bialek, W. (2009). Optimizing information flow in small genetic networks. *Physical Review E*, **80**(3), 1–18.

Tostevin, F. and Ten Wolde, P. R. (2009). Mutual information between input and output trajectories of biochemical networks. *Physical Review Letters*, **102**(21), 218101.

Vaga, S. (2013). *Investigation of budding yeast signaling transmission and integration by high-throughput quantitative mass spectrometry*. Ph.D. thesis, Diss., Eidgenössische Technische Hochschule ETH Zürich, Nr. 21523, 2013.

Vaga, S., Bernardo-Faura, M., Cokelaer, T., Maiolica, A., Barnes, C. A., Gillet, L. C., Hegemann, B., Drogen, F., Sharifian, H., Klipp, E., *et al.* (2014). Phosphoproteomic analyses reveal novel cross-modulation mechanisms between two signaling pathways in yeast. *Molecular systems biology*, **10**(12).

Van Kampen, N. (2011). *Stochastic Processes in Physics and Chemistry*. Elsevier.

Vinga, S. (2013). Information theory applications for biological sequence analysis. *Briefings in bioinformatics*, page bbt068.

Voliotis, M., Perrett, R. M., McWilliams, C., McArdle, C. A., and Bowsher, C. G. (2014). Information transfer by leaky, heterogeneous, protein kinase signaling systems. *Proceedings of the National Academy of Sciences*, **111**(3), E326–E333.

Wallace, E., Gillespie, D., Sanft, K., and Petzold, L. (2012). Linear noise approximation is valid over limited times for any chemical system that is sufficiently large. *IET systems biology*, **6**(4), 102–115.

Waltermann, C. and Klipp, E. (2011). Information theory based approaches to cellular signaling. *Biochimica et Biophysica Acta (BBA)-General Subjects*, **1810**(10), 924–932.

Watson, J. D., Crick, F. H., *et al.* (1953). Molecular structure of nucleic acids. *Nature*, **171**(4356), 737–738.

West, A. H. and Stock, A. M. (2001). Histidine kinases and response regulator proteins in two-component signaling systems. *Trends in biochemical sciences*, **26**(6), 369–376.

Whitmarsh, A. J. and Davis, R. J. (1998). Structural organization of MAP-kinase signaling modules by scaffold proteins in yeast and mammals. *Trends in biochemical sciences*, **23**(12), 481–485.

Widmann, C., Gibson, S., Jarpe, M. B., and Johnson, G. L. (1999). Mitogen-activated protein kinase: conservation of a three-kinase module from yeast to human. *Physiological reviews*, **79**(1), 143–180.

Wiles, A. (1995). Modular elliptic curves and Fermat's last theorem. *Annals of Mathematics*, pages 443–551.

Witzel, F., Maddison, L. E., and Blüthgen, N. (2012). How scaffolds shape MAPK signaling: what we know and opportunities for systems approaches. *Frontiers in physiology*, **3**, 475.

Wurgler-Murphy, S. M., Maeda, T., Witten, E. A., and Saito, H. (1997). Regulation of the Saccharomyces cerevisiae HOG1 mitogen-activated protein kinase by the PTP2 and PTP3 protein tyrosine phosphatases. *Molecular and Cellular Biology*, **17**(3), 1289–1297.

Yeger-Lotem, E., Sattath, S., Kashtan, N., Itzkovitz, S., Milo, R., Pinter, R. Y., Alon, U., and Margalit, H. (2004). Network motifs in integrated cellular networks of transcription–regulation and protein–protein interaction. *Proceedings of the National Academy of Sciences of the United States of America*, **101**(16), 5934–5939.

Yu, H., Kim, P. M., Sprecher, E., Trifonov, V., and Gerstein, M. (2007). The importance of bottlenecks in protein networks: correlation with gene essentiality and expression dynamics. *PLoS computational biology*, **3**(4), e59.

Yu, R. C., Pesce, C. G., Colman-Lerner, A., Lok, L., Pincus, D., Serra, E., Holl, M., Benjamin, K., Gordon, A., and Brent, R. (2008). Negative feedback that improves information transmission in yeast signalling. *Nature*, **456**(7223), 755–761.

Zalatan, J. G., Coyle, S. M., Rajan, S., Sidhu, S. S., and Lim, W. A. (2012). Conformational control of the Ste5 scaffold protein insulates against MAP kinase misactivation. *Science*, **337**(6099), 1218–1222.

Zheng, Y. and Kwoh, C. K. (2006). Cancer classification with microRNA expression patterns found by an information theory approach. *Journal of computers*, **1**(5), 30–39.

Zi, Z., Liebermeister, W., and Klipp, E. (2010). A quantitative study of the Hog1 MAPK response to fluctuating osmotic stress in Saccharomyces cerevisiae. *PloS one*, **5**(3), e9522.

Danksagungen / Acknowledgements

Das Ende einer Arbeit ist möglicherweise immer der Anfang einer Arbeit. Ich möchte den Einschnitt an dieser Stelle jedoch dazu nutzen, kurz innezuhalten und nach all diesen Seiten noch einpaar weitere, noch wichtige Worte aufzuschreiben.

Mein größter Dank gebührt meiner Freundin Katharina. Ich möchte ihr für all die Zeit, Verständnis und Geduld danken, für all die Unterstützung und Kraft die sie in mich steckt. Es ist schön, dich zu haben und die Welt mit dir zu teilen - ihre schönen und auch ihre stressigeren Seiten.

Desweiteren ist es mir ein Bedürftnis, meiner Betreuerin Edda Klipp zu danken, die entscheidenden Anteil an dieser Arbeit hat. Vielen Dank für eine stets offene Tür, ein inspirierendes und gesundes Arbeitsumfeld und die Möglichkeit, Wissenschaft zu machen und so mein Leben mit interessanten Inhalten und tollen Leuten zu füllen.

Zu diesen Leuten gehören ganz besonders Katja Tummler, Matthias Reis, Wolfgang Giese, Björn Goldenbogen und Marcus Krantz[68], denen ich für ihre Freundschaft und Hilfe danken möchte. Ihr habt die letzten Jahre erst zu der schönen Zeit gemacht, die sie waren. Dies gilt auch für die vielen weiteren aktuellen und ehemaligen Mitglieder des Labs, die ich alle ins Herz geschlossen habe.

Nicht zuletzt möchte ich auch den Gutachtern dieser Arbeit sowie den Mitgliedern der Promotionskommission danken. Für ihre Zeit und Mühe, für ihr Interesse und Verständnis, aber auch für die interessanten Diskussionen und anregenden Lehrinhalte der letzten Jahre.

Der größte Dank jedoch gehört meinen Eltern, die mich seit 30 Jahren auf meinem Weg begleiten, mir jeden Rückhalt geben und mir all dies hier erst ermöglicht haben. Vielen Dank!

Diese Arbeit wurde gefödert durch die Deutsche Forschungsgemeinschaft (DFG) im Rahmen des Graduiertenkollegs “Computational Systems Biology” (GRK 1772) sowie das Bundesministerium für Bildung und Forschung im Kontext des Sonderforschungsbereiches SFB 740, “From molecules to modules”. Ich möchte mich für diese Unterstützung herzlich bedanken, da ohne sie meine Arbeit nicht möglich gewesen wäre.

[68] Besonders dankbar bin ich für die gemeinsame Arbeit an unserem Scaffolding-Projekt, welches durch dich initiiert und erdacht wurde, bis es zu unserem gemeinsamen Werk wurde. Bitte verzeih alle Verzögerungen, die ich dir aufbürde.

A. Moment Equations for the crosstalk model of chapter 4

In this section, we give the complete set of differential equations for the crosstalk model, described in chapter 4 as well as the parameters used. The equations have been simulated and plotted with the D2D ("Data2Dynamics") framework in MATLAB R2015a (see section 2.3).

Parameters for the crosstalk model

The parameters values used for simulation of the model are given in Tab. 4.

Initial states

The model does not consider production and degradation of components. Mass in the model is conserved and for purposes of dimension reduction we set initial states of certain species and compounds to a very low level, while the others define the total mass of the species. The initial states for all species are listed in Tab. 5.

The model was run for $t_{\text{off}} = 5\text{min}$, before the different stress scenarios were applied for $t_{\text{on}} = 45\text{min}$ as in the stress duration of the experiments. For this, the species "NaCl" and "alpha" are initially set to zero and at 5min switched "on" to a value of 1, indicating the addition of salt and/or pheromone stress in the experiment.

ODEs for the crosstalk model

The complete set of ODEs that govern the crosstalk model are stated in Fig. 28 to 30.[69]

[69]Note, that the system is split over several figures for better readability.

Name	Value
k_p_Sln1	1e+02 $\cdot s^{-1}$
k_p_Ssk1	1e+05 $\cdot (\mathrm{mM} \cdot s)^{-1}$
k_p_Ssk2	1e+02 $\cdot (\mathrm{mM} \cdot s)^{-1}$
k_p_Ste20	0.00038 $\cdot s^{-1}$
k_p_Ste20_Ste50_Ste11	0.25 $\cdot s^{-1}$
k_p_Ste50_Ste11_Fus3	1.7e+02 $\cdot s^{-1}$
k_p_Ypd1	1e+06 $\cdot (\mathrm{mM} \cdot s)^{-1}$
k_p_Far1	0.77 $\cdot (\mathrm{mM} \cdot s)^{-1}$
k_p_Hog1	13.0 $\cdot (\mathrm{mM} \cdot s)^{-1}$
k_p_Pbs2	1e+05 $\cdot (\mathrm{mM} \cdot s)^{-1}$
k_p_Pbs2$_P$_Ste50_Ste11$_P$	1e+03 $\cdot s^{-1}$
k_dep_Hog1	2.9 $\cdot s^{-1}$
k_dep_Pbs2$_P$	1e+02 $\cdot s^{-1}$
k_dep_Ssk1	1.0 $\cdot s^{-1}$
k_dep_Ssk2	0.17 $\cdot s^{-1}$
k_dep_Ste20	12.0 $\cdot s^{-1}$
k_dep_Ste50_Ste11	0.0031 $\cdot s^{-1}$
k_dep_Far1	0.12 $\cdot s^{-1}$
k_dep_Fus3	3.1e-05 $\cdot s^{-1}$
k_act_Ptp2	1.4 $\cdot s^{-1}$
k_act_Sho1	1e+03 $\cdot s^{-1}$
k_basal_deg	0.09 $\cdot s^{-1}$
k_comp_Pbs2_Ste50_Ste11	7.1e+02 $\cdot (\mathrm{mM} \cdot^2 s)^{-1}$
k_comp_Ste20_Ste50_Ste11	1.6 $\cdot (\mathrm{mM} \cdot s)^{-1}$
k_comp_Ste50_Ste11	0.24 $\cdot (\mathrm{mM} \cdot s)^{-1}$
k_comp_Ste50_Ste11_Fus3	49 $\cdot (\mathrm{mM} \cdot s)^{-1}$
k_deact_Sho1	1.1e-05 $\cdot s^{-1}$
k_decomp_Pbs2_Ste50_Ste11	1.2e+02 $\cdot s^{-1}$
k_decomp_Pbs2_Ste50_Ste11$_P$	1e-05 $\cdot s^{-1}$
k_decomp_Ste20_Ste50_Ste11	1.2e-05 $\cdot s^{-1}$
k_decomp_Ste20_Ste50_Ste11$_P$	0.0062 $\cdot s^{-1}$
k_decomp_Ste50_Ste11	4e+02 $\cdot s^{-1}$
k_decomp_Ste50_Ste11_Fus3	2.4e-05 $\cdot s^{-1}$
k_decomp_Ste50_Ste11_Fus3$_P$	0.42 $\cdot s^{-1}$

Table 4: Parameter set of the rates used in the equations of Fig. 28 to 30.

Name	Initial value	Name	Initial value
Far1	0.74 mM	$Sho1_P$	10^{-5}mM
$Far1_P$	10^{-5}mM	Ste11	0.013 mM
Fus3	0.22 mM	Ste20	10 mM
$Fus3_P$	10^{-5}mM	$Ste20_P$	10^{-5}mM
Sln1	10^{-5}mM	Ste50	0.035 mM
$Sln1_P$	0.016 mM	FB1	0.0 (a.u.)
Ypd1	10^{-5}mM	FB2	0.0 (a.u.)
$Ypd1_P$	0.14 mM	FB3	0.0 (a.u.)
Ssk1	10^{-5}mM	$FB4_{inact}$	1.0 (a.u.)
$Ssk1_P$	0.039 mM	$FB4_{act}$	0.0 (a.u.)
Ssk2	0.0004 mM	Ste50_Ste11	10^{-5}mM
$Ssk2_P$	10^{-5}mM	$Ste50\_Ste11_P$	10^{-5}mM
Pbs2	0.051 mM	Ste20_Ste50_Ste11	10^{-5}mM
$Pbs2_P$	10^{-5}mM	$Ste20\_Ste50\_Ste11_P$	10^{-5}mM
Hog1	0.17 mM	Ste50_Ste11_Pbs2	10^{-5}mM
$Hog1_P$	10^{-5}mM	$Ste50\_Ste11\_Pbs2_P$	10^{-5}mM
Ptp2	0.005 mM	Ste50_Ste11_Fus3	10^{-5}mM
Sho1	0.057 mM	$Ste50\_Ste11\_Fus3_P$	10^{-5}mM

Table 5: Initial values for all species considered in the equations (Fig. 28 to 30).

$$\frac{d\text{Sln1}}{dt} = +\text{k\_p\_Ypd1} \cdot \text{Sln1}_P \cdot \text{Ypd1} - \text{k\_p\_Sln1} \cdot \text{Sln1} \cdot (1 - \text{NaCl} + \text{FB4}_{act})$$
$$\frac{d\text{Sln1}_P}{dt} = -\text{k\_p\_Ypd1} \cdot \text{Sln1}_P \cdot \text{Ypd1} + \text{k\_p\_Sln1} \cdot \text{Sln1} \cdot (1 - \text{NaCl} + \text{FB4}_{act})$$
$$\frac{d\text{Ypd1}}{dt} = -\text{k\_p\_Ypd1} \cdot \text{Sln1}_P \cdot \text{Ypd1} + \text{k\_p\_Ssk1} \cdot \text{Ssk1} \cdot \text{Ypd1}_P$$
$$\frac{d\text{Ypd1}_P}{dt} = +\text{k\_p\_Ypd1} \cdot \text{Sln1}_P \cdot \text{Ypd1} - \text{k\_p\_Ssk1} \cdot \text{Ssk1} \cdot \text{Ypd1}_P$$
$$\frac{d\text{Ssk1}}{dt} = -\text{k\_p\_Ssk1} \cdot \text{Ssk1} \cdot \text{Ypd1}_P + \text{k\_dep\_Ssk1} \cdot \text{Ssk1}_P$$
$$\frac{d\text{Ssk1}_P}{dt} = +\text{k\_p\_Ssk1} \cdot \text{Ssk1} \cdot \text{Ypd1}_P - \text{k\_dep\_Ssk1} \cdot \text{Ssk1}_P$$

Fig. 27: ODEs for the Sln1 phosphorelay of the model.

$$\frac{d\text{Ssk2}}{dt} = -\text{k\_p\_Ssk2} \cdot \text{Ssk1} \cdot \text{Ssk2} + \text{k\_dep\_Ssk2} \cdot \text{Ssk2}_P$$
$$\frac{d\text{Ssk2}_P}{dt} = \text{k\_p\_Ssk2} \cdot \text{Ssk1} \cdot \text{Ssk2} - \text{k\_dep\_Ssk2} \cdot \text{Ssk2}_P$$
$$\frac{d\text{Pbs2}}{dt} = \text{k\_dep\_Pbs2}_P \cdot \text{Pbs2}_P - \text{k\_p\_Pbs2} \cdot \text{Pbs2} \cdot \text{Ssk2}_P - \text{k\_comp\_Pbs2\_Ste50\_Ste11} \cdot \ldots$$
$$\ldots \cdot \text{Pbs2} \cdot \text{Ste50\_Ste11}_P \cdot \text{Sho1}_P + \text{k\_decomp\_Pbs2\_Ste50\_Ste11} \cdot \text{Ste50\_Ste11\_Pbs2}$$
$$\frac{d\text{Pbs2}_P}{dt} = -\text{k\_dep\_Pbs2}_P \cdot \text{Pbs2}_P + \text{k\_p\_Pbs2} \cdot \text{Pbs2} \cdot \text{Ssk2}_P$$
$$+ \text{k\_decomp\_Pbs2\_Ste50\_Ste11}_P \cdot \text{Ste50\_Ste11\_Pbs2}_P$$
$$\frac{d\text{Hog1}}{dt} = \text{k\_dep\_Hog1} \cdot \text{Hog1}_P \cdot \text{Ptp2} - \text{k\_p\_Hog1} \cdot \text{Hog1} \cdot \text{Pbs2}_P$$
$$\frac{d\text{Hog1}_P}{dt} = -\text{k\_dep\_Hog1} \cdot \text{Hog1}_P \cdot \text{Ptp2} + \text{k\_p\_Hog1} \cdot \text{Hog1} \cdot \text{Pbs2}_P$$
$$\frac{d\text{Ptp2}}{dt} = \text{k\_act\_Ptp2} \cdot \text{Hog1}_P - \text{k\_basal\_deg} \cdot \text{Ptp2}$$
$$\frac{d\text{FB1}}{dt} = (\text{Hog1}_P - \text{FB1})$$
$$\frac{d\text{FB2}}{dt} = (\text{FB1} - \text{FB2})/25$$
$$\frac{d\text{FB3}}{dt} = (\text{FB2} - \text{FB3})/25$$
$$\frac{d\text{FB4}_{inact}}{dt} = -\text{FB3} \cdot \text{FB4}_{inact}/5$$
$$\frac{d\text{FB4}_{act}}{dt} = \text{FB3} \cdot \text{FB4}_{inact}/5$$
$$\frac{d\text{Sho1}}{dt} = -\text{k\_act\_Sho1} \cdot \text{Sho1} \cdot NaCl + \text{k\_deact\_Sho1} \cdot \text{Sho1}_P$$
$$\frac{d\text{Sho1}_P}{dt} = \text{k\_act\_Sho1} \cdot \text{Sho1} \cdot NaCl - \text{k\_deact\_Sho1} \cdot \text{Sho1}_P$$

Fig. 28: ODEs for the MAPK cascade, the delayed feedback (FB) as well as the Sho-branch.

$$\frac{d\text{Fus3}}{dt} = \text{k\_dep\_Fus3} \cdot \text{Fus3}_P - \text{k\_comp\_Ste50\_Ste11\_Fus3} \cdot \text{Fus3} \cdot \text{Ste50\_Ste11}_P \cdot (alpha) + \text{k\_decomp\_Ste50\_Ste11\_Fus3} \cdot \text{Ste50\_Ste11\_Fus3}$$

$$\frac{d\text{Fus3}_P}{dt} = -\text{k\_dep\_Fus3} \cdot \text{Fus3}_P + \text{k\_decomp\_Ste50\_Ste11\_Fus3}_P \cdot \text{Ste50\_Ste11\_Fus3}_P$$

$$\frac{d\text{Far1}}{dt} = \text{k\_dep\_Far1} \cdot \text{Far1}_P - \text{k\_p\_Far1} \cdot \text{Far1} \cdot \text{Fus3}_P$$

$$\frac{d\text{Far1}_P}{dt} = -\text{k\_dep\_Far1} \cdot \text{Far1}_P + \text{k\_p\_Far1} \cdot \text{Far1} \cdot \text{Fus3}_P$$

Fig. 29: ODEs for the pheromone pathway.

$$
\begin{aligned}
\frac{d\text{Ste20}}{dt} &= \text{k\_dep\_Ste20} \cdot \text{Ste20}_P - \text{k\_p\_Ste20} \cdot \text{Ste20} \cdot (alpha + NaCl) \\
\frac{d\text{Ste20}_P}{dt} &= -\text{k\_comp\_Ste20\_Ste50\_Ste11} \cdot \text{Ste20}_P \cdot \text{Ste50\_Ste11} + \text{k\_decomp\_Ste20\_Ste50\_Ste11} \cdot \text{Ste20\_Ste50\_Ste11} \\
&\quad - \text{k\_dep\_Ste20} \cdot \text{Ste20}_P + \text{k\_p\_Ste20} \cdot \text{Ste20} \cdot (alpha + NaCl) + \text{k\_decomp\_Ste20\_Ste50\_Ste11}_P \cdot \text{Ste20\_Ste50\_Ste11}_P \\
\frac{d\text{Ste50}}{dt} &= \text{k\_decomp\_Ste50\_Ste11} \cdot \text{Ste50\_Ste11} \cdot (\text{Hog1}_P/(\text{Hog1} + \text{Hog1}_P)) \cdot (\text{Fus3}_P/(\text{Fus3} + \text{Fus3}_P)) \\
&\quad - \text{k\_comp\_Ste50\_Ste11} \cdot \text{Ste11} \cdot \text{Ste50} \\
\frac{d\text{Ste11}}{dt} &= \text{k\_decomp\_Ste50\_Ste11} \cdot \text{Ste50\_Ste11} \cdot (\text{Hog1}_P/(\text{Hog1} + \text{Hog1}_P)) \cdot (\text{Fus3}_P/(\text{Fus3} + \text{Fus3}_P)) \\
&\quad - \text{k\_comp\_Ste50\_Ste11} \cdot \text{Ste11} \cdot \text{Ste50} \\
\frac{d\text{Ste50\_Ste11}}{dt} &= -\text{k\_decomp\_Ste50\_Ste11} \cdot \text{Ste50\_Ste11} \cdot (\text{Hog1}_P/(\text{Hog1} + \text{Hog1}_P)) \cdot (\text{Fus3}_P/(\text{Fus3} + \text{Fus3}_P)) \\
&\quad + \text{k\_comp\_Ste50\_Ste11} \cdot \text{Ste11} \cdot \text{Ste50} + \text{k\_dep\_Ste50\_Ste11} \cdot \text{Ste50\_Ste11}_P \\
&\quad + \text{k\_decomp\_Ste20\_Ste50\_Ste11} \cdot \text{Ste20\_Ste50\_Ste11} - \text{k\_comp\_Ste20\_Ste50\_Ste11} \cdot \text{Ste20}_P \cdot \text{Ste50\_Ste11} \\
\frac{d\text{Ste50\_Ste11}_P}{dt} &= -\text{k\_dep\_Ste50\_Ste11} \cdot \text{Ste50\_Ste11}_P - \text{k\_comp\_Ste50\_Ste11\_Fus3} \cdot \text{Fus3} \cdot \text{Ste50\_Ste11}_P \cdot (alpha) \\
&\quad + \text{k\_decomp\_Ste50\_Ste11\_Fus3} \cdot \text{Ste50\_Ste11\_Fus3} + \text{k\_decomp\_Ste20\_Ste50\_Ste11}_P \cdot \text{Ste20\_Ste50\_Ste11}_P \\
&\quad - \text{k\_comp\_Pbs2\_Ste50\_Ste11} \cdot \text{Pbs2} \cdot \text{Ste50\_Ste11}_P \cdot \text{Sho1}_P + \text{k\_decomp\_Pbs2\_Ste50\_Ste11}_P \cdot \text{Ste50\_Ste11\_Pbs2}_P \\
&\quad + \text{k\_decomp\_Ste50\_Ste11\_Fus3}_P \cdot \text{Ste50\_Ste11\_Fus3}_P + \text{k\_decomp\_Pbs2\_Ste50\_Ste11} \cdot \text{Ste50\_Ste11\_Pbs2} \\
\frac{d\text{Ste20\_Ste50\_Ste11}}{dt} &= -\text{k\_p\_Ste20\_Ste50\_Ste11} \cdot \text{Ste20\_Ste50\_Ste11} - \text{k\_decomp\_Ste20\_Ste50\_Ste11} \cdot \text{Ste20\_Ste50\_Ste11} \\
&\quad + \text{k\_comp\_Ste20\_Ste50\_Ste11} \cdot \text{Ste20}_P \cdot \text{Ste50\_Ste11} \\
\frac{d\text{Ste20\_Ste50\_Ste11}_P}{dt} &= \text{k\_p\_Ste20\_Ste50\_Ste11} \cdot \text{Ste20\_Ste50\_Ste11} - \text{k\_decomp\_Ste20\_Ste50\_Ste11}_P \cdot \text{Ste20\_Ste50\_Ste11}_P \\
\frac{d\text{Ste50\_Ste11\_Pbs2}}{dt} &= \text{k\_comp\_Pbs2\_Ste50\_Ste11} \cdot \text{Pbs2} \cdot \text{Ste50\_Ste11}_P \cdot \text{Sho1}_P - \text{k\_p\_Pbs2}_P\text{\_Ste50\_Ste11}_P \cdot \text{Ste50\_Ste11\_Pbs2} \\
&\quad - \text{k\_decomp\_Pbs2\_Ste50\_Ste11} \cdot \text{Ste50\_Ste11\_Pbs2} \\
\frac{d\text{Ste50\_Ste11\_Pbs2}_P}{dt} &= \text{k\_p\_Pbs2}_P\text{\_Ste50\_Ste11}_P \cdot \text{Ste50\_Ste11\_Pbs2} - \text{k\_decomp\_Pbs2\_Ste50\_Ste11}_P \cdot \text{Ste50\_Ste11\_Pbs2}_P \\
\frac{d\text{Ste50\_Ste11\_Fus3}}{dt} &= -\text{k\_p\_Ste50\_Ste11\_Fus3} \cdot \text{Ste50\_Ste11\_Fus3} + \text{k\_comp\_Ste50\_Ste11\_Fus3} \cdot \text{Fus3} \cdot \text{Ste50\_Ste11}_P \cdot (alpha) \\
&\quad - \text{k\_decomp\_Ste50\_Ste11\_Fus3} \cdot \text{Ste50\_Ste11\_Fus3} \\
\frac{d\text{Ste50\_Ste11\_Fus3}_P}{dt} &= \text{k\_p\_Ste50\_Ste11\_Fus3} \cdot \text{Ste50\_Ste11\_Fus3} - \text{k\_decomp\_Ste50\_Ste11\_Fus3}_P \cdot \text{Ste50\_Ste11\_Fus3}_P
\end{aligned}
$$

Fig. 30: ODEs for the species that are shared between pheromone- and Hog-pathway in the model.

B. Moment Equations for the Models of chapter 5

B.1. Moments for the mixed channel

Here, we give the full sets of equations governing the first two moments of the models implemented in chapter 5 of this work. The simulations were carried out in MATLAB R2015a, using the ODE45 solver.

First, the equation systems (47) and (48) in Fig. 31 describes the dynamics of the mixed signaling system for the expected value of $N = 6$ species $\{P_i, P_i^A\}$, where $i = 1, 2, 3$, as well as their second moments, from which (together with the first) variances and covariances can be calculated. As described in the text (see section 5.2.1), the ODE system comprises $\frac{N(N+3)}{2} = 27$ equations and utilizes 6 parameters $\{k_{i,forw}, k_{i,rev}\}$, where $i = \{1, 2, 3\}$.
The input strength (stress level) for the first signaling tier has been varied with

$$\begin{aligned} X &= k_{1,forw} \\ &= [0.001, 0.0025, 0.005, 0.01, 0.02, 0.04, 0.08, 0.16, \ldots \\ &\qquad \ldots, 0.32, 0.64, 1.28, 2.54, 5.12, 10] \cdot s^{-1}. \end{aligned}$$

For initial values as well as further variation of parameters, see chapter 5 (particularly Tab. 3 for the standard models) in order to recreate the results from the text.

The equations can be derived as explained in section 2.2.3. Yet, since these equations only depend on the stoichiometry of the system, giving us its master equation and subsequently the moment generating function (and thus its closure), automatic generation of the equations (for example from an SBML description of the reaction system) is recommended to avoid mistakes in the lengthy calculations.

$$
\begin{aligned}
\frac{dE\left[P_1P_1\right]}{dt} &= 2k_{1,rev}E\left[P_1P_1^A\right] + k_{1,rev}E\left[P_1^A\right] - 2k_{1,forw}E\left[P_1P_1\right] + k_{1,forw}E\left[P_1\right]\\
\frac{dE\left[P_1P_1^A\right]}{dt} &= -k_{1,rev}E\left[P_1P_1^A\right] + k_{1,rev}E\left[P_1^AP_2^A\right] - k_{1,rev}E\left[P_1^A\right] + k_{1,forw}E\left[P_1P_1\right] - k_{1,forw}E\left[P_1P_1^A\right] - k_{1,forw}E\left[P_1\right]\\
\frac{dE\left[P_1P_2\right]}{dt} &= 2k_{2,forw}E\left[P_1\right]E\left[P_1^A\right]E\left[P_2\right] - k_{2,forw}E\left[P_1P_1^A\right]E\left[P_2\right] - k_{2,forw}E\left[P_1P_2\right]E\left[P_1^A\right] - k_{2,forw}E\left[P_2P_1^A\right]E\left[P_1\right] + k_{1,rev}E\left[P_2P_1^A\right] + k_{2,rev}E\left[P_1P_2^A\right] - k_{1,forw}E\left[P_1P_2\right]\\
\frac{dE\left[P_1P_2^A\right]}{dt} &= -2k_{2,forw}E\left[P_1\right]E\left[P_1^A\right]E\left[P_2\right] + k_{2,forw}E\left[P_1P_1^A\right]E\left[P_2\right] + k_{2,forw}E\left[P_1P_2\right]E\left[P_1^A\right] + k_{2,forw}E\left[P_2P_1^A\right]E\left[P_1\right] + k_{1,rev}E\left[P_1^AP_2^A\right] - k_{2,rev}E\left[P_1P_2^A\right] - k_{1,forw}E\left[P_1P_2^A\right]\\
\frac{dE\left[P_1P_3\right]}{dt} &= 2k_{3,forw}E\left[P_1\right]E\left[P_2^A\right]E\left[P_3\right] - k_{3,forw}E\left[P_1P_2^A\right]E\left[P_3\right] - k_{3,forw}E\left[P_1P_3\right]E\left[P_2^A\right] - k_{3,forw}E\left[P_3P_2^A\right]E\left[P_1\right] + k_{1,rev}E\left[P_3P_1^A\right] + k_{3,rev}E\left[P_1P_3^A\right] - k_{1,forw}E\left[P_1P_3\right]\\
\frac{dE\left[P_1P_3^A\right]}{dt} &= -2k_{3,forw}E\left[P_1\right]E\left[P_2^A\right]E\left[P_3\right] + k_{3,forw}E\left[P_1P_2^A\right]E\left[P_3\right] + k_{3,forw}E\left[P_1P_3\right]E\left[P_2^A\right] + k_{3,forw}E\left[P_3P_2^A\right]E\left[P_1\right] + k_{1,rev}E\left[P_1^AP_3^A\right] - k_{3,rev}E\left[P_1P_3^A\right] - k_{1,forw}E\left[P_1P_3^A\right]\\
\frac{dE\left[P_1^AP_1^A\right]}{dt} &= -2k_{1,rev}E\left[P_1^AP_2^A\right] + k_{1,rev}E\left[P_1^A\right] + 2k_{1,forw}E\left[P_1P_1^A\right] + k_{1,forw}E\left[P_1\right]\\
\frac{dE\left[P_1^AP_2^A\right]}{dt} &= -2k_{2,forw}E\left[P_1^A\right]^2E\left[P_2\right] + k_{2,forw}E\left[P_1^AP_2^A\right]E\left[P_2\right] + 2k_{2,forw}E\left[P_2P_1^A\right]E\left[P_1^A\right] - k_{1,rev}E\left[P_1^AP_2^A\right] - k_{2,rev}E\left[P_1^AP_2^A\right] + k_{1,forw}E\left[P_1P_2^A\right]\\
\frac{dE\left[P_1^AP_3^A\right]}{dt} &= -2k_{3,forw}E\left[P_1^A\right]E\left[P_2^A\right]E\left[P_3\right] + k_{3,forw}E\left[P_1^AP_2^A\right]E\left[P_3\right] + k_{3,forw}E\left[P_3P_1^A\right]E\left[P_2^A\right] + k_{3,forw}E\left[P_3P_2^A\right]E\left[P_1^A\right] - k_{1,rev}E\left[P_1^AP_3^A\right] - k_{3,rev}E\left[P_1^AP_3^A\right] + k_{1,forw}E\left[P_1P_3^A\right]\\
\frac{dE\left[P_2P_1^A\right]}{dt} &= 2k_{2,forw}E\left[P_1^A\right]^2E\left[P_2\right] - k_{2,forw}E\left[P_1^AP_2^A\right]E\left[P_2\right] - 2k_{2,forw}E\left[P_2P_1^A\right]E\left[P_1^A\right] - k_{1,rev}E\left[P_2P_1^A\right] + k_{2,rev}E\left[P_1^AP_2^A\right] + k_{1,forw}E\left[P_1P_2\right]\\
\frac{dE\left[P_2P_2\right]}{dt} &= 4k_{2,forw}E\left[P_1^A\right]E\left[P_2\right]^2 - 4k_{2,forw}E\left[P_2P_1^A\right]E\left[P_2\right] - 2k_{2,forw}E\left[P_2P_2\right]E\left[P_1^A\right] + 2k_{2,rev}E\left[P_2P_2^A\right] + k_{2,rev}E\left[P_2^A\right] + k_{2,forw}E\left[P_2P_1^A\right]\\
\frac{dE\left[P_2P_2^A\right]}{dt} &= -2k_{2,forw}E\left[P_1^A\right]E\left[P_2\right]^2 + 2k_{2,forw}E\left[P_1^A\right]E\left[P_2\right]E\left[P_2^A\right] - k_{2,forw}E\left[P_1^AP_2^A\right]E\left[P_2\right] + 2k_{2,forw}E\left[P_2P_1^A\right]E\left[P_2\right] - k_{2,forw}E\left[P_2P_1^A\right]E\left[P_2^A\right] + k_{2,forw}E\left[P_2P_2\right]E\left[P_1^A\right]\\
&\quad - k_{2,forw}E\left[P_2P_2^A\right]E\left[P_1^A\right] - k_{2,rev}E\left[P_2P_2^A\right] + k_{2,rev}E\left[P_2^AP_2^A\right] - k_{2,rev}E\left[P_2^A\right] - k_{2,forw}E\left[P_2P_1^A\right]\\
\frac{dE\left[P_2P_3\right]}{dt} &= 2k_{2,forw}E\left[P_1^A\right]E\left[P_2\right]E\left[P_3\right] + 2k_{3,forw}E\left[P_2\right]E\left[P_2^A\right]E\left[P_3\right] - k_{2,forw}E\left[P_2P_1^A\right]E\left[P_3\right] - k_{2,forw}E\left[P_2P_3\right]E\left[P_1^A\right] - k_{2,forw}E\left[P_3P_1^A\right]E\left[P_2\right] - k_{3,forw}E\left[P_2P_2^A\right]E\left[P_3\right]\\
&\quad - k_{3,forw}E\left[P_2P_3\right]E\left[P_2^A\right] - k_{3,forw}E\left[P_3P_2^A\right]E\left[P_2\right] + k_{2,rev}E\left[P_3P_2^A\right] + k_{3,rev}E\left[P_2P_3^A\right]\\
\frac{dE\left[P_2P_3^A\right]}{dt} &= 2k_{2,forw}E\left[P_1^A\right]E\left[P_2\right]E\left[P_3^A\right] - 2k_{3,forw}E\left[P_2\right]E\left[P_2^A\right]E\left[P_3\right] - k_{2,forw}E\left[P_1^AP_3^A\right]E\left[P_2\right] - k_{2,forw}E\left[P_2P_1^A\right]E\left[P_3^A\right] - k_{2,forw}E\left[P_2P_3^A\right]E\left[P_1^A\right] + k_{3,forw}E\left[P_2P_2^A\right]E\left[P_3\right]\\
&\quad + k_{3,forw}E\left[P_2P_3\right]E\left[P_2^A\right] + k_{3,forw}E\left[P_3P_2^A\right]E\left[P_2\right] + k_{2,rev}E\left[P_2^AP_3^A\right] - k_{3,rev}E\left[P_2P_3^A\right]
\end{aligned}
\tag{47}
$$

Fig. 31: Moment closure ODEs of the mixed channel (continued on page 154)

$$
\begin{aligned}
\frac{dE\left[P_2^A P_2^A\right]}{dt} &= -4k_{2,forw}E\left[P_1^A\right]E\left[P_2\right]E\left[P_2^A\right] + 2k_{2,forw}E\left[P_1^A P_2^A\right]E\left[P_2\right] + 2k_{2,forw}E\left[P_2P_1^A\right]E\left[P_2^A\right] + 2k_{2,forw}E\left[P_2P_2^A\right]E\left[P_1^A\right] - 2k_{2,rev}E\left[P_2^A P_2^A\right] + k_{2,rev}E\left[P_2^A\right] + k_{2,forw}E\left[P_2P_1^A\right] \\
\frac{dE\left[P_2^A P_3^A\right]}{dt} &= -2k_{2,forw}E\left[P_1^A\right]E\left[P_2\right]E\left[P_3^A\right] - 2k_{3,forw}E\left[P_2^A\right]^2E\left[P_3\right] + k_{2,forw}E\left[P_1^A P_3^A\right]E\left[P_2\right] + k_{2,forw}E\left[P_2P_1^A\right]E\left[P_3^A\right] + k_{2,forw}E\left[P_2P_3^A\right]E\left[P_1^A\right] + k_{3,forw}E\left[P_2^A P_2^A\right]E\left[P_3\right] \\
&\quad + 2k_{3,forw}E\left[P_3P_2^A\right]E\left[P_2^A\right] - k_{2,rev}E\left[P_2^A P_3^A\right] - k_{3,rev}E\left[P_2^A P_3^A\right] \\
\frac{dE\left[P_3P_1^A\right]}{dt} &= 2k_{3,forw}E\left[P_1^A\right]E\left[P_2^A\right]E\left[P_3\right] - k_{3,forw}E\left[P_1^A P_2^A\right]E\left[P_3\right] - k_{3,forw}E\left[P_3P_1^A\right]E\left[P_2^A\right] - k_{3,forw}E\left[P_3P_2^A\right]E\left[P_1^A\right] - k_{1,rev}E\left[P_3P_1^A\right] + k_{3,rev}E\left[P_1^A P_3^A\right] + k_{1,forw}E\left[P_1P_3\right] \\
\frac{dE\left[P_3P_2^A\right]}{dt} &= -2k_{2,forw}E\left[P_1^A\right]E\left[P_2\right]E\left[P_3\right] + 2k_{3,forw}E\left[P_2^A\right]^2E\left[P_3\right] + k_{2,forw}E\left[P_2P_1^A\right]E\left[P_3\right] + k_{2,forw}E\left[P_2P_3\right]E\left[P_1^A\right] + k_{2,forw}E\left[P_3P_1^A\right]E\left[P_2\right] - k_{3,forw}E\left[P_2^A P_2^A\right]E\left[P_3\right] \\
&\quad - 2k_{3,forw}E\left[P_3P_2^A\right]E\left[P_2^A\right] - k_{2,rev}E\left[P_3P_2^A\right] + k_{3,rev}E\left[P_2^A P_3^A\right] \\
\frac{dE\left[P_3P_3\right]}{dt} &= 4k_{3,forw}E\left[P_2^A\right]E\left[P_3\right]^2 - 4k_{3,forw}E\left[P_3P_2^A\right]E\left[P_3\right] - 2k_{3,forw}E\left[P_3P_3\right]E\left[P_2^A\right] + 2k_{3,rev}E\left[P_3P_3^A\right] + k_{3,rev}E\left[P_3^A\right] + k_{3,forw}E\left[P_3P_2^A\right] \\
\frac{dE\left[P_3P_3^A\right]}{dt} &= -2k_{3,forw}E\left[P_2^A\right]E\left[P_3\right]^2 + 2k_{3,forw}E\left[P_2^A\right]E\left[P_3\right]E\left[P_3^A\right] - k_{3,forw}E\left[P_2^A P_3^A\right]E\left[P_3\right] + 2k_{3,forw}E\left[P_3P_2^A\right]E\left[P_3\right] - k_{3,forw}E\left[P_3P_2^A\right]E\left[P_3^A\right] + k_{3,forw}E\left[P_3P_3\right]E\left[P_2^A\right] \\
&\quad - k_{3,forw}E\left[P_3P_3^A\right]E\left[P_2^A\right] - k_{3,rev}E\left[P_3P_3^A\right] + k_{3,rev}E\left[P_3^A P_3^A\right] - k_{3,rev}E\left[P_3^A\right] - k_{3,forw}E\left[P_3P_2^A\right] \\
\frac{dE\left[P_3^A P_3^A\right]}{dt} &= -4k_{3,forw}E\left[P_2^A\right]E\left[P_3\right]E\left[P_3^A\right] + 2k_{3,forw}E\left[P_2^A P_3^A\right]E\left[P_3\right] + 2k_{3,forw}E\left[P_3P_2^A\right]E\left[P_3^A\right] + 2k_{3,forw}E\left[P_3P_3^A\right]E\left[P_2^A\right] - 2k_{3,rev}E\left[P_3^A P_3^A\right] + k_{3,rev}E\left[P_3^A\right] + k_{3,forw}E\left[P_3P_2^A\right] \\
\frac{dE\left[P_1\right]}{dt} &= k_{1,rev}E\left[P_1^A\right] - k_{1,forw}E\left[P_1\right] \\
\frac{dE\left[P_1^A\right]}{dt} &= -k_{1,rev}E\left[P_1^A\right] + k_{1,forw}E\left[P_1\right] \\
\frac{dE\left[P_2\right]}{dt} &= k_{2,rev}E\left[P_2^A\right] - k_{2,forw}E\left[P_2P_1^A\right] \\
\frac{dE\left[P_2^A\right]}{dt} &= -k_{2,rev}E\left[P_2^A\right] + k_{2,forw}E\left[P_2P_1^A\right] \\
\frac{dE\left[P_3\right]}{dt} &= k_{3,rev}E\left[P_3^A\right] - k_{3,forw}E\left[P_3P_2^A\right] \\
\frac{dE\left[P_3^A\right]}{dt} &= -k_{3,rev}E\left[P_3^A\right] + k_{3,forw}E\left[P_3P_2^A\right]
\end{aligned}
\tag{48}
$$

Fig. 31: Moment closure ODEs of the mixed channel (continued from page 153).

B.2. Moments for the insulated channel

As for the mixed channel, the following equation systems (49)-(53) of Fig. 32 are the basis for all simulations of the insulated channel. Since the number of different scaffold configurations for three signaling tiers (and thus the number species) is $N = 2^3 = 8$, the system is comprised of $\frac{N(N+3)}{2} = 44$ ODEs that represent the first and second moments. The number of parameters is equal to the mixed model, i.e. 6 parameters $\{k_{i,forw}, k_{i,rev}\}$, for $i = \{1, 2, 3\}$ where $k_{1,forw}$ is the input strength.

For parametrization, see chapter 5 (Tab. 3 defines the standard models) and the input strength $k_{1,forw}$ as defined in Appendix B.1.

$$
\begin{aligned}
\frac{dE\left[P_{(0,0,0)}P_{(0,0,0)}\right]}{dt} &= 2\cdot k_{1,rev}\cdot E\left[P_{(0,0,0)}P_{(1,0,0)}\right] + k_{1,rev}\cdot E\left[P_{(1,0,0)}\right] + 2\cdot k_{2,rev}\cdot E\left[P_{(0,0,0)}P_{(0,1,0)}\right] + k_{2,rev}\cdot E\left[P_{(0,1,0)}\right] + 2\cdot k_{3,rev}\cdot E\left[P_{(0,0,0)}P_{(0,0,1)}\right] \\
&\quad + k_{3,rev}\cdot E\left[P_{(0,0,1)}\right] - 2\cdot k_{1,forw}\cdot E\left[P_{(0,0,0)}P_{(0,0,0)}\right] + k_{1,forw}\cdot E\left[P_{(0,0,0)}\right] \\
\frac{dE\left[P_{(0,0,0)}P_{(0,0,1)}\right]}{dt} &= k_{1,rev}\cdot E\left[P_{(1,0,0)}P_{(0,0,1)}\right] + k_{1,rev}\cdot E\left[P_{(0,0,0)}P_{(1,0,1)}\right] + k_{2,rev}\cdot E\left[P_{(0,1,0)}P_{(0,0,1)}\right] + k_{2,rev}\cdot E\left[P_{(0,0,0)}P_{(0,1,1)}\right] - k_{3,rev}\cdot E\left[P_{(0,0,0)}P_{(0,0,1)}\right] \\
&\quad + k_{3,rev}\cdot E\left[P_{(0,0,1)}P_{(0,0,1)}\right] - k_{3,rev}\cdot E\left[P_{(0,0,1)}\right] - k_{1,forw}\cdot E\left[P_{(0,0,0)}P_{(0,0,1)}\right] - k_{1,forw}\cdot E\left[P_{(0,0,0)}P_{(0,0,1)}\right] \\
\frac{dE\left[P_{(0,0,0)}P_{(0,1,0)}\right]}{dt} &= k_{1,rev}\cdot E\left[P_{(0,0,0)}P_{(1,1,0)}\right] + k_{1,rev}\cdot E\left[P_{(1,0,0)}P_{(0,1,0)}\right] - k_{2,rev}\cdot E\left[P_{(0,0,0)}P_{(0,1,0)}\right] + k_{2,rev}\cdot E\left[P_{(0,1,0)}P_{(0,1,0)}\right] - k_{2,rev}\cdot E\left[P_{(0,1,0)}\right] \\
&\quad + k_{3,rev}\cdot E\left[P_{(0,1,0)}P_{(0,0,1)}\right] + k_{3,rev}\cdot E\left[P_{(0,0,0)}P_{(0,1,1)}\right] - k_{1,forw}\cdot E\left[P_{(0,0,0)}P_{(0,1,0)}\right] - k_{1,forw}\cdot E\left[P_{(0,0,0)}P_{(0,1,0)}\right] - k_{3,forw}\cdot E\left[P_{(0,0,0)}P_{(0,1,0)}\right] \\
\frac{dE\left[P_{(0,0,0)}P_{(0,1,1)}\right]}{dt} &= k_{1,rev}\cdot E\left[P_{(1,0,0)}P_{(0,1,1)}\right] + k_{1,rev}\cdot E\left[P_{(0,0,0)}P_{(1,1,1)}\right] + k_{2,rev}\cdot E\left[P_{(0,1,0)}P_{(0,1,1)}\right] - k_{2,rev}\cdot E\left[P_{(0,0,0)}P_{(0,1,1)}\right] + k_{3,rev}\cdot E\left[P_{(0,0,1)}P_{(0,1,1)}\right] \\
&\quad - k_{3,rev}\cdot E\left[P_{(0,0,0)}P_{(0,1,1)}\right] - k_{1,forw}\cdot E\left[P_{(0,0,0)}P_{(0,1,1)}\right] - k_{1,forw}\cdot E\left[P_{(0,0,0)}P_{(0,1,1)}\right] + k_{3,forw}\cdot E\left[P_{(0,0,0)}P_{(0,1,0)}\right] \\
\frac{dE\left[P_{(0,0,0)}P_{(1,0,0)}\right]}{dt} &= -k_{1,rev}\cdot E\left[P_{(0,0,0)}P_{(1,0,0)}\right] + k_{1,rev}\cdot E\left[P_{(1,0,0)}P_{(1,0,0)}\right] - k_{1,rev}\cdot E\left[P_{(1,0,0)}\right] + k_{2,rev}\cdot E\left[P_{(0,0,0)}P_{(1,1,0)}\right] + k_{2,rev}\cdot E\left[P_{(1,0,0)}P_{(0,1,0)}\right] \\
&\quad + k_{3,rev}\cdot E\left[P_{(1,0,0)}P_{(0,0,1)}\right] + k_{3,rev}\cdot E\left[P_{(0,0,0)}P_{(1,0,1)}\right] + k_{1,forw}\cdot E\left[P_{(0,0,0)}P_{(0,0,0)}\right] - k_{1,forw}\cdot E\left[P_{(0,0,0)}P_{(1,0,0)}\right] \\
&\quad - k_{1,forw}\cdot E\left[P_{(0,0,0)}\right] - k_{2,forw}\cdot E\left[P_{(0,0,0)}P_{(1,0,0)}\right] \\
\frac{dE\left[P_{(0,0,0)}P_{(1,0,1)}\right]}{dt} &= k_{1,rev}\cdot E\left[P_{(1,0,0)}P_{(1,0,1)}\right] - k_{1,rev}\cdot E\left[P_{(0,0,0)}P_{(1,0,1)}\right] + k_{2,rev}\cdot E\left[P_{(0,1,0)}P_{(1,0,1)}\right] + k_{2,rev}\cdot E\left[P_{(0,0,0)}P_{(1,1,1)}\right] + k_{3,rev}\cdot E\left[P_{(0,0,1)}P_{(1,0,1)}\right] \\
&\quad - k_{3,rev}\cdot E\left[P_{(0,0,0)}P_{(1,0,1)}\right] - k_{1,forw}\cdot E\left[P_{(0,0,0)}P_{(1,0,1)}\right] + k_{1,forw}\cdot E\left[P_{(0,0,0)}P_{(0,0,1)}\right] - k_{2,forw}\cdot E\left[P_{(0,0,0)}P_{(1,0,1)}\right] \\
\frac{dE\left[P_{(0,0,0)}P_{(1,1,0)}\right]}{dt} &= -k_{1,rev}\cdot E\left[P_{(0,0,0)}P_{(1,1,0)}\right] + k_{1,rev}\cdot E\left[P_{(1,0,0)}P_{(1,1,0)}\right] - k_{2,rev}\cdot E\left[P_{(0,0,0)}P_{(1,1,0)}\right] + k_{2,rev}\cdot E\left[P_{(0,1,0)}P_{(1,1,0)}\right] + k_{3,rev}\cdot E\left[P_{(0,0,1)}P_{(1,1,0)}\right] \\
&\quad + k_{3,rev}\cdot E\left[P_{(0,0,0)}P_{(1,1,1)}\right] - k_{1,forw}\cdot E\left[P_{(0,0,0)}P_{(1,1,0)}\right] + k_{1,forw}\cdot E\left[P_{(0,0,0)}P_{(0,1,0)}\right] + k_{2,forw}\cdot E\left[P_{(0,0,0)}P_{(1,0,0)}\right] - k_{3,forw}\cdot E\left[P_{(0,0,0)}P_{(1,1,0)}\right] \\
\frac{dE\left[P_{(0,0,0)}P_{(1,1,1)}\right]}{dt} &= k_{1,rev}\cdot E\left[P_{(1,0,0)}P_{(1,1,1)}\right] - k_{1,rev}\cdot E\left[P_{(0,0,0)}P_{(1,1,1)}\right] + k_{2,rev}\cdot E\left[P_{(0,1,0)}P_{(1,1,1)}\right] - k_{2,rev}\cdot E\left[P_{(0,0,0)}P_{(1,1,1)}\right] + k_{3,rev}\cdot E\left[P_{(0,0,1)}P_{(1,1,1)}\right] \\
&\quad - k_{3,rev}\cdot E\left[P_{(0,0,0)}P_{(1,1,1)}\right] - k_{1,forw}\cdot E\left[P_{(0,0,0)}P_{(1,1,1)}\right] + k_{1,forw}\cdot E\left[P_{(0,0,0)}P_{(0,1,1)}\right] + k_{2,forw}\cdot E\left[P_{(0,0,0)}P_{(1,0,1)}\right] + k_{3,forw}\cdot E\left[P_{(0,0,0)}P_{(1,1,0)}\right] \\
\frac{dE\left[P_{(0,0,1)}P_{(0,0,1)}\right]}{dt} &= 2\cdot k_{1,rev}\cdot E\left[P_{(0,0,1)}P_{(1,0,1)}\right] + k_{1,rev}\cdot E\left[P_{(1,0,1)}\right] + 2\cdot k_{2,rev}\cdot E\left[P_{(0,0,1)}P_{(0,1,1)}\right] + k_{2,rev}\cdot E\left[P_{(0,1,1)}\right] - 2\cdot k_{3,rev}\cdot E\left[P_{(0,0,1)}P_{(0,0,1)}\right] \\
&\quad + k_{3,rev}\cdot E\left[P_{(0,0,1)}\right] - 2\cdot k_{1,forw}\cdot E\left[P_{(0,0,1)}P_{(0,0,1)}\right] + k_{1,forw}\cdot E\left[P_{(0,0,1)}\right] \\
\frac{dE\left[P_{(0,0,1)}P_{(0,1,1)}\right]}{dt} &= k_{1,rev}\cdot E\left[P_{(0,0,1)}P_{(1,1,1)}\right] + k_{1,rev}\cdot E\left[P_{(1,0,1)}P_{(0,1,1)}\right] - k_{2,rev}\cdot E\left[P_{(0,0,1)}P_{(0,1,1)}\right] + k_{2,rev}\cdot E\left[P_{(0,1,1)}P_{(0,1,1)}\right] - k_{2,rev}\cdot E\left[P_{(0,1,1)}\right] \\
&\quad - k_{3,rev}\cdot E\left[P_{(0,0,1)}P_{(0,1,1)}\right] - k_{3,rev}\cdot E\left[P_{(0,0,1)}P_{(0,1,1)}\right] - k_{1,forw}\cdot E\left[P_{(0,0,1)}P_{(0,1,1)}\right] - k_{1,forw}\cdot E\left[P_{(0,0,1)}P_{(0,1,1)}\right] + k_{3,forw}\cdot E\left[P_{(0,1,0)}P_{(0,0,1)}\right]
\end{aligned}
\tag{49}
$$

Fig. 32: Moment closure ODEs of the insulated channel (continued on page 157).

$$
\begin{aligned}
\frac{dE\left[P_{(0,0,1)}P_{(1,0,1)}\right]}{dt} &= -k_{1,rev}\cdot E\left[P_{(0,0,1)}P_{(1,0,1)}\right] + k_{1,rev}\cdot E\left[P_{(1,0,1)}P_{(1,0,1)}\right] - k_{1,rev}\cdot E\left[P_{(1,0,1)}\right] + k_{2,rev}\cdot E\left[P_{(0,0,1)}P_{(1,1,1)}\right] + k_{2,rev}\cdot E\left[P_{(1,0,1)}P_{(0,1,1)}\right] \\
&\quad - k_{3,rev}\cdot E\left[P_{(0,0,1)}P_{(1,0,1)}\right] - k_{3,rev}\cdot E\left[P_{(0,0,1)}P_{(1,0,1)}\right] + k_{1,forw}\cdot E\left[P_{(0,0,1)}P_{(0,0,1)}\right] - k_{1,forw}\cdot E\left[P_{(0,0,1)}P_{(1,0,1)}\right] \\
&\quad - k_{1,forw}\cdot E\left[P_{(0,0,1)}\right] - k_{2,forw}\cdot E\left[P_{(0,0,1)}P_{(1,0,1)}\right] \\
\frac{dE\left[P_{(0,0,1)}P_{(1,1,0)}\right]}{dt} &= -k_{1,rev}\cdot E\left[P_{(0,0,1)}P_{(1,1,0)}\right] + k_{1,rev}\cdot E\left[P_{(1,1,0)}P_{(1,0,1)}\right] - k_{2,rev}\cdot E\left[P_{(0,0,1)}P_{(1,1,0)}\right] + k_{2,rev}\cdot E\left[P_{(1,1,0)}P_{(0,1,1)}\right] - k_{3,rev}\cdot E\left[P_{(0,0,1)}P_{(1,1,0)}\right] \\
&\quad + k_{3,rev}\cdot E\left[P_{(0,0,1)}P_{(1,1,1)}\right] + k_{1,forw}\cdot E\left[P_{(0,1,0)}P_{(0,0,1)}\right] - k_{1,forw}\cdot E\left[P_{(0,0,1)}P_{(1,1,0)}\right] + k_{2,forw}\cdot E\left[P_{(1,0,0)}P_{(0,0,1)}\right] - k_{3,forw}\cdot E\left[P_{(0,0,1)}P_{(1,1,0)}\right] \\
\frac{dE\left[P_{(0,0,1)}P_{(1,1,1)}\right]}{dt} &= -k_{1,rev}\cdot E\left[P_{(0,0,1)}P_{(1,1,1)}\right] + k_{1,rev}\cdot E\left[P_{(1,0,1)}P_{(1,1,1)}\right] - k_{2,rev}\cdot E\left[P_{(0,0,1)}P_{(1,1,1)}\right] + k_{2,rev}\cdot E\left[P_{(0,1,1)}P_{(1,1,1)}\right] - k_{3,rev}\cdot E\left[P_{(0,0,1)}P_{(1,1,1)}\right] \\
&\quad - k_{3,rev}\cdot E\left[P_{(0,0,1)}P_{(1,1,1)}\right] - k_{1,forw}\cdot E\left[P_{(0,0,1)}P_{(1,1,1)}\right] + k_{1,forw}\cdot E\left[P_{(0,0,1)}P_{(0,1,1)}\right] + k_{2,forw}\cdot E\left[P_{(0,0,1)}P_{(1,0,1)}\right] + k_{3,forw}\cdot E\left[P_{(0,0,1)}P_{(1,1,0)}\right] \\
\frac{dE\left[P_{(0,1,0)}P_{(0,0,1)}\right]}{dt} &= k_{1,rev}\cdot E\left[P_{(0,0,1)}P_{(1,1,0)}\right] + k_{1,rev}\cdot E\left[P_{(0,1,0)}P_{(1,0,1)}\right] - k_{2,rev}\cdot E\left[P_{(0,1,0)}P_{(0,0,1)}\right] + k_{2,rev}\cdot E\left[P_{(0,1,0)}P_{(0,1,1)}\right] - k_{3,rev}\cdot E\left[P_{(0,1,0)}P_{(0,0,1)}\right] \\
&\quad + k_{3,rev}\cdot E\left[P_{(0,0,1)}P_{(0,1,1)}\right] - k_{1,forw}\cdot E\left[P_{(0,1,0)}P_{(0,0,1)}\right] - k_{1,forw}\cdot E\left[P_{(0,1,0)}P_{(0,0,1)}\right] - k_{3,forw}\cdot E\left[P_{(0,1,0)}P_{(0,0,1)}\right] \\
\frac{dE\left[P_{(0,1,0)}P_{(0,1,0)}\right]}{dt} &= 2\cdot k_{1,rev}\cdot E\left[P_{(0,1,0)}P_{(1,1,0)}\right] + k_{1,rev}\cdot E\left[P_{(1,1,0)}\right] - 2\cdot k_{2,rev}\cdot E\left[P_{(0,1,0)}P_{(0,1,0)}\right] + k_{2,rev}\cdot E\left[P_{(0,1,0)}\right] + 2\cdot k_{3,rev}\cdot E\left[P_{(0,1,0)}P_{(0,1,1)}\right] \\
&\quad + k_{3,rev}\cdot E\left[P_{(0,1,1)}\right] - 2\cdot k_{1,forw}\cdot E\left[P_{(0,1,0)}P_{(0,1,0)}\right] + k_{1,forw}\cdot E\left[P_{(0,1,0)}\right] - 2\cdot k_{3,forw}\cdot E\left[P_{(0,1,0)}P_{(0,1,0)}\right] + k_{3,forw}\cdot E\left[P_{(0,1,0)}\right] \\
\frac{dE\left[P_{(0,1,0)}P_{(0,1,1)}\right]}{dt} &= k_{1,rev}\cdot E\left[P_{(1,1,0)}P_{(0,1,1)}\right] + k_{1,rev}\cdot E\left[P_{(0,1,0)}P_{(1,1,1)}\right] - k_{2,rev}\cdot E\left[P_{(0,1,0)}P_{(0,1,1)}\right] - k_{2,rev}\cdot E\left[P_{(0,1,0)}P_{(0,1,1)}\right] - k_{3,rev}\cdot E\left[P_{(0,1,0)}P_{(0,1,1)}\right] \\
&\quad + k_{3,rev}\cdot E\left[P_{(0,1,1)}P_{(0,1,1)}\right] - k_{3,rev}\cdot E\left[P_{(0,1,1)}\right] - k_{1,forw}\cdot E\left[P_{(0,1,0)}P_{(0,1,1)}\right] - k_{1,forw}\cdot E\left[P_{(0,1,0)}P_{(0,1,1)}\right] + k_{3,forw}\cdot E\left[P_{(0,1,0)}P_{(0,1,0)}\right] \\
&\quad - k_{3,forw}\cdot E\left[P_{(0,1,0)}P_{(0,1,1)}\right] - k_{3,forw}\cdot E\left[P_{(0,1,0)}\right] \\
\frac{dE\left[P_{(0,1,0)}P_{(1,0,1)}\right]}{dt} &= k_{1,rev}\cdot E\left[P_{(1,1,0)}P_{(1,0,1)}\right] - k_{1,rev}\cdot E\left[P_{(0,1,0)}P_{(1,0,1)}\right] - k_{2,rev}\cdot E\left[P_{(0,1,0)}P_{(1,0,1)}\right] + k_{2,rev}\cdot E\left[P_{(0,1,0)}P_{(1,1,1)}\right] + k_{3,rev}\cdot E\left[P_{(1,0,1)}P_{(0,1,1)}\right] \\
&\quad - k_{3,rev}\cdot E\left[P_{(0,1,0)}P_{(1,0,1)}\right] - k_{1,forw}\cdot E\left[P_{(0,1,0)}P_{(1,0,1)}\right] + k_{1,forw}\cdot E\left[P_{(0,1,0)}P_{(0,0,1)}\right] - k_{2,forw}\cdot E\left[P_{(0,1,0)}P_{(1,0,1)}\right] - k_{3,forw}\cdot E\left[P_{(0,1,0)}P_{(1,0,1)}\right] \\
\frac{dE\left[P_{(0,1,0)}P_{(1,1,0)}\right]}{dt} &= -k_{1,rev}\cdot E\left[P_{(0,1,0)}P_{(1,1,0)}\right] + k_{1,rev}\cdot E\left[P_{(1,1,0)}P_{(1,1,0)}\right] - k_{1,rev}\cdot E\left[P_{(1,1,0)}\right] - k_{2,rev}\cdot E\left[P_{(0,1,0)}P_{(1,1,0)}\right] - k_{2,rev}\cdot E\left[P_{(0,1,0)}P_{(1,1,0)}\right] \\
&\quad + k_{3,rev}\cdot E\left[P_{(0,1,0)}P_{(1,1,1)}\right] + k_{3,rev}\cdot E\left[P_{(1,1,0)}P_{(0,1,1)}\right] + k_{1,forw}\cdot E\left[P_{(0,1,0)}P_{(0,1,0)}\right] - k_{1,forw}\cdot E\left[P_{(0,1,0)}P_{(1,1,0)}\right] - k_{1,forw}\cdot E\left[P_{(0,1,0)}\right] \\
&\quad + k_{2,forw}\cdot E\left[P_{(1,0,0)}P_{(0,1,0)}\right] - k_{3,forw}\cdot E\left[P_{(0,1,0)}P_{(1,1,0)}\right] - k_{3,forw}\cdot E\left[P_{(0,1,0)}P_{(1,1,0)}\right] \\
\frac{dE\left[P_{(0,1,0)}P_{(1,1,1)}\right]}{dt} &= k_{1,rev}\cdot E\left[P_{(1,1,0)}P_{(1,1,1)}\right] - k_{1,rev}\cdot E\left[P_{(0,1,0)}P_{(1,1,1)}\right] - k_{2,rev}\cdot E\left[P_{(0,1,0)}P_{(1,1,1)}\right] - k_{2,rev}\cdot E\left[P_{(0,1,0)}P_{(1,1,1)}\right] - k_{3,rev}\cdot E\left[P_{(0,1,0)}P_{(1,1,1)}\right] \\
&\quad + k_{3,rev}\cdot E\left[P_{(0,1,1)}P_{(1,1,1)}\right] - k_{1,forw}\cdot E\left[P_{(0,1,0)}P_{(1,1,1)}\right] + k_{1,forw}\cdot E\left[P_{(0,1,0)}P_{(0,1,1)}\right] + k_{2,forw}\cdot E\left[P_{(0,1,0)}P_{(1,0,1)}\right] \\
&\quad - k_{3,forw}\cdot E\left[P_{(0,1,0)}P_{(1,1,1)}\right] + k_{3,forw}\cdot E\left[P_{(0,1,0)}P_{(1,1,0)}\right]
\end{aligned}
\tag{50}
$$

Fig. 32: Moment closure ODEs of the insulated channel (continued from page 156).

$$
\begin{aligned}
\frac{dE\left[P_{(0,1,1)}P_{(0,1,1)}\right]}{dt} &= 2\cdot k_{1,rev}\cdot E\left[P_{(0,1,1)}P_{(1,1,1)}\right] + k_{1,rev}\cdot E\left[P_{(1,1,1)}\right] - 2\cdot k_{2,rev}\cdot E\left[P_{(0,1,1)}P_{(0,1,1)}\right] + k_{2,rev}\cdot E\left[P_{(0,1,1)}\right] - 2\cdot k_{3,rev}\cdot E\left[P_{(0,1,1)}P_{(0,1,1)}\right] \\
&\quad + k_{3,rev}\cdot E\left[P_{(0,1,1)}\right] - 2\cdot k_{1,forw}\cdot E\left[P_{(0,1,1)}P_{(0,1,1)}\right] + k_{1,forw}\cdot E\left[P_{(0,1,1)}\right] + 2\cdot k_{3,forw}\cdot E\left[P_{(0,1,0)}P_{(0,1,1)}\right] + k_{3,forw}\cdot E\left[P_{(0,1,0)}\right] \\
\frac{dE\left[P_{(0,1,1)}P_{(1,1,1)}\right]}{dt} &= -k_{1,rev}\cdot E\left[P_{(0,1,1)}P_{(1,1,1)}\right] + k_{1,rev}\cdot E\left[P_{(1,1,1)}P_{(1,1,1)}\right] - k_{1,rev}\cdot E\left[P_{(1,1,1)}\right] - k_{2,rev}\cdot E\left[P_{(0,1,1)}P_{(1,1,1)}\right] - k_{2,rev}\cdot E\left[P_{(0,1,1)}P_{(1,1,1)}\right] \\
&\quad - k_{3,rev}\cdot E\left[P_{(0,1,1)}P_{(1,1,1)}\right] - k_{3,rev}\cdot E\left[P_{(0,1,1)}P_{(1,1,1)}\right] + k_{1,forw}\cdot E\left[P_{(0,1,1)}P_{(0,1,1)}\right] - k_{1,forw}\cdot E\left[P_{(0,1,1)}P_{(1,1,1)}\right] - k_{1,forw}\cdot E\left[P_{(0,1,1)}\right] \\
&\quad + k_{2,forw}\cdot E\left[P_{(1,0,1)}P_{(0,1,1)}\right] + k_{3,forw}\cdot E\left[P_{(0,1,0)}P_{(1,1,1)}\right] + k_{3,forw}\cdot E\left[P_{(1,1,0)}P_{(0,1,1)}\right] \\
\frac{dE\left[P_{(1,0,0)}P_{(0,0,1)}\right]}{dt} &= -k_{1,rev}\cdot E\left[P_{(1,0,0)}P_{(0,0,1)}\right] + k_{1,rev}\cdot E\left[P_{(1,0,0)}P_{(1,0,1)}\right] + k_{2,rev}\cdot E\left[P_{(0,0,1)}P_{(1,1,0)}\right] + k_{2,rev}\cdot E\left[P_{(1,0,0)}P_{(0,1,1)}\right] - k_{3,rev}\cdot E\left[P_{(1,0,0)}P_{(0,0,1)}\right] \\
&\quad + k_{3,rev}\cdot E\left[P_{(0,0,1)}P_{(1,0,1)}\right] + k_{1,forw}\cdot E\left[P_{(0,0,0)}P_{(0,0,1)}\right] - k_{1,forw}\cdot E\left[P_{(1,0,0)}P_{(0,0,1)}\right] - k_{2,forw}\cdot E\left[P_{(1,0,0)}P_{(0,0,1)}\right] \\
\frac{dE\left[P_{(1,0,0)}P_{(0,1,0)}\right]}{dt} &= k_{1,rev}\cdot E\left[P_{(1,0,0)}P_{(1,1,0)}\right] - k_{1,rev}\cdot E\left[P_{(1,0,0)}P_{(0,1,0)}\right] + k_{2,rev}\cdot E\left[P_{(0,1,0)}P_{(1,1,0)}\right] - k_{2,rev}\cdot E\left[P_{(1,0,0)}P_{(0,1,0)}\right] + k_{3,rev}\cdot E\left[P_{(1,0,0)}P_{(0,1,1)}\right] \\
&\quad + k_{3,rev}\cdot E\left[P_{(0,1,0)}P_{(1,0,1)}\right] + k_{1,forw}\cdot E\left[P_{(0,0,0)}P_{(0,1,0)}\right] - k_{1,forw}\cdot E\left[P_{(1,0,0)}P_{(0,1,0)}\right] - k_{2,forw}\cdot E\left[P_{(1,0,0)}P_{(0,1,0)}\right] - k_{3,forw}\cdot E\left[P_{(1,0,0)}P_{(0,1,0)}\right] \\
\frac{dE\left[P_{(1,0,0)}P_{(0,1,1)}\right]}{dt} &= -k_{1,rev}\cdot E\left[P_{(1,0,0)}P_{(0,1,1)}\right] + k_{1,rev}\cdot E\left[P_{(1,0,0)}P_{(1,1,1)}\right] + k_{2,rev}\cdot E\left[P_{(1,1,0)}P_{(0,1,1)}\right] - k_{2,rev}\cdot E\left[P_{(1,0,0)}P_{(0,1,1)}\right] - k_{3,rev}\cdot E\left[P_{(1,0,0)}P_{(0,1,1)}\right] \\
&\quad + k_{3,rev}\cdot E\left[P_{(1,0,1)}P_{(0,1,1)}\right] + k_{1,forw}\cdot E\left[P_{(0,0,0)}P_{(0,1,1)}\right] - k_{1,forw}\cdot E\left[P_{(1,0,0)}P_{(0,1,1)}\right] - k_{2,forw}\cdot E\left[P_{(1,0,0)}P_{(0,1,1)}\right] + k_{3,forw}\cdot E\left[P_{(1,0,0)}P_{(0,1,0)}\right] \\
\frac{dE\left[P_{(1,0,0)}P_{(1,0,0)}\right]}{dt} &= -2\cdot k_{1,rev}\cdot E\left[P_{(1,0,0)}P_{(1,0,0)}\right] + k_{1,rev}\cdot E\left[P_{(1,0,0)}\right] + 2\cdot k_{2,rev}\cdot E\left[P_{(1,0,0)}P_{(1,1,0)}\right] + k_{2,rev}\cdot E\left[P_{(1,1,0)}\right] + 2\cdot k_{3,rev}\cdot E\left[P_{(1,0,0)}P_{(1,0,1)}\right] \\
&\quad + k_{3,rev}\cdot E\left[P_{(1,0,1)}\right] + 2\cdot k_{1,forw}\cdot E\left[P_{(0,0,0)}P_{(1,0,0)}\right] + k_{1,forw}\cdot E\left[P_{(0,0,0)}\right] - 2\cdot k_{2,forw}\cdot E\left[P_{(1,0,0)}P_{(1,0,0)}\right] + k_{2,forw}\cdot E\left[P_{(1,0,0)}\right] \\
\frac{dE\left[P_{(1,0,0)}P_{(1,0,1)}\right]}{dt} &= -k_{1,rev}\cdot E\left[P_{(1,0,0)}P_{(1,0,1)}\right] - k_{1,rev}\cdot E\left[P_{(1,0,0)}P_{(1,0,1)}\right] + k_{2,rev}\cdot E\left[P_{(1,1,0)}P_{(1,0,1)}\right] + k_{2,rev}\cdot E\left[P_{(1,0,0)}P_{(1,1,1)}\right] - k_{3,rev}\cdot E\left[P_{(1,0,0)}P_{(1,0,1)}\right] \\
&\quad + k_{3,rev}\cdot E\left[P_{(1,0,1)}P_{(1,0,1)}\right] - k_{3,rev}\cdot E\left[P_{(1,0,1)}\right] + k_{1,forw}\cdot E\left[P_{(0,0,0)}P_{(1,0,1)}\right] + k_{1,forw}\cdot E\left[P_{(1,0,0)}P_{(0,0,1)}\right] \\
&\quad - k_{2,forw}\cdot E\left[P_{(1,0,0)}P_{(1,0,1)}\right] - k_{2,forw}\cdot E\left[P_{(1,0,0)}P_{(1,0,1)}\right] \\
\frac{dE\left[P_{(1,0,0)}P_{(1,1,0)}\right]}{dt} &= -k_{1,rev}\cdot E\left[P_{(1,0,0)}P_{(1,1,0)}\right] - k_{1,rev}\cdot E\left[P_{(1,0,0)}P_{(1,1,0)}\right] - k_{2,rev}\cdot E\left[P_{(1,0,0)}P_{(1,1,0)}\right] + k_{2,rev}\cdot E\left[P_{(1,1,0)}P_{(1,1,0)}\right] - k_{2,rev}\cdot E\left[P_{(1,1,0)}\right] \\
&\quad + k_{3,rev}\cdot E\left[P_{(1,0,0)}P_{(1,1,1)}\right] + k_{3,rev}\cdot E\left[P_{(1,1,0)}P_{(0,1,1)}\right] + k_{1,forw}\cdot E\left[P_{(0,0,0)}P_{(1,1,0)}\right] + k_{1,forw}\cdot E\left[P_{(1,0,0)}P_{(0,1,0)}\right] + k_{2,forw}\cdot E\left[P_{(1,0,0)}P_{(1,0,0)}\right] \\
&\quad - k_{2,forw}\cdot E\left[P_{(1,0,0)}P_{(1,1,0)}\right] - k_{2,forw}\cdot E\left[P_{(1,0,0)}\right] - k_{3,forw}\cdot E\left[P_{(1,0,0)}P_{(1,1,0)}\right] \\
\frac{dE\left[P_{(1,0,0)}P_{(1,1,1)}\right]}{dt} &= -k_{1,rev}\cdot E\left[P_{(1,0,0)}P_{(1,1,1)}\right] - k_{1,rev}\cdot E\left[P_{(1,0,0)}P_{(1,1,1)}\right] + k_{2,rev}\cdot E\left[P_{(1,1,0)}P_{(1,1,1)}\right] - k_{2,rev}\cdot E\left[P_{(1,0,0)}P_{(1,1,1)}\right] - k_{3,rev}\cdot E\left[P_{(1,0,0)}P_{(1,1,1)}\right] \\
&\quad + k_{3,rev}\cdot E\left[P_{(1,0,1)}P_{(1,1,1)}\right] + k_{1,forw}\cdot E\left[P_{(0,0,0)}P_{(1,1,1)}\right] + k_{1,forw}\cdot E\left[P_{(1,0,0)}P_{(0,1,1)}\right] - k_{2,forw}\cdot E\left[P_{(1,0,0)}P_{(1,1,1)}\right] \\
&\quad + k_{2,forw}\cdot E\left[P_{(1,0,0)}P_{(1,0,1)}\right] + k_{3,forw}\cdot E\left[P_{(1,0,0)}P_{(1,1,0)}\right]
\end{aligned}
\tag{51}
$$

Fig. 32: Moment closure ODEs of the insulated channel (continued from page 157).

$$
\begin{aligned}
\frac{dE\left[P_{(1,0,1)}P_{(0,1,1)}\right]}{dt} &= k_{1,rev}\cdot E\left[P_{(1,0,1)}P_{(1,1,1)}\right] - k_{1,rev}\cdot E\left[P_{(1,0,1)}P_{(0,1,1)}\right] + k_{2,rev}\cdot E\left[P_{(0,1,1)}P_{(1,1,1)}\right] - k_{2,rev}\cdot E\left[P_{(1,0,1)}P_{(0,1,1)}\right] - k_{3,rev}\cdot E\left[P_{(1,0,1)}P_{(0,1,1)}\right] \\
&\quad - k_{3,rev}\cdot E\left[P_{(1,0,1)}P_{(0,1,1)}\right] + k_{1,forw}\cdot E\left[P_{(0,0,1)}P_{(0,1,1)}\right] - k_{1,forw}\cdot E\left[P_{(1,0,1)}P_{(0,1,1)}\right] - k_{2,forw}\cdot E\left[P_{(1,0,1)}P_{(0,1,1)}\right] + k_{3,forw}\cdot E\left[P_{(0,1,0)}P_{(1,0,1)}\right] \\
\frac{dE\left[P_{(1,0,1)}P_{(1,0,1)}\right]}{dt} &= -2\cdot k_{1,rev}\cdot E\left[P_{(1,0,1)}P_{(1,0,1)}\right] + k_{1,rev}\cdot E\left[P_{(1,0,1)}\right] + 2\cdot k_{2,rev}\cdot E\left[P_{(1,0,1)}P_{(1,1,1)}\right] + k_{2,rev}\cdot E\left[P_{(1,1,1)}\right] - 2\cdot k_{3,rev}\cdot E\left[P_{(1,0,1)}P_{(1,0,1)}\right] \\
&\quad + k_{3,rev}\cdot E\left[P_{(1,0,1)}\right] + 2\cdot k_{1,forw}\cdot E\left[P_{(0,0,1)}P_{(1,0,1)}\right] + k_{1,forw}\cdot E\left[P_{(0,0,1)}\right] - 2\cdot k_{2,forw}\cdot E\left[P_{(1,0,1)}P_{(1,0,1)}\right] + k_{2,forw}\cdot E\left[P_{(1,0,1)}\right] \\
\frac{dE\left[P_{(1,0,1)}P_{(1,1,1)}\right]}{dt} &= -k_{1,rev}\cdot E\left[P_{(1,0,1)}P_{(1,1,1)}\right] - k_{1,rev}\cdot E\left[P_{(1,0,1)}P_{(1,1,1)}\right] - k_{2,rev}\cdot E\left[P_{(1,0,1)}P_{(1,1,1)}\right] + k_{2,rev}\cdot E\left[P_{(1,1,1)}P_{(1,1,1)}\right] - k_{2,rev}\cdot E\left[P_{(1,1,1)}\right] \\
&\quad - k_{3,rev}\cdot E\left[P_{(1,0,1)}P_{(1,1,1)}\right] - k_{3,rev}\cdot E\left[P_{(1,0,1)}P_{(1,1,1)}\right] + k_{1,forw}\cdot E\left[P_{(0,0,1)}P_{(1,1,1)}\right] + k_{1,forw}\cdot E\left[P_{(1,0,1)}P_{(0,1,1)}\right] \\
&\quad + k_{2,forw}\cdot E\left[P_{(1,0,1)}P_{(1,0,1)}\right] - k_{2,forw}\cdot E\left[P_{(1,0,1)}P_{(1,1,1)}\right] - k_{2,forw}\cdot E\left[P_{(1,0,1)}\right] + k_{3,forw}\cdot E\left[P_{(1,1,0)}P_{(1,0,1)}\right] \\
\frac{dE\left[P_{(1,1,0)}P_{(0,1,1)}\right]}{dt} &= -k_{1,rev}\cdot E\left[P_{(1,1,0)}P_{(0,1,1)}\right] + k_{1,rev}\cdot E\left[P_{(1,1,0)}P_{(1,1,1)}\right] - k_{2,rev}\cdot E\left[P_{(1,1,0)}P_{(0,1,1)}\right] - k_{2,rev}\cdot E\left[P_{(1,1,0)}P_{(0,1,1)}\right] + k_{3,rev}\cdot E\left[P_{(0,1,1)}P_{(1,1,1)}\right] \\
&\quad - k_{3,rev}\cdot E\left[P_{(1,1,0)}P_{(0,1,1)}\right] + k_{1,forw}\cdot E\left[P_{(0,1,0)}P_{(0,1,1)}\right] - k_{1,forw}\cdot E\left[P_{(1,1,0)}P_{(0,1,1)}\right] + k_{2,forw}\cdot E\left[P_{(1,0,0)}P_{(0,1,1)}\right] \\
&\quad + k_{3,forw}\cdot E\left[P_{(0,1,0)}P_{(1,1,0)}\right] - k_{3,forw}\cdot E\left[P_{(1,1,0)}P_{(0,1,1)}\right] \\
\frac{dE\left[P_{(1,1,0)}P_{(1,0,1)}\right]}{dt} &= -k_{1,rev}\cdot E\left[P_{(1,1,0)}P_{(1,0,1)}\right] - k_{1,rev}\cdot E\left[P_{(1,1,0)}P_{(1,0,1)}\right] - k_{2,rev}\cdot E\left[P_{(1,1,0)}P_{(1,0,1)}\right] + k_{2,rev}\cdot E\left[P_{(1,1,0)}P_{(1,1,1)}\right] + k_{3,rev}\cdot E\left[P_{(1,0,1)}P_{(1,1,1)}\right] \\
&\quad - k_{3,rev}\cdot E\left[P_{(1,1,0)}P_{(1,0,1)}\right] + k_{1,forw}\cdot E\left[P_{(0,1,0)}P_{(1,0,1)}\right] + k_{1,forw}\cdot E\left[P_{(0,0,1)}P_{(1,1,0)}\right] + k_{2,forw}\cdot E\left[P_{(1,0,0)}P_{(1,0,1)}\right] \\
&\quad - k_{2,forw}\cdot E\left[P_{(1,1,0)}P_{(1,0,1)}\right] - k_{3,forw}\cdot E\left[P_{(1,1,0)}P_{(1,0,1)}\right] \\
\frac{dE\left[P_{(1,1,0)}P_{(1,1,0)}\right]}{dt} &= -2\cdot k_{1,rev}\cdot E\left[P_{(1,1,0)}P_{(1,1,0)}\right] + k_{1,rev}\cdot E\left[P_{(1,1,0)}\right] - 2\cdot k_{2,rev}\cdot E\left[P_{(1,1,0)}P_{(1,1,0)}\right] + k_{2,rev}\cdot E\left[P_{(1,1,0)}\right] + 2\cdot k_{3,rev}\cdot E\left[P_{(1,1,0)}P_{(1,1,1)}\right] \\
&\quad + k_{3,rev}\cdot E\left[P_{(1,1,1)}\right] + 2\cdot k_{1,forw}\cdot E\left[P_{(0,1,0)}P_{(1,1,0)}\right] + k_{1,forw}\cdot E\left[P_{(0,1,0)}\right] + 2\cdot k_{2,forw}\cdot E\left[P_{(1,0,0)}P_{(1,1,0)}\right] \\
&\quad + k_{2,forw}\cdot E\left[P_{(1,0,0)}\right] - 2\cdot k_{3,forw}\cdot E\left[P_{(1,1,0)}P_{(1,1,0)}\right] + k_{3,forw}\cdot E\left[P_{(1,1,0)}\right] \\
\frac{dE\left[P_{(1,1,0)}P_{(1,1,1)}\right]}{dt} &= -k_{1,rev}\cdot E\left[P_{(1,1,0)}P_{(1,1,1)}\right] - k_{1,rev}\cdot E\left[P_{(1,1,0)}P_{(1,1,1)}\right] - k_{2,rev}\cdot E\left[P_{(1,1,0)}P_{(1,1,1)}\right] - k_{2,rev}\cdot E\left[P_{(1,1,0)}P_{(1,1,1)}\right] - k_{3,rev}\cdot E\left[P_{(1,1,0)}P_{(1,1,1)}\right] \\
&\quad + k_{3,rev}\cdot E\left[P_{(1,1,1)}P_{(1,1,1)}\right] - k_{3,rev}\cdot E\left[P_{(1,1,1)}\right] + k_{1,forw}\cdot E\left[P_{(0,1,0)}P_{(1,1,1)}\right] + k_{1,forw}\cdot E\left[P_{(1,1,0)}P_{(0,1,1)}\right] + k_{2,forw}\cdot E\left[P_{(1,0,0)}P_{(1,1,1)}\right] \\
&\quad + k_{2,forw}\cdot E\left[P_{(1,1,0)}P_{(1,0,1)}\right] + k_{3,forw}\cdot E\left[P_{(1,1,0)}P_{(1,1,0)}\right] - k_{3,forw}\cdot E\left[P_{(1,1,0)}P_{(1,1,1)}\right] - k_{3,forw}\cdot E\left[P_{(1,1,0)}\right] \\
\frac{dE\left[P_{(1,1,1)}P_{(1,1,1)}\right]}{dt} &= -2\cdot k_{1,rev}\cdot E\left[P_{(1,1,1)}P_{(1,1,1)}\right] + k_{1,rev}\cdot E\left[P_{(1,1,1)}\right] - 2\cdot k_{2,rev}\cdot E\left[P_{(1,1,1)}P_{(1,1,1)}\right] + k_{2,rev}\cdot E\left[P_{(1,1,1)}\right] - 2\cdot k_{3,rev}\cdot E\left[P_{(1,1,1)}P_{(1,1,1)}\right] \\
&\quad + k_{3,rev}\cdot E\left[P_{(1,1,1)}\right] + 2\cdot k_{1,forw}\cdot E\left[P_{(0,1,1)}P_{(1,1,1)}\right] + k_{1,forw}\cdot E\left[P_{(0,1,1)}\right] + 2\cdot k_{2,forw}\cdot E\left[P_{(1,0,1)}P_{(1,1,1)}\right] + k_{2,forw}\cdot E\left[P_{(1,0,1)}\right] \\
&\quad + 2\cdot k_{3,forw}\cdot E\left[P_{(1,1,0)}P_{(1,1,1)}\right] + k_{3,forw}\cdot E\left[P_{(1,1,0)}\right]
\end{aligned}
\tag{52}
$$

Fig. 32: Moment closure ODEs of the insulated channel (continued from page 158).

$$
\begin{aligned}
\frac{dE\left[P_{(0,0,0)}\right]}{dt} &= k_{1,rev}\cdot E\left[P_{(1,0,0)}\right] + k_{2,rev}\cdot E\left[P_{(0,1,0)}\right] + k_{3,rev}\cdot E\left[P_{(0,0,1)}\right] - k_{1,forw}\cdot E\left[P_{(0,0,0)}\right] \\
\frac{dE\left[P_{(0,0,1)}\right]}{dt} &= k_{1,rev}\cdot E\left[P_{(1,0,1)}\right] + k_{2,rev}\cdot E\left[P_{(0,1,1)}\right] - k_{3,rev}\cdot E\left[P_{(0,0,1)}\right] - k_{1,forw}\cdot E\left[P_{(0,0,1)}\right] \\
\frac{dE\left[P_{(0,1,0)}\right]}{dt} &= k_{1,rev}\cdot E\left[P_{(1,1,0)}\right] - k_{2,rev}\cdot E\left[P_{(0,1,0)}\right] + k_{3,rev}\cdot E\left[P_{(0,1,1)}\right] - k_{1,forw}\cdot E\left[P_{(0,1,0)}\right] - k_{3,forw}\cdot E\left[P_{(0,1,0)}\right] \\
\frac{dE\left[P_{(0,1,1)}\right]}{dt} &= k_{1,rev}\cdot E\left[P_{(1,1,1)}\right] - k_{2,rev}\cdot E\left[P_{(0,1,1)}\right] - k_{3,rev}\cdot E\left[P_{(0,1,1)}\right] - k_{1,forw}\cdot E\left[P_{(0,1,1)}\right] + k_{3,forw}\cdot E\left[P_{(0,1,0)}\right] \\
\frac{dE\left[P_{(1,0,0)}\right]}{dt} &= -k_{1,rev}\cdot E\left[P_{(1,0,0)}\right] + k_{2,rev}\cdot E\left[P_{(1,1,0)}\right] + k_{3,rev}\cdot E\left[P_{(1,0,1)}\right] + k_{1,forw}\cdot E\left[P_{(0,0,0)}\right] - k_{2,forw}\cdot E\left[P_{(1,0,0)}\right] \\
\frac{dE\left[P_{(1,0,1)}\right]}{dt} &= -k_{1,rev}\cdot E\left[P_{(1,0,1)}\right] + k_{2,rev}\cdot E\left[P_{(1,1,1)}\right] - k_{3,rev}\cdot E\left[P_{(1,0,1)}\right] + k_{1,forw}\cdot E\left[P_{(0,0,1)}\right] - k_{2,forw}\cdot E\left[P_{(1,0,1)}\right] \\
\frac{dE\left[P_{(1,1,0)}\right]}{dt} &= -k_{1,rev}\cdot E\left[P_{(1,1,0)}\right] - k_{2,rev}\cdot E\left[P_{(1,1,0)}\right] + k_{3,rev}\cdot E\left[P_{(1,1,1)}\right] + k_{1,forw}\cdot E\left[P_{(0,1,0)}\right] + k_{2,forw}\cdot E\left[P_{(1,0,0)}\right] - k_{3,forw}\cdot E\left[P_{(1,1,0)}\right] \\
\frac{dE\left[P_{(1,1,1)}\right]}{dt} &= -k_{1,rev}\cdot E\left[P_{(1,1,1)}\right] - k_{2,rev}\cdot E\left[P_{(1,1,1)}\right] - k_{3,rev}\cdot E\left[P_{(1,1,1)}\right] + k_{1,forw}\cdot E\left[P_{(0,1,1)}\right] + k_{2,forw}\cdot E\left[P_{(1,0,1)}\right] + k_{3,forw}\cdot E\left[P_{(1,1,0)}\right]
\end{aligned}
\tag{53}
$$

Fig. 32: Moment closure ODEs of the insulated channel (continued from page 159). This part of the ODE system depicts also the moment closure after the first moment, i.e. the expectation values of the species.

Selbständigkeitserklärung

Ich, Friedemann Uschner, erkläre hiermit, dass ich die vorliegende Arbeit selbständig verfasst und noch nicht für andere Prüfungen eingereicht habe. Sämtliche Quellen einschließlich Internetquellen, die unverändert oder abgewandelt wiedergegeben werden, insbesondere Quellen für Texte, Grafiken, Tabellen und Bilder, sind als solche kenntlich gemacht. Mir ist bekannt, dass bei Verstößen gegen diese Grundsätze ein Verfahren wegen Täuschungsversuchs bzw. Täuschung eingeleitet wird.