Fakultät für Maschinenwesen
Technische Universität München

Technische Universität München

Professur für Systembiotechnologie

Model-based Characterization of Intrinsic Noise in Multistable Genetic Circuits

Sayuri Katharina Hortsch

Vollständiger Abdruck der von der Fakultät für Maschinenwesen der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktor-Ingenieurs (Dr.-Ing.)

genehmigten Dissertation.

Vorsitzende:

Prof. Dr. rer. nat. Sonja Berensmeier

Prüfende der Dissertation:

1. Prof. Dr.-Ing. Andreas Kremling
2. Prof. Dr. rer. nat. Christina Kuttler

Die Dissertation wurde am 02.07.2018 bei der Technischen Universität München eingereicht und durch die Fakultät für Maschinenwesen am 03.09.2018 angenommen.

Meiner Familie

Danksagung

Die vorliegende Arbeit entstand während meiner Zeit als wissenschaftliche Mitarbeiterin an der Professur für Systembiotechnologie der Technischen Universität München. Diese Zeit war geprägt von der Beschäftigung mit vielen spannenden Forschungsthemen, von neuartigen interessanten Herausforderungen und Erfahrungen, und von den vielen wunderbaren Menschen, die ich kennenlernen durfte.

Dies alles wäre nicht möglich gewesen ohne meinen Doktorvater, Herrn Prof. Dr.-Ing. Andreas Kremling. Vielen Dank für Ihr großes Vertrauen, für Ihre Unterstützung und für den nötigen Raum zu meiner fachlichen und persönlichen Weiterentwicklung. Wenn ich mich festgefahren hatte, brachten Sie mich mit Geduld und Zuversicht dazu, den weiteren Weg doch zu meistern.

Ganz herzlich möchte ich auch meiner Mentorin und Zweitprüferin, Frau Prof. Dr. Christina Kuttler, meinen außerordentlichen Dank aussprechen. Sie waren schon während des Studiums eine wichtige Ansprechpartnerin für mich, vom Proseminar bis hin zu meiner Masterarbeit. Vielen Dank, dass Sie mir immer so hilfsbereit mit Rat und Tat zur Seite standen!

Frau Prof. Dr. Sonja Berensmeier danke ich herzlich für die Übernahme des Prüfungsvorsitzes.

Bei den Partnern des e:biofilm-Projekts, vor allem bei Prof. Dr. Irene Wagner-Döbler und bei Dr. Michael Reck vom HZI Braunschweig, möchte ich mich für die Zusammenarbeit und die vielen spannenden und sehr angenehmen Projekttreffen bedanken. Es war eine Freude, Einblicke in so viele verschiedene Bereiche zu bekommen und darüber diskutieren zu können.

Nun zum vielleicht schönsten Bestandteil meines Arbeitsalltags – meinen lieben Kollegen aus der Systembiotechnologie, der Selektiven Trenntechnik und der Bioverfahrenstechnik: Danke für den tollen Zusammenhalt über die Grenzen der Uni hinweg, den intensiven Austausch, Eure Hilfe, Eure Ratschläge und den vielen Spaß, den wir immer hatten. Ihr habt mir jeden Tag versüßt und Ihr wisst ja, wie sehr ich Süßes schätze! Ein Extra-Dankeschön gebührt Dr. Alberto Marín-Sanguino für das Korrekturlesen von Teilen der Arbeit und für seine sonstigen fachlichen und auch nicht-fachlichen Hilfestellungen.

Während der turbulenten Promotionsphase gab es eine Konstante, und sie geht weit über diese Zeit hinaus: Meine Familie. Danke an meine Eltern, dass Ihr immer für mich da seid, dass Ihr immer mit mir mitgefiebert, mitgelitten und mitgejubelt habt, und Euch immer begeistern konntet für das, was ich mache, egal, was es war. Und an Ralf – meine Dankbarkeit für Deine immerwährende Unterstützung und Geduld, mit der Du mich durch die Promotion begleitet hast, lässt sich kaum in Worte fassen.

Abstract

In biological cells, the stochasticity of biochemical reactions inevitably causes so-called cellular *noise*, resulting in fluctuating copy numbers of mRNA and protein species. In multistable gene expression systems, large noise may lead to random transitions between distinct cellular states and thus to heterogeneous population behavior. The aim of this work was to use mathematical methods to predict and to gain control over the magnitude and the effects of gene expression noise.

In a preliminary study, an analytical comparison between deterministic and stochastic model descriptions was performed and it was shown that the reliability of deterministic models in the context of gene expression is limited: Under common conditions, neither the average cellular behavior, nor the preferred states of the system were depicted correctly. In particular, deterministic bistability was not always associated with stochastic bimodality (and thus with phenotypic heterogeneity). Criteria for the occurrence of such discrepancies could be identified and visualized through a systematic graphical model analysis. Among these criteria are large protein bursts combined with nonlinear reaction kinetics. The usage of deterministic approaches for model simplification in order to approximately quantify cellular noise can thus lead to poor estimates. This is, for example, the case with the classical linear noise approximation (LNA).

Due to that, a novel method called hybrid LNA (hLNA) was developed for the evaluation of noise. This method is able to predict the variance of mRNA and protein fluctuations, based on the kinetic parameters of the underlying regulatory system, with significantly increased accuracy. Using this method, every stable expression state can be analyzed separately, so that multistable systems can be characterized as well. Moreover, new definitions for the average protein burst size and frequency were derived, which are again suitable for the characterization of multistable systems. Based on those measures, the temporal structure of fluctuations can be investigated. Additionally, a model reduction method was proposed based on the burst characteristics, which enables an *a priori* estimation of the location of modes in a protein distribution. The quality of all newly developed methods was verified through simulation studies.

This framework of novel methods was subsequently used to comprehensively characterize noise in two bistable motifs, namely in an autoregulatory gene expression system and in the genetic toggle switch. It was systematically evaluated how the strength and structure of mRNA and protein fluctuations may be manipulated using specific genetic modifications of the regulatory system. That way, the robustness of single expression states could be changed, which led to various, partially complex heterogeneous patterns of population behavior.

All in all, this work demonstrated the significance of intrinsic noise in gene expression systems, regarding the choice and reliability of mathematical modeling approaches as well as the biological functionality of a system. The presented methods are not only suited to incorporate the effects of cellular noise into the study of natural systems, but also to support the creation of robust synthetic constructs, for example for applications in biotechnology.

Zusammenfassung

In biologischen Zellen verursacht die Stochastizität biochemischer Reaktionen zwangsläufig sogenanntes zelluläres *Rauschen*. Dieses äußert sich u. a. in fluktuierenden Kopienzahlen von mRNA- und Proteinspezies. In multistabilen Genexpressionssystemen kann starkes Rauschen zu zufallsgetriebenen Übergängen zwischen verschiedenen Zellzuständen und damit zu heterogenem Populationsverhalten führen. Das Ziel der Arbeit bestand darin, mit Hilfe mathematischer Methoden die Stärke und die Auswirkungen von Genexpressions-Rauschen vorhersagbar und besser kontrollierbar zu machen.

Als Voruntersuchung wurde ein analytischer Vergleich deterministischer und stochastischer Modellbeschreibungen vollzogen und gezeigt, dass die Aussagekraft deterministischer Modelle im Kontext der Genexpression eingeschränkt ist: Unter üblichen Bedingungen bildeten sie weder das durchschnittliche Zellverhalten, noch bevorzugte Zellzustände korrekt ab. Insbesondere war deterministische Bistabilität nicht immer mit stochastischer Bimodalität (und damit mit phänotypischer Heterogenität) assoziiert. Kriterien für das Auftreten solcher Diskrepanzen konnten durch eine systematische graphische Modellanalyse identifiziert und veranschaulicht werden. Unter diesen Kriterien sind vor allem große Protein-*Bursts*, kombiniert mit nichtlinearen Reaktionskinetiken, zu nennen. Werden demnach deterministische Ansätze zur Modelleinfachung genutzt, um zelluläres Rauschen approximativ zu quantifizieren, kann dies zu gravierenden Fehleinschätzungen führen. Dies ist beispielsweise bei der klassischen *Linear Noise Approximation* (LNA) der Fall.

Daher wurde zur Quantifizierung des Rauschens eine neue Methode, die Hybrid-LNA (hLNA) entwickelt. Diese kann, basierend auf den kinetischen Parametern des betrachteten Regulationssystems, die sich ergebende Varianz der mRNA- und Proteinfluktuationen mit deutlich gesteigerter Genauigkeit vorhersagen. Dabei kann jeder stabile Expressionszustand getrennt untersucht werden, so dass auch multistabile Systeme charakterisiert werden können. Desweiteren wurden neue Definitionen für die durchschnittliche Protein-*Burst*-Größe und –Frequenz hergeleitet, die ebenfalls für die Charakterisierung multistabiler Systeme geeignet sind. Damit kann auch die zeitliche Struktur von Fluktuationen untersucht werden. Darüberhinaus wurde basierend auf den *Burst*-Charakteristiken eine Methode zur Modellreduktion vorgeschlagen, mit deren Hilfe eine *a-priori*-Abschätzung der Lage der Modalwerte einer Proteinverteilung möglich ist. Die Qualität der neu entwickelten Methoden wurde durch Simulationsstudien verifiziert.

Das erarbeitete Methodenspektrum wurde anschließend zur umfassenden Charakterisierung des Rauschens in zwei bistabilen Motiven verwendet, nämlich in einem autoregulatorischen Genexpressionssystem und im genetischen *Toggle-Switch*. Es konnte systematisch untersucht werden, wie die Stärke und Struktur von mRNA- und Proteinfluktuationen durch spezifische genetische Eingriffe in das Regulationssystem manipuliert werden können. Dadurch ließ sich sogar gezielt die Robustheit einzelner Expressionszustände beeinflussen, was wiederum zu unterschiedlichem, teils komplexem heterogenem Populationsverhalten führte.

Insgesamt konnte in dieser Arbeit die Signifikanz intrinsischen Rauschens in Genexpressionssystemen verdeutlicht werden. Dies betrifft sowohl die Wahl und Zuverlässigkeit mathematischer Modellierungsansätze, als auch die biologische Funktionalität des Systems. Die vorgestellten Methoden eignen sich nicht nur, um die Effekte zellulären Rauschens in die Studie natürlicher Systeme einzubeziehen, sondern auch, um die Schaffung robuster synthetischer Konstrukte, beispielsweise für Anwendungen in der Biotechnologie, zu unterstützen.

Contents

List of Figures

List of Tables

List of Variables and Functions

$\mathbf{A}$	Stoichiometrix matrix of reaction system
$\mathbf{a}_j$	j-th column of $\mathbf{A}$
a_{ij}	(i,j)-th entry of $\mathbf{A}$
a	Scaled basal transcription rate
$\hat{a}$	Unscaled basal transcription rate
B	Stochastic protein burst size
c_m	Deterministic mRNA concentration
c_m^*	mRNA concentration in the deterministic steady state
c_s	Deterministic protein concentration
c_s^*	Protein concentration in the deterministic steady state
$\mathbf{D}(\mathbf{x})$	Diffusion tensor in a Fokker-Planck equation, evaluated at state $\mathbf{x}$
$\mathbf{D}^*$	Local diffusion tensor in the linear noise approximation
d_m	Unscaled mRNA degradation rate constant
d_s	Unscaled protein degradation rate constant
$\mathbb{E}$	Expected value operator
F	Scaled transcriptional propensity function
$\hat{F}$	Unscaled transcriptional propensity function
$\mathcal{F}$	Scaled transcription rate
g	Scaled translation rate constant
$\hat{g}$	Unscaled translation rate constant
g^*	Local translational sensitivity
G	Scaled translational propensity function
$\bar{G}$	Scaled effective translation rate in terms of copy numbers
$\mathcal{G}$	Scaled translation rate
$\bar{\mathcal{G}}$	Scaled effective translation rate
Geo	Geometric probability mass function
H	Function proportional to F that is independent of burst size
$\mathcal{H}$	Function proportional to $\mathcal{F}$ that is independent of burst size
f^*	Local transcriptional sensitivity
$\mathbf{J}(\mathbf{x})$	Jacobian matrix, evaluated at state $\mathbf{x}$
$\mathbf{J}^*$	Local Jacobian matrix
k_j	Deterministic rate constant of j-th reaction
M	Stochastic process describing mRNA copy numbers
M_1	Stochastic copy number of mRNA 1 in the genetic toggle switch
M_2	Stochastic copy number of mRNA 2 in the genetic toggle switch
m^*	mRNA copy number in the deterministic steady state
$\bar{m}$	Current average mRNA copy number
$p_{\mathbf{n}}(\tau)$	Probability of system being state $\mathbf{n}$ at time τ
p_b^B	Probability of protein burst with size $B = b$
$p_{m,s}$	Joint PMF of mRNA and protein evaluated at $M = m$ and $S = s$

p_{m_1,m_2,s_1,s_2}	Joint PMF of mRNA and protein species in the genetic toggle switch
p_m^M	Marginal PMF of mRNA evaluated at $M = m$
p^{M*}	Local unimodal mRNA PMF
p_s^S	Marginal PMF of protein evaluated at $S = s$
p^{S*}	Local unimodal protein PMF
$p_s^{S_B}$	PMF of burst proteins evaluated at $S_b = s$
$P(\mathbf{x}, \tau)$	Probability density function, evaluated for state $\mathbf{x}$ at time τ
R^*	Combined parameter decisive for the effect of partial shifts in reaction time scales
r^*	Local protein-to-mRNA ratio
S	Stochastic process describing protein copy numbers
S_1	Stochastic copy number of protein 1 in the genetic toggle switch
S_2	Stochastic copy number of protein 2 in the genetic toggle switch
S_b	Stochastic copy number of burst proteins (during burst)
s^*	Protein copy number in the deterministic steady state
$\mathring{s}$	Mode (or minimum) of a protein distribution
$\bar{s}$	Current average protein copy number
V	Volume of the regarded system
Var	Variance operator
$\mathbf{v}(\mathbf{x})$	Drift vector in a Fokker-Planck equation, evaluated at state $\mathbf{x}$
w_j	Propensity of j-th reaction
$\mathbf{w}$	Vector of propensities
w_m^*	Stationary propensity of reactions on mRNA level
w_s^*	Stationary propensity of reactions on protein level
$\tilde{\mathbf{w}}$	Vector of propensities in the thermodynamic limit
$\mathcal{Z}$	(Probability) generating function
α	Average protein burst size
α^*	Local average protein burst size
β^*	Effective protein burst size
β_{ij}	Absolute value of stoichiometric coefficient of i-th educt in j-th reaction
γ^*	Observable protein peak height
γ_{ij}	Stoichiometric coefficient of i-th product in j-th reaction
δ	Time-scale parameter relating the two subsystems in the genetic toggle switch
$\eta(N)$	Fano factor of the stochastic process N
$\eta^*(N)$	Local Fano factor of the stochastic process N
ν	Time-scale parameter relating mRNA and protein dynamics
κ_j	Stochastic reaction constant of j-th reaction
π	Propensity of state transition
Ψ^*	Combined parameter for the analysis of the toggle switch
ρ_0^*	Average protein level at the beginning of a non-overlapping burst
ρ_m^*	Average protein level at the end of a non-overlapping burst
σ	Standard deviation operator
$\mathbf{\Sigma}$	Approximate variance matrix
$\mathbf{\Sigma}^*$	Approximate local variance matrix
τ	Scaled time variable
ω	Average protein burst frequency
ω^*	Local average protein burst frequency

List of Abbreviations

B. subtilis	*Bacillus subtilis*
CFPE	chemical Fokker-Planck equation
CME	chemical master equation
CV	coefficient of variation
E. coli	*Escherichia coli*
EMRE	effective mesoscopic rate equation
fb	feedback
FPE	Fokker-Planck equation
FPT	first passage time
hLNA	hybrid linear noise approximation
IOS	inverse omega square
LHS	left hand side (of an equation)
LNA	linear noise approximation
MFPT	mean first passage time
mRNA	messenger ribonucleic acid
NB	negative binomial
ODE	ordinary differential equation
PDE	partial differential equation
PMF	probability mass function
PSS	pseudo steady state
RE	rate equation
RHS	right hand side (of an equation)
SEM	standard error of the mean
S. mutans	*Streptococcus mutans*

1. Introduction: Noise and Heterogeneity

Looking back at the research done in the field of microbial or cellular biology, it seems that the original approach of studying metabolic pathways, regulatory networks, and signaling cascades based on ensemble measurements has shaped the picture of single cells as small machines or factories, which generate defined outputs from specific input signals. The ambition to gain full control over these machines – for example for their use in biotechnology – has also motivated the development of apt mathematical models: In accordance with experimental observations, they were able to describe population-averaged kinetics of growth, product formation, intracellular reactions, etc. [Almquist et al., 2014; Kompala et al., 1984; Koutinas et al., 2011]. The models were usually deterministic, often phenomenological, and sometimes motivated by concepts from control theory [Alon, 2007; Iglesias and Ingalls, 2009]. These experimental and theoretical tools have indeed boosted our understanding of complex cellular networks. Moreover, they have enabled systematic modifications of these networks or the model-guided construction of synthetic circuits according to various goals and requirements in research and industry [He et al., 2016]. However, working with cells (or more generally, with biological entities) has never really become completely unproblematic, as they remained somehow capricious. The true extent of cellular individualism has only come to light in the last few decades through the development of experimental techniques which allowed measurements on a single-cell level instead of using ensemble averages. These experiments proved a significant level of heterogeneity in all kinds of cellular behavior. This can partly be attributed to an inhomogeneous environment, but it is also a consequence of an inherent stochasticity present inside cells [Elowitz et al., 2002; McAdams and Arkin, 1997]. What is the origin of this stochasticity? By regarding biochemical processes as collections of random reactive collisions between molecules, their stochastic nature becomes obvious [Gillespie, 1976]. Fluctuations in the copy numbers of cellular components, referred to as *cellular noise*, are thus an inevitable trait. Although the same basically holds for chemical reactions occurring in a large, well-mixed reaction vessel, their reproducibility is usually much higher, showing a difference in the *significance* of fluctuations. The main reason for this difference is the sparseness of many cellular reactants like genes, mRNA species, and protein species, in contrast to the common abundance of reactive species in chemical experiments [Kærn et al., 2005]. In the low copy number regime, even fluctuations in the order of single molecules may have great impact on the dynamics of the system. Additional stochasticity is caused by cellular dynamics like growth, replication and cell division [Huh and Paulsson, 2011]. Therefore, the copy number or concentration of a specific protein is never uniform across the cells of a clonal population. Instead, it follows a distribution that might be relatively broad, and which might even have a complex shape with multiple peaks, indicating the presence of phenotypic subpopulations. These could not be observed through population-averaged measurements [Dubnau and Losick, 2006].

This insight has led to a change in the perception of cells as being fully controllable devices and maybe to the impression that one might never be able to understand them completely. Although it would be quite wrong to state that all methods and results based on ensemble averaged experiments are void, single-cell experiments can provide additional – or even crucial

– insight into cellular mechanisms and strategies, which have remained hidden for a long time [Altschuler and Wu, 2010]. In fact, it is fair to assume that phenotypic heterogeneity[1], like any other cellular trait, has been shaped by evolution and can thus be an advantageous property [Ackermann, 2015; Cerulus et al., 2016; Dubnau and Losick, 2006] – at least from the perspective of a cellular population.

One of the evolutionary goals that a population might pursue using heterogeneity is the anticipation of environmental fluctuations: The split into phenotypically different subpopulations may provide a certain amount of cells that are disadvantaged under most conditions (for example with respect to growth rate), but which may be the only ones surviving sudden external stresses, as they have invested most of their energy and resources into defense mechanisms[2]. Such slowly growing, resilient subpopulations are known as *persister cells* [Balaban et al., 2004], and the overall strategy of reducing the current fitness of a population in order to increase its chance of long-term survival or its overall productivity is called *bet hedging* [Grimbergen et al., 2015; Veening et al., 2008]. Another strategy that exploits heterogeneity is *division of labour* [Ackermann, 2015]. In this case, subpopulations are formed which are specialized onto the provision of certain goods. Interactions between these subpopulations then result in collective functionality. This is of particular relevance if some important metabolic pathways are incompatible, or if the simultaneous activation of multiple pathways would culminate in a severe metabolic burden. In contrast to heterogeneity caused by genetic variability, isogenic heterogeneity is self-sustaining, since eliminated phenotypes can be regenerated [Ackermann, 2015]. Sometimes, phenotypic diversification occurs irreversibly, constituting a process of cell fate decision [Huang, 2009].

Nowadays, cell-to-cell variability is undesired in most industrial applications, as it hampers reproducibility. Current biotechnological research therefore aims at the reduction of cellular noise [Delvigne et al., 2017]. However, most studies in this applied field deal, if at all, with environmental heterogeneity in bioreactors [Kuschel et al., 2017; Lara et al., 2006; Lemoine et al., 2017], while not paying regard to the omnipresence of intracellular stochasticity [Delvigne and Goffin, 2014; Lencastre Fernandes et al., 2011]. This is partially due to fact that single-cell studies are more elaborate and less standardized than the majority of established population-averaged methods. Besides the larger expenditure of time and costs for the performance and evaluation of experiments[3], the corresponding mathematical approaches are tedious as well: While deterministic models of moderate complexity can often be analyzed in a straightforward manner, supported by graphical methods that provide intuitive insight into certain basic system properties, this does not necessarily hold for stochastic descriptions [Tolle and Le Novere, 2006]. Nevertheless, a lot of fundamental research concerning the generation and propagation of fluctuations on a genetic and regulatory level has already been conducted in the last years [Chalancon et al., 2012], elucidating fundamental connections between properties of a regulatory network and its noise. Many of those examples will be presented throughout this study. To date, cellular noise is still hard to predict, and the real potential of heterogeneity is far from being fully explored. The vision however exists that noise will once be made controllable through targeted genetic manipulations. One day, it might even become possible to transfer the advantages of heterogeneity in natural systems to biotechnological applications in order to improve their performance and robustness [Delvigne et al., 2017; González-Cabaleiro et al., 2017].

[1] The term *phenotypic heterogeneity* has been introduced in a review by M. Ackermann and denotes diversity between isogenic cells in the same microenvironment [Ackermann, 2015].

[2] This phenomenon is a major issue in medicine, as it is known to increase bacterial drug resistance.

[3] A short overview of experimental single-cell methods is given e.g. in [Ackermann, 2015; Delvigne et al., 2017; Huang, 2009]. Among these are flow cytometry [Baert et al., 2015], microfluidics coupled to time-lapse microscopy [Grünberger et al., 2015], and single-cell sequencing [Winkel et al., 2016].

In order to better understand the intracellular origins of phenotypic heterogeneity, one should look at the level of gene expression and regulation. Gene expression occurs steadily in cells to ensure the supply of important proteins (like enzymes) against the backdrop of protein degradation and dilution due to growth. Many genetic circuits are additionally regulated, so that they can dynamically adapt the intensity of protein production to the current needs of the cell, e.g. in response to the environment. Some circuits show a gradual response according to the strength of the input signal [Igoshin et al., 2008]. Others operate almost exclusively in one of two (or more) distinct expression states, while switches between these states are possible in reaction to specific triggers [Novick and Weiner, 1957]. In contrast to circuits with gradual response, they are able to sustain more than one stable phenotype under identical environmental conditions [Dubnau and Losick, 2006]. Additionally, they often show a behavior called *hysteresis*, which denotes a kind of robustness against external fluctuations preventing uncontrolled ON-OFF-switches [Savageau, 1999]. Many natural genetic circuits exhibit this all-or-nothing behavior, because it allows a strong, fast, and robust response. Moreover, such circuits are increasingly used in biotechnology: With their help, cells can e.g. be artificially switched from growth to production phase at a certain stage of the bioprocess for increasing its overall efficiency [Anesiadis et al., 2008; Roquet and Lu, 2014]. Such approaches form a contrast to traditional static metabolic engineering strategies. However, the noise that is inevitably associated with biochemical processes might hamper the controlled usage of such cellular switches. This is depicted in Fig 1.1:

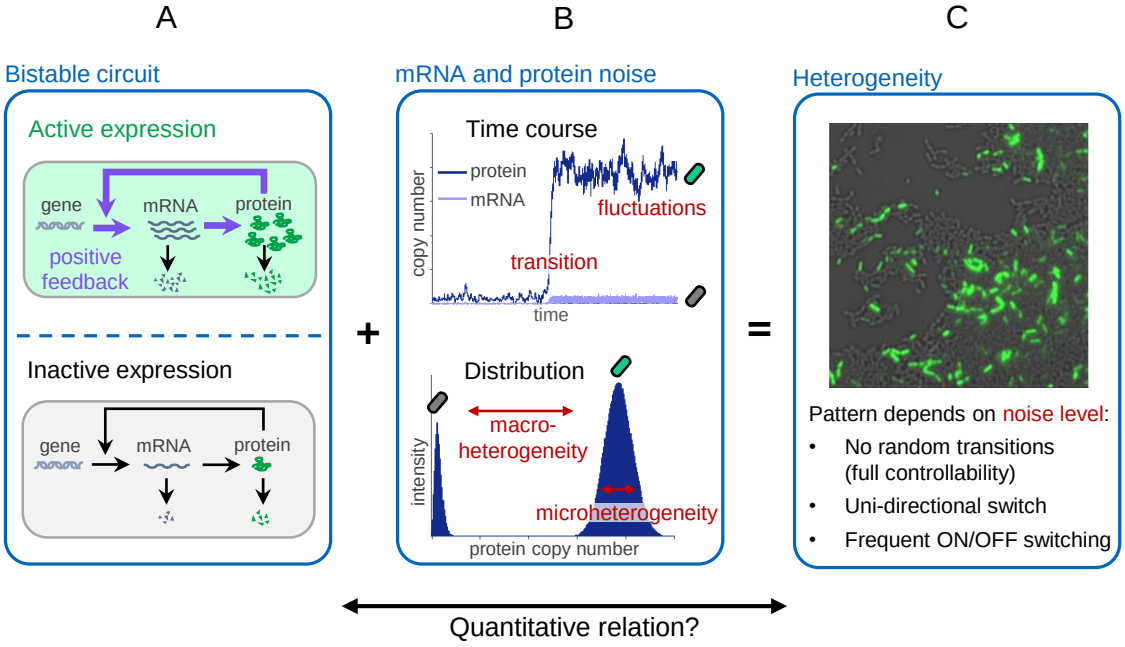


Figure 1.1.: **Emergence of phenotypic heterogeneity from noisy circuits.** In (A), a circuit with a positive feedback loop is shown, which has two different stable expression states. The stochasticity of processes like transcription, translation or molecule degradation generates fluctuations in mRNA and protein copy numbers as shown in the simulated time course in (B) (upper plot). Here, noise is large enough to cause a spontaneous transition from the inactive to the active expression state through a random activation of the positive feedback loop. The corresponding population distribution (lower plot) exhibits both micro- and macroheterogeneity. Macroheterogeneity can be observed using suitable fluorescent reporter strains, cf. (C) (image section taken from [Reck et al., 2015]). However, not all bistable circuits result in phenotypic heterogeneity, since the probability of random transitions depends on the noise levels of each expression state. Depending on the intensity of fluctuations, transitions might be lacking, uni-directional, or bi-directional. Although a direct connection between circuit properties and heterogeneity patterns exists, it cannot be easily determined. This limits the possibility to predict and to control heterogeneity.

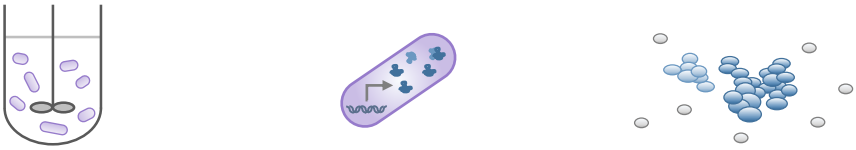
In Fig 1.1 (A), a simplified circuit is illustrated which shows either high or low gene expression and can therefore act as a cellular ON-OFF-switch. This binary behavior is due to the involvement of a positive feedback loop, in which the target protein stimulates its own production, thereby intrinsically maintaining the active or inactive expression state. If the circuit acted deterministically, the cell would rest in one of the stable states until an external perturbation triggers a state-transition. However, the noise within the circuit leads to temporal fluctuations in the copy numbers of cellular components. This is illustrated in Fig 1.1 (B). As long as these fluctuations occur within the attractor region of an expression state, they simply broaden the peak representing the phenotype in the population distribution. This temporal heterogeneity is also denoted as *microheterogeneity* [Huang, 2009]. Noise can, however, be large enough to generate random transitions between expression levels, resulting in *macroheterogeneity* [Huang, 2009], i.e. in the coexistence of different phenotypic subpopulations[4].

While noise-driven macroheterogeneity may be advantageous in nature, the majority of current industrial applications requires almost noise-free circuits with full external controllability. It is thus important to consider the impact of noise on the functionality of a circuit and to study how different noise levels are generated (different "patterns" of heterogeneity are listed in Fig 1.1 (C)). In general, intrinsic noise is solely determined by the architecture of the circuit itself and by the dynamics of its reactions and can in principle be altered through suitable genetic manipulations. In order to systematically analyze and to understand this rather direct, but intricate connection, mathematical modeling of the regarded system is highly valuable. For capturing the stochasticity of biochemical reactions, *mesoscopic* models are advantageous, which provide a description of the system at single-species resolution. Their complexity is higher than that of macroscopic models depicting population dynamics, but lower than that of microscopic models, with which molecular dynamics are described based on quantum physical principles [Grima et al., 2011; Tolle and Le Novere, 2006], cf. Fig 1.2. The latter are usually too involved to efficiently study biochemical reactions with several regulatory interactions, however, the analysis of mesoscopic models may be challenging as well. Despite a multitude of studies about circuit noise that make use of stochastic models, the analytical connection between circuit properties and noise is still rather unclear, especially in the context of macroheterogeneous systems. This is partially due to the lack of suitable theoretical methods.

This motivates the overall goal of the study at hand – namely the development and application of a theoretical framework for the comprehensive and predictive characterization of intrinsic noise based on the underlying circuit. Special attention will be paid to systems with multiple expression states. Based on the gained insights, one should be able to identify strategies for the targeted adjustment of noise patterns. Since noise is an important trigger of spontaneous cell state transitions, this would also be a first step towards a better understanding and control of phenotypic heterogeneity. The obtained results are expected to improve our current understanding of natural mechanisms leading to cell-to-cell variability. Moreover, they should enable an *a-priori* assessment of whether a given circuit is robust enough for industrial applications. Last but not least, they should help in the design of synthetic circuits from a more "stochastic" perspective, since it is nowadays primarily guided by deterministic considerations only.

As already stated, the theoretical framework needs to be mainly based on mesoscopic, i.e. on stochastic modeling approaches. In this context, two major challenges are expected: The first

[4]The dynamics of a circuit can be figuratively thought of as marbles moving on a "potential landscape" [Hsu et al., 2016; Tang et al., 2017; Waddington, 1957], whose valleys represent the favoured expression states. The marbles tend to roll downhill, but are also subject to noise-driven random movement that may cause spontaneous crossing of hills.



Model classification	Macroscopic	Mesoscopic	Microscopic
Resolution	Population dynamics	Species/particle dynamics	Molecular dynamics
Application	Systems with large volume and large species copy numbers e.g. microbial growth	Systems with small volume and small species copy numbers e.g. regulated gene expression	Systems with very few molecules and known atomic composition e.g. protein-protein interactions
Approaches (examples)	Rate equations (homogenous), Reaction-diffusion equations (heterogenous)	Chemical master equation (homogenous)	Quantum physical approaches

Figure 1.2.: **Classification of models for the dynamic description of biochemical processes.** Modeling approaches are grouped according to their resolution. From left to right, the complexity of the approach increases, so that the complexity of the described systems needs to be reduced accordingly.

concerns the handling of these stochastic models themselves, since their analysis is generally known to be intricate. In order to obtain closed-form solutions for the regarded noise characteristics, which show their dependence on the properties of the circuit, approximations and simplifications are often necessary. These might severely reduce the quality of results. Many of the existing approaches use an underlying deterministic description, but the applicability of deterministic models to the analysis of highly stochastic systems is questionable and needs to be critically assessed. The second challenge concerns the definition and extraction of noise characteristics in a macroheterogeneous regime: While noise can usually be quantified using measures like the mean value and variance of a probability distribution, in case of macroheterogeneity, the characterization needs to be performed for every phenotypic subpopulation separately, i.e. for every peak in the probability distribution. This is non-trivial, since from a mathematical perspective, such "local" noise measures may be rather artificial. Currently existing approaches are not fully reliable; their quality depends severely on the mathematical properties of the described system. The development of novel, more suitable approaches might therefore be required.

The remaining chapters are organized as follows: In Chapter 2, the theoretical background of this study is provided, including the most important mathematical concepts and tools that are used and will partially be refined in the course of the study. Chapter 3 studies mathematical connections and discrepancies between common deterministic and stochastic modeling approaches in the context of gene expression and regulation. Based on that, the validity, interpretability, and reliability of deterministic descriptions is critically assessed. These results provide a basis for the development of novel methods in Chapter 4, with which expression-state dependent noise levels in gene regulatory circuits can be characterized and analyzed. In the subsequent two chapters, the methods are used to study noise patterns in two basic regulatory motifs that can serve as biological ON-OFF-switches: Single-gene autoregulation (Chapter 5) and the genetic toggle switch (Chapter 6). Each of the Chapters 3-6 is closed with a short summary and discussion. Chapter 7 then provides an overall discussion of the obtained results, followed by the main conclusions and possible future directions.

2. Theoretical Background and Methods

This chapter introduces the most important terminology, notations, definitions, and established methods that are used in this work for the modeling and analysis of biochemical processes. They will be embedded in a broader theoretical context to allow comparisons with alternative methods and to give a rough outline of current mathematical developments in this field. The main focus will be on mesoscopic models, which provide an appropriate level of complexity for the description of intracellular processes on a single-cell level, cf. Fig 1.2. In contrast to deterministic (macroscopic) models based e.g. on rate equations (REs), they are able to describe stochastic effects without going into too much molecular detail [Grima et al., 2011]. With their help, different aspects of noise in genetic circuits have already been successfully studied in the last years and decades. Some of the main results will be presented in this chapter.

Throughout this work, it is assumed that the interior of cells is spatially homogeneous, i.e. that all intracellular reactants are diffusing freely in the cytoplasm. This is certainly a strong simplification, as any cell – be it prokaryotic or eukaryotic – is in fact packed with macromolecules, which cause molecular crowding effects and which often form spatial clusters [Roberts et al., 2011]. Besides that, noise contributions from the environment and from cellular dynamics (growth, replication, and division) will mostly be ignored. Although all these neglected effects are expected to actually have a significant impact on cellular behavior, it is valuable to first understand the reduced problem (which will itself turn out to be quite complex) before trying to get the full picture – the basic principles derived from simplified systems should be present and relevant in real, complex systems as well, albeit superposed by further effects. In fact, numerous results derived from spatially homogeneous mesoscopic models could be verified experimentally, especially in the field of synthetic biology. More details will be given in Section 2.2.

The chapter will start with the introduction of important stochastic modeling approaches, then proceed to the quantification and characterization of noise in the context of gene expression and regulation. The last part will deal with a comparison of deterministic and stochastic descriptions of heterogeneity.

2.1. Stochastic modeling and analysis of biochemical systems

2.1.1. The chemical master equation

The *chemical master equation* (CME, also called *master equation*) is nowadays regarded as the "gold standard" for describing biochemical processes inside living cells [Meister et al., 2014], although it has met some criticism in the past. Conditions under which one of the most restrictive assumptions of this approach, namely spatial homogeneity, is physically justified, could be identified for chemical reactions in gas phase [Gillespie, 1976, 1977, 1992] or in solution [Gillespie, 2009]. These conditions ensure the dominance of the frequency of non-reactive molecular collisions over reactive ones [Nicolis et al., 1974], thereby creating a large enough mixing effect. Even though the conditions do not exactly hold in the intracellular space, the CME is still the

most accurate description available that allows the examination of stochastic cellular processes with reasonable complexity: The additional inclusion of spatial gradients would lead to models on the level of molecular dynamics [Grima, 2010], which would just be too detailed to efficiently depict and analyze complex cellular processes [Gibson and Bruck, 2000].

Formulation of the CME

There are several detailed derivations of the master equation available in literature [Anderson and Kurtz, 2015; Gillespie, 1976; McQuarrie, 1967; van Kampen, 2007]. This section will thus quickly proceed to the general formulation of the CME and just provide some additional comments.

Let us consider a system with constant volume V that contains molecular populations of M different reactive species. As the system is assumed to be spatially homogeneous, its current state can be characterized by the copy numbers of all species, which are collected in a discrete-valued state vector $\mathbf{N}(\tau) = (N_1(\tau)\ N_2(\tau)\ ...\ N_M(\tau))^\top$, where τ denotes time. The state vector changes whenever a (bio-)chemical reaction occurs, according to the stoichiometry of the reaction. As the reactions happen randomly, $\mathbf{N}(\tau)$ is a stochastic process.

Its probabilistic dynamics can be mathematically described using the CME, which is basically a differential form of the Chapman-Kolmogorov equation for memoryless Markov processes: Let $\mathbf{n} \in (\mathbb{N}_0)^M$ be a possible state of the stochastic process and $p_\mathbf{n}(\tau)$ be the probability that $\mathbf{N}(\tau) = \mathbf{n}$. The CME can initially be written as:

$$\dot{p}_\mathbf{n}(\tau) = \sum_{\mathbf{n}'} \left(\pi(\mathbf{n}, \mathbf{n}')\ p_{\mathbf{n}'}(\tau) - \pi(\mathbf{n}', \mathbf{n})\ p_\mathbf{n}(\tau) \right), \tag{2.1}$$

where $\dot{p}$ is the time derivative of p, $\pi(\mathbf{n}_2, \mathbf{n}_1)$ is the transition probability per infinitesimal unit time from state $\mathbf{n}_1$ to $\mathbf{n}_2$, and the summation runs over all possible states. According to this formulation, the probability of being in state $\mathbf{n}$ is enhanced through reactions that transfer the system from other states to $\mathbf{n}$, and reduced through reactions away from $\mathbf{n}$. Non-negativity of the state space is ensured by the conditions $p_\mathbf{n}(0) = 0\ \forall\ \mathbf{n} \ngeq 0$ and $\pi(\mathbf{n}_1, \mathbf{n}_2) = 0$ if $\mathbf{n}_i \ngeq 0$, $i = 1, 2$. The CME in the given form is discrete in state space, but continuous in time (there is also a version of the CME with continuous state space, but in the context of small species populations, it is much more accurate to take the discreteness of copy numbers into account).

The transition probability $\pi(.,.)$ can be further specified: This function only regards transitions caused by single reactions, as the infinitesimal time unit $\delta\tau$ is chosen sufficiently small so that the probability of having more than one reaction in a time interval of length $\delta\tau$ is negligible. The "fundamental hypothesis" of the CME [Gillespie, 1976] is that the probability of the j-th reaction to occur within the interval $[\tau, \tau + \delta\tau]$ is proportional to $\delta\tau$:

$$\begin{aligned} w_j(\mathbf{n})\,\delta\tau\ :=\ & \text{the probability that the } j\text{-th reaction will occur in } [\tau, \tau + \delta\tau], \\ & \text{if the system is in state } \mathbf{n} \text{ at time } \tau. \end{aligned} \tag{2.2}$$

$w_j(\mathbf{n})$ is also called the *propensity* of the j-th reaction[1]. If R reaction channels are considered and the stoichiometric matrix is denoted by $\mathbf{A}$ with columns $\mathbf{a}_1, ..., \mathbf{a}_R$, then the CME can be

[1] Non-negativity of the state space is ensured by the condition $(\mathbf{n} \ngeq 0)\ \vee\ (\mathbf{n} + \mathbf{a}_j \ngeq 0)\ \Rightarrow\ w_j(\mathbf{n}) = 0$.

re-written as

$$\dot{p}_{\mathbf{n}}(\tau) = \sum_{j=1}^{R} \left(w_j(\mathbf{n} - \mathbf{a}_j) \cdot p_{\mathbf{n}-\mathbf{a}_j}(\tau) \; - \; w_j(\mathbf{n}) \cdot p_{\mathbf{n}}(\tau) \right). \tag{2.3}$$

In fact, assumption (2.2) is the most critical part in the formulation of the master equation (at least for systems that are spatially homogeneous). For certain physical conditions, D. T. Gillespie has indeed managed to rigorously prove that the assumption is valid: He derived explicit expressions for $w_j(\mathbf{n})$ which only depended on some physical parameters and natural constants [Gillespie, 1976, 1977, 2009]. In common applications of the CME, however, the propensities are pragmatically formulated based on phenomenological constants that are inferred from experimental data. This formulation additionally includes the stoichiometry of an elementary reaction in the following way:

Let $X_i,\ i \in \{1, ..., M\}$, be the i-th reaction component and let the j-th reaction be given by

$$\sum_{i=1}^{M} \beta_{ij} \cdot X_i \quad \longrightarrow \quad \sum_{i=1}^{M} \gamma_{ij} \cdot X_i, \tag{2.4}$$

where $\beta_{ij},\ \gamma_{ij}$ are the stoichiometric coefficients of the educts and products of the reaction. Then, $\mathbf{a}_j = (\gamma_{ij} - \beta_{ij})_{i=1,...,M}$ and

$$w_j(\mathbf{n}) = \kappa_j \cdot \prod_{i=1}^{M} \binom{n_i}{\beta_{ij}}. \tag{2.5}$$

Here, κ_j is the (phenomenological) stochastic reaction constant. The product term reflects the combinatorial chance that all educt molecules collide randomly, which depends on the stoichiometry of the reaction. Note that this dependence is only valid for *elementary* reactions and not for reactions that are composed of multiple steps. In the latter case, the stoichiometry of the rate-limiting step would be decisive for the combinatorial term, cf. [McQuarrie, 1967, p. 415][2]. By inserting Eq (2.5) into the CME formulation (2.3), we are now able to describe the considered reaction system stochastically, based on stoichiometries and some phenomenological constants, similar to the formulation of deterministic ordinary differential equation (ODE) models. The connection between stochastic and deterministic reaction constants will be pointed out in Section 2.1.3.

Solving the CME using generating functions

In case the state variables can take on infinitely many values $\mathbf{n}$, the master equation is actually an infinite-dimensional system of differential equations, which usually cannot be solved in a straightforward manner. However, it can be converted into a single partial differential equation (PDE) using so-called *generating functions* [Kimmel and Axelrod, 2015; Scott, 2006; van Kampen, 2007; Walczak et al., 2012]. For simple cases, this PDE can be solved exactly, and backtransformation yields a solution of the master equation.

[2]Here, an example is given in the context of deterministic rate equations, but the same holds for the corresponding stochastic formulation.

The generating function $\mathcal{Z}$ of a time-dependent array $a_{\mathbf{n}}(\tau), \mathbf{n} \in (\mathbb{N}_0)^M$ is defined as[3]

$$\mathcal{Z} : [0,1]^M \times \mathbb{R} \;\rightarrow\; \mathbb{R}, \qquad \mathcal{Z}(\mathbf{z},\tau) = \sum_{n_1=0}^{\infty} ... \sum_{n_M=0}^{\infty} a_{n_1,...,n_M}(\tau) \cdot z_1^{n_1} \cdot ... \cdot z_M^{n_M}. \tag{2.6}$$

The array can be retrieved using the backtransformation:

$$a_{n_1,...,n_M}(\tau) \;=\; \frac{1}{n_1! \,...\, n_M!} \frac{\partial^{n_1+...+n_M}}{\partial z_1^{n_1} \,...\, \partial z_M^{n_M}} \mathcal{Z}(\mathbf{z},\tau)\Big|_{\mathbf{z}=\mathbf{0}}. \tag{2.7}$$

If $a_{\mathbf{n}}(\tau)$ is the probability mass function (PMF) $p_{\mathbf{n}}(\tau)$ of a discrete, non-negative random process $\mathbf{N}(\tau)$, then

$$\mathcal{Z}(\mathbf{z},\tau) = \mathbb{E}\left[z_1^{N_1(\tau)} \cdot ... \cdot z_M^{N_M(\tau)}\right] = \mathbb{E}\left[e^{\mathbf{r}^\top \mathbf{N}(\tau)}\right] =: \mathcal{M}(\mathbf{r},\tau), \tag{2.8}$$

where $\mathbb{E}$ denotes the expected value operator (cf. next section) and $r_i := \ln(z_i)$. In this case, $\mathcal{Z}$ is usually called the *probability generating function* due to Eq (2.7), while $\mathcal{M}$ is called the *moment generating function*. The latter term refers to the fact that the $(k_1, ..., k_M)$-th mixed moment of $\mathbf{N}(\tau)$, defined as

$$\mu_{\,k_1,...,k_M}(\tau) := \mathbb{E}\left[N_1^{k_1}(\tau) \cdot ... \cdot N_M^{k_M}(\tau)\right] = \sum_{n_1=0}^{\infty} ... \sum_{n_M=0}^{\infty} n_1^{k_1} \cdot ... \cdot n_M^{k_M}\; p_{n_1,...,n_M}(\tau), \tag{2.9}$$

is given by:

$$\mu_{\,k_1,...,k_M}(\tau) \;=\; \frac{\partial^{k_1+...+k_M}}{\partial r_1^{k_1} \,...\, \partial r_M^{k_M}} \mathcal{M}(\mathbf{r},\tau)\Big|_{\mathbf{r}=\mathbf{0}}. \tag{2.10}$$

Going back to Eq (2.3), the transformation of both sides of the equation using their respective generating functions yields a PDE, which can possibly be solved explicitly based on the method of characteristics [Scott, 2006; Walczak et al., 2012], so that a closed-form solution of $p_{\mathbf{n}}(\tau)$ is obtained. Depending on the goal of the analysis, it might also be sufficient to regard the system only in its *steady state* (i.e. under the condition $\dot{p}_{\mathbf{n}}(\tau) = 0$), which facilitates the solution of the equation. This method will be applied a few times throughout this study.

However, in most cases, it remains impossible to find a closed-form solution of the (dynamic or stationary) probability distribution p. Even for the mean vectors and/or variance-covariance matrices of the system state (cf. next section), analytical expressions cannot always be found. This has created the need for numerical and for approximate analytic approaches, a selection of which will be presented later in the chapter.

2.1.2. Mean values, variances, and covariances

Moments are important and intuitive characteristics of a probability distribution, since with their help, the average behavior of a stochastic process and the width and shape of its probability distribution can be described. In order to calculate these measures, we first look at the general definition of an *expected value* (*mean value*, *average*):

[3] The conversion is equivalent to a Z-transform, the discrete analogue of the Laplace-transform.

Let $h : \mathbb{R}^M \to \mathbb{R}$ be an arbitrary function. The expected value of $h(\mathbf{N}(\tau))$ is defined as

$$\mathbb{E}[h(\mathbf{N})(\tau)] := \sum_{n_1=0}^{\infty} ... \sum_{n_M=0}^{\infty} h(\mathbf{n})\, p_{\mathbf{n}}(\tau). \tag{2.11}$$

In the context of biochemical processes, $p_{\mathbf{n}}(\tau)$ is in general unknown, while its time derivative is specified by the master equation. This allows the formulation of the derivative $\dot{\mathbb{E}}[h(\mathbf{N})(\tau)] = \sum_{n_1=0}^{\infty} ... \sum_{n_M=0}^{\infty} h(\mathbf{n})\, \dot{p}_{\mathbf{n}}(\tau)$. The ODE describing the average dynamics of the i-th component N_i is for example given by:

$$\begin{aligned}
\dot{\mathbb{E}}[N_i] &= \sum_{\mathbf{n}\in\mathbb{N}_0^M} n_i\, \dot{p}_{\mathbf{n}} = \sum_{\mathbf{n}\in\mathbb{Z}^M} n_i\, \dot{p}_{\mathbf{n}} \\
&= \sum_{\mathbf{n}\in\mathbb{Z}^M} \left(n_i \sum_{j=1}^{R} \left[\, w_j(\mathbf{n}-\mathbf{a}_j)\cdot p_{\mathbf{n}-\mathbf{a}_j} \;-\; w_j(\mathbf{n})\cdot p_{\mathbf{n}} \,\right] \right) \\
&= \sum_{j=1}^{R} \sum_{\mathbf{n}\in\mathbb{Z}^M} \left[\, (n_i + a_{ij})\cdot w_j(\mathbf{n})\cdot p_{\mathbf{n}} \;-\; n_i \cdot w_j(\mathbf{n})\cdot p_{\mathbf{n}} \,\right] \\
&= \sum_{j=1}^{R} \left(a_{ij} \cdot \mathbb{E}[w_j(\mathbf{N})] \right).
\end{aligned} \tag{2.12}$$

For the sake of simplicity, the time variable τ has been omitted. The second equality is ensured by the condition $p_{\mathbf{n}} = 0 \;\forall\, \mathbf{n} \ngeq 0$. The fourth equality is obtained by an index shift.

In an analogous manner, one can formulate an ODE for the variance of the i-th component, $\mathrm{Var}(N_i) := \mathbb{E}[(N_i - \mathbb{E}[N_i])^2]$, and for the covariance between the i-th and j-th component, $\mathrm{Cov}(N_i, N_j) := \mathbb{E}[(N_i - \mathbb{E}[N_i])(N_j - \mathbb{E}[N_j])]$, using the relations $\dot{\mathrm{Var}}(N_i) = \dot{\mathbb{E}}[N_i^2] - 2\,\dot{\mathbb{E}}[N_i]\,\mathbb{E}[N_i]$ and $\dot{\mathrm{Cov}}(N_i, N_j) = \dot{\mathbb{E}}[N_i\, N_j] - \dot{\mathbb{E}}[N_i]\,\mathbb{E}[N_j] - \mathbb{E}[N_i]\,\dot{\mathbb{E}}[N_j]$.

We can now assess the solvability of Eq (2.12): In case all propensity functions are *linear* (which holds true if all reactions in the system are of zeroth or first order), the equality $\mathbb{E}[w_j(\mathbf{N})] = w_j(\mathbb{E}[\mathbf{N}])$ gives the ODE $\dot{\mathbb{E}}[N_i] = \sum_{j=1}^{R} \left(a_{ij} \cdot w_j(\mathbb{E}[\mathbf{N}]) \right),\ i = 1, ..., M$, or:

$$\dot{\mathbb{E}}[\mathbf{N}] \;=\; \mathbf{A}\, \mathbf{w}(\mathbb{E}[\mathbf{N}]). \tag{2.13}$$

Here, $\mathbf{A}$ is again the stoichiometric matrix defined in Section 2.1.1 and $\mathbf{w} = (w_j)_{j=1,...,R}$ is the vector of reaction propensity functions. Eq (2.13) is a system of first-order linear ODEs with constant coefficients, for which a closed-form solution exists.

If w_j is however nonlinear, in general, $\mathbb{E}[w_j(\mathbf{N})] \neq w_j(\mathbb{E}[\mathbf{N}])$. In this case, multidimensional Taylor expansion of $w_j(\mathbf{N})$ at $\mathbb{E}[\mathbf{N}]$ would show that the ODE of the k-th moment (the first moment being the mean value) depends on the $(k+1)$-st moment as well. This finally results in an infinite-dimensional system of ordinary differential equations, making an exact solution of mean values, variances, etc. generally intractable [McQuarrie, 1967; van Kampen, 2007].

2.1.3. Connection to macroscopic models

At this point, it is useful to regard the macroscopic description of chemical reaction systems based on deterministic rate laws, and to specify some of the connections to the parameters of the CME. This will help in introducing the stochastic modeling approaches presented in the

subsequent sections.

The macroscopic description is based on the law of mass action (originally postulated and studied by Guldberg and Waage in the 1860s). Usually, concentrations $c_i = \frac{n_i}{V}$ instead of molecule numbers n_i are regarded [Gillespie, 1976] and treated as continuous variables [Gillespie, 1977]. This conversion enables the description of the system using intensive (i.e. size-independent) variables and parameters, but the assumption of continuity is only justified in case copy numbers and the system's volume are large. For the elementary reaction (2.4), the deterministic dynamics of c_i are by default described by:

$$\dot{c}_i \;=\; \sum_{j=1}^{R} \left(k_j \left(\gamma_{ij} - \beta_{ij}\right) \prod_{l=1}^{M} c_l^{\beta_{lj}} \right) \;=\; \sum_{j=1}^{R} \left(k_j \, a_{ij} \prod_{l=1}^{M} c_l^{\beta_{lj}} \right). \tag{2.14}$$

k_j is the deterministic reaction rate constant that depends on physical (e.g. thermodynamic) properties of the reaction and of its environment. The relation between k_j and the stochastic constant κ_j can be formulated as:

$$\kappa_j \;=\; k_j \cdot V \cdot \prod_{l=1}^{M} \frac{\beta_{lj}!}{V^{\beta_{lj}}}. \tag{2.15}$$

This is a generalization of the formula derived in [Gillespie, 1977] and it shows that κ_j is an extensive parameter which depends on the volume V. This dependence is in turn determined by stoichiometry in a quite intuitive manner: For a given system state $\mathbf{n}$, zero-order reactions are more probable in large systems, while the chance of molecular collisions that are required for reactions of second or higher order are reduced in larger volumes.

Insertion of Eq (2.15) into Eq (2.5) shows that the j-th reaction propensity can be written as:

$$w_j(\mathbf{n}) = \frac{k_j}{V^{\sum_{l=1}^{M} \beta_{lj} - 1}} \prod_{l=1}^{M} \frac{n_l!}{(n_l - \beta_{lj})!}. \tag{2.16}$$

Furthermore, one can state that for systems with large size and copy numbers,

$$\frac{1}{V} \cdot w_j(\mathbf{n}) \;=\; k_j \prod_{l=1}^{M} \frac{n_l}{V} \left(\frac{n_l - 1}{V}\right) \left(\frac{n_l - 2}{V}\right) ... \left(\frac{n_l - \beta_{lj} + 1}{V}\right) \;=\; k_j \prod_{l=1}^{M} c_l^{\beta_{lj}} \;+\; \mathcal{O}(V^{-1}) \tag{2.17}$$

with $\mathbf{c} = \frac{\mathbf{n}}{V}$. The rate equation (2.14) then emerges in the *thermodynamic limit* $n_i \to \infty$, $V \to \infty$, s.t. $c_i = \frac{n_i}{V} = \text{const.}$:

$$\dot{\mathbf{c}} \;=\; \lim_{V \to \infty} \frac{1}{V} \, \mathbf{A} \, \mathbf{w}(\mathbf{c} \cdot V). \tag{2.18}$$

This relation provides a connection to the stochastic formulation. However, it is nontrivial to determine under which conditions the macroscopic formulation gives a good description of the regarded system. This question will be addressed in Chapters 3 and 4.

2.1.4. Simulating the master equation: the Gillespie algorithm

The *stochastic simulation algorithm* [Gillespie, 1976], also known as the *Gillespie algorithm*, is probably the most prominent method with which the CME can be treated computationally. It is a Monte-Carlo-based approach that generates *exact trajectories* of the CME, instead of providing solutions of the whole distribution $p_{\mathbf{n}}(\tau)$. The basic idea behind it is probably explained best by looking at the algorithm itself:

The Gillespie algorithm (pseudoalgorithm)

- Initialization:
  Initialize the state vector $\mathbf{n}$ and calculate the propensities $w_j(\mathbf{n})$ of all reactions. Initialize time by setting $\tau = 0$.
- Monte-Carlo step:
  Randomly choose a time step $\Delta\tau$ (time until the next reaction event) and a number $J \in \{1, ..., R\}$ (number of the reaction channel that fires next). This is done by drawing two independent, uniformly distributed random variables $U_1 \sim \mathcal{U}([0,1])$ and $U_2 \sim \mathcal{U}([0,1])$, by then calculating

$$\Delta\tau = \frac{1}{\sum_{j=1}^{N} w_j(\mathbf{n})} \cdot \ln\left(\frac{1}{U_1}\right), \tag{2.19}$$

  and by choosing J which satisfies the condition

$$\sum_{j=1}^{J-1} w_j(\mathbf{n}) \;<\; U_2 \cdot \sum_{j=1}^{R} w_j(\mathbf{n}) \;\leq\; \sum_{j=1}^{J} w_j(\mathbf{n}) \tag{2.20}$$

- Update:
  Set $\tau \leftarrow \tau + \Delta\tau$ and $\mathbf{n} \leftarrow \mathbf{n} + \mathbf{a}_J$.
- Iteration:
  Go back to the Monte-Carlo step unless the stopping criterion is fulfilled.

The Monte-Carlo step uses the principle of inverse transform sampling in order to choose the time steps and reaction channels consistently with the CME. Basically, Eq (2.19) shows that the time interval between two subsequent reactions is exponentially distributed, and Eq (2.20) says that the probability of a reaction channel to be the next that fires is proportional to the relative size of its propensity. For a rigorous derivation, see the original publication [Gillespie, 1976]. The algorithm shown above is called the "direct method", in contrast to the "first-reaction method", also presented in [Gillespie, 1976]. The direct method is computationally more efficient, however, the idea of the first-reaction method has been modified later to yield the *next-reaction method* [Gibson and Bruck, 2000]. This version is usually even faster, because part of the information is continually updated so that after initialization, only one random variable needs to be generated per iteration. In spite of their differences in computational efficiency, all the above mentioned approaches yield equivalent outputs. Since in this work, simulations are only used for validating approximate analytical methods, the choice of the algorithm plays a minor role.

2.1.5. Fokker-Planck-type approaches

The Fokker-Planck equation (FPE) is a partial differential equation that describes the evolution of a probability density function P, which is continuous in time *and* in space. Its general formula in the multivariate case with state vector $\mathbf{x}$ reads:

$$\frac{\partial P(\mathbf{x},\tau)}{\partial \tau} = -\sum_{i=1}^{M} \frac{\partial}{\partial x_i}[v_i(\mathbf{x})\, P(\mathbf{x},\tau)] + \frac{1}{2}\sum_{i,k=1}^{M} \frac{\partial^2}{\partial x_i \partial x_k}\,[d_{ik}(\mathbf{x})\, P(\mathbf{x},\tau)]. \tag{2.21}$$

v_i and d_{ik} must be real-valued, differentiable functions and for every $\mathbf{x}$, $d_{ik}(\mathbf{x})$ needs to be symmetric and positive definite. $\mathbf{v}(\mathbf{x}) := (v_i(\mathbf{x}))_{i=1,\dots,M}$ is called *drift vector* and $\mathbf{D}(\mathbf{x}) := (d_{ik}(\mathbf{x}))_{i,k=1,\dots,M}$ is referred to as *diffusion tensor*.

A special case of Eq (2.21) is the *linear* FPE[4], which possesses a drift term that is linear in $\mathbf{x}$ and a constant diffusion tensor and which therefore has the form:

$$\frac{\partial P(\mathbf{x},\tau)}{\partial \tau} = -\sum_{i,k=1}^{M} v_{ik} \frac{\partial}{\partial x_i}[x_k\, P(\mathbf{x},\tau)] + \frac{1}{2}\sum_{i,k=1}^{M} d_{ik}\, \frac{\partial^2}{\partial x_i \partial x_k}\, P(\mathbf{x},\tau). \tag{2.22}$$

It turns out that the solution of the linear FPE is a multivariate Gaussian distribution for every τ, which is fully characterized by its current mean and variance [van Kampen, 2007].

