Journal of Process Control 42 (2016) 14–34

Contents lists available at ScienceDirect

Journal of Process Control

journal homepage: www.elsevier.com/locate/jprocont

Review

Robustness analysis, prediction, and estimation for uncertain biochemical networks: An overview☆

Stefan Streif[a,*], Kwang-Ki K. Kim[d], Philipp Rumschinski[c], Masako Kishida[c], Dongying Erin Shen[b], Rolf Findeisen[c], Richard D. Braatz[b]

[a] Technische Universität Chemnitz, 09107 Chemnitz, Germany
[b] Massachusetts Institute of Technology, Cambridge, MA 02139, USA
[c] Otto-von-Guericke-Universität Magdeburg, 39106 Magdeburg, Germany
[d] Georgia Institute of Technology, Atlanta, GA 30308, USA

ARTICLE INFO

Article history:
Received 6 October 2014
Received in revised form 14 March 2016
Accepted 18 March 2016
Available online 19 April 2016

Keywords:
Biochemical reaction network
Complex dynamical system
Estimation
Robustness
Systems and control theory

ABSTRACT

Mathematical models of biochemical reaction networks are important tools in systems biology and systems medicine, e.g., to analyze disease causes or to make predictions for the development of effective treatments. Models are also used in synthetic biology for the design of circuits that perform specialized tasks. Prediction, analysis and design require plausible and reliable models, that is, models must reflect the properties of interest of the considered biochemical networks. One remarkable property of biochemical networks is *robust functioning* over a wide range of perturbations and environmental conditions. The intrinsic robustness of a network should be reflected into its associated mathematical model. The description and analysis of robustness in biochemical reaction networks are challenging, however, because accounting explicitly for the various types of structural, parametric and data uncertainty in the description of the models is not straightforward. Furthermore, system properties are typically inherently uncertain and often only given by qualitative or verbal descriptions that impede a straightforward and comprehensive mathematical analysis. In the first part of this overview article, network functions and behaviors of interest are formally defined, and different classes of uncertainties and perturbations are consistently described. The second part reviews frequently used mathematical formulations and presents the authors' recent developments for robustness analysis, estimation, and model-based prediction. One biochemical network model is used to illustrate the capabilities of various methods to deal with the different types of uncertainties and robustness requirements.

© 2016 Elsevier Ltd. All rights reserved.

Contents

1. Introduction 15
   1.1. Challenge 1: Specification of uncertainties and system behaviors 16
   1.2. Challenge 2: Robustness analysis, estimation, and prediction methods 16
   1.3. Outline of this article 17
2. Biochemical network models 17
   2.1. Conservation relations 18
   2.2. Discrete-time approximations 18
   2.3. Linearization 19

☆ A preliminary version of this article was presented at the Computer Applications in Biotechnology and 10th International Symposium on Dynamics and Control of Process Systems, Mumbai, India, December 16–20, 2013, see [120].

* Corresponding author.
E-mail addresses: stefan.streif@etit.tu-chemnitz.de (S. Streif), kwangki.kim@ece.gatech.edu (K.-K.K. Kim), philipp.rumschinski@ovgu.de (P. Rumschinski), masako.kishida@ovgu.de (M. Kishida), dongying@mit.edu (D.E. Shen), rolf.findeisen@ovgu.de (R. Findeisen), braatz@mit.edu (R.D. Braatz).

http://dx.doi.org/10.1016/j.jprocont.2016.03.004
0959-1524/© 2016 Elsevier Ltd. All rights reserved.

3. Uncertainty descriptions and robustness specifications 19
3.1. Model structure uncertainties and model hypotheses 20
3.2. Set-based uncertainties 20
3.2.1. Interval uncertainties 20
3.2.2. General set-based uncertainties 20
3.2.3. Parameter or data relations 20
3.3. Probabilistic uncertainties 20
3.4. Qualitative information on system behavior 21
3.4.1. Steady-state location 21
3.4.2. Stability, instability, and oscillation 21
3.4.3. Adaptation and inverse response 21
3.4.4. Conditional and temporal observations 21
3.5. Problem statements for robustness analysis 22
4. Analysis of local perturbations and network structures 22
4.1. Sensitivity analysis 22
4.1.1. Local sensitivity analysis and metabolic control analysis 22
4.1.2. Global sensitivity analysis 23
4.2. Bifurcation analysis 23
4.3. Kinetic perturbations 23
4.4. Qualitative behavior and its dependence on network structure and feedback loops 24
4.4.1. Structural robustness 24
4.4.2. Monotone systems 24
5. Set-based uncertainties and analysis 24
5.1. Interval analysis 24
5.1.1. Stability analysis using interval matrices 24
5.2. Linear and semidefinite relaxations 25
5.2.1. Structural uncertainties and model invalidation 25
5.2.2. Uncertainty propagation and parameter estimation 26
5.2.3. Stability and instability 26
5.3. Mixed-integer linear relaxations 26
5.3.1. Inner approximations 26
5.3.2. Including qualitative information in robustness analysis 27
5.4. Methods for continuous-time dynamical systems 27
5.5. Approaches using skewed structured singular values 27
5.5.1. Bounds on uncertainty propagation 27
5.5.2. Inner approximation 28
6. Probabilistic uncertainties and analysis 28
6.1. Approximate analysis using the Fokker-Planck equation 28
6.1.1. Stochastic stability analysis 28
6.1.2. Sensitivity analysis 29
6.2. Polynomial chaos theory 29
6.2.1. Propagation of probabilistic uncertainty 29
6.2.2. Steady-state probability distributions 29
6.2.3. Stochastic stability 30
6.2.4. Scalable uncertainty quantification 31
7. Discussion and conclusions 31
Acknowledgements 31
References 31

1. Introduction

Biochemical reaction networks form the structural basis of most cellular processes such as metabolism, signal transduction, and gene expression. In these networks, many species dynamically interact and are transformed by biochemical reactions to perform and maintain biological functions. Intertwined and possibly redundant feedback and feedforward mechanisms give rise to complex dynamical behaviors while their lack or improper functioning can result in dysfunctions or diseases. To minimize such risks, biological networks must perform their tasks reliably under various changes of the cellular environment and conditions [19,58,74,80]. This property is generally called *robustness* and refers to the *persistence of a behavior* or the *insensitivity of function characteristics* in the presence of (external or internal) perturbations [130]. Typical examples of *biological behaviors* that are robust to environmental changes are oscillations or multistability in the cell cycle, in apoptosis, or in adaptation [121,2]. The readers are referred to [1] for other examples and mechanisms of cellular regulation.

Robust functioning is of particular interest in synthetic biology or metabolic engineering. One core task in synthetic biology is the design of motifs or building blocks that perform a function robustly when connected into larger networks and under various perturbations of the cellular environment [99]. Function characteristics of interest include certain types of dynamic behavior such as the time derivation of inputs as required for adaptation to constant stimuli, logical combinations of different inputs, oscillatory behavior or multistability [113]. Note that these behaviors are described in a verbal and qualitative manner.

Besides the *robustness of qualitative behaviors* or function characteristics, *quantitative predictions* of system responses have become increasingly important, especially in therapy design and synthetic biology [145]. Often mathematical models are developed and employed to analyze and quantitatively predict, estimate, and control the response of the considered systems with respect to applied

inputs or environmental changes [73]. The main challenges are not only to consider the various external perturbations of these systems, but also to take into account the various uncertainties that arise in the models [113]. The uncertainty in models of biological networks can arise due to uncertain data and parameters, incomplete knowledge of the reaction kinetics, neglected intermediate reaction steps, and unmodeled transport phenomena such as diffusion and delays.

This article provides an overview on the modeling of biochemical reaction networks with a focus on describing and analyzing robustness, taking qualitative as well as quantitative aspects into account. More precisely, robustness analysis requires a formal specification and definition of the analyzed behavior or function characteristics, and of the uncertainties with respect to which robustness is to be analyzed. Robustness can then be analyzed and quantified by determining the allowable uncertainties for which the desired system behaviors or function characteristics are still observed. Because descriptions of biological functions are inherently uncertain, robustness analysis is inevitably linked to uncertainty analysis. To this end, two key challenges are outlined:

1 specification of the different types of uncertainties encountered in the description of system behaviors, functions, and in the data and models;
2 methods for analyzing and quantifying robustness under consideration of uncertainties using prediction and estimation.

While the first challenge is mainly of a conceptual nature, the second is still an open field of research concerning the development of *computationally efficient* methods for analysis and predictions.

1.1. Challenge 1: Specification of uncertainties and system behaviors

Mathematical models of robust biological systems should exhibit appropriate levels of robustness when analyzed [59,58,80]. For robustness analysis and prediction of system responses to perturbations and uncertainties, two crucial ingredients are required: first, a clear *description of external or internal perturbations* under which the biological system (represented by the model) should function robustly; second, a clear *formulation of the behavior* or *function characteristics* that are to be analyzed for its robustness. However, considering only (external or internal) perturbations is far too limited for the analysis of biological models because the models itself are inherently uncertain. The model uncertainty is due to the fact that experimental data are sparse, limited and incomplete, since measurement techniques are mostly indirect and have very low accuracy [101,113,131]. These factors result in large uncertainties of the absolute quantities of the measured physical or chemical entities or species. In addition, due to low sampling times and missing normalization standards for absolute quantification, the data are usually neither quantitative nor time-resolved. Often, the data available for model construction are supplemented by *qualitative information* such as conditional or temporal statements or *if-then* observations [24,103,107].

For the analysis of robustness and for prediction and estimation, different type of uncertainties must be modeled and considered:

1 *external perturbations*;
2 *uncertainties and qualitative formulation of the investigated model behaviors or function characteristics*;
3 *uncertain measurement data*;
4 *structural uncertainties* due to incomplete knowledge of the reaction kinetics or intermediate reaction steps;
5 *parametric uncertainties* due to unknown reaction rate constants.

As stated in [50], if a given model proves to be fragile against biologically probable perturbations, then it is important to establish whether the lack of robustness is a property of the cellular function itself or is caused by an uncertain or incomplete model. In either case, more detailed analysis is required to gain insight into the source of the lacking robustness. For synthetic networks or therapeutic treatments, a redesign should be taken into account if model-based predictions do not guarantee robust performance. Methods must be chosen for robustness analysis, robust prediction, and estimation that can account for the encountered and largely different types of uncertainties. Currently no suitable tools exist that can handle all of the described uncertainties simultaneously.

1.2. Challenge 2: Robustness analysis, estimation, and prediction methods

Methods for the analysis of robustness and robust prediction for biochemical reaction networks should allow nonlinearities to be taken into account. The methods should be able to handle different types of uncertainties and to make robust statements on network performance, qualitative behavior, and the influence of uncertainties (see Fig. 1). In particular, the interest is in making dynamical predictions of system outputs under uncertainties and perturbations (left to right in Fig. 1). To quantify robustness of function characteristics, parameter estimation can be used (right to left in Fig. 1) where the volume of the consistent and robust parameter set could serve as a measure of robustness [21].

Robustness analysis, robust estimation, and prediction are classical topics in control engineering, for example, see [147]. Most existing methods were developed for linear systems, whereas realistic biochemical networks are nonlinear. Methods that can handle nonlinear systems are often limited or assume that the steady-state is not affected by uncertainties or perturbations. In contrast to most technical systems where the robustness of stability is the main objective, the robustness of instability is also important in biological systems, because instability is related to biological behaviors such as oscillations or multistability [6,136]. The direct application of classical systems and control methods is therefore limited. In addition, the different types of uncertainties encountered in biological and medicinal research differ compared with technical systems. For example, in human-made technical systems, sensors can often be placed as desired or uncertainties can often be avoided by a suitable design. In contrast, it is believed that biological systems evolved to withstand environmental changes by selection and evolution, not by aiming to cancel the effects of uncertainties itself, but rather by aiming to achieve maximum performance or survival. Due to the limitations of most control theories developed for technical systems, this article describes several recent approaches that are tailored to the needs of biochemical networks.

In the *qualitative* or *coarse-grained* approach, uncertainties are considered implicitly by constructing *qualitative models* without attempting to describe all molecular details, often even neglecting dynamics. Qualitative modeling (such as interaction graphs or Boolean networks) has become an important framework for analyzing signaling networks (for a review, see, for example, [107]). These approaches are applicable to analyze large scale networks in a qualitative setting [114]. However, the models usually fail to make quantitative dynamical predictions, which may restrict their applicability, such as for the development of therapeutic treatments.

Generally speaking, an exclusive qualitative or static description of biological networks will not be sufficient to fully understand how qualitative and quantitative dynamical behaviors arise, or to design suitable intervention strategies to combat diseases. Dynamic and quantitative models may be needed, but making definite predictions on the behavior of such models is difficult

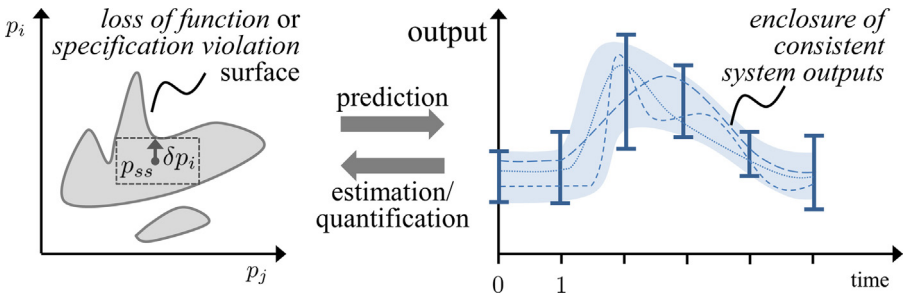

Fig. 1. Robustness analysis and quantification are closely related to the prediction of uncertainty propagation and estimation. Prediction can be used to analyze the influence of uncertainties and perturbations on the model outputs/response. Estimation can be used to determine, for example, parametric uncertainties (gray-shaded areas in the figure on the left) such that certain qualitative or quantitative specifications of system behaviors or properties are satisfied (light blue-shaded areas on the right). Robustness of system behavior can be either quantified with respect to some nominal parameters p_{ss} and the distance δp_i from the surface where the specifications are no longer satisfied, or the volume of the consistent parameter set or some approximation of the set such as the dashed box. To simplify the presentation, only the set-based description is shown in the figure. The ideas are transferable to probabilistic uncertainties and predictions. For such cases, constraint violation is formulated in terms of probabilities.

due to uncertain model components, such as unknown or imprecisely known parameters or kinetic rate expressions, and due to the uncertainties in the experimental data. Most methods for the analysis of uncertain models rely on extensive sampling via Monte Carlo methods. Due to a finite sample size and the complexities and nonlinearities of the models, at best probabilistic statements can be made about system properties, parameter values, or model invalidity using sampling-based methods.

1.3. Outline of this article

In the last decade, several approaches have been developed that can handle or overcome some of the aforementioned challenges. Section 2 reviews modeling of biochemical reaction networks by ordinary differential or difference equations. This article does not review qualitative or structural modeling frameworks or methods allowing for qualitative predictions using these models; for reviews of such methods, see [107,140,10,24,114]. This article does not cover Monte Carlo simulation or related analysis methods. Such methods are, for example, reviewed in [108].

Section 3 describes methods for describing and capturing uncertainties in the models as well as in the descriptions of investigated system behaviors (mainly probabilistic and set-based descriptions). Because the presented methods have individual and different advantages and limitations, and no method can deal with all uncertainty types, the different types are carefully classified. We discuss the selection of suitable methods and the associated types of statements that can be made for robustness analysis and model verification.

Section 4 reviews *classical approaches* for uncertainty analysis and extensions thereof. These methods are often restricted to local or structural analysis, but can still give valuable insights into the system. However, these methods are not well-suited for quantitative predictions and analysis in the case of large uncertainties. Section 5 reviews set-based approaches that can efficiently deal with set-based uncertainty descriptions. Well-known and computationally efficient approaches such as interval analysis fall within this class. Interval analysis methods can produce results that are too conservative, and less restrictive methods are also outlined in Section 5.

While set-based approaches allow robust and guaranteed statements, the results can be conservative. This conservatism can be reduced by employing approaches that provide statements in terms of probabilities and probability distributions with the premises that definite and guaranteed statements are only possible asymptotically. Several existing probabilistic methods are reviewed in Section 6.

Section 7 discusses various problems that have not yet been solved in the current literature and provides an outlook for future research.

A simple example of a reversible enzymatic transformation is used throughout the manuscript to illustrate the different methods and to highlight their advantages and disadvantages. The small reaction network was taken from [35,125] and is often found embedded in (larger) reaction networks.

2. Biochemical network models

This article considers a wide class of biological systems, including metabolic, signal transduction, and gene regulation networks. Most of these processes can be formally modeled as biochemical reaction networks. Although many biological processes are of an inherently stochastic and discrete nature, such as the binding of a ligand to a cellular membrane, the approximation of these processes by deterministic continuous-time models is often acceptable and often the preferred choice for modeling and analysis.

Biochemical networks have two main elements, namely, species and reactions. The biochemical species

$$X_1, X_2, \ldots, X_{n_x}$$

represent ensembles of chemically identical molecules in a specific cell compartment. These species are interconverted by chemical reactions of the general form

$$s_{1,j}^{(s)} X_1 + s_{2,j}^{(s)} X_2 + \cdots + s_{n_x,j}^{(s)} X_{n_x} \longrightarrow s_{1,j}^{(p)} X_1 + s_{2,j}^{(p)} X_2 + \cdots + s_{n_x,j}^{(p)} X_{n_x}, \tag{1}$$

where $j \in \{1, 2, \ldots, n_v\}$ is the reaction index, and the non-negative integers $s_{i,j}^{(s)}, s_{i,j}^{(p)}$ are the stoichiometric coefficients of the substrate and product species, respectively.

The structural information of the reaction network is usually collected in the stoichiometric $S \in \mathbb{R}^{n_x \times n_v}$ matrix with entries given by

$$S_{i,j} := s_{i,j}^{(p)} - s_{i,j}^{(s)}, \quad i = 1, \ldots, n_x, \quad j = 1, \ldots, n_v. \tag{2}$$

For simplicity of presentation, the system's state vector $x(t) \in \mathbb{R}^{n_x}$ comprises the species' concentrations[1] $[X_i]$ and is denoted by

$$x := ([X_1], [X_2], \ldots, [X_{n_x}])^\top \in \mathbb{R}^{n_x}. \tag{3}$$

The reaction kinetics are given by rate functions, which depend on the state x, time-invariant parameters $p \in \mathbb{R}^{n_p}$, and time-varying

[1] Concentrations are always greater than or equal to zero.

signals $w(t) \in \mathbb{R}^{n_w}$. In the context of biochemical networks, $w(t)$ can represent internal and/or external perturbations, changes of the cellular environment or stimuli, or control inputs such as temperature or added nutrients or substrates in fed-batch processes.

The transformation of the species is described by a vector of reaction rates that is given by

$$v(x,p,w) := (v_1(x,p,w), \ldots, v_{n_v}(x,p,w))^\top \in \mathbb{R}^{n_v}. \tag{4}$$

Typically, reaction rates are polynomial or rational expressions arising from the law of mass action or the Michaelis–Menten mechanism [1]. Some of the analysis frameworks are restricted to certain types of equations such as polynomial equations. In principle and under rather mild assumptions, other nonlinear expressions including quasi-polynomial kinetics and non-algebraic functions can be converted into polynomial form using state immersion (see references cited by [37,91,95]) or by functional approximation techniques such as Taylor series (see references cited by [68]).

The model outputs include measured variables and are also used to quantitatively describe the uncertain/robust behavior, or to qualitatively characterize robustness. Consider outputs of the form $y(t) = h(x,p,w) \in \mathbb{R}^{n_y}$ that are nonlinear functions of the states, parameters, and inputs. With this notation, the biochemical reaction network is concisely written as

$$\dot{x} = Sv(x,p,w) = f(x,p,w), \tag{5a}$$

$$y = h(x,p,w), \tag{5b}$$

where $\dot{x}$ refers to the ordinary derivative with respect to time and $f(\cdot)$ is introduced to simplify notation and is nonlinear in general.

Example *(Model definition).* In this work, a classical example for a biochemical network is used to illustrate and compare the different presented methods, namely the so called *reversible covalent modification system* [35]. The chemical reactions for this network are given by (cf. Eq. (1)):

$$\begin{aligned}
A + E_1 &\xrightarrow{p_1} C_1 \\
C_1 &\xrightarrow{p_2} E_1 + A^* \\
C_1 &\xrightarrow{p_3} A + E_1 \\
A^* + E_2 &\xrightarrow{p_4} C_2 \\
C_2 &\xrightarrow{p_5} E_2 + A \\
C_2 &\xrightarrow{p_6} A^* + E_2
\end{aligned} \tag{6}$$

Here enzymes E_1 and E_2 convert the protein between its two states A and A^* with intermediate complexes C_1 and C_2. Note that all stoichiometric coefficients are 1 in this example. The system's state and parameter vector are given by (cf. Eq. (3))

$$x := ([A],[A^*],[C_1],[C_2],[E_1],[E_2])^\top \in \mathbb{R}^6, \tag{7}$$

$$p := (p_1,p_2,p_3,p_4,p_5,p_6)^\top \in \mathbb{R}^6. \tag{8}$$

Applying mass-action kinetics gives the reaction rate vector (cf. Eq. (4))

$$v := \begin{bmatrix} p_1 x_1 x_5 \\ p_2 x_3 \\ p_3 x_3 \\ p_4 x_2 x_6 \\ p_5 x_4 \\ p_6 x_4 \end{bmatrix}. \tag{9}$$

With that, the model of the biochemical network (cf. Eq. (5a)) is given by:

$$\dot{x} = \begin{bmatrix} -1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & -1 & 0 & 1 \\ 1 & -1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 & -1 \\ -1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 1 & 1 \end{bmatrix} \begin{bmatrix} v_1(x,p) \\ v_2(x,p) \\ v_3(x,p) \\ v_4(x,p) \\ v_5(x,p) \\ v_6(x,p) \end{bmatrix}. \tag{10}$$

The output vector y depends on the question and will be introduced below.

The rest of this section considers some mathematical preliminaries that are needed to describe uncertainties in the models.

2.1. Conservation relations

A common feature in biochemical networks is conservation relations among the state variables x of the form

$$x_{T,j} = \sum_{i=1}^{n_x} l_{j,i} x_i, \quad j = 1, \ldots, n_c \tag{11}$$

with non-negative coefficients $l_{j,i}$ [41]. Such relations reduce the number of degrees of freedom and usually the system of differential equations is treated in its reduced form with $n_x - n_c$ state variables.

Example *(Conservation relations).* Under the assumption that there are no species removed ore added to the considered biochemical network (6), i.e. there is no inflow and outflow of mass, the following conservation relations exist:

$$A_T = [A] + [A^*] + [C_1] + [C_2] = x_1 + x_2 + x_3 + x_4 \tag{12}$$

$$E_{T,1} = [E_1] + [C_1] = x_5 + x_3 \tag{13}$$

$$E_{T,2} = [E_2] + [C_2] = x_6 + x_4. \tag{14}$$

With these, the model can be represented in the equivalent form with state and reaction vectors as follows:

$$x := ([A],[A^*],[C_1])^\top \in \mathbb{R}^3, \tag{15}$$

$$v := \begin{bmatrix} p_1 x_1 (E_{T,1} - x_3) \\ p_2 x_3 \\ p_3 x_3 \\ p_4 x_2 (E_{T,2} - A_T + x_1 + x_2 + x_3) \\ p_5 (A_T - x_1 - x_2 - x_3) \\ p_6 (A_T - x_1 - x_2 - x_3) \end{bmatrix}, \tag{16}$$

$$S := \begin{bmatrix} -1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & -1 & 0 & 1 \\ 1 & -1 & -1 & 0 & 0 & 0 \end{bmatrix}. \tag{17}$$

2.2. Discrete-time approximations

The discrete-time version of the continuous-time system (5) is employed in later sections. The representation

$$0 = F(x(k+1), x(k), p, w(k)), \tag{18a}$$

$$0 = H(y(k), x(k), p, w(k)), \tag{18b}$$

will be used to account for both implicit and explicit integration schemes, where $k \in \mathbb{N}$ is the time index with associated time points

$t_k \in \mathbb{R}$. The functions $F(\cdot)$ and $H(\cdot)$ represent the discrete-time versions of $f(\cdot)$ and $h(\cdot)$, respectively.

Example *(Discrete-time approximation).* Using a first-order explicit approximation of the time derivative, i.e.

$$\frac{d}{dt}x(t) = f(x(t), p, w(t)) \tag{19}$$

$$\approx \frac{x(t_{k+1}) - x(t_k)}{t_{k+1} - t_k} = f(x(t_k), p, w(t_k)), \tag{20}$$

the explicit Euler forward approximation is obtained:

$$0 = -x(t_{k+1}) + x(t_k) + (t_{k+1} - t_k) f(x(t_k), p, w(t_k)). \tag{21}$$

Here t_k denotes time at the kth sampling instance. Note that other, e.g. implicit, higher-order approximation schemes or even exact-discretization methods could be used to improve accuracy or numerical stability [102].

2.3. Linearization

For several subsequent analyses, the linearization of the nonlinear dynamical system (5) at a steady-state will be used. To shorten the notation, the steady state will be written as

$$\xi_{ss} := (x_{ss}^\top, p_{ss}^\top, w_{ss}^\top)^\top, \tag{22}$$

for fixed parameter values $p = p_{ss}$ and constant perturbations $w = w_{ss}$. The steady-states are given by

$$f(\xi_{ss}) = 0. \tag{23}$$

The linearization of (5) around (22) is then given by

$$\delta\dot{x} = A(\xi_{ss})\delta x + B_w(\xi_{ss})\delta w + B_p(\xi_{ss})\delta p, \tag{24a}$$

$$\delta y = C(\xi_{ss})\delta x + D_w(\xi_{ss})\delta w + D_p(\xi_{ss})\delta p, \tag{24b}$$

where the matrices

$$A(\tilde{\xi}) := \left.\frac{\partial f}{\partial x}\right|_{\tilde{\xi}}, \tag{25a}$$

$$B_w(\tilde{\xi}) := \left.\frac{\partial f}{\partial w}\right|_{\tilde{\xi}}, \tag{25b}$$

$$B_p(\tilde{\xi}) := \left.\frac{\partial f}{\partial p}\right|_{\tilde{\xi}}, \tag{25c}$$

$$C(\tilde{\xi}) := \left.\frac{\partial h}{\partial x}\right|_{\tilde{\xi}}, \tag{25d}$$

$$D_w(\tilde{\xi}) := \left.\frac{\partial h}{\partial w}\right|_{\tilde{\xi}}, \tag{25e}$$

$$D_p(\tilde{\xi}) := \left.\frac{\partial h}{\partial p}\right|_{\tilde{\xi}}, \tag{25f}$$

are evaluated at the point $\tilde{\xi}$, e.g. at the steady-state ξ_{ss}. This formulation also considers perturbations of the nominal parameter values p_{ss} by δp, through the matrices B_p and D_p.

3. Uncertainty descriptions and robustness specifications

Robustness is typically analyzed and quantified by determining the allowable uncertainties for which a desired system behavior or function characteristic can be guaranteed (see Fig. 1). In addition, uncertainty can enter the analysis from many sources such as the quantitative or even qualitative description of the robustness property or function characteristic. Other sources of uncertainty are associated with limited structural knowledge and data. For the model-based analysis of biological networks and for the development of therapeutic treatments, models must be plausible, predictive, and describe the observed robustness property. A major problem in relating predictions derived from a model-based analysis to the actual biological system are model and data uncertainties. Note that the term *data uncertainty* is used in a rather broad sense in this article, to refer not only to the sparse and uncertain measurement data, but also to qualitative observations and information from expert knowledge. It is important to point out that *model structure uncertainties*, as explained in the next subsection, are either due to inconsistencies of the model and the underlying system for which data has been gathered, or model structure uncertainties are simply a consequence of data uncertainty.

To be able to make correct statements and predictions despite these uncertainties, suitable methods that can capture and handle uncertainties must be selected, which are presented in Sections 4–6. This section presents the required formulation of the different uncertainties (see Fig. 2) and analysis questions, and illustrates the uncertainties with some observed biological

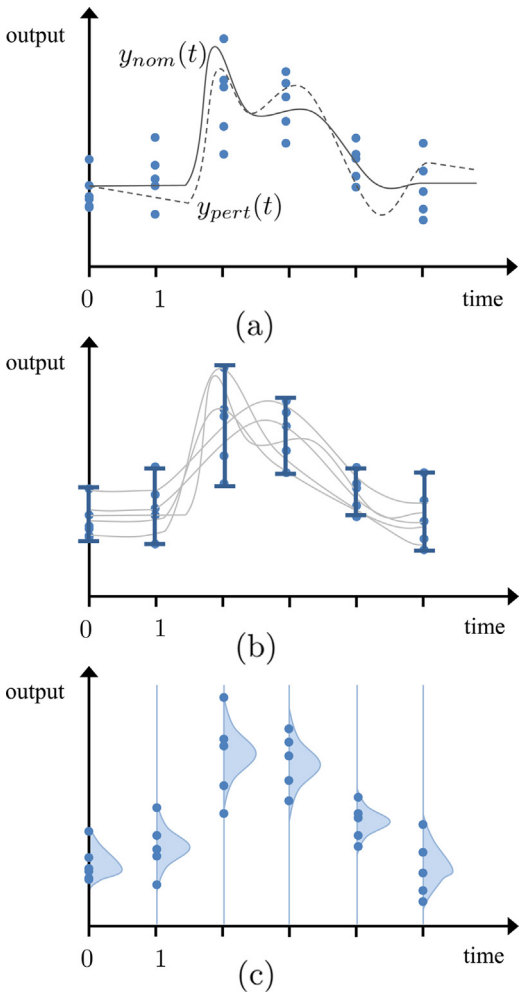

Fig. 2. Different quantitative data uncertainties at time points 0, 1, ... and system outputs/response. (a) Pointwise measurements at the time points. The curves show the effect of a parameter perturbation on the system output $y_{nom}(t)$ resulting in $y_{pert}(t)$. (b) A set-based or worst-case uncertainty description, which in this case was derived from (a) by the minimum and maximum measurement values. Multiple trajectories (gray solid curves) are consistent with the measurements. (c) Under certain conditions, if the number of samples is sufficiently large, then the probability density of the underlying uncertainty distribution may be assumed or derived.

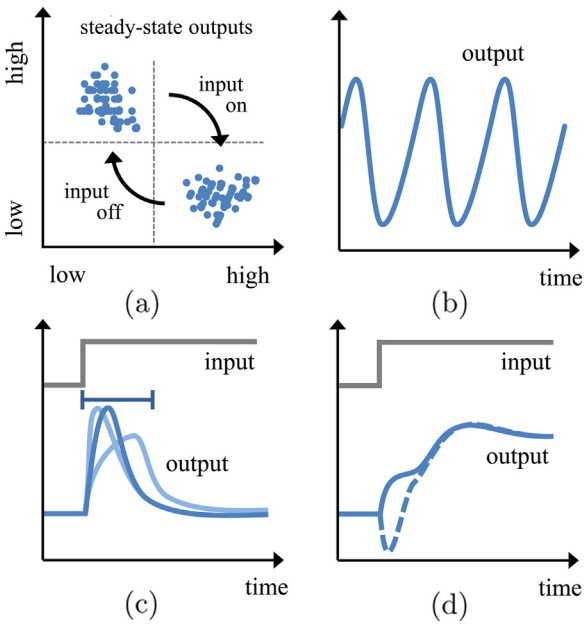

Fig. 3. Qualitative biological system behaviors. (a) Steady-state pattern in the presence or absence of inputs, i.e., stimuli. (b) Stable oscillations like a cell cycle or circadian rhythm. (c and d) Qualitatively different transient system outputs. In both cases, a persistent stimulus is applied as an input. In (c), the output transiently increases and then decreases again to converge to the prestimulus level even though the stimulus persists, which is called *adaptation* in biology. Due to a low sampling frequency, the time when the maximum occurs may not be known, which then corresponds to temporal uncertainties as shown by the horizontal bar. In (d), the output increases to an elevated steady-state level without (solid line) or with (dashed line) initial inverse response behavior.

behaviors and states commonly considered analysis questions (see Fig. 3).

3.1. Model structure uncertainties and model hypotheses

It is often unclear whether relevant species or reactions and interactions between species have been missed in the construction of a model for a biochemical reaction network. The structure of the models is uncertain in the sense that either the stoichiometric coefficients in S are not precisely known, some reactions $v(\cdot)$ might be absent or present, or the mathematical expression of the reaction rates (i.e., involved species and reaction kinetics) are unknown.

The uncertain structure often leads to different model hypotheses that need to be compared against the data. Models with structures that are inconsistent with the data can then be ruled out or may indicate that further cycles of iterative modeling and testing of consistency are required.

Methods to test model consistency while taking uncertainties into account are presented in Section 5, which are based on the comparison of model outputs with data. For comparison purposes, it is important to formulate the uncertain data such that a comparison can be made based on unbiased tests. These criteria can be formulated mathematically as *constraints* on the model outputs, which is discussed in detail below.

3.2. Set-based uncertainties

Due to limited measurement precision, low sampling frequency, and low number of experimental replicates (see Fig. 2a), it is not always possible to have precise knowledge of the actual probabilistic uncertainty distribution. Error bars, as depicted in Fig. 2b, derived from standard deviations or worst-case approximations, can be helpful. Such data correspond to uncertainty intervals or more generally, uncertainty sets. This set-based uncertainty description is also encountered when parameter values cannot be specified exactly. Indeed, parameters are usually highly uncertain and possible ranges can span several orders of magnitude. Methods to deal with set-based uncertainty descriptions are presented in Section 5.

3.2.1. Interval uncertainties

Often uncertainties are described by upper and lower bounds (see Fig. 2). Such bounds on uncertain parameters p or uncertain initial conditions x_0 are represented by

$$\begin{aligned} &\underline{p}_i \le p_i \le \overline{p}_i, \quad i = 1, \ldots, n_p, \\ &\underline{x}_{0,i} \le x_{0,i} \le \overline{x}_{0,i}, \quad i = 1, \ldots, n_x. \end{aligned} \tag{26}$$

Interval uncertainties are a special class of set-based uncertainties.

3.2.2. General set-based uncertainties

To describe such uncertainties more generally, assume that the variables such as parameters, states, inputs, disturbances or measurements, are all collected in the vector ξ and take values from a set defined by possibly nonlinear inequalities $g_i(\cdot) \ge 0$, i.e.,

$$\xi \in \mathcal{S}_\xi := \{\widetilde{\xi} \in \mathbb{R}^{n_\xi} : g_i(\widetilde{\xi}) \ge 0, i = 1, 2, \ldots, n_m\}. \tag{27}$$

Note that equalities can be handled by introducing $g_i(\cdot) \ge 0$ and $-g_i(\cdot) \ge 0$. The set-based uncertainty description also allows for the next type of data.

3.2.3. Parameter or data relations

This special class of set-based uncertainties is encountered when there are known relations between measurements of different outputs or of the same output at different time points. A common example is when an observed peak of a biological output is known to be twice as high after a stimulus compared to its value before the stimulus. In general, this results in relations between (unknown or uncertain) variables and can be represented as a manifold, which can be implicitly expressed by (27).

3.3. Probabilistic uncertainties

The set-based uncertainty description only describes the *possibility* for a parameter value, but does not make any statements about the *probability* that particular parameter values are taken. For certain cases, a set-based description is reasonable and less assumptive especially if only few replicates are available to make meaningful statements on probabilistic measurement uncertainties or parameter distributions. On the other hand, due to rapidly improving measurement and high-throughput techniques, statistics and thus probabilistic descriptions become more and more available.[2]

In case meaningful statistics or distributions can be derived, as in Fig. 2c, these uncertainties are referred to as *probabilistic uncertainties*, which usually can be characterized by (joint) probability densities PDF(ξ). Chance constraints are another way of handling probabilistic uncertainties [14,98], which can be formalized by

$$v \ge \mathrm{Prob}(\xi \in \mathcal{S}_\xi) := \int_{\mathcal{S}_\xi} \mathrm{PDF}(\xi) d\xi, \tag{28}$$

[2] When a probabilistic uncertainty description has a finite support, then it can be related directly to the set-based uncertainty description defined over the same finite support. For a probabilistic uncertainty description with infinite support, such as the Gaussian normal distribution, such relationships are only approximate.

where $\mathcal{S}_\xi$ is as in (27), and ν is the maximally allowed *risk level* associated with the considered (joint) chance constraint.

Methods that can handle probabilistic uncertainties are presented in Section 6.

3.4. Qualitative information on system behavior

Very often the system behavior of interest is not described by quantitative measures or values. Instead, so-called *qualitative information*[3] is available on how the system behaves (see Fig. 3). To analyze robustness of this kind of system behaviors requires a different definition, which is given in this subsection. A thorough formulation and description can be found in [103]. In many cases, the qualitative data directly follow from the lack of data and that absolute and quantitative measurement techniques are either not available or too expensive.

3.4.1. Steady-state location

Due to lack of affordable measurement techniques with fast sampling and the long time required in many biological experiments, only measurements of the steady-state before and after a perturbation or stimulus may be available (see Fig. 3a), in which case only a particular qualitative pattern of (some) outputs is available. Because quantification standards may not be available, the data are normalized, e.g. with respect to the maximum. Different levels are then defined such as *high* or *low* (see Fig. 3a) to verbalize the qualitative system behavior as, for instance, gene activation patterns [24].

This information can be written similarly as (27),

$$\xi_{ss} \in \mathcal{S}_{ss} := \{\widetilde{\xi} \in \mathbb{R}^{n_\xi} : f(\widetilde{\xi}) = 0, g_{ss,j}(\widetilde{\xi}) \geq 0, \quad j = 1, \ldots, n_{ss}\}. \tag{29}$$

The first constraint $f(\cdot) = 0$ is the steady-state condition, cf. (22), $g_{ss,j}(\cdot) \geq 0$ is the location information, and n_{ss} is the number of steady-state conditions. Alternatively, probabilistic uncertainties as in (28) could be used.

3.4.2. Stability, instability, and oscillation

In addition to the location of a steady-state, the dynamical properties of the biochemical reaction network are of interest. Typically, biological functions are related to certain qualitative dynamical behaviors such as multistability, i.e., the existence of several stable steady-states, limit cycle oscillations, or non-periodic oscillations (see Fig. 3b). Examples are the cell cycle, apoptosis, or circadian rhythm [1,113]. From a dynamical systems point of view, bistability, multistability, and sustained oscillations are related to stability and instability of an equilibrium point.

A steady-state ξ_{ss} is locally stable if all eigenvalues of the Jacobian A (see also (24)) have negative real parts. Here stability is meant in the sense of Hurwitz or local asymptotical stability, which can be checked by the characteristic equation:

$$q(s, \xi_{ss}) = \det(sI - A(\xi_{ss})) = \sum_{i=0}^{n_x} c_i(\xi_{ss}) s^i = 0. \tag{30}$$

The question then is whether the family of polynomials $q(s, \xi_{ss})$ with coefficients $c_i(\xi_{ss})$ is (robustly) stable for all uncertainties $\xi_{ss} \in \mathcal{X}_{ss}$, where $\mathcal{X}_{ss}$ is a subset of $\mathbb{R}^{n_\xi}$. Answering this question is equivalent to assessing whether $q(s, \xi_{ss}) \neq 0$ for all $s \in \mathbb{C}$ with $\mathrm{Real}(s) \geq 0$.

Other ways besides (30) exist to check Hurwitz stability, e.g., see [44]. For nonlinear systems, various conditions exist that can be used for the analysis of stability of nonlinear systems, e.g., see standard textbooks such as [56].

3.4.3. Adaptation and inverse response

Besides stability- and instability-related behaviors, the qualitative dynamical responses to perturbations of external conditions or stimuli are of interest (Fig. 3c). Most prominent examples where robustness has been observed *in vivo* and *in silico* are excitation and adaptation in bacterial chemotaxis [2,11] and archaeal phototaxis [121]. In this context, adaptation denotes the property in which an observed output initially changes in response to a stimulus, but then returns to the value before the stimulus, even though the stimulus persists. Adaptation is important to keep cells fit in changing environments by maintaining homeostasis under perturbations, or by expanding the dynamic range of sensory receptors. From a systems theoretic point of view, the conditions for exact (or perfect) adaptation have been extensively investigated (see [138,112,144] and references within): if a system adapts to a class of input signals, then it necessarily contains a subsystem that is capable of generating signals of this class, which is known under the term *internal model principle* [112,5].

For constant inputs, adaptation is obviously related to the steady-state gain of the system. The relation between the output deviation from steady-state and the stimulus deviation from steady-state in the Laplace domain is characterized by the transfer function

$$G(s) = C(\xi_{ss})(sI - A(\xi_{ss}))^{-1} B_w(\xi_{ss}) \tag{31}$$

(assuming $p = 0$ and $D_w(\xi_{ss}) = 0$) as computed from the matrices in (5), with the complex variable $s \in \mathbb{C}$. A system with scalar input $w(t)$ and scalar output $y(t)$ has perfect local adaptation at ξ_{ss} if and only if $G(s) = 0$ for $s = 0$. For $A(\xi_{ss})$ with no zero eigenvalues, which would hold for any system that is locally asymptotically stable about the steady-state, this condition is equivalent to

$$\det G(0) = 0. \tag{32}$$

For systems with uncertainty, local adaptation can hold only if this condition holds for all systems within the set of allowable uncertainties.

In a similar manner, *inverse response behavior* (see Fig. 3d) could be analyzed with the linearization and is related to unstable zero dynamics, i.e., zeros of the transfer function $G(s)$ in the open right-half plane. Inverse response behavior in living cells has been reported e.g. in [38].

Robustness or invariance of concentrations of particular species despite the variations of the concentrations of other species is another important class that has been considered, e.g., by [116,109]. Robustness or invariance of concentration can be more general and can also refer to invariance of the considered concentrations on short or long time scales. In contrast, adaptative system behavior is usually associated with an initial excitation, i.e., large changes of the concentrations of one or several species.

3.4.4. Conditional and temporal observations

Very often, only a limited amount of quantitative and temporally resolved data are available for model construction and estimation. Instead, only qualitative statements such as *if a stimulus is given, then the concentration increases transiently, before returning to its prestimulus level even though the stimulus persists*. Such information is often also provided by experts and experimentalists from

[3] Actually, mixed classes of *qualitative* and *quantitative* data are usually encountered.

biological insight or knowledge. Such constraints can be formally captured using conditional statements such as

$$\begin{aligned} &\text{IF} \quad (g_1(\xi) \geq 0 \quad \text{AND} \quad g_2(\xi) \geq 0) \quad \text{OR} \\ &\quad (g_3(\xi) \geq 0 \quad \text{AND} \quad \ldots) \quad \text{OR} \quad \ldots \\ &\quad\quad \text{THEN} \quad (g_4(\xi) \geq 0 \quad \text{AND} \quad \ldots). \end{aligned} \tag{33}$$

Generally, the conditions can be constraints involving variables at the same or even different time points, which allows the formulation of *temporal uncertainties*, meaning that the time point an event happens is not exactly known. This situation is illustrated in Fig. 3c where the time point when the transient response peaks is temporally uncertain.

As shown by [103], such a formulation allows the description of many qualitative observations, biological knowledge, and data using Boolean logic [100,54,13]. The robustness of such qualitative behaviors, such as having oscillations or transient peaks, is most often of interest.

3.5. Problem statements for robustness analysis

Robustness analysis can be informally stated as *quantification of the perturbations that a system can tolerate before losing a specific function* [58,74,115,80,90]. More formally, the remainder of this article considers the question of robustness analysis as (see also Fig. 1):

1 *Quantification of robustness by estimation* of parameter sets or distributions that are consistent with uncertain data (Sections 3.2–3.3) and qualitative specifications of system behavior (Section 3.4).
2 *Prediction of uncertainty propagation* and how uncertainties affect robustness and output specifications.

The volume of the consistent parameter set is an immediate measure of robustness. As illustrated and discussed by [21], however, a large volume of the robust parameter set does not by itself imply large robustness. For instance, the set might be very thin where a small perturbation into one direction of the parameter space leads to a loss of function, while the system can still be robust for perturbations in other directions (see also Fig. 1). Thus, the geometry and topology of the robust set contain very important information on robustness [21].

Sections 4–6 present different methods to tackle these questions for quantitative dynamical models of biochemical reaction networks. The main focus is on set-based methods that provide guarantees, and methods that allow probabilistic statements with reasonable computational cost.

4. Analysis of local perturbations and network structures

This section reviews methods from systems and control theory that are suited to analyze the influence and propagation of uncertainties in biochemical reaction networks. In particular, *perturbations* of variables such as the parameters are considered, which then lead to perturbations of the *nominal system behavior*. The output trajectory resulting from the perturbation is denoted by $y_{pert}(t)$, as illustrated in Fig. 2.

The presented methods allow a characterization of steady-state solutions and a quantification of robustness by the distance between nominal parameters and the parameters for which constraints on the qualitative or quantitative behavior are violated (see Fig. 1). While these methods are conceptually simple and well-known, their applicability for the analysis of the usually large uncertainties is limited, because these methods focus on nominal points that are often assumed to be invariant to the perturbations, or because the methods rely on linearizations which are valid approximations only for small perturbations.

In addition to these methods, methods are reviewed that allow statements about the qualitative behavior to be drawn from the nominal network structure or its perturbation. These methods are, at least to a large extent, independent of parameters. Due to this fact, the methods are well-suited for qualitative analyses, but are limited for quantitative predictions.

4.1. Sensitivity analysis

One question is how the nominal dynamical system behavior or the steady-state changes in response to perturbations of the parameters.

For reviews of applications and the general methods, see, e.g., [148,125,47,106]. Sensitivity analysis provides a good starting point to identify the parameters and corresponding key factors that have strong impact on the output. This analysis can provide valuable insights about how robust the biological responses are with respect to parameter changes. In general, sensitivity analysis methods can be classified as *local* and *global* as detailed further below. The sensitivity with respect to probabilistic uncertainties is addressed in Section 6.

4.1.1. Local sensitivity analysis and metabolic control analysis

Local sensitivity analysis concentrates on a nominal point in the parameter space, such as a nominal operating condition or steady-state ξ_{ss}. To approximate the perturbed output trajectory $y_{nom}(t)$ (see Fig. 1) in the case of a small perturbation δp_j of parameter p_j, the first-order sensitivity system can be used, which is obtained from the linearization (24) of (5) in which the input perturbation δw is set to zero:

$$\frac{d}{dt}\frac{\partial x}{\partial p_j} = A(\xi(t))\frac{\partial x}{\partial p_j} + B_{p_j}(\xi(t)), \tag{34a}$$

$$\frac{\partial y}{\partial p_j} = C(\xi(t))\frac{\partial x}{\partial p_j} + D_{p_j}(\xi(t)). \tag{34b}$$

The sensitivity equations (34) describe a linear time-varying system of ordinary differential equations (ODEs) in which the time-dependent matrices $A(\xi(t))$, $B_{p_j}(\xi(t))$, $C(\xi(t))$, $D_{p_j}(\xi(t))$ (cf. (25)) are evaluated along the nominal trajectories $\xi(t)$ of the states, parameters and disturbances.

The sensitivity of the steady-state ξ_{ss} can be obtained from the steady-state solution of (34) if A is Hurwitz:

$$\frac{\partial x}{\partial p_j} = -A(\xi_{ss})^{-1}B_{p_j}(\xi_{ss}) \tag{35}$$

and

$$\frac{\partial y}{\partial p_j} = -C(\xi_{ss})A(\xi_{ss})^{-1}B_{p_j}(\xi_{ss}) + D_{p_j}(\xi_{ss}), \tag{36}$$

where the matrices are evaluated at ξ_{ss}. The latter equation can then be used to extrapolate the shift of the steady-state output δy_{ss} for finite perturbations δp_j:

$$\delta y_{ss} = (-C(\xi_{ss})A(\xi_{ss})^{-1}B_{p_j}(\xi_{ss}) + D_{p_j}(\xi_{ss}))\delta p_j. \tag{37}$$

Note that instead of the *absolute sensitivity* (35), often scaled or normalized sensitivities such as $(p_{ss,j}/x_{ss})(\partial x/\partial p_j)$ are more meaningful since they measure the relative change of the steady-state [106].

Local sensitivity analysis provides due to the underlying linearization a first-order approximation of the effect of parameter

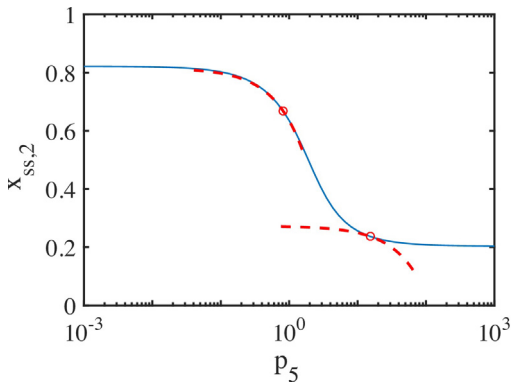

Fig. 4. Illustration of the approximation of the steady state of $[A^*]$, i. e., $x_{ss,2}$, for different values of p_5. Red dashed lines indicate the predicted steady states determined from the linear sensitivity analysis.

Table 1
Nominal parameterization.

Parameter	Value [a.u.]	Parameter	Value [a.u.]
p_1	10	A_T	1
p_2	10	$E_{T,1}$	0.1
p_3	1	$E_{T,2}$	0.2
p_4	10		
p_5	1		
p_6	1		

perturbations. Thus, the local parametric sensitivity analysis results should be used with care for the prediction of large perturbations, as illustrated in Fig. 4. Note that several extensions have been presented, such as for oscillating systems [148,141,129].

A prominent tool for sensitivity analysis of biochemical networks is *Metabolic Control Analysis*, see e.g. [53]. Within this framework, the so called elasticity and control coefficients together with the connectivity and summation theorems provide useful means to analyze relationships between changes in system variables with respect to changes of the parameters. It has been shown in [48] that Metabolic Control Analysis is closely related to standard sensitivity analysis as presented above, but with the addition of constraints resulting due to the summation and connectivity theorems, as well as the explicit decomposition of systems behavior into the component parts.

Example *(Local sensitivity analysis).* For the reduced model (16)–(17), the parameterization given in Table 1 was chosen. With help of (24)–(25) the linearization was determined in dependence of the steady-state ξ_{ss} and the parameter p_6. The sensitivities of the states are derived according to (35) and depicted in Fig. 4 for state x_1 for several choices of p_6. Additionally, the figure depicts the dependency of the steady state $x_1 = A$ in dependence of k_6. As one can see the local sensitivity information can be misleading, since relatively small changes in p_6 change drastically the steady state value $\xi_{ss,1}$, although the sensitivity might suggest otherwise. For this reason, higher-order sensitivities (e.g., [46,20,118] and global sensitivity methods have been considered. The latter are described in the next subsection. Note that the steady state behavior of the model under uncertainties is studied further below.

4.1.2. Global sensitivity analysis

Global sensitivity analysis aims to predict model behavior either for larger parameter values, or for local sensitivities averaged over a domain in parameter values. Often statistical methods are used to guide the sampling of values from within the specified domains in the parameter space. An introduction to global sensitivity analysis is available from [84].

In [117,125] a global sensitivity analysis method based on an input-output control engineering view was presented. The idea employs a combination of observability and controllability Gramians (also see [111]), the so-called *cross Gramian*, and an empirical extension for nonlinear systems. While this approach allows larger parameter perturbations to be considered, the statements are still made with respect to a nominal operating point.

4.2. Bifurcation analysis

Bifurcation analysis based on numerical continuation has often be used to measure robustness in cases when only few parameters are assumed uncertain and varied [58,80,90]. As discussed by [136], a major limitation is that bifurcation surfaces (see Fig. 1) can usually not be computed explicitly in a high-dimensional parameter space. In addition, continuation methods may miss parts of the bifurcation surface, even if only one or two parameters are uncertain. To deal with uncertainty in multiple parameters, structured singular value has been suggested as an analysis tool [110,58,80]. However, most approaches based on the structured singular value do not directly take the uncertainty in the location of the steady-state due to parameter variations into account (e.g., see discussion by [135]). More recently, researchers have shown how to use structured singular analysis to assess the effects of parametric uncertainties on the qualitative character of the local dynamics [69].

The authors of [87] use normal vectors on manifolds of critical points to measure the distance between these manifolds and equilibrium solutions. This approach allows the characterization of an equilibrium solution by their parametric distance to manifolds at which the behavior of the system changes qualitatively, i.e., bifurcation points, or points at which state variable constraints or output constraints are violated. Statements are only made with respect to a nominal operating point.

4.3. Kinetic perturbations

A kinetic perturbation is a modification of the network's reaction rate vector from $v(\cdot)$ to $\tilde{v}(\cdot)$. The key restriction is that the steady-state reaction rates should remain unchanged, i.e., $v(\xi_{ss}) = \tilde{v}(\xi_{ss})$. Kinetic perturbations change the slope of the reaction rates at a steady-state, but leave the nominal steady-state and the stoichiometric matrix unchanged. Defining a suitable perturbation matrix $\Delta \in \mathbb{R}^{n_r \times n_x}$, the change in the slope of the reaction rate at steady-state ξ_{ss} is given by

$$\delta v := \frac{\partial \tilde{v}}{\partial x}(\xi_{ss}) - \frac{\partial v}{\partial x}(\xi_{ss}) = \operatorname{diag}(\tilde{v}(\xi_{ss}))\Delta \operatorname{diag}(\xi_{ss})^{-1}.$$

Considering the linear approximation (24) of the network at steady-state, the Jacobian is

$$\tilde{A}(\delta v) = S\frac{\partial v}{\partial x}(\xi_{ss}) + \operatorname{diag}(\tilde{v}(\xi_{ss}))\Delta \operatorname{diag}(\xi_{ss})^{-1}.$$

Using the notion of kinetic perturbations, [137] studied the robustness problem of finding a perturbation δv such that $\tilde{A}(\delta v)$ has eigenvalues on the imaginary axis, i.e., where the qualitative behavior of the system changes.

In a similar approach, [138] investigated the adaptation problem for a network with scalar input and output. Then the adaptation problem is to find Δ such that (32) is satisfied and $\tilde{A}(\delta v)$ is Hurwitz. Both the robustness and adaptation problem are solved by robust control techniques [147].

4.4. Qualitative behavior and its dependence on network structure and feedback loops

In principle, using parametric sensitivity (as discussed in previous subsections) as a measure of robustness is based on the assumption that the underlying model structure is exactly known and that all relevant perturbations can be represented by changes in the model parameters. However, model structures are in general uncertain due to incomplete knowledge of the reaction kinetics, neglected intermediate reaction steps, and unmodeled transport phenomena such as diffusion and delays [49].

A large number of methods are available to analyze the robustness of the qualitative system behavior with respect to perturbation or addition/removal of interactions between the species in the network. Below is a review of a selection of methods for structural network analysis.

4.4.1. Structural robustness

The authors of [50,130,49] used linear systems analysis, transfer functions, and structured singular values to analyze perturbations that affect the model structure, and [51] considered structural uncertainty as unmodeled dynamics and used transfer function analysis to compute robustness with respect to structural changes in reaction networks. In addition, it was shown that robustness analysis can be used to validate/invalidate a hypothesized model structure and to detect structural fragilities.

4.4.2. Monotone systems

For certain classes of biochemical reaction networks, model structures (or reaction structures) can be related to dynamical system properties such as multi-stability, in which the network can alternate between mutually exclusive states over time due to small perturbations in parameters or noise. For monotonic reaction rates, [22] analyzed associated graphical models to characterize structures of biochemical reaction networks in terms of dynamical properties including multi-stability and convergence to an equilibrium point. Monotonicity of reaction rates means that, for the reaction network dynamics (5), the nonlinear function $v(x, p)$ corresponding to a reaction rate vector satisfies the relations

$$\frac{\partial v_i(x,p)}{\partial x_j} = \begin{cases} \geq 0 & \text{if } s_{i,j}^{(s)} > 0, \quad \forall i = 1,\ldots,n_r, \\ = 0 & \text{if } s_{i,j}^{(s)} = 0, \quad \forall j = 1,\ldots,n_x, \end{cases}$$

for all p.[4] These results follow from extensive studies of monotone systems [28,6,7].

In [132,64,65] the authors discovered that the existence of positive and negative feedback loops in a reaction network plays a key role in robustifying a dynamical reaction network system against both parametric and structural perturbations.

5. Set-based uncertainties and analysis

The methods presented in Section 4 enable the analysis of perturbations of a small number of parameters, and for the most part do not easily allow the consideration of more general set-based uncertainties. In the analysis of biochemical networks, however, it is important to consider simultaneous perturbations and uncertainties in all parameters, and to derive rigorous enclosures of all solutions for iterative modeling or classification of motifs. This section presents methods that can deal with set-based uncertainties and provide bounds on all solutions (see Fig. 2b), as well as how to employ set-based approaches for robustness analysis.

5.1. Interval analysis

Interval analysis was introduced by [88] as an approach to bound rounding and truncation errors in mathematical computations. Due to its general simplicity and computational efficiency, as well as many sophisticated improvements, interval analysis has gained much attention. Several reviews have been published [89,42,52], with many discussing applications to robust prediction, estimation and control.

This approach considers interval uncertainties (26), which result in different possible output trajectories. Guaranteed bounds on the outputs,

$$\underline{y}_i \leq y_i \leq \overline{y}_i, \quad i = 1,\ldots,n_y, \tag{38}$$

can be computed using interval functions and interval arithmetic. Interval arithmetic is a logical extension of standard arithmetic. Operations like addition and subtraction are simply defined by operations on the lower and upper bounds.

Interval analysis is a simple but very computationally efficient technique with numerous applications in mathematics, computer science, and engineering (see reviews cited above), but the computation of the tightest possible interval solution set that completely contains all solutions is difficult, primarily due to correlations among uncertain variables. This difficulty is discussed in more detail next.

5.1.1. Stability analysis using interval matrices

Stability and instability are related to many biological functions as highlighted in Section 3.4. Consider a system (5) with box-shaped parametric uncertainties,

$$\mathcal{P} := [\underline{p}_1, \overline{p}_1] \times \cdots \times [\underline{p}_{n_p}, \overline{p}_{n_p}], \tag{39}$$

as defined in (26). A standard approach for robustness analysis of nonlinear systems in the presence of such parametric uncertainties is to examine the linearization (24) and its associated characteristic equation. In mass action networks, the parametric uncertainties appear affinely in the entries of the Jacobian A, denoted by $A(p)$, and consequently polynomially in the coefficients $c_i(p)$ of the each term of the characteristic equation (30).

A common approach to bound $c_i(p)$ is to relax the polynomial dependencies and assume each $c_i(p)$ as an independent uncertainty interval. Lower and upper bounds for $c_i(p)$ can then be computed using interval arithmetic and the polynomial becomes consequentially an interval polynomial. In 1978, Kharitonov proposed a theorem on robust Hurwitz stability of interval polynomials that reduces the stability analysis to the stability of four deterministic polynomials, where each polynomial corresponds to a certain combination of extreme values $\underline{c}_i(p)$ and $\overline{c}_i(p)$ of each coefficient [57].

Interval polynomials and the well-known Kharitonov theorem is thus an approach for the stability analysis of biochemical reaction networks with interval uncertainties. Albeit this approach is elegant, a drawback is that it neglects the parameter correlation in each coefficient $c_i(p)$. In general, algorithms that are based on positivity of the Hurwitz determinant associated with (30) suffer from the fact that the order of the polynomials $c_i(p)$ grow polynomially in the number of states, n_x, and in the order of the entries in $A(p)$. This approach can lead to high-order polynomials $c_i(p)$, and neglecting the parameter correlations by an over approximation using interval analysis is highly conservative. In realistic reaction networks, usually this conservatism is too large for this approach to produce

[4] Since monotonicity of reaction rates are required for all p in the set, this property might be called *robust* monotonicity.

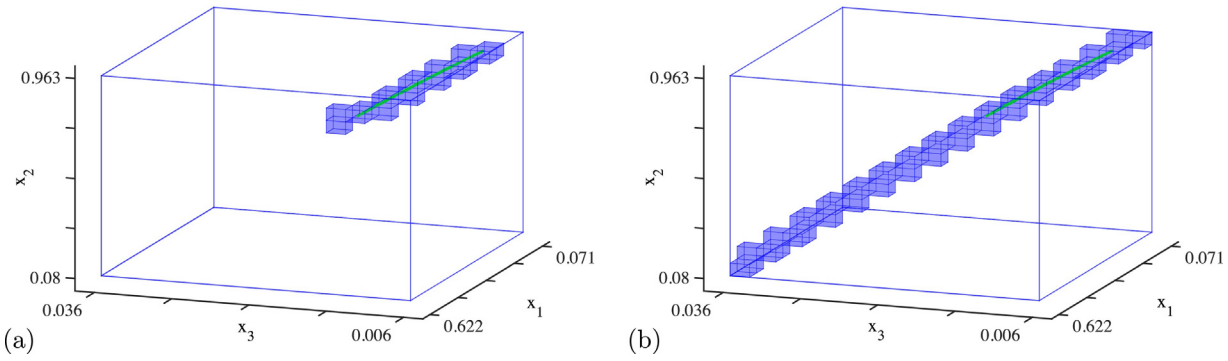

Fig. 5. The figure depicts the derived outer-bounds on the steady-states if the parameters p_4 and p_6 are unknown-but-bounded, i.e. $p_4 \in [3.33, 30]$, $p_6 \in [0.33, 3]$: (a) Feasibility formulation, (b) interval analysis. Both results were obtained with the toolbox ADMIT, in case of the feasibility formulation the SDP was solved with MOSEK 7.0. Green dots represent Monte Carlo samples uniformly taken from the previously given intervals and $x \in [0, 1]^3$.

useful stability analysis results for the original nonlinear dynamical system.

It is therefore important to apply other appropriate methods that take care of the dependence of the parameters to remove conservativeness due to correlation of the uncertain parameters whenever possible (for specific numerical examples, see [34] and citations therein). Root locus is a general method but too computationally expensive when the number of parameters is large. The generalized Kharitonov [12] method is more general but still only applies to limited parameter dependencies. Numerous other approaches for robust Hurwitz stability that handle more general polynomial parameter dependencies have been presented in the systems and control communities, for example, see [146] and references within. In the following, other set-based approaches for stability analysis and for addressing other questions stated in Section 3 are presented. These methods reduce this conservatism by improved relaxations.

5.2. Linear and semidefinite relaxations

The general idea behind the methods presented next is to construct a *feasibility problem* (or sometimes called a *constraint satisfaction problem*) and to derive the set of all feasible solutions, that is, not just one feasible solution. This feasibility problem is a set of nonlinear equations and inequalities, concisely written as

$$\begin{aligned} \text{find} \quad & \xi \\ \text{subject to} \quad & g_i(\xi) \geq 0, \quad i = 1, \ldots, n_g \end{aligned} \tag{40}$$

where the vector $\xi = (x, p, w)$ contains all time-varying and time-invariant variables that appear in the problem. The constraints $g_i(\xi)$ are used to represent the nonlinear dynamics (5) after a suitable time discretization on a finite-time horizon $t_0, t_1, \ldots, t_{n_t}$, as well as all set-based uncertainties and relation of variables and other information on the outputs in the form of (27).

Due to nonlinearities and nonconvexities, the solution set of (40) cannot be derived directly. Therefore, relaxations into linear and semidefinite feasibility problems are applied. The basic idea of relaxations is to replace nonlinearities with simpler expressions. For example, a linear relaxation introduces variables that are linear in the relaxation (or lifting) variables. The resulting relaxed problems can be solved efficiently, and due to this relaxation procedure, each solution of the original nonlinear feasibility problem is also a solution of the relaxed feasibility problem. The converse is, however, not true.

Interval analysis (see Section 5.1) follows the same idea and relaxes a nonlinear expression by over approximating it by an interval. As discussed, interval-based relaxations can be quite conservative. The following approach produces, in our experience, tighter bounds [123].

With the assumptions of real-valued,[5] bounded, and non-negative[6] variables ξ, as well as polynomial or rational expressions[7] $g_i(\xi) \geq 0$, the constraints $g_i(\cdot)$ in (40) can be reformulated in terms of a symmetric, rank-one matrix $X = \xi\xi^\top$ composed of monomials needed to represent the inequalities. This representation enables the nonconvexities to be captured in form of a rank-one condition on X. By relaxing the condition $\text{rank}(X) = 1$ with the weaker constraint $X \succeq 0$, a convex semidefinite program (SDP) is obtained:

$$\begin{aligned} \text{find} \quad & X \\ \text{subject to} \quad & \text{trace}(Q_i X) \geq 0, \quad i = 1, \ldots, n_{g_{SDP}} \\ & X \succeq 0 \\ & X_{1,1} = 1 \end{aligned} \tag{41}$$

where $X_{1,1}$ is the element in the first row and column of matrix X. The trace constraints (where $n_{g_{SDP}} \geq n_m$) are obtained from rewriting the constraints $g_i(\cdot)$ in terms of X, and the $X_{1,1} = 1$ accounts for constant terms in the inequalities.[8]

To deal with larger biochemical reaction networks with more constraints and variables, the SDP can be relaxed to a linear program (LP) by replacing $X \succeq 0$ by the weaker constraints $X \geq 0$ and by assuming symmetry of X. More details can be found elsewhere [123,103,125].

The relaxed SDP or LP contains information about the model dynamics (5) and the set-based uncertainties (27). In general, this approach is very flexible and allows different robust analysis questions to be tackled as done further below. Formulation of feasibility problems with associated constraints on data and models with set-based uncertainties, the involved relaxation steps, and the respective solution of the problems by outer approximations can be performed in ADMIT, which is a freely available toolbox [123].

5.2.1. Structural uncertainties and model invalidation

To prove that a model hypothesis, reformulated as a feasibility problem (40), is inconsistent with some qualitative behavior or

[5] Section 5.3 shows that mixed real and integer-valued variables can be handled if qualitative (e.g., conditional and temporal) constraints are needed.

[6] Due to the boundedness of the variables, the non-negativity of the variables poses no limitation and can be obtained by suitable translation.

[7] Assuming polynomial expressions poses few limitations, because different solutions exist to convert the rational and transcendental into polynomial form, see Section 2.

[8] Usually additional constraints, such as McCormick's relaxation of bilinear terms like $\xi_i = p_1 p_2$, are added to tighten the solution set as discussed and detailed by [119].

with output constraints despite set-based uncertainties, it is necessary to determine if a solution for the feasibility problem (40) exists. An efficient approach [103] in this case is to consider the Lagrangian dual formulation of the semidefinite or linear relaxation. The weak duality theorem and the relaxation process guarantee that, if the objective of the dual program is unbounded, then (40) does not admit a solution and hence is inconsistent. This approach enables entire families of models to be ruled out and thus deals efficiently with structural uncertainties.

5.2.2. Uncertainty propagation and parameter estimation

The feasibility approach also enables the derivation of outer approximations of consistent parameter sets or state variable sets and can therefore be used to address the robust estimation of states and parameters. Outer approximations of uncertain variables can be obtained if the feasibility problem is replaced by an optimization problem in which the single variables are minimized or maximized. A tighter lower bound of variable ξ_j can be obtained by the formulation

$$\begin{aligned} &\text{infimum} \quad \xi_j \\ &\text{subject to} \quad g_i(\xi) \geq 0, \quad i = 1, \ldots, n_g \end{aligned} \tag{42}$$

followed by application of the same relaxations as above. In this way, box-shaped outer approximations can be easily determined on state variables of parameters.

Outer approximations provide an intuitive measure of robustness—the larger the volume of the outer approximation, the more robust the system. As discussed in Section 3.5, robustness measures purely based on volume should not be overinterpreted because the shape of the robust set also contains important robustness information. The shape can be approximated by partitioning the initial uncertainties into regions that are independently checked for feasibility (that is, whether each region contains a solution or not). With suitable recursive algorithms, this approach allows the derivation of an outer approximation of the variables of interest. Using the partitioning approach, guaranteed bounds on steady-states were derived by [39]. The approach has been used for set-based parameter estimation [101] and the analysis of system responses to perturbations [125]. In the following two examples the steady states and the consistent parameters are analyzed.

Example *(Set-based analysis of steady-states).* Interval analysis (Section 5.1) and the feasibility problem formulation (Section 5.2) were employed to derive outer-bounds on the steady states of model (15)–(17) for unknown-but-bounded parameters p_4 and p_6. For p_4 and p_6 the intervals $[3.33, 30]$ and $[0.33, 3]$ are chosen, respectively. The results of both methods are depicted in Fig. 5 as well as some Monte Carlo samples for comparison. Both estimations were obtained with ADMIT [123]. It should be noted that the focus of ADMIT is the formulation and solution of feasibility problems and, therefore, it might be possible to improve the interval analysis results by employing a more specialized implementation.

Example *(Set-based parameter estimation).* Using the feasibility formulation from Section 5.2 outer-bounds on the parameters were derived for given measurement data. The measurement data was obtained by simulating the model (15)–(17) with the parameterization provided in Table 1 and the initial condition $x(0) = [0.5\ 0.25\ 0.1]^\top$. Afterwards, a relative error of 1% was added to the simulated trajectory to get the upper and lower bounds of the measurement intervals. To estimate the parameters, model (15)–(17) was discretized employing the implicit Euler integration with a step size of 0.5. The implementation was done in the toolbox ADMIT. The result of the parameter estimation procedure is depicted in Fig. 6, when in addition to the initial condition five measurement intervals are considered. The involved semidefinite programs were solved with the solver SEDUMI.

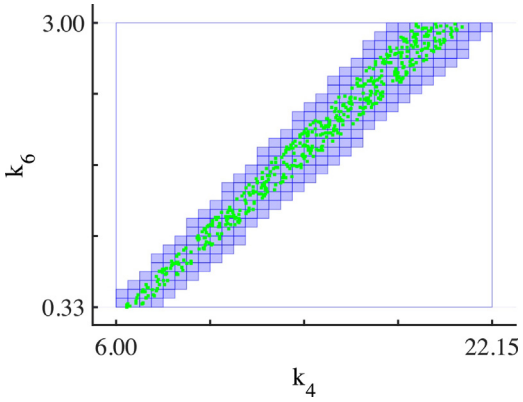

Fig. 6. Outer-approximation of the consistent parameters employing the feasibility formulation. Results obtained with the toolbox ADMIT using the solver SEDUMI. Green dots represent Monte Carlo samples leading to consistent trajectories.

5.2.3. Stability and instability

Biological networks should perform their functions robustly in spite of the presence of uncertainties. To check the robustness of stability- and instability-related behaviors, (30) and its associated Hurwitz determinant could be used to test local asymptotic stability, similar as for the interval analysis approach, but this approach suffers from similar conservatism as the interval polynomial approach.

Linear programming can be employed in a feedback loop-breaking approach to obtain conditions for the non-existence of local bifurcations under a parametric uncertainty [136]. The conditions are checked computationally by applying the Positivstellensatz, which is relaxed into a linear program similar as above. A solution to the linear program yields a robustness certificate for the considered dynamical behavior and lower robustness bound corresponding to a level of parametric uncertainty up to which no local bifurcations can occur. A less conservative alternative approach for identifying parameter sets in which local bifurcations cannot exist is based on structured singular value analysis [69].

5.3. Mixed-integer linear relaxations

Semidefinite and linear relaxations enable the determination of *outer approximations* of parameters and state variables consistent with formulated *quantitative robustness constraints*. The methods presented in this subsection use mixed-integer programming[9] and allow the determination of *inner approximations* and the analysis of *qualitative information* (which was introduced in Section 3.4).

5.3.1. Inner approximations

Robustness can be quantified by the volume of the outer approximation of the consistent sets, but this measure can be misleading for two reasons. First, the outer approximation is conservative and usually contains inconsistent parameterizations. Second, even if the outer approximation is tight, the consistent parameter set might not be simply connected such as discussed by [21] (and also discussed in Section 3.5). An important analysis problem is to find *inner approximating sets* for which it is guaranteed that all parameter combinations from this set lead to consistent solutions. This

[9] Mixed-integer programming can handle also allows discrete values for parameters, states, and input variables.

determination will then help to elucidate the geometry and topology of the consistent parameter set.

A method to determine inner approximations was proposed in [124]. It borrowed ideas from and built on a framework presented by [103] that allows qualitative information to be considered. In this approach, the constraints $g_i(\xi) \geq 0$ in (27) are reformulated by associating a binary variable b_i with each constraint such that each binary variable is constrained to be true if and only if the associated inequality is satisfied. This approach leads to an equivalent mixed-integer nonlinear programming problem. An inner approximation is then obtained by adding a logical combination using *OR*-statements and by checking if the feasibility problem has a solution or not. Sets of parameters or initial conditions for which the feasibility problem provides no solution then give an inner approximation. The relaxation of the mixed-integer nonlinear feasibility problem into a mixed-integer linear feasibility problem allows the inner approximations to be determined efficiently.

5.3.2. Including qualitative information in robustness analysis

Qualitative, temporal, and conditional statements of the general form given in (33) were considered in [103]. In general, *AND* combinations of constraints (e.g., $g_1(\xi) \geq 0$ *AND* $g_2(\xi) \geq 0$) can be readily combined with any quantitative information because such constraints can simply be included in the feasibility problem. *OR*-combinations (e.g., $g_1(\xi) \geq 0$ *OR* $g_2(\xi) \geq 0$), on the other hand, have to be treated differently. Similar to the inner approximation idea, one approach is to introduce binary variables (e.g., b_1) with constraints imposed such that the binary variables have values of *true* (i.e., equal to 1) if a condition (e.g., $g_1(\xi) \geq 0$) is satisfied. By suitably combining the binary variables in linear constraints, the statements associated with g_1 and g_2 can be enforced to be true, which then allows the estimation of consistent parameters.

The inclusion of integer variables is computationally demanding, but very versatile. By suitable formulation of qualitative information and statements, using Boolean algebra and reformulation in terms of the mixed-integer program, many different types of uncertainties can be considered such as temporal uncertainties (i.e., the statement associated with $g_a(\xi)$ is true at time t_1 *OR* at time t_2 *OR* ...), or more general logical combinations (statement 1 *OR* statement 2 is true if statement 3 is true). This approach has been used for the robustness analysis of adaptation networks [103].

5.4. Methods for continuous-time dynamical systems

Continuous-time systems can be addressed within the presented framework by discrete-time approximations, e.g., obtained by numerical integration. Due to the discretization error, the consistent parameter sets of continuous-time and discrete-time model do not necessarily overlap and, thus, wrong conclusions on model robustness are possible [102]. One assumption that still allows the direct application of the presented methods is to assume that the time derivatives of the state variables are available as measurements [31], but this assumption often does not hold in practice.

Interval analysis methods that can deal with continuous-time dynamics rely on higher-order Taylor approximations and are usually termed *verified integration* (for example, see [79,94,15]). The tightness of the over approximations produced by these methods heavily depend on the underlying integration algorithms. The over approximations typically grow in conservatism over time and can become arbitrarily conservative at intermediate to late times, even to the point of growing exponentially fast over time for problems in which the volume of the true set is constant. Verified integration can easily be used for prediction of uncertainty propagation, but analysis considerations such as stability can be difficult to assess.

There exist other methods that employ semidefinite programming and the related *Sum-of-Squares* approach, such as *barrier certificates* [4,96,97] and *occupation measures* [122]. These approaches allow the continuous-time dynamics to be considered directly without numerical integration and allow the certification of parameter regions as inconsistent with data and model. Occupation measures have been used [122] to derive both polynomial inner- and outer-approximations of the consistent parameter sets for continuous-time nonlinear systems.

An advantage of approaches based on barrier certificate and occupation measure is that these approaches directly incorporate the right-hand side of the differential equations between the measurement points without using any approximation by time discretization. However, both approaches are computationally demanding for realistic systems since the number of decision variables increases quickly as the number of variables and relaxation order increases (see discussion and references in [122]).

5.5. Approaches using skewed structured singular values

As stated in Section 2, many biochemical reaction networks can be described by polynomial or rational functions. Such systems together with parametric uncertainties, noise, and disturbances can be expressed in the form of a linear fractional transformation (LFT) by using block diagram algebra or multidimensional realization algorithms, assuming that the input-output mapping is well-defined for the nominal system [147,40,82]. The resulting LFT consists of two terms: a term for the nominal system and a term for the uncertain/varying portion of the system.

5.5.1. Bounds on uncertainty propagation

If a system output is expressed by an LFT, the system's parameter set and the output set can be related by the skewed structured singular value along with the scaled main loop theorem [29,105,147]. This theorem states that the maximum norm of the LFT over an uncertainty set is the corresponding skewed structured singular value. With this theorem, the skewed structured singular value can be used to analyze the effects of uncertainties on the output of systems with very general types of parameter dependence, which can reduce conservatism compared to other methods.

To see how the parametric uncertainties propagate to the output, consider a discrete-time system with a scalar output (18), which corresponds to continuous-time system (5). In order to derive lower and upper bounds on the output variables $y(k)$, an LFT that expresses the output $y(k)$ is constructed using a constant matrix N, and a perturbation matrix Δ that contains information on the interval uncertainties.[10] For systems with parametric uncertainties, Δ is a block-diagonal matrix where each block is a scalar times the identity matrix.

By using the main loop theorem, the lower and upper bounds on the output can be expressed as

$$\underline{y}(k) = -\nu_{\Delta}(N_l) + M, \tag{43a}$$

$$\overline{y}(k) = \nu_{\Delta}(N_u) - M, \tag{43b}$$

where $M > 0$ is any large enough scalar value, and N_l and N_u are constant matrices coming from the matrix N of the LFT and M. Variations that balance the conservativeness and computational complexities of the algorithm have been developed [72].

In this approach, once the system is written in LFT form, the computational difficulty and robustness are the same as that of computing lower/upper bounds on the skewed structured singular values, which can be computed in polynomial time. For further discussions, see [72]. The computational cost of bounds on the

[10] For systems with ellipsoidal uncertainties, the skewed spherical structured singular value should be used instead of the skewed structured singular value [67,70].

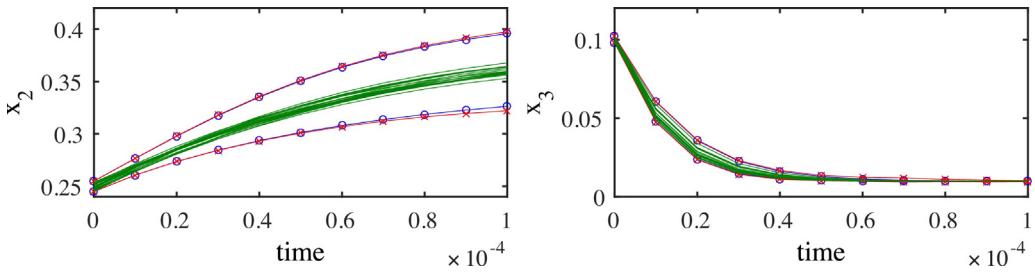

Fig. 7. The figures depict the derived outer-bounds on the trajectories of the states x_2 and x_3 with parameters and initial conditions given in Table 2. The trajectories for x_1 are not shown for lack of space. Blue bounds are based on ν lower bounds, and red bounds are based on ν upper bounds. Green lines represent trajectories generated by Monte Carlo samples uniformly taken from the uncertainty set.

Table 2
Parameters.

Known parameters	Value	Uncertain parameters	Value range
A_T	1	p_1	[90,000, 110,000]
$E_{T,1}$	0.01	p_4	[45,000, 55,000]
$E_{T,2}$	0.01	p_5	[900, 1100]
p_2	1000	$x_1(0)$	[0.49, 0.51]
p_3	1	$x_2(0)$	[0.245, 0.255]
p_6	1	$x_3(0)$	[0.098, 0.102]

skewed structured singular values depends on the size of the LFT. For this purpose, it is desired to employ an LFT of minimal order. Although the construction of a minimal LFT is equivalent to a minimal multidimensional realization and is non-trivial, there exist order reduction algorithms [83,23,36,104] that can be applied.

Example *(Set-based estimation of state trajectories).* In this example, the approach based on the skewed structured singular value discussed in Section 5.5.1 is employed to derive outer-bounds on the state trajectories of the model (15)–(17). The model (15)–(17) is first discretized employing the explicit Euler integration as in Example 3 with a step size of 0.00001, and the bounds on the state trajectories are obtained using (43) for the discretized model with the parameterization and the set of initial conditions provided in Table 2. The results are depicted in Fig. 7. The computation of the value of ν is NP-hard [17], and its polynomial-time upper and lower bounds were obtained with the Skew Mu Toolbox (SMT) [30].

5.5.2. Inner approximation

The skewed structured singular value can also be used to find a parameter set that is consistent with the given specification of the system output/response. A simple approach consists of two steps for the output expressed by an LFT. The first step is to decide the *shape* of the box of the allowable uncertain set (e.g., the ratio of δp_1 and δp_2 in Fig. 1). The shape can be chosen based on the relative expected variations in the parameters, the magnitude of the nominal parameters, or other methods [68]. This step is necessary because the skewed structured singular value gives only one value, and cannot compute multiple values at once. In the second step, the skewed structured singular value is used to stretch or shrink the chosen box equally in all directions to find a maximum volume uncertain set that guarantees the satisfaction of the specification on the system output/response [68].

As an alternative to the two-step method by skewed structured singular values, the problem of finding a consistent parameter set can also be solved by using D,G-K iteration via reformulation of the problem as a constant-matrix μ synthesis problem. This reformulation allows the simultaneous optimization over the parameter center, box shape, and box size [71].

6. Probabilistic uncertainties and analysis

As more measurements can be made for biochemical experiments as a result of technological advancements, parameters are more often reported in probabilistic form, which increases the importance of methods that analyze the influence and propagation of probabilistic uncertainties (