The FPE formalism is often used to approximate the CME. In the following, two common types of FPEs are presented. The first type, sometimes referred to as the *chemical* FPE (CFPE) [Grima et al., 2011], results from the *Kramers-Moyal expansion* of the CME. The second, called the *linear noise approximation* (LNA), has been proposed in [van Kampen, 2007]. As the LNA is used later in this study, a stronger focus will be put on it. Afterwards, a brief look is taken at some critical reviews and comparisons of the two approaches.

The chemical Fokker-Plank equation

Broadly speaking, the Kramers-Moyal expansion of the CME (2.3) is obtained by performing a multivariate Taylor expansion with respect to the state variable $\mathbf{n}$. This is only possible if $\mathbf{n}$ is regarded as continuous vector and if the domain of the propensities w_j is extended to $\mathbb{R}^{\mathbb{M}}$ accordingly. Additionally, the probability mass function $p_{\cdot}(\tau)$ needs to be replaced by a probability density function $P(.,\tau)$. All these steps are obviously critical if molecular copy numbers are small. The Kramers-Moyal expansion is therefore usually classified as a "large number approximation"[5]. The CFPE is obtained by truncating the Kramers-Moyal expansion at second order: If the function $w_j(.) \cdot P(.,\tau)$ is smooth, its Taylor expansion to second order yields

[4] Note that in this context, the term *linear* does not refer to linearity in P (which holds per definition for every FPE), but defines properties of v_i and d_{ik}.

[5] In the context of the Kramers-Moyal expansion and of FPEs, copy numbers are often converted to concentrations, which are then assumed to be continuous. This assumption is reasonable in the case of a large system size (concerning volume and copy numbers), but otherwise, the justification of performing interpolations remains equally problematic after conversion as in the original state space. We thus refrain from any conversion at this stage.

$$
\begin{aligned}
& w_j(\mathbf{n}-\mathbf{a}_j)\ P(\mathbf{n}-\mathbf{a}_j,\tau) \\
\approx\ & w_j(\mathbf{n})\ P(\mathbf{n},\tau)\ -\ \sum_{i=1}^{M}\frac{\partial}{\partial n_i}(a_{ij}\,w_j(\mathbf{n})\,P(\mathbf{n},\tau))\ +\ \frac{1}{2}\sum_{i,k=1}^{M}\frac{\partial^2}{\partial n_i\partial n_k}(a_{ij}a_{kj}\,w_j(\mathbf{n})\,P(\mathbf{n},\tau)).
\end{aligned}
\tag{2.23}
$$

Inserting this expression into the CME (2.3) results in the CFPE

$$
\frac{\partial}{\partial\tau}P(\mathbf{n},\tau)\approx-\sum_{i=1}^{M}\frac{\partial}{\partial n_i}\left(\sum_{j=1}^{R}a_{ij}\,w_j(\mathbf{n})\,P(\mathbf{n},\tau)\right)+\frac{1}{2}\sum_{i,k=1}^{M}\frac{\partial^2}{\partial n_i\partial n_k}\left(\sum_{j=1}^{R}a_{ij}a_{kj}\,w_j(\mathbf{n})\,P(\mathbf{n},\tau)\right)
\tag{2.24}
$$

(cf. [Gillespie, 1996, 2000]). Compared with the general formulation of the FPE (2.21), $v_i(\mathbf{n}) = \sum_{j=1}^{R} a_{ij}\,w_j(\mathbf{n})$ and $d_{ik}(\mathbf{n}) = \sum_{j=1}^{R} a_{ij}a_{kj}\,w_j(\mathbf{n})$. The main criticism against the CFPE concerns the physical justification of the performed truncation.

The linear noise approximation

The LNA is another large number approximation. Its derivation by van Kampen is based on a Taylor series expansion, the so-called *system size expansion*, which explicitly includes the size of the system V as a parameter. Truncation of the Taylor series is justified if V is large. The LNA is built on the assumption that the state vector fluctuates stochastically around a macroscopic solution $\boldsymbol{\phi}(\tau)\cdot V$, and that the fluctuations are in the order of $V^{\frac{1}{2}}$. The random state vector can hence be decomposed into a deterministic and a stochastic part and can be written as:

$$
\mathbf{N}(\tau)=\boldsymbol{\phi}(\tau)\cdot V+\mathbf{Y}(\tau)\cdot V^{\frac{1}{2}}.
\tag{2.25}
$$

$\mathbf{N}$, $\boldsymbol{\phi}$ and $\mathbf{Y}$ are treated as continuous, and $\boldsymbol{\phi}$ solves the macroscopic RE (2.18):

$$
\dot{\boldsymbol{\phi}}=\mathbf{A}\tilde{\mathbf{w}}(\boldsymbol{\phi}),
\tag{2.26}
$$

where $\tilde{\mathbf{w}}(\boldsymbol{\phi}) := \lim_{V\to\infty}\frac{1}{V}\,\mathbf{w}(\boldsymbol{\phi}\cdot V)$. According to Eq (2.17), $\tilde{\mathbf{w}}(\boldsymbol{\phi}) = \frac{1}{V}\,\mathbf{w}(\boldsymbol{\phi}\cdot V)+\mathcal{O}(V^{-1})$.

The whole stochasticity of $\mathbf{N}$ is now contained in $\mathbf{Y}$, which basically captures the deviation of the original stochastic process from the deterministic solution. Its probability density function $\tilde{P}$ is connected to that of $\mathbf{N}$ via the relation $P(\mathbf{n},\tau) = P(\boldsymbol{\phi}\cdot V+\mathbf{y}\cdot V^{\frac{1}{2}},\tau) =: \tilde{P}(\mathbf{y},\tau)$, where in the first step, Eq (2.25) has been used. In [van Kampen, 2007], the derivation of the LNA is shown in detail. In the Supporting Information of [Elf and Ehrenberg, 2003], a concise and explicit derivation for the multivariate case is given. Here, only the main steps are summarized:

First, the variant of the CME for $\mathbf{N}$ which is continuous in time and space is transformed into a CME in terms of $\mathbf{Y}$. Taylor expansion of $\mathbf{w}(\mathbf{n})$ to second order around $\boldsymbol{\phi}\cdot V$, which corresponds to an approximation up to order V^0, eventually yields the LNA:

$$
\frac{\partial\tilde{P}(\mathbf{y},\tau)}{\partial\tau}\ =\ -\sum_{i,k=1}^{M}\sum_{j=1}^{R}a_{ij}\,\frac{\partial}{\partial\phi_k}\tilde{w}_j(\boldsymbol{\phi})\,\frac{\partial}{\partial y_i}[y_k\,\tilde{P}(\mathbf{y},\tau)]\ +\ \frac{1}{2}\sum_{i,k=1}^{M}\sum_{j=1}^{R}a_{ij}a_{kj}\tilde{w}_j(\boldsymbol{\phi})\,\frac{\partial^2}{\partial y_i\partial y_k}\tilde{P}(\mathbf{y},\tau).
\tag{2.27}
$$

This is a linear FPE, whose general form was given in Eq (2.22). Here, $v_{ik} = \sum_{j=1}^{R} a_{ij} \frac{\partial}{\partial \phi_k} \tilde{w}_j(\boldsymbol{\phi}) = [\frac{\partial}{\partial \boldsymbol{\phi}} \mathbf{A}\tilde{\mathbf{w}}(\boldsymbol{\phi})]_{ik}$ is exactly the (i,k)-th entry of the Jacobian (cf. Section 2.3.1) of the macroscopic rate equation (2.26) and $d_{ik} = \sum_{j=1}^{R} a_{ij} a_{kj} \tilde{w}_j(\boldsymbol{\phi}) = [\mathbf{A} \cdot \mathrm{diag}(\tilde{\mathbf{w}}(\boldsymbol{\phi})) \cdot \mathbf{A}^\top]_{ik}$. Note that these coefficients are time-variant because of their dependence on $\boldsymbol{\phi}$. The solution is Gaussian distributed at any point in time and therefore fully characterized by the dynamics of its first two moments.

The expected value and variance of this Gaussian distribution in its stationary state can be analytically determined. One can show that $\mathbb{E}[\mathbf{Y}] = \mathbf{0}$ and that $\tilde{\boldsymbol{\Sigma}} := \mathrm{Var}(\mathbf{Y})$ obeys the so-called *Lyapunov equation*

$$\mathbf{J}^* \cdot \tilde{\boldsymbol{\Sigma}} + \tilde{\boldsymbol{\Sigma}} \cdot (\mathbf{J}^*)^\top + \tilde{\mathbf{D}}^* = \mathbf{0}, \tag{2.28}$$

where $\mathbf{J}^* := \frac{\partial}{\partial \boldsymbol{\phi}} \mathbf{A}\tilde{\mathbf{w}}(\boldsymbol{\phi^*})$ is the Jacobian of the macroscopic ODE evaluated at its stable fixed point $\boldsymbol{\phi}^*$, and $\tilde{\mathbf{D}}^* := \mathbf{A} \cdot \mathrm{diag}(\tilde{\mathbf{w}}(\boldsymbol{\phi}^*)) \cdot \mathbf{A}^\top$. The variance of the original state vector $\mathbf{N}$ is then given by $\boldsymbol{\Sigma} := \mathrm{Var}(\mathbf{N}) = V \cdot \mathrm{Var}(\mathbf{Y})$. It solves the Lyapunov equation

$$\mathbf{J}^* \cdot \boldsymbol{\Sigma} + \boldsymbol{\Sigma} \cdot (\mathbf{J}^*)^\top + \mathbf{D}^* = \mathbf{0} \tag{2.29}$$

with $\mathbf{D}^* = \tilde{\mathbf{D}}^* \cdot V$. The Lyapunov equation can always be solved, although the process might be tedious, depending on the complexity of the system.

Eq (2.27) shows that the reaction rates only enter the LNA in a form that is linearized around the deterministic solution (i.e. only in form of their zeroth and first derivatives evaluated at $\boldsymbol{\phi}$). This allows the formulation and interpretation of the stationary LNA from a different perspective: Since systems with nonlinear propensities are in general very difficult to analyze, the idea is to linearize them. The point around which this linearization is performed is chosen as the macroscopic fixed point $\boldsymbol{\phi}^* \cdot V$, which serves as an approximation of the stationary mean value of $\mathbf{N}$. The variance of the linearized system $\boldsymbol{\Sigma}$, which can be determined analytically, is then used as an approximation of the variance of the original, nonlinear system. Performing the according calculation steps automatically ends up in the Lyapunov equation (2.29) [Tomioka et al., 2004].

The stationary LNA returns the correct mean and variance of $\mathbf{N}$ if all reaction propensities are linear functions of the system state. However, unlike the distribution obtained with the LNA, the real distribution is usually only Gaussian in the thermodynamic limit (cf. [Walczak et al., 2012][6]).

Comparison of the approaches

The CFPE with its nonlinear drift term is apparently more complex from a mathematical point of view than the LNA. In fact, it can be shown that linearization of the propensity functions in the CFPE exactly yields the LNA formulation [Gardiner, 2009; Wallace et al., 2012]. While it is always possible to solve the LNA, this is not the case for the CFPE. However, it would be wrong to immediately conclude that the CFPE is more *accurate* than the LNA:

For example, van Kampen has remarked that the truncation of the Kramers-Moyal expansion,

[6] In this paper, the fact is illustrated using a simple birth-death process that results in a Poisson distributed state space. The mean and variance is correctly predicted by LNA, independently of the size of the system. For large protein numbers, the Poisson distribution approximates a Gaussian.

which results in the CFPE, does not have any physical foundation due to the lack of a small system parameter that would make higher orders of the Kramers-Moyal expansion negligible [van Kampen, 2007]. In contrast to that, the LNA mathematically emerges from a Taylor expansion with respect to $V^{-\frac{1}{2}}$, showing that its results may be considered reliable if the LNA is applied, e.g., to chemical processes in large, well-stirred reaction vessels.

However, van Kampen's reasoning did not put an end to the discussion: Conditions were identified under which the CFPE (and the related chemical Langevin equation) does very well have a physical meaning [Gillespie, 2000; Wallace et al., 2012]: For example, in case a time interval can be defined during which every reaction channel is expected to fire several times without any propensity changing its value significantly, the CFPE is shown to be a valid approximation. A large system size usually helps to fulfill this condition, but it is not a necessary criterion.

For modeling intracellular processes, one should refrain from concentrating too much on scenarios close to the thermodynamic limit anyway, as the size of the system is naturally limited[7,8]. Instead, one should assess how good the approximate approaches perform in smaller-scale systems. Under this aspect, R. Grima *et al.* have compared the accuracy of the CFPE and of the LNA in monostable systems [Grima et al., 2011]. They have shown that the CFPE estimates the mean and variance of a distribution accurately to order $V^{-\frac{3}{2}}$. The estimation of the variance using the LNA has the same accuracy, while the mean is only predicted accurately to order $V^{-\frac{1}{2}}$, even in cases where the applicability of the CFPE cannot be rigorously justified. This is one important example of recent papers that have studied the reliability of the two approaches in systems with limited volume. Other considerations, for example concerning thermodynamic aspects, can be found in [Ceccato and Frezzato, 2018; Horowitz, 2015].

For the intended goal of this study, it is important to obtain closed-form solutions of the probability distribution or, at least, of its mean and variance, as with their help, parametric connections between the properties of a system and its noise level can be found. Therefore, the usage of the LNA is preferred, but one needs to bear in mind that a critical evaluation of the quality of results is absolutely necessary.

2.1.6. Other approaches

In this section, we will briefly mention some further approximate approaches that facilitate the analysis of stochastic systems. Some of them are used in this work, some of them will be revisited later for the discussion of the methodology that will be newly developed.

Higher orders of the system size expansion

Through the inclusion of more terms than those used for LNA, mean and variance estimates can be improved, which becomes relevant in the mesoscopic regime. At least, this statement holds for unimodal systems (the multimodal case will be discussed at the end of the chapter). This idea has been pursued by Grima and colleagues, who have incorporated terms up to order $V^{-\frac{1}{2}}$ to obtain the so-called *effective mesoscopic rate equations* (EMREs). They give a more accurate description of average concentration dynamics than the common REs [Grima, 2010] by adding a systematically derived mesoscopic correction term. The term vanishes in the limit of large volumes. In an analogous manner, an improved estimation of the variance-covariance

[7] The volume of an *Escherichia coli* (*E. coli*) cell comprises about 10^{-15} liters [Scott, 2006].
[8] Remark: In the thermodynamic limit, the CFPE becomes equal to the LNA [Grima et al., 2011].

matrix has been achieved by including terms up to order V^{-1}. This approach has been denoted the *Inverse Omega Square* (IOS) method [Thomas et al., 2012].

Moment closure techniques

Like the EMRE approach and the IOS method, moment closure techniques focus on the estimation of (central) moments (means, variances, etc.) of a distribution. As already mentioned in Section (2.1.2), the exact dynamics of the k-th moment can only be formulated in dependence of the $(k+1)$-st moment, unless the system comprises zero- or first-order reactions only. Moment closure techniques assume that moments higher than a certain order can approximately be formulated in terms of lower moments only. The formulations usually follow the example of prominent distributions like the normal distribution, where all cumulants except for the first and second (mean and variance) are simply equal to zero [Gómez-Uribe and Verghese, 2007]. This artificial limitation of the number of moment equations makes them analytically or numerically solvable. If moments higher than the first are included in the moment closure technique, the expression obtained for the average systems behavior usually differs from the corresponding RE formulation, where the impact of fluctuations is fully neglected. Again, however, major problems arise when moment closure techniques are applied to multimodal systems [Gómez-Uribe and Verghese, 2007; Lakatos et al., 2015; Schnoerr et al., 2015]. Moreover, they are often said to constitute an *ad hoc* approximation, whose accuracy cannot be reliably predicted [Grima, 2012; Schnoerr et al., 2015].

Model reductions based on separation of scales

The following approaches split the state variables into at least two groups: One that needs to be modeled in detail, while the other can be described in a simplified manner, thereby reducing the overall complexity of the model. The criteria used for the classification of the variables are related to the presence of different *scales*.

Model reduction based on time-scale separation is a concept well-known from ODE modeling: If the dynamics of certain components within the system are much faster than the dynamics of all other components, one can regard the two time-scales separately. On the "fast" time-scale, all slow components can be assumed to have constant values, as they do not significantly change while the fast subsystem relaxes to its *pseudo steady state* (PSS). On the "slow" timescale, all fast components can be expressed as functions of the slow ones and therefore be eliminated, since they are assumed to adapt infinitely quickly to their PSS that is dictated by the dynamics of the slow variables. In [Rao and Arkin, 2003], the same principle was applied to stochastic systems. The elimination of fast variables (through conditioning on slow variables and subsequent averaging) leads to a system consisting of non-elementary reactions, whose stochastic reaction "constants" κ_j might actually be complex functions of the slow system variables. They might, e.g., have the form of typical Michaelis-Menten or Hill-type kinetics. This is of great value in the modeling of biochemical networks that are too complex to be described in terms of elementary reactions. Similar approaches can be found in [Cao et al., 2005; Haseltine and Rawlings, 2005; Kepler and Elston, 2001].

The method proposed in [Elf and Ehrenberg, 2003] first eliminates fast variables by setting them into their macroscopic PSS, then applies the LNA to the slow variables only. Interestingly, all these variables do not necessarily describe the original reaction components, but transformed

versions thereof (so-called relaxation modes), which can be better grouped according to their time-scales.

Another kind of model reduction exploits the molecular abundance of some components. In [Lin and Galla, 2016], a piecewise deterministic Markov process was developed, in which the dynamics of mRNA (low copy number species) are modeled as a stochastic switching process, while the protein level is described deterministically using ODEs. The relevant set of ODEs changes in accordance with the current state of the mRNA.

In a similar way, slow switches of a low copy number species (here: promoter state) are modeled stochastically in [Thomas et al., 2014]. The high-copy number species (mRNA and protein) are then described with the help of stationary LNAs, which are conditioned on the promoter state. In contrast to the previously mentioned approach in [Lin and Galla, 2016], an approximate stochastic distribution of the full system state can thus be obtained.

An approach where the description of components with small noise is reduced to the dynamics of their mean values, whereas the other variables are simulated in detail using the Gillespie algorithm, was proposed and discussed in [Hellander and Lötstedt, 2007; Jahnke, 2011].

2.2. Noise in genetic circuits

Gene expression is a highly stochastic process due to the involvement of reactive species that occur in very low copy numbers. The amount of active promoters of a gene of interest usually ranges between zero and eight per bacterial cell, and the abundance of the corresponding mRNA often lies in the same order of magnitude [Milo et al., 2016]. Therefore, mRNA and protein time courses typically show significant fluctuations, which influence the robustness of a circuit and generate variability within isogenic populations. These fluctuations can be characterized mathematically with the help of stochastic descriptions, some of which have been introduced in the preceding section.

2.2.1. Measures of variability

Until now, we have discussed how the variance of fluctuations can be determined at least approximately. A better appraisal of variability can be achieved when the variance is related to the mean value of fluctuations. One common measure with which the noise of a stochastic process N is quantified is the so-called *Fano factor* (or *noise strength*)

$$\eta(N) := \frac{\mathrm{Var}(N)}{\mathbb{E}[N]}. \tag{2.30}$$

The Fano factor enables to compare the distribution of N with a Poisson distribution, where the variance is always equal to the mean and therefore $\eta(N) = 1$. Typical birth-death processes are Poisson distributed, e.g. the simple formation and degradation of mRNA molecules, if promoter dynamics are neglected and other regulatory mechanisms are absent. All deviations from $\eta(N) = 1$ therefore indicate some kind of upstream stochasticity or the involvement of regulation. Processes with $\eta(N) < 1$ are called sub-Poissonian, while processes with $\eta(N) > 1$ are referred to as super-Poissonian.

An alternative measure is the *coefficient of variation* (CV):

$$\mathrm{CV}(N) := \frac{\sigma(N)}{\mathbb{E}[N]}, \tag{2.31}$$

where $\sigma(N) := \sqrt{\mathrm{Var}(N)}$ is the standard deviation of N.

There are different opinions about which measure is used best. While the Fano factor is often favored due to its relation to Poisson processes [Chong et al., 2014; Elf and Ehrenberg, 2003; Thattai and van Oudenaarden, 2001; Thomas et al., 2013], the CV is said to be better suited to identify contributions from different noise sources [Bar-Even et al., 2006; Paulsson, 2005]. The CV is furthermore regarded as a "dimensionless" [Scott, 2006], "unambiguous" [Kærn et al., 2005] measure, accounting for the fact that fluctuations become negligible in large systems with high molecular abundance. The discussion in [Kærn et al., 2005] implies that the CV is better suited to compare noise levels in systems of differing size but with the same average concentration of the target component. The Fano factor is more appropriate for comparing fluctuations around different mean values within the same system and is therefore valuable for studying phenotypic heterogeneity. Besides this aspect, the Fano factor better describes the influence of translational bursts [Kærn et al., 2005; Ozbudak et al., 2002; Thattai and van Oudenaarden, 2001], which will be introduced in the following section. As the present study will mainly deal with heterogeneous systems and with the role of bursts, the Fano factor is preferred.

2.2.2. Translational bursts

Translational bursts (or *protein bursts*) emerge when several protein molecules are translated from a single, short-lived mRNA molecule [Friedman et al., 2006] and constitute a major source of gene expression noise. They can be quantified using the *burst size* and the *burst frequency*: The burst size is the average number of proteins synthesized from a transcript. The burst frequency is the average number of translational bursts (i.e. of mRNA transcription events) per protein lifetime. The scaling of the burst frequency with the average lifespan of a protein ensures that the average protein level is now given by the product of burst size and burst frequency. In order to see how these measures can be calculated, let us have a look at the reaction scheme of linear gene expression in Fig 2.1 (A) (cf. [Thattai and van Oudenaarden, 2001]). Here, m and s denote the copy numbers of the regarded mRNA and protein species in a single cell, respectively. Transcription occurs with constant rate $\hat{a}$, while the translational propensity is proportional to the mRNA level with proportionality constant $\hat{g}$. The degradation of mRNA and protein is assumed to be linear as well with constants d_m and d_s, respectively. The average burst size α and the average burst frequency ω are then given by [Thattai and van Oudenaarden, 2001]:

$$\alpha := \frac{\hat{g}}{d_m} \qquad \text{and} \qquad \omega := \frac{\hat{a}}{d_s}. \tag{2.32}$$

One can verify that the average protein level is indeed equal to the product of α and ω. In Section 4.5, the definitions are extended to more complex reaction schemes. Based on the derivations shown there, it is also possible to reconstruct the emergence of the formulae in (2.32).

If the lifespan of a transcript, which determines the duration of a burst, was short enough to neglect the degradation of proteins in that time interval (i.e. if $d_s \ll d_m$, which is typically the case [Milo et al., 2016]), the burst size α would correspond to the average amplitude of protein peaks that are generated through translational bursting [Thattai and van Oudenaarden, 2001],

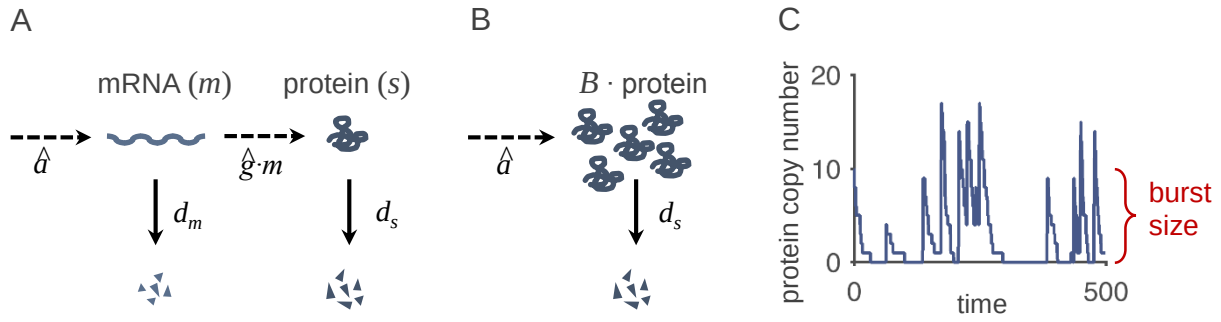


Figure 2.1.: **Protein bursts in a linear scheme of gene expression.** (A) Full reaction scheme of a linear gene expression model. Interactions are indicated by dashed arrows, conversions by solid arrows. (B) Reduced scheme exploiting protein burst dynamics. The number of protein molecules produced per burst equals B, which is a geometrically distributed variable whose mean corresponds to the burst size. (C) Protein bursts form peaks in a protein time course, which was obtained through a stochastic simulation of the full reaction scheme. The average amplitude of the peaks is approximately equal to the burst size, in this case $\alpha = 10$. The parameters are: $\hat{a} = 0.02$, $\hat{g} = 10$, $d_m = 1$, $d_s = 0.1$.

see Fig 2.1 (C). The actual number of translated proteins per mRNA is a random variable, which is denoted by B. In [McAdams and Arkin, 1997], it has been theoretically shown that its distribution is geometric, which means that the probability p_b^B of having a burst of size b is equal to

$$p_b^B = \left(\frac{\alpha}{1+\alpha}\right)^b \cdot \left(\frac{1}{1+\alpha}\right). \tag{2.33}$$

The impact of translational bursting on the Fano factor has been demonstrated mathematically in [Thattai and van Oudenaarden, 2001]: According to the linear expression model in Fig 2.1 (A), where promoter dynamics are neglected, the stationary mRNA distribution is Poissonian. Therefore, $\eta(M) = 1$, where M is the random mRNA copy number. The protein Fano factor $\eta(S) = \frac{\mathrm{Var}(S)}{\mathbb{E}[S]}$ can be calculated based on the CME, following the approach from Section 2.1.2. It is possible in this case to obtain an explicit solution, as all reaction propensities are constant or linear. The result is:

$$\eta(S) = 1 + \frac{\alpha}{1 + \frac{d_s}{d_m}} \approx 1 + \alpha. \tag{2.34}$$

The given approximation is again valid if protein degradation is much slower than mRNA degradation. The formula shows that the protein distribution is super-Poissonian and that the burst size has substantial influence on the Fano factor. This fact has an interesting consequence: If the transcription rate was decreased and if at the same time, the translation rate was increased, so that the average protein level is kept constant, the protein noise level would be amplified [Kærn et al., 2005; Kierzek et al., 2001; Thattai and van Oudenaarden, 2001]. This manipulation causes an increase in the burst size and a corresponding decrease in the burst frequency. Remarkably, this theoretical fact could be verified experimentally by varying inductor concentrations to modulate transcription initiation rates and by altering either the ribosomal binding site or the initiation codon to change the translational efficiency in a linear expression system [Ozbudak et al., 2002]. Through a similar manipulation of a bistable system that regulates genetic competence in *Bacillus subtilis* (*B. subtilis*), it has been shown that artificial burst size reduction led to fewer random cell-state transitions and hence to fewer entries of cells into the

competent state [Maamar et al., 2007].

Even higher experimental resolution was achieved in [Cai et al., 2006] using microfluidic devices, which allowed the tracking of protein production events in single cells at single molecule resolution. Indeed, protein bursts could be observed, and the geometric distribution of B could be confirmed[9].

Based on these burst statistics and provided that mRNA half-life is so short that protein translations from one mRNA molecule can be lumped into a single burst event, a reduction of the full model in Fig 2.1 (A) can be performed, which is illustrated in Fig 2.1 (B). Here, the mRNA level is eliminated, and each protein production event is assumed to result in a geometrically distributed increase in protein molecules. Based on this model reduction, the probability distribution of the protein copy number could be explicitly calculated from the corresponding CME and it was shown to follow a negative binomial distribution [Aquino et al., 2012; Paulsson and Ehrenberg, 2000; Shahrezaei and Swain, 2008] (or, depending on model formulation, a Gamma distribution [Friedman et al., 2006], which is the continuous equivalent).

A mechanism equivalent to translational bursting that originates from the interplay between promoter and mRNA dynamics can lead to the emergence of mRNA bursts. This mechanism is more common in eukaryotes, while protein bursts are predominant in prokaryotes [Kærn et al., 2005]. In this work, a stronger focus will be on prokaryotic systems and therefore on protein bursts, but the mathematical tools will be transferable to the mRNA level.

Ever since the impact of bursts on gene expression noise has been recognized, their biological role has been intensely discussed. It has for example been postulated that inefficient translation, although being energetically unfavourable, is a strategy of reducing fluctuations in cyclic adenosine monophosphate (cAMP) in *E. coli* [Ozbudak et al., 2002]. However, in other contexts, noise from bursts might be exploited to generate population heterogeneity like in the lysis-lysogeny decision of the bacteriophage λ [Ozbudak et al., 2002], or in the previously mentioned development of competence in *B. subtilis* [Maamar et al., 2007]. In theses cases, the translational efficiency is observed to be high, leading to large protein fluctuations. However, bursting is not the only mechanism that predefines the magnitude of noise: The topology of a circuit as well as the distribution of reaction time scales are further important determinants that will be addressed next.

2.2.3. The influence of circuit topology and dynamics

As we have seen, the combination of different time-scales influences the emergence of noise through the generation or attenuation of bursts. The same principle holds during further propagation of noise through the gene expression system. In general, slow downstream reactions can act as low-pass filters, partially averaging out fast fluctuations of upstream components [Kærn et al., 2005; Paulsson, 2004; Pedraza and van Oudenaarden, 2005].

Another efficient way of noise reduction is negative feedback regulation. In [Becskei and Serrano, 2000], this fact has been demonstrated experimentally by comparing noise levels of systems with and without negative regulation. In accordance with this, the theoretical protein Fano factor was shown to be lower in negatively autoregulated systems than in unregulated ones [Paulsson, 2004; Simpson et al., 2003; Singh, 2011; Thattai and van Oudenaarden, 2001].

[9] Actually in this publication, the burst size was determined to be exponentially distributed. The exponential distribution is the continuous equivalent of the geometric distribution, which is why the observation made in [Cai et al., 2006] is in accordance with the aforementioned calculation in [McAdams and Arkin, 1997].

The opposite holds for positive feedback regulation: It increases noise and may additionally generate bistability, potentially leading to bimodal protein distributions and to transitions between different expression states (a closer look at this fact will be taken in Section 2.3). Therefore, positive autoregulation often plays a role in the generation of subpopulations and in cellular decision making [Kærn et al., 2005; Tsimring, 2014]. Typical, well-studied examples are the *lac* operon in *E. coli* and the development of competence in *B. subtilis* [Eldar and Elowitz, 2010; Süel et al., 2006].

The connection of multiple topological motifs and different time-scales can lead to more complex and fine-tuned noise regulation [Brandman et al., 2005; Cagatay et al., 2009; Hornung and Barkai, 2008; Huang et al., 2014].

2.2.4. Intrinsic vs. extrinsic noise

The stochasticity of a gene expression system is not only determined by the randomness of its reactions, but there are further factors that may influence its noise: For example, other intracellular components (like metabolites or global regulators) can transfer their fluctuations to the circuit via direct or indirect interactions. Moreover, the stochasticity of cellular dynamics like growth, chromosomal replication, or cell division is a source of noise which affects all cellular processes. Finally, even environmental fluctuations may be propagated into the cell.

All these noise sources can be classified as either being *intrinsic* or *extrinsic*, depending on whether they lie inside the borders of the regarded system or not. Since these borders can be defined individually (e.g. around a specific genetic circuit [Swain et al., 2002], or identical to the cell membrane [González-Cabaleiro et al., 2017]), a uniform classification does not exist [Huang, 2009; Kærn et al., 2005; Paulsson, 2005].

Fluctuations that are experimentally measured are always superpositions of different noise contributions. In order to differentiate between them, interesting experimental approaches exist, which are then usually supported by and validated with statistical analysis and mathematical models. For example, M. Elowitz *et al.* used two different fluorophores under the control of the same regulatory sequence to discriminate gene expression-specific noise from extrinsic noise sources [Elowitz et al., 2002]. This discrimination was based on a comparison of fluctuation patterns of the two fluorophores, coupled with statistical correlation studies.

In this work, we focus on intrinsic noise that is generated and propagated within the regarded genetic circuit. All other sources of noise are considered extrinsic. In the final outlook, the incorporation of cellular dynamics will shortly be discussed.

2.3. Deterministic vs. stochastic descriptions of heterogeneity

Until now, we have not paid much attention to the presence or absence of phenotypic heterogeneity. This topic will be considered here, both from a deterministic as well as from a stochastic perspective. More specifically, the concepts of “bistability” and “bimodality” are introduced. It will then be shown which impact phenotypic heterogeneity may have on the applicability and on the reliability of stochastic modeling approaches.

2.3.1. Deterministic bistability

Deterministic models ignore all the stochasticity within a system. Probably the most prominent approach in this field, with which temporal, but not spatial dynamics can be described, consists in a system of autonomous[10] ordinary differential equations of first order, which has the general explicit form[11]

$$\dot{\mathbf{c}} = \mathbf{f}(\mathbf{c}). \tag{2.35}$$

Here, $\mathbf{c} \in \mathbb{D} \subset \mathbb{R}^M$ is a vector of state variables and $\mathbf{f} : \mathbb{D} \to \mathbb{R}^M$ is a vector-valued function that specifies the dynamics of the system. Additionally, we assume that the *initial condition* of the ODE system is given by $\mathbf{c}(0) = \mathbf{c}_0$. If $\mathbf{f}$ fulfills certain requirements[12], the existence and uniqueness of a solution of the ODEs is guaranteed, which means that future dynamics are unequivocally determined by present conditions. Obtaining an explicit analytical solution for the system of ODEs might, however, be tedious or even impossible, depending on the structure of $\mathbf{f}$. In many cases, though, one is only interested in the long-time behavior of a system, which can be studied using dynamical systems theory. The most important definitions and principles in this context are the following (see [Guckenheimer and Holmes, 1983] for details):

A *stationary state* (or *steady state*, *fixed point*) $\mathbf{c}^*$ is defined by the condition $\dot{\mathbf{c}} = \mathbf{0}$, i.e. it is a state that is never left once reached. According to Eq (2.35), the stationary states are the roots of $\mathbf{f}$. One of their most important characteristics is their stability, which tells us how the system qualitatively behaves in a domain around the fixed point, e.g. after a (sufficiently small) perturbation:

The steady state $\mathbf{c}^*$ is called *stable*, if for any neighborhood $U_0 \subset \mathbb{D}$, there exists a neighborhood $U_1 \subset U_0$ such that every solution $\mathbf{c}(\tau)$ with $\mathbf{c}_0 \in U_1$ is defined and $\mathbf{c}(\tau) \in U_0 \quad \forall\, \tau > 0$. In case U_1 can be chosen such that additionally, $\lim_{\tau\to\infty} \mathbf{c}(\tau) = \mathbf{c}^*$, the fixed point is called *asymptotically stable*. A fixed point is called *unstable*, if it is not stable.

The stability of a fixed point $\mathbf{c}^*$ can in most cases be determined using the Jacobian matrix $\mathbf{J}^*$ of the right hand-side of Eq (2.35), evaluated at $\mathbf{c}^*$:

$$\mathbf{J}^* := \frac{\partial \mathbf{f}}{\partial \mathbf{c}}\Big|_{\mathbf{c}=\mathbf{c}^*} = \begin{pmatrix} \frac{\partial f_1}{\partial c_1} & \frac{\partial f_1}{\partial c_2} & \cdots & \frac{\partial f_1}{\partial c_M} \\ \frac{\partial f_2}{\partial c_1} & \frac{\partial f_2}{\partial c_2} & \cdots & \frac{\partial f_2}{\partial c_M} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_M}{\partial c_1} & \frac{\partial f_M}{\partial c_2} & \cdots & \frac{\partial f_M}{\partial c_M} \end{pmatrix}\Bigg|_{\mathbf{c}=\mathbf{c}^*} \tag{2.36}$$

In case the real parts of the eigenvalues of the Jacobian are all negative, $\mathbf{c}^*$ is asymptotically stable. If at least one of the eigenvalues has a positive real part, it is unstable[13]. This finally brings us to the definition of *bistability*, which denotes the presence of exactly two asymptotically stable stationary states. This concept is generalized by *multistability*, referring to the occurrence of more than one asymptotically stable fixed point. As already indicated, it only depends on

[10] The term *autonomous* refers to the fact that the function $\mathbf{f}$ in Eq (2.35) does not explicitly depend on τ.

[11] This is a generalized form of Eq (2.18), which has described processes consisting of elementary reactions only.

[12] cf. the central theorems by Peano and Picard-Lindelöf. In the following, we assume that $\mathbf{f}$ is globally Lipschitz continuous, from which follows the existence and uniqueness of the solution [Murray and Miller, 1976].

[13] The given Jacobian-based criterion originates from a linearization of the differential equation (2.35). In case $\mathbf{f}$ is nonlinear and the largest real part of eigenvalues is equal to zero, no direct statement can be made about the stability of the fixed point, as it then depends on higher derivatives of $\mathbf{f}$.

the initial condition which of the two steady states will be assumed in the long run. Transitions between them are excluded, unless a suitable change in parameter values occurs (cf. the theory of bifurcations), which is driven by extrinsic forces.

From a fully deterministic viewpoint, a bistable genetic circuit (i.e. a circuit with two stable gene expression states like the one shown in Fig 1.1) would not be able to generate population heterogeneity under uniform environmental conditions, since all cells would always behave in a completely synchronized fashion.

2.3.2. Stochastic bimodality

We already know that in a stochastic framework, the future behavior of a system is not uniquely determined by present conditions. As a consequence, the solution of a stochastic model with a given initial condition cannot be described by a single trajectory (like in the deterministic case); instead, it is usually given in the form of the dynamic probability distribution p over the whole state space. *Bimodality* means that this probability distribution possesses two maxima (also called *modes*), each representing a preferred system state, and is therefore an indication of true phenotypic heterogeneity. Random transitions between the phenotypes are in this case not excluded. Unfortunately, simple mathematical criteria for the detection of bimodality are lacking, in contrast to the methods available for the treatment of bistability.

Bimodality (or, in general, multimodality) is not only a theoretical trait, it can also be observed experimentally, using suitable measurement techniques: If, for example, a target protein can be quantified experimentally on a single-cell level in a large enough population, the measured distribution should resemble the probability distribution obtained by a suitable stochastic description (cf. the frequentist interpretation of probabilities). Instead of regarding the steady-state distribution across a whole population at one specific point in time, one might also consider the distribution in one cell over a long enough time period in order to detect heterogeneity. Based on the assumption of *ergodicity*, it is usually expected that the obtained distributions are identical [Altschuler and Wu, 2010; Huang, 2009; Patrascioiu, 1987; Thomas et al., 2013].

2.3.3. Comparison between bistability and bimodality

Fig 2.2 illustrates and compares the concepts of bistability and bimodality, using the example of a simple bistable genetic circuit with a positive feedback loop, which has been modeled deterministically and stochastically[14]. In (A), protein trajectories obtained with deterministic and with stochastic simulations are shown (the stochastic simulation was generated using the Gillespie algorithm, cf. Sec 2.1.4). In the deterministic case, the trajectories are smooth, and the initial condition determines which of the two stable steady states is assumed in the long run. By contrast, the stochastic trajectory is dominated by random fluctuations, which may be large enough to promote noise-driven system-state transitions. Panel (B) shows a method to graphically identify deterministic fixed points. The system regarded here is one-dimensional and can be described by an ODE of the form

$$\dot{n} = r_{\text{synthesis}}(n) - r_{\text{degradation}}(n). \tag{2.37}$$

[14]The question which stochastic model corresponds to a given deterministic model has been answered in Section 2.1.3.

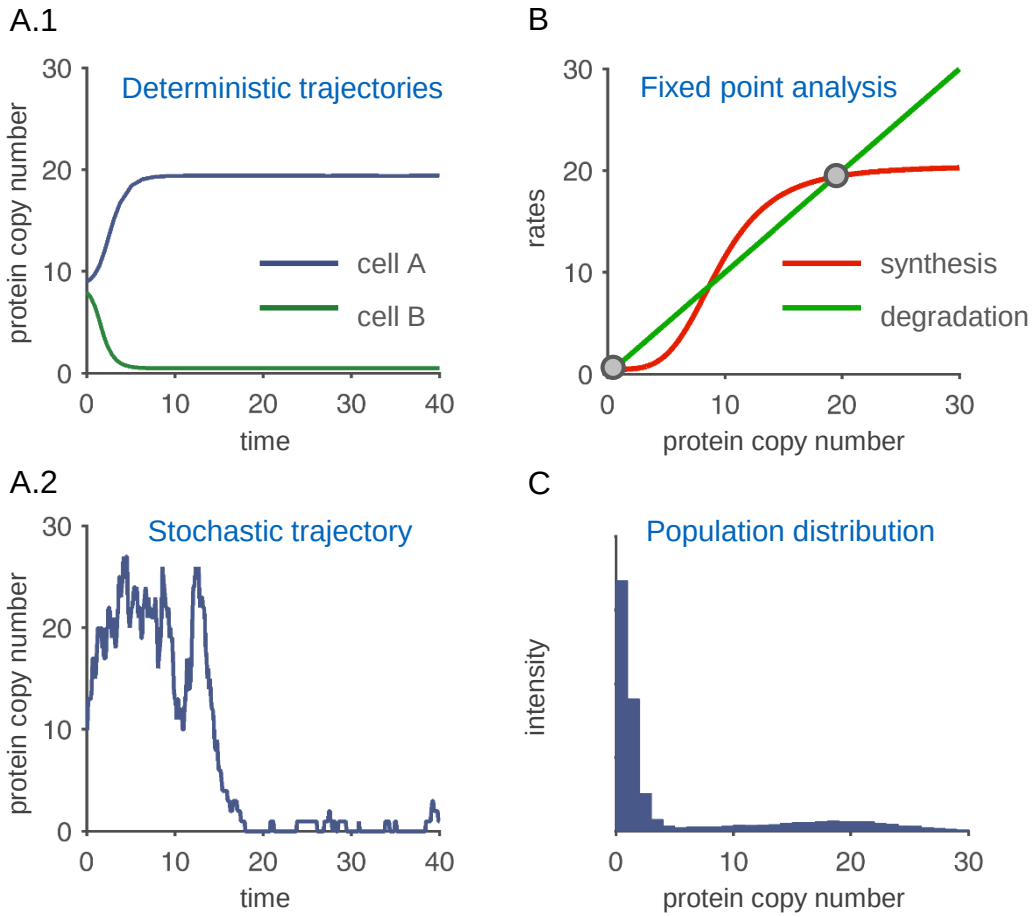


Figure 2.2.: **Bistability versus bimodality.** (A.1) and (A.2) show deterministic and stochastic trajectories of a bistable system. In the deterministic simulation, the initial condition determines whether the high or low expression state is reached. The stochastic trajectory shows large fluctuations and a random transition from the high to the low expression state. (B) Graphical fixed point analysis. Stable steady states are marked with filled circles. (C) Population distribution obtained from stochastic simulations. Here, $r_{\text{synthesis}}(n) = 0.5 + 20 \cdot n^4 \cdot (9.5^4 + n^4)^{-1}$, $r_{\text{degradation}}(n) = n$.

In this example, n is the deterministic protein copy number[15], which determines the rates of its own synthesis and degradation. The equation tells us that fixed points occur when $r_{\text{synthesis}}(n)$ and $r_{\text{degradation}}(n)$ are identical. By visualizing the graphs of both rates in one common plot as shown in panel (B), the stationary states can thus be found at the intersection points, whose number and locations depend on the shapes of the two graphs. Using the Jacobian-based stability analysis described before, one can show that in the case illustrated here (sigmoid synthesis rate and linear degradation rate, generating three fixed points), the two outer stationary states are stable, while the middle one is unstable – a clear indication of bistability. The graphical analysis also shows that this bistability is basically generated through a nonlinear, positively autoregulated (i.e. monotonically increasing) synthesis rate like in Fig 1.1. In (C), the stochastically simulated protein distribution is shown to be bimodal. When comparing the locations of stochastic modes and deterministic stable fixed points, bimodality simply seems to be a "noisy"

[15]In ODE modeling, state variables are usually given in terms of concentrations, while the usage of copy numbers is prevalent in stochastic descriptions of low-abundance species. For the sake of comparability, however, we only use copy numbers in this example.

version of bistability, at least in this example.

It is thus reasonable to ask whether bistability and bimodality always coincide. If this was the case, one could first use deterministic modeling to reliably (and comparatively easily) detect bimodality. In case of interest, one could then further specify population distributions and dynamics with the help of stochastic approaches. In [Karmakar and Bose, 2007], two scenarios in which the introduction of only one noisy variable or parameter into an otherwise deterministic, bistable model led to bimodality are illustrated: In the first scenario, noise was assumed to be generated within the circuit, leading to fluctuations of a reaction component that might reach the basin of attraction of another stable fixed point. In the second, noise enters the system from outside the regarded circuit, resulting in fluctuations of a circuit parameter which span the region around a bifurcation point. Although many other theoretical and experimental examples have shown that bistability may be important for the generation of bimodality, it is evident that if the level of internal and external noise was too low to generate random state transitions, bistability would occur without bimodality.

Unfortunately, things are even more complicated, as a number of recent theoretical and experimental studies have demonstrated the existence of bimodal distributions whose deterministic counterpart was monostable [Artyomov et al., 2007; McSweeney and Popovic, 2014; Ochab-Marcinek and Tabaka, 2010; Qian et al., 2009; Samoilov et al., 2005; Shu et al., 2011; To and Maheshri, 2010]. After a review of these works, all of which concentrate on special cases, it is hardly possible to derive universal conditions under which such qualitative deviations occur, since there are many different ways in which bimodality can be generated. Two of them should nevertheless be pointed out: It has been shown that bimodality without bistability can be generated in closed-loop and even in open-loop systems (i.e. in systems with or without feedback) through slow binary switches of an upstream component (e.g. of the promoter or operator state) [Karmakar and Bose, 2004, 2007; Kepler and Elston, 2001; Qian et al., 2009; Thomas et al., 2014], which mathematically lead to a superposition of two unimodal distributions. The reason for monostability is that the binary switching behavior is simply averaged out in a deterministic description. A second, remarkable situation in which bimodality occurs in spite of monostability is assessed in [Bishop and Qian, 2010]: In this study, a feedback system is regarded whose nonlinearity is too weak to generate deterministic bistability, but whose probability mass function has, under certain parametric conditions, two modes: one corresponding to the deterministic stable state and one state where the molecule copy number equals zero. This situation seems to occur when the zero state is almost absorbing, i.e. if the probability of reactions away from this state is comparatively low.

To conclude, bistability may serve as an important indicator of bimodality, but true phenotypic heterogeneity is only reliably detected using stochastic approaches. Moreover, the observed discrepancies between bistability and bimodality motivate a critical evaluation of deterministic approaches in the context of highly stochastic mesoscopic systems, especially since they are often used in the formulation of approximate stochastic system descriptions, like in the LNA.

2.3.4. Multimodality as a major challenge in modeling

In several passages of Section 2.1.6, it has been mentioned that the presented approximate approach is not applicable to multimodal systems. For example, the LNA *per se* is not suited to describe multimodality, as the obtained probability distribution is necessarily Gaussian and hence unimodal. However, even exact calculations of mean values and variances lose all signifi-

cance in the multimodal case, cf. Fig 2.3. This might also hold for burst characteristics. It would thus make sense to find a possibility to characterize noise for each expression state individually.

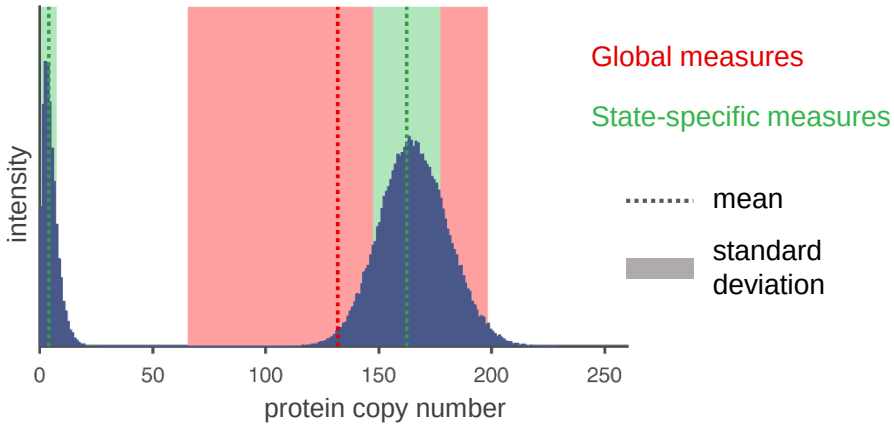


Figure 2.3.: **State-specific noise measures.** The actual mean value and standard deviation (in mean ± standard deviation) of the bimodal distribution is visualized in red. Compared to that, expression-state specific measures (shown in light green) are much more informative.

The way how this could be achieved is certainly dependent on the way how bimodality is generated. For example in [Thomas et al., 2014], a monostable system is analyzed whose multimodality is caused by rare switches of a binary upstream component. The system is modeled using LNAs which are conditioned on the stochastically switching state of this component. This ansatz is able to give a quite complete picture of the system dynamics, including state transitions. However, the approach cannot be applied to other origins of multimodality, in particular to multimodality that is due to feedback regulation and therefore associated with multistability.

In the latter case, one could associate every peak of a distribution with a stable deterministic steady state and perform one LNA per fixed point in order to obtain local variance estimates for each mode. This was done, e.g., in [Tomioka et al., 2004]. However, the probability of transitions between different modes is not fully quantifiable. Moreover, it is hard to predict *a priori* whether the deterministic variable is a good approximation of the local mean value around which fluctuations occur. According to the decomposition in Eq (2.25), this is however an important prerequisite, which has a severe impact on the quality of the LNA. The EMRE and IOS methods, which generally give improved mean and variance estimates in the unimodal case, cannot be applied to multimodal systems, as they approximate the means and variances of the whole distribution instead of yielding local estimates for each mode. All in all, there is currently no satisfying approach with which noise can be reliably quantified in systems that are multistable and multimodal.

3. Bridging the gap between deterministic and stochastic models

As indicated in Section 2.3, deterministic models are typically much easier to analyze than their stochastic counterparts, which is one reason for their great popularity in theoretical biology. In general, they are assumed to give an accurate description of a dynamical system in the thermodynamic (or macroscopic) limit, when random fluctuations become negligible. But what about systems with smaller volume (like prokaryotic cells or eukaryotic cellular compartments), where fluctuations might play a significant role?

In order to answer this question, one first needs to clarify how the quality of a deterministic model can be assessed, i.e. what the stochastic equivalent of a deterministic variable is. One might suppose that ODE models capture the average behavior of a system, implying that deterministic variables are comparable to stochastic mean values. But if so, how can the occurrence of deterministic bistability be explained? In Fig 2.2, it appeared as if deterministic variables were rather associated with stochastic modes. But numerous examples from literature have shown that bistability does not always coincide with bimodality.

The resolution of this ambiguity is not only important in order to assess whether a deterministic description is appropriate or not; it is also crucial for the application of the LNA (and hence for the analysis of noise using this approximate approach), where a symmetric distribution of fluctuations around the deterministic solution is assumed. In the following, a simple gene expression system with feedback regulation and with protein bursts is systematically analyzed under these aspects and some general rules are deduced. In contrast to previous studies on this topic (e.g. in [Grima et al., 2011]), multimodal systems will be regarded as well. The results have partially been published in advance in [Hahl and Kremling, 2016] and in [Hortsch and Kremling, 2018b].

3.1. Deterministic and stochastic model of a gene regulatory system with feedback

The system that will be described and analyzed is similar to that in Fig 2.1 (A), but with autoregulated transcription. The reactions of the full model, which includes mRNA and protein dynamics, can be described as follows:

$$
\begin{array}{rcl}
\emptyset & \xrightarrow{\hat{F}(s)} & \text{mRNA} \\
\text{mRNA} & \xrightarrow{d_m} & \emptyset \\
\text{mRNA} & \xrightarrow{\hat{g}} & \text{mRNA + Protein} \\
\text{Protein} & \xrightarrow{d_s} & \emptyset
\end{array}
\tag{3.1}
$$

$\hat{F} : \mathbb{N}_0 \to \mathbb{R}_{\geq 0}$ is a monotonically increasing function that, depending on its shape, can specify different types of autoregulation. For example, non-cooperative feedback is typically described by a function that can either be interpolated linearly or by a Michaelis-Menten-type saturation function. Cooperativity, where several protein molecules of the same species create a synergistic, positive effect on their own expression, may be modeled using a sigmoid Hill-type function. Obviously, $\hat{F}$ is a non-elementary process consisting of a series of biochemical reactions (e.g. potential interactions among target protein molecules to form complexes, binding of these proteins to the activator site, assembly, binding/unbinding and action of RNA polymerase and of other regulators), which are lumped into one function using time-scale separation (cf. the model reduction techniques in Section 2.1.6). Using a similar argument, promoter states are assumed to change quickly enough to be averaged out. In order to reduce the number of parameters and to increase the comparability between different systems, the reactions are scaled: The dimensionless time variable $\tau := t \cdot d_s$ is chosen such that proteins degrade with proportionality constant 1, where t is the original process time. $F : \mathbb{N}_0 \to \mathbb{R}_{\geq 0}$, $F(.) := \frac{1}{d_m}\hat{F}(.)$ now defines the mode of autoregulation, and $g := \frac{1}{d_s}\hat{g}$ is the scaled translation constant. Finally, the parameter $\nu := \frac{d_s}{d_m}$ relates the time-scales of mRNA and protein to each other. By changing ν, the scaled propensities of mRNA transcription and degradation events are altered proportionally. The list of scaled reactions reads:

$$\begin{array}{lcl} \emptyset & \xrightarrow{\frac{1}{\nu}\cdot F(s)} & \text{mRNA} \\ \text{mRNA} & \xrightarrow{\frac{1}{\nu}} & \emptyset \\ \text{mRNA} & \xrightarrow{g} & \text{mRNA + Protein} \\ \text{Protein} & \xrightarrow{1} & \emptyset \end{array} \tag{3.2}$$

This reaction scheme is illustrated in Fig 3.1 (A).

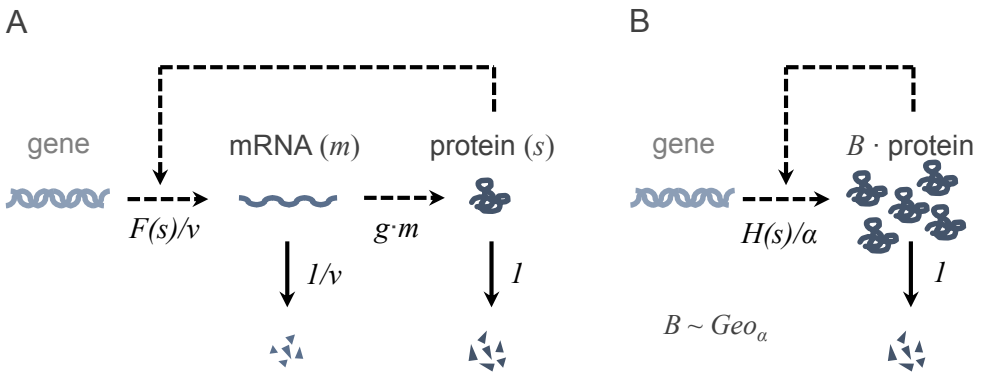


Figure 3.1.: **Reaction scheme of a simple autoregulatory circuit.** Visualizations of (A) the full reaction scheme (3.2), which encompasses the mRNA and protein level, and (B) the reduced reaction scheme (3.3) that takes protein bursts into account while having eliminated the mRNA level. For both systems, the reference time scale is determined by protein degradation.

As the chance of finding an exact solution to a CME usually decreases with the number of involved reactive species, a model reduction based on the stochastics of protein bursts is performed, which is inspired by [Aquino et al., 2012] and by [Friedman et al., 2006] and which has been shortly introduced in Section 2.2.2. The idea is that only protein formation and

degradation events are described, while the mRNA dynamics are implicitly included in the protein burst statistics. As already mentioned, the condensation of translation events into a single bursty protein production event is only possible if $\nu \ll 1$, i.e. if the average mRNA half-life is extremely short. The reactions now read:

$$\begin{array}{rcll} \emptyset & \xrightarrow{\frac{1}{\nu}\cdot F(s)=\frac{1}{\alpha}\cdot H(s)} & B\cdot\text{Protein}, & B\sim \text{Geo}_\alpha \\ \text{Protein} & \xrightarrow{1} & \emptyset & \end{array} \tag{3.3}$$

cf. the illustration in Fig 3.1 (B). The average burst size in the reduced model is given by $\alpha = \frac{\hat{g}}{d_m} = g\cdot\nu$, cf. Eq (2.32). Like before, the random variable B is the actual burst size of the reaction, which is geometrically distributed. $\text{Geo}_\alpha(b) := (\frac{\alpha}{1+\alpha})^b \cdot \frac{1}{1+\alpha}$ denotes the geometric probability mass function evaluated at b. In order to investigate the influence of the average burst size on the accuracy of deterministic models without changing the corresponding ODEs, we furthermore define the function $H : \mathbb{N}_0 \to \mathbb{R}$, $H :\equiv \frac{\alpha}{\nu} F = g\cdot F$. An increase in α under constant H and ν corresponds to an increase in g and a simultaneous, inverse proportional decrease in F, as it was experimentally studied e.g. in [Ozbudak et al., 2002; Thattai and van Oudenaarden, 2001], cf. Section 2.2.2. The reduced system (3.3) can be stochastically simulated using the Gillespie algorithm, which can handle the random stoichiometric coefficient B by drawing a geometrically distributed random variable with mean value α. For analytical calculations, however, it might be useful to integrate the stochasticity of the burst size into the reaction propensity. This results in an infinite number of reactions leading to protein production, which differ in the number of produced molecules:

$$\begin{array}{rcll} \emptyset & \xrightarrow{\text{Geo}_\alpha(b)\cdot\frac{1}{\alpha}\cdot H(s)} & b\cdot\text{Protein}, & b\in\mathbb{N}_0 \\ \text{Protein} & \xrightarrow{1} & \emptyset & \end{array} \tag{3.4}$$

All parameters in the model are non-negative. Furthermore, the volume V of the system is assumed to be constant.

3.1.1. CME formulation of the reduced system and its stationary solution

The basic stochastic description which our subsequent analysis will be based on is obtained by inserting the propensities of the reduced model (3.4) into the general formulation of the CME, which is given in Eq (2.3):

$$\dot{p}_s = \sum_{b=0}^{s} \left(\text{Geo}_\alpha(b)\cdot\frac{1}{\alpha}\cdot H(s-b)\, p_{s-b} \right) - \frac{1}{\alpha}\cdot H(s)\, p_s + (s+1)\, p_{s+1} - s\, p_s. \tag{3.5}$$

For the sake of simplicity, the time variable τ has been omitted.

The stationary[1] PMF p_s can be obtained using generating functions [Aquino et al., 2012], see also Section 2.1.1. The calculations are shown in the Appendix, Section A.1, and yield the

[1]Since stochastic descriptions will mostly be regarded in their stationary state, the probabilistic steady state will not be specifically labeled as such. In contrast, deterministic fixed points are marked with an asterisk. The reason for this seeming inconsistency shall become clearer in the following chapters, when analyses of local stationary fluctuations around fixed points are performed. All the obtained values and measures will then be labeled with an asterisk as well to indicate their close association with the fixed point.

following recursive formulation:

$$\begin{aligned} p_1 &= \frac{1}{1+\alpha}\, H(0)\, p_0, \\ (s+1)\, p_{s+1} &= \frac{1}{1+\alpha}\left(\frac{H(s)}{s} + \alpha\right) s\, p_s, \quad s \in \mathbb{N}. \end{aligned} \tag{3.6}$$

The explicit steady-state-distribution can now be determined easily:

$$p_s = \frac{H(0)}{s}\, p_0 \left(\frac{1}{1+\alpha}\right)^s \prod_{i=1}^{s-1} \left(\frac{H(i)}{i} + \alpha\right). \tag{3.7}$$

3.1.2. Determination of moments and modes

In order to later be able to compare deterministic fixed points with some of the most important characteristics of a PMF, the locations of mean values and of modes are determined. Since the size of fluctuations might influence these potential relations, also the variance is regarded.

The dynamics of the mean value and variance of the stochastic process S describing the protein copy number are given by:

$$\dot{\mathbb{E}}[S] = \mathbb{E}[H(S)] - \mathbb{E}[S], \tag{3.8}$$

$$\dot{\mathrm{Var}}(S) = 2\,\mathrm{Cov}(S, H(S)) + (\mathbb{E}[H(S)] + \mathbb{E}[S]) + 2\,\alpha\,\mathbb{E}[H(S)] - 2\,\mathrm{Var}(S), \tag{3.9}$$

cf. Appendix, Section A.2. When the stationary state is reached, the following conditions[2] hold:

$$\mathbb{E}[S] = \mathbb{E}[H(S)], \tag{3.10}$$

$$\mathrm{Var}(S) = \mathbb{E}[S]\,(1+\alpha) + \mathrm{Cov}(S, H(S)). \tag{3.11}$$

The formulae show that if H is constant, i.e. if feedback regulation is missing, the relation $\eta(S) = (1+\alpha)$ is obtained, in accordance with Eq (2.34). In case of positive feedback, which is described by a monotonically increasing function H, $\mathrm{Cov}(S, H(S)) > 0$ holds, so that the protein noise level is augmented compared to a system without feedback. In an analogous manner, one can deduce that negative feedback reduces noise.

Next, the maxima and minima of the stationary PMF p_s are determined. In general, a local maximum (mode) occurs at a positive value $s > 0$ if $p_{s-1} \leq p_s \geq p_{s+1}$. Accordingly, a minimum fulfills the condition $p_{s-1} \geq p_s \leq p_{s+1}$. Using these conditions and the recursive formulation of the stationary distribution in Eq (3.6), which can be re-written as $p_{s+1} = \frac{H(s)+\alpha\, s}{s+\alpha+1+\alpha\, s}\, p_s$, one can conclude that a mode/minimum occurs at a positive value $\mathring{s} > 0$ if:

$$(\mathring{s}-1) + \alpha + 1 \begin{matrix} \leq \\ \geq \end{matrix} H(\mathring{s}-1) \qquad \text{and} \qquad \mathring{s} + \alpha + 1 \begin{matrix} \geq \\ \leq \end{matrix} H(\mathring{s}). \tag{3.12}$$

A necessary condition for the occurrence of an extreme value is thus given by:

$$\mathring{s} = \lceil \sigma \rceil \quad \text{where} \quad \sigma \in \mathbb{R}, \quad \sigma + \alpha + 1 = H(\sigma). \tag{3.13}$$

[2] They can also be obtained using the stationary solution p_s defined in Eqs (3.6) or (3.7).

Here, H is a smooth, reasonable[3] interpolation of the originally discrete function, defined on the nonnegative real line. If possible, the interpolation should be linear.

A further mode/minimum is located at $\mathring{s} = 0$ if $p_0 \gtreqless p_1$, i.e. if

$$1 + \alpha \gtreqless H(0). \tag{3.14}$$

3.1.3. RE model of the reduced system

After having obtained formulae that specify the location of stochastic means and modes, we will now take a look at the deterministic formulation. In the ODE modeling of cellular systems, variables are usually given in terms of concentrations (instead of copy numbers) and treated as continuous. Let c_s denote the deterministic concentration of the regarded protein in a cell. Now, the scaled deterministic rates of protein production and degradation need to be specified, which are related to the stochastic propensities. For this purpose, one should consider that $\frac{1}{\alpha} H(s)$ actually plays the role of an extensive stochastic reaction constant of a zeroth-order reaction, which is however dependent on the current state of the system due to the PSS assumption. It is therefore related to an intensive[4], deterministic expression $\frac{1}{\alpha} \mathcal{H}(c_s)$, specified by a function $\mathcal{H} : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$, in a way that $H(c_s \cdot V) =: \mathcal{H}(c_s) \cdot V$, cf. Eq (2.15). Here, the continuous interpolation of H is used again. $\frac{1}{\nu} \cdot \mathcal{F} := \frac{1}{\alpha} \cdot \mathcal{H}$ is thus the deterministic rate of protein production. Since protein degradation is a first-order process, it is described by the same reaction constant as in the stochastic model.

Now the dynamics of c_s can be formulated. Since stochasticity is neglected, the number of proteins produced per burst is equal to the average burst size α. This gives the ODE

$$\dot{c}_s \;=\; \alpha \cdot \frac{1}{\alpha} \mathcal{H}(c_s) \;-\; c_s \;=\; \mathcal{H}(c_s) \;-\; c_s \tag{3.15}$$

with the steady state condition

$$c_s^* \;=\; \mathcal{H}(c_s^*). \tag{3.16}$$

Note that this deterministic formulation is not influenced by the presence of protein bursts.

In case Eq (3.16) cannot be solved for c_s^*, the number and location of fixed points can be determined at least graphically by drawing the graphs of both sides of the equation in one plot and by identifying the intersection points (cf. Fig 2.2 (B)). In the top row of Fig 3.2, four exemplary plots are shown[5] that differ in the shape of $\mathcal{H}$. Provided that basal protein production occurs (i.e. $\mathcal{H}(0) > 0$), systems without autoregulation (panel (A)) or with non-cooperative feedback (panel (B)) can only have one stable fixed point. Under cooperative autoregulation, the system might either be mono- or bistable (panels (C) and (D)).

[3] Sometimes, mechanistic considerations could help to find a reasonable, simple interpolation function - e.g. a Michaelis-Menten-type function.

[4] One should bear in mind that H depends on the volume of the system, while $\mathcal{H}$ does not, when systems with changing volume are considered.

[5] In these plots, the amount of protein is given in copy numbers. The deterministic fixed points $s^* = c_s^* \cdot V$ then fulfill the condition $s^* := H(s^*)$ (note that s^* might assume non-integer values). This conversion is done in order to allow direct comparisons to the locations of modes and mean values in the following section.

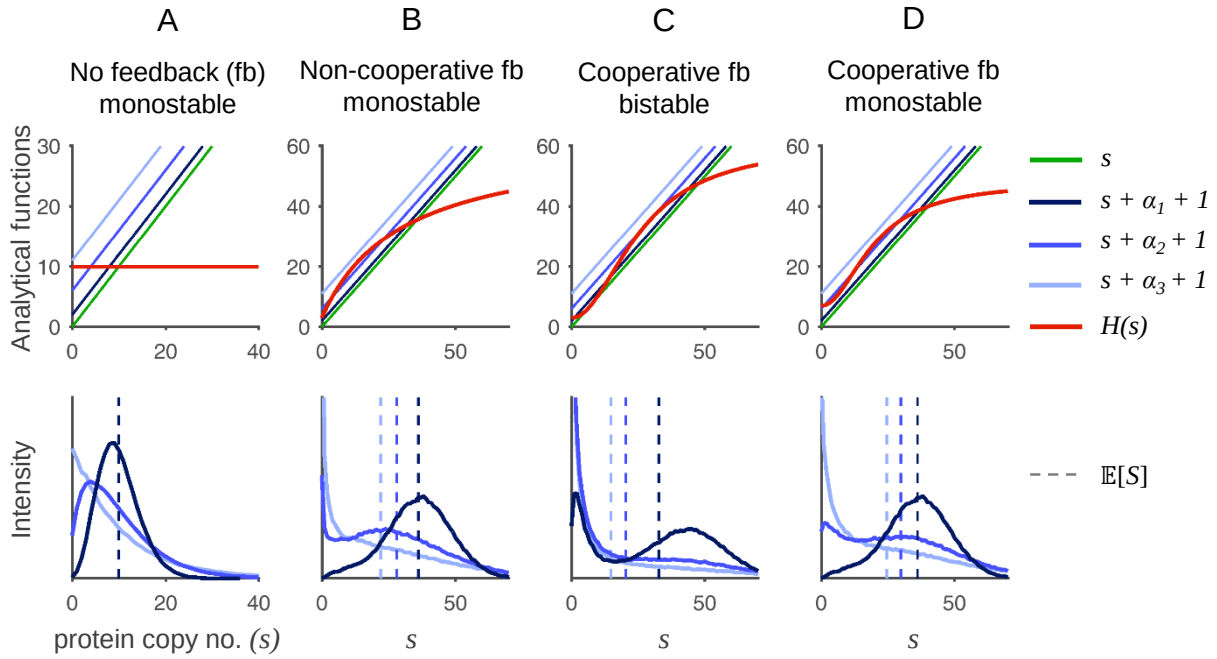


Figure 3.2.: **Analytical comparison of stable fixed points and modes.** From left to right, the type of feedback is varied. The top row shows the graphical determination of the locations of fixed points and modes. The deterministic fixed points can be read from the value of s at the intersection points between the graph of $H(s)$ (red) and the graph of the identity (green). The system shown in (C) is bistable, all others are monostable. The intersection points of the red line and the blue lines $.+\alpha+1$ indicate the locations of the extreme values of p_s. Three different values of α are shown: $\alpha_1 = 1$ (very small bursts, dark blue), $\alpha_2 = 5$ (medium-size bursts, mid blue), and $\alpha_3 = 10$ (strong bursts, light blue). Depending on the burst size, the location and number of modes may change. The second row shows the histograms of the protein distribution obtained from protein time-course simulations using the Gillespie algorithm. The distribution is shown for each burst size (same color as above). Average values are marked by dashed lines. The protein synthesis rates are defined by: (A) $H(s) = 10$, (B) $H(s) = 3 + 60\,\frac{s}{s+30}$, (C) $H(s) = 3 + 58\,\frac{s^2}{s^2+680}$, (D) $H(s) = 7 + 41.25\,\frac{s^2}{s^2+400}$. The figure is adapted from [Hahl and Kremling, 2016].

3.2. What are the stochastic equivalents of deterministic variables?

Sometimes, deterministic models are regarded as descriptions of the average behavior of a system. In order to assess the validity of this statement, let us revisit the dynamics of the mean value in Eq (3.8) and compare them to the ODE formulation, Eq (3.15). Apparently, $c_s(\tau)$ can be identified with $\frac{\mathbb{E}[S]}{V}(\tau)$ if H (and, accordingly, $\mathcal{H}$) is a linear function:

$$\frac{\dot{\mathbb{E}}[S]}{V} = \frac{1}{V}\left(\mathbb{E}[H(S)] - \mathbb{E}[S]\right) = \frac{1}{V}H(\mathbb{E}[S]) - \frac{\mathbb{E}[S]}{V} = \mathcal{H}\left(\frac{\mathbb{E}[S]}{V}\right) - \frac{\mathbb{E}[S]}{V}. \tag{3.17}$$

Linearity of H was used in the crucial second step in order to set $\mathbb{E}[H(.)] = H(\mathbb{E}[\,.\,])$. This formula is independent of α. Combined with Eq (3.11), this shows that if H is linear, the variance of protein fluctuations can be modulated by changing the average burst size α without altering the average protein level.

The relation (3.17) does usually not hold if H is nonlinear (cf. Jensen's inequality[6] to get a rough idea of the bias). In order to see what causes the deviation between c_s and $\frac{\mathbb{E}[S]}{V}$ in this case, one may formulate the Taylor expansion of $H(S)$ at $\mathbb{E}[S]$ and calculate the expected value thereof:

$$\mathbb{E}[H(S)] \;=\; H(\mathbb{E}[S]) + \sum_{r=2}^{\infty} \left(\frac{1}{r!} \cdot \mathbb{E}[(S - \mathbb{E}[S])^r] \cdot H^{(r)}(\mathbb{E}[S]) \right). \tag{3.18}$$

$H^{(r)}$ denotes the r-th derivative of H. The expression $\mathbb{E}[(S-\mathbb{E}[S])^r]$ is the r-th central moment of S (the second central moment being the variance). The deterministic variable is hence only a good approximation of the mean value if the summation term almost vanishes, which is the case if in the global range of protein fluctuations, the function H is almost linear. Our calculation of $\mathrm{Var}(S)$ in Eq (3.11) has shown that this range of fluctuations increases with the average burst size α, so that large bursts may promote even greater deviations of fixed points from the mean value.

The statement about the role of linearity can be generalized: Eqs (2.12) and (2.18) show that the deterministic vector of variables exactly describes the average dynamics of a multivariate stochastic process $\mathbf{N}$, if all propensity functions w_j can be interpolated by a linear function[7]:

$$\dot{\mathbf{c}} \overset{(2.18)}{=} \lim_{V\to\infty} \frac{1}{V}\, \mathbf{A}\, \mathbf{w}(\mathbf{c}\cdot V) \overset{\text{lin.}}{=} \mathbf{A}\, \mathbf{w}(\mathbf{c}) \mathrel{\widehat{=}} \mathbf{A}\, \mathbf{w}\left(\frac{\mathbb{E}[\mathbf{N}]}{V}\right) \overset{\text{lin.}}{=} \frac{\mathbf{A}\, \mathbb{E}[\mathbf{w}(\mathbf{N})]}{V} \overset{(2.12)}{=} \frac{\dot{\mathbb{E}}[\mathbf{N}]}{V}. \tag{3.19}$$

Note that linearity of H excludes the possibility of the system to be bistable or bimodal.

Let us next compare the necessary condition for the occurrence of deterministic fixed points in Eq (3.16) with that of stationary modes and minima in Eq (3.13). The mathematical structure of the two conditions is very similar and shows that the smaller α is, the better is the correspondence between $c_s^* \cdot V$ and $\mathring{s}$. One can furthermore use Eq (3.12) and deterministic stability analysis to show that for sufficiently small bursts, the modes of the probability distribution are associated with stable fixed points, while the minima correspond to the unstable ones. But what if bursts are large? This question will be answered in the following.

3.3. Disrupting the connection between bistability and bimodality

This section aims at demonstrating that, as a consequence of the burst-dependent discrepancy between stable steady states and modes, bistable, but unimodal, and monostable, but bimodal systems may emerge. It is shown how such systems can be constructed with the help of graphical methods, based on the rather simple, but flexible structure of the general autoregulatory circuit.

According to Eq (3.13), the location of modes and minima can be determined graphically by evaluating the intersection points between the graphs of $H(s)$ and of $s+\alpha+1$. For this reason, the latter, linear graph is added to the plots in the top row of Fig 3.2. Three different values of α are chosen in order to demonstrate the influence of the burst size. We will go through panels (A) to (D) to see what effect an increase in α might have, depending on the feedback

[6] In the stochastic context, Jensen's inequality states that $\mathbb{E}[H(S)] \leq H(\mathbb{E}[S])$ if H is concave, and that $\mathbb{E}[H(S)] \geq H(\mathbb{E}[S])$ if H is convex.

[7] For ensuring linearity of the propensity function, the combination of the (possibly state-dependent) reaction "parameter" $\kappa_j(s)$ with the product term taking the stoichiometry of the reaction into account needs to be linear, cf. Eq (2.5).

mechanism.

In panel (A), H is a constant function. For $\alpha = 1$, the mode corresponds well to the stable deterministic fixed point. An increase in α, however, shifts the location of the mode towards lower protein copy numbers until finally, the intersection point vanishes. In this case, $H(0) < 1 + \alpha$, so that the only mode is located at $\mathring{s} = 0$, cf. Eq (3.14). In the bottom row of Fig 3.2, the corresponding protein distributions (obtained through simulations using the Gillespie algorithm) are visualized in order to verify the analytical results. As expected, they clearly show the shifting of the mode due to the increase in the burst size, resulting in a highly skewed distribution. Moreover, the variance of the distribution is apparently increased through bursting, in accordance with Eq (3.11). The locations of the mean values of all three distributions are identical (i.e. not influenced by bursting) and correspond exactly to the deterministic fixed point, which is due to the linearity of H.

In panel (B), H is a Michaelis-Menten-type saturation function, which is suitable to describe non-cooperative feedback. In case of the smallest burst size that is shown here, the monostable system has only one mode. A sufficiently large increase in α, however, does not only lead to a shift of this mode like in panel (A), but also to the emergence of a second intersection point between $H(s)$ and $s + \alpha + 1$, which corresponds to a *de novo* formation of a minimum in the probability distribution. At the same time, another mode is created at $\mathring{s} = 0$. The distribution is thus bimodal, although the deterministic description postulates monostability. A further increase in α lets both intersection points vanish, so that the only mode is now located at $\mathring{s} = 0$. However, this mode cannot be associated with the deterministic stable state anymore. These qualitative changes in the probability distribution share some similarities with phenomena observed in deterministic bifurcation analysis – however, the "stochastic bifurcation parameter" α considered here does not have any effect on the deterministic model.

The location of the mean value in this example deviates from the location of the fixed point due to the nonlinearity of H. The deviation is enlarged through greater bursts, which amplify the range of protein fluctuations, making them cover a larger nonlinear domain of H.

The behavior of the system represented in panel (C) with cooperative feedback, described by a sigmoid function H, is as follows: The parameters are chosen such that the deterministic system is bistable. Accordingly, the stochastic distribution is bimodal for small burst sizes. However, an increase in α finally leads to an unimodal distribution with a mode located at $\mathring{s} = 0$.

In panel (D), the parameters of H are varied so that the deterministic system is monostable, but in a certain range of α, the distribution becomes bimodal. In contrast to panel (B), both modes occur at positive protein copy numbers.

We can conclude that sufficiently large bursts lead to a skewed protein distribution, whose shape might contrast quantitatively and qualitatively with the behavior predicted by deterministic modeling. The graphical illustrations in the top row of Fig 3.2 facilitate the construction of systems where bistability and bimodality do not coincide. Note that depending on the properties of the circuit, bursts do not really need to be "large" (from a biological perspective) in order to generate fundamental differences between deterministic and stochastic system properties.

From a more general point of view, we have just studied the influence of nonlinearities (represented by H) and of the magnitude of stoichiometric coefficients (represented by α) on the relation between deterministic and stochastic models. The opportunity to flexibly choose these functions and parameters makes the genetic circuit considered here a powerful object of study.

3.4. Partial "rehabilitation" of deterministic rate equations

The insights that have been gained until now have challenged the role of rate equations in the context of mesoscopic systems, as well as the role of bistability in cellular regulation and signaling, since monostable systems may generate bimodal distributions as well. In this section, two aspects are highlighted that justify the application of rate equations under special conditions: The first one regards the macroscopic limit, which has already been thoroughly studied in other works, using quite sophisticated approaches[8]. In the following, a good agreement between stable fixed points and modes in large systems is demonstrated in a quite straightforward manner at the example of the autoregulatory circuit. This connection can also be visualized with the graphical method used in Fig 3.2. The second aspects concerns the robustness of a bimodal distribution, which is shown to be particularly high when the system is bistable.

3.4.1. The thermodynamic limit

Despite the possibly large deviations between deterministic and stochastic model behavior in systems with limited volume, a good agreement between stochastic stationary modes and stable deterministic fixed points exists in the thermodynamic limit $V \to \infty,\ s \to \infty$, s.t. $\frac{s}{V}$ is constant. In order to show this, a system is considered whose original volume V is increased k-fold ($k \geq 1$), i.e. $V_k = k \cdot V$. The burst size α as well as the deterministic behavior of the system should be kept constant. The deterministic rate of protein production $\mathcal{H}_k$ in the enlarged system is therefore equal to $\mathcal{H}$:

$$\dot{c_s} \;=\; \mathcal{H}_k(c_s) \;-\; c_s \;=\; \mathcal{H}(c_s) \;-\; c_s. \tag{3.20}$$

The propensity of protein production H_k changes in accordance with the volume V_k: Using the relation $c_s = \frac{s_k}{V_k}$, one obtains $H_k(s_k) = H_k(c_s \cdot k \cdot V) = k \cdot V \cdot \mathcal{H}_k(c_s) = k \cdot V \cdot \mathcal{H}(c_s) = k \cdot V \cdot \frac{H(c_s \cdot V)}{V} = k \cdot H(\frac{s_k}{k})$. Next, the location of modes is determined in the system with changed size: Let $\mathring{s}_k > 0$ be an extreme value of the PMF. According to Eq (3.13), it fulfills:

$$\begin{aligned} \mathring{s}_k = \lceil \sigma_k \rceil \quad \text{where} \quad \sigma_k \in \mathbb{R}, \quad \sigma_k \;+\; \alpha \;+\; 1 \;=\; H_k(\sigma_k) \;=\; k \cdot H\left(\frac{\sigma_k}{k \cdot V} \cdot V\right) \\ \Leftrightarrow \qquad \frac{\sigma_k}{k \cdot V} + \frac{\alpha + 1}{k \cdot V} = \mathcal{H}\left(\frac{\sigma_k}{k \cdot V}\right) \end{aligned} \tag{3.21}$$

Taking the limit $k \to \infty$ in the last equation gives (using the continuity of $\mathcal{H}$):

$$\lim_{k\to\infty} \frac{\sigma_k}{k \cdot V} = \mathcal{H}\left(\lim_{k\to\infty} \frac{\sigma_k}{k \cdot V}\right) \tag{3.22}$$

Eq (3.16) shows that the last equation is solved by a deterministic fixed point, i.e. $\exists\ c_s^* \in \mathbb{R}_{\geq 0}$: $\lim_{k\to\infty} \frac{\sigma_k}{k \cdot V} = c_s^*$.

Furthermore, using

$$c_s^* = \lim_{k\to\infty} \frac{\sigma_k}{k \cdot V} \leq \lim_{k\to\infty} \frac{\lceil \sigma_k \rceil}{k \cdot V} \leq \lim_{k\to\infty} \frac{\sigma_k + 1}{k \cdot V} = c_s^* \tag{3.23}$$

[8] e.g. approaches that use the chemical Langevin equation, whose validity needs to be justified itself [Ceccato and Frezzato, 2018; Gillespie, 2009], or considerations based on stochastic convergence [Kurtz, 1972].

and the squeeze theorem, one obtains

$$c_s^* = \lim_{k\to\infty} \frac{\mathring{s}_k}{k \cdot V}, \tag{3.24}$$

the right hand-side (RHS) of the equation being an extreme value of the probability distribution in the thermodynamic limit.

Besides that, a mode occurs at $\mathring{s}_k = 0$, if $k \cdot H(0) = H_k(0) \leq 1 + \alpha$, which in the thermodynamic limit reduces to

$$0 \leq H(0) \leq \lim_{k\to\infty} \frac{1+\alpha}{k} = 0. \tag{3.25}$$

This means that a mode at zero only occurs in the macroscopic limit if basal protein production is lacking. In this case, the deterministic model possesses a fixed point at $c_s^* = 0$ as well.

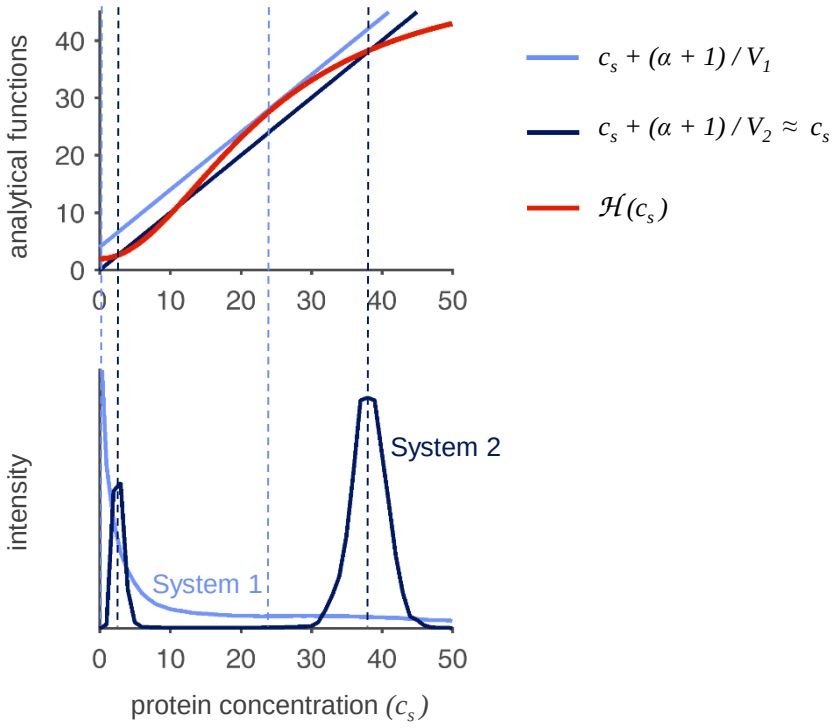


Figure 3.3.: **Influence of system size on the relation between stable fixed points and modes.** Two systems are compared that differ in their volume, while the protein concentrations at the deterministic fixed points are identical. The volume V_1 of system 1 (graphs in light blue) is 50-fold smaller than the volume V_2 of system 2 (graphs in dark blue). In the upper plot, the location of modes is graphically determined like in Fig 3.2; they correspond to the outer intersection points of the blue lines with the red line. The modes of system 2 are almost identical to the stable fixed points, since the expression $\frac{\alpha+1}{V_2}$ is negligible. The distributions in the bottom plot were obtained using the Gillespie algorithm and confirm these results: the distribution of the larger system shows two distinct peaks whose locations match those of the stable deterministic fixed points, while the modes of the small system are shifted to smaller concentrations, and the distribution is much broader. The parameters and functions used here are: $\mathcal{H}(c_s) = 2 + 50\frac{c_s^2}{c_s^2+550}$, $\alpha = 3$, $V_1 = 1$, $V_2 = 50$. The figure has been modified from [Hahl and Kremling, 2016].

We can thus conclude that in the thermodynamic limit, the deterministic fixed points correspond to the extreme values of the protein distribution (except for minima at $c_s = 0$, which, however, do not hold any special significance). One can furthermore show as before that stable

steady states are associated with modes, while unstable fixed points correspond to minima in the distribution. The finding is illustrated in Fig 3.3, where the concentration distributions of two deterministically equivalent systems with different size are compared. While the distribution of the system with small volume is not clearly bimodal, the distribution of the system with larger volume shows two narrow peaks, whose locations coincide with those of the deterministic stable states. One can imagine that in this system, random transitions between the two expression states are rare, so that its behavior is expected to be almost deterministic. The main reason for the low relative noise level is that bursts in protein production (or, to be more general, jumps in stochastic trajectories due to large stoichiometric coefficients) become negligible in the macroscopic limit, leading to fluctuations that are small compared to their (local) average.

3.4.2. On the robustness of expression states

Going back to smaller volumes, the robustness of expression states in bimodal systems will now be examined. It can be qualitatively assessed by regarding the local shape of the protein distribution: A narrow peak that does hardly overlap with peaks of other expression states suggests that the corresponding state is stably maintained (cf. the protein distribution of the large-scale system in Fig 3.3). In order to find out which circuit properties lead to robustness, the recursive formulation of the stationary PMF in Eq (3.6) is rewritten as:

$$p_{s+1} - p_s \;=\; \frac{1}{1+s} \cdot \frac{H(s) - (s+\alpha+1)}{1+\alpha} \cdot p_s. \tag{3.26}$$

If s is held fixed, one can deduce that the relative local change in the PMF $\frac{|p_{s+1}-p_s|}{p_s}$ is large if

- α is small, while
- $|H(s) - (s+\alpha+1)|$ is large.

If those criteria are fulfilled in a domain around a mode, the corresponding expression state is expected to be robust.

As far as the first criterion is concerned, we know that if the average burst size α is small in relation to the protein copy number of the regarded mode, the deterministic fixed point and the mode would be located closely to each other. We can thus already conclude that a circuit with very high robustness is expected to be described well by the variables of a deterministic model.

Now the second criterion is regarded. From Eq (3.12), it can be deduced that in a domain around $\mathring{s}$, the conditions $H(s) \geq (s+\alpha+1)\ \forall s < \mathring{s}$ and $H(s) \leq (s+\alpha+1)\ \forall s \geq \mathring{s}$ hold true, where the right hand-side (RHS) of the inequalities is linear with slope 1. Considering that H is non-decreasing, the robustness of an expression state is therefore largest if H is locally constant[9]. Interestingly, if the very same condition is fulfilled in a domain around a stable fixed point c_s^*, the time it requires for a locally perturbed system to return to c_s^* is minimized: The eigenvalue of the deterministic ODE (3.15) linearized around c_s^* is equal to $\lambda = \frac{d}{dc_s}\mathcal{H}(c_s^*) - 1 < 0$. Its absolute value $|\lambda|$ characterizes the velocity with which the system relaxes to its steady state. It is maximized if $\frac{d}{dc_s}\mathcal{H}(c_s^*) = 0$, since the derivative of $\mathcal{H}$ is necessarily nonnegative. Stochastic and deterministic robustness thus seem to be closely connected.

Taking both conditions together, the highest degree of robustness in bimodality (under constant volume) appears to be achieved by deterministic bistability that ensures a fast relaxation of the system inside each attractor region.

[9] In Section 5.1, a similar result is deduced using an approach that will be developed in Section 4.3.

In the following, further consequences of the two robustness criteria are highlighted: Three scenarios are studied each of which compares two related systems that are bimodal, but *monostable*. The results are visualized in Fig 3.4.

In the first scenario (panel (A)), the burst size α is varied, while the expression $H(s) - (s + \alpha + 1)$ and the location of modes and minima are held constant by simply adjusting the basal protein production rate. As expected, bimodality is more pronounced in the system with smaller burst size (system 1, dark colors), which becomes apparent when comparing the simulated protein distributions (plot in middle row). Morever, also the simulated protein time courses (lower plots) show that system 1 switches between missing and active protein production, while it is hardly possible to distinguish different expression states in the protein fluctuation pattern of system 2.

In panel (B), two systems with equal burst size are compared, which are again monostable, but bimodal. Once more, it is taken care that the extreme values of the probability distributions occur at identical values. In one case, the positive feedback is cooperative (sigmoid H), in the other case non-cooperative (concave H). This generates a local difference in the values of $|H(s) - (s + \alpha + 1)|$. The expression is always larger in the case of cooperative feedback[10], which can be recognized in the analytical plot, and which can be mathematically shown to hold in general, see the Appendix, Section A.3.1. As a consequence, the protein distribution of the cooperative feedback system has much sharper peaks. The time courses show that this system mainly stays in its high expression state, while the other system fluctuates between both states.

Panel (C) again compares two monostable, but bimodal systems with identical locations of modes and minima. In contrast to (A), the basal protein production rate is equal, and in contrast to (B), feedback is non-cooperative in both systems. The systems differ in their average burst size and in the parameters of the Michaelis-Menten-type function H. This affects both robustness criteria in a way that the effects counteract each other (cf. Appendix A.3.2). Explicit calculations are therefore required to determine which effect is prevailing. Interestingly, this example shows that large bursts may even increase the robustness of bimodality by enabling feedback with stronger nonlinearity.

Although in these three scenarios, only monostable systems have been regarded, it turned out that the circuit which exhibits stronger features of a bistable system with potent attractors (in particular small bursts and cooperativity, or at least a high degree of nonlinearity) tends to be the one whose expression states are more robust.

3.5. Short discussion and summary

In several classical works, it has been shown that chemical reaction systems behave almost deterministically in the thermodynamic (or macroscopic) limit, as fluctuations become negligibly small compared to the absolute copy numbers of reactive species [Gillespie, 2001, 2007, 2009; Kurtz, 1972; van Kampen, 2007]. Because of that, a direct comparison between deterministic variables and the stochastic process is possible[11]. However, in systems with small volume, the

[10] Note that in a domain around the positive-valued mode, the propensity of protein production is usually rather flat if feedback is cooperative. This has previously been identified as a stabilizing property.

[11] In [Kurtz, 1972], the probability of a deviation between the stochastic process and the deterministic value that is larger than an arbitrarily small threshold value is shown to converge to zero in the large volume limit. In the works by Gillespie, the chemical Langevin equation is used as a link between stochastic and deterministic descriptions.

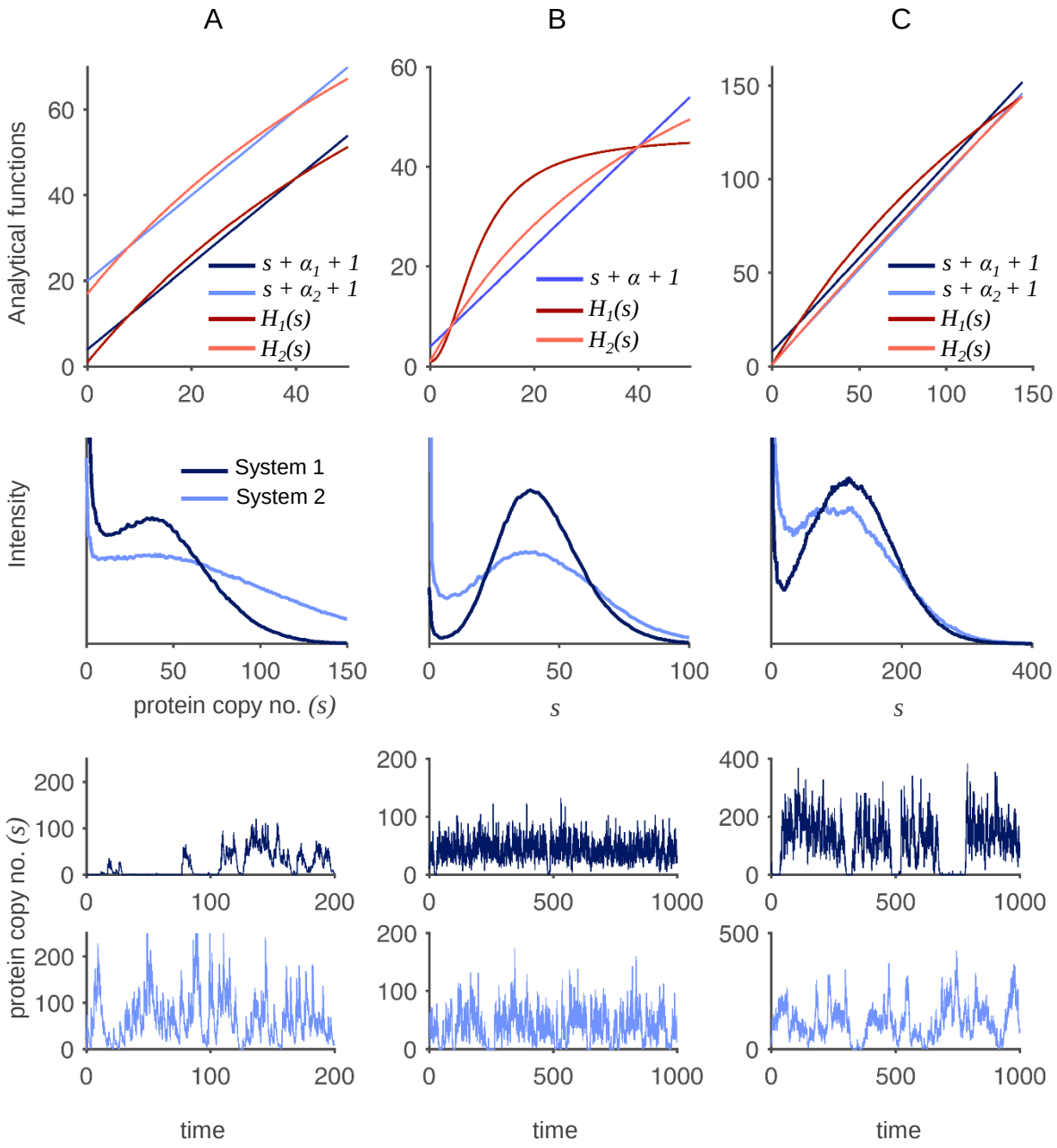


Figure 3.4.: **Robustness of bimodality in different autoregulatory systems.** In each panel, the robustness of the expression states of two bimodal, but monostable regulatory systems is compared based on protein distributions and exemplary protein time-courses. (A) Comparison of two non-cooperative feedback regulations with differing burst sizes $\alpha_1 < \alpha_2$ and with functions $H_1(s) = a + v\frac{s}{s+K}$ and $H_2 = (\alpha_2 - \alpha_1) + H_1$ that have identical shape but different intercept. (B) Comparison of cooperative ($H_1(s) = a + v_1\frac{s^2}{s^2+K_1}$) and non-cooperative ($H_2(s) = a + v_2\frac{s}{s+K_2}$) regulation in two systems with identical burst size α. (C) Comparison of two non-cooperative regulatory systems $H_i(s) = a + v_i\frac{s}{s+K_i}$ with identical basal protein expression but with differing burst sizes $\alpha_1 > \alpha_2$. In all cases (A)-(C), the system illustrated in dark colors shows clearer bimodality. Parameters were chosen such that (A) $\alpha_1 = 1$, $\alpha_2 = 20$, $s_{min} = 8$, $s_{max} = 40$, (B) $\alpha = 4$, $s_{min} = 4$, $s_{max} = 40$, (C) $\alpha_1 = 8$, $\alpha_2 = 2$, $s_{min} = 15$, $s_{max} = 120$, where s_{min} and s_{max} set the locations of the two modes (cf. Appendix A.3). The figure is adapted from [Hahl and Kremling, 2016].

probability distribution is typically much broader and it is thus not trivial to decide which characteristic measure (e.g. mean value or mode) of the distribution the deterministic variable should be compared with, and even whether such a comparison is justified at all. Further works which have pointed out significant discrepancies between bistability and bimodality (cf. Section 2.3.3) have created further doubt about the validity of deterministic modeling in the context of cellular systems.

In this study, a simple autoregulatory genetic circuit was modeled both stochastically (using the CME) and deterministically (using REs). The flexibility to choose various kinds of propensity functions and parameters allowed to systematically study the impact of different features (e.g., nonlinearity of reaction propensity functions or the magnitude of stoichiometric coefficients) on the comparability of the two model types. The analysis has been supported by graphical methods that allow a more intuitive understanding of the findings. They also clarify why in the thermodynamic limit, potential deviations between the outcomes of the two modeling approaches vanish.

To quickly summarize the main results, we have seen that only if all reaction propensities are linear, the average behavior of the system is exactly described by the deterministic variable – independent of the size of the system. As soon as nonlinearity is involved (at least in the range of the stochastic fluctuations), the equivalence between average and deterministic systems behavior breaks down. This observation has already been made e.g. in [van Kampen, 2007]. The comparability between stable fixed points and modes depends on the occurrence of translational bursts (or more generally, on the involvement of large stoichiometric coefficients, be they random or fixed), since large burst sizes promote highly asymmetric fluctuation patterns, thereby increasing the skewness of the protein distribution. Modes are thus shifted away from the stable fixed points. In combination with nonlinear propensities, bursts can even alter the number of modes, which can be analytically reproduced and visualized like in Fig 3.2. The graphical method that is used here to locate modes is similar to the classical graphical approach to locate fixed points. This suggests that parameter changes which lead to deterministic bifurcations may also lead to stochastic bifurcations, in the sense that the number of modes is altered. However, additional stochastic bifurcation parameters could be identified which do not have an impact on the deterministic system – in our case, the average burst size α.

How do the findings presented here fit into the list of former publications about parallels and discrepancies between deterministic and stochastic models in mesoscopic regimes – especially concerning the relation between bistability and bimodality (cf. Section 2.3.3)? As already stated, a fully generalizable rule for this relation cannot be deduced due to the variety of possible origins of bimodality. However, the methods that have been presented in this section may at least unify some of the observations[12]:

The study that is perhaps most closely related to the one at hand has analyzed a specific phosphorylation-dephosphorylation cycle with autocatalytic kinase, which is basically a non-cooperative feedback system [Bishop and Qian, 2010]. This system was shown to be deterministically monostable, but bimodal: One mode corresponded to the (only) deterministic fixed point, while the other occurred at the zero state, which was explained by its weakly absorbing property. Although the described system is not a genetic circuit, the principles of the graphical determination of modes shown in Fig 3.2 can be directly applied to it, revealing that the

[12]Since the focus in this study was on one-component feedback systems, the effect of slow upstream binary switches (cf. [Karmakar and Bose, 2004, 2007; Kepler and Elston, 2001; Qian et al., 2009; Thomas et al., 2014]) did not play a role.

phosphorylation-dephosphorylation cycle is comparable to the system presented in Fig 3.2 (B), where bimodality with a mode at the zero state has been observed as well. While the study in [Bishop and Qian, 2010] was limited to that special case, we could in addition show that monostable, bimodal systems can be constructed where *both* modes lie in the positive-valued range (cf. Fig 3.2 (D), medium-size burst). In this case, bimodality is more "related" to deterministic bistability, in the sense that a moderate change in the parameters (but not in the structure or shape) of the rates would result in actual bistability, not least because of the sigmoidal shape of the protein production rate. Although an absorbing zero state is missing in that system, the analytical plots showed that the principles underlying Fig 3.2 (B) and Fig 3.2 (D) are closely related to each other. Moreover, the generation of a unimodal distribution from bistable systems could be explained likewise, cf. Fig 3.2 (C). In all cases, the occurrence of relatively large stoichiometric coefficients led to a pronounced asymmetry of the peaks, which either resulted in local accumulations within the protein distribution forming novel peaks, or to the gradual coalescence of different peaks. The question what a "relatively large stoichiometric coefficient" is can only be answered in the context of the regarded system. Sometimes, reactions generating only one molecule might be sufficient to cause stochastic bifurctions (like in the situation depicted in [Bishop and Qian, 2010]).

All these results have shown that deterministic descriptions need to be treated cautiously in the context of cellular systems. But are there still conditions under which their usage is justified? In the last part of this study, it was shown that high robustness of expression states is often based on principles of bistability. Therefore, many important regulatory systems are bistable, e.g. the circuit triggering genetic competence in *B. subtilis* [Maamar et al., 2007], or the lysis-lysogeny switch in the bacteriophage λ [Bednarz et al., 2014]. In bistable systems, the strongest effects of hysteresis are expected (even in a stochastic regime), which is of particular importance in cases where cellular state transitions are supposed to be controlled by external inductors. However, one must not ignore the fact that a combination of large bursts and non-cooperative feedback is capable of creating rather robust bimodality, too (cf. Fig 3.4 (C)). The potential discrepancy between deterministic stable states and stochastic means or modes is highly relevant for the application of the LNA. This will be addressed more thoroughly in the following study.

This discussion will now be closed with some thoughts about the thermodynamic limit, which should offer additional views on the biological meaning of the discussed problems. As stated by Gillespie, "[...] the thermodynamic limit is not a limit that the system actually approaches as a consequence of its natural temporal evolution, nor even as a result of the experimenter's direct intervention; rather, the thermodynamic limit is a kind of *gedanken limit*, an idealized state that turns out to be useful because it often provides a convenient approximation to macroscopic systems" [Gillespie, 2009]. Let us now perform a "gedanken experiment", visualized in Fig 3.5, in which all cells of a large isogenic population are lyzed while maintaining the functionality of all biochemical processes. Let us first assume that experimental measurements are performed immediately after lysis and some rapid mixing. In this case, the measured concentration would correspond to the stochastic population average. This situation resembles the performance of classical ensemble measurements, which are not able to detect heterogeneity. Next, we assume that the biochemical reactions would continue after lysis under perfect mixing conditions, before measurements are made. The reactive species would synchronize and behave as predicted by the deterministic (macroscopic) model, since this would be the experimental implementation of the thermodynamic limit[13]. If the underlying system is bistable, the final concentrations

[13]It is, however, highly unrealistic that cell lysis is performed without altering cellular functions, which is why this experiment remains purely hypothetical.

that are measured might be different when the experiment is repeated, depending on the initial conditions. Last but not least, experimental measurements without lysis, which are performed on a single-cell level, reveal the full heterogeneity of a population and can only be described using stochastic models. A fine compartmentalization of a biological system into single cells apparently creates individual cellular behavior, which can differ greatly from that of another individuum as well as from the behavior within the hypothetical broth of lyzed cells. Processes inside single cells thus have reduced reproducibility and controllability, but the resulting intercellular heterogeneity provides greater flexibility and new opportunities to the population as a whole.

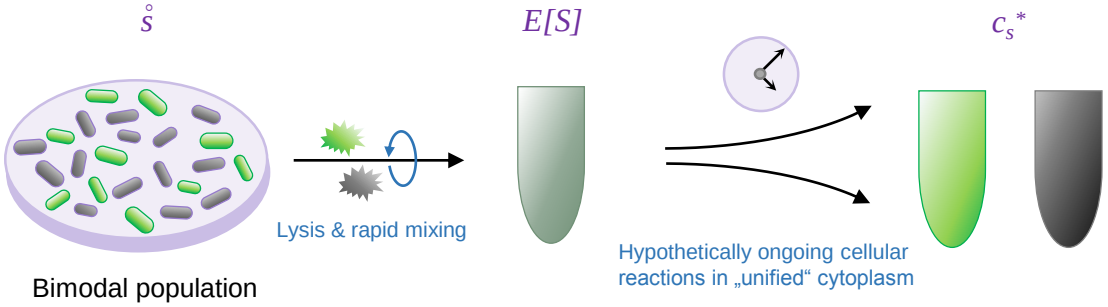


Figure 3.5.: **Modes, means, and fixed points in a hypothetical experiment.** Modes $\mathring{s}$ can be determined by measurements on a single-cell level, as the measured protein distribution over the whole population should resemble the probability distribution of a suitable stochastic model. The mean value $\mathbb{E}[S] \cdot V^{-1}$ of the distribution is obtained after unification of the cytoplasms of all cells and rapid mixing. If the biochemical reactions continued in the unified, well-mixed cytoplasm, almost deterministic model behavior would emerge. The observed stationary concentration c_s^* might be similar to one of the modes in the single-cell experiment, but the values do not need to be identical.

4. Method development for the analysis of noise in multistable circuits

In the previous study, the robustness of expression states was estimated based on the shape of the corresponding peaks in the probability mass function (Section 3.4.2). This method requires the solvability of the master equation and is therefore restricted to systems with few components and with few reactions of limited complexity. In the following, an alternative method with much wider applicability is proposed, which is inspired by the linear noise approximation (Section 2.1.5), but produces results of higher accuracy: It uses an alternative, newly developed deterministic model formulation that better captures mesoscale dynamics. The so-called *hybrid linear noise approximation* (hLNA) is used to approximate the local variances of every peak in a multimodal distribution. The prediction of the variance is solely based on circuit-specific parameters and functions, so that quantitative connections between circuit properties and the robustness of expression states can be deduced. Besides the width of fluctuations, also their skewness and temporal structure will be regarded. This is done by a characterization of translational bursts, whose average size may vary in dependence on the gene expression level as well.

The present chapter deals with the development of methods to characterize noise in a mesoscopic, multistable regime at the example of autoregulatory gene expression. Furthermore, the quality of the methods is evaluated. The obtained results and their biological implications will then be elucidated in more detail in Chapter 5.

4.1. Model system: Two-stage gene expression with autoregulation

A single-gene autoregulatory circuit with scaled reaction propensities is used as model system, which is very similar to the full gene expression system (3.2) of the previous study. This means that mRNA formation and degradation will be explicitly regarded in order to study the interplay of the mRNA and the protein level in detail. The translational propensity function $g \cdot m$ is now replaced by a more general, potentially nonlinear function $G(m)$ to further enhance the flexibility of the description. G is assumed to be strictly monotonically increasing and to suffice $G(0) = 0$. In this study, we allow the function F to describe positive, missing, or even negative autoregulation. The latter is described through a monotonically decreasing function. The reaction system can thus be written as:

$$\begin{array}{rcl} \emptyset & \xrightarrow{\frac{1}{\nu} \cdot F(s)} & \text{mRNA} \\ \text{mRNA} & \xrightarrow{\frac{1}{\nu}} & \emptyset \\ \emptyset & \xrightarrow{G(m)} & \text{Protein} \\ \text{Protein} & \xrightarrow{1} & \emptyset \end{array} \tag{4.1}$$

and be visualized like in Fig 4.1 (A).

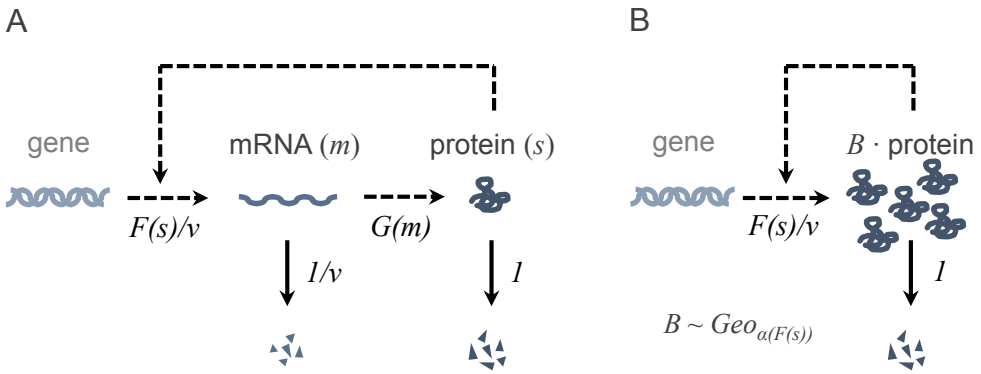


Figure 4.1.: **Reaction scheme of a simple autoregulatory circuit with nonlinear translation.** Visualizations of (A) the full reaction scheme (4.1) and (B) the reduced reaction scheme (4.28) that is formulated using state-dependent protein bursts. The reference time scale of both systems is again determined by protein degradation.

The corresponding CME is given by:

$$\begin{aligned}\dot{p}_{m,s} &= \big[F(s)\,p_{m-1,s} - F(s)\,p_{m,s} + (m+1)\,p_{m+1,s} - m\,p_{m,s}\big]\cdot\frac{1}{\nu}\\ &\quad + G(m)\,p_{m,s-1} - G(m)\,p_{m,s} + (s+1)\,p_{m,s+1} - s\,p_{m,s},\end{aligned} \tag{4.2}$$

where $p_{m,s}$ is the probability that the system is in state $(M,S)=(m,s)$. M is the stochastic process describing the mRNA copy number.

The system of ODEs depicting average dynamics reads:

$$\begin{aligned}\dot{\mathbb{E}}[M] &= \Big(\mathbb{E}[F(S)] - \mathbb{E}[M]\Big)\cdot\frac{1}{\nu}\\ \dot{\mathbb{E}}[S] &= \mathbb{E}[G(M)] - \mathbb{E}[S].\end{aligned} \tag{4.3}$$

The calculations as well as the dynamics and steady-state conditions of the variances $\mathrm{Var}(M)$ and $\mathrm{Var}(S)$ and of the covariance $\mathrm{Cov}(M,S)$ are given in the Appendix, Section B.1.

4.2. Deterministic description: From rate equations to a novel hybrid modeling approach

In the LNA, the dynamics of random variables are described as uniform stochastic fluctuations around the solution of a deterministic description (cf. Section 2.1.5). However, as we have seen in the previous study, rate equations may deviate significantly from the average behavior of a system, which would unavoidably hamper the quality of the LNA and of the variance estimates that are based on it. After formulating the classical REs for the reaction system (4.1), an alternative way of deterministically modeling average dynamics will be derived, and the idea will then be extended to multistable systems.

4.2.1. Formulation in terms of classical rate equations

For the deterministic formulation in terms of rate equations, the variable c_m, which denotes the deterministic mRNA concentration, and the function $\mathcal{G} : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$, $\mathcal{G}(c_m) := \frac{G(c_m \cdot V)}{V} \ \forall\, c_m \in \mathbb{R}_{\geq 0}$ are introduced. In the latter definition, the originally discrete function G has been interpolated smoothly and, if possible, linearly, in analogy to the definition of the interpolated function H and of $\mathcal{H}$ in the preceding chapter. All the other parameters and functions have already been defined in the previous study in Section 3.1.3. The system of rate equations reads:

$$\begin{aligned} \dot{c}_m &= (\mathcal{F}(c_s) - c_m) \cdot \frac{1}{\nu} \\ \dot{c}_s &= \mathcal{G}(c_m) - c_s. \end{aligned} \tag{4.4}$$

As shown in Chapter 3, this formulation only provides a good description of mRNA and protein dynamics if all reactions are linear, or if the reaction volume and the copy numbers of the reactive components are large enough so that fluctuations become negligible. These conditions are usually not fulfilled in single cells, especially regarding the copy numbers of mRNA species. Therefore, an alternative modeling approach is developed, which is referred to as *hybrid deterministic model*: The approach takes discrete fluctuations in M into account, although being fully deterministic.

4.2.2. Derivation of the hybrid deterministic model

As a first step towards improving the accuracy of the deterministic model, the differential equations (4.3) of the mean values are rewritten as

$$\begin{aligned} \dot{c}_m &= \Big(\frac{1}{V}\,\mathbb{E}[F(S)] - c_m\Big) \cdot \frac{1}{\nu} = \Big(\frac{1}{V}\sum_{r=0}^{\infty} F(r) \cdot p_r^S - c_m\Big) \cdot \frac{1}{\nu} \\ \dot{c}_s &= \frac{1}{V}\,\mathbb{E}[G(M)] - c_s = \frac{1}{V}\sum_{n=0}^{\infty} G(n) \cdot p_n^M - c_s, \end{aligned} \tag{4.5}$$

where $c_m := \frac{\mathbb{E}[M]}{V}$ and $c_s := \frac{\mathbb{E}[S]}{V}$ are the expected time-dependent mRNA and protein concentrations. $p_n^M := \sum_{s=0}^{\infty} p_{n,s}$ and $p_r^S := \sum_{m=0}^{\infty} p_{m,r}$ are the marginal PMFs of mRNA and protein with mean values $c_m \cdot V$ and $c_s \cdot V$, respectively.

In case F and G are linear, p^S and p^M do not need to be further specified, since $\frac{1}{V} \cdot \mathbb{E}[F(S)] = \frac{1}{V} \cdot F(\mathbb{E}[S]) = \mathcal{F}(\frac{\mathbb{E}[S]}{V}) = \mathcal{F}(c_s)$ and, analogously, $\mathbb{E}[G(M)] = \mathcal{G}(c_m)$ hold, so that the average dynamics of mRNA and protein concentrations are exactly given by the rate equations (4.4). However, if F and/or G are nonlinear, these equations are not valid. Moreover, the joint and marginal probability distributions might be multimodal. In order to obtain *local* information about noise for every gene expression state, we interpret multimodal distributions as superpositions of two or more unimodal ones, each one of them representing one cellular state, cf. Fig 4.2. As these local[1] (unimodal) probability mass functions p_r^{S*} and p_n^{M*} are typically just purely hypothetical constructs, they cannot be determined analytically. We hence need to search for

[1]From now on, all stationary, local – i.e. expression-state-dependent – parameters, functions, and measures (probability distributions, expectations, variances, etc.) will be marked with an asterisk, just like deterministic fixed points. The reason is that every state-dependent measure is associated with exactly one stable fixed point, if the LNA or variants thereof are used.

approximations, which might be pragmatic, but in the end expected to improve the quality of the deterministic description compared to REs.

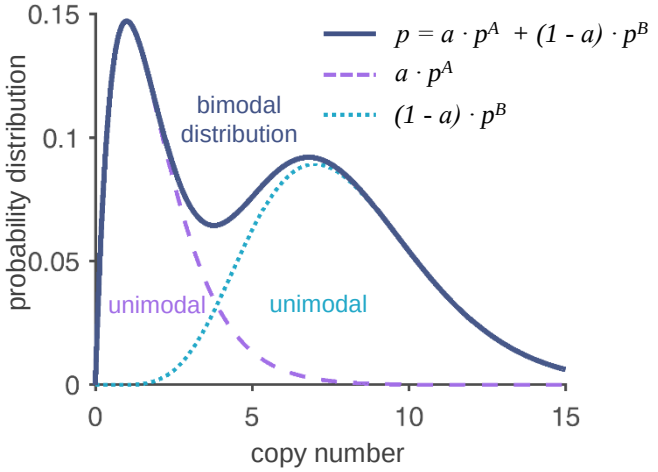


Figure 4.2.: **Bimodal distribution as a superposition of two unimodal ones.** Interpretation of a univariate bimodal PMF p as a superposition of two unimodal ones, p^A and p^B, each representing a specific cellular state. In this example, the bimodal distribution consists of two gamma distributions. However usually, an unambiguous analytical decomposition is not possible.

First, an approximation of the local mRNA distribution p^{M*} is defined. Using generating functions, the CME of the stationary marginal mRNA distribution

$$0 = \dot{p}_n^{M*} = \sum_{s=0}^{\infty} \left(F(s)\, p_{n-1,s}^* - F(s)\, p_{n,s}^* + (n+1)\, p_{n+1,s}^* - n\, p_{n,s}^* \right) \cdot \frac{1}{\nu} \tag{4.6}$$

can be solved, first yielding the recursion formula

$$(n+1)\, p_{n+1}^{M*} = \sum_{s=0}^{\infty} F(s)\, p_{n,s}^* = \mathbb{E}^*[F(S)|M=n] \cdot p_n^{M*}, \tag{4.7}$$

which can then be explicitly written as

$$\begin{aligned} p_0^{M*} &= \left(\sum_{n=0}^{\infty} \frac{\prod_{l=0}^{n-1} \mathbb{E}^*[F(S)|M=l]}{n!} \right)^{-1}, \\ p_n^{M*} &= \frac{\prod_{l=0}^{n-1} \mathbb{E}^*[F(S)|M=l]}{n!} \cdot p_0^{M*}. \end{aligned} \tag{4.8}$$

The calculation is given in the Appendix, Section B.2. At this point, the problem is encountered that the conditional expectation is unknown, unless $F \equiv a$ is constant (at least in the range of local protein fluctuations). In the latter case, $\mathbb{E}^*[F(S)|M=l] = a \;\; \forall\, l \in \mathbb{N}_0$ and $p_n^{M*} = \frac{a^n}{n!}\, e^{-a}$, or equivalently,

$$p_n^{M*} = \frac{\mathbb{E}^*[M]^n}{n!}\, e^{-\mathbb{E}^*[M]}, \tag{4.9}$$

i.e. M is Poisson distributed. Due to the lack of a better, but still tractable *a-priori*-approximation for systems with non-constant F or for systems that are not in their stationary state, Eq (4.9)

will be generally used. Of course, the quality of this approximation needs to be evaluated, which will be done in Section 4.4. The differential equation of the protein dynamics, Eq (4.5), is now replaced by:

$$\dot{c}_s = \frac{1}{V}\sum_{n=0}^{\infty} G(n) \cdot \frac{(c_m V)^n}{n!} e^{-c_m V} - c_s =: \bar{\mathcal{G}}(c_m) - c_s. \tag{4.10}$$

$\bar{\mathcal{G}}$ can be interpreted as an effective translation rate, which emerges from a local averaging of the function G that takes mRNA fluctuations around $c_m \cdot V$ into account. In order to express $\bar{\mathcal{G}}$ in terms of copy numbers $m = c_m \cdot V$, set $\bar{G}(m) := \bar{\mathcal{G}}(\frac{m}{V}) \cdot V = \sum_{n=0}^{\infty} G(n) \frac{m^n}{n!} e^{-m} \; \forall\, m \in \mathbb{R}_{\geq 0}$. Note that in these function definitions, the usage of arbitrary[2] continuous interpolations is completely avoided, since only evaluations of G at integer values enter the formulae. This guarantees the uniqueness of Eq (4.10), which is not given in the RE formulation. A visual comparison of the translation rates in the RE-based (G) and in the hybrid model ($\bar{G}$) is made in Fig 4.3. Here, G was chosen to be a concave function. Obviously, the curvature of $\bar{G}$ is diminished compared to that of G due to local averaging[3].

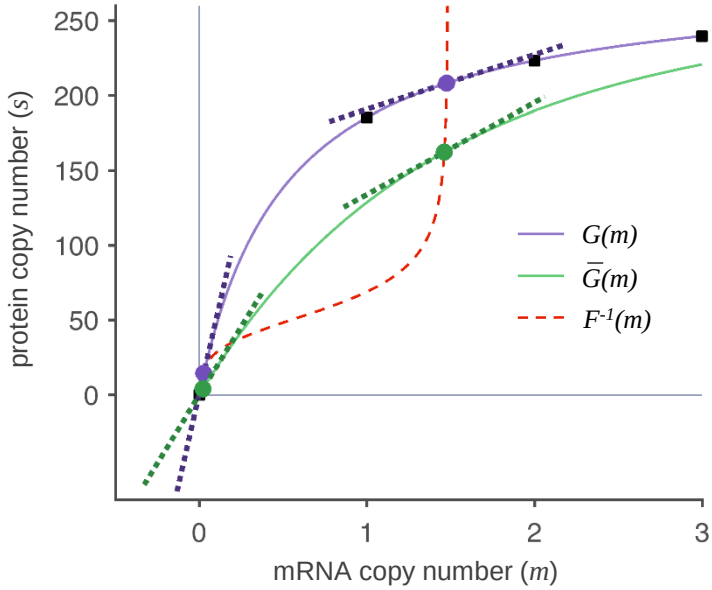


Figure 4.3.: **Translation rates according to the classical and hybrid deterministic model.** While the translation rate G in the classical RE model (purple line) is an interpolation of the discrete function values (black squares), the translation rate $\bar{G}$ of the hybrid deterministic model (green line) is a local average (weighted sum following a Poisson distributed mRNA level with mean m). The exemplary stable fixed points (outer intersection points with F^{-1}, red dashed line) of the classical (purple filled circles) and hybrid (green filled circles) models diverge significantly. In addition, the local derivatives of G and $\bar{G}$ (dashed colored lines) at the respective fixed points are different, which has a great impact on the calculation of the Fano factors using the LNA, cf. Section 4.3. The functions and parameters used in this example are: $F(s) = 0.022 + 1.46\,\frac{s^4}{s^4+58^4}$, $G(m) = 281\,\frac{m}{m+0.51}$, $\nu = 0.01$. The figure has been adapted from [Hortsch and Kremling, 2018b].

It would now be logical to repeat the procedure in order to define an effective transcription rate $\bar{\mathcal{F}}$ as well. However, it is known (and will also later be shown) that the local protein distribution p^{S*} may vary considerably between different circuits, so that it cannot be simply approximated *a priori*, especially if regulatory mechanisms like feedback are involved. Depending on the variance

[2]The interpolation is arbitrary if it cannot be fully deduced based on reliable mechanistic considerations.
[3]Fig 4.3 can be basically interpreted as a visualization of Jensen's inequality for concave functions.

of p^{S*}, the averaging effect on F might be more or less pronounced. As the average protein copy number is typically quite large (at least compared to mRNA copy numbers), the small noise assumption is applied to the protein level, resulting in the classical macroscopic formulation.

All in all, the hybrid deterministic model is given by:

$$\begin{aligned} \dot{c}_m &= (\mathcal{F}(c_s) - c_m) \cdot \frac{1}{\nu} \\ \dot{c}_s &= \bar{\mathcal{G}}(c_m) - c_s, \quad \bar{\mathcal{G}}(c_m) \;=\; \frac{1}{V} \sum_{n=0}^{\infty} G(n)\, \frac{(c_m\, V)^n}{n!}\, e^{-c_m V}. \end{aligned} \tag{4.11}$$

The first ODE defines via c_m the dynamically changing average of the approximate mRNA distribution, which is used in the second ODE for the calculation of the effective translation rate.

4.2.3. Fixed points of the deterministic models and their stability

Here, the main properties of the two deterministic approaches according to the theory of dynamical systems are summarized.

The stationary mRNA and protein concentrations are given by $c_m^* = \mathcal{F}(c_s^*)$ and by

$$c_s^* \;=\; \mathcal{G}(\mathcal{F}(c_s^*)) \qquad \text{in case of model (4.4)}, \tag{4.12}$$

$$c_s^* \;=\; \bar{\mathcal{G}}(\mathcal{F}(c_s^*)) \;=\; \frac{1}{V} \sum_{n=0}^{\infty} G(n)\, \frac{(\mathcal{F}(c_s^*) \cdot V)^n}{n!}\, e^{-\mathcal{F}(c_s^*) \cdot V} \qquad \text{in case of model (4.11)}. \tag{4.13}$$

The stability behavior of (c_m^*, c_s^*) is now determined as explained in Section 2.3.1. Here, (c_m^*, c_s^*) may be a fixed point either of the classical or of the hybrid deterministic model. The Jacobian $\mathbf{J}^*$ evaluated at (c_m^*, c_s^*) reads

$$\mathbf{J}^* \;=\; \begin{pmatrix} -\frac{1}{\nu} & f^* \cdot \frac{1}{\nu} \\ g^* & -1 \end{pmatrix}, \tag{4.14}$$

where $f^* := \frac{d\,\mathcal{F}(c_s)}{d\,c_s}|_{c_s = c_s^*}$ and

$$g^* := \frac{d\,\mathcal{G}(c_m)}{d\,c_m}|_{c_m = c_m^*} \qquad \text{in case of model (4.4)}, \tag{4.15}$$

$$g^* := \frac{d\,\bar{\mathcal{G}}(c_m)}{d\,c_m}|_{c_m = c_m^*} = \sum_{n=0}^{\infty} (G(n+1) - G(n))\, \frac{(c_m^* V)^n}{n!} e^{-c_m^* V} \qquad \text{in case of model (4.11)}. \tag{4.16}$$

As expected, a difference between $\mathcal{G}$ and $\bar{\mathcal{G}}$ can cause a significant difference in the steady states and in the local derivatives, cf. Fig 4.3.

If the trace of $\mathbf{J}^*$ is negative and the determinant is positive, (c_m^*, c_s^*) is asymptotically stable[4]. The first condition is always fulfilled, while the second condition holds iff

$$f^* \cdot g^* \;<\; 1. \tag{4.17}$$

[4]This is a criterion applicable to 2×2-matrices which ensures negativity of the real parts of all eigenvalues.

In this case, the expression $(\text{trace}(\mathbf{J}^*))^2 - 4 \cdot \det(\mathbf{J}^*)$ can be shown to be always positive, so that the eigenvalues turn out to be real-valued and (c_m^*, c_s^*) is thus a stable node.

4.3. The hybrid linear noise approximation

The hybrid LNA is performed like the classical LNA (cf. Section 2.1.5), but using the hybrid deterministic model instead of REs.

This implies that the stable fixed points of the hybrid model $(c_m^* V,\, c_s^* V) =: (m^*, s^*)$, defined in Eq (4.13), serve as an approximation of the local mean values of mRNA and protein fluctuations. The local stationary variances $\text{Var}^*(M)$ and $\text{Var}^*(S)$ around the fixed point $\boldsymbol{\phi}^* \cdot V = (m^*, s^*)^\top$ are calculated by solving the Lyapunov equation (2.29) and by extracting the diagonal entries from the obtained variance-covariance matrix $\boldsymbol{\Sigma}^*$. In that context, the local Jacobian $\mathbf{J}^*$ of the hybrid model is used, cf. Eqs (4.14) and (4.16).

As final noise measures, the expression-state specific Fano factors $\eta^*(M) = \frac{\text{Var}^*(M)}{m^*}$ and $\eta^*(S) = \frac{\text{Var}^*(S)}{s^*}$ are deployed, which are obtained through division of the local variance estimates by the corresponding local mean value estimates.

In general, solving the Lyapunov equation might be tedios, but to two-dimensional reaction systems like the one analyzed here, the method proposed in [Tomioka et al., 2004] can be applied. The corresponding calculation steps are shown in the Appendix, Section B.3, and yield:

$$\eta^*(M) = 1 + \frac{f^*}{(g^*)^{-1} - f^*} \left(\frac{\nu}{1+\nu} + \frac{1}{1+\nu} \cdot\, r^* \frac{f^*}{g^*} \right) \tag{4.18}$$

$$\eta^*(S) = 1 + \frac{f^*}{(g^*)^{-1} - f^*} \left(\frac{1}{1+\nu} + \frac{\nu}{1+\nu} \cdot \frac{1}{r^* \frac{f^*}{g^*}} \right) \quad \text{if } f^* \neq 0, \tag{4.19}$$

$$\eta^*(S) = 1 + \frac{\nu}{1+\nu} \frac{(g^*)^2}{r^*} \quad \text{if } f^* = 0. \tag{4.20}$$

f^* and g^* have been defined in the previous section. $r^* = \frac{s^*}{m^*}$ denotes the stationary ratio of protein and mRNA copy numbers. Note that f^*, g^*, and r^* are parameters whose values may depend on the gene expression state, so that differences in local noise levels within a multistable system are possible. The interpretation of these formulae from a biological point of view will be done in Section 5.1 in the next chapter.

4.4. Comparing the quality of the classical and hybrid LNA

We have seen that the hybrid methods are based on some assumptions, which need to be critically examined. First of all, one should note that the hLNA is only applicable to systems whose multimodality can be explained by hybrid-deterministic multistability, since every expression state needs to be associated with a stable fixed point. Chapter 3 has shown that this condition is not always fulfilled, at least for deterministic models based on rate equations. The situation for hybrid-deterministic approaches is similar (however improved), which will be discussed in more detail in Section 4.4.3.

As a second preliminary note, it should be pointed out that the hybrid approach only differs from the classical one in the usage of an effective (locally averaged) translation rate. If

$G(m) = g \cdot m$ is linear, the identity $\mathcal{G}(c_m) = g \cdot c_m = g \frac{1}{V} \sum_{n=0}^{\infty} n \frac{(c_m V)^n}{n!} e^{-c_m V} = \bar{\mathcal{G}}(c_m) \ \forall\, c_m \in \mathbb{R}_{\geq 0}$ shows that the two deterministic models (4.4) and (4.11) and the corresponding two LNA approaches are equivalent. The following discussion about differences in quality is thus only relevant for systems with a nonlinear translational propensity function – a condition which will be of major interest, since it enables burst sizes to be state-dependent, as will be shown in Section 4.5.

As already stated, for multimodal distributions that originate from multistability, an analytical and biologically reasonable approach by which the distribution can be unequivocally decomposed into unimodal ones does not exist. As a consequence, the "local" measures by which we like to quantify state-specific noise are only hypothetical values, which hampers rigorous error estimation. The following results are therefore mostly restricted to semi-quantitative or qualitative statements and, due to the additional intractability of the full CME, to comparisons with values extracted from stochastic simulations.

4.4.1. Quality of local mean value estimates

First of all, the quality of the local mean value estimates is assessed. They are approximated by the stable fixed points m^* and s^* of the classical or hybrid deterministic models, which are determined by Eqs (4.12) and (4.13), respectively.

In the special case of a constant transcriptional propensity $F \equiv a$, the average mRNA level $\mathbb{E}[M]$ according to the CME (4.2) is correctly given by the value of $m^* = a$ in both deterministic frameworks. Moreover, we have seen that the stationary distribution of the mRNA copy number is exactly Poisson (see Eq (4.9)). As a consequence, also the value of s^* according to the hybrid model is equal to the average protein copy number:

$$s^* \;=\; \bar{G}(a) \;=\; \sum_{n=0}^{\infty} G(n)\, \frac{a^n}{n!}\, e^{-a} \;=\; \mathbb{E}[G(M)] \;=\; \mathbb{E}[S]. \tag{4.21}$$

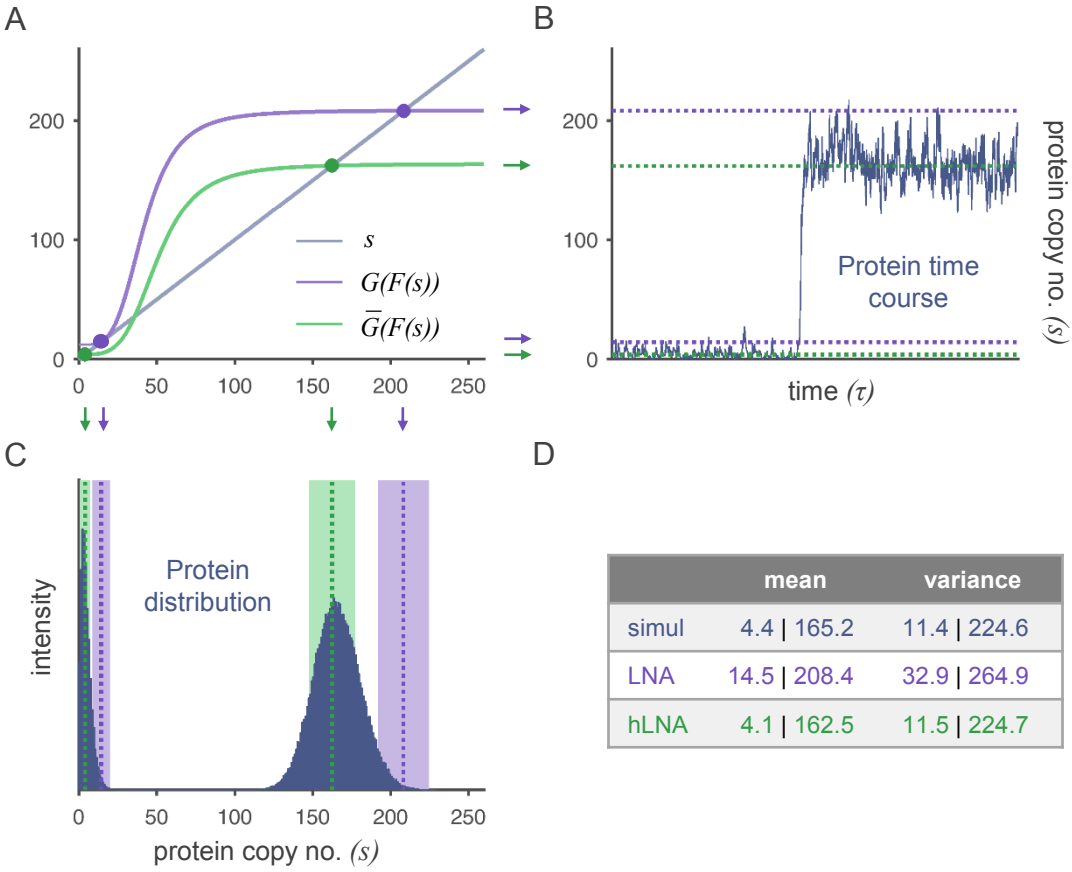
The hybrid fixed point thus provides the exact location of the means, as desired (still under the condition $F \equiv a$). By contrast, the value of s^* according to the RE model is given by the point evaluation $G(a)$. Here, G is a (maybe rather arbitrary) continuous interpolation of the original translational propensity function, which had only been defined on the space of non-negative, integer mRNA copy numbers. If G is nonlinear, $G(a)$ usually differs from $\bar{G}(a)$ and hence from $\mathbb{E}[S]$ due to the missing local averaging effect, cf. Fig 4.3. Therefore, it would certainly provide a worse estimate of the stochastic mean in *every* system without feedback regulation.

One can more generally state that the hybrid deterministic model leads to reliable results if the transcriptional propensity is almost constant in the range of local protein fluctuations: First of all, the local stationary mRNA distribution is then well approximated by a Poisson distribution, which is important for the hybrid formulation of $\dot{c}_s$. Second, also the relation $F(\mathbb{E}^*[S]) = \mathbb{E}^*[F(S)]$ is approximately fulfilled, which has been used in the formulation of $\dot{c}_m$.

This statement also holds for multistable systems, as shown in Fig 4.4: Here, a bistable system with a Michaelis-Menten-type translational propensity function G and with a Hill-type transcriptional propensity function $\frac{1}{\nu} \cdot F$ is studied. In panel (A), the graphical determination of fixed points according to the classical and hybrid ODE model is visualized, showing large differences in their locations. Due to its sigmoid shape, the function F is almost constant in a domain around the fixed points. Visual inspection of panels (B) and (C), which show a simulated

protein trajectory and its histogram, respectively, suggests that the hybrid deterministic fixed points are much better estimates of local averages. This is confirmed by panel (D), which lists the values of the two fixed points according to both deterministic model types together with estimates that have been numerically extracted from the stochastic simulations. In order to obtain these numerical estimates, phases in which the system is clearly in one of the two expression states have been visually grouped across a large collection of protein trajectories, before state-specific local mean values were computed.

For systems with more sensitive or with locally highly nonlinear feedback regulation, the quality of the hybrid ansatz is expected to be reduced, especially in the presence of large protein fluctuations. However, further qualitatively distinct systems were tested and it was observed that in each case, the hybrid approach clearly outperformed the classical one. This will be shown and discussed in Fig 4.5. All in all, the higher accuracy of the hybrid approach is attributable to the local averaging of G, which generates a more realistic effective translation rate, even if the true mRNA distribution is not exactly Poisson.



	mean	variance
simul	4.4 | 165.2	11.4 | 224.6
LNA	14.5 | 208.4	32.9 | 264.9
hLNA	4.1 | 162.5	11.5 | 224.7

Figure 4.4.: **Comparison of the classical and hybrid linear noise approximation.** (A) Graphical determination of stable fixed points according to Eqs (4.12) and (4.13). Intersection points marked in purple and green indicate the stable fixed points of the classical and hybrid deterministic model, respectively. (B) Simulated protein time course showing a random transition from the inactive to the active expression state. The locations of stable fixed points according to the analysis in (A) are indicated by dashed lines. (C): Histogram of the simulated time course. In addition to the estimated local means, the standard deviations obtained with classical and hybrid LNA are visualized by colored areas (mean ± standard deviation). (D): Quantitative comparison of estimates with values extracted from the simulation. Functions and parameters are the same as in Fig 4.3. Figure modified from [Hortsch and Kremling, 2018b].

4.4.2. Quality of the estimation of local Fano factors

Reliable estimates of local averages are crucial for the quality of the linear noise approximation, as they constitute the points around which the propensity functions are linearized. After having compared the classical with the hybrid deterministic model, we would thus expect that the local variances calculated via hLNA are better estimates than those based on standard (RE-based) LNA. Indeed, this is the case in the example visualized in Fig 4.4. The numerical estimates of Fano factors, listed in panel (D), were obtained in analogy to the mean value estimates: After identifying and grouping intervals in the simulated trajectories according to the current expression state, the empirical variances and means of each group were computed and the Fano factor was obtained by dividing the two measures. In order to further validate the quality of the hLNA, four qualitatively different exemplary reaction systems were studied, which are defined by the following functions and parameters:

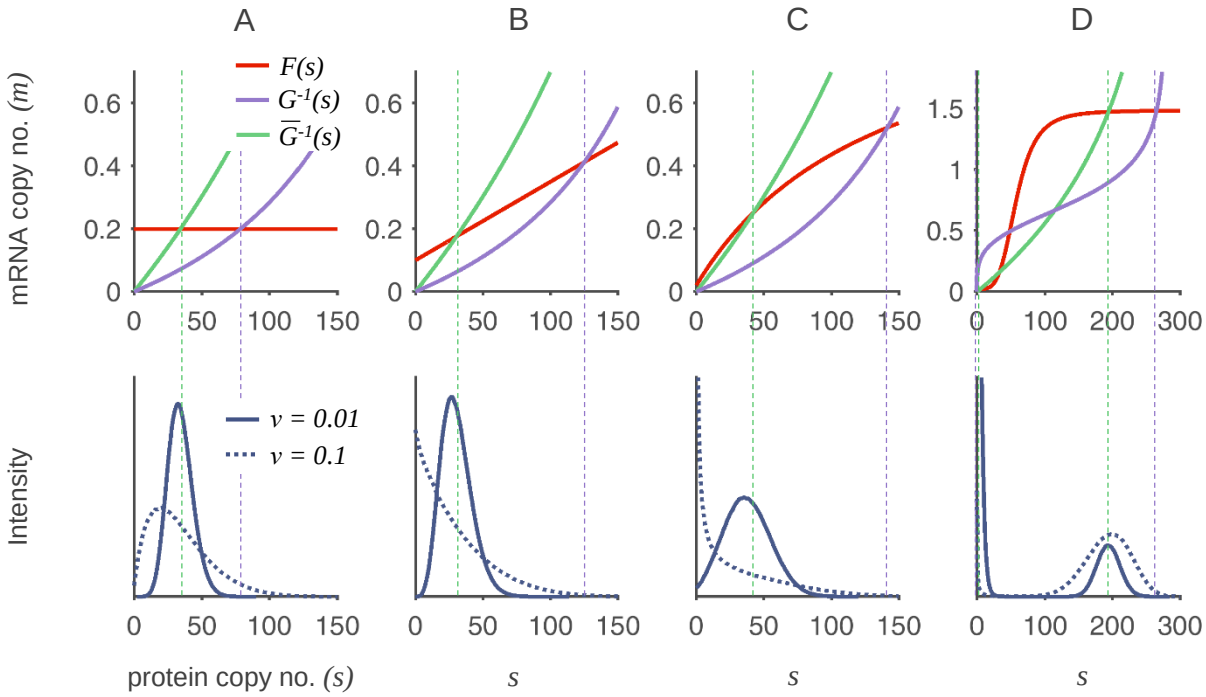
(A)	constant F, concave G:	$F(s) = 0.2,$	$G(m) = 281\,\frac{m}{m+0.51}$
(B)	linear F, concave G:	$F(s) = 0.1 + 0.0025 \cdot s,$	$G(m) = 281\,\frac{m}{m+0.51}$
(C)	concave F and G:	$F(s) = 0.02 + \frac{s}{s+140},$	$G(m) = 281\,\frac{m}{m+0.51}$
(D)	sigmoid F and G:	$F(s) = 0.022 + 1.46\,\frac{s^4}{s^4+58^4},$	$G(m) = 281\,\frac{m^4}{m^4+0.73^4}$

The corresponding deterministic phase plots are illustrated in the top row of Fig 4.5. The parameter ν is varied ($\nu = 0.01$ and $\nu = 0.1$), which does not affect the location of the deterministic fixed points, but the shape of the probability mass functions (cf. the second row of plots in Fig 4.5 showing simulated protein distributions). The tables in the lower part of the figure contain estimates of s^* and $\eta^*(S)$ obtained by classical and hybrid LNA. Moreover, computational estimates and their standard errors of means (SEMs) are given which have been extracted from 10^2 protein time courses with final reaction time $\tau_f = 10^3$. In case of panel (D), the system is bistable, so that the two expression states needed to be evaluated separately as in Fig 4.4. For $\nu = 0.1$, it was observed that the system switches quickly to the high expression state, so that the numerical estimates for the low expression state are of reduced reliability. In the last row of Fig 4.5, the mRNA Fano factor $\eta^*(M)$ (obtained from simulations) is given in order to check whether the approximation of the mRNA distribution with a Poisson PMF is justified.

The system in panel (A) can be regarded as a control for the hybrid deterministic formulation, as F is constant and therefore, the fixed point of the model is expected to exactly match the mean value. This is indeed the case, while the fixed point according to the classical approach deviates significantly. Despite the nonlinearity of G, the hLNA is able to accurately predict the Fano factor as well, suggesting that the approximation error made through the linearization of propensity functions is small.

In (B), F is a linear, increasing function. The linearity of F ensures the commutativity with the expected value operator, which is used in both the classical and the hybrid formulation of $\dot{c}_m$. Compared with panel (A), the quality of the hybrid approach is expected to be reduced, as under feedback regulation, the mRNA distribution generally deviates from a Poisson distribution. The empirical mRNA Fano factor however suggests that this deviation is moderate (although slightly larger for $\nu = 0.1$). As a consequence, the quality of the hybrid LNA is still high, which is not the case for the estimates based on classical LNA.

The function F in panel (C) is a concave saturation curve. Interestingly, the simulated



ν	0.01	0.1	0.01	0.1	0.01	0.1	0.01	0.1
s^* simul.	34.3 ± 0.4	34.2 ± 1.2	30.4 ± 0.6	29.1 ± 1.5	37.9 ± 1.3	20.8 ± 2.8	5.0 ± 0.3 / 194.2 ± 1.4	3.3 ± 2.0 / 193.9 ± 1.7
s^* LNA	78.7		125.1		141.7		0.0 / 265.3	
s^* hLNA	34.3		30.5		42.6		4.8 / 194.2	
$\eta^*(S)$ simul.	2.5 ± 0.1	14.6 ± 0.7	4.2 ± 0.2	23.2 ± 1.5	7.7 ± 0.6	42.0 ± 3.0	1.5 ± 0.1 / 3.4 ± 0.4	5.7 ± 0.4 / 11.5 ± 6.3
$\eta^*(S)$ LNA	3.0	19.5	3.3	16.3	2.2	9.2	1.0 / 4.2	1.0 / 30.1
$\eta^*(S)$ hLNA	2.4	14.3	4.2	24.5	6.5	37.1	1.4 / 3.2	4.5 / 20.6
$\eta^*(M)$ simul.	1.00 ± 0.00	1.00 ± 0.01	1.01 ± 0.00	1.06 ± 0.02	1.03 ± 0.01	1.19 ± 0.03	1.03 ± 0.10 / 1.03 ± 0.01	1.02 ± 0.03 / 1.03 ± 0.01

Figure 4.5.: **Validation of the quality of the hybrid linear noise approximation in different regulatory circuits.** Four qualitatively different systems are analyzed, which are specified in the text: (A) constant F, concave G; (B) linear, increasing F, concave G; (C) concave F and G; (D) sigmoid F and G. In the last case, the system is bistable. The s-m-plots in the top row illustrate the deviations in the fixed points between the classical and hybrid deterministic model. The plots underneath show the protein distributions for $\nu = 0.01$ and for $\nu = 0.1$. The tables contain the estimated fixed points and Fano factors of both approaches (purple: classical LNA, green: hybrid LNA) as well as values extracted from simulations (blue; estimates ± SEMs of 10^2 repeats with final scaled process time $\tau_f = 10^3$ are shown). In (D), estimates for both steady states are given (upper right values: high expression state; lower left values: low expression state). Functions and parameters are listed in the text. The figure is adapted from [Hortsch and Kremling, 2018b].

estimates of s^* are highly different for $\nu = 0.01$ and for $\nu = 0.1$, a phenomenon which has not been observed for linear F. This can be explained as follows: Similar to the averaging of G within the range of mRNA fluctuations, protein noise actually leads to a local averaging of F (which was not included in the hybrid deterministic model formulation). One can show that the range of protein fluctuations is augmented by an increase[5] in ν, which thus enhances this averaging effect. When we imagine that the red line (F) in the phase plot of panel (C) is averaged to various degrees, we can indeed qualitatively reconstruct how the intersection point with the green line ($\bar{G}^{-1}$) is shifted towards the values extracted from the simulations. An extension of the hybrid approach that calculates an effective transcription rate $\bar{F}$ (in a way similar to the calculation of $\bar{G}$) is thus expected to yield improved estimates. However, as mentioned before, the range of protein fluctuations is highly variable and cannot be estimated *a priori* (cf. the highly different values of $\eta^*(S)$ in these examples). Nevertheless, the values listed in panel (C) show that all estimates obtained with the hybrid methods are in a realistic range and clearly outperform the values obtained with REs and with the classical LNA.

In (D), a hypothetical, bistable system with sigmoid F and G is shown. Interestingly, the averaging effect on G is so strong that the effective rate $\bar{G}$ is not a sigmoid function anymore. Due to its sigmoidal shape, F is almost constant in a domain around the fixed points (like in Fig 4.4), leading to good mean and variance estimates when the hLNA method is used.

In all examples, we have seen that the mRNA Fano factor is still close to one (especially in the systems where $\nu = 0.01$), so that the Poisson distribution appears to be a valid approximation.

After this simulation-based evaluation of the hLNA, let us quickly revisit the formulae describing the local Fano factors in Eqs (4.18)-(4.20). Based on these expressions, the following differences between the results of the classical and hybrid LNA can be identified: First of all, the classical LNA uses a local derivative g^* of the continuous function $\mathcal{G}$, which depends strongly on the chosen interpolation. In contrast, g^* in the hybrid approach is a difference quotient (cf. Eq (4.16)) that is unique, since it only uses the original, discrete function evaluations of G. While the function F is the same in both approaches, the values of the local protein-to-mRNA ratio r^* and of the local derivative f^* may differ greatly, since they depend on the locations of the corresponding fixed points.

In summary, the quality of the hybrid approaches is highest for systems without feedback and is not significantly affected by nonlinearity of G, which is the main advantage over the classical approaches. Moreover, the assumption that the mRNA distribution can be approximately described by a Poisson PMF seems to be legitimate even under feedback regulation. Compared with the RE-based approach, the hybrid approach is thus better suited to handle the involvement of mRNA as a low-copy-number species that may participate in reactions with nonlinear propensity functions.

4.4.3. Multistability vs multimodality

Due to the difference between G and $\bar{G}$, classical and hybrid deterministic descriptions of the same genetic circuit may predict qualitatively different systems behavior, which is illustrated in Fig 4.6: In panel (A), the RE model is monostable, whereas the hybrid model is bistable, in panel (B), the opposite holds. The number of stable fixed points according to the hybrid approach coincides with the number of modes in the simulated protein distribution. It needs

[5] This fact can be read from the numerically extracted estimates of $\eta^*(S)$ and will be addressed analytically in Chapter 5.

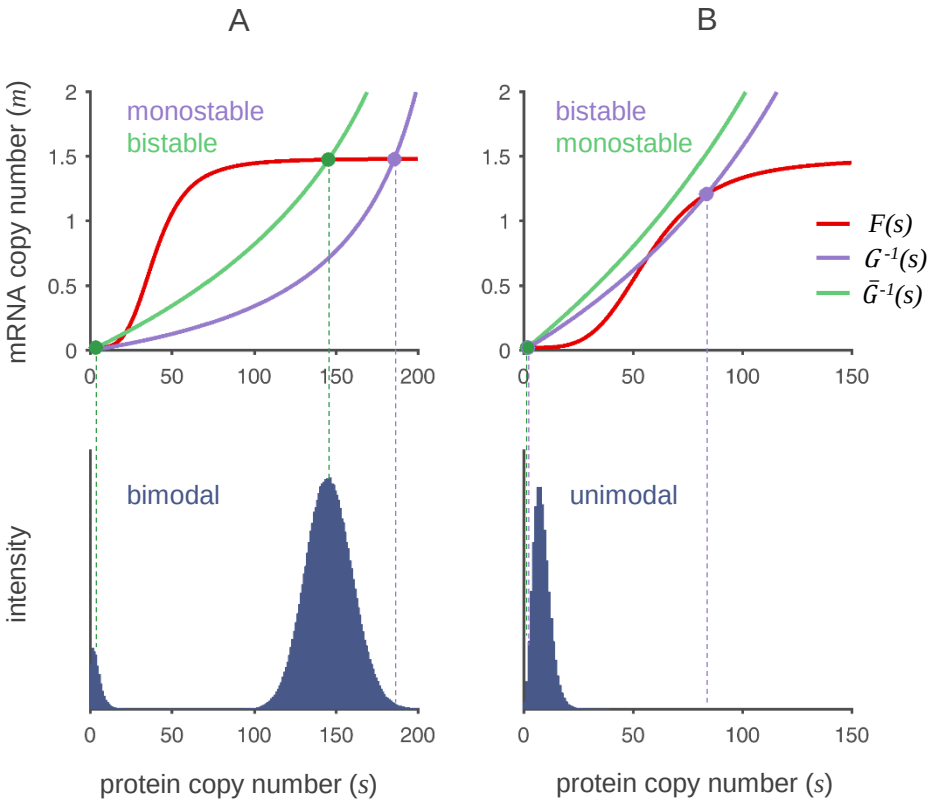


Figure 4.6.: **Differences in system behavior between classical and hybrid deterministic models.** In (A) and (B), two different systems are considered with sigmoid F and saturated G. The upper part shows the phase plots of the deterministic models. Purple filled circles indicate the locations of the stable fixed points according to the classical rate laws, while green filled circles mark those of the hybrid deterministic model. Due to the difference between G and $\bar{G}$, the number of fixed points differs. The lower plots show simulated protein distributions. In (A), the system is bimodal. According to the hybrid model, it is bistable, while the classical model predicts monostability. $F(s) = 0.022 + 1.46\frac{s^4}{s^4+45^4}$, $G(m) = 250\frac{m}{m+0.51}$. In (B), the protein distribution is unimodal, in accordance with the monostability of the hybrid model. However, the classical approach predicts bistability. $F(s) = 0.1 + 1.46\frac{s^4}{s^4+58^4}$, $G(m) = 281\frac{m}{m+2.57}$. The figure is modified from [Hortsch and Kremling, 2018b].

to be mentioned, however, that the burst sizes (which will be defined in the following sections) in these examples are small. Under large bursts, the asymmetry of fluctuations may lead to stochastic bifurcations, which can neither be captured by a RE-based nor by a hybrid deterministic model. This phenomenon has been extensively discussed in Chapter 3. In this case, the hLNA approach is not applicable, since modes and stable fixed points are not associated anymore. Apart from this scenario, one may conclude that the hybrid model is able to predict multimodality more reliably than the classical RE model. As a consequence, the applicability of the hLNA is wider than that of the original approach.

4.5. Definition of state-dependent burst characteristics

Besides the Fano factor, translational bursting is an important characteristic of gene expression noise, which provides information of the temporal structure of fluctuations (see Section 2.2.2 in the Theoretical Background). Since conventional definitions of burst sizes and burst frequencies

only consider systems in which those measures are independent of the gene expression state (as it has also been assumed in Chapter 3), the definitions need to be extended. In this context, will see how state-dependence in bursting is generated biologically.

4.5.1. State-dependent burst size

In general, the burst size is the average number of proteins translated per mRNA, and can be calculated by dividing the effective rate of translation by the rate of mRNA degradation, cf. Eq (2.32) [Friedman et al., 2006; Thattai and van Oudenaarden, 2001]. However, this ratio may vary with the current mRNA level, making the burst size dependent on the state of the system. In order to take this dependence into account, let $\bar{m} \in \mathbb{R}_{\geq 0}$ be the average mRNA level around which the copy number currently fluctuates at time τ (note that the system does not need to be in stationary state). $\alpha(\bar{m})$ is then defined as the number of proteins that are expected to be produced during a potential burst starting in the infinitesimal interval $[\tau, \tau + \delta\tau]$. The burst is triggered by the formation of an additional mRNA molecule and ends with its degradation. During that time, the distribution of the remaining mRNA molecules is assumed to be unaffected[6], so that the average mRNA copy number then becomes equal to $\bar{m} + 1$. In consistence with the derivation of the hybrid deterministic model, the distribution of the other mRNA molecules is assumed to be Poisson, i.e. during the burst, $(M-1) \sim \mathrm{Pois}_{\bar{m}}$, where Pois denotes the Poisson PMF, whose average is given by the subscript.

In order to determine the average burst size, let $[\tau_0, \tau_m]$ be the lifespan of the mRNA molecule that causes the burst. Without loss of generality, let $\tau_0 = 0$. Let furthermore S_B denote the stochastic number of burst proteins translated from that specific mRNA molecule and $p_b^{S_B}(\tau)$ the probability that $S_B(\tau) = b$ at time $\tau \in [0, \tau_m]$. The production of burst proteins is a Poisson counting process with randomly distributed propensity $\frac{G(M)}{M}$, therefore

$$p_b^{S_B}(\tau) \;=\; \mathbb{E}_M^*\left[\mathrm{Pois}_{\frac{G(M)}{M}\cdot\tau}(b)\right] \;=\; \mathbb{E}_M^*\left[\frac{(\frac{G(M)}{M}\cdot\tau)^b}{b!}\, e^{-\frac{G(M)}{M}\cdot\tau}\right]. \tag{4.22}$$

The subscript at $\mathbb{E}$ specifies the random variable the operator refers to. The average number of mRNA-specific translation events during the lifespan of the mRNA molecule, which is exponentially distributed with mean ν, equals[7]:

$$\begin{aligned}
\alpha(\bar{m}) \;&=\; \mathbb{E}_{\tau_m}\left[\mathbb{E}_{S_B}\left[\mathbb{E}_M^*\left[\mathrm{Pois}_{\left(\frac{G(M)}{M}\cdot\tau_m\right)}(S_B)\right]\right]\right] \;=\; \mathbb{E}_{\tau_m}\left[\mathbb{E}_M^*\left[\mathbb{E}_{S_B}\left[\mathrm{Pois}_{\left(\frac{G(M)}{M}\cdot\tau_m\right)}(S_B)\right]\right]\right] \\
&= \mathbb{E}_{\tau_m}\left[\mathbb{E}_M^*\left[\frac{G(M)}{M}\cdot\tau_m\right]\right] \;=\; \mathbb{E}_{\tau_m}\left[\sum_{n=0}^{\infty}\frac{G(n+1)}{n+1}\,\frac{\bar{m}^n}{n!}e^{-\bar{m}}\cdot\tau_m\right] \\
&= \int_0^{\infty}\sum_{n=0}^{\infty}\frac{G(n+1)}{n+1}\,\frac{\bar{m}^n}{n!}e^{-\bar{m}}\cdot\frac{\tau_m}{\nu}\,e^{-\frac{\tau_m}{\nu}}\,d\tau_m \\
&= \sum_{n=0}^{\infty}\frac{G(n+1)}{n+1}\,\frac{\bar{m}^n}{n!}e^{-\bar{m}}\,\nu \;=\; \begin{cases}\frac{\bar{G}(\bar{m})}{\bar{m}}\,\nu & \text{if } \bar{m} > 0\\ G(1)\,\nu & \text{if } \bar{m} = 0.\end{cases}
\end{aligned} \tag{4.23}$$

If $G(m) = g \cdot m$ was linear, the formula would reduce to $\alpha(\bar{m}) = g \cdot \nu$. The burst size is thus

[6] The validity and implications of this assumption will be considered later in Section 4.5.5.

[7] ν is the inverse of the degradation constant in the scaled system.

only state-dependent if the translational propensity G is nonlinear.

4.5.2. State-dependent burst frequency

The burst frequency is the mean number of bursts (i.e., of transcription events) occurring in an interval $[0, \tau_s]$, whose average length corresponds to the average lifespan of a protein. Under feedback regulation, the number of transcription events depends on the protein level, so that the burst frequency becomes state-dependent. Let $\bar{s}$ be the current mean copy number around which protein fluctuations take place. Assuming that the average transcriptional propensity does not significantly change during an interval of length τ_s, this propensity is simply approximated by the point evaluation $\frac{1}{\nu} \cdot F(\bar{s})$, like in the formulation of the hybrid deterministic model. The lifespan of a protein is exponentially distributed with mean 1. The state-dependent burst frequency can now be determined in analogy to the calculation of the state-dependent burst size:

$$\omega(\bar{s}) \;=\; \int_0^\infty \frac{1}{\nu}\, F(\bar{s})\, \tau_s\, e^{-\tau_s}\, d\,\tau_s \;=\; \frac{F(\bar{s})}{\nu}. \tag{4.24}$$

$\alpha(\bar{m})$ and $\omega(\bar{s})$ are generalizations of the expressions formulated in [Friedman et al., 2006; Thattai and van Oudenaarden, 2001], where only systems without feedback ($F \equiv a$) and with linear propensities ($G(m) = g \cdot m$) are studied, leading to state-independent burst measures.

4.5.3. Stationary burst measures

Next, let us regard local bursts in stationary state. Let (m^*, s^*) denote a stable fixed point of the hybrid deterministic model, around which the system fluctuates. It can be assumed that the fixed point is positive-valued, since otherwise, bursts would not occur at all. Since the steady state condition implies that $\bar{G}(m^*) = s^*$ and $F(s^*) = m^*$ hold, the following relations for the local stationary burst characteristics $\alpha^* := \alpha(m^*)$ and $\omega^* := \omega(s^*)$ are obtained:

$$\alpha^* = \frac{\bar{G}(m^*)}{m^*} \cdot \nu = \frac{s^*}{m^*} \cdot \nu = r^* \cdot \nu, \qquad \omega^* = \frac{F(s^*)}{\nu} = \frac{m^*}{\nu} = \frac{s^*}{r^* \cdot \nu}. \tag{4.25}$$

If s^* is held fixed, the burst properties can be fully characterized in terms of the quantities r^* and ν, i.e. in terms of the stationary protein-to-mRNA ratio and the relative time-scale of mRNA dynamics. In this case, it is sufficient to concentrate on the determination of the burst size, since the burst frequency follows directly from the relation $s^* = \alpha^* \cdot \omega^*$. This means that in order to keep s^* constant, any change of α^* needs to be balanced by an inverse change of ω^*.

4.5.4. Relation between stationary burst characteristics and stationary propensities and its usage for a graphical determination of the burst size

The occurrence of bursts can be related to differences in the average propensities of reactions on the mRNA versus the protein level: In the stationary state, the average propensities of transcription and mRNA degradation, $\mathbb{E}^*[\frac{1}{\nu} \cdot F(S)]$ and $\mathbb{E}^*[\frac{1}{\nu} \cdot M]$, are necessarily identical. We denote this value by w_m^*, since it reflects mRNA dynamics. In line with this, the average propensities of translation and protein degradation, $\mathbb{E}^*[G(M)]$ and $\mathbb{E}^*[S]$, are equal as well and their value is denoted by w_s^*. Usage of the same approximations as in the hybrid deterministic model yields the relations $w_m^* \approx \frac{1}{\nu} \cdot F(s^*) = \frac{1}{\nu} \cdot m^*$ and $w_s^* \approx \bar{G}(m^*) = s^*$. Since $\bar{G}$ is invertible,

the following two conditions for w_m^* and w_s^* are obtained:

$$w_m^* = \frac{1}{\nu} \cdot F(w_s^*), \qquad w_m^* = \frac{1}{\nu} \cdot \bar{G}^{-1}(w_s^*). \tag{4.26}$$

A connection between propensities and stationary burst characteristics is now given by:

$$\alpha^* = r^* \cdot \nu = s^*/m^* \cdot \nu = w_s^*/w_m^*. \tag{4.27}$$

This relation illustrates that in the stationary state, bursts with large size occur if reactions on the protein level are much more probable and therefore expected to be much more frequent than reactions on the mRNA level.

The steady state propensities can be graphically determined by visualizing the two equations in (4.26) in a w_s^*-w_m^*-plot, and by reading the values at the intersection points of the two graphs. The stationary burst size α^* is then equal to the inverse slope of the line connecting the intersection point with the origin. The plot helps to visualize how the shapes of F and $\bar{G}$ and the interplay of relative reaction time-scales affect the burst size. Examples are shown later in Fig 5.4 (B) and in Fig 5.5.

4.5.5. Quality of burst measures

The state-dependent burst measures provide information about the expected size and frequency of protein bursts based on the current average system state $(\bar{m}, \bar{s})$, under the premise that the average translational propensity per mRNA and the average transcriptional propensity do not significantly change during a burst or during the lifespan of a protein. Is this premise justified?

First of all, the average translational propensity per mRNA, $\mathbb{E}^*\left[\frac{G(M+1)}{M+1}\right]$, is obviously constant if G is linear. However, even if this is not the case, the expression is expected to remain unchanged during a burst if the system is stationary and if F is constant: In this case, the mRNA distribution is not influenced by the production of burst proteins.

Apparently, constant F also leads to a constant average transcriptional propensity $\mathbb{E}^*\left[\frac{F(S)}{\nu}\right]$. Moreover, it is the main criterion ensuring a high quality of the hybrid deterministic approach, whose principles have been used again in order to locally approximate the averaged reaction propensities with effective reaction rates $\frac{\bar{G}(\bar{m})}{\bar{m}}$ and $\frac{1}{\nu}F(\bar{s})$. Taken together, both burst measures $\alpha(\bar{m})$ and $\omega(\bar{s})$ are expected to be of high accuracy if the effect of feedback is locally weak.

What if F is not constant? In this case, the mRNA level is regulated and may systematically change during a burst. If additionally G is nonlinear, the effective rate of translation per mRNA might accordingly increase or decrease, leading to larger or smaller bursts than predicted by $\alpha(\bar{m})$.

Concerning $\omega(\bar{s})$, the relation $\mathbb{E}^*[F(\,.\,)]) = F(\mathbb{E}^*[\,.\,])$ is usually not fulfilled, so that the burst size may be over- or underestimated, depending on the local curvature of F.

In contrast to the simulation-based validation of estimated Fano factors, the comparison between analytically approximated burst sizes and values extracted from experimental or simulated time course data is difficult, since the production of burst proteins is superposed by degradation events and by other, overlapping bursts. In Section 5.2.1, a procedure for the extraction of burst characteristics from simulations is explained – however, its applicability is limited and the obtained values are only approximations as well, so that the procedure is not suitable for system-

atically and comprehensively evaluating the quality of the burst measures. An "indirect" way of quality control however arises from the model reduction explained in the following section, which is performed with the help of the state-dependent burst measures.

4.6. Model reduction and determination of modes in the burst limit

In Chapter 3, we have seen that based on a reduced model, which only describes burst-like protein production and simple degradation without explicitly considering the mRNA level (cf. the system given in (3.4)), modes in the protein distribution can be determined. Here, it will be examined whether and how this procedure can be applied in case of state-dependent burst sizes.

4.6.1. Model reduction

Like in the model reduction explained in Section 3.1, the condition $\nu \ll 1$ needs to be fulfilled in order to describe a translational burst as an instantaneous increase in the protein time course: In this case, mRNA molecules are degraded very fast compared to the protein, so that all translation events belonging to one burst occur in a temporally highly condensed manner (on a time-scale determined by protein degradation). The general form of the reduced model should be:

$$\boxed{\begin{array}{lll} \emptyset & \xrightarrow{\frac{1}{\nu}\cdot F(s)\cdot p_b^B} & b\cdot\text{Protein}, \qquad b\in\mathbb{N}_0 \\ \text{Protein} & \xrightarrow{\quad 1 \quad} & \emptyset \end{array}} \tag{4.28}$$

What remains to be defined is p_b^B, the probability of a burst of size $B = b$. Note that according to the notation in Section 4.5, $B = S_B(\tau_m)$.

We already know that if $G(m) = g \cdot m$ is linear, the average burst size $\alpha = g \cdot \nu$ does not depend on the current mRNA level and B is geometrically distributed [McAdams and Arkin, 1997], cf. Section 2.2.2. We will first review how the geometric distribution can be deduced, before the calculations are transferred to systems with expression-state-dependent bursts. Let us recall the reaction in which the burst proteins are generated in case of linear G:

$$\emptyset \quad \xrightarrow{g=\frac{\alpha}{\nu}} \quad \text{Burst proteins.} \tag{4.29}$$

The burst ends with the degradation of the responsible mRNA molecule (propensity equal to $1/\nu$). Comparing the two reactions, namely protein formation and mRNA degradation, the probability that the next reaction happening will be translation is equal to $\frac{g}{1/\nu+g}$, while the probability that it will be mRNA degradation is equal to $\frac{1/\nu}{1/\nu+g}$. A burst of size b is hence generated when b protein formation events occur before the mRNA molecule is degraded. As a consequence,

$$p_b^B = \left(\frac{g}{\frac{1}{\nu}+g}\right)^b \frac{\frac{1}{\nu}}{\frac{1}{\nu}+g} = \left(\frac{\alpha}{1+\alpha}\right)^b \frac{1}{1+\alpha} = \text{Geo}_\alpha(b), \tag{4.30}$$

i.e. the burst size is indeed geometrically distributed for linear G.

If G is nonlinear, the burst size $\alpha(\bar{m})$ varies with the current average mRNA copy number $\bar{m}$. Since our goal is to fully eliminate the mRNA level, assumptions about $\bar{m}$ need to be made for every point in time τ, based solely on the knowledge of the current protein copy number s. The assumption which will be used is that in the time interval between two consecutive bursts, $\bar{m} = F(s)$ holds. There are two scenarios in which this equality is valid: In the first, the overall mRNA dynamics are much faster than protein degradation, so that the mRNA distribution adapts instantaneously to the current protein copy number in the period between two bursts (pseudo steady state). In the second scenario, F is constant in the range of local protein fluctuations, meaning that the local mRNA distribution is not influenced by the protein level at all. Furthermore, it will be assumed that the mRNA distribution does not change significantly during a burst (like in the preceding chapter) – either again due to the missing dependence of the average mRNA level on the protein level, or due to a certain delay in the mRNA dynamics[8].

Under these assumptions, the translation of burst proteins occurs with propensity $\frac{\bar{G}(F(s))}{F(s)} = \frac{1}{\nu} \cdot \alpha(F(s))$, and the burst size is again geometrically distributed:

$$p_b^B = \left(\frac{\alpha(F(s))}{1+\alpha(F(s))}\right)^b \frac{1}{1+\alpha(F(s))} = \mathrm{Geo}_{\alpha(F(s))}(b). \tag{4.31}$$

The CME of the reduced model then reads:

$$\begin{aligned} \dot{p}_s \;=& \sum_{b=0}^{s} \frac{1}{\nu} \cdot F(s-b) \cdot \mathrm{Geo}_{\alpha(F(s-b))}(b) \cdot p_{s-b} \;-\; \frac{1}{\nu} \cdot F(s)\, p_s \\ &+\; (s+1)\, p_{s+1} \;-\; s\, p_s. \end{aligned} \tag{4.32}$$

In case the above assumptions do not hold, the general quality of the reduced model needs to be critically evaluated, which will be done in Section 4.7.

4.6.2. Solving the CME of the reduced model

We already know from Chapter 3 the recursive and explicit solutions of the CME in case the mean burst size α is constant, namely:

$$\begin{aligned} p_1 \;&=\; \frac{\alpha}{1+\alpha}\, \frac{F(0)}{\nu}\, p_0 \\ (s+1)\, p_{s+1} \;&=\; \frac{\alpha}{1+\alpha} \left(\frac{F(s)}{s\,\nu} + 1\right) s\, p_s, \quad s \in \mathbb{N} \end{aligned} \tag{4.33}$$

and

$$p_s \;=\; \frac{F(0)}{s\,\nu}\, p_0 \left(\frac{\alpha}{1+\alpha}\right)^s \prod_{i=1}^{s-1} \left(\frac{F(i)}{i\,\nu} + 1\right) \tag{4.34}$$

(cf. Eqs (3.6) and (3.7), whose notations have been adapted to those of the present chapter). As we have seen, fixed average burst sizes are obtained in systems without feedback (F constant) or with linear translational propensity function G.

[8] Note that this delay needs to be so small that in between bursts, the pseudo steady state assumption is still applicable.

For bursts that depend on the protein level, an explicit calculation is not possible. Inspired by the above equations, one may, however, assume that in this case, the approximate relations

$$\begin{aligned} p_1 &= \frac{\alpha(F(0))}{1 + \alpha(F(0))} \frac{F(0)}{\nu} p_0 \\ (s+1)\, p_{s+1} &= \frac{\alpha(F(s))}{1 + \alpha(F(s))} \left(\frac{F(s)}{s\,\nu} + 1 \right) s\, p_s, \quad s \in \mathbb{N} \end{aligned} \tag{4.35}$$

and

$$p_s = \frac{F(0)}{s\,\nu} p_0 \frac{\alpha(F(0))}{1+\alpha(F(0))} \prod_{i=1}^{s-1} \left[\frac{\alpha(F(i))}{1+\alpha(F(i))} \cdot \left(\frac{F(i)}{i\,\nu} + 1 \right) \right] \tag{4.36}$$

hold. As a control, these expressions may be inserted into the stationary form of the reduced CME (4.32). It turns out that they approximately fulfill the steady state condition if $\alpha(F(s)) \approx \alpha(F(s+1)) \;\; \forall\, s \in \mathbb{N}_0$.

4.6.3. Estimation of modes

Based on Eq (4.35), we may – in analogy to the previous study – conclude that the locations of positive modes approximately fulfill the condition

$$s = \lceil \sigma \rceil \qquad \text{where} \quad \sigma \in \mathbb{R}, \quad \sigma + \alpha(F(\sigma)) + 1 = \frac{\alpha(F(\sigma))\, F(\sigma)}{\nu} = \bar{G}(F(\sigma)), \tag{4.37}$$

cf. Eq (3.13). A further mode is located at $s = 0$ if $p_0 > p_1$, i.e. if

$$\alpha(F(0)) + 1 > \frac{\alpha(F(0))\, F(0)}{\nu} = \bar{G}(F(0)). \tag{4.38}$$

Like in Chapter 3, plotting the left and right hand-side of the equation in (4.37) helps to understand the connection between circuit properties and the location of modes in the protein distribution, which provide important indications on skewness. Since the LHS is not necessarily linear anymore, but dependent on $\alpha(F(\sigma))$, the degree of skewness may vary between different expression states.

Using the determination of modes, potential discrepancies between multistability and multimodality, which would hamper the applicability of the hLNA, can be identified.

4.6.4. Applying the hLNA to the reduced model

The hLNA method can also be applied to the reduced model. Note that the stoichiometric matrix of the reduced reaction system (4.28) is in this case infinite-dimensional, since the actual jump size of a burst event can in principle be equal to any nonnegative integer. The calculations are shown in the Appendix, Section B.4. It turns out that the hybrid deterministic model reads:

$$\dot{c}_s = \bar{\mathcal{G}}(\mathcal{F}(c_s)) - c_s. \tag{4.39}$$

This formulation corresponds to the ODE which is obtained by assuming a PSS for c_m in the full hybrid deterministic model, Eq (4.11).

The Fano factor of the reduced model obeys:

$$\eta^*(S) \;=\; 1 + \frac{1}{(g^*)^{-1} - f^*}\left(f^* + \frac{r^*\,\nu}{g^*}\right). \tag{4.40}$$

Here, $g^* := \frac{d\,\bar{G}(c_m)}{d\,c_m}|_{c_m=\mathcal{F}(c_s^*)}$ and $r^* := \frac{s^*}{F(s^*)}$. If the condition $\nu \ll 1$ is used to simplify Eqs (4.19) and (4.20), which are the formulae of the protein Fano factors according to the full model, this yields:

$$\eta^*(S) \;=\; 1 + \frac{1}{(g^*)^{-1} - f^*}\left(f^* + \frac{g^*\,\nu}{r^*}\right). \tag{4.41}$$

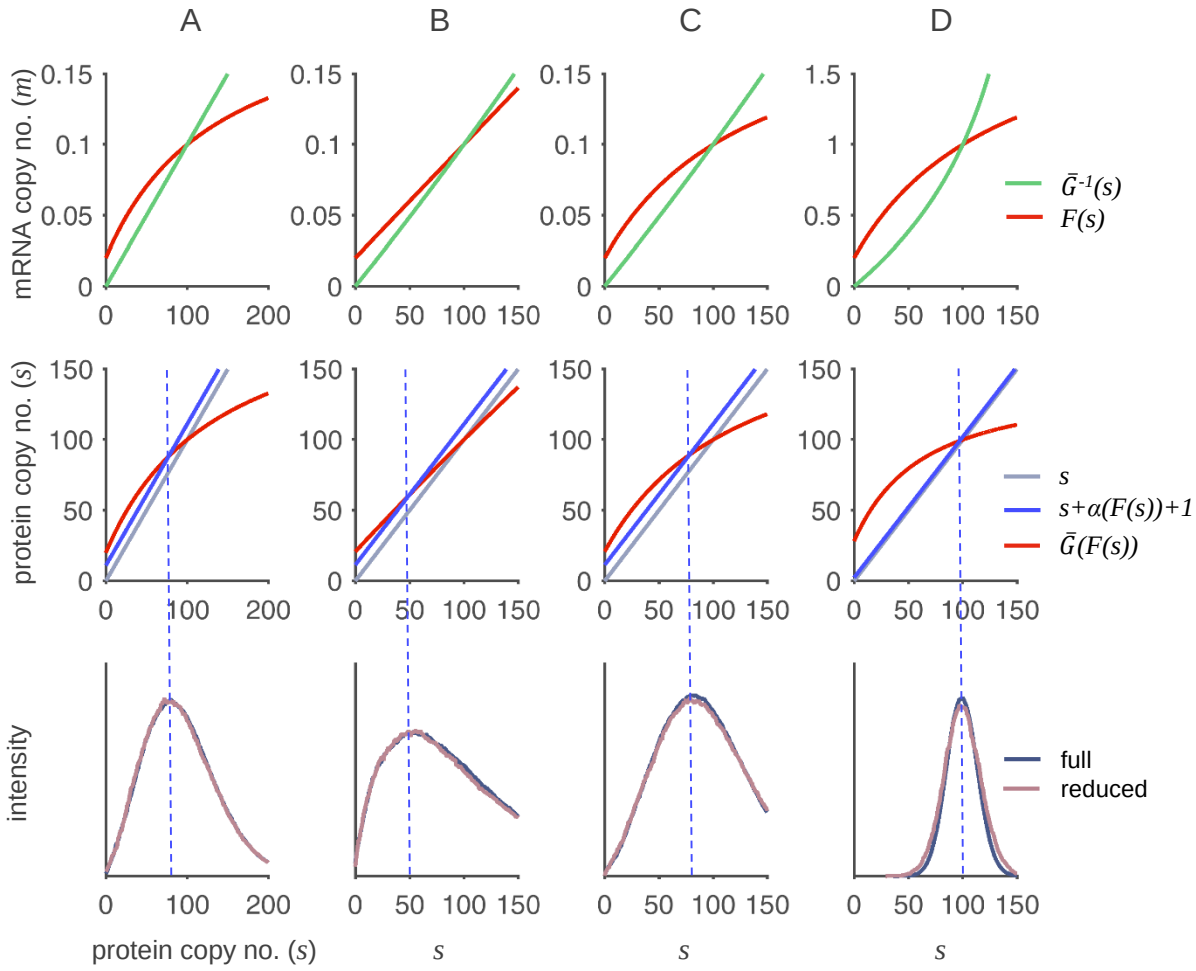
The expression differs from Eq (4.40) if $g^* \neq r^*$. This discrepancy will be discussed in the next section, where the quality of model reduction is assessed.

4.7. Quality of model reduction and of the determination of modes

In order to be able to perform the model reduction, the condition $\nu \ll 1$ is crucial, since it justifies the temporal condensation of multiple translation events into a single reaction with burst-like protein production. Besides that, there are other criteria which should be fulfilled:

The proposed techniques of model reduction and of mode determination make use of the newly defined state-dependent burst size, whose quality is hence decisive. In Section 4.5.5, we have discussed that the formula (4.23), with which the burst size is calculated, is accurate if G is linear, or if F is constant and the mRNA level is stationary. Both conditions also allow the mRNA level to be easily eliminated from the model. If G was nonlinear and F was non-constant, the mRNA level could still be averaged out by assuming a pseudo-stationary mRNA distribution. This assumption is justified if the mRNA dynamics adapt quickly to that of the protein while the protein copy number changes due to degradation. At the same time, the mRNA distribution should remain unaffected during protein bursts to ensure the geometric distribution of the burst size. These two requirements on the mRNA dynamics seem contradictory, but they are still feasible, since delay effects may only come into play during the high-frequency-production of burst proteins. However, as soon as the mRNA level is sensitive to the sudden increase in the protein level, the burst size distribution is not geometric anymore. In the case of positive feedback, for example, the mean mRNA copy number $\bar{m}$ would actually rise – at least slightly – during burst protein production. If $\bar{G}$ was locally concave, the effective average translation rate from a single mRNA, $\frac{\bar{G}(\bar{m})}{\bar{m}}$, would decrease accordingly. The real protein burst size would thus be smaller than predicted by the geometric distribution. The opposite holds if $\bar{G}$ was locally convex. The local protein distribution of the reduced model would therefore be too broad or too narrow, compared with the distribution of the full model.

Indeed, this fact comes to light as a quantifiable systematic error in model reduction when the protein Fano factors of the full and reduced models are compared: Eqs (4.40) and (4.41) show that $\eta^*(S)$ of the reduced model is systematically biased if $g^* \neq r^*$, which is in general the case if $\bar{G}$ is nonlinear. In Fig 4.7, a comparison between simulated protein distributions according to the full and the reduced model are shown for four qualitatively different systems. Here, $\nu = 0.01 \ll 1$. The following circuits have been tested:



	calc.	simul.	calc.	simul.	calc.	simul.	calc.	simul.	
$\eta^*(S)$	20.4	21.3 ± 1.2	42.1	38.4 ± 3.6	18.0	19.5 ± 1.1	1.89	2.04 ± 0.12	full
	20.6	21.5 ± 1.3	46.3	42.5 ± 3.8	19.8	20.8 ± 1.1	2.78	2.82 ± 0.60	reduced

Figure 4.7.: **Quality of model reduction and mode prediction.** Four qualitatively different reaction systems are evaluated, which are specified in the text. Top row: s-m-phase-plots according to the hybrid deterministic model. Second row: Plots for the determination of modes. Third row: Histograms of the full and reduced models. To generate the distributions, protein time courses of each (full or reduced) system have been simulated with final reaction time $\tau_f = 10^3$. Protein Fano factors are given in the tables, which were obtained either analytically ("calc.") using Eq (4.19) for the full model and Eq (4.40) for the reduced model, or computed based on the simulated distribution ("simul."), together with the SEMs. Histograms of full and reduced systems are nearly congruent if $\bar{G}$ is linear as in system (A), while the distribution of the reduced system is broader if $\bar{G}$ is concave as in (D). Nonlinearity of F does not hamper the quality of model reduction (cf. (B) and (C)). The prediction of modes worked reliably in all four cases.

(A) concave F, linear G: $F(s) = 0.02 + 0.192\,\frac{s}{s+140}$, $G(m) = 1000 \cdot m$

(B) linear F, concave G: $F(s) = 0.02 + 0.0008 \cdot s$, $G(m) = 1153\,\frac{m}{m+0.1}$

(C) concave F and G: $F(s) = 0.02 + 0.192\,\frac{s}{s+140}$, $G(m) = 1153\,\frac{m}{m+0.1}$

(D) concave F and G: $F(s) = 0.2 + 1.92\,\frac{s}{s+140}$, $G(m) = 170.3\,\frac{m}{m+0.1}$

In all systems, the average protein copy number s^* is approximately equal to 100. In systems (A) to (C), $m^* \approx 0.1$, while in (D), $m^* \approx 1$. The top row of Fig 4.7 shows the protein-mRNA phase plots of all four systems. In these plots, $\bar{G}$, but not G is visualized, since the latter function would mostly lie outside the range of the values shown. The seeming linearity of the effective translation rate in panels (B) and (C) is due to local averaging and due to the very low mean mRNA copy number. In panel (D), the curvature of $\bar{G}$ is comparatively larger. The middle row of plots illustrates the condition (4.37) for the determination of modes, which are approximately located at the intersection point of the red and the blue line. The plots in the third row show the simulated protein distributions according to the full and the reduced model. For both model types, the protein Fano factors are given in the table below: In each panel, the column on the left contains the values calculated using Eqs (4.19) and Eq (4.40), the one on the right contains the values extracted from the histograms. Differences between the analytically determined and the numerically estimated values are mostly attributable to the strong feedback present in all four circuits.

In the system illustrated in Fig 4.7 (A), the linearity of G makes the average burst size independent of the mRNA level. As a consequence, the simulated protein distributions according to the full and reduced model are almost identical. Minor differences emerge due to the stronger temporal condensation of translation events in the reduced model and the resulting negligence of protein degradation during a burst. This leads to a slight overestimation of noise, which is why the Fano factors are not exactly equal. The prediction of modes in panel (A) is accurate.

In panel (B), the protein distribution of the reduced model is slightly broader than that of the full model. This is due to the fact that G and, as a consequence, $\bar{G}$ is concave, so that $g^* < r^*$ holds. According to the systematic deviation between Eqs (4.40) and (4.41), the reduced model is indeed expected to overestimate the protein fluctuations. The effect is however moderate in the regarded case since $\bar{G}$ is almost linear in the range of mRNA fluctuations (although G is highly nonlinear). The prediction of the mode still works very well.

The system visualized in panel (C) combines the concave transcription rate of panel (A) with the concave translation rate of (B). Again, the distribution of the reduced model is slightly broader, but mode prediction is reliable. Apparently, nonlinearity of F does not hamper the quality of model reduction, as it does not make use of the average burst frequency.

In Fig 4.7 (D), m^* has been increased, while s^* was kept constant. The burst size is therefore reduced, so that the mode of the protein distribution is located closely to the mean value. In contrast to panels (B) and (D), concavity of $\bar{G}$ is more pronounced, and as expected, the distribution of the reduced model is now significantly broader than that of the full model. Nevertheless, mode prediction works reliably.

Obviously, the correspondence between the distributions of the full and reduced model is good except for systems in which positive (or negative) feedback is combined with strong nonlinearity of $\bar{G}$. Although one might expect that the quality of the determination of modes is closely associated with the quality of model reduction, the estimation of modes was surprisingly accurate in all studied cases.

4.8. Short discussion and summary

Since the classical LNA approach for the approximation of local variances often performs poorly in mesoscopic systems where low-copy-number-species like mRNA are involved (cf. Figs 4.4 - 4.6), a novel approach has been developed that is based on an improved underlying deterministic description. This hybrid deterministic model takes the stochasticity and discreteness of mRNA copy numbers into account and thereby generates variables that are usually closer to the local stochastic mean values. It should be pointed out that the deterministic description is only used for linearization of the stochastic model and not for replacing a part of the stochastic variables, as it was done e.g. in [Hellander and Lötstedt, 2007; Jahnke, 2011; Lin and Galla, 2016] (cf. the reduced models presented in Section 2.1.6). The assumptions made during the establishment of the hybrid deterministic model also served as a basis for burst quantification and model reduction. As a result, profound connections between different noise measures can now be identified and biologically interpreted. This will be elaborated on in the next chapter.

The development of the hybrid approaches and of state-dependent burst characterization was based on several assumptions, which suggest that the methods work best if the transcriptional propensity function is locally almost constant in a domain around the stable fixed points. Importantly, this criterion is not only fulfilled in systems without feedback, but also in bistable systems with cooperative autoregulation, cf. Fig 4.4. Moreover, simulation studies have shown that even if the transcriptional propensity function is locally increasing, the mean value and Fano factor estimates and the distributions obtained through model reduction are still rather accurate. In all cases, the hybrid approach led to much better noise estimates than the classical LNA approach through the improved description of locally averaged systems behavior.

Approaches other than the ones proposed, for example effective mesoscopic rate equations (EMREs) [Grima, 2010] and the inverse omega square (IOS) method [Thomas et al., 2012] (cf. Sec 2.1.6) have also led to a description of means and variances that are more accurate than classical LNA. Compared with those methods, the hLNA approach is much more pragmatic, as it simply assumes the mRNA copy number to be Poisson distributed. In return, it is easily applied without the need for extensive Taylor expansion. Moreover – and this might be its main advantage – it is suitable for analyzing multimodal systems.

However, in order to apply the hybrid approaches, this multimodality needs to be associated with hybrid deterministic multistability. In this case, the multimodal distribution cannot be easily decomposed into unimodal PMFs, which aggravated the rigorous assessment of the quality of estimates[9]. This is why the evaluation was mainly based on simulations.

In cases where multimodality and (hybrid deterministic) multistability are divergent, (which happens especially in the presence of large translational bursts), the hLNA method is not capable of characterizing the noise in all peaks of the protein distribution anymore. However, it seems that the locations of modes are predicted reliably, provided that the requirements for model reduction are fulfilled (in particular $\nu \ll 1$). An *a priori* comparison of the number and location of modes and stable hybrid deterministic fixed points can thus tell in advance whether the application of the hLNA is appropriate.

[9]By contrast, a multimodal distribution that is generated by a slowly switching component (e.g. promoter state) can be modeled as a superposition of unimodal distributions with the help of conditional probabilities [Hasenauer et al., 2014; Thomas et al., 2014].

4.9. Outlook: Transferability and possible extensions of the approaches

Even though the model system used here comprised only one mRNA and protein species, the approach can be extended to multi-gene systems, see Chapter 6. Furthermore, systems with more than two stable expression states can be analyzed as long as the hybrid deterministic model reflects this multimodality. The range of reaction components and reactions is not restricted either: For example, nonlinear degradation rates can be treated in an analogous fashion, additional intermediate reactions or reactants can be included, the promoter state can be integrated, and transcriptional bursts can be analyzed in rather the same way as translational bursts.

A major restriction of the hLNA approach originates from the assumption that the reacting species, whose large noise level makes the calculation of effective rates necessary, is Poisson-distributed. This approximation might work well for the local mRNA level, but is not reasonable for, e.g., the protein distribution, for which we have simply applied the small noise assumption. This is not always justified, as Fig 4.5 (C) has shown. For simple reaction systems, the protein distribution has been proven to be well approximated by a negative binomial (NB) distribution [Aquino et al., 2012; Ochab-Marcinek and Tabaka, 2010; Paulsson and Ehrenberg, 2000; Shahrezaei and Swain, 2008]. It therefore seems natural to just replace the Poisson distribution accordingly. However, unlike in the Poisson case, the variance of an NB distribution is decoupled from the mean value – it can take any value larger than the average. The NB PMF with mean value s^* and Fano factor η^* is given by

$$\mathrm{NB}_{s^*,\eta^*}(r) = \frac{\Gamma(r+\frac{s^*}{\eta^*-1})}{\Gamma(\frac{s^*}{\eta^*-1})\,\Gamma(r+1)} \left(\frac{1}{\eta^*}\right)^{\frac{s^*}{\eta^*-1}} \left(1-\frac{1}{\eta^*}\right)^{r}. \tag{4.42}$$

Trying to estimate the variance using an NB-based hLNA just ends up in a dilemma, since the hLNA requires the formulation of a hybrid deterministic model, which in turn requires a variance estimate for the NB distribution. One possible way to overcome this predicament is to start with some initial variance estimate and to then use an iterative cycle that subsequently adapts the mean and the average value until the chosen values do not lead to obvious inconsistencies anymore. A computational procedure that implements this cycle has been developed and tested in a first attempt. The pseudoalgorithm and some preliminary results are given in Section B.5 of the Appendix. In one test system, the obtained average and noise estimates could be improved significantly compared to the values obtained with Poisson-based hLNA. The NB-based method seems to be promising, but the conditions under which the algorithm converges still need to be evaluated.

5. Design of noise patterns in a single-gene autoregulatory system

In the following, we will discuss the biological implications of the results obtained in the preceding chapter, continuing with the same model system. It has been introduced and formulated in Section 4.1. First of all, the formula of the mRNA and protein Fano factors will be analyzed with regard to their dependence on circuit parameters and functions. Then, we will see how the burst characteristics can – under certain conditions – be extracted from time course data, which helps to interpret them. The dependence of burst sizes and frequencies on circuit parameters will be discussed as well. Finally, connections between Fano factors and protein burst characteristics will be drawn. Based on these biological insights, circuits with different noise patterns will be designed *in silico*. With the term "noise pattern", the qualitative distribution of noise levels among different expression states (e.g. large protein noise in the low expression state, small noise in the high expression state) is described. Since the noise level is a qualitative measure for the robustness of an expression state, we can make rough predictions on the phenotypic heterogeneity of a population. Throughout this chapter, it is assumed that the multimodality of the considered systems is associated with hybrid-deterministic multistability (cf. Chapter 4), so that the hLNA method can be applied.

The single-gene autoregulatory feedback system (4.1) studied here is a common regulatory motif, for which several biological examples exist. One of those, namely the regulatory system in *Streptococcus mutans* (*S. mutans*) that triggers genetic competence, will be part of the subsequent discussion. In Chapter 6, we will then apply the methods to a two-gene system, the genetic toggle-switch.

The results presented in the following have mostly been published in [Hortsch and Kremling, 2018b].

5.1. Interpreting the influence of circuit properties on Fano factors

The Fano factors given in Eqs (4.18) - (4.20) depend on the four parameters f^*, g^*, r^*, and ν. f^* and g^* are measures for the local sensitivities of the transcriptional and translational propensity functions with respect to fluctuations on the protein and mRNA level, respectively. Note that if $G = g \cdot m$ is linear, the equality $r^* = g^* = g$ holds, so that the degree of freedom in the formulae is reduced. In the following, we will examine the influence of each of the circuit parameters on the local noise level.

Transcriptional sensitivity f^*:

First, we regard the effect of f^*, which specifies the mode and sensitivity of autoregulation. $f^* = 0$ means that local feedback is missing. We already know that without feedback, the stationary mRNA level is Poisson distributed. In accordance to that, Eq (4.18) reduces to

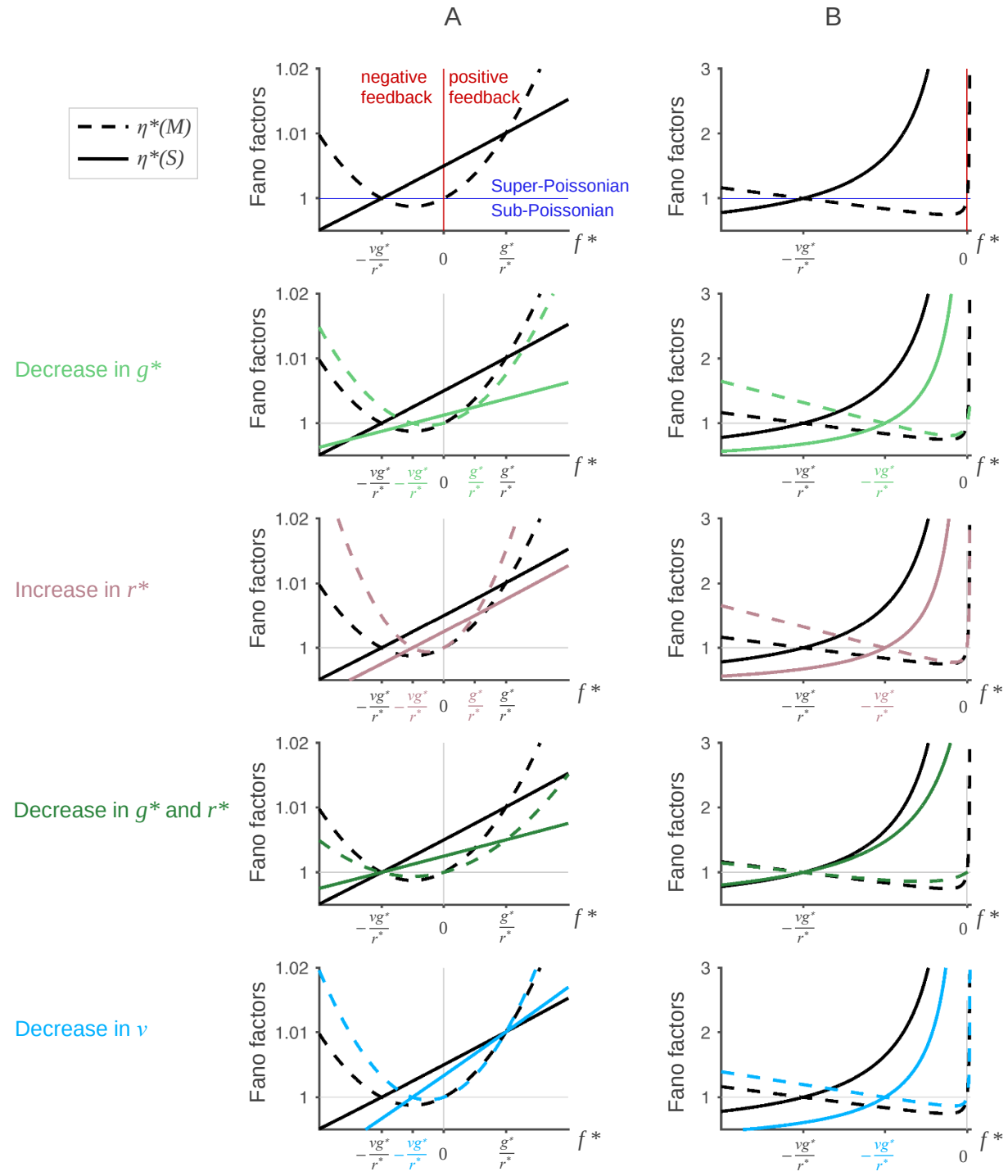


Figure 5.1.: **Dependence of Fano factors on circuit properties.** In the first row, the dependence of $\eta^*(M)$ (dashed line) and $\eta^*(S)$ (solid line) on the transcriptional sensitivity f^* with fixed g^*, r^*, and ν is shown for two different sets of parameters: (A) $g^* = 1$, $r^* = 100$, $\nu = 1$ and (B) $g^* = 100$, $r^* = 100$, $\nu = 0.5$. The lower plots show variations in the remaining parameters (colored lines) in order to illustrate their effect. Compared to the first row, the following changes are made: In the second row, the translational sensitivity g^* is halved. In the third row, the stationary protein-to-mRNA ratio r^* is doubled. In the fourth row, both g^* and r^* are replaced by one-half (panel (A)) or one-tenth (panel (B)) of their original values. In the last row, the time-scale parameter ν is halved.

$\eta^*(M) = 1$. At the same time, the protein distribution is super-Poissonian ($\eta^*(S) > 1$) due to noise propagated from the mRNA to the protein level via translation. Positive feedback ($f^* > 0$) increases both the mRNA and protein Fano factor due to cyclic forward propagation of noise, until finally, both Fano factors tend to infinity in the limit[1] $f^* \to (g^*)^{-1}$. In the case of negative feedback ($f^* < 0$), up- or downward fluctuations on the protein level cause an inverse effect on the mRNA copy number, whose fluctuations are then immediately passed back onto the protein level. Altogether, this results in the attenuation of protein noise. It can even reach sub-Poissonian levels ($\eta^*(S) < 1$) for strong negative feedback with $\frac{\nu}{1+\nu}$ as a lower bound. The effect of autorepression on mRNA noise is slightly more complex: Reducing f^* (< 0) first causes a decrease in mRNA noise as well, but once f^* has crossed a certain negative threshold[2], the mRNA Fano factor starts to rise again. Apparently, strong protein noise attenuation can be achieved at the expense of enlarged mRNA fluctuations[3]. From the mRNA perspective, this can be regarded as over-compensation.

In the first row of Fig 5.1, the graphs of $\eta^*(M)$ and $\eta^*(S)$ are plotted over f^*, while all other circuit parameters are held constant. In panels (A) and (B), the graphs belonging to two different parameter sets are illustrated. Based on this representation, one can determine the regions where fluctuations are super- or sub-Poissonian. Depending on the circuit parameters, either one or two reference points are available, which can help to draw the graphs and to thereby roughly predict the influence of the other parameters: First of all, $\eta^*(M) = \eta^*(S) = 1$ is fulfilled for $R^* := r^* \frac{f^*}{g^*} = -\nu$. Moreover, $\eta^*(M) = \eta^*(S) > 1$ holds in case $R^* = 1$, but this condition is only relevant if the stability criterion is fulfilled, i.e. if $f^* < (g^*)^{-1}$ holds at the same time. In the first row of Fig 5.1 (A), $f^* = \frac{g^*}{r^*} = \frac{1}{100} < 1 = (g^*)^{-1}$, so that the second reference point is available, but this is not the case in Fig 5.1 (B), where $f^* = \frac{g^*}{r^*} = 1 > 0.01 = (g^*)^{-1}$. As a result, the protein Fano factor in the latter system is always larger than the mRNA Fano factor under positive or sufficiently weak negative feedback.

Translational sensitivity g^* (in case $g^* \neq r^*$):

Now, the influence of the remaining parameters on noise is studied. Let us first assume that the two parameters g^* and r^* are uncoupled, so that they can be considered separately. This is only possible if G is nonlinear.

The effect of g^* under positive feedback is intuitive: An increase in g^* augments noise propagation through the circuit and therefore leads to larger mRNA and protein Fano factors.

If feedback regulation is negative, mRNA fluctuations are first transferred to the protein level, from where they repress themselves. This way, $\eta^*(M)$ is reduced, and the effect becomes stronger the larger g^* is. However, concerning $\eta^*(S)$, overregulation through negative feedback is possible: If the sensitivity of autorepression is very strong ($f^* < -\frac{2\nu g^*}{r^* - \nu (g^*)^2}$) and if $\frac{\nu (g^*)^2}{r^*} < 1$ holds simultaneously, an increase in g^* leads to an increase in $\eta^*(S)$.

The effect of a discrete reduction of g^* on the Fano factors under different feedback sensitivities is indicated by the colored graphs in the second row of Fig 5.1.

[1] Stability analysis in Section 4.2.3 has shown that the condition $f^* < (g^*)^{-1}$ holds in any case.

[2] namely $-\frac{1}{g^*} \cdot \left(\sqrt{1 + \frac{(g^*)^2 \nu}{r^*}} - 1 \right)$

[3] Interestingly, the transition of protein noise from super- to sub-Poissonian levels occurs at the same negative feedback strength $f^* = -\frac{\nu g^*}{r^*}$ as the transition of mRNA noise from the sub- to the super-Poissonian region.

Stationary protein-to-mRNA ratio r^* (in case $g^* \neq r^*$):

An increase in r^* reduces $\eta^*(S)$, but augments $\eta^*(M)$ (cf. the third row in Fig 5.1) – a relation with quite surprising simplicity.

Here, it was assumed that the parameters g^* and r^* are completely independent of each other. However, a selective modulation of only one of these parameters is biologically challenging, as they are in fact somehow coupled via the mechanisms underlying the translational propensity function. This motivates us to consider the case when g^* and r^* are fully dependent of each other:

Translational sensitivity and protein-to-mRNA ratio in case $g^* = r^*$:

We first study the situation where $g^* = r^*$, which is necessarily the case when G is linear. An increase in this parameter leads to larger $\eta^*(S)$ if $f^* > -\nu$, i.e. in case of positive feedback or in case the negative feedback from the protein to the mRNA level is too weak to compensate for the increased noise propagation from the mRNA to the protein level. For $f^* < -\nu$, protein noise is reduced through an increase in $g^* = r^*$.

Furthermore, an increase in $g^* = r^*$ causes a rise in $\eta^*(M)$ if $f^* > 0$ or if $f^* < -\nu$ and to a decrease otherwise. This again tells us that under strong negative feedback, protein noise reduction is associated with mRNA noise amplification.

In Fig 5.1 (B), $g^* = r^* = 100$ holds (except for the second and third row, where one of the parameters has been varied). Such a parameter set might belong to a system with a linear translational propensity function. The effect of a change in $g^* = r^*$, which has just been described, is illustrated in the fourth row. In contrast, $1 = g^* \neq r^* = 100$ holds in Fig 5.1 (A), indicating that $\bar{G}$ is in this case nonlinear. However, modulating translation may affect both parameters in a similar manner, which is why in the fourth row of panel (A), the effect of a proportional change in the two parameters is shown.

Time-scale parameter ν:

The impact of ν can be read quite directly from the formula, as the terms in brackets have the form of a convex combination of two values, namely 1 and $R^* = r^* \frac{f^*}{g^*}$ in the case of $\eta^*(M)$, and 1 and $(R^*)^{-1}$ in the case of $\eta^*(S)$. The point between these two edges is determined via ν. When mRNA dynamics are accelerated ($\nu \downarrow$), protein noise is attenuated if $R^* < 1$ and mRNA noise is reduced if $0 < R^* < 1$. Otherwise, the fluctuations are amplified (cf. the fifth row of Fig 5.1). Note that in Fig 5.1 (B), where $g^* = r^* = 100$, $R^* < 1$ holds for all f^*.

The combined parameter R^* contains information on the relative capabilities of reactions (transcription vs. translation) to propagate noise. It is decisive for the effect of the time-scale parameter ν, which determines the capability of reactions to average out upstream fluctuations (noise filtering through slow reactions, cf. Section 2.2.3). Note that in a bimodal system, the value of R^* may severely differ between the two stable expression states, so that the qualitative effect of a change in ν on the noise level of those states might even be opposite.

5.2. Interpretation of burst measures

5.2.1. Extraction of burst parameters from time course data

Before discussing the effect of circuit properties on burst characteristics in the same way as it was done for the Fano factors, we will first regard the biological meaning of α^* and ω^* *per se*. In Section 2.2.2, it has been mentioned that it is indeed possible to measure translational bursts experimentally. However, Paulsson has argued that even if a mathematical analysis predicts large burst sizes, bursts in the form of distinct peaks in the protein time course are not always observable [Paulsson, 2005]. It therefore makes sense at this point to see whether and how burst measures can be inferred from time course data in order to be able to interpret them. We will start with the simple case of a gene expression system without feedback, where all propensities are linear, and then successively increase the mathematical complexity of the system.

Regulatory systems without feedback and with linear propensities

In [Thattai and van Oudenaarden, 2001], a schematic, idealized protein time course is shown, where instantaneous increases in the protein copy number are followed by periods of exponential protein degradation. The amplitude of the peaks is equal to the burst size α^*, while the time interval between two consecutive bursts is given by the inverse of the transcriptional propensity, $\left(\frac{1}{\nu} \cdot F(s^*)\right)^{-1} = \frac{1}{\omega^*}$ (see also Fig 2.1 (C)). Due to our choice of the reference time-scale (it is defined by protein degradation), it is equal to the inverse of the burst frequency.

This intuitive interpretation of the burst measures and their seemingly easy extraction from protein time courses turns out to be difficult when non-idealized data (be them experimental or simulated) are regarded: First of all, the burst amplitude is randomly distributed. Second, the peak height might not correspond to the actual number of produced proteins (which would be the burst size), since some of the molecules might already have been degraded during the burst. Third, bursts might overlap, so that the contributions of single bursts cannot be clearly distinguished.

Concerning the first two problems, a mathematical analysis can be performed in order to relate peak heights to burst sizes. We start with systems in which the translational propensity is linear, $G(m) = g \cdot m$, and consider a burst which does not overlap with other bursts. Hence, if $[\tau_0, \tau_m]$ is the lifespan of the mRNA that causes the burst, $m(\tau) = 1$ for $\tau_0 \leq \tau \leq \tau_m$. Without loss of generality, let $\tau_0 = 0$ and let $s(0) =: s_0,\ s_0 \in \mathbb{N}_0$. For $\tau \in [0, \tau_m]$, the CME of the protein dynamics read

$$\begin{aligned} \dot{p}_s(\tau) &= g \cdot (p_{s-1} - p_s) + (s+1)\, p_{s+1} - s\, p_s, \\ p_s(0) &= \delta_s(\{s_0\}). \end{aligned} \tag{5.1}$$

The ODE of the expected value $\mathbb{E}[\, S(\tau)]$ is given by

$$\begin{aligned} \dot{\mathbb{E}}[\, S(\tau)\,] &= g - \mathbb{E}[\, S(\tau)\,], \\ \mathbb{E}[\, S(0)\,] &= s_0. \end{aligned} \tag{5.2}$$

Its solution evaluated at $\tau = \tau_m$ reads:

$$\mathbb{E}[\, S(\tau_m)\,] \;=\; g \;-\; (g \;-\; s_0)\cdot e^{-\tau_m}. \tag{5.3}$$

The lifespan τ_m of the mRNA molecule is exponentially distributed with expected value ν. At the end of the burst, which corresponds to the moment of mRNA degradation, the protein copy number is thus on average given by:

$$\begin{aligned}\rho_m^*(s_0) \;&:=\; \mathbb{E}\,[\,\mathbb{E}[\, S(\tau_m)\,|\,\tau_m\,]] \;=\; \int_0^\infty \mathbb{E}[\, S(\tau_m)\,]\cdot\frac{1}{\nu}\cdot e^{-\frac{\tau_m}{\nu}}\,\mathrm{d}\tau_m \\ &= \frac{g\cdot\nu \;+\; s_0}{1+\nu} \;=\; \frac{\alpha^* \;+\; s_0}{1+\nu}.\end{aligned} \tag{5.4}$$

Note that $g\cdot\nu = r^*\cdot\nu = \alpha^*$ in systems with linear translation. Averaging out the stochasticity of $S(0)$ yields

$$\rho_m^* \;:=\; \mathbb{E}[\,\mathbb{E}[\,\rho_m^*(S(0))\,|\,S(0)\,]] \;=\; \frac{\alpha^* \;+\; \rho_0^*}{1+\nu} \tag{5.5}$$

with $\rho_0^* := \mathbb{E}[S(0)]$. The expression $\frac{\alpha^*}{1+\nu}$ is the average number of burst proteins that are available at the end of the burst. We will call this term the *effective* burst size β^*:

$$\beta^* \;:=\; \frac{\alpha^*}{1+\nu}. \tag{5.6}$$

$\frac{\rho_0^*}{1+\nu}$ is the part of initial proteins that is still left after the burst.

α^* can now be calculated as:

$$\alpha^* \;=\; \rho_m^*\cdot(1+\nu) \;-\; \rho_0^* \;=\; \gamma^* \;+\; \rho_m^*\,\nu. \tag{5.7}$$

with $\gamma^* := \rho_m^* - \rho_0^*$ being the average *observable* peak height.

Fig 5.2 visualizes the connections between α^*, β^*, and γ^*. If ν is very small, protein degradation becomes negligible and the burst size α^* is almost equal to the effective burst size β^* and to the peak height γ^*. Furthermore, when the protein level has dropped to zero before the next burst is initiated, $\rho_m^* = \beta^* = \gamma^* = \frac{\alpha^*}{1+\nu}$. This situation occurs frequently[4] if $\omega^* \ll \frac{1}{\ln(\alpha^*)}$. If additionally, ν is known, the burst size could be read directly from the protein time course under these circumstances. Otherwise, the burst characteristics are obtained following this procedure:

[4]The given upper boundary of ω^* is derived from an idealized time course in the following way: When – starting from zero protein molecules – a burst with size α^* occurs at time $\tau = 0$, the subsequent protein decay follows on average the function $\alpha^*\cdot\exp(-\tau)$. After the time $\frac{1}{\omega^*}$, just before the next burst occurs, it has dropped to the level $\alpha^*\cdot\exp(-\frac{1}{\omega^*})$, which is smaller than 1 if $\omega^* < \frac{1}{\ln(\alpha^*)}$.

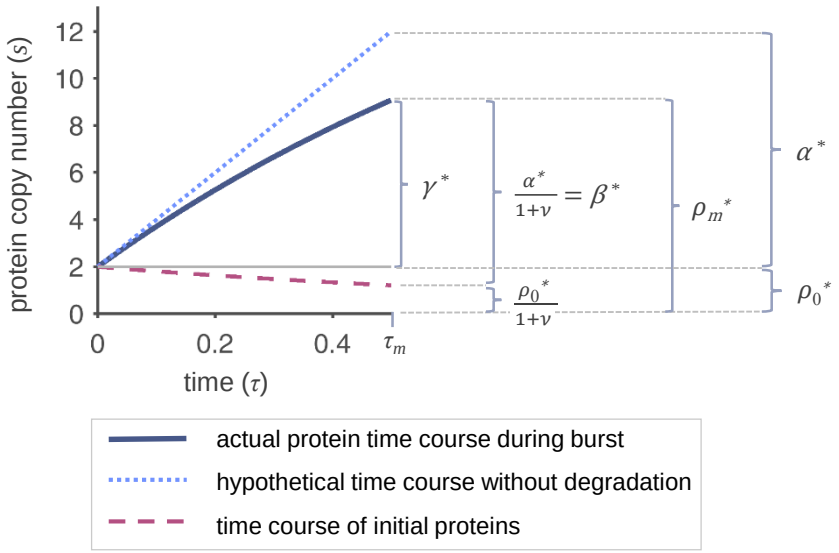


Figure 5.2.: **Connections between the average burst size and the shape of protein peaks.** The blue solid line visualizes the expected protein time course during a burst with parameters $g = G(1) = 20$, $\rho_0^* = 2$, $\tau_m = \nu = 0.5$. It is described by the function $g - (g - \rho_0^*) \cdot e^{-\tau}$, $0 \leq \tau \leq \tau_m$. The parameter ν was chosen quite large in order to visualize the effect of protein degradation. Without degradation, the average protein time course would follow the light blue dotted line, given by the function $\rho_0^* + g \cdot \tau$. The purple line shows the mean degradation of the initially available protein molecules, described by $\rho_0^* \cdot e^{-\tau}$.

Procedure for extracting burst measures from protein time courses

- Determination of the burst frequency ω^*:

  Count the number of mRNA formation events and divide it by the length of the regarded interval. The result is an estimate of ω^*.

- Determination of the burst size α^*:

  Exclude all bursts that overlap with others (i.e. all bursts where more than one mRNA molecule is transiently present) by scanning the mRNA trajectory. Based on all the other, non-overlapping bursts, determine the average of the protein copy number at the beginning and at the end of the burst. These are estimates of ρ_0^* and of ρ_m^*, respectively. Use Eq (5.7) to calculate an estimate of α^*.

- Determination of ν:

  The average mRNA lifespan in the scaled system is equal to $\frac{1}{\nu}$ and can be inferred from the mRNA time course. However, note that for time scaling, the unscaled average lifespan of protein d_s needs to be known as well. It can be inferred from the dynamics of protein degradation between bursts, which on average follows the function $s(t) = s(0) \cdot e^{-t \cdot d_s}$.

- Control:

  Extract the average protein level s^* from the time course and compare it with the product $\alpha^* \cdot \omega^*$.

In principle, the procedure should work whenever G is linear (otherwise, the relation $g \cdot \nu = \alpha^*$ does not hold – the nonlinear case will be examined separately) and if there are enough bursts that do not overlap with others. Table 5.1 lists estimates of burst measures obtained with the

above procedure for a number of qualitatively different systems:

In the first set of systems, $m^* = 0.1$ and $s^* = 10$ holds. The functions F and G are given by:

(A.1)	No feedback:	$F(s) = 0.1,$	$G(m) = 100 \cdot m$
(B)	Positive linear feedback:	$F(s) = 0.02 + 0.008 \cdot s,$	$G(m) = 100 \cdot m$
(C.1)	No feedback, convex G:	$F(s) = 0.1,$	$G(m) = 90.9 \cdot m^2$
(D)	No feedback, concave G:	$F(s) = 0.1,$	$G(m) = 398 \cdot \frac{m}{m+2.9}$
(E)	Negative feedback, convex F:	$F(s) = \frac{2}{s+10},$	$G(m) = 100 \cdot m$
(F)	Positive feedback, concave F:	$F(s) = 0.05 + 3 \cdot \frac{s}{s+590},$	$G(m) = 100 \cdot m$

Two systems with $m^* = 0.01$ and $s^* = 10$ are additionally regarded:

(A.2)	No feedback:	$F(s) = 0.01,$	$G(m) = 1000 \cdot m$
(C.2)	No feedback, convex G:	$F(s) = 0.01,$	$G(m) = 9.9 \cdot \frac{m^2}{0.01}$

In each case, two different values of ν were tested: 1 and 0.01. For every circuit, 10^3 stochastic mRNA and protein time courses were generated, starting from $m = 0$ and $s = 10$ until the final reaction time $\tau_f = 2 \cdot 10^3 \cdot \nu$ was reached. From each trajectory, estimates of α^* and ω^* were extracted. The values given in the table list the empirical means and SEMs of the estimates.

Table 5.1.: Estimates of α^* and ω^* extracted from simulated mRNA and protein time course data.

	$\nu = 1$		$\nu = 0.01$	
type	α^* $[\cdot 10^2]$	ω^* $[\cdot 10^{-1}]$	α^* $[-]$	ω^* $[\cdot 10^1]$
analytical	**1.00**	**1.00**	**1.00**	**1.00**
(A.1) no feedback (fb)	0.95 ± 0.04	1.00 ± 0.07	0.92 ± 0.10	1.00 ± 0.07
(B) pos. linear fb	0.84 ± 0.07	0.99 ± 0.43	0.88 ± 0.10	0.99 ± 0.40
(C.1) no fb, convex G	0.87 ± 0.04	1.00 ± 0.07	0.83 ± 0.10	1.00 ± 0.07
(D) no fb, concave G	0.97 ± 0.04	1.00 ± 0.07	0.94 ± 0.10	1.00 ± 0.07
(E) neg. fb, convex F	0.99 ± 0.04	1.40 ± 0.06	0.93 ± 0.10	1.02 ± 0.07
(F) pos. fb, concave F	0.87 ± 0.05	0.90 ± 0.13	0.90 ± 0.10	1.01 ± 0.18
analytical	**10.00**	**0.10**	**10.0**	**0.10**
(A.2) no feedback	9.93 ± 1.32	0.10 ± 0.02	9.86 ± 2.38	0.10 ± 0.02
(C.2) no fb, convex G	9.85 ± 1.37	0.10 ± 0.02	9.92 ± 2.42	0.11 ± 0.02

(A.1) and (A.2) are systems without feedback and with linear propensity functions. The estimates of ω^* are accurate, while α^* is slightly underestimated, which can be explained as follows: mRNA molecules with a lifespan longer than average tend to produce an above-average amount of protein. At the same time, they are the molecules which are more likely to be involved in overlapping bursts. The elimination of those overlaps thus leads to a certain bias and to the observed underestimation. The probability of overlaps depends of the burst frequency, which is smaller in system (A.2), therefore resulting in a more reliable estimation of α^*.

Regulatory systems with feedback and with linear propensities

Before regarding the estimates for systems with feedback in Table 5.1, let us first have a look at the simulated mRNA and protein trajectories in Fig 5.3. The time course in panel (A) belongs to a system without feedback, the one in panel (B) to a system with positive feedback, where $F(s) = a + f \cdot s,\ f > 0$ is a linear function (and obviously, $f^* = f$). At first glance, the bursts in the protein time course of panel (B) seem to be have larger amplitudes and reduced frequency than those in panel (A), although the calculated local characteristics α^* and ω^* of the two systems are identical. Since all propensities are linear, they also correspond exactly to the average ones: $\mathbb{E}[\alpha(M)] = \mathbb{E}\left[\frac{\nu \cdot g \cdot (M+1)}{M+1}\right] = \alpha^*$ and $\mathbb{E}[\omega(S)] = \mathbb{E}\left[\frac{a+f \cdot S}{\nu}\right] = \frac{a+f \cdot s^*}{\nu} = \omega^*$.

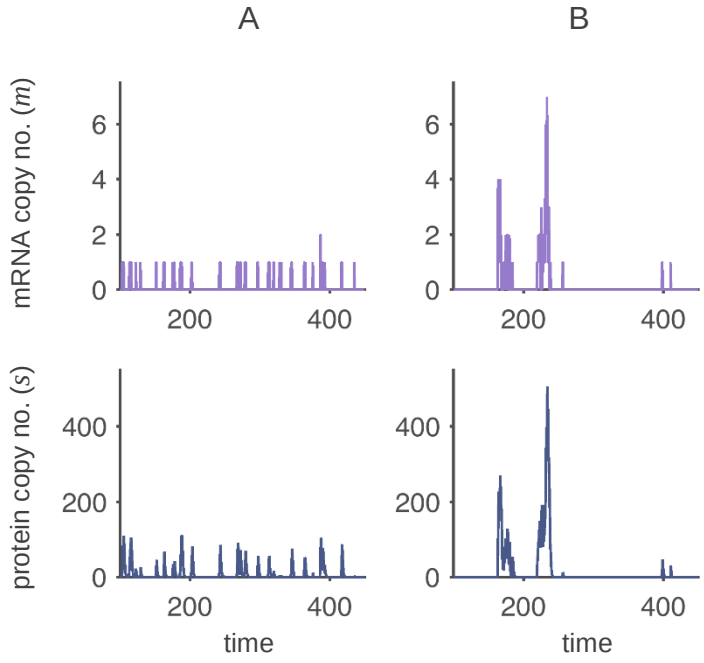


Figure 5.3.: **Bursts of systems with linear propensities without or with feedback.** (A) In the system without feedback, bursts are homogeneously distributed over time, as can be seen from the mRNA time course (top) and the protein time course (bottom). (B) In the system with positive, linear feedback, mRNA transcription events occur in temporal clusters, which promotes super-bursts. The simulated circuits correspond to systems (A.1) and (B) in Table 5.1. In both systems, $m^* = 0.1$, $s^* = 10$, $\alpha^* = 100$, $\omega^* = 0.1$, and $\nu = 1$. Figure taken from [Hortsch and Kremling, 2018b].

The reason for this seeming discrepancy can be found by looking at the state-dependent burst frequency $\omega(\bar{s})$: While it is constant in the system without feedback, it depends on the protein level when positive feedback regulation is present: $\omega(\bar{s}) = \frac{F(s^*)}{\nu} + \frac{F'(s^*)}{\nu}(\bar{s} - s^*) = \omega^* + \frac{f^*}{\nu}(\bar{s} - s^*)$. This causes an inhomogeneous distribution of bursts over time, as they tend to accumulate and overlap when $\bar{s} > s^*$, while their frequency is below average when $\bar{s} < s^*$ (the opposite is true for negative feedback). The mRNA time course in panel (B) shows that the transcription events are indeed temporally clustered. We can thus conclude that the seemingly large protein peaks do not result from an increase in the burst size, but from overlapping bursts.

The proposed procedure for the extraction of estimates ignores these overlaps, so that the number of evaluable bursts is reduced and the estimate of α^* tends to be underestimated even more. The estimate of ω^* is still expected to be correct, but with a larger SEM due to the irregular distribution of transcription events. The values in Table 5.1 referring to system (B) confirm all these expectations.

Regulatory systems with nonlinear propensities

In the derivation of Eq (5.7), the fact that the burst size is given by the product $g \cdot \nu$ for systems with linear translational propensity has been used. In order to handle nonlinear translation, the calculations need to be adapted. We can still only evaluate non-overlapping bursts, during which the translational propensity is equal to $G(1)$. The general formulation of the CME which describes protein formation and degradation during a non-overlapping burst is given by:

$$\dot{p}_s(\tau) \;=\; G(1) \cdot (p_{s-1} - p_s) \;+\; (s+1)\, p_{s+1} \;-\; s\, p_s. \tag{5.8}$$

Repeating the calculations as in the case of linear translation, one finally obtains – instead of Eq (5.5) – the relation:

$$\rho_m^* = \frac{G(1) \cdot \nu + \rho_0^*}{1+\nu} \tag{5.9}$$

and Eq (5.7) becomes

$$\alpha^* \cdot \frac{G(1)}{r^*} \;=\; G(1) \cdot \nu \;=\; \rho_m^* \cdot (1+\nu) \;-\; \rho_0^*. \tag{5.10}$$

Here, the problem occurs that $G(1)$ might be unknown. For $m^* \ll 1$, however, one can start from the premise that the deviation of Eq (5.10) from Eq (5.7) is moderate since

$$\begin{aligned} r^* = \frac{\bar{G}(m^*)}{m^*} &= \frac{\sum_{n=0}^{\infty} G(n) \frac{(m^*)^n}{n!} e^{-m^*}}{m^*} = \sum_{n=0}^{\infty} \frac{G(n+1)}{n+1} \frac{(m^*)^n}{n!} e^{-m^*} \\ &\approx \sum_{n=0}^{\infty} \frac{G(n+1)}{n+1} \frac{(0)^n}{n!} e^{-0} = G(1) \end{aligned} \tag{5.11}$$

holds[5]. For average mRNA levels far below 1, the proposed procedure for extracting burst measures is thus still applicable. This is confirmed by the evaluation of system (C.2) with its concave translational propensity function, cf. Table 5.1. In system (C.1), the average mRNA level is comparably higher and the estimated burst sizes are therefore indeed of reduced quality. The same is basically expected for system (D), where the translational propensity function is convex instead of concave; however, the estimates appear accurate. This might be the result of an overestimation of the burst size, caused by the convex structure of $\bar{G}$, which compensates for the underestimation made through the elimination of overlapping bursts.

Note that the inference of ω^* is not impaired by nonlinearity of G.

Now we consider the case in which G is linear, but F is not. The estimation of α^* then works properly, while the inference of ω^* is hampered by the fact that the *actual* average burst size indeed deviates from that value: $\mathbb{E}[F(S)] \neq F(\mathbb{E}[S]) \neq F(s^*) = \omega^* \cdot \nu$. The tendency of the deviations can be predicted for concave or for convex F using Jensen's inequality (cf. Table 5.1, systems (E) and (F)). As the listed values show, the deviation becomes larger if ν is increased, since this leads to amplified protein fluctuations and therefore to a higher impact of nonlinearity.

All in all, we can conclude that peaks in the protein time course are an indication of the presence of translational bursts. Their average height is actually related to the burst size, but one must keep in mind that a part of the burst proteins might have been degraded immediately.

[5] By the way, this would not be the case if r^* was calculated based on the classical deterministic model.

Degradation can happen so fast (compared to mRNA half-life) that the protein peaks almost vanish. Furthermore, when the burst frequency is regular but high, single peaks might become indistinct, leading to an optical smoothing of the protein time course. Therefore, protein production that mechanistically occurs in bursts may in the end not appear burst-like, which has already been remarked in [Paulsson, 2005]. In contrast to that, we have shown that positive feedback might cause misleadingly large and rare superbursts. It is therefore not straightforward to "read" burst characteristics from any protein time course data directly. However, this section has hopefully further enhanced their interpretability.

5.2.2. Dependence of burst measures on circuit properties

Eqs (4.25) formalize the connections between circuit and burst properties. The average burst size α^* can be modified by changing the average protein-to-mRNA-ratio r^* or the time-scale parameter ν. The average protein level s^* can be maintained by an inverse modulation of the average burst frequency ω^*. Let us now look at the possible modifications in detail:

First, let us generate larger burst sizes α^* by increasing r^*, while keeping ν and s^* constant. This means that the scaled translational propensity function G is increased, while the transcriptional propensity is decreased through a reduction of F, which also reduces the burst frequency ω^*. As a consequence, the average mRNA level m^* is lowered. More burst proteins are thus on average read from a single mRNA molecule through increased translational activity.

Next, α^* is increased through an increase in ν, while r^* and s^* stay fixed. The dynamics of mRNA are thus slowed down compared to those of the proteins. This not only means that burst events occur less frequently, but also that mRNA half-life is prolonged, so that again, more proteins are on average translated from one mRNA molecule. However, this also implies that a larger amount of burst proteins are degraded during the burst. Although the two kinds of modifications – increase in r^* or in ν – may have the same effect on α^*, the effect on the peak height γ^* is thus less pronounced if ν is enhanced.

5.3. Influence of bursts on noise

Until now, it has remained rather unclear what the real effect of burst-like protein production on intrinsic circuit noise might be. In the previous section, it has been shown that fast protein degradation may lower the effect of translational bursts, while positive feedback may enhance it through the generation of superbursts. This section therefore deals with the establishment of a mathematical connection between burst characteristics and the variance and shape of fluctuations.

5.3.1. Links between Fano factors and burst characteristics

First, we return to the very simple relation between bursts and noise strengths, which has been set up in [Thattai and van Oudenaarden, 2001] and which has already been given in the Theoretical Background, Eq (2.34):

$$\eta^*(S) \;=\; 1 \;+\; \frac{\alpha^*}{1 \;+\; \nu} \;=\; 1 \;+\; \beta^*. \tag{5.12}$$

This equation holds only if transcription is (at least locally) unregulated, i.e. if $f^* = 0$, and if the translational propensity is linear. If ν is small enough to be negligible, a direct relation between α^* and the protein Fano factor exists. If f^* is unequal to zero due to transcriptional feedback, the equation

$$\eta^*(S) \;=\; 1 \;+\; \frac{\alpha^* \;+\; f^* \cdot r^*}{(1 \;-\; f^* \cdot r^*) \cdot (1 \;+\; \nu)}. \tag{5.13}$$

holds instead. This equation shows that $\eta^*(S)$ can be quite flexibly adjusted via feedback without changing α^* or ω^*. At first glance, the burst characteristics thus seem to be of minor importance for the magnitude of the Fano factor and *vice versa*. However, in the previous Section 5.2.1, we have seen that feedback regulation influences the temporal homogeneity of burst events, leading to super-bursts in the case of positive autoregulation. The change in the Fano factor might thus be explainable with the help of some "extended" bursting properties. In order to prove this, we take into account the state-dependence of $\alpha(\bar{m}) = \frac{\bar{G}(\bar{m})}{\bar{m}}\nu$ and $\omega(\bar{s}) = \frac{F(\bar{s})}{\nu}$: The local derivatives of these functions are:

$$\begin{aligned} \alpha'^* &:= \frac{d\,\alpha(\bar{m})}{d\,\bar{m}}|_{\bar{m}=m^*} = \frac{g^*\,m^* - \bar{G}(m^*)}{(m^*)^2}\nu = \frac{g^*}{\omega^*} - \frac{\alpha^*}{\omega^*\,\nu} \\ \omega'^* &:= \frac{d\,\omega(\bar{s})}{d\,\bar{s}}|_{\bar{s}=s^*} = \frac{f^*}{\nu}. \end{aligned} \tag{5.14}$$

They can be solved for f^* and g^*, resulting in:

$$f^* \;=\; \omega'^* \cdot \nu \qquad \text{and} \qquad g^* \;=\; \alpha^*/\nu \;+\; \alpha'^* \cdot \omega^*. \tag{5.15}$$

The first relation in Eqs (5.15) shows how the presence of feedback generates a state-dependent (and thus inhomogeneous) burst frequency ($\omega'^* \neq 0$), which has already been observed in the previous chapter and which is now captured quantitatively. We can combine this information with the previous analysis of the impact of f^* on the Fano factor (Section 5.1) to tell that the generation of super-bursts through positive feedback enhances $\eta^*(S)$. On the other hand, negative feedback regulation quickly compensates for protein levels that are above or below average through a reduced or increased number of burst events, thereby reducing the protein Fano factor[6].

The second equation in (5.15) shows that the translational sensitivity g^* is related to the average burst size and to its behavior in a domain around the steady state, specified by α^* and α'^*. As already shown in Section 4.5, the burst size is state-dependent ($\alpha'^* \neq 0$) if G is nonlinear: If, for example, G is a concave function, $\alpha'^* < 0$ holds, i.e. a rise in the mRNA level is accompanied by a decrease in protein burst sizes. Compared with an equivalent system with the same average burst size α^*, but with linear or convex G, the capability to propagate mRNA fluctuations to the protein level – expressed by the translational sensitivity g^* – is therefore reduced. We know from Section 5.1 that, in case of positive feedback, this also reduces the protein and mRNA Fano factors.

Now that f^* and g^* have been expressed in terms of extended burst properties, they can be replaced in Eq (4.19):

$$\eta^*(S) \;=\; 1 \;+\; \frac{(\alpha'^* \cdot \omega^* \cdot \nu \;+\; \alpha^*) \cdot (\alpha'^* \cdot \omega^* \cdot \nu \;+\; \alpha^* \cdot (1 \;+\; \omega'^*))}{\alpha^* \cdot (1 \;+\; \nu) \cdot (1 \;-\; \omega'^* \cdot (\alpha'^* \cdot \omega^* \cdot \nu \;+\; \alpha^*))}. \tag{5.16}$$

[6]This may happen at the expense of an enlarged mRNA Fano factor, as we have seen in Section 5.1.

This formula captures the relation between $\eta^*(S)$ and protein fluctuation patterns and shows how the inhomogeneities in the burst properties affect the Fano factor. Compared to that, Eq (4.19) is better suited for studying the dependence of the Fano factor on the basic circuit properties.

Note that like in Eq (5.12), ν (or r^*, after a reformulation that uses the relation $\alpha^* = r^*\nu$) cannot be eliminated from the above expression, so that the Fano factor cannot be exclusively expressed in terms of extended burst characteristics. On the other hand, it is impossible to deduce the burst characteristics from the sole knowledge of $\eta^*(S)$. This shows that both the burst size and the Fano factor contain some unique information on the characteristics of protein noise. It therefore makes sense to look at both kinds of noise measures and to combine the obtained information to get a better understanding of intrinsic stochasticity.

5.3.2. Links between the skewness of the protein distribution and bursts

The relation between bursts and the location of modes has already been discussed in the last two chapters for the case $\nu \ll 1$. It has turned out that the burst size determines whether the protein level fluctuates quite symmetrically around the mean value (small α^*, large ω^*), or whether the protein trajectory remains below-average for most of the time, showing only rare peak-like excursions to high copy numbers (large α^*, small ω^*), which causes strong right-skewness of the protein distribution.

5.4. The effect of genetic manipulations

In this section, it is demonstrated how the obtained, quite formal results about the connection between circuit properties and noise can be interpreted from a more biological point of view. More specifically, the effect of concrete genetic manipulations on bursts and Fano factors is assessed at the example of artificial protein overexpression. In order to better relate the performed manipulations to circuit properties and parameters, it is helpful to first consider the regulatory system in terms of the unscaled reactions, cf. the reaction system (3.1).

The example shown here focuses on a system without feedback and with linear translation. In this case, the transcription and translation propensities of the original, unmanipulated system are given by $\hat{F} \equiv \hat{a}$ and by $\hat{G}(m) = \hat{g} \cdot m$, respectively. The degradation reactions of mRNA and protein are linear with constants d_m and d_s. Then, $a = \frac{\hat{a}}{d_m}$, $r^* = \frac{\hat{g}}{d_s}$, $\nu = \frac{d_s}{d_m}$, $\alpha^* = \frac{\hat{g}}{d_m}$, and $\omega^* = \frac{\hat{a}}{d_s}$.

Now we assume that the steady state protein level s^* is artificially raised k-fold ($k > 1$) through a manipulation on the genetic level. In terms of the original parameter space, the following four basic modulations of the system are possible:

1. k-fold increase in the transcription constant $\hat{a}$,
2. k-fold reduction of the mRNA degradation constant d_m,
3. k-fold increase in the translation constant $\hat{g}$,
4. k-fold reduction of the protein degradation constant d_s.

The first two kinds of modulations are made on the mRNA level, while the others affect reactions on the protein level. From a biotechnological point of view, the first modification could be achieved by increasing the strength of the promoter, while the third may be accomplished

through a stronger ribosomal binding site or through altered codon usage. The degradation of mRNA or protein molecules might be decelerated through the usage of special tags [Cameron and Collins, 2014; Carrier and Keasling, 1997; Rauhut and Klug, 1999].

Table 5.2 summarizes the effects of the modulations, which are listed in the left column and which are also expressed in terms of the scaled parameters (second column). The third column indicates changes in α^* and ω^*, respectively. In the fourth column, the resulting protein Fano factors are given as functions of the characteristics α^*_{ref} and ν_{ref} of the original reaction system without overexpression, which has the reference Fano factor $1 + \frac{\alpha^*_{\text{ref}}}{1+\nu_{\text{ref}}}$. In columns 1 to 3, only magnitudes that have changed are listed. $\uparrow$ and $\downarrow$ indicate a k-fold and $\frac{1}{k}$-fold change compared with the reference system, respectively.

The third column shows that in all cases, either the burst size or the burst frequency is increased k-fold: Modulations 1 and 4 lead to a change in ω^*, while modulations 2 and 3 affect α^*. Interestingly, the resulting protein Fano factors are all different. In particular, an increase of the transcription rate (modulation 1) does not alter the Fano factor at all.

When modulations on the mRNA level (1 and 2) and on the protein level (3 and 4) are considered separately, one realizes that increases in the burst size lead to larger changes in the Fano factor than increases in the burst frequency. This becomes particularly obvious for very small values of ν. Moreover, changes of either α^* or ω^* on the protein level affect $\eta^*(S)$ more severely than the corresponding change on the mRNA level.

Table 5.2.: Effects of changes in reaction rate constants on burst characteristics and on the protein Fano factor.

Changed parameter	**Changed parameters (scaled)**	**Changes in burst characteristics**	**Fano factor $\eta(S)$**
$\hat{a} \uparrow$	$a \uparrow$	$\omega^* \uparrow$	$1 + \frac{\alpha^*_{\text{ref}}}{1+\nu_{\text{ref}}}$
$d_m \downarrow$	$a \uparrow$, $\nu \uparrow$	$\alpha^* \uparrow$	$1 + \frac{k \cdot \alpha^*_{\text{ref}}}{1+k \cdot \nu_{\text{ref}}}$
$\hat{g} \uparrow$	$r^* \uparrow$	$\alpha^* \uparrow$	$1 + \frac{k \cdot \alpha^*_{\text{ref}}}{1+\nu_{\text{ref}}}$
$d_s \downarrow$	$r^* \uparrow$, $\nu \downarrow$	$\omega^* \uparrow$	$1 + \frac{k \cdot \alpha^*_{\text{ref}}}{k+\nu_{\text{ref}}}$

5.5. Overview of graphical methods used in circuit design

In Section 4.5.4 and in Section 4.6.3, the graphical determination of the burst size and of the location of modes has been described. Using a protein-mRNA-phase plot, further circuit properties can be visualized which have an influence on the Fano factors (cf. Eqs (4.18)-(4.20)). The three graphical methods are summarized in Fig 5.4. They support the construction of circuits with a certain qualitative noise pattern. The insights will be used in the following application.

Let us first have a look at the protein-mRNA phase plot in panel (A). The red graph is the mRNA nullcline, the green line marks the protein nullcline. The outer intersection points (grey circles) correspond to the stable deterministic fixed points. This kind of plot is commonly used in the course of deterministic model analysis (cf. Fig 2.2 (B)), but it also contains some information on the noise in the system: The dashed lines indicate the derivatives f^* and g^*. Furthermore, the stationary protein-to-mRNA ratio r^* can be read at each stable steady state. Except for ν, all parameters that have an influence on the local mRNA and protein Fano factors can thus be visualized, cf. Eqs (4.18)-(4.20). Moreover, the parameters of different stable states

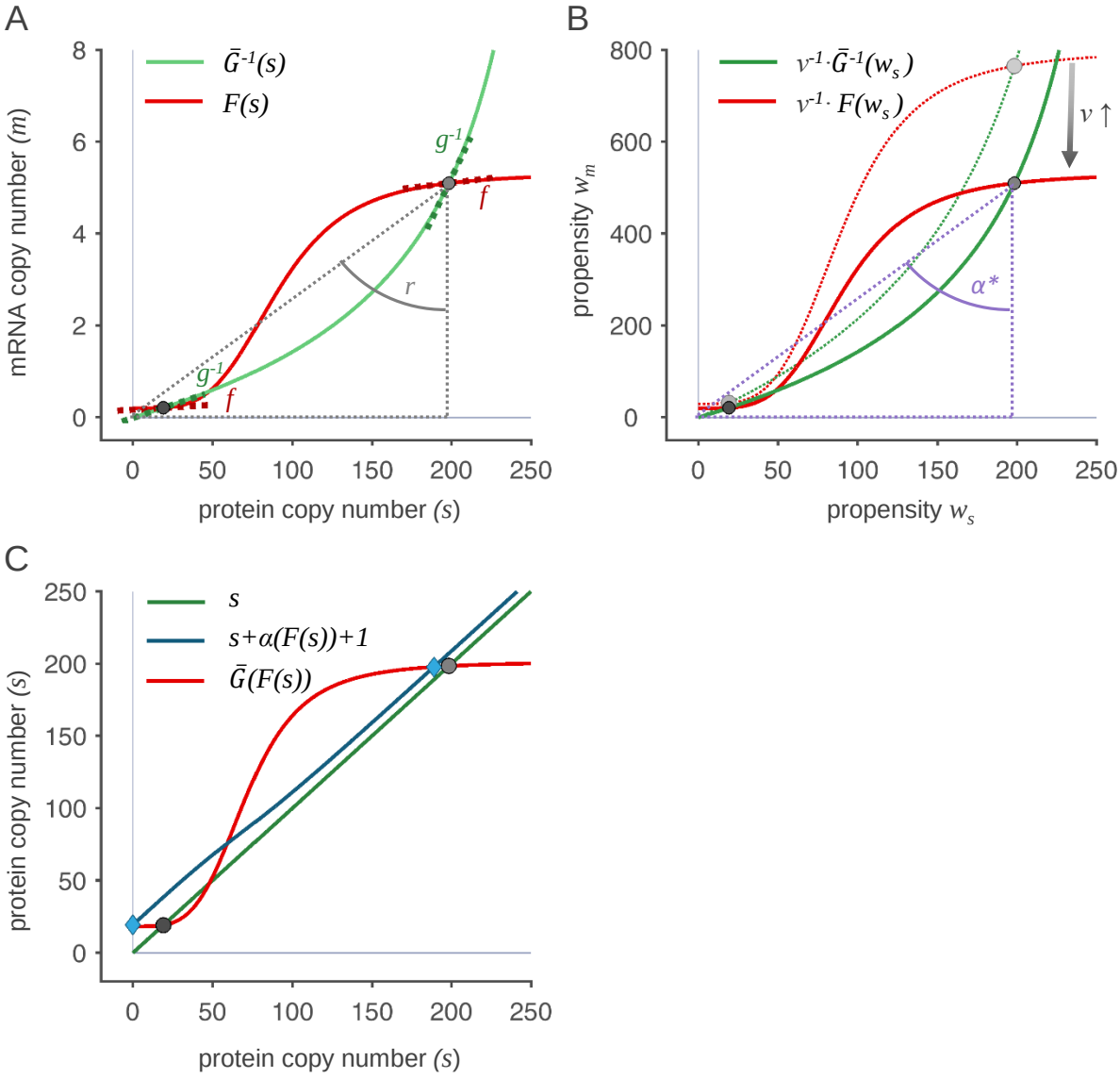


Figure 5.4.: **Graphical methods for state-dependent burst characterization.** (A) Deterministic phase plot supporting the analysis and adjustment of Fano factors. (B) Propensity plot for the determination of burst sizes. (C) Plot for the determination of modes. The functions and parameters used are: $F(s) = 0.2 + 5.12\,\frac{s^4}{s^4+91^4}$, $G(m) = 290\,\frac{m}{m+2}$. In panel (A), $\nu = 0.01$. In panel (B), ν is increased from $\nu = 0.010$ to $\nu = 0.015$. In (C), $\nu = 0.2$ was chosen. The figure is adapted from [Hortsch and Kremling, 2018b].

can be compared directly. One can thus make qualitative statements about how to modulate the shapes of F and G in order to change the noise pattern. Obviously, nonlinearity of G drastically increases the flexibility to adjust the noise pattern, since g^* and r^* can be chosen rather independently of each other and differently in the two expression states. However of course, mechanistic restrictions still need to be taken into account, which might limit the possible range of those parameters. This plot also shows that strong cooperativity in feedback regulation allows locally small values of f^* in both expression states, which makes them robust. The effect of locally flat transcriptional propensities has already been discussed in Section 3.4.2 without using the hLNA, based on an explicit solution of the protein PMF.

The propensity plot in panel (B) illustrates the influence of ν. It is basically very similar to the phase plot, but the scaling is different, see Eqs (4.26). As already described in Section 4.5.4, α^* is equal to the w_s^*-to-w_m^*-ratio at the outer intersection points, which still correspond to the stable stationary states. A variation of ν vertically stretches or compresses the plotted functions. This has a linear effect on w_m, while the protein dynamics are unchanged. This plot immediately shows that nonlinearity of G is also important in order to generate different average bursts sizes in the two expression states.

The third plot can be used if $\nu \ll 1$. It compares the locations of the stable steady states (grey circles) with the locations of the modes (blue diamonds). The larger the local burst size $\alpha(F(s^*))$ is, the larger are the deviations between modes and stable fixed points.

5.6. Application: Generating different noise patterns in a bistable feedback system

The previous results are now used for the *in silico* design of different noise patterns in single-gene autoregulatory circuits, which can be described by the reaction system (4.1). The noise level of a stationary state has an influence on its robustness, i.e. on the duration how long this state is on average maintained. There are, however, further important determinants of robustness that are not regarded (for example the basins of attraction of the stable expression states). For a full quantitative prediction of dynamic population heterogeneity (e.g. in terms of mean first passage times [Ghusinga et al., 2017]), the analysis of noise patterns is therefore insufficient. The assessment of whether the desired population behavior is achieved at least qualitatively will be done using stochastic simulations. As an example, we define our goal to be the construction of a bimodal system whose low-expression (inactive) state has large noise, while noise in the high-expression (active) state is small. With this noise pattern, we try to generate a population in which cells can randomly jump from the inactive to the active state, but not back. Such unidirectional switches are encountered in several cellular decision-making processes, where the generation of two or more phenotypic subpopulations is advantageous in order to reduce overall risks or costs. One example is the development of competence in *S. mutans* [Reck et al., 2015].

5.6.1. System with linear translation rate

First, a system with linear translational propensity $G(m) = g \cdot m$ is regarded. Bistability is generated through a sigmoid transcriptional propensity $F(s) = a + v \frac{s^h}{K_s^h + s^h}$, where a and $a + v$ correspond to basal and maximum transcription, respectively, K_s is the microscopic dissociation constant and h is the Hill coefficient.

Due to the restriction $g = g^* = r^*$, the local Fano factors can only be adapted using three parameters, ν, g, and f^*. In the two expression states, the parameters ν and g are identical and therefore not suitable to generate different noise levels. The burst size $\alpha^* = g \cdot \nu$ is state-independent as well. A decoupling of the two stationary noise levels can only be achieved through a difference in f^*. For the generation of a noise-driven, unidirectional switch to the active state, the transcriptional propensity should be sensitive to protein fluctuations in the inactive state, so that super-bursts are promoted, but it should be rather saturated in the active state.

5.6.2. Systems with nonlinear translation rate

Now, G is allowed to be a nonlinear function. The basic shape of F remains unchanged. As already mentioned, nonlinearity of G does not only decouple the parameters g^* and r^* from each other, it also allows them to vary between different stationary states. The only parameter that is still constant throughout the system is the time-scale parameter ν. In Fig 5.4, the graphical evaluation of noise characteristics has been performed using a strictly concave function G, which can be shown to result in strict concavity of $\bar{G}$ (see Section C.1 in the Appendix). In this case, g^* and r^* obviously decrease with increasing m^*. It can be concluded immediately that the burst size in the low-expression state is larger than in the high-expression state. The decrease in g^* also favors the reduction of the protein Fano factor in the active state, but this tendency is counteracted by the decrease in r^*. Eq (4.19) shows that if the ratio $\frac{g^*}{r^*}$ decreases with m^*, which is for example the case if $G(m) = u\,\frac{m}{K_m+m}$ has the form of a typical Michaelis-Menten function (cf. Section C.1), $\eta^*(S)$ will be smaller in the active than in the inactive state.

With this choice of G, we thus expect that in the inactive state, rare but large bursts occur that may randomly activate the expression circuit. In the active state, the fluctuations are supposed to be moderate and more symmetrical. The following simulation study will show whether this assumption is correct.

5.6.3. Stochastic simulations of a bistable system

In Fig 5.5, simulated protein trajectories of three different bistable systems with cooperative autostimulation are shown. All of them start in the inactive expression state. The distribution of the population, consisting of 500 cells, is indicated by the color plot. Additionally, one exemplary trajectory is highlighted in white in order to show the structure of fluctuations.

In panel (A), the translational propensity function G is linear. The burst size α^* is therefore constant throughout the system. Protein dynamics are rather slow, resulting in a small average burst size. The sensitivity of feedback in the inactive state is small as well. As a consequence, none of the simulated expression systems is activated. In panel (B), ν is increased compared with (A). As a consequence, the burst sizes α^* and the Fano factors $\eta^*(S)$ are enhanced at both fixed points. The resulting fluctuations in the inactive state are now large enough to randomly activate some of the simulated circuits. However, the active state is too noisy to be robustly maintained. The circuits thus switch between the two states in an irregular and uncontrolled manner. The biological functionality of such a system is questionable. In panel (C), a concave, saturated translational propensity function has been chosen. Both the burst size and the protein Fano factor are significantly larger in the inactive than in the active state. They lead to random activation of the circuit in a subpopulation of cells, which then remains in the high-expression state. This results in a clear bimodal distribution of the population.

In all three systems, the transcriptional sensitivity f^* at the stable fixed points is small due to the sigmoid shape of F. We thus expect the results of hLNA and burst characterization to have high accuracy. All in all, the desired noise-driven, unidirectional switching was achieved through a targeted modulation of noise patterns within a bistable circuit, which has been confirmed through stochastic simulations. A more quantitative view on the robustness of the expression states in the considered systems is given in the Appendix, Section C.2, where the mean first passage times extracted from the simulations are listed.

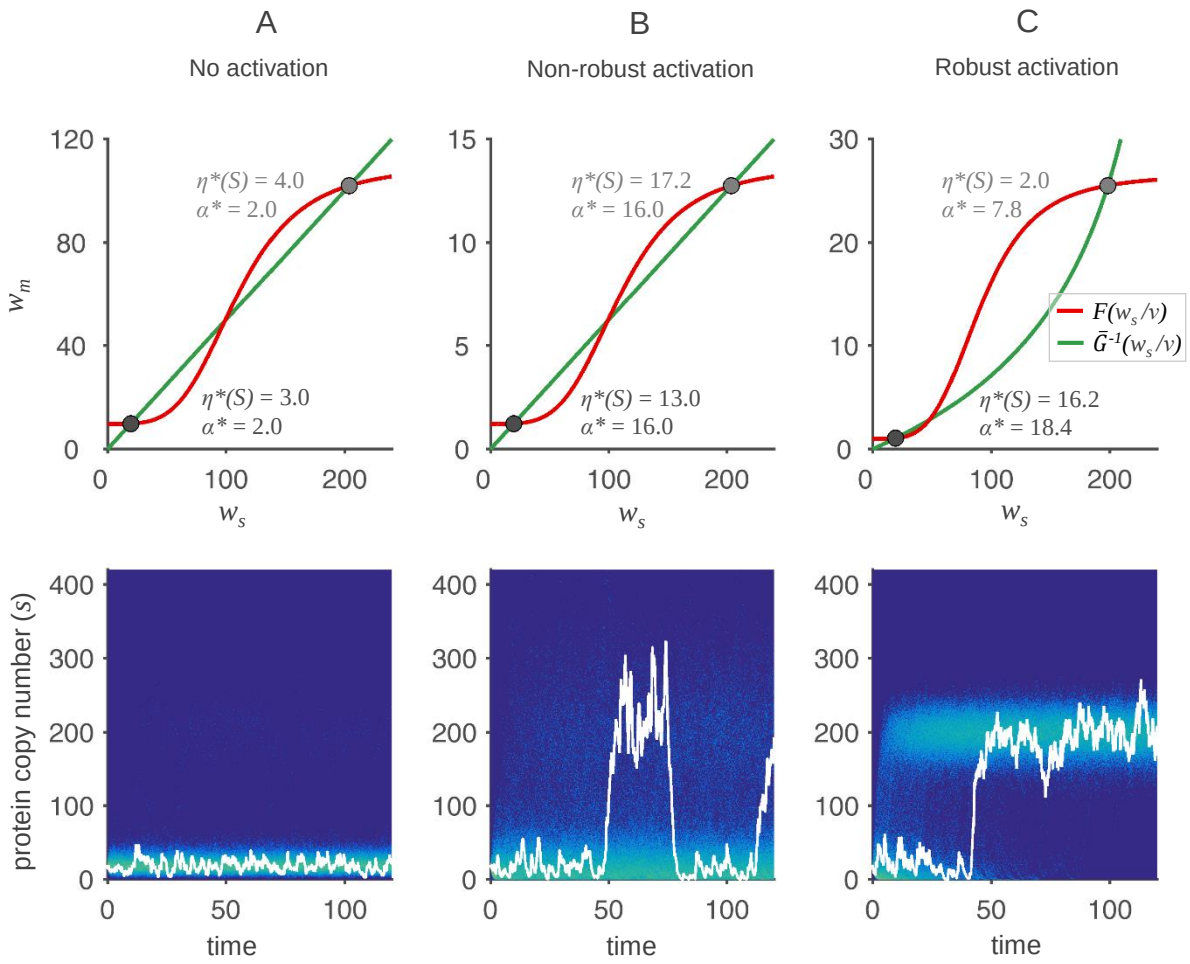


Figure 5.5.: **Stochastic simulations of bistable autostimulatory reaction systems.** The upper row shows propensity plots of three different regulatory systems, whose active expression states have similar protein levels. The bottom row visualizes the protein dynamics of the simulated population, consisting of 500 circuits, by color plots. All simulations were initialized in the inactive expression state. Additionally, one representative time course is highlighted in white. (A) System with linear translational propensity function. The noise in the inactive expression state is too small to randomly activate the system. $F(s) = 0.49 + 5\,\frac{s4}{s^4+110^4}$, $G(m) = 40\,m$, $\nu = 0.05$. (B) Bursts are generated through decelerated mRNA dynamics (increase in ν), leading to random switches between expression states. Parameters as in (A), except for $\nu = 0.4$. (C) System with saturated translational propensity. Bursts occur in the inactive state, but are diminished in the active state, leading to a bimodal, robust activation of the system. The parameters are equal to those in Fig 5.4 (C): $F(s) = 0.2 + 5.12\,\frac{s4}{s^4+91^4}$, $G(m) = 290\,\frac{m}{m+2}$, $\nu = 0.2$. The figure is modified from [Hortsch and Kremling, 2018b].

5.7. Short discussion and summary

In this chapter, we have used our previously developed tools

- to identify mathematical connections between different noise measures (variance and skewness of fluctuations and burst characteristics),
- to further examine and interpret the connection between circuit properties and intrinsic noise,
- and to use this information to find strategies for the modulation of noise patterns through *in silico* circuit design.

The obtained results are in accordance with the insights gained in previous studies; for example, the role of positive or negative feedback in noise amplification or attenuation, and the impact of slow reactions acting as noise filters. Moreover, some facts about protein burst generation could be confirmed, cf. Sections 2.2.2 and 2.2.3. However, the present study goes beyond these aspects. The novel definitions of state-dependent noise strengths and burst characteristics, combined with their graphical visualization, has paved the way for the analysis of complex noise patterns, where different expression states are associated with different noise properties.

For example, it could be shown that in order to better adjust the noise measures for every expression state individually, a nonlinear translation rate is helpful, which makes the interplay of mRNA and protein dynamics more flexible, especially concerning the generation of protein bursts and concerning the capability to propagate fluctuations. In our example, it turned out that the bistable combination of a Michaelis-Menten type translation function with a sigmoid transcriptional propensity supports uni-directional noise-driven switches from the inactive to the active expression state: Large bursts are generated when protein expression is low, while noise is efficiently suppressed in the high-expression state. Biologically speaking, such a saturated translation rate can be the result of indirect global regulation, e.g. limitation of cellular resources like ribosomes, amino acids, or tRNAs [Mather et al., 2013].

This provides a possible explanation of how some random decision-making processes can robustly proceed. In *S. mutans*, for example, the activation of a central regulator, ComS, occurs in a bimodal fashion [Reck et al., 2015]. ComS makes the cell enter a regulatory cascade that finally leads to the genetically competent state. This state is quite risky, as it renders cells susceptible to environmental stress. Therefore, it indeed makes sense that only a subpopulation of *S. mutans* cells becomes activated. The observed bimodality has previously simply been attributed to the presence of a positive feedback loop in the regulation of ComS and to the bistability it causes according to a RE-based model [Son et al., 2012]. However, stochastic simulations of *comS* expression with linear translation propensities like those in Fig 5.5 showed either a lack in random activation (panel A) or uncontrolled switches between the two expression states (panel B). This switching behavior would not be functional, since the high ComS expression state needs to be maintained for a certain period to ensure the proceeding of the competence cascade, before it is deactivated in a controlled manner. The generation of a more delicate noise pattern ensuring uni-directional switches – maybe due to global regulation, as mentioned above – is therefore necessary. In fact, genetic competence is a costly process, as the competence cascade involves the expression of multiple genes. As a consequence, the activation of ComS might indeed be accompanied by a limitation of cellular resources, which would automatically stabilize the active state.

Besides translation, nonlinearity of any other reaction enhances the flexibility of a circuit to arrange its intrinsic noise pattern. Here, we have only considered linear mRNA and protein degradation, but nonlinear relations are indeed realistic, for example if proteolytic machineries are involved [Kasper et al., 2014] or if cooperative stabilizing effects in multimers occur [Buchler et al., 2005]. Based on the previous results, it might now be easier to imagine how altered mRNA degradation kinetics in one of the expression states would change the average burst size and hence the fluctuation pattern.

One should note that although we have associated small local fluctuations with the robustness of an expression state, this relation is purely qualitative, since the local noise level is not the only determinant for system state transitions: For example, it might be valuable to relate the basins of attraction of all fixed points, which are also determined by the reaction propensity functions, to

the noise pattern. Moreover, it is *per se* difficult to define how to quantify robustness: Basically, one could call a stationary state robust, if the mean first passage time (MFPT) away form this state is large. However, the reference time-scale needs to be defined on which the MFPTs of different expression states or circuits should be compared – this might be the original process time, or time normalized with respect to mRNA or protein degradation, or with respect to the average burst frequency. In this study, we have only compared the robustness of different expression states within the same autoregulatory circuit referring to the scaled process time, and we could verify through stochastic simulations that the targeted creation of noise patterns has indeed led to the desired qualitative behavior.

6. Engineering the genetic toggle switch

The novel methods are next applied to the genetic toggle switch, a system comprising two genes whose gene products mutually inhibit each other. Its experimental construction was first published in [Gardner et al., 2000]. In this early seminal work of synthetic biology, the bistability predicted by a deterministic model could be confirmed. This chapter deals with a more stochastic view on the circuit. In particular, it will assess whether and how the involvement of two proteins increases the flexibility of noise adjustment. It will quickly turn out that the doubled amount of reactive species and reaction channels makes the relation between circuit parameters and Fano factors rather intricate. Therefore, three different model formulations of the toggle switch with increasing complexity will be successively analyzed in order to identify more and more contributions to circuit noise in a step-wise manner. Some of the results are published in [Hortsch and Kremling, 2018a].

6.1. Model formulation

6.1.1. Full model

The full model is schematically shown in Fig 6.1 (A). It can be divided into two subsystems 1 and 2, each comprising one of the mRNA and protein species with copy numbers m_i and s_i, $i = 1, 2$, respectively. Transcription is negatively regulated by the protein of the other subsystem. The reference time-scale is set by the degradation kinetics of protein 1. The propensities of the other subsystem are scaled with the parameter δ, which corresponds to the degradation constant of protein 2. Through a change in δ, the time-scale of all reactions in the second subsystem can be uniformly altered and thereby shifted relative to subsystem 1. The notation within the subsystems is consistent with that of the previous chapters. The reactions can be written as:

$$
\begin{aligned}
\emptyset &\xrightarrow{\frac{1}{\nu_1} \cdot F_1(s_2)} \text{mRNA 1} \\
\text{mRNA 1} &\xrightarrow{\frac{1}{\nu_1}} \emptyset \\
\emptyset &\xrightarrow{G_1(m_1)} \text{Protein 1} \\
\text{Protein 1} &\xrightarrow{1} \emptyset \\
\\
\emptyset &\xrightarrow{\frac{1}{\nu_2} \cdot F_2(s_1) \cdot \delta} \text{mRNA 2} \\
\text{mRNA 2} &\xrightarrow{\frac{1}{\nu_2} \cdot \delta} \emptyset \\
\emptyset &\xrightarrow{G_2(m_2) \cdot \delta} \text{Protein 2} \\
\text{Protein 2} &\xrightarrow{\delta} \emptyset
\end{aligned}
\tag{6.1}
$$

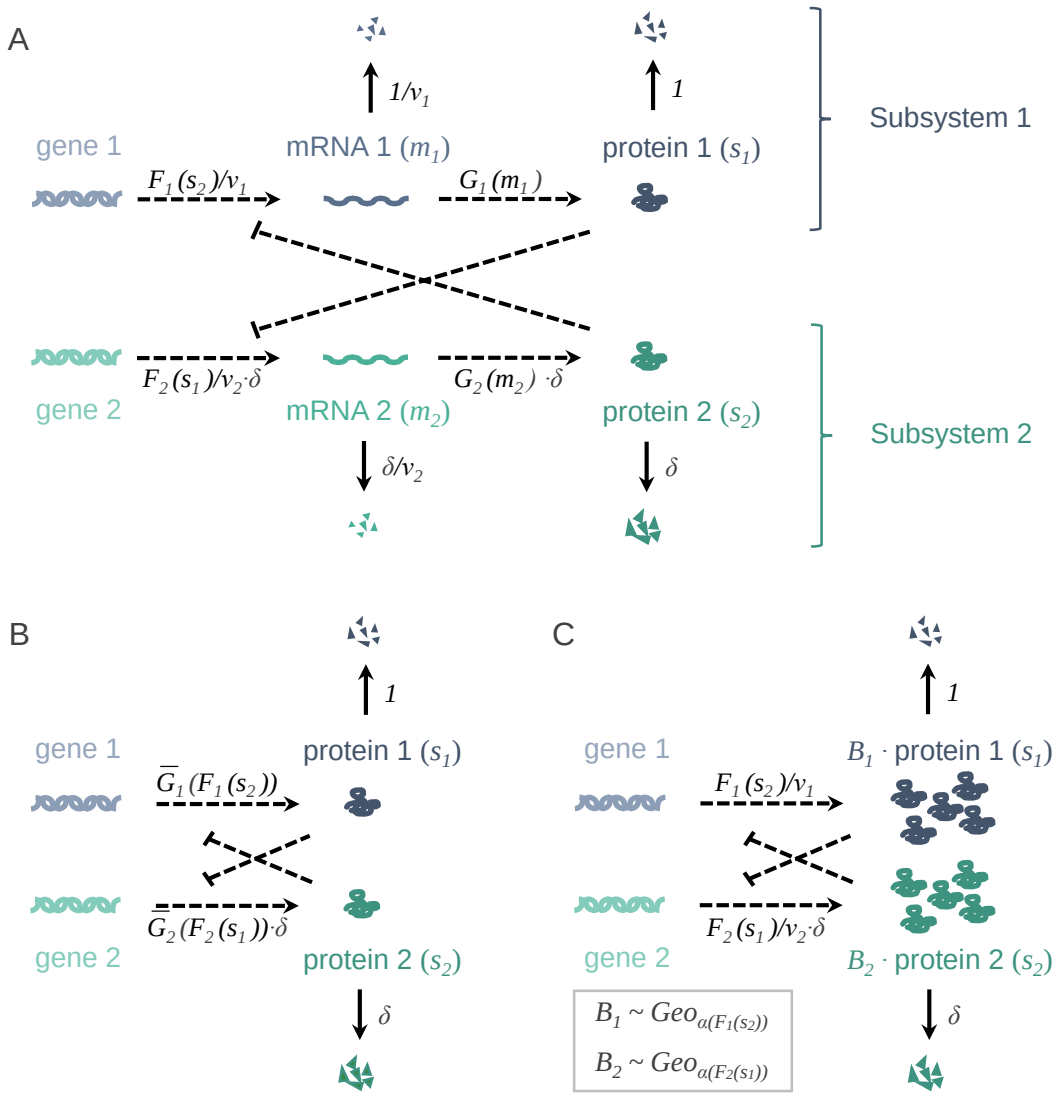


Figure 6.1.: **Schematic representations of the genetic toggle switch.** (A) Full model. (B) Reduced model assuming infinitely fast mRNA dynamics. (C) Reduced protein bursting model.

and the corresponding CME reads:

$$
\begin{aligned}
&\dot{p}_{m_1,s_1,m_2,s_2} = \\
&\left[F_1(s_2)\,p_{m_1-1,s_1,m_2,s_2} - F_1(s_2)\,p_{m_1,s_1,m_2,s_2} + (m_1+1)\,p_{m_1+1,s_1,m_2,s_2} - m_1\,p_{m_1,s_1,m_2,s_2}\right]\cdot\frac{1}{\nu_1} \\
&+ G_1(m_1)\,p_{m_1,s_1-1,m_2,s_2} - G_1(m_1)\,p_{m_1,s_1,m_2,s_2} + (s_1+1)\,p_{m_1,s_1+1,m_2,s_2} - s_1\,p_{m_1,s_1,m_2,s_2} \\
&+\Big(\left[F_2(s_1)\,p_{m_1,s_1,m_2-1,s_2} - F_2(s_1)\,p_{m_1,s_1,m_2,s_2} + (m_2+1)\,p_{m_1,s_1,m_2+1,s_2} - m_2\,p_{m_1,s_1,m_2,s_2}\right]\cdot\frac{\delta}{\nu_2} \\
&+ G_2(m_2)\,p_{m_1,s_1,m_2,s_2-1} - G_2(m_2)\,p_{m_1,s_1,m_2,s_2} + (s_2+1)\,p_{m_1,s_1,m_2,s_2+1} - s_2\,p_{m_1,s_1,m_2,s_2}\Big)\cdot\delta,
\end{aligned}
\tag{6.2}
$$

where p_{m_1,s_1,m_2,s_2} is the probability that the system is in state $(M_1, S_1, M_2, S_2) = (m_1, s_1, m_2, s_2)$.

The hybrid deterministic formulation, in which local averaging of the translational propensity

function is performed for both subsystems (again under the assumption that the two mRNA species are Poisson distributed), is given by:

$$\begin{aligned}
\dot{c}_{m_1} &= (\mathcal{F}_1(c_{s_2}) - c_{m_1}) \cdot \frac{1}{\nu_1} \\
\dot{c}_{s_1} &= \bar{\mathcal{G}}_1(c_{m_1}) - c_{s_1} \\
\dot{c}_{m_2} &= (\mathcal{F}_2(c_{s_1}) - c_{m_2}) \cdot \frac{1}{\nu_2} \cdot \delta \\
\dot{c}_{s_2} &= (\bar{\mathcal{G}}_2(c_{m_2}) - c_{s_2}) \cdot \delta,
\end{aligned} \tag{6.3}$$

with $\bar{\mathcal{G}}_i(c_{m_i}) := \frac{1}{V} \sum_{n=0}^{\infty} G_i(n) \frac{(c_{m_i} V)^n}{n!} e^{-c_{m_i} V}$, $i = 1, 2$. Its derivation is analogous to the one performed for the autoregulatory circuit in Section 4.2.2.

6.1.2. Model reduction 1: Fast mRNA dynamics

The first model simplification that is performed in order to reduce the complexity of the description is an elimination of the mRNA level: It is assumed that the dynamics of both mRNA species are infinitely fast, which corresponds to the limit $\nu_1 \to 0$, $\nu_2 \to 0$, while the ratio $\frac{\nu_1}{\nu_2}$ is kept constant.

On a time-scale determined by the variable τ/ν_1, on which the mRNA species are formed and degraded, the protein copy number hardly changes, so that the mRNA dynamics can be approximately depicted by the following conditional CMEs (cf. the derivation in the Appendix, Section D.1):

$$\frac{d\, p_{m_1|s_2}}{d\,(\tau/\nu_1)} = F_1(s_2)\, p_{m_1-1|s_2} - F_1(s_2)\, p_{m_1|s_2} + (m_1+1)\, p_{m_1+1|s_2} - m_1\, p_{m_1|s_2} \tag{6.4}$$

$$\frac{d\, p_{m_2|s_1}}{d\,(\tau/\nu_1)} = \big[F_2(s_1)\, p_{m_2-1|s_1} - F_2(s_1)\, p_{m_2|s_1} + (m_2+1)\, p_{m_2+1|s_1} - m_2\, p_{m_2|s_1}\big] \cdot \frac{\nu_1}{\nu_2}, \tag{6.5}$$

while the protein species are assumed to stagnate in their current state. The dynamics of the two mRNA species are nearly independent of each other and the molecules are transcribed with "constant" propensities $F_1(s_2)$ and $F_2(s_1)$, respectively. In the stationary state, they are thus Poisson distributed with mean values $\mathbb{E}[M_1|S_2 = s_2] = F_1(s_2)$ and $\mathbb{E}[M_2|S_1 = s_1] = F_2(s_1)$.

Back on the time-scale of protein dynamics, the marginal CME of the protein species can be calculated, using the above pseudo-stationary distributions of the mRNA species:

$$\begin{aligned}
\dot{p}_{s_1,s_2} = {} & \bar{G}_1(F_1(s_2))\, p_{s_1-1,s_2} - \bar{G}_1(F_1(s_2))\, p_{s_1,s_2} + (s_1+1)\, p_{s_1+1,s_2} - s_1\, p_{s_1,s_2} \\
& + \big[\, \bar{G}_2(F_2(s_1))\, p_{s_1,s_2-1} - \bar{G}_2(F_2(s_1))\, p_{s_1,s_2} + (s_2+1)\, p_{s_1,s_2+1} - s_2\, p_{s_1,s_2} \,\big] \cdot \delta,
\end{aligned} \tag{6.6}$$

cf. the calculation steps in Section D.1.

The reaction scheme is thus given by:

$$\boxed{\begin{array}{rcl} \emptyset & \xrightarrow{\bar{G}_1(F_1(s_2))} & \text{Protein 1} \\ \text{Protein 1} & \xrightarrow{1} & \emptyset \\ & & \\ \emptyset & \xrightarrow{\bar{G}_2(F_2(s_1))\cdot\delta} & \text{Protein 2} \\ \text{Protein 2} & \xrightarrow{\delta} & \emptyset \end{array}} \tag{6.7}$$

It is visualized in Fig 6.1 (B). The hybrid ODE reads:

$$\begin{aligned} \dot{c}_{s_1} &= \bar{\mathcal{G}}_1(\mathcal{F}_1(c_{s_2})) - c_{s_1} \\ \dot{c}_{s_2} &= (\bar{\mathcal{G}}_2(\mathcal{F}_2(c_{s_1})) - c_{s_2})\cdot\delta. \end{aligned} \tag{6.8}$$

Note that this kind of model reduction does not consider translational bursts. It will be used to examine the interplay between the two subsystems on a basic level. The influence of bursts is then considered in the next model version:

6.1.3. Model reduction 2: Translational bursting model

Like in the previous chapter, the bursting model provides a good system description if $\nu_i \ll 1$, $i = 1, 2$, since then, the translation events belonging to one burst can be condensed into a single reaction. In this case, the burst sizes B_1 and B_2 are approximately geometrically distributed: $B_1 \sim \text{Geo}_{\,\alpha(F_1(s_2))}$, $B_2 \sim \text{Geo}_{\,\alpha(F_2(s_1))}$.

The list of reactions reads:

$$\boxed{\begin{array}{rcl} \emptyset & \xrightarrow{F_1(s_2)\,\frac{1}{\nu_1}} & B_1 \cdot \text{Protein 1} \\ \text{Protein 1} & \xrightarrow{1} & \emptyset \\ & & \\ \emptyset & \xrightarrow{F_2(s_1)\cdot\frac{1}{\nu_2}\cdot\delta} & B_2 \cdot \text{Protein 2} \\ \text{Protein 2} & \xrightarrow{\delta} & \emptyset \end{array}} \tag{6.9}$$

The CME reads, in analogy to Eq (4.32):

$$\begin{aligned} \dot{p}_{s_1,s_2} \;=\; & \sum_{b=0}^{s_2} \frac{1}{\nu_1}\cdot F_1(s_2-b)\cdot \text{Geo}_{\alpha(F_1(s_2-b))}(b)\cdot p_{s_1,s_2-b} \;-\; \frac{1}{\nu_1}\cdot F_1(s_2)\, p_{s_1,s_2} \\ & +\; (s_1+1)\, p_{s_1+1,s_2} \;-\; s_1\, p_{s_1,s_2} \\ & +\Big(\sum_{b=0}^{s_1} \frac{1}{\nu_2}\cdot F_2(s_1-b)\cdot \text{Geo}_{\alpha(F_2(s_1-b))}(b)\cdot p_{s_1-b,s_2} \;-\; \frac{1}{\nu_2}\cdot F_2(s_1)\, p_{s_1,s_2} \\ & +\; (s_2+1)\, p_{s_1,s_2+1} \;-\; s_2\, p_{s_1,s_2}\Big)\cdot\delta. \end{aligned} \tag{6.10}$$

The reaction scheme is depicted in Fig 6.1 (C). Like in Section 4.6.4, the hybrid deterministic

formulation of the reduced bursting model is the same as the formulation in the fast mRNA limit, i.e. it is identical to Eq (6.8).

6.2. Deterministic model behavior

In order to be able to apply the hLNA to the full model description, the behaviour of the corresponding hybrid deterministic formulation (6.3) should be analyzed first. The deterministic fixed points in terms of copy numbers are defined by the conditions:

$$\begin{aligned} m_i^* &= F_i(s_j^*), \\ s_i^* &= \bar{G}_i(F_i(s_j^*)) = \bar{G}_i(F_i(\bar{G}_j(F_j(s_i^*)))) \qquad \text{for } i,j = 1,2,\ i \neq j. \end{aligned} \tag{6.11}$$

The steady-state conditions for the proteins also apply to the reduced model versions. The protein-protein phase plot in Fig 6.2 shows that the system can have three fixed points, if the functions are chosen suitably.

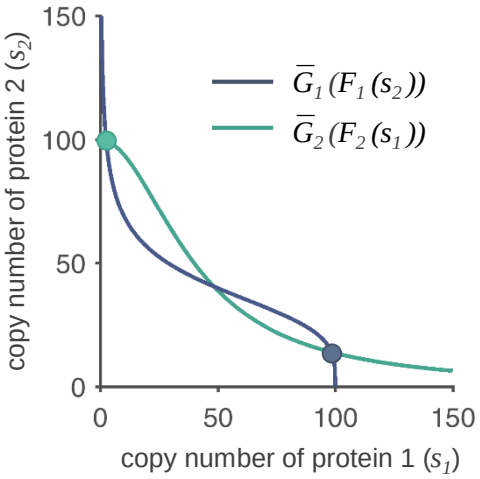


Figure 6.2.: **Deterministic protein-protein phase plot of the toggle switch.** The two graphs indicate the nullclines of protein 1 (dark blue) and protein 2 (green) according to the deterministic models of the full and the reduced systems. The two outer intersection points mark the stable fixed points.

The Jacobian of the RHS of the full deterministic description (6.3) is given by:

$$\mathbf{J}^* = \begin{pmatrix} -\frac{1}{\nu_1} & 0 & 0 & f_1^* \frac{1}{\nu_1} \\ g_1^* & -1 & 0 & 0 \\ 0 & f_2^* \frac{\delta}{\nu_2} & -\frac{\delta}{\nu_2} & 0 \\ 0 & 0 & g_2^* \delta & -\delta \end{pmatrix}. \tag{6.12}$$

Here, $f_1^* := \frac{dF_1}{ds_2}(s_2^*)$, $f_2^* := \frac{dF_2}{ds_1}(s_1^*)$, $g_1^* := \frac{d\bar{G}_1}{dm_1}(m_1^*)$, and $g_2^* := \frac{d\bar{G}_2}{dm_2}(m_2^*)$, which are expression-state specific parameters. f_1^* and f_2^* are non-positive since both proteins act as repressors. The characteristic polynomial $\mathcal{P}$ of $\mathbf{J}^*$ with variable λ reads:

$$0 = \left(-\frac{1}{\nu_1} - \lambda\right) \cdot (-1-\lambda) \cdot \left(-\frac{\delta}{\nu_2} - \lambda\right) \cdot (-\delta - \lambda) - \frac{f_1^*\, f_2^*\, g_1^*\, g_2^*\, \delta^2}{\nu_1 \nu_2} =: \mathcal{P}(\lambda). \tag{6.13}$$

The solutions of this equation are the eigenvalues of $\mathbf{J}^*$ and therefore crucial for the stability of the fixed point. The first part of the expression is a quartic polynomial, whose roots correspond to the diagonal entries of $\mathbf{J}^*$. It is shifted downwards by $\frac{f_1^*\, f_2^*\, g_1^*\, g_2^*\, \delta^2}{\nu_1 \nu_2}$ to yield the characteristic polynomial. If the shift is small enough, all real-valued roots λ are negative. This holds until $\lambda = 0$ solves Eq (6.13), i.e. as long as $\frac{f_1^*\, f_2^*\, g_1^*\, g_2^*\, \delta^2}{\nu_1 \nu_2} < \frac{\delta^2}{\nu_1\, \nu_2} \;\Leftrightarrow\; f_1^*\, f_2^*\, g_1^*\, g_2^* < 1$. This is a necessary condition for the stability of a fixed point.

In Fig 6.2, the two outer intersection points are stable, the middle one is unstable. The stable fixed points therefore either exhibit high expression of protein 1, which represses protein 2, or the opposite scenario where high levels of protein 2 restrict the formation of protein 1.

6.3. Calculations of Fano factors and burst characterization

We have seen in the study of the single-gene autoregulatory system that great adjustability of noise patterns can be reached through nonlinearity of translation. In order to examine the increase in flexibility through the usage of two instead of one gene, the translational propensity functions are henceforward assumed to be linear, i.e. $G_i(m_i) = \bar{G}_i(m_i) = g_i \cdot m_i,\ i = 1, 2$. The hLNA approach is thus equivalent to the classical LNA, which is now expected to be reliable.

6.3.1. The protein Fano factors of the full model

In order to estimate the Fano factors, the Lyapunov equation (2.29) needs to be solved again, this time for a system with four reactive species and with eight reactions. The calculations are more tedious than those in the previous study, where only one gene was involved. The required matrices are deduced in the Appendix, Section D.2. The local mean values are again approximated by the stable fixed points, cf. Eq (6.11). By setting $\Psi^* := g_1 g_2 f_1^* f_2^*$, the resulting Fano factor of protein 1 according to the full model can be written as[1]:

$$\begin{aligned}
\eta^*(S_1) \;=\; 1 + \Big(& g_1\, \nu_1 \Big[(1+\delta)(1+\delta\,\nu_1)(1+\delta^{\nu_1}\!/_{\nu_2})(\delta+\nu_2)(\delta+\delta\,\nu_2) \\
& - \Psi^* \big((1+\delta^{\nu_1}\!/_{\nu_2})(\delta+\nu_2)(\delta+\delta\nu_2) + \delta^2 \nu_2^2 (1+\nu_1+\delta\nu_1+\delta^{\nu_1}\!/_{\nu_2}) \big) \Big] \\
& + \Psi^*\, \delta^2 \Big[(1+\delta\,\nu_1)(1+\delta^{\nu_1}\!/_{\nu_2})(\delta+\delta\,\nu_2) \\
& + \Psi^*\, \nu_2\, (1+\delta\,\nu_1+\delta^{\nu_1}\!/_{\nu_2})(1+\nu_1+\delta\,\nu_1+\delta^{\nu_1}\!/_{\nu_2}) \Big] \\
& + (g_1 f_1^*)^2\, \delta\, (1+g_2\nu_2) \Big[(1+\delta\,\nu_1)(1+\delta^{\nu_1}\!/_{\nu_2})(\delta+\delta\,\nu_2) \\
& + \nu_2\, (1+\delta\,\nu_1+\delta^{\nu_1}\!/_{\nu_2})(1+\nu_1+\delta\,\nu_1+\delta^{\nu_1}\!/_{\nu_2}) \Big] s_2^*/s_1^* \\
& + (g_1 f_1^*)^2\, \delta\, (1-\Psi^*)\, \nu_2^2 (1+\nu_1+\delta\,\nu_1+\delta^{\nu_1}\!/_{\nu_2}) s_2^*/s_1^* \Big)\, \frac{1}{(1-\Psi^*)\, Q},
\end{aligned}$$

$$\begin{aligned}
Q \;=\; & (1+\nu_1)(1+\delta)(1+\delta\,\nu_1)(1+\delta^{\nu_1}\!/_{\nu_2})(\delta+\nu_2)(\delta+\delta\,\nu_2) \\
& - \Psi^*\, \delta^2\, \nu_2\, (1+\nu_1+\delta\,\nu_1+\delta^{\nu_1}\!/_{\nu_2})^2.
\end{aligned} \tag{6.14}$$

[1] Deterministic stability analysis has shown that $\Psi^* < 1$ holds for every stable fixed point.

The Fano factors of the remaining reactive species are formulated analogously. The structure of Eq (6.14) is rather intricate, especially due to the complexity of the denominator, so that it would be difficult to directly deduce general rules about the connection between circuit properties and noise. We will therefore study the Fano factors of the reduced models and of certain special cases of the full model, which will altogether provide some fundamental insights into the intrinsic noise of a toggle switch.

6.3.2. Noise under fast mRNA dynamics

The Fano factor of protein 1 according to the reduced system (6.7) is given by:

$$\begin{aligned}\eta^*(S_1) &= 1 + \frac{g_1 f_1^*}{1-\Psi^*}\left(g_2 f_2^* \frac{\delta}{1+\delta} + g_1 f_1^* \frac{s_2^*}{s_1^*}\frac{1}{1+\delta}\right) \\ &= 1 + \frac{\Psi^*}{1-\Psi^*}\left(\frac{\delta}{1+\delta} + R^* \frac{1}{1+\delta}\right), \end{aligned} \tag{6.15}$$

where $R^* := \frac{g_1 f_1^* s_2^*}{g_2 f_2^* s_1^*}$. The second equation holds if $f_2^* \neq 0$. Accordingly, the Fano factor of protein 2 reads:

$$\begin{aligned}\eta^*(S_2) &= 1 + \frac{g_2 f_2^*}{1-\Psi^*}\left(g_1 f_1^* \frac{1}{1+\delta} + g_2 f_2^* \frac{s_1^*}{s_2^*}\frac{\delta}{1+\delta}\right) \\ &= 1 + \frac{\Psi^*}{1-\Psi^*}\left(\frac{\delta}{1+\delta} + R^* \frac{1}{1+\delta}\right)\cdot\frac{1}{R^*}. \end{aligned} \tag{6.16}$$

These formulae are either obtained by taking the limit $\nu_1 \to 0$, $\nu_2 \to 0$, $\frac{\nu_1}{\nu_2} = \text{const.}$ in Eq (6.14), or by performing the LNA based on the CME given in Eq (6.6), which is easily calculated due to the limited size and complexity of the formulation. Based on these simplified Fano factors, the following insights can be gained:

Local effects of circuit parameters on Fano factors

Let us first interpret the single terms in Eq (6.15) (the structure of $\eta^*(S_2)$ can be explained analogously): In case $f_1^* = 0$, i.e. if protein 2 does not influence system 1 (at least locally around the fixed point), constant production and linear degradation of protein 1 results in a Poisson distribution of its copy numbers and therefore, $\eta^*(S_1) = 1$. If $f_1^* < 0$, but $f_2^* = 0$, the feedback loop would not be closed, but subsystem 2 would act upstream of subsystem 1, increasing the protein Fano factor by the term $\frac{(g_1 f_1^*)^2}{1-\Psi^*}\frac{s_2^*}{s_1^*}\frac{1}{1+\delta}$. Interestingly, this term would vanish if the dynamics of the second subsystem were extremely fast ($\delta \to \infty$). Now let us assume that $f_2^* < 0$. The term $\frac{\Psi^*}{1-\Psi^*}\frac{\delta}{1+\delta}$ describes the noise that is propagated through the whole loop. This expression now becomes largest if δ is large.

One might have realized that the structure of Eqs (6.15) and (6.16) is similar to that of Eqs (4.18) and (4.19). This is because the number of species and reaction channels and their topological connection is virtually identical. As a consequence, the effects of circuit parameters on noise will show certain similarities to those identified in Section 5.1. Note that in the toggle switch, f_1^* and f_2^* are actually non-positive, but they always occur in multiplicative pairs, so that the negative signs are cancelled out. This facilitates the analysis of the formulae. As before, both stable fixed points (s_1^*, s_2^*) will be held fixed when parametric variations are studied.

The formulae of the Fano factors immediately show that an increase in g_1 or in g_2 would enhance $\eta^*(S_1)$ and $\eta^*(S_2)$. However, note that if the change in g_i is balanced by an inverse scaling of F_i in order to maintain the average protein level, the products $g_i f_i^*,\ i = 1, 2$ and thus also the Fano factors would remain constant according to this model.

The expression-state-specific parameters f_1^* and f_2^* can be adjusted by selecting repressor systems with suitable inhibition strength, cooperativity, etc. In general, locally small transcriptional sensitivities $|f_1^*|$ and $|f_2^*|$ limit the propagation of noise and therefore the Fano factors.

The effect of δ depends on the value of the combined parameter R^* and can be easily determined when the expression in brackets in Eqs (6.15) and (6.16) is interpreted as a convex combination: If $R^* < 1$, a reduction (increase) in δ attenuates (amplifies) noise of *both* proteins, while the opposite holds if $R^* > 1$. A shift in δ can thus modulate the robustness of an expression state. The decisive parameter combination R^*compares the local capabilities of system 1 and system 2 to propagate noise, so that the influence of δ can be interpreted as follows: Noise is reduced if the reactions with high sensitivity towards upstream noise are decelerated. Note that R^* is dependent on the expression state, so that a change in δ may even have opposite effects on the local noise levels within the bistable system.

Since we are only regarding positive feedback, a local change of any of the circuit parameters affects the local noise levels of both proteins in qualitatively the same way[2]. Therefore, it is not possible to locally reduce the fluctuations of one protein while augmenting fluctuations of the other. This facilitates statements about changes in the robustness of an expression state, which is shaped by the Fano factors of both proteins.

Generation of asymmetric noise levels

How can the interplay of the two subsystems now support the creation of asymmetric noise patterns between the two stable expression states, according to the simplified model?

We have previously seen that, in the case of bistable single-gene autoregulation with linear translation, a difference between the Fano factors of the two stable expression states can only be achieved through a difference in f^*. In case of the toggle-switch, the shape of the transcriptional propensities can of course create such an imbalance, too: If both transcriptional sensitivities $|f_1^*|$ and $|f_2^*|$ are larger in one state than in the other, this would lead to a clear difference in the noise levels. However, the shapes of F_1 and F_2 are not completely flexible, since they are mechanistically constrained and, at the same time, need to ensure bistability. Often, $|f_1^*|$ is larger in one state than in the other, while $|f_2^*|$ behaves opposite, so that direct statements about the relative robustness of the expression states cannot be made. A further way to generate asymmetric noise is, though, available, even if the two subsystems are symmetric in the sense that $g_1 \cdot F_1 = g_2 \cdot F_2$. In order to show this, let us denote the two stable fixed points by state a and state b, and the according transcriptional sensitivities f_i^* by f_1^a, f_1^b, f_2^a, and f_2^b, respectively. Deterministic model analysis shows that due to the symmetry, $s_1^a = s_2^b$ and $s_1^b = s_2^a$. As a consequence, $g_1\, f_1^a = g_2\, f_2^b$, $g_1\, f_1^b = g_2\, f_2^a$, and $\Psi^a = \Psi^b =: \Psi$ hold. The only real asymmetry in this example may be caused by δ: The decisive parameter combinations R^a and R^b fulfill the relation:

$$R^a \;=\; \frac{g_1\, f_1^a \cdot s_2^a}{g_2\, f_2^a \cdot s_1^a} = \frac{g_2\, f_2^b \cdot s_1^b}{g_1\, f_1^b \cdot s_2^b} \;=\; \frac{1}{R^b}. \tag{6.17}$$

[2] As already mentioned, this does not necessarily hold across different expression states.

One can conclude that a shift in δ would inevitably change the noise levels of state a and b in opposite ways unless $R^a = R^b = 1$.

R^a can now be calculated for a common structure of F_i, namely

$$F_i(s_j) = \frac{1}{g_i} \cdot \frac{c}{1 + \left(\frac{s_j}{K}\right)^h}. \tag{6.18}$$

This expression typically models the kinetics of cooperative inhibition and was also used in [Gardner et al., 2000]. c is the maximum protein production rate in the absence of the repressor, K corresponds to the amount of repressor molecules that causes a 50 percent inhibition, and h is the cooperativity parameter. The derivative of $g_i\, F_i$ at a fixed point (s_i^*, s_j^*) reads:

$$g_i\, f_i^* = -\frac{c}{\left(1 + \left(\frac{s_j^*}{K}\right)^h\right)^2} \cdot \frac{h}{K} \cdot \left(\frac{s_j^*}{K}\right)^{h-1} = -\frac{h}{c\,K} \cdot (g_i\, F_i(s_j^*))^2 \cdot \left(\frac{s_j^*}{K}\right)^{h-1} = -\frac{h}{c\,K^h}(s_i^*)^2 (s_j^*)^{h-1}.$$

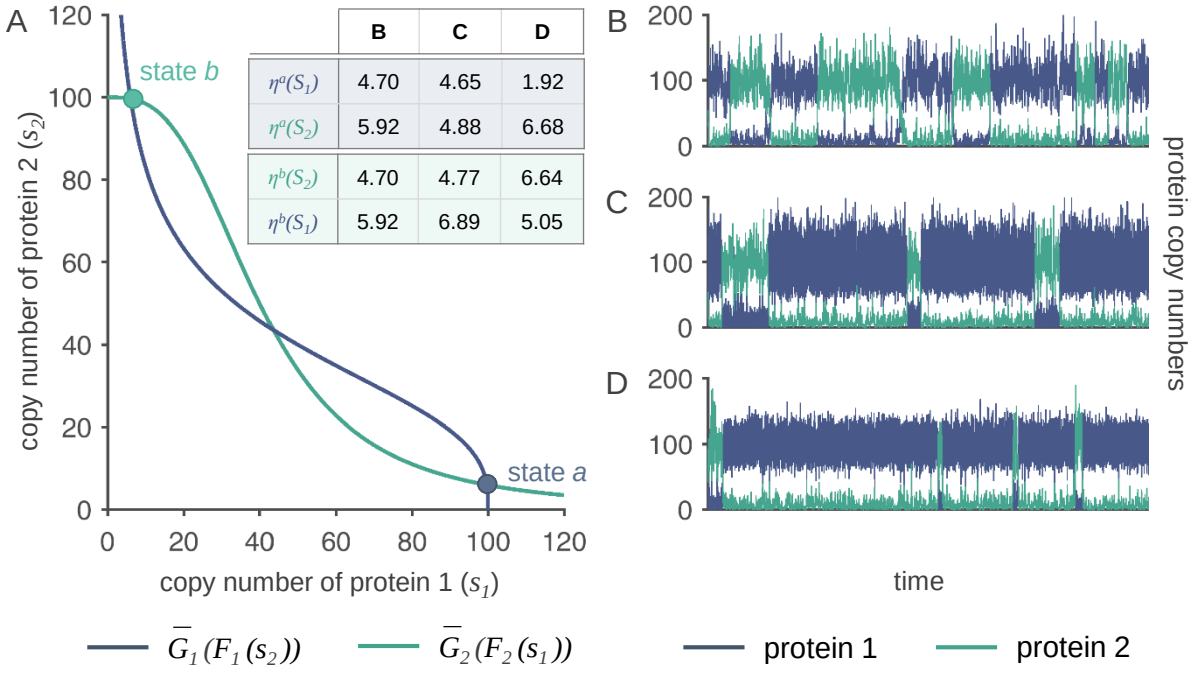
This results in the combined parameter:

$$R^a = \left(\frac{s_2^a}{s_1^a}\right)^{h-2}. \tag{6.19}$$

Without loss of generality, let $s_1^a > s_2^a$, i.e. let state a be characterized by dominant expression of protein 1. Obviously, R^a is equal to 1 if $h = 2$. In this case, $\eta^a(S_1) = \eta^a(S_2) = \eta^b(S_1) = \eta^b(S_2) = 1 + \frac{\Psi}{1-\Psi}$. Otherwise, asymmetric noise patterns can be created through a suitable choice of δ.

Fig 6.3 shows the time courses of three different toggle switches, which all have the same deterministic protein-protein phase plot that is illustrated in panel (A). In the toggle switch simulated in panel (B), all scaled parameters and functions of the two subsystems are identical, and $\delta = 1$. In accordance to that, the fluctuations are symmetric as well, without any of the states being preferred. In panel (C), the parameters have been maintained except for $\delta = 0.05$. This means that all processes in system 2 occur more slowly than in system 1. Compared with the circuit simulated in panel (B), the Fano factors of proteins 1 and 2 in state a are reduced, while they are increased in state b. This is plausible since $R^a < 1$ holds in this example (note that $h = 3$ and cf. Eq (6.19)), so that a reduction of δ leads to a stabilization of state a. One should mention that the simulations and calculated Fano factors in this figure were obtained using the full instead of the reduced model description. The principle deduced from the simplified model thus seems to be valid in the full model as well.

In the case of single-gene autoregulation, the decoupling of Fano factors in the two expression states could only be achieved through expression-state-specific circuit parameters. In the toggle switch, the dynamical interplay between the two subsystems, which is modulated through the state-*independent* parameter δ, can cause asymmetric noise despite all other parameters being symmetric. Interestingly, the way how the parameter needs to be shifted to create a certain noise pattern depends on the cooperativity of repression. These results demonstrate how the adjustability of noise is increased through the involvement of a second gene expression system.



	B	C	D
$\eta^a(S_1)$	4.70	4.65	1.92
$\eta^a(S_2)$	5.92	4.88	6.68
$\eta^b(S_2)$	4.70	4.77	6.64
$\eta^b(S_1)$	5.92	6.89	5.05

Figure 6.3.: **Creation of asymmetric noise patterns through state-independent circuit parameters.** (A) Protein-protein phase plot valid for all three systems whose simulated time courses are plotted on the right. Simulations were created with the Gillespie algorithm based on the full model description. The Fano factors of both proteins in the two expression states, calculated by numerically solving the Lyapunov equation of the full model, are given in the table. (B) Fully symmetric toggle switch with function parameters $c_1 = c_2 = 2.5$, $K_1 = K_2 = 40$, $h_1 = h_2 = 3$, $g_1 = g_2 = 40$, $\nu_1 = \nu_2 = 0.1$, $\delta = 1$ (cf. main text). For the burst sizes, $\alpha_1 = \alpha_2 = 4$ hold. (C) All parameter values as in (B), except for $\delta = 0.05$. This shift leads to a relative destabilization of state b. (D) Compared to (C), g_i and c_i are changed, while the products $g_i \cdot c_i$ are maintained: $g_1 = 10$, $g_2 = 60$, $c_1 = 10$, $c_2 = 5/3$. This changes the burst sizes to $\alpha_1 = 1$ and $\alpha_2 = 6$. The reduction of $\eta^a(S_1)$ further stabilizes state a. All time courses are shown in the time interval $[0, 2 \cdot 10^3/\delta]$; simulations have been initialized at the unstable fixed point.

6.3.3. Noise under bursting conditions

The Fano factor of protein 1 in the translational bursting model (6.9) can either be obtained by performing the LNA like in Section B.4 of the Appendix or by taking the limits $\nu_i \to 0$, $i = 1, 2$, in Eq (6.14), while $\frac{\nu_1}{\nu_2}$, $g_i \cdot \nu_i$, and $\frac{F_i}{\nu_i}$ are kept constant. Both approaches yield:

$$\begin{aligned}\eta^*(S_1) &= (1+\alpha_1) + \frac{g_1 f_1^*}{1-\Psi^*}\left((1+\alpha_1)\, g_2\, f_2^* \frac{\delta}{1+\delta} + (1+\alpha_2)\, g_1\, f_1^* \frac{s_2^*}{s_1^*}\frac{1}{1+\delta}\right) \\ &= (1+\alpha_1)\left[1 + \frac{\Psi^*}{1-\Psi^*}\left(\frac{\delta}{1+\delta} + \frac{1+\alpha_2}{1+\alpha_1} R^* \frac{1}{1+\delta}\right)\right]. \end{aligned} \tag{6.20}$$

$$\begin{aligned}\eta^*(S_2) &= (1+\alpha_2) + \frac{g_2 f_2^*}{1-\Psi^*}\left((1+\alpha_2)\, g_1\, f_1^* \frac{1}{1+\delta} + (1+\alpha_1)\, g_2\, f_2^* \frac{s_1^*}{s_2^*}\frac{\delta}{1+\delta}\right) \\ &= (1+\alpha_2)\left[1 + \frac{\Psi^*}{1-\Psi^*}\left(\frac{1+\alpha_1}{1+\alpha_2}\frac{1}{R^*}\frac{\delta}{1+\delta} + \frac{1}{1+\delta}\right)\right]. \end{aligned} \tag{6.21}$$

These formulae have the same basic structure as Eqs (6.15) and (6.16), and they demonstrate the additional impact of the burst sizes $\alpha_1 = g_1\nu_1$ and $\alpha_2 = g_2\nu_2$ on noise[3].

In contrast to the previous model simplification, an increase in g_i that is balanced by an inverse scaling of F_i leads to noise amplification because of the increased burst size. Moreover, the burst sizes influence the qualitative and quantitative effect of a shift in δ, as the decisive parameter composition is now given by $\frac{1+\alpha_2}{1+\alpha_1} \cdot R^*$. The fast mRNA limit can be regarded as a special case of the translational bursting model with $\alpha_i = 0$.

In Fig 6.3 (D), the asymmetry of the noise pattern in panel (C) is further modified through the generation of different burst sizes in the two subsystems, namely $\alpha_1 = 1$ and $\alpha_2 = 6$. The nullclines in the protein-protein phase plot are still symmetric. The simulated time courses suggest that state a is further stabilized through the modification of the burst sizes, but this observation is less intuitive, since the reduction of α_1 (compared with panel (C)) tends to reduce noise in both expression states, while the increase in α_2 has an opposite effect. By analyzing the structures of Eqs (6.20) and (6.21) more closely, one may nevertheless comprehend why state a is stabilized, although the argumentation is only qualitative and not predictive: To do so, we first note that in this example, $\delta = 0.05$ is again very small. In order to simplify the considerations, let us thus regard Eqs (6.20) and (6.21) in the limit $\delta \to 0$. This yields:

$$\eta^*(S_1) \approx (1+\alpha_1)\left[1 + \frac{\Psi^*}{1-\Psi^*} \cdot \frac{1+\alpha_2}{1+\alpha_1} R^*\right] \tag{6.22}$$

$$\eta^*(S_2) \approx (1+\alpha_2)\left[1 + \frac{\Psi^*}{1-\Psi^*}\right]. \tag{6.23}$$

Since $\eta^*(S_2)$ is in this case independent of R^*, the Fano factors $\eta^a(S_2)$ and $\eta^b(S_2)$ are expected to be similar[4]. This is indeed confirmed by the LNA-based estimates listed in Fig 6.3 (A). The real effect of $\frac{1+\alpha_2}{1+\alpha_1} R^*$ is thus reflected by $\eta^*(S_1)$. Again, $R^a < 1$ and $R^b = \frac{1}{R^a}$ hold. Since $\frac{1+\alpha_2}{1+\alpha_1} > 1$, the stabilizing effect of a reduction in δ on state a is diminished through the asymmetric burst sizes. However, at the same time, the destabilizing effect on state b is augmented over-proportially (note that $\frac{1+\alpha_2}{1+\alpha_1}R^b \gg \frac{1+\alpha_1}{1+\alpha_2}R^b = (\frac{1+\alpha_2}{1+\alpha_1}R^a)^{-1}$ and that $\eta^b(S_1) \gg (1+\alpha_1)$), which in the end might lead to the relative stabilization of state a.

All in all, the results show that the average burst sizes α_1 and α_2 have a severe impact on protein noise. By modulating promoter strengths and ribosomal binding sites, the overall circuit noise can thus be efficiently modified without altering average protein levels. In order to keep the noise level low, large bursts should be avoided.

6.3.4. Special cases in the full model

In the previous two model simplifications, the assumptions $\nu_1 \ll 1$ and $\nu_2 \ll 1$ have been made. In order to assess the effect of these two parameters, let us now consider two special limit cases of the full model.

[3] Here, the asterisks are omitted, since the burst sizes α_i are independent of the expression state.
[4] This was also the case in panel (C).

Case $\delta \to \infty$:

In this case, the dynamics of system 2 become infinitely fast compared to those of system 1. Taking the limit in the Fano factor of the full model, given in Eq (6.14), yields:

$$\begin{aligned}\eta^*(S_1) &= 1 + \frac{\Psi^*}{1-\Psi^*}\left(\frac{1}{1+\nu_1} + \frac{\nu_1}{1+\nu_1}\cdot\frac{1}{g_2 f_1^* f_2^*}\right) \\ &= \left(1+\frac{\alpha_1}{1+\nu_1}\right) + \frac{\Psi^*}{1-\Psi^*}\left(1+\frac{\alpha_1}{1+\nu_1} - \frac{\nu_1}{1+\nu_1}\right). \end{aligned} \tag{6.24}$$

The formulation in the first row is equivalent to that of the protein Fano factor in a single-gene autoregulatory system, cf. Eq (4.19). This is because the fast dynamics of subsystem 2 are condensed into a single positive feedback reaction, given by the function $F = F_1(g_2 \cdot F_2(.))$. This implies that autoregulation can even be interpreted as a special case of the toggle-switch.

The effect of ν_1 is thus comparable with the effect of ν in the autoregulatory system with linear translation. The parameter combination $f^* = g_2 f_1^* f_2^*$ determines whether a shift in ν_1 leads to a local increase or decrease of the Fano factor. Usually, the average protein copy number is larger than the average copy number of the corresponding mRNA, so that $\frac{1}{g_2 f_1^* f_2^*} = \frac{g_1}{\Psi^*} > g_1 = \frac{s_1^*}{m_1^*} > 1$. An increase in ν_1, which corresponds to a slow-down of mRNA dynamics in system 1, would therefore in general augment the fluctuations of protein 1.

The formulation in the second row of Eq (6.24) allows an interpretation in terms of protein bursts. For comparison, we also take the limit $\delta \to \infty$ in the Fano factor (6.20) of the reduced bursting model, resulting in the expression:

$$\eta^*(S_1) = (1+\alpha_1) + \frac{\Psi^*}{1-\Psi^*}(1+\alpha_1). \tag{6.25}$$

If the burst size α_1 was chosen identically in Eqs (6.24) and (6.25), the Fano factor according to Eq (6.24) would be smaller, because this formula takes protein degradation into account, which reduces the effective burst amplitude (cf. Section 5.2.1). One can conclude that although an increase in ν_1 augments the burst size α_1 and thus also the Fano factor, this kind of noise amplification is not as strong the one achieved through an increase in g_1 (see also Section 5.2.2).

Case $\delta \to 0$:

When system 2 is slowed down, the Fano factor is given by:

$$\eta^*(S_1) = \left(1+\frac{\alpha_1}{1+\nu_1}\right) + \frac{\Psi^*}{1-\Psi^*}\cdot R^*\left(1+\frac{\alpha_2}{1+\nu_2} - \frac{\Psi^*\nu_2}{1+\nu_2}\right). \tag{6.26}$$

g_2, like g_1, is typically larger than 1. Using the same argument as above, one can show that an increase in ν_1 and/or ν_2 augments $\eta^*(S_1)$, but again, the effect is less pronounced than an amplification of burst sizes through g_1 or g_2.

After having regarded these two limit cases, it seems that slow mRNA dynamics tend to make the system noisier. Concerning the asymmetry of fluctuation patterns, a simulation of a toggle switch was performed, whose parameters are equivalent to those of panel (C), but with changed values of ν_i, so that the burst sizes α_1 and α_2 of panel (D) are generated. It turned out that the calculated Fano factors as well as the fluctuation patterns were similar to that of panel (D), but the asymmetry was slightly less pronounced (data not shown).

6.4. Short discussion and summary

This study has shown that the toggle switch offers a broader range of possibilities for the adaptation of noise patterns than single-gene autoregulation. The basic principles of noise generation through bursts and noise propagation through sensitive transcriptional and translational propensities are maintained. However, the time-scale parameter δ, which relates the dynamics of the two subsystems to each other, provides an additional means to create asymmetric noise patterns.

The influence of circuit parameters has been studied with the help of several different, simplified model versions, since the mRNA and protein Fano factors of the detailed model turned out to be too complex to be analyzed comprehensively and comprehensibly. Based on a model reduction that assumed infinitely fast mRNA dynamics, fundamental effects concerning the interplay between the two subsystems on circuit noise were studied. In this context, a special role could be attributed to differences in time-scales of the two subsystems and to the degree of cooperativity in mutual transcriptional repression. The bursting model then elucidated the additional impact of protein bursts and demonstrated that the previous elimination of the mRNA level (which is quite commonly applied in deterministic modeling) led to a quantitative underestimation of noise. The qualitative principles, however, remained valid. The effect of time-scale-distributions within the subsystems (interplay of mRNA and protein dynamics) was then analyzed by regarding special cases of the full model. Although the Fano factors obtained with the reduced models may not be accurate, the analysis of the corresponding formula led to a better understanding of circuit noise. The qualitative design of noise patterns can now first be performed based on these insights, then be tested with the help of the Fano factors according to the full model (which can be numerically evaluated), or with stochastic simulations.

In this study, is was assumed that translation is a linear function of the mRNA copy number, which is why the hLNA approach was equivalent to the classical LNA. Nonlinearity of translation would be an additional effective tool to modify noise levels as it would generate state-dependent burst sizes, which were previously shown to have a strong influence on protein noise.

In general, the assembly of a genetic toggle switch can be done quite flexibly: A lot of repressor-promoter combinations are available that differ in the strength and cooperativity of the repressor and in the strength of the promoter. Moreover, ribosomal binding sites can be varied or genetically modified, and even mRNA and protein half-lives can be altered to a certain extent. A model-driven selection of synthetic building blocks can thus lead to the construction of a bistable switch with desired deterministic and stochastic properties.

7. Discussion and Conclusions

7.1. Overall discussion of results

The main goal of this study was the development and subsequent usage of mathematical methods, with which noise patterns in multistable genetic circuits can be quantitatively analyzed. In particular, the connection between expression-state-specific noise levels and the dynamic properties of the underlying circuit should be clarified.

For the characterization of noise, two different types of measures were used: First of all, the so-called Fano factors provided information on the variances of target mRNA and protein fluctuations. Second, characteristics of protein bursts were used to take the temporal structure of protein fluctuations into account. All these measures were chosen to be state-dependent (or, say, subpopulation-specific).

The quantification of mRNA and protein Fano factors was performed using the hybrid linear noise approximation (hLNA), which was developed in Chapter 4. It is a variant form of the classical LNA approach proposed in [van Kampen, 2007], but better suited to model mesoscopic systems with reactive species occurring in low copy numbers. This is of particular importance in the context of gene expression, where components of low molecular abundance like mRNA species are involved. In the derivation of the hLNA, the main advantages of the classical LNA approach were maintained. Among these is the feasibility of finding approximate, but explicit formulae, which describe all state-specific mRNA and protein variances in dependence of relevant properties of the circuit. The (mainly simulation-based) assessment of the quality of the formulae demonstrated that they reliably described the impact of feedback-regulated transcription, of translation, and of target molecule degradation on noise. In particular, it could be shown that the hLNA approach significantly outperformed the classical LNA in all considered examples, cf. Fig 4.5.

In addition to the variance of fluctuations, their structure was characterized using novel definitions of state-dependent protein burst sizes and frequencies, which provide an extension to the static definitions in [Thattai and van Oudenaarden, 2001]. The link between those burst characteristics and the occurrence of peaks in protein time courses was discussed for better interpretability in Section 5.2.1.

The intuitive understanding of noise was further improved by drawing connections between Fano factors and burst characteristics, cf. Section 5.3.1. For example, it was shown that fast protein degradation reduces the observable amplitude of protein peaks, so that the corresponding Fano factor might turn out to be small in spite of bursty protein production. Moreover, small but frequent bursts were demonstrated to cause smaller Fano factors than large, rare bursts. While these observations had in principle already been made for simple, linear regulatory systems [Ozbudak et al., 2002; Paulsson, 2005; Thattai and van Oudenaarden, 2001], further important insights were gained by looking at bursts under feedback regulation: It was shown that positive feedback resulted in a self-enforcing accumulation of burst events, which amplified the variance

of fluctuations, while under negative feedback, bursts were triggered in order to balance below-average protein levels, which led to noise attenuation. A further, novel aspect was that besides the burst frequency, also the burst size might vary in unison with local fluctuations. This occurs due to translation that is non-proportional to the mRNA level. For instance, concavity of the translational propensity function was demonstrated to enable local noise reduction (at least under non-negative feedback), since above-average mRNA copy numbers were compensated for by a reduced amount of produced proteins, while low mRNA levels were accompanied by stronger translational activity.

The formula-based study of the connection between circuit properties and noise was furthermore supported by graphical illustrations, which are related to the phase plots commonly deployed in deterministic model analysis for specifying the location and stability of fixed points. By contrast, the graphical tools developed in this study helped to determine state-specific noise measures qualitatively (like the Fano factor in Fig 5.4 (A)) or even quantitatively (like the burst size in Fig 5.4 (B)). The latter plot highlighted the role of relative time-scales in the interplay of mRNA and protein dynamics on protein bursting. Through the direct visual connection between noise measures and the graphs of the propensity functions, which reflect the dynamical properties of the regulatory system, the plots were suitable to support the *in silico* design of circuits with desired noise patterns. In particular, not only the local noise characteristics could be read for every stable expression state, but noise levels of different states within the same circuit could be compared directly. A further graphical method, shown in Fig 5.4 (C), allowed the determination of modes in the protein distribution in dependence of the protein burst size, under the premise that mRNA half-life is very short. The plot not only illustrated the impact of bursts on the asymmetry (skewness) of fluctuations, it also explained how this asymmetry might lead to the disappearance or to the *de novo* formation of modes, a phenomenon that could not be foreseen with any kind of deterministic model. The way how the dependence of stochastic circuit behavior on the burst size was studied graphically resembles deterministic bifurcation analysis, although the burst size as a "stochastic bifurcation parameter" did not influence the stationary state of the macroscopic model[1].

The observed deviation between deterministic and stochastic model behavior was studied more extensively in Chapter 3. There, the discrepancy between deterministic variables and stochastic modes under bursting conditions, potentially leading to systems that are bistable but unimodal, or monostable but bimodal, was systematically analyzed and explained. This suggested a differentiation between circuits that are bimodal due to bistability, circuits whose bimodality is at least associated with principles of bistability (e.g. cooperative feedback), although the deterministic model postulates monostability, and circuits that are bimodal without exhibiting any features of bistability, where one of the modes emerges at the zero state due to a local accumulation of the skewed probability distribution. On the other hand, unimodality of bistable systems, caused by a coalescence of modes, could be traced back to the same origin, namely to the asymmetry of fluctuations due to large bursts. Once more, the usage of the above mentioned graphical method led to a more intuitive understanding of all these observations, cf. Fig 3.2. It should however be noted again that the scenarios considered here are not the only ones in which deviations between bistability and bimodality may occur. Here, the reader is referred to the discussion in Section 3.5.

Chapter 3 also studied the connection between deterministic variables and the average behav-

[1] This holds if the burst frequency was accordingly modulated to maintain the locally average protein expression level.

ior of the system. The main reason for potential deviations was identified to be the nonlinearity of reaction propensities, which hampers its commutativity with the expected value operator, as already reported in [van Kampen, 2007]. This fact basically motivated the formulation of effective deterministic rates in the development of the hybrid deterministic model (cf. Chapter 4), which served as a basis for the hLNA and for the state-dependent characterization of protein bursts. This hybrid approach might be considered pragmatic, since it simply assumes the copy number of any mRNA species to be Poisson distributed. However, it is easily and broadly applicable, and the novel deterministic variables turned out to well approximate the local (i.e. expression-state-specific) mean values – an important prerequisite for the quality of the hLNA. In Section B.5, possible extensions of the hybrid ansatz were discussed, especially concerning the replacement of the Poisson distribution with a more accurate one and concerning the expansion of this idea to the protein level. In other publications, different (and maybe more rigorous) ansatzes were chosen in order to improve the quality of deterministic descriptions in the context of mesoscopic systems [Gómez-Uribe and Verghese, 2007; Grima, 2010]. However, these approaches are not applicable to multistable systems, since they approximate real (and not local) mean values, and are therefore not suitable for studying macroheterogeneity.

The hybrid methods developed for the state-specific characterization of noise were applied to two bistable regulatory motifs, namely the single-gene autoregulatory feedback loop in Chapter 5 and the genetic toggle switch in Chapter 6. In case of the autoregulatory system, the dependence of the Fano factor on the circuit properties could be discussed in a straightforward manner based on the formulae resulting from the hLNA. However, it took a lot more effort in case of the genetic toggle switch to analytically solve the Lyapunov equation for the approximate variances. Morever, the obtained expressions were of much higher complexity. This demonstrated that although the (classical or hybrid) LNA approach is in principle applicable to systems of any size, the interpretability of results might be hampered by the number of involved components. Numerical evaluations of the Lyapunov equation under systematic variation of parameters are however possible for larger systems as well, and they are expected to be much more efficient and easier to perform than the extraction of values from a huge number of stochastic simulations, especially in the context of multimodal systems.

The problem of high complexity in the analysis of the genetic toggle switch was circumvented by regarding certain limit cases that simplified the structure of the Fano factors. In this context, the previous analysis of the single-gene autoregulatory system turned out to be helpful, as many of the basic principles could be shown to hold in the two-gene system as well. One of these principles that were identified in the autoregulatory circuit was the importance and efficiency of nonlinear rates in the state-specific adaptation of noise levels. In particular, a nonlinear relation between the mRNA level and translational activity, whose local effect on Fano factors has just been discussed, was also demonstrated to enable a flexible adjustment of noise patterns across different expression states. The example considered at the end of Chapter 4 showed that a translational propensity which is a concave function of the mRNA level is highly sensitive to mRNA fluctuations in the inactive state, while its saturation impedes noise propagation in the active state. Concerning the protein time courses, significant protein bursts could be observed in the OFF state, while their size was severely reduced after activation. It is indeed justified to postulate nonlinear translation propensities, as experimental studies have shown that the mRNA level is often nonproportional to translational activity, possibly due to limitations of cellular resources [Mather et al., 2013; Mauro and Edelman, 2002; Rodnina, 2016]. Nonlinearity may also occur in other reactions like degradation [Buchler et al., 2005], which could be analyzed analogously using the same new methods proposed in this work. Compared to single-gene

autoregulation, the flexibility to adjust noise patterns was shown to be higher in the toggle switch due to the involvement and close interaction of two gene expression systems. A uniform shift in the reaction time-scales in only one of the subsystems was demonstrated to enable asymmetric noise levels between the two stable expression states, although the toggle switch was otherwise constructed symmetrically and with constant average burst sizes (cf. Fig 6.3). However, the analysis of the reduced bursting model once again illustrated the predominant role of bursts in noise regulation and suggested that an efficient adaptation of noise patterns in the toggle-switch may be achieved through state-dependent burst characteristics.

Although noise in gene expression and regulation has been theoretically studied in many other publications as well (cf. Section 2.2.3), the work at hand differs in the approaches that were used or developed and partially also in its goals: Throughout the present study, all stochastic approaches were derived from the chemical master equation in order to obtain a detailed characterization of noise. In contrast, many other publications directly started with a description using the Langevin formalism [Ozbudak et al., 2002; Simpson et al., 2003; Zheng et al., 2011], where Gaussian white noise terms are added to deterministic rate equations. The latter ansatz obscures the real structure of fluctuations, but turned out to be valuable for studying approximate effects of noise in larger reaction systems [Ceccato and Frezzato, 2018]. While in principle, the hLNA can also be regarded as a deterministic description with an added Gaussian noise term, the particular formulation of the hybrid deterministic equations was motivated by the trial to achieve a closer approximation to the CME. In addition, the analysis of the temporal structure of fluctuations was derived from the CME as well and revealed a fluctuation pattern which might significantly deviate from white noise due to high asymmetry. The combination of information about the variance and structure of fluctuations could characterize circuit noise in a comprehensive manner, which was one of the main focuses of this study. Besides that, the analytical studies of connections between properties of the circuit and the different noise measures form a contrast to those that are mainly simulation-based [Cagatay et al., 2009; Jaruszewicz and Lipniacki, 2013; Maamar et al., 2007]. The latter studies are again valuable for studying more complex regulatory systems, while the analytic approach helped in the derivation of more general results and insights. On the other extreme, several studies dealt with exact solutions of the chemical master equation [Friedman et al., 2006; Houchmandzadeh and Vallade, 2015; Paulsson and Ehrenberg, 2000; Shahrezaei and Swain, 2008; Walczak et al., 2012]. While providing fundamental knowledge, e.g. about the analytical protein distribution in basic regulatory systems, their approaches are usually restricted to systems with very low complexity, whereas the approximate approaches presented here are more flexible.

An improved, generalized understanding of the emergence and propagation of circuit noise is beneficial in various applications: Using the methods proposed here, one may study and compare different circuit architectures and interpret the findings with respect to the functionality of the noise pattern that is generated. This would be an important step towards the identification of evolutionarily shaped design principles of noise and heterogeneity [Lehner, 2008]. In this work, the circuit triggering the entry of *S. mutans* cells into the competence cascade has been discussed. Its activation only occurs in a subpopulation of isogenic cells, but seems to be unidirectional, as observed experimentally [Reck et al., 2015]. This was postulated to be the result of an asymmetrical distribution of noise levels between the two stable expression states, and it was explained how this could be achieved on a genetic level. Here, the fact was used that in general, the strength of fluctuations in a certain expression state is associated with its robustness. However, a generalized quantitative analysis of robustness (e.g. in terms of first passage times) based on noise patterns has not yet been established. An accurate prediction or

design of population heterogeneity is therefore not possible yet, but the local noise levels provide at least good qualitative indications on robustness, as confirmed by the simulations in Fig 5.5.

The examination of noise patterns is also expected to be valuable in the field of biotechnology [Levchenko and Nemenman, 2014]. For example, artificial cellular switches may become better controllable through model-guided genetic manipulations that lead to noise reduction [Cardelli et al., 2016]. As shown before for monostable systems, efficient noise attenuation can be achieved through a relative increase in transcriptional activity, which is balanced by reduced translation [Ozbudak et al., 2002], since this suppresses protein bursts[2]. The present work has demonstrated that under positive feedback regulation, one additionally needs to take into account the possible generation of super-bursts, which might destabilize the system. They can be partially prevented through feedback regulation with high cooperativity, whose strongly sigmoid transcriptional propensity function is locally flat in the two stable expression states, ensuring a homogeneous temporal distribution of bursts.

Similar considerations can be made in applications of synthetic biology [Bandiera et al., 2016; Murphy et al., 2010]. Until now, the design of artificial circuits has mostly been based on deterministic considerations only, which worked well in many cases, e.g. in the construction of the genetic toggle switch [Gardner et al., 2000]. However, since these circuits have never been subject to evolutionary pressure, designing their noise levels might be equally important in order to ensure functionality [Oyarzún et al., 2015]. This does not only hold for bistable systems, which have been considered here, but also for circuits showing oscillatory behaviour [Elowitz and Leibler, 2000; Potvin-Trottier et al., 2016].

Synthetic circuits are probably the ones suited best for verifying the results of this or of similar studies experimentally: They are expected to show only few interference with other intracellular processes, although some global regulators, environmental triggers, and, of course, cell cycle dynamics inevitably affect the dynamics of any circuit. Using apt single-cell analysis techniques like microfluidics coupled to time-lapse microscopy, at least extracellular disturbances can be minimized and the population could be kept in a stationary state [Grünberger et al., 2015]. However, the volume of cells constantly varies because of growth and division – a fact that has been neglected in this study, not because it is considered unimportant, but because it severely aggravates the analytical treatment of models: The constant but stochastic change in cellular volume, the discrete doubling of gene dosage during replication, and the random distribution of cellular components on daughter cells in the course of division [Huh and Paulsson, 2011] – all these factors influence the stochasticity of the circuit and keep the metabolic state of a cell changing. Even average concentrations of cellular components are usually subject to some regular deterministic fluctuations [Bierbaum and Klumpp, 2015], so that the analysis of stationary states is in principle not possible. Due to the complexity of the processes, such studies need to be simulation-based, which aggravates the deduction of general rules from the obtained results. It therefore makes sense to first analytically study the isolated gene regulatory system in stationary state, as it was done here, and to then check which of the observations can still be confirmed based on simulations including the cell cycle and based on experimental data. Several interesting studies have already been performed that modeled gene expression under the influence of cellular dynamics [Gomez et al., 2014; Lu et al., 2004; Luo et al., 2013]. However, the mechanisms regulating cell cycle itself and the various mutual interactions with

[2]This constellation of strong transcription and weak translation, which leads to high mRNA levels, is actually highly inefficient from an energetic perspective, but is justified under the aspect of noise reduction and is therefore expected to be relevant in many natural circuits, too.

gene expression[3] are highly intricate and not fully understood even in *E. coli* [Adiciptaningrum et al., 2016; Taheri-Araghi et al., 2015; Wallden et al., 2016; Wang and Levin, 2009; Wang et al., 2010; Willis and Huang, 2017]. Further experimental and theoretical research is therefore still required.

Besides the negligence of cell growth and division, further model simplifications have been made in this study, among which was the description of reactions as non-elementary processes. In particular, transcription and translation are actually events that are composed of many single steps, which were assumed to be fast enough to be lumped into a single process. However, it is known that this assumption is not always valid: For example, irregular elongation dynamics were shown to be a potential trigger of transcriptional bursts [Chong et al., 2014]. In this context, one can also revisit the two model reductions made in the analysis of the genetic toggle switch: In the first, the mRNA dynamics were assumed to be fast enough to be averaged out – a conjecture that is rather frequently made, at least in deterministic modeling [Alon, 2007; Gardner et al., 2000; Süel et al., 2006]. However, the second model reduction showed that if mRNA dynamics are only faster than those of protein degradation, but not of protein formation, protein bursts emerge that have a major impact on noise. The omission of intermediate components or reaction steps may thus distort the true behavior and properties of the system. But again, even the most simplified description of the toggle switch provided beneficial insights into some basic regulatory principles of the circuit, which were shown to also hold in the more complex model formulations.

Another problem occurs when using non-elementary reactions: All these events actually require some time, so that the effect of feedback regulation is always subject to a certain temporal delay. An inclusion of these delays into the model is possible, but again, this renders the mathematical analysis more complicate. Nevertheless, this would be a valuable further step in the examination of gene expression noise, since interesting fluctuation patterns exhibiting e.g. oscillatory behavior might occur [Bratsun et al., 2005].

A further model simplification, which has already been mentioned in the beginning of Chapter 2, is the assumption of spatial homogeneity – actually a prerequisite for the application of the chemical master equation. Although in this work, the CME was used as the gold standard, to which all other approaches were compared, one should keep in mind that even the CME, no matter how detailed its formulation is chosen, would never be capable of describing the system accurately. Only a comparison with experimental data will show which of the simplifying assumptions are valid and which are not, but since this would entail a lot of new challenges (uncertainty of measured data, superposition of intrinsic noise with other biological sources of stochasticity, parameter estimation in a stochastic regime, etc.), this was beyond the scope of this study.

[3]The production of proteins may influence the state of the cell either directly through the properties of the protein (e.g. toxicity, or a role in metabolism as an enzyme) or indirectly through metabolic burdening. On the other hand, the changing volume and composition of a cell during the cell cycle affects the reaction propensities and makes them time-dependent. All these mutual interactions depend on the average growth rate and therefore also on the environment [Hintsche and Klumpp, 2013; Klumpp and Hwa, 2014; Marguerat and Bähler, 2012].

7.2. Conclusions and Future Directions

This study has dealt with the detailed characterization of intrinsic noise in multistable gene expression systems. For this characterization, a theoretical framework was established and applied, which enables an expression-state-specific description and prediction of mRNA and protein noise based on the stochastic reaction propensity functions of the circuit. This framework thus allows the quantitative examination of microheterogeneities. Since local noise, if large enough, may lead to random state transitions, qualitative statements about macroheterogeneity could be deduced as well. This, on the one hand, allows the prediction of noise patterns based on the topological and dynamic features of a given circuit; on the other, it helps in the design of circuits with defined stochastic properties. In particular, manipulations on the genetic level can be identified which lead to the desired modulations of noise levels.

A comprehensive characterization of intrinsic noise was ensured through the combination of different noise measures: Population distributions of mRNA and protein copy numbers were specified via the mean values and variances of the single peaks in the probability mass functions. Additionally, the temporal structure of protein fluctuations and their skewness were described with the help of translational bursting properties. To this end, novel mathematical methods were developed, which provide closed-form approximations of the state-specific noise measures, showing their dependence on the reaction propensities. This dependence could also be depicted graphically through illustrations that resemble deterministic phase plots. Through this visual connection, the intuitive understanding of the emergence and propagation of circuit noise was promoted. Additionally, connections between all the different noise measures were examined in order to enhance their interpretability. Based on these insights, statements were derived about how fluctuations within expression states can be influenced locally through genetic manipulations, and how the noise levels of different states can be adjusted rather independently of each other in order to create desired noise patterns. The framework is basically applicable to all kinds of genetic circuits, although the obtained closed-form solutions might become intricate for systems with many reactions and reactive species. In this work, it was tested on two common regulatory motifs, namely a single-gene expression system with autoregulated transcription, and the genetic toggle switch, where two gene products mutually inhibit each other.

Compared to simple models of linear gene expression, the regarded circuits exhibited two additional features, namely feedback regulation and a nonlinear translational propensity function. These features turned out to have a significant influence on local noise levels by making protein burst frequencies and sizes vary in accordance with the current state of the system, thereby amplifying or attenuating noise propagation. Nonlinear translation was also shown to enable highly different bursting regimes in the two expression states of a bistable system. Through this asymmetric distribution of noise levels, unidirectional state-transitions are promoted. This might play a role in cellular decision-making, like in the bimodal development of genetic competence in isogenic *S. mutans* populations.

The newly developed mathematical methods succeeded in overcoming the two expected challenges mentioned in the beginning (cf. Chapter 1); namely finding a way to characterize intrinsic noise in an expression-state (or subpopulation-specific) manner, and formulating these local noise characteristics in terms of (approximate) closed-form expressions: First of all, each mode in the probabilistic state distribution was associated with one stable fixed point of a corresponding deterministic description. Local simplifications of the stochastic model around the fixed points then led to closed-form noise estimates. The deterministic description used here was newly de-

veloped in order to appropriately describe the dynamics of mesoscopic systems. This resulted in an improved quality of noise estimates compared to classical approaches like the linear noise approximation, which is based on rate equations.

Despite the improvement of the deterministic model, it could be shown that deviations between deterministic bistability and stochastic bimodality may occur: A systematic study supported by graphical methods demonstrated that translational bursts may cause the coalescence of multiple modes or the formation of new modes through stochastic bifurcations, which cannot be captured by any kind of deterministic model. In these cases, a state-dependent characterization of noise using the novel theoretical framework is not possible, due to the missing link between modes and fixed points.

After having established a basic, comprehensive framework for characterizing intrinsic noise patterns, the following steps are recommended for future work: First of all, the gained knowledge might be used to further establish approaches with which transition probabilities between different expression states can be predicted. This would allow a full quantitative description and prediction of phenotypic heterogeneity. To this end, it might be valuable to relate the information about local noise levels to more global properties of the circuit, e.g. to the basins of attraction of the corresponding stable fixed points.

Furthermore, the reliability of the established mathematical framework should be verified experimentally, using suitable single-cell techniques like microfluidics and different synthetic constructs for comparison and validation. The inevitable overlap of intrinsic noise with other contributions to cellular noise, in particular with cell-cycle-dynamics, needs to be resolved. To that end, established experimental approaches for the discrimination of intrinsic and extrinsic noise may be used. Moreover, model-based methods, with which the cell cycle dynamics can either be integrated into the stochastic description of the circuit, or eliminated from the experimental measurements, should make experimental and theoretical data better comparable. This would show whether the description of intrinsic noise is accurate, or if it needs modification. Extensions to or modifications of the newly established methods are thinkable and have already been discussed, e.g. possible extensions of the principles of the hybrid LNA. Moreover, the level of detail in the description of the circuit may be increased, e.g. through the inclusion of promoter dynamics or through the introduction of time delays for reactions that actually consist of several reactive steps. After having obtained a full description of cell-associated noise sources, it could be further extended in order to study the effect of external disturbances. This would finally yield a full stochastic description of the genetic circuit with which its functionality and robustness can be assessed.

Bibliography

Ackermann, M. (2015). A functional perspective on phenotypic heterogeneity in microorganisms. *Nat. Rev. Microbiol.*, 13(8):497–508.

Adiciptaningrum, A., Osella, M., Moolman, M. C., Cosentino Lagomarsino, M., and Tans, S. J. (2016). Stochasticity and homeostasis in the E. coli replication and division cycle. *Sci. Rep.*, 5(1):18261.

Almquist, J., Cvijovic, M., Hatzimanikatis, V., Nielsen, J., and Jirstrand, M. (2014). Kinetic models in industrial biotechnology – Improving cell factory performance. *Metab Eng*, 24:38–60.

Alon, U. (2007). *An Introduction to Systems Biology: Design Principles of Biological Circuits*, volume 10. Chapman & Hall/CRC.

Altschuler, S. J. and Wu, L. F. (2010). Cellular Heterogeneity: Do Differences Make a Difference? *Cell*, 141(4):559–563.

Anderson, D. F. and Kurtz, T. G. (2015). *Stochastic Analysis of Biochemical Systems*. Springer International Publishing, Cham.

Anesiadis, N., Cluett, W. R., and Mahadevan, R. (2008). Dynamic metabolic engineering for increasing bioprocess productivity. *Metab. Eng.*, 10(5):255–266.

Aquino, T., Abranches, E., and Nunes, A. (2012). Stochastic single-gene autoregulation. *Phys. Rev. E. Stat. Nonlin. Soft Matter Phys.*, 85(6 Pt 1):061913.

Artyomov, M. N., Das, J., Kardar, M., and Chakraborty, A. K. (2007). Purely stochastic binary decisions in cell signaling models without underlying deterministic bistabilities. *Proc. Natl. Acad. Sci. U. S. A.*, 104(48):18958–63.

Baert, J., Kinet, R., Brognaux, A., Delepierre, A., Telek, S., Sørensen, S. J., Riber, L., Fickers, P., and Delvigne, F. (2015). Phenotypic variability in bioprocessing conditions can be tracked on the basis of on-line flow cytometry and fits to a scaling law. *Biotechnol. J.*, 10(8):1316–1325.

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

Bandiera, L., Furini, S., and Giordano, E. (2016). Phenotypic Variability in Synthetic Biology Applications: Dealing with Noise in Microbial Gene Expression. *Front. Microbiol.*, 7:479.

Bar-Even, A., Paulsson, J., Maheshri, N., Carmi, M., O'Shea, E., Pilpel, Y., and Barkai, N. (2006). Noise in protein expression scales with natural protein abundance. *Nat. Genet.*, 38(6):636–43.

Becskei, A. and Serrano, L. (2000). Engineering stability in gene networks by autoregulation. *Nature*, 405(6786):590–3.

Bednarz, M., Halliday, J. A., Herman, C., and Golding, I. (2014). Revisiting bistability in the lysis/lysogeny circuit of bacteriophage lambda. *PLoS One*, 9(6):e100876.

Bierbaum, V. and Klumpp, S. (2015). Impact of the cell division cycle on gene circuits. *Phys.*

Biol., 12(6):066003.

Bishop, L. M. and Qian, H. (2010). Stochastic bistability and bifurcation in a mesoscopic signaling system with autocatalytic kinase. *Biophys. J.*, 98(1):1–11.

Brandman, O., Ferrell, J. E., Li, R., and Meyer, T. (2005). Interlinked fast and slow positive feedback loops drive reliable cell decisions. *Science*, 310(5747):496–8.

Bratsun, D., Volfson, D., Tsimring, L. S., and Hasty, J. (2005). Delay-induced stochastic oscillations in gene regulation. *Proc. Natl. Acad. Sci. U. S. A.*, 102(41):14593–8.

Buchler, N. E., Gerland, U., and Hwa, T. (2005). Nonlinear protein degradation and the function of genetic circuits. *Proc. Natl. Acad. Sci. U. S. A.*, 102(27):9559–9564.

Cagatay, T., Turcotte, M., Elowitz, M. B., Garcia-Ojalvo, J., and Süel, G. M. (2009). Architecture-dependent noise discriminates functionally analogous differentiation circuits. *Cell*, 139(3):512–22.

Cai, L., Friedman, N., and Xie, X. S. (2006). Stochastic protein expression in individual cells at the single molecule level. *Nature*, 440(7082):358–362.

Cameron, D. E. and Collins, J. J. (2014). Tunable protein degradation in bacteria. *Nat. Biotechnol.*, 32(12):1276–81.

Cao, Y., Gillespie, D. T., and Petzold, L. R. (2005). The slow-scale stochastic simulation algorithm. *J. Chem. Phys.*, 122(10):14116.

Cardelli, L., Csikász-Nagy, A., Dalchau, N., Tribastone, M., and Tschaikowski, M. (2016). Noise Reduction in Complex Biological Switches. *Sci. Rep.*, 6(1):20214.

Carrier, T. and Keasling, J. (1997). Controlling Messenger RNA Stability in Bacteria: Strategies for Engineering Gene Expression. *Biotechnol. Prog.*, 13(6):699–708.

Ceccato, A. and Frezzato, D. (2018). Remarks on the chemical Fokker-Planck and Langevin equations: Nonphysical currents at equilibrium. *J. Chem. Phys.*, 148(6):064114.

Cerulus, B., New, A. M., Pougach, K., and Verstrepen, K. J. (2016). Noise and Epigenetic Inheritance of Single-Cell Division Times Influence Population Fitness. *Curr. Biol.*, 26(9):1138–47.

Chalancon, G., Ravarani, C. N. J., Balaji, S., Martinez-Arias, A., Aravind, L., Jothi, R., and Madan Babu, M. (2012). Interplay between gene expression noise and regulatory network architecture. *Trends Genet.*, 28:221–232.

Chong, S., Chen, C., Ge, H., and Xie, X. S. (2014). Mechanism of transcriptional bursting in bacteria. *Cell*, 158(2):314–26.

Delvigne, F., Baert, J., Sassi, H., Fickers, P., Grünberger, A., and Dusny, C. (2017). Taking control over microbial populations: Current approaches for exploiting biological noise in bioprocesses. *Biotechnol. J.*, 12(7):1600549.

Delvigne, F. and Goffin, P. (2014). Microbial heterogeneity affects bioprocess robustness: Dynamic single-cell analysis contributes to understanding of microbial populations. *Biotechnol. J.*, 9(1):61–72.

Dubnau, D. and Losick, R. (2006). Bistability in bacteria. *Mol. Microbiol.*, 61(3):564–572.

Eldar, A. and Elowitz, M. B. (2010). Functional roles for noise in genetic circuits. *Nature*, 467(7312):167–73.

Elf, J. and Ehrenberg, M. (2003). Fast evaluation of fluctuations in biochemical networks with

the linear noise approximation. *Genome Res.*, 13(11):2475–2484.

Elowitz, M. B. and Leibler, S. (2000). A synthetic oscillatory network of transcriptional regulators. *Nature*, 403(6767):335–338.

Elowitz, M. B., Levine, A. J., Siggia, E. D., and Swain, P. S. (2002). Stochastic gene expression in a single cell. *Science (80-.).*, 297(5584):1183–1186.

Friedman, N., Cai, L., and Xie, X. S. (2006). Linking stochastic dynamics to population distribution: an analytical framework of gene expression. *Phys. Rev. Lett.*, 97(16):168302.

Gardiner, C. (2009). *Stochastic Methods: A Handbook for the Natural and Social Sciences*, volume 13. Springer Berlin.

Gardner, T. S., Cantor, C. R., and Collins, J. J. (2000). Construction of a genetic toggle switch in Escherichia coli. *Nature*, 403(6767):339–342.

Ghusinga, K. R., Dennehy, J. J., and Singh, A. (2017). First-passage time approach to controlling noise in the timing of intracellular events. *Proc. Natl. Acad. Sci. U. S. A.*, 114(4):693–698.

Gibson, M. A. and Bruck, J. (2000). Efficient exact stochastic simulation of chemical systems with many species and many channels. *J. Phys. Chem. A*, 104(9):1876–1889.

Gillespie, D. T. (1976). A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *J. Comput. Phys.*, 22(4):403–434.

Gillespie, D. T. (1977). Exact stochastic simulation of coupled chemical reactions. *J. Phys. Chem.*, 81(25):2340–2361.

Gillespie, D. T. (1992). A rigorous derivation of the chemical master equation. *Phys. A Stat. Mech. its Appl.*, 188(1):404–425.

Gillespie, D. T. (1996). The multivariate Langevin and Fokker–Planck equations. *Am. J. Phys.*, 64(10):1246–1257.

Gillespie, D. T. (2000). The chemical Langevin equation. *J. Chem. Phys.*, 113(1):297.

Gillespie, D. T. (2001). Approximate accelerated stochastic simulation of chemically reacting systems. *J. Chem. Phys.*, 115(4):1716.

Gillespie, D. T. (2007). Stochastic simulation of chemical kinetics. *Annu. Rev. Phys. Chem.*, 58:35–55.

Gillespie, D. T. (2009). Deterministic limit of stochastic chemical kinetics. *J. Phys. Chem. B*, 113(6):1640–4.

Gomez, D., Marathe, R., Bierbaum, V., and Klumpp, S. (2014). Modeling stochastic gene expression in growing cells. *J. Theor. Biol.*, 348:1–11.

Gómez-Uribe, C. A. and Verghese, G. C. (2007). Mass fluctuation kinetics: Capturing stochastic effects in systems of chemical reactions through coupled mean-variance computations. *J. Chem. Phys.*, 126(2):024109.

González-Cabaleiro, R., Mitchell, A. M., Smith, W., Wipat, A., and Ofiţeru, I. D. (2017). Heterogeneity in Pure Microbial Systems: Experimental Measurements and Modeling. *Front. Microbiol.*, 8:1813.

Grima, R. (2010). An effective rate equation approach to reaction kinetics in small volumes: Theory and application to biochemical reactions in nonequilibrium steady-state conditions. *J. Chem. Phys.*, 133(3):035101.

Grima, R. (2012). A study of the accuracy of moment-closure approximations for stochastic chemical kinetics. *J. Chem. Phys.*, 136(15):154105.

Grima, R., Thomas, P., and Straube, A. V. (2011). How accurate are the nonlinear chemical Fokker-Planck and chemical Langevin equations? *J. Chem. Phys.*, 135(8).

Grimbergen, A. J., Siebring, J., Solopova, A., and Kuipers, O. P. (2015). Microbial bet-hedging: the power of being different. *Curr. Opin. Microbiol.*, 25:67–72.

Grünberger, A., Probst, C., Helfrich, S., Nanda, A., Stute, B., Wiechert, W., von Lieres, E., Nöh, K., Frunzke, J., and Kohlheyer, D. (2015). Spatiotemporal microbial single-cell analysis using a high-throughput microfluidics cultivation platform. *Cytom. Part A*, 87(12):1101–1115.

Guckenheimer, J. and Holmes, P. (1983). *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields*, volume 42 of *Applied Mathematical Sciences*. Springer New York, New York, NY.

Hahl, S. K. and Kremling, A. (2016). A Comparison of Deterministic and Stochastic Modeling Approaches for Biochemical Reaction Systems: On Fixed Points, Means, and Modes. *Front. Genet.*, 7:157.

Haseltine, E. L. and Rawlings, J. B. (2005). Approximate simulation of coupled fast and slow reactions for stochastic chemical kinetics. *J. Chem. Phys.*, 6959(2002):164115–14116.

Hasenauer, J., Wolf, V., Kazeroonian, A., and Theis, F. J. (2014). Method of conditional moments (MCM) for the Chemical Master Equation. *J. Math. Biol.*, 69(3):687–735.

He, F., Murabito, E., and Westerhoff, H. V. (2016). Synthetic biology and regulatory networks: where metabolic systems biology meets control engineering. *J. R. Soc. Interface*, 13(117):20151046.

Hellander, A. and Lötstedt, P. (2007). Hybrid method for the chemical master equation. *J. Comput. Phys.*, pages 1–31.

Hintsche, M. and Klumpp, S. (2013). Dilution and the theoretical description of growth-rate dependent gene expression. *J. Biol. Eng.*, 7(1):22.

Hornung, G. and Barkai, N. (2008). Noise propagation and signaling sensitivity in biological networks: a role for positive feedback. *PLoS Comput. Biol.*, 4(1):e8.

Horowitz, J. M. (2015). Diffusion approximations to the chemical master equation only have a consistent stochastic thermodynamics at chemical equilibrium. *J. Chem. Phys.*, 143(4):044111.

Hortsch, S. K. and Kremling, A. (2018a). Adjusting Noise in the Genetic Toggle Switch through Stochastic Circuit Design. *IFAC-PapersOnLine*, 51(19):68–71.

Hortsch, S. K. and Kremling, A. (2018b). Characterization of noise in multistable genetic circuits reveals ways to modulate heterogeneity. *PLoS One*, 13(3).

Houchmandzadeh, B. and Vallade, M. (2015). Exact results for a noise-induced bistable system. *Phys. Rev. E*, 91(2):022115.

Hsu, C., Jaquet, V., Maleki, F., and Becskei, A. (2016). Contribution of Bistability and Noise to Cell Fate Transitions Determined by Feedback Opening. *J. Mol. Biol.*, 428(20):4115–4128.

Huang, L., Yuan, Z., Liu, P., and Zhou, T. (2014). Feedback-induced counterintuitive correlations of gene expression noise with bursting kinetics. *Phys. Rev. E*, 90(5):052702.

Huang, S. (2009). Non-genetic heterogeneity of cells in development: more than just noise. *Development*, 136(23):3853–62.

Huh, D. and Paulsson, J. (2011). Non-genetic heterogeneity from stochastic partitioning at cell division. *Nat. Genet.*, 43(2):95–100.

Iglesias, P. A. and Ingalls, B. P. (2009). *Control Theory and Systems Biology*. MIT Press.

Igoshin, O. A., Alves, R., and Savageau, M. A. (2008). Hysteretic and graded responses in bacterial two-component signal transduction. *Mol. Microbiol.*, 68(5):1196–215.

Jahnke, T. (2011). On Reduced Models for the Chemical Master Equation. *Multiscale Model. Simul.*, 9(4):1646–1676.

Jaruszewicz, J. and Lipniacki, T. (2013). Toggle switch: noise determines the winning gene. *Phys. Biol.*, 10(3):035007.

Kærn, M., Elston, T. C., Blake, W. J., and Collins, J. J. (2005). Stochasticity in gene expression: from theories to phenotypes. *Nat. Rev. Genet.*, 6(6):451–464.

Karmakar, R. and Bose, I. (2004). Graded and binary responses in stochastic gene expression Stochastic model of transcription factor- regulated gene expression. *Phys. Biol*, 1(4):197–204.

Karmakar, R. and Bose, I. (2007). Positive feedback, stochasticity and genetic competence. *Phys. Biol.*, 4(1):29–37.

Kasper, J. R., Andrews, E. C., and Park, C. (2014). Product inhibition in native-state proteolysis. *PLoS One*, 9(10):e111416.

Kepler, T. B. and Elston, T. C. (2001). Stochasticity in transcriptional regulation: Origins, consequences, and mathematical representations. *Biophys. J.*, 81(6):3116–3136.

Kierzek, A. M., Zaim, J., and Zielenkiewicz, P. (2001). The effect of transcription and translation initiation frequencies on the stochastic fluctuations in prokaryotic gene expression. *J. Biol. Chem.*, 276(11):8165–72.

Kimmel, M. and Axelrod, D. E. (2015). *Branching Processes in Biology*, volume 19 of *Interdisciplinary Applied Mathematics*. Springer New York, New York, NY.

Klumpp, S. and Hwa, T. (2014). Bacterial growth: global effects on gene expression, growth feedback and proteome partition. *Curr. Opin. Biotechnol.*, 28:96–102.

Kompala, D. S., Ramkrishna, D., and Tsao, G. T. (1984). Cybernetic modeling of microbial growth on multiple substrates. *Biotechnol. Bioeng.*, 26(11):1272–1281.

Koutinas, M., Kiparissides, A., Silva-Rocha, R., Lam, M. C., Martins dos Santos, V. A., de Lorenzo, V., Pistikopoulos, E. N., and Mantalaris, A. (2011). Linking genes to microbial growth kinetics - An integrated biochemical systems engineering approach. *Metab. Eng.*, 13(4):401–413.

Kurtz, T. G. (1972). The Relationship between Stochastic and Deterministic Models for Chemical Reactions. *J. Chem. Phys.*, 57(7):2976.

Kuschel, M., Siebler, F., and Takors, R. (2017). Lagrangian Trajectories to Predict the Formation of Population Heterogeneity in Large-Scale Bioreactors. *Bioeng. (Basel, Switzerland)*, 4(2).

Lakatos, E., Ale, A., Kirk, P. D. W., and Stumpf, M. P. H. (2015). Multivariate moment closure techniques for stochastic kinetic models. *J. Chem. Phys.*, 143(9):094107.

Lara, A. R., Galindo, E., Ramírez, O. T., and Palomares, L. A. (2006). Living With Heterogeneities in Bioreactors: Understanding the Effects of Environmental Gradients on Cells. *Mol. Biotechnol.*, 34(3):355–382.

Lehner, B. (2008). Selection to minimise noise in living systems and its implications for the evolution of gene expression. *Mol. Syst. Biol.*, 4:170.

Lemoine, A., Delvigne, F., Bockisch, A., Neubauer, P., and Junne, S. (2017). Tools for the determination of population heterogeneity caused by inhomogeneous cultivation conditions. *J. Biotechnol.*, 251:84–93.

Lencastre Fernandes, R., Nierychlo, M., Lundin, L., Pedersen, A. E., Puentes Tellez, P. E., Dutta, A., Carlquist, M., Bolic, A., Schäpper, D., Brunetti, A. C., Helmark, S., Heins, A. L., Jensen, A. D., Nopens, I., Rottwitt, K., Szita, N., van Elsas, J. D., Nielsen, P. H., Martinussen, J., Sørensen, S. J., Lantz, A. E., and Gernaey, K. V. (2011). Experimental methods and modeling techniques for description of cell population heterogeneity. *Biotechnol. Adv.*, 29(6):575–599.

Levchenko, A. and Nemenman, I. (2014). Cellular noise and information transmission. *Curr. Opin. Biotechnol.*, 28:156–164.

Lin, Y. T. and Galla, T. (2016). Bursting noise in gene expression dynamics: linking microscopic and mesoscopic models. *J. R. Soc. Interface*, 13(114):20150772.

Lu, T., Volfson, D., Tsimring, L., and Hasty, J. (2004). Cellular growth and division in the Gillespie algorithm. *Syst. Biol. (Stevenage).*, 1(1):121–128.

Luo, R., Ye, L., Tao, C., and Wang, K. (2013). Simulation of E. coli Gene Regulation including Overlapping Cell Cycles, Growth, Division, Time Delays and Noise. *PLoS One*, 8(4):e62380.

Maamar, H., Raj, A., and Dubnau, D. (2007). Noise in gene expression determines cell fate in Bacillus subtilis. *Science*, 317(5837):526–9.

Marguerat, S. and Bähler, J. (2012). Coordinating genome expression with cell size. *Trends Genet.*, 28(11):560–565.

Mather, W. H., Hasty, J., Tsimring, L. S., and Williams, R. J. (2013). Translational Cross Talk in Gene Networks. *Biophys. J.*, 104(11):2564–2572.

Mauro, V. P. and Edelman, G. M. (2002). The ribosome filter hypothesis. *Proc. Natl. Acad. Sci. U. S. A.*, 99(19):12031–6.

McAdams, H. H. and Arkin, A. (1997). Stochastic mechanisms in gene expression. *Proc. Natl. Acad. Sci. U. S. A.*, 94(3):814–819.

McQuarrie, D. A. (1967). Stochastic approach to chemical kinetics. *J. Appl. Probab.*, 4(03):413–478.

McSweeney, J. K. and Popovic, L. (2014). Stochastically-induced bistability in chemical reaction systems. *Ann. Appl. Probab.*, 24(3):1226–1268.

Meister, A., Du, C., Li, Y. H., and Wong, W. H. (2014). Modeling stochastic noise in gene regulatory systems. *Quant. Biol.*, 2(1):1–29.

Milo, R., Phillips, R., and Orme, N. (2016). *Cell biology by the numbers*. Taylor & Francis Ltd.

Murphy, K. F., Adams, R. M., Wang, X., Balázsi, G., and Collins, J. J. (2010). Tuning and controlling gene expression noise in synthetic gene networks. *Nucleic Acids Res.*, 38(8):2712–26.

Murray, F. J. F. J. and Miller, K. S. (1976). *Existence theorems for ordinary differential equations*. R.E. Krieger Pub. Co.

Nicolis, G., Allen, P., and Van Nypelseer, A. (1974). Some Remarks on the Theory of Fluctua-

tions around Nonequilibrium States. *Prog. Theor. Phys.*, 52(5):1481–1497.

Novick, A. and Weiner, M. (1957). Enzyme Induction as an all-or-none-phenomenon. *Proc. Natl. Acad. Sci.*, 43(7):553–566.

Ochab-Marcinek, A. and Tabaka, M. (2010). Bimodal gene expression in noncooperative regulatory systems. *Proc. Natl. Acad. Sci. U. S. A.*, 107(51):22096–101.

Oyarzún, D. A., Lugagne, J. B., and Stan, G. B. V. (2015). Noise propagation in synthetic gene circuits for metabolic control. *ACS Synth. Biol.*, 4(2):116–125.

Ozbudak, E. M., Thattai, M., Kurtser, I., Grossman, A. D., and van Oudenaarden, A. (2002). Regulation of noise in the expression of a single gene. *Nat. Genet.*, 31(1):69–73.

Patrascioiu, A. (1987). The Ergodic Hypothesis. *Los Alamos Sci.*, 1(15):263–279.

Paulsson, J. (2004). Summing up the noise in gene networks. *Nature*, 427(6973):415–418.

Paulsson, J. (2005). Models of stochastic gene expression. *Phys. Life Rev.*, 2(2):157–175.

Paulsson, J. and Ehrenberg, M. (2000). Random signal fluctuations can reduce random fluctuations in regulated components of chemical regulatory networks. *Phys. Rev. Lett.*, 84(23):5447–5450.

Pedraza, J. M. and van Oudenaarden, A. (2005). Noise propagation in gene networks. *Science*, 307(5717):1965–9.

Potvin-Trottier, L., Lord, N. D., Vinnicombe, G., and Paulsson, J. (2016). Synchronous long-term oscillations in a synthetic gene circuit. *Nature*, 538(7626):514–517.

Qian, H., Shi, P.-Z., and Xing, J. (2009). Stochastic bifurcation, slow fluctuations, and bistability as an origin of biochemical complexity. *Phys. Chem. Chem. Phys.*, 11(24):4861–70.

Rao, C. V. and Arkin, A. P. (2003). Stochastic chemical kinetics and the quasi-steady-state assumption: Application to the Gillespie algorithm. *J. Chem. Phys.*, 118(11):4999–5010.

Rauhut, R. and Klug, G. (1999). mRNA degradation in bacteria. *FEMS Microbiol. Rev.*, 23(3):353–370.

Reck, M., Tomasch, J., and Wagner-Döbler, I. (2015). The Alternative Sigma Factor SigX Controls Bacteriocin Synthesis and Competence, the Two Quorum Sensing Regulated Traits in Streptococcus mutans. *PLoS Genet.*, 11(7):e1005353.

Roberts, E., Magis, A., Ortiz, J. O., Baumeister, W., and Luthey-Schulten, Z. (2011). Noise Contributions in an Inducible Genetic Switch: A Whole-Cell Simulation Study. *PLoS Comput. Biol.*, 7(3):e1002010.

Rodnina, M. V. (2016). The ribosome in action: Tuning of translational efficiency and protein folding. *Protein Sci.*, 25(8):1390–1406.

Roquet, N. and Lu, T. K. (2014). Digital and analog gene circuits for biotechnology. *Biotechnol. J.*, 9(5):597–608.

Samoilov, M., Plyasunov, S., and Arkin, A. P. (2005). Stochastic amplification and signaling in enzymatic futile cycles through noise-induced bistability with oscillations. *Proc. Natl. Acad. Sci.*, 102(7):2310–2315.

Savageau, M. A. (1999). Design of gene circuitry by natural selection: analysis of the lactose catabolic system in Escherichia coli. *Biochem. Soc. Trans.*, 27(2):264–70.

Schnoerr, D., Sanguinetti, G., and Grima, R. (2015). Comparison of different moment-closure

approximations for stochastic chemical kinetics. *J. Chem. Phys.*, 143(143):185101.

Scott, M. (2006). Tutorial : Genetic circuits and noise. *Lect. Notes*, (July):1–19.

Shahrezaei, V. and Swain, P. S. (2008). Analytical distributions for stochastic gene expression. *Proc. Natl. Acad. Sci. U. S. A.*, 105(45):17256–61.

Shu, C.-C., Chatterjee, A., Dunny, G., Hu, W.-S., and Ramkrishna, D. (2011). Bistability versus bimodal distributions in gene regulatory processes from population balance. *PLoS Comput. Biol.*, 7(8):e1002140.

Simpson, M. L., Cox, C. D., and Sayler, G. S. (2003). Frequency domain analysis of noise in autoregulated gene circuits. *Proc. Natl. Acad. Sci.*, 100(8):4551–4556.

Singh, A. (2011). Negative feedback through mRNA provides the best control of gene-expression noise. *IEEE Trans. Nanobioscience*, 10(3):194–200.

Son, M., Ahn, S.-J., Guo, Q., Burne, R. A., and Hagen, S. J. (2012). Microfluidic study of competence regulation in Streptococcus mutans: environmental inputs modulate bimodal and unimodal expression of comX. *Mol. Microbiol.*, 86(2):258–72.

Süel, G. M., Garcia-Ojalvo, J., Liberman, L. M., and Elowitz, M. B. (2006). An excitable gene regulatory circuit induces transient cellular differentiation. *Nature*, 440(7083):545–550.

Swain, P. S., Elowitz, M. B., and Siggia, E. D. (2002). Intrinsic and extrinsic contributions to stochasticity in gene expression. *Proc. Natl. Acad. Sci. U. S. A.*, 99(20):12795–800.

Taheri-Araghi, S., Bradde, S., Sauls, J. T., Hill, N. S., Levin, P. A., Paulsson, J., Vergassola, M., and Jun, S. (2015). Cell-size control and homeostasis in bacteria. *Curr. Biol.*, 25(3):385–391.

Tang, Y., Yuan, R., Wang, G., Zhu, X., and Ao, P. (2017). Potential landscape of high dimensional nonlinear stochastic dynamics with large noise. *Sci. Rep.*, 7(1):15762.

Thattai, M. and van Oudenaarden, A. (2001). Intrinsic noise in gene regulatory networks. *Proc. Natl. Acad. Sci. U. S. A.*, 98(15):8614–9.

Thomas, P., Matuschek, H., and Grima, R. (2012). Computation of biochemical pathway fluctuations beyond the linear noise approximation using iNA. In *Proc. - 2012 IEEE Int. Conf. Bioinforma. Biomed. BIBM 2012*, pages 192–196.

Thomas, P., Matuschek, H., and Grima, R. (2013). How reliable is the linear noise approximation of gene regulatory networks? *BMC Genomics*, 14:S5.

Thomas, P., Popović, N., and Grima, R. (2014). Phenotypic switching in gene regulatory networks. *Proc. Natl. Acad. Sci. U. S. A.*, 111(19):6994–9.

To, T.-L. and Maheshri, N. (2010). Noise can induce bimodality in positive transcriptional feedback loops without bistability. *Science*, 327(5969):1142–5.

Tolle, D. P. and Le Novere, N. (2006). Particle-based stochastic simulation in systems biology. *Curr. Bioinform.*, 1(3):315–320.

Tomioka, R., Kimura, H., J Kobayashi, T., and Aihara, K. (2004). Multivariate analysis of noise in genetic regulatory networks. *J. Theor. Biol.*, 229(4):501–21.

Tsimring, L. S. (2014). Noise in biology. *Rep. Prog. Phys.*, 77(2):026601.

van Kampen, N. G. (2007). *Stochastic processes in physics and chemistry*. North-Holland, 3rd edition.

Veening, J.-W., Smits, W. K., and Kuipers, O. P. (2008). Bistability, Epigenetics, and Bet-

Hedging in Bacteria. *Annu. Rev. Microbiol.*, 62(1):193–210.

Waddington, C. H. (1957). *The strategy of the genes. A discussion of some aspects of theoretical biology.* London: George Allen and Unwin, Ltd.

Walczak, A. M., Mugler, A., and Wiggins, C. H. (2012). Analytic methods for modeling stochastic regulatory networks. *Methods Mol. Biol.*, 880:273–322.

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

Wallden, M., Fange, D., Lundius, E. G., Baltekin, Ö., and Elf, J. (2016). The Synchronization of Replication and Division Cycles in Individual E. coli Cells. *Cell*, 166(3):729–739.

Wang, J. D. and Levin, P. A. (2009). Metabolism, cell growth and the bacterial cell cycle. *Nat. Rev. Microbiol.*, 7(11):822–827.

Wang, P., Robert, L., Pelletier, J., Dang, W. L., Taddei, F., Wright, A., and Jun, S. (2010). Robust growth of Escherichia coli. *Curr. Biol.*, 20(12):1099–103.

Willis, L. and Huang, K. C. (2017). Sizing up the bacterial cell cycle. *Nat. Rev. Microbiol.*, 15(10):606–620.

Winkel, M., Salman-Carvalho, V., Woyke, T., Richter, M., Schulz-Vogt, H. N., Flood, B. E., Bailey, J. V., and Mußmann, M. (2016). Single-cell Sequencing of Thiomargarita Reveals Genomic Flexibility for Adaptation to Dynamic Redox Conditions. *Front. Microbiol.*, 7:964.

Zheng, X.-D., Yang, X.-Q., and Tao, Y. (2011). Bistability, probability transition rate and first-passage time in an autoactivating positive-feedback loop. *PLoS One*, 6(3):e17104.

A. Calculations for Chapter 3

A.1. Generating functions for the solution of the stationary CME

Generating functions will be used to find a closed-form expression of p_s, which is defined through the CME (3.5). The CME is only considered in the stationary state ($\dot{p}_s = 0$), where it fulfills:

$$\sum_{b=0}^{\infty} \left(\mathrm{Geo}_\alpha(b) \cdot \frac{1}{\alpha} \cdot H(s-b) \cdot p_{s-b} \right) - \frac{1}{\alpha} \cdot H(s)\, p_s = -\left((s+1)\, p_{s+1} - s\, p_s \right)$$

$$\Leftrightarrow \quad \left((\mathrm{Geo}_\alpha - \delta_\cdot(\{0\})) * H\, p \right)(s) = -\alpha \left((s+1)\, p_{s+1} - s\, p_s \right). \tag{A.1}$$

Here, $H(s) = 0 \ \forall\, s < 0$, and $*$ denotes a discrete convolution. In the following, $\mathcal{Z}_f$ will denote the generating function of the array $a_{\mathbf{s}} := f(\mathbf{s})$, $\mathbf{s} \in \mathbb{N}_0$, which is specified by the real-valued function f. The generating function of the LHS of the equation is given by:

$$\begin{aligned}
&\mathcal{Z}_{[\mathrm{Geo}_\alpha - \delta_\cdot(\{0\})] * H\, p}(z) = \left(\mathcal{Z}_{\mathrm{Geo}_\alpha}(z) - \mathcal{Z}_{\delta_\cdot(\{0\})}(z) \right) \cdot \mathcal{Z}_{H\, p}(z) = \left(\mathcal{Z}_{\mathrm{Geo}_\alpha}(z) - 1 \right) \cdot \left(\mathcal{Z}_H * \mathcal{Z}_p \right)(z) \\
&= \left(\sum_{s=0}^{\infty} \frac{1}{1+\alpha} \left(\frac{\alpha}{1+\alpha} \right)^s z^s - 1 \right) \cdot \left(\mathcal{Z}_H * \mathcal{Z}_p \right)(z) \\
&= \left(\frac{1}{1+\alpha} \sum_{s=0}^{\infty} \left(\frac{\alpha\, z}{1+\alpha} \right)^s - 1 \right) \cdot \left(\mathcal{Z}_H * \mathcal{Z}_p \right)(z) \\
&= \left(\frac{1}{1+\alpha} \cdot \frac{1}{1 - \frac{\alpha\, z}{1+\alpha}} - 1 \right) \cdot \left(\mathcal{Z}_H * \mathcal{Z}_p \right)(z) = -\frac{\alpha\,(1-z)}{\alpha\,(1-z)+1} \cdot \left(\mathcal{Z}_H * \mathcal{Z}_p \right)(z).
\end{aligned} \tag{A.2}$$

The generating function of the RHS of Eq (A.1) divided by $-\alpha$ reads:

$$\sum_{s=0}^{\infty} z^s \left((s+1)\, p_{s+1} - s\, p_s \right) = \frac{1-z}{z} \sum_{s=0}^{\infty} z^s\, s\, p_s = (1-z)\, \partial_z\, \mathcal{Z}_p(z). \tag{A.3}$$

All in all, the Z-transform of the whole steady state equation reads

$$-\alpha\,(1 - z)\, \partial_z \mathcal{Z}_p(z) = -\frac{\alpha\,(1-z)}{\alpha\,(1-z)+1} \cdot \left(\mathcal{Z}_H * \mathcal{Z}_p \right)(z) \tag{A.4}$$

or, after some simplification:

$$(\alpha\,(1-z) + 1)\, \partial_z \mathcal{Z}_p(z) = \left(\mathcal{Z}_H * \mathcal{Z}_p \right)(z). \tag{A.5}$$

One can easily show that $\partial_z \mathcal{Z}_p(z) = \mathcal{Z}_{(s+1)\, p_{s+1}}$. Backtransformation thus yields

$$\alpha\,(s+1)\, p_{s+1} - \alpha\, s\, p_s + (s+1)\, p_{s+1} = H(s)\, p_s, \tag{A.6}$$

from which follows the simplified recursive steady state condition in Eq (3.6).

A.2. Mean and variance of the stochastic distribution

The calculation of the dynamics of the mean value and variance of the protein copy number S is performed as explained in Section 2.1.2 (cf. Eq (2.12) for the general calculation):

$$\begin{aligned}
\dot{\mathbb{E}}[S] &= \sum_{s=0}^{\infty} s\,\dot{p}_s \\
&= \sum_{s=0}^{\infty}\left[\sum_{b=0}^{\infty}\left(\mathrm{Geo}_\alpha(b)\,\frac{1}{\alpha}\,H(s-b)\,s\,p_{s-b}\right) - \frac{1}{\alpha}\,H(s)\,s\,p_s + s\,(s+1)\,p_{s+1} - s^2\,p_s\right] \\
&= \sum_{s=0}^{\infty}\left[\sum_{b=0}^{\infty}\left(\mathrm{Geo}_\alpha(b)\,\frac{1}{\alpha}\,H(s)\,(s+b)\,p_s\right) - \frac{1}{\alpha}\,H(s)\,s\,p_s + (s-1)\,s\,p_s - s^2\,p_s\right] \\
&= \sum_{s=0}^{\infty}\left[\left(\sum_{b=0}^{\infty} b\,\mathrm{Geo}_\alpha(b)\right)\frac{1}{\alpha}\,H(s)\,p_s + \left(\sum_{b=0}^{\infty}\mathrm{Geo}_\alpha(b)\right)\frac{1}{\alpha}\,H(s)\,s\,p_s - \frac{1}{\alpha}\,H(s)\,s\,p_s - s\,p_s\right] \\
&= \sum_{s=0}^{\infty}\left[\alpha\,\frac{1}{\alpha}\,H(s)\,p_s - s\,p_s\right] = \mathbb{E}[H(S)] - \mathbb{E}[S].
\end{aligned} \tag{A.7}$$

The third equality was obtained through an index shift in s (note that $p_s = 0$ for negative s). In the fifth equality, the fact that $\mathrm{Geo}_\alpha(b)$ is a probability distribution with mean α was used. In a similar manner, the variance of S is computed. The ODE of the second moment $\mathbb{E}[S^2]$ is given by:

$$\begin{aligned}
\dot{\mathbb{E}}[S^2] &= \sum_{s=0}^{\infty} s^2\,\dot{p}_s \\
&= \sum_{s=0}^{\infty}\left[\sum_{b=0}^{\infty}\left(\mathrm{Geo}_\alpha(b)\,\frac{1}{\alpha}\,H(s-b)\,s^2\,p_{s-b}\right) - \frac{1}{\alpha}\,H(s)\,s^2\,p_s + s^2\,(s+1)\,p_{s+1} - s^3\,p_s\right] \\
&= \sum_{s=0}^{\infty}\left[\sum_{b=0}^{\infty}\left(\mathrm{Geo}_\alpha(b)\,\frac{1}{\alpha}\,H(s)\,(s+b)^2\,p_s\right) - \frac{1}{\alpha}\,H(s)\,s^2\,p_s + (s-1)^2\,s\,p_s - s^3\,p_s\right] \\
&= \sum_{s=0}^{\infty}\left[\left(\sum_{b=0}^{\infty} b^2\,\mathrm{Geo}_\alpha(b)\right)\frac{1}{\alpha}\,H(s)\,p_s + 2\left(\sum_{b=0}^{\infty} b\,\mathrm{Geo}_\alpha(b)\right)\frac{1}{\alpha}\,H(s)\,s\,p_s + \right. \\
&\quad \left. + \left(\sum_{b=0}^{\infty}\mathrm{Geo}_\alpha(b)\right)\frac{1}{\alpha}\,H(s)\,s^2\,p_s - \frac{1}{\alpha}\,H(s)\,s^2\,p_s - 2\,s^2\,p_s + s\,p_s\right] \\
&= \sum_{s=0}^{\infty}\left[(2\,\alpha+1)\,H(s)\,p_s + 2\,H(s)\,s\,p_s - 2\,s^2\,p_s + s\,p_s\right] \\
&= (2\,\alpha+1)\,\mathbb{E}[H(S)] + 2\,\mathbb{E}[S\,H(S)] - 2\,\mathbb{E}[S^2] + \mathbb{E}[S]
\end{aligned} \tag{A.8}$$

The ordinary differential equation of the variance $\mathrm{Var}(S) = \mathbb{E}[S^2] - \mathbb{E}[S]^2$ reads:

$$\begin{aligned}
\dot{\mathrm{Var}}(S) &= \dot{\mathbb{E}}[S^2] - 2\,\mathbb{E}[S]\,\dot{\mathbb{E}}[S] \\
&= 2\,\mathrm{Cov}(S, H(S)) + (\mathbb{E}[H(S)] + \mathbb{E}[S]) + 2\,\alpha\,\mathbb{E}[H(S)] - 2\,\mathrm{Var}(S).
\end{aligned} \tag{A.9}$$

A.3. Comparing the robustness of expression states in different bimodal systems

In Section 3.4.2, three different pairs of bimodal systems are compared. Here, we perform this comparison mathematically for two of the bimodal pairs (corresponding to panels (B) and (C) in Fig 3.4).

A.3.1. Comparison of cooperative and non-cooperative systems

Let

$$\begin{aligned} H_1(s) &= a + v_1 \frac{s^h}{s^h + K_1}, \quad h \geq 2, \quad \text{and} \\ H_2(s) &= a + v_2 \frac{s}{s + K_2} \end{aligned} \tag{A.10}$$

be two functions expressing a saturated feedback mechanism, the first one being cooperative, the second non-cooperative. Moreover, let $(.+\alpha+1)$ and H_1 possess two positive intersection points s_{min} and s_{max}, which are identical to the only two positive intersection points of $(.+\alpha+1)$ and H_2. In this section, we show that

$$|H_1(s) - (s+\alpha+1)| \;\geq\; |H_2(s) - (s+\alpha+1)| \tag{A.11}$$

holds for all $s \geq 0$.

The proof is given as follows: Since H_2 is concave, for the above conditions to be met, it can easily be derived that

- $H_2(s) < s+\alpha+1 \quad \forall s \in [0, s_{min})$
- $H_2(s) > s+\alpha+1 \quad \forall s \in (s_{min}, s_{max})$
- $H_2(s) < s+\alpha+1 \quad \forall s \in (s_{max}, \infty)$.

Therefore, it is sufficient to prove that

- $H_1(s) < H_2(s) \quad \forall s \in (0, s_{min})$
- $H_1(s) > H_2(s) \quad \forall s \in (s_{min}, s_{max})$
- $H_1(s) < H_2(s) \quad \forall s \in (s_{max}, \infty)$.

As a first step, we prove that $v_1 < v_2$. Solving both conditions $H_1(s_{min}) = H_2(s_{min})$ and $H_1(s_{max}) = H_2(s_{max})$ for K_1 and equating the resulting expressions leads to

$$\frac{v_1}{v_2} \;=\; \frac{s_{max}^h \;-\; s_{min}^h}{s_{max}^h \;-\; s_{min}^h \;+\; K_2\,(s_{max}^{h-1} \;-\; s_{min}^{h-1})}. \tag{A.12}$$

This fraction is smaller than 1.

First, this fact is used in order to prove that 0, s_{min} and s_{max} are the only intersection points between H_1 and H_2. Reformulating the equation $H_1(s) = H_2(s)$ leads to the following conditions for nonzero solutions:

$$g(s) := s^{h-1}\,[(v_2 - v_1)\,s \;-\; v_1\,K_2] \;+\; v_2\,K_1 \;=\; 0 \tag{A.13}$$

The extreme values of g fulfill the condition

$$g'(s) := s^{h-2}\left[h\,(v_2 - v_1)\,s\;-\;(h-1)\,v_1\,K_2\right] \;=\; 0 \tag{A.14}$$

which has at most one positive solution. From Rolle's theorem, one can conclude that g has at most two positive roots. Hence, the only intersection points in positive space correspond to s_{min} and s_{max}.

Due to the fact that for $s \to \infty$, H_1 and H_2 converge to $a + v_1$ and $a + v_2$, respectively, the relation $v_2 > v_1$ and the fact that $H_1(s) \neq H_2(s)$ for all $s > s_{max}$ shows that one the one hand, $H_2(s) > H_1(s)$ necessarily holds for all $s \in (s_{max}, \infty)$.

On the other hand, as $H_1'(0) = 0 < \frac{v_2}{K_2} = H_2'(0)$, the inequality $H_1(s) < H_2(s)$ is fulfilled in the interval $(0, s_{min})$.

Finally, it remains to prove that $H_2 < H_1$ in (s_{min}, s_{max}). It is sufficient to show that $H_1'(s_{min}) \neq H_2'(s_{min})$ or $H_1'(s_{max}) \neq H_2'(s_{max})$, so that H_1 and H_2 truly intersect. Suppose that

$$v_1 \frac{K_1\,h\,s_{min}^{h-1}}{(s_{min}^h + K_1)^2} = H_1'(s_{min}) = H_2'(s_{min}) = v_2 \frac{K_2}{(s_{min} + K_2)^2}. \tag{A.15}$$

Then, by exploiting the relation $H_1(s_{min}) = H_2(s_{min})$, this equation can be reformulated as

$$\frac{K_1\,v_2\,h}{K_2\,v_1} = s_{min}^{h-1}. \tag{A.16}$$

In an analogous manner, the assumption $H_1'(s_{max}) = H_2'(s_{max})$ leads to the contradictory equation

$$\frac{K_1\,v_2\,h}{K_2\,v_1} = s_{max}^{h-1}, \tag{A.17}$$

which is why the original assumption was wrong and $H_1'(s_{min}) \neq H_2'(s_{min})$ or $H_1'(s_{max}) \neq H_2'(s_{max})$ holds true. □

A.3.2. Comparison of non-cooperative systems with differing burst sizes

Let

$$\begin{aligned} H_1(s) &= a + v_1 \frac{s}{s + K_1} \quad \text{and} \\ H_2(s) &= a + v_2 \frac{s}{s + K_2} \end{aligned} \tag{A.18}$$

be two functions describing Michaelis-Menten type kinetics. Let furthermore $\alpha_1 > \alpha_2$ and let the equations $H_1(s) = s + \alpha_1 + 1$ and $H_2(s) = s + \alpha_2 + 1$ have an identical set of solutions $\{s_{min}, s_{max}\}$. Our aim is to show that

$$\left|H_1(s)\;-\;(s + \alpha_1 + 1)\right| \;>\; \left|H_2(s) - (s + \alpha_2 + 1)\right|, \qquad s \notin \{s_{min}, s_{max}\} \tag{A.19}$$

holds true, which counteracts the effect of the difference in the burst sizes with regard to the robustness of bimodality.

First, note that for $l \in \{1, 2\}$,

- $H_l(s) < s + \alpha_l + 1 \quad \forall s \in [0, s_{min})$
- $H_l(s) > s + \alpha_l + 1 \quad \forall s \in (s_{min}, s_{max})$
- $H_l(s) < s + \alpha_l + 1 \quad \forall s \in (s_{max}, \infty)$.

Therefore, it suffices to show that

- $H_2(s) - s + \alpha_2 + 1 > H_1(s) - s + \alpha_1 + 1 \quad \forall s \in (0, s_{min})$
- $H_2(s) - s + \alpha_2 + 1 < H_1(s) - s + \alpha_1 + 1 \quad \forall s \in (s_{min}, s_{max})$
- $H_2(s) - s + \alpha_2 + 1 > H_1(s) - s + \alpha_1 + 1 \quad \forall s \in (s_{max}, \infty)$.

K_1, K_2, v_1, and v_2 can be determined analytically, using the relations above. The results are:

$$\begin{aligned} K_1 &= \frac{s_{min}\, s_{max}}{\alpha_1 + 1 - a}, \\ K_2 &= \frac{s_{min}\, s_{max}}{\alpha_2 + 1 - a}, \\ v_1 &= \frac{(s_{min} + \alpha_1 + 1 - a)\,(s_{max} + \alpha_1 + 1 - a)}{\alpha_1 + 1 - a}, \\ v_2 &= \frac{(s_{min} + \alpha_2 + 1 - a)\,(s_{max} + \alpha_2 + 1 - a)}{\alpha_2 + 1 - a} \end{aligned} \tag{A.20}$$

The evaluation of the first derivative of H_1 and H_2 at the intersection points s_{min} and s_{max} reads

$$\begin{aligned} H_l'(s_{min}) &= \frac{s_{max}\, s_{min} \;+\; s_{max}\,(\alpha_l + 1 - a)}{s_{max}\, s_{min} \;+\; s_{min}\,(\alpha_l + 1 - a)}, \\ H_l'(s_{max}) &= \frac{s_{max}\, s_{min} \;+\; s_{min}\,(\alpha_l + 1 - a)}{s_{max}\, s_{min} \;+\; s_{max}\,(\alpha_l + 1 - a)}, \end{aligned} \tag{A.21}$$

for $l \in \{1, 2\}$. The dependence of the expressions on α_l shows that

$$\begin{aligned} H_1'(s_{min}) &> H_2'(s_{min}) \quad \text{and} \\ H_1'(s_{max}) &< H_2'(s_{max}). \end{aligned} \tag{A.22}$$

The equation $H_2'(s) - H_1'(s) = 0$ has only one positive solution (analytical calculation not shown), so that s_{min} and s_{max} are the only intersection points of $H_1(s) - s - \alpha_1 - 1$ and $H_2(s) - s - \alpha_2 - 1$ according to Rolle's theorem. Hence, one can conclude from the derivatives at the intersection points that the above conditions are fulfilled. □

B. Calculations for Chapter 4

B.1. Means, variances, and covariances of the stochastic distribution

The mean value of M using the CME (4.2) is calculated as follows (cf. the theory in Section 2.1.2):

$$\begin{aligned}\dot{\mathbb{E}}[M] =& \sum_{s=0}^{\infty}\sum_{m=0}^{\infty}\Big(\left[m\,F(s)\,p_{m-1,s} - m\,F(s)\,p_{m,s} + m\,(m+1)\,p_{m+1,s} - m^2\,p_{m,s}\right]\cdot\frac{1}{\nu}\\ &+ \left[m\,G(m)\,p_{m,s-1} - m\,G(m)\,p_{m,s} + m\,(s+1)\,p_{m,s+1} - m\,s\,p_{m,s}\right]\Big)\\ =& \sum_{s=0}^{\infty}\sum_{m=0}^{\infty}\Big(\big(F(s) - m\big)\,p_{m,s}\Big)\cdot\frac{1}{\nu}\\ =& \Big(\mathbb{E}[F(S)] - \mathbb{E}[M]\Big)\cdot\frac{1}{\nu}.\end{aligned}$$

In an analogue manner,

$$\begin{aligned}\dot{\mathrm{Var}}(M) &= \Big(\mathbb{E}[F(S)] + \mathbb{E}[M] - 2\,\mathrm{Var}(M) + 2\,\mathrm{Cov}(M,F(S))\Big)\cdot\frac{1}{\nu}\\ \dot{\mathbb{E}}[S] &= \mathbb{E}[G(M)] - \mathbb{E}[S]\\ \dot{\mathrm{Var}}(S) &= \mathbb{E}[G(M)] + \mathbb{E}[S] - 2\,\mathrm{Var}(S) + 2\,\mathrm{Cov}(S,G(M))\\ \dot{\mathrm{Cov}}(M,S) &= \frac{1}{\nu}\cdot\mathrm{Cov}(S,F(S)) + \mathrm{Cov}(M,G(M)) - \frac{1+\nu}{\nu}\cdot\mathrm{Cov}(M,S)\end{aligned}\tag{B.1}$$

For calculating stationary distributions, the above ODEs are all set to 0 and the obtained system of equations is solved, which leads to the implicit relations:

$$\begin{aligned}\mathbb{E}[M] &= \mathbb{E}[F(S)]\\ \mathrm{Var}(M) &= \mathbb{E}[M] + \mathrm{Cov}(M,F(S))\\ \mathbb{E}[S] &= \mathbb{E}[G(M)]\\ \mathrm{Var}(S) &= \mathbb{E}[S] + \mathrm{Cov}(S,G(M))\\ \mathrm{Cov}(M,S) &= \frac{\mathrm{Cov}(S,F(S)) + \nu\cdot\mathrm{Cov}(M,G(M))}{1+\nu}.\end{aligned}\tag{B.2}$$

B.2. Generating functions for the calculation of the stationary marginal mRNA distribution

The stationary marginal mRNA distribution is given in Eq (4.6). Transformation of this equation using generating functions yields:

$$\begin{aligned} 0 &= \sum_{s=0}^{\infty} \left[F(s) \sum_{n=0}^{\infty} (z-1)\, z^n\, p_{n,s}^* \right] - (z-1)\, \frac{\partial}{\partial z} \left(\sum_{n=0}^{\infty} z^n\, p_n^{M*} \right), \\ \Leftrightarrow \quad 0 &= \sum_{s=0}^{\infty} \left[F(s) \sum_{n=0}^{\infty} z^n\, p_{n,s}^* \right] - \frac{\partial}{\partial z} \left(\sum_{n=0}^{\infty} z^n\, p_n^{M*} \right) \end{aligned} \tag{B.3}$$

Backtransformation leads to the recursive version of the solution given in Eq (4.7).

B.3. Application of the classical and hybrid linear noise approximation to the full reaction system

Here, the LNA is applied to the full reaction system (4.1). The following formulae are valid for both the classical and the hybrid LNA; the difference lies in the values of the deterministic stable fixed points m^* and s^* (and therefore in the stationary protein-to-mRNA ratio r^*), as well as in those of the local derivatives f^* and g^*.

In order to formulate the vectors and matrices required for the Lyapunov equation (2.29), with which the local variance-covariance-matrix can be estimated, we first define the stoichimetric matrix and the vector of stationary propensities:

$$\mathbf{A} = \begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{pmatrix} \tag{B.4}$$

$$\mathbf{w}(m^*, s^*) = \begin{pmatrix} \frac{1}{\nu}\, m^* & \frac{1}{\nu}\, m^* & s^* & s^* \end{pmatrix}^\top \tag{B.5}$$

The stationary Jacobian matrix $\mathbf{J}^* = (v_{i,l}^*)_{\substack{i=1,\dots,M,\\ l=1,\dots,M}}$ has already been given in Eq (4.14). The diffusion matrix $\mathbf{D}^* = \mathbf{D}(m^*, s^*)$ reads:

$$\mathbf{D}^* = \mathbf{A} \cdot \mathrm{diag}(\mathbf{w}(m^*, s^*)) \cdot \mathbf{A}^\top = \begin{pmatrix} 2\,\frac{1}{\nu}\, m^* & 0 \\ 0 & 2\, s^* \end{pmatrix} \tag{B.6}$$

Since $\mathbf{D}^* = (d_{i,l}^*)_{\substack{i=1,\dots,M,\\ l=1,\dots,M}}$ is a diagonal matrix and $M = 2$, the method proposed in [Tomioka et al., 2004] can be applied which gives a general, direct formulation of Σ^*:

$$
\begin{aligned}
\mathbf{\Sigma}^* = & -\frac{1}{2}\frac{d^*_{1,1}}{v^*_{1,1}+v^*_{2,2}}\left(\begin{pmatrix}1&0\\0&0\end{pmatrix}+\frac{1}{v^*_{1,1}\,v^*_{2,2}-v^*_{1,2}\,v^*_{2,1}}\begin{pmatrix}(v^*_{2,2})^2 & -v^*_{2,1}\,v^*_{2,2}\\ -v^*_{2,1}\,v^*_{2,2} & (v^*_{2,1})^2\end{pmatrix}\right)\\
& -\frac{1}{2}\frac{d^*_{2,2}}{v^*_{1,1}+v^*_{2,2}}\left(\begin{pmatrix}0&0\\0&1\end{pmatrix}+\frac{1}{v^*_{1,1}\,v^*_{2,2}-v^*_{1,2}\,v^*_{2,1}}\begin{pmatrix}(v^*_{1,2})^2 & -v^*_{1,2}\,v^*_{1,1}\\ -v^*_{1,2}\,v^*_{1,1} & (v^*_{1,1})^2\end{pmatrix}\right)\\
= & \; m^*\frac{1}{1+\nu}\left(\begin{pmatrix}1&0\\0&0\end{pmatrix}+\frac{\nu}{1-f^*\,g^*}\begin{pmatrix}1 & g^*\\ g^* & (g^*)^2\end{pmatrix}\right)\\
& + s^*\frac{\nu}{1+\nu}\left(\begin{pmatrix}0&0\\0&1\end{pmatrix}+\frac{1}{\nu\,(1-f^*\,g^*)}\begin{pmatrix}(f^*)^2 & f^*\\ f^* & 1\end{pmatrix}\right).
\end{aligned}
\tag{B.7}
$$

A reformulation of $\mathbf{\Sigma}^*_{1,1}$ and $\mathbf{\Sigma}^*_{2,2}$ yields the following approximate expressions for $\mathrm{Var}^*(M)$ and $\mathrm{Var}^*(S)$:

$$
\begin{aligned}
\mathrm{Var}^*(M) &\approx m^*\left[1+\frac{f^*}{(g^*)^{-1}-f^*}\left(\frac{\nu}{1+\nu}+\frac{1}{1+\nu}\cdot\, r^*\,\frac{f^*}{g^*}\right)\right]\\
\mathrm{Var}^*(S) &\approx s^*\left[1+\frac{1}{(g^*)^{-1}-f^*}\left(\frac{f^*}{1+\nu}+\frac{\nu}{1+\nu}\cdot\frac{g^*}{r^*}\right)\right].
\end{aligned}
\tag{B.8}
$$

B.4. Application of the hybrid linear noise approximation to the reduced system

The mean value of the CME of the reduced reaction scheme (4.32) obeys (with $F(s) = 0$ if $s < 0$):

$$
\begin{aligned}
\dot{\mathbb{E}}[S] &= \sum_{s=0}^{\infty}\left[\sum_{b=0}^{\infty}\left(\frac{1}{\nu}\cdot s\;F(s-b)\;\mathrm{Geo}_{\alpha(F(s-b))}(b)\;p_{s-b}\right)-\frac{1}{\nu}\cdot s\;F(s)\;p_s+s\,(s+1)\,p_{s+1}-s^2\,p_s\right]\\
&= \sum_{s=0}^{\infty}\sum_{b=0}^{\infty}\left(\frac{1}{\nu}\cdot b\;F(s)\,\mathrm{Geo}_{\alpha(F(s))}(b)\;p_s\right)-\sum_{s=0}^{\infty}s\,p_s=\sum_{s=0}^{\infty}\left(\frac{1}{\nu}\cdot F(s)\;\alpha(F(s))\;p_s\right)-\mathbb{E}[S]\\
&= \sum_{s=0}^{\infty}\bar{G}(F(s))\;p_s-\mathbb{E}[S]=\mathbb{E}[\bar{G}(F(S))]-\mathbb{E}[S].
\end{aligned}
\tag{B.9}
$$

For the formulation of the hybrid deterministic model, the small noise assumption is made for the protein level again, so that the deterministic formulation

$$
\dot{c}_s = \bar{\mathcal{G}}(\mathcal{F}(c_s))-c_s \tag{B.10}
$$

is obtained (it also holds for constant F). Its fixed points c^*_s correspond to those of the full hybrid deterministic model. In the subsequent calculations for the hLNA, linearization is thus performed around the same protein level.

Next, the vectors and matrices that are required for LNA are calculated. Since bursts of any size can occur, the stochastic matrix $\mathbf{A}$ has an infinite number of columns.

$$\mathbf{A} = \begin{pmatrix} -1 & 0 & 1 & 2 & 3 & \dots \end{pmatrix}$$

$$\mathbf{w}(s^*) = \begin{pmatrix} s^* & \frac{1}{\nu} \cdot \mathrm{Geo}_{r^*\nu}(0) \cdot F(s^*) & \frac{1}{\nu} \cdot \mathrm{Geo}_{r^*\nu}(1) \cdot F(s^*) & \frac{1}{\nu} \cdot \mathrm{Geo}_{r^*\nu}(2) \cdot F(s^*) & \dots \end{pmatrix}^\top$$

$$\mathbf{J}^* = g^* \cdot f^* - 1$$

$$\begin{aligned} \mathbf{D}(s^*) &= s^* + \sum_{b=0}^{\infty} \frac{1}{\nu} b^2 \, \mathrm{Geo}_{r^*\nu}(b) \cdot F(s^*) \\ &= s^* + r^* (1 + 2\, r^* \nu) \cdot F(s^*) \\ &= 2 \cdot s^* \cdot (1 + r^* \nu) \end{aligned} \tag{B.11}$$

Here, $g^* := \frac{d\,\bar{G}(c_m)}{d\,c_m}|_{c_m = \mathcal{F}(c_s^*)}$ and $r^* := \frac{s^*}{F(s^*)}$. The one-dimensional equation can be written as $\mathbf{\Sigma} = -\frac{\mathbf{D}^*}{2\,\mathbf{J}^*}$, so that the approximate local variance of the protein copy number reads

$$\mathrm{Var}^*(S) = s^* \cdot \frac{1 + r^* \nu}{1 - f^* g^*} = s^* \left[1 + \frac{1}{(g^*)^{-1} - f^*} \left(f^* + \frac{r^* \nu}{g^*} \right) \right]. \tag{B.12}$$

B.5. Hybrid linear noise approximation with NB distributed protein copy number

The NB-based iterative hLNA can be implemented as follows:

Pseudoalgorithm

- Initialization:
  Set $\bar{F}(s) := F(s) \;\forall\, s \in \mathbb{R}_{\geq 0}$ and $\bar{G}(m) := \sum_{n=0}^{\infty} G(n) \cdot \mathrm{Pois}_m(n) \;\forall\, m \in \mathbb{R}_{\geq 0}$.

- Update stationary states:
  Choose m^* and s^* as solutions of

  $$\begin{aligned} s^* &= \bar{G}(\bar{F}(s^*)) \\ m^* &= \bar{F}(s^*) \end{aligned}$$

  and return the values.

- Update Fano factors:
  Calculate the derivatives $\bar{F}'(s)$ and $\bar{G}'(m)$ and the Fano factors

  $$\begin{aligned} \eta^*(M) &= 1 + \frac{\bar{F}'(s^*)}{(\bar{G}'(m^*))^{-1} - \bar{F}'(s^*)} \left(\frac{\nu}{1+\nu} + \frac{1}{1+\nu} \cdot \frac{s^* \cdot \bar{F}'(s^*)}{m^* \cdot \bar{G}'(m^*)} \right) \\ \eta^*(S) &= 1 + \frac{1}{(\bar{G}'(m^*))^{-1} - \bar{F}'(s^*)} \left(\frac{\bar{F}'(s^*)}{1+\nu} + \frac{\nu}{1+\nu} \cdot \frac{m^* \cdot \bar{G}'(m^*)}{s^*} \right) \end{aligned} \tag{B.13}$$

Return the Fano factors.

- Update effective rates:
  If $\eta^*(M) > 1$ set

$$\bar{G}(m) = \sum_{n=0}^{\infty} G(n) \cdot \mathrm{NB}_{m,\,\eta^*(M)}(n)$$

  else $$\bar{G}(m) = \sum_{n=0}^{\infty} G(n) \cdot \mathrm{Pois}_m(n)$$

  If $\eta^*(S) > 1$ set

$$\bar{F}(s) = \sum_{r=0}^{\infty} F(r) \cdot \mathrm{NB}_{s,\,\eta^*(S)}(r)$$

  else $$\bar{F}(s) = \sum_{r=0}^{\infty} F(r) \cdot \mathrm{Pois}_s(r)$$

  Go back to step where stationary states are updated unless the stopping criterion is fulfilled.

The PMF $\mathrm{NB}_{x,\eta}$ of the negative binomial distribution with mean x and Fano factor η has been defined in Eq (4.42). Pois_x denotes the PMF of the Poisson distribution with mean x. The stopping criterion could for example be fulfilled when the values of m^*, s^*, $\eta^*(M)$, and $\eta^*(S)$ stop changing significantly between consecutive iterations, indicating consistency between all four estimates.

The idea behind this algorithm is to use the mRNA and protein variances that have been obtained with the Poisson-based hLNA (which was developed in Section 4.3) in order to formulate improved approximations of local mRNA and protein PMFs. Based on these PMFs, the formulation of the locally averaged rates $\bar{F}$ and $\bar{G}$ is adapted. Usually, negative binomial distributions are chosen as a basis for the approximate PMFs. However, since NB distributions are always super-Poissonian, this procedure only works if $\eta^*(M)$ and $\eta^*(S)$ are larger than 1. Otherwise, the Poisson distribution is used instead[1]. If the system is multistable, the algorithm needs to be run for each state separately, and in each iteration it needs to be taken care that the correct stable steady state m^*, s^* is selected.

The algorithm however raises some questions: First of all, it is not guaranteed that mRNA and protein distributions can in any case be well approximated by an NB distribution. This might be a problem since the chosen approximate PMFs have a major impact on the estimated fixed points and Fano factors, whose consistency is the only criterion by which their quality can be intrinsically assessed. It might thus be reasonable to somehow take the uncertainty of the approximate PMFs into account.

Second, the consistency between the local fixed point and Fano factor estimates is never checked simultaneously: In the i-th iteration, the current values of $m^*_{\langle i \rangle}$ and $s^*_{\langle i \rangle}$ are used to

[1] In case of the mRNA distribution, the Poisson PMF has been shown to be a good approximation anyway. In case of the protein distribution, using the Poisson distribution usually yields an effective rate $\bar{F}$ which is close to F. This is because the protein copy number is typically high so that the local averaging effect of a Poisson distribution is small. The Poisson PMF is therefore expected to yield good estimates in case of sub-Poissonian distributions as well.

obtain new estimates of $\eta^*(M)_{\langle i \rangle}$ and $\eta^*(S)_{\langle i \rangle}$, but the effective rates based on these Fano factors might yield updated stable states $m^*_{\langle i+1 \rangle}$ and $s^*_{\langle i+1 \rangle}$ that differ from the previous ones. This shows that the estimates in the i-th iteration were not consistent, and it is not guaranteed that this inconsistency is reduced in the course of iterations. In other words, the convergence of the algorithm is not ensured. Moreover, it is unclear whether the algorithm is robust, i.e. whether its convergence depends on the initialization of the algorithm or not.

In order to perform some preliminary tests, the algorithm was applied to the system in Fig 4.5 (C) for $\nu = 0.01$ and for $\nu = 0.1$, where the Poisson-based hLNA yielded inaccurate estimates (in particular concerning s^*). It has been discussed that the reason for the observed deviations might be large protein noise, which is neglected in the Poisson-based hLNA formulation, although it may have a strong impact on the effective transcription rate. The results of the NB-based hLNA are summarized in Table B.1: It shows the mean and Fano factor estimates during the first iterations ("iter"), where the 0-th iteration corresponds to the values obtained from Poisson-based hLNA, with which the algorithm is initialized. The values extracted from stochastic simulations (with SEMs) are given in the last row ("true").

Table B.1.: Mean and Fano factor estimates according to the iterative NB-based hLNA.

	$\nu = 0.01$				$\nu = 0.1$			
iter	s^* $[\cdot 10]$	$\eta^*(S)$	m^* $[\cdot 10^{-1}]$	$\eta^*(M)$	s^* $[\cdot 10]$	$\eta^*(S)$ $[\cdot 10]$	m^* $[\cdot 10^{-1}]$	$\eta^*(M)$
0	4.258	6.503	2.532	1.026	4.258	3.710	2.532	1.168
1	3.889	7.037	2.315	1.029	2.377	5.291	1.438	1.254
2	3.856	7.084	2.296	1.029	1.821	5.347	1.120	1.255
3	3.853	7.088	2.294	1.029	1.812	5.345	1.114	1.254
10	3.853	7.088	2.294	1.029	1.812	5.346	1.114	1.254
true	3.79 ± 0.13	7.72 ± 0.63	2.26 ± 0.08	1.03 ± 0.01	2.08 ± 0.28	4.20 ± 0.30	1.26 ± 0.18	1.19 ± 0.03

Obviously, convergence happened very fast in these two examples: After ten iterations, the values stayed constant. In the system where $\nu = 0.01$, all estimates could be improved compared to Poisson-based hLNA. For $\nu = 0.1$, the local mean value estimates moved closer to the true values in the course of the iterations, while the mRNA and protein Fano factors were both slightly overestimated. Choosing different, quite arbitrary initial PMFs lead in both systems to the same estimates after few iterations (data not shown), indicating that the algorithm worked robustly in these cases.

All in all, the idea behind the algorithm seems to be promising, but a lot of further studies about convergence and quality still need to be performed.

C. Calculations for Chapter 5

C.1. Noise in systems with saturated translational propensity

In order to compare the local Fano factors of different stable expression states in a bistable system with concave translational propensity function, the dependence of $\bar{G}'(m)$ and of $\frac{\bar{G}'(m)\cdot m}{\bar{G}(m)}$ on m is studied. Based on that, the values of g_i^* and $\frac{g_i^*}{r_i^*}$ $i = 1, 2,$ at the two fixed points $(m_1^*, s_1^*) < (m_2^*, s_2^*)$ can be compared (cf. Section 5.6.2).

First, the behavior of $\bar{G}'(m)$ is regarded. Since G is concave,

$$G(n+1) \geq \frac{1}{2}\,G(n) + \frac{1}{2}\,G(n+2) \quad \forall\, n \in \mathbb{N}_0. \tag{C.1}$$

holds. Therefore,

$$\bar{G}''(m) = \sum_{n=0}^{\infty}(G(n+2) - 2G(n+1) + G(n))\,\frac{m^n}{n!}\,e^{-m} \leq 0 \quad \forall\, m \in \mathbb{R}_{\geq 0}. \tag{C.2}$$

As a consequence, $g_1^* \geq g_2^*$. (In an analogous manner, it can be proven that $\bar{G}$ is strictly concave, convex, or strictly convex, whenever G is.)

The monotonicity of $\frac{\bar{G}'(m)\cdot m}{\bar{G}(m)}$ depends on the choice of G. Here, we set

$$G(m) = u\,\frac{m}{K_m + m} \tag{C.3}$$

as in the main text. In this case, one obtains:

$$\begin{aligned}
\frac{\bar{G}'(m)\cdot m}{\bar{G}(m)} &= \frac{\sum_{n=0}^{\infty}(G(n+1) - G(n))\,\frac{m^n}{n!}\,e^{-m}}{\sum_{n=0}^{\infty}\frac{G(n+1)}{n+1}\,\frac{m^n}{n!}\,e^{-m}} \\
&= \frac{\sum_{n=0}^{\infty}\frac{K_m}{(K_m+n+1)(K_m+n)}\,\frac{m^n}{n!}\,e^{-m}}{\sum_{n=0}^{\infty}\frac{1}{K_m+n+1}\,\frac{m^n}{n!}\,e^{-m}} \\
&= \frac{\sum_{n=0}^{\infty}h(n)\cdot h(n+1)\,\frac{m^n}{n!}\,e^{-m}}{\sum_{n=0}^{\infty}h(n+1)\,\frac{m^n}{n!}\,e^{-m}}\cdot K_m
\end{aligned} \tag{C.4}$$

with $h(n) := \frac{1}{K_m+n}$. Let $N := \sum_{n=0}^{\infty}h(n+1)\,\frac{m^n}{n!}\,e^{-m}$ be the denominator of the above expression. Then:

$$\begin{aligned}
&\frac{1}{K_m}\,\frac{d}{dm}\left(\frac{\bar{G}'(m)\cdot m}{\bar{G}(m)}\right) \\
= \;& \frac{\left(\sum_{n=0}^{\infty}(h(n+1)\,h(n+2) - h(n)\,h(n+1))\,\frac{m^n}{n!}\,e^{-m}\right)\left(\sum_{n=0}^{\infty}h(n+1)\,\frac{m^n}{n!}\,e^{-m}\right)}{N^2} \\
- \;& \frac{\left(\sum_{n=0}^{\infty}(h(n+2) - h(n+1))\,\frac{m^n}{n!}\,e^{-m}\right)\left(\sum_{n=0}^{\infty}h(n)\,h(n+1)\,\frac{m^n}{n!}\,e^{-m}\right)}{N^2}
\end{aligned}$$

$$
\begin{aligned}
&= \frac{\left(\sum_{n=0}^{\infty} h(n+1)\,h(n+2)\,\frac{m^n}{n!}\,e^{-m}\right)\left(\sum_{n=0}^{\infty} h(n+1)\,\frac{m^n}{n!}\,e^{-m}\right)}{N^2} \\
&\quad - \frac{\left(\sum_{n=0}^{\infty} h(n+2)\,\frac{m^n}{n!}\,e^{-m}\right)\left(\sum_{n=0}^{\infty} h(n)\,h(n+1)\,\frac{m^n}{n!}\,e^{-m}\right)}{N^2} \\
&= \frac{\left(\sum_{n=0}^{\infty}\sum_{k=0}^{n} h(k+1)\,h(k+2)\,h(n-k+1)\,\frac{m^n}{k!(n-k)!}\,e^{-2m}\right)}{N^2} \\
&\quad - \frac{\left(\sum_{n=0}^{\infty}\sum_{k=0}^{n} h(k+2)\,h(n-k)\,h(n-k+1)\,\frac{m^n}{k!(n-k)!}\,e^{-2m}\right)}{N^2} \\
&= \frac{1}{N^2}\sum_{n=0}^{\infty}\sum_{k=0}^{n} h(k+2)\,h(n-k+1)\,\underbrace{[h(k+1)\;-\;h(n-k)]}_{\substack{=\frac{1}{K_m+k+1}-\frac{1}{K_m+n-k} \\ =h(k+1)\,h(n-k)\,(n-2k-1)}}\binom{n}{k}2^{-n}\,\frac{(2m)^n}{n!}\,e^{-2m} \\
&= \frac{1}{N^2}\sum_{n=0}^{\infty}\sum_{k=0}^{n} \underbrace{h(k+1)\,h(k+2)\,h(n-k)\,h(n-k+1)}_{:=H(k)}\,(n-k)\binom{n}{k}2^{-n}\,\frac{(2m)^n}{n!}\,e^{-2m} \\
&\quad - \frac{1}{N^2}\sum_{n=0}^{\infty}\sum_{k=0}^{n} H(k)\,(k+1)\binom{n}{k}2^{-n}\,\frac{(2m)^n}{n!}\,e^{-2m} \\
&= \frac{1}{N^2}\sum_{n=0}^{\infty}\sum_{j=-1}^{n-1} H(j)\,(j+1)\binom{n}{j+1}2^{-n}\,\frac{(2m)^n}{n!}\,e^{-2m} \\
&\quad - \frac{1}{N^2}\sum_{n=0}^{\infty}\sum_{k=0}^{n} H(k)\,(k+1)\binom{n}{k}2^{-n}\,\frac{(2m)^n}{n!}\,e^{-2m} \\
&= -\frac{1}{N^2}\sum_{n=0}^{\infty} H(n)\,(n+1)\,2^{-n}\,\frac{(2m)^n}{n!}\,e^{-2m} \\
&\quad + \frac{1}{N^2}\sum_{n=0}^{\infty}\sum_{k=0}^{n-1} H(k)\,(k+1)\left(\binom{n}{k+1}-\binom{n}{k}\right)2^{-n}\,\frac{(2m)^n}{n!}\,e^{-2m} \\
&< \frac{1}{N^2}\sum_{n=0}^{\infty}\sum_{k=0}^{\lfloor\frac{n}{2}\rfloor-1} H(k)\,(k+1)\left(\binom{n}{k+1}-\binom{n}{k}\right)2^{-n}\,\frac{(2m)^n}{n!}\,e^{-2m} \\
&\quad + \frac{1}{N^2}\sum_{n=0}^{\infty}\sum_{k=\lceil\frac{n}{2}\rceil}^{n-1} H(k)\,(k+1)\left(\binom{n}{k+1}-\binom{n}{k}\right)2^{-n}\,\frac{(2m)^n}{n!}\,e^{-2m} \\
&= \frac{1}{N^2}\sum_{n=0}^{\infty}\sum_{k=0}^{\lfloor\frac{n}{2}\rfloor-1} H(k)\cdot(k+1)\left(\binom{n}{k+1}-\binom{n}{k}\right)2^{-n}\,\frac{(2m)^n}{n!}\,e^{-2m} \\
&\quad - \frac{1}{N^2}\sum_{n=0}^{\infty}\sum_{j=0}^{\lfloor\frac{n}{2}\rfloor-1} H(j)\cdot(n-j)\left(\binom{n}{j+1}-\binom{n}{j}\right)2^{-n}\,\frac{(2m)^n}{n!}\,e^{-2m}. \qquad \text{(C.5)}
\end{aligned}
$$

During this calculation, we twice substituted $j = n-k-1$ and used the symmetry of the binomial coefficient. For $k = 0, ..., \lfloor\frac{n}{2}\rfloor - 1$, $\left(\binom{n}{k+1}-\binom{n}{k}\right) > 0$. Furthermore, $k+1 < n-k$. Therefore, the above expression is negative. From this follows that $\frac{\bar{G}'(m)\cdot m}{\bar{G}(m)}$ is monotonously decreasing and that $\frac{g_1^*}{r_1^*} \geq \frac{g_2^*}{r_2^*}$.

C.2. Evaluation of first passage times for the systems in Fig 5.5

The three regulatory systems considered in Fig 5.5 were analytically shown to exhibit different noise patterns, from which qualitative statements about their relative robustness were deduced that were verified through stochastic simulations. Here, first passage times (FPTs) of noise-driven transitions from the inactive (OFF) to the active (ON) state as well as from the ON to the OFF state are determined computationally based on the simulations. For each system and switching direction, 10^3 trajectories have been generated using the Gillespie algorithm, starting from the OFF ($s^* = 20$) and the ON ($s^* = 200$) state, respectively. The first passage times τ_{FPT} correspond to the time points when the trajectory reaches the other opposite expression state for the first time. For different reaction times τ_f, the percentage of cells for which $\tau_{\mathrm{FPT}} < \tau_f$ holds is evaluated (these are the cells that have switched during $[0, \tau_f]$). The faster this percentage rises, the less robust is the state. The mean FPT (MFPT) of a switch is determined by taking the average of the FPTs (given in Table C.1 together with the standard error of the mean). Extreme robustness of an expression state might impede the determination of the MFPT (note that MFPTs might be infinite). This case is indicated with the entry "n.a.". Table C.1 summarizes the results for the three bistable systems illustrated in Fig 5.5. It shows that in the systems visualized in panels (A) and (B), where G is linear, the OFF state is more robust than the ON state. In order to achieve the opposite, one might try to adjust the parameters in F accordingly. However, the range of modulations is limited as bistability needs to be preserved. Much greater flexibility is ensured by allowing nonlinear translation propensities (system in panel (C)) that generate state-dependent burst sizes.

Table C.1.: Percentage of cells that have left their original state within $[0, \tau_f]$ and MFPTs

	Switch	**Cells having switched during** $[0, \tau_f]$ [%]				**MFPT** $[\cdot 10^2]$
(A)	τ_f $[\cdot 10^2]$	1.3	2.6	5.1	10.3	
	OFF $\rightarrow$ ON	0.00	0.00	0.01	0.01	n.a.
	ON $\rightarrow$ OFF	0.29	0.50	0.78	0.95	3.49 ± 3.03
(B)	τ_f $[\cdot 10]$	0.8	1.56	3.12	6.24	
	OFF $\rightarrow$ ON	0.02	0.05	0.11	0.23	2.24 ± 2.34
	ON $\rightarrow$ OFF	0.28	0.51	0.77	0.94	0.22 ± 0.20
(C)	τ_f $[\cdot 10]$	0.72	1.44	2.88	5.76	
	OFF $\rightarrow$ ON	0.19	0.52	0.81	0.97	0.19 ± 0.15
	ON $\rightarrow$ OFF	0.00	0.00	0.00	0.00	n.a.

In the cases shown here, FPTs are useful for comparing the robustness of different stable expression states within the same regulatory system. If the aim is to compare different systems, the reference time scale needs to be defined (e.g. the original process time, or time normalized with respect to mRNA or protein degradation rates).

D. Calculations for Chapter 6

D.1. Stochastic description of the fast mRNA limit

On the time-scale defined by τ/ν_1, the CME (6.2) reduces in the limit $\nu_1 \to 0$, $\nu_2 \to 0$, $\frac{\nu_1}{\nu_2} = \text{const.}$ to:

$$\begin{aligned}
&\frac{d\,p_{m_1,s_1,m_2,s_2}}{d(\tau/\nu_1)} = \\
&F_1(s_2)\,p_{m_1-1,s_1,m_2,s_2} - F_1(s_2)\,p_{m_1,s_1,m_2,s_2} + (m_1+1)\,p_{m_1+1,s_1,m_2,s_2} - m_1\,p_{m_1,s_1,m_2,s_2} \\
+\Big(&\big[F_2(s_1)\,p_{m_1,s_1,m_2-1,s_2} - F_2(s_1)\,p_{m_1,s_1,m_2,s_2} + (m_2+1)\,p_{m_1,s_1,m_2+1,s_2} - m_2\,p_{m_1,s_1,m_2,s_2}\big]\cdot\frac{\delta\nu_1}{\nu_2}.
\end{aligned} \tag{D.1}$$

Summation over m_2 and s_1 yields:

$$\frac{d\,p_{m_1,s_2}}{d(\tau/\nu_1)} = F_1(s_2)\,p_{m_1-1,s_2} - F_1(s_2)\,p_{m_1,s_2} + (m_1+1)\,p_{m_1+1,s_2} - m_1\,p_{m_1,s_2}. \tag{D.2}$$

Using Bayes' theorem, the joint PMF can be re-written as $p_{m_1,s_2} = p_{m_1|s_2}\,p_{s_2}$ if $p_{s_2} \neq 0$. The derivative of the LHS fulfills:

$$\frac{d\,p_{m_1,s_2}}{d(\tau/\nu_1)} = \frac{d\,p_{m_1|s_2}}{d(\tau/\nu_1)}\cdot p_{s_2} + \frac{d\,p_{s_2}}{d(\tau/\nu_1)}\cdot p_{m_1|s_2} \approx \frac{d\,p_{m_1|s_2}}{d(\tau/\nu_1)}\cdot p_{s_2}, \tag{D.3}$$

since the protein dynamics are assumed to be constant on the fast time-scale. Cancelling p_{s_2} on both sides of Eq (D.2) yields Eq (6.4). Analogously, Eq (6.5) is obtained.

On the time-scale of protein dynamics, the mRNA species are assumed to be in their pseudo steady-state, which means that they are Poisson distributed and $\mathbb{E}[M_1|S_2 = s_2] = F_1(s_2)$ and $\mathbb{E}[M_2|S_1 = s_1] = F_2(s_1)$, as explained in the main text. Therefore, the joint PMF of S_1 and S_2 reduces to:

$$\begin{aligned}
\dot{p}_{s_1,s_2} = &\sum_{m_1=0}^{\infty}\sum_{m_2=0}^{\infty}\Big(G_1(m_1)\,p_{m_1,m_2|s_1-1,s_2}\,p_{s_1-1,s_2} - G_1(m_1)\,p_{m_1,m_2|s_1,s_2}\,p_{s_1,s_2} \\
&+ (s_1+1)\,p_{m_1,m_2|s_1+1,s_2}\,p_{s_1+1,s_2} - s_1\,p_{m_1,m_2|s_1,s_2}\,p_{s_1,s_2} \\
&+ \big[G_2(m_2)\,p_{m_1,m_2|s_1,s_2-1}\,p_{s_1,s_2-1} - G_2(m_2)\,p_{m_1,m_2|s_1,s_2}\,p_{s_1,s_2} \\
&+ (s_2+1)\,p_{m_1,m_2|s_1,s_2+1}\,p_{s_1,s_2+1} - s_2\,p_{m_1,m_2|s_1,s_2}\,p_{s_1,s_2}\big]\cdot\delta\Big) \\
= &\quad \mathbb{E}[G_1(M_1)|s_1-1,s_2]\,p_{s_1-1,s_2} - \mathbb{E}[G_1(M_1)|s_1,s_2]\,p_{s_1,s_2} \\
&+ (s_1+1)\,p_{s_1+1,s_2} - s_1\,p_{s_1,s_2} \\
&+ \big[\mathbb{E}[G_2(M_2)|s_1,s_2-1]\,p_{s_1,s_2-1} - \mathbb{E}[G_2(M_2)|s_1,s_2]\,p_{s_1,s_2} \\
&+ (s_2+1)\,p_{s_1,s_2+1} - s_2\;p_{s_1,s_2}\big]\cdot\delta\Big).
\end{aligned} \tag{D.4}$$

Inserting the PSS-distributions of the mRNA copy numbers determined above leads to Eq (6.6).

D.2. Application of the LNA to the genetic toggle switch

The stoichiometric matrix, the vector of stationary propensities, and the diffusion matrix $\mathbf{D}^*$ read:

$$\mathbf{A} = \begin{pmatrix} 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 \end{pmatrix}$$

$$\mathbf{w}(m_1^*, s_1^*, m_2^*, s_2^*) = \left(\tfrac{1}{\nu_1} m_1^* \quad \tfrac{1}{\nu_1} m_1^* \quad s_1^* \quad s_1^* \quad \tfrac{\delta}{\nu_2} m_2^* \quad \tfrac{\delta}{\nu_2} m_2^* \quad \delta\, s_2^* \quad \delta\, s_2^* \right)^\top,$$

$$\mathbf{D}^* = \begin{pmatrix} 2\,\tfrac{1}{\nu_1} m_1^* & 0 & 0 & 0 \\ 0 & 2\, s_1^* & 0 & 0 \\ 0 & 0 & 2\,\tfrac{\delta}{\nu_2} m_2^* & 0 \\ 0 & 0 & 0 & 2\,\delta\, s_2^* \end{pmatrix}. \tag{D.5}$$

The Jacobian matrix $\mathbf{J}^*$ is given in Eq (6.12). Now, the Lyapunov equation (2.29) needs to be solved for $\mathbf{\Sigma}^*$. Due to its symmetry, it is basically an inhomogeneous linear system of 10 equations with 10 unknown variables (these are the variances and covariances) $\mathbf{C} \cdot \mathbf{x} = \mathbf{b}$ with

$$\mathbf{C} = \begin{pmatrix} -\frac{1}{\nu_1} & 0 & 0 & \frac{f_1^*}{\nu_1} & 0 & 0 & 0 & 0 & 0 & 0 \\ g_1 & -\left(\frac{1}{\nu_1}+1\right) & 0 & 0 & 0 & 0 & \frac{f_1^*}{\nu_1} & 0 & 0 & 0 \\ 0 & \frac{f_2^*\cdot\delta}{\nu_2} & -\left(\frac{1}{\nu_1}+\frac{\delta}{\nu_2}\right) & 0 & 0 & 0 & 0 & 0 & \frac{f_1^*}{\nu_1} & 0 \\ 0 & 0 & g_2\cdot\delta & -\left(\frac{1}{\nu_1}+\delta\right) & 0 & 0 & 0 & 0 & 0 & \frac{f_1^*}{\nu_1} \\ 0 & g_1 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & g_1 & 0 & \frac{f_2^*\cdot\delta}{\nu_2} & -\left(1+\frac{\delta}{\nu_2}\right) & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & g_1 & 0 & g_2\cdot\delta & -(1+\delta) & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{f_2^*\cdot\delta}{\nu_2} & 0 & -\frac{\delta}{\nu_2} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \frac{f_2^*\cdot\delta}{\nu_2} & g_2\cdot\delta & -\left(\frac{\delta}{\nu_2}+\delta\right) & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & g_2\cdot\delta & \delta \end{pmatrix},$$

$$\mathbf{x} = \begin{pmatrix} \mathrm{Var}(M_1) \\ \mathrm{Cov}(M_1, M_2) \\ \mathrm{Cov}(M_1, S_1) \\ \mathrm{Cov}(M_1, S_2) \\ \mathrm{Var}(M_2) \\ \mathrm{Cov}(M_2, S_1) \\ \mathrm{Cov}(M_2, S_2) \\ \mathrm{Var}(S_1) \\ \mathrm{Cov}(S_1, S_2) \\ \mathrm{Var}(S_2) \end{pmatrix}, \qquad \mathbf{b} = - \begin{pmatrix} \frac{m_1^*}{\nu_1} \\ 0 \\ 0 \\ 0 \\ s_1^* \\ 0 \\ 0 \\ \frac{m_2^*\cdot\delta}{\nu_2} \\ 0 \\ s_2^*\cdot\delta \end{pmatrix}, \tag{D.6}$$

for which a unique solution exists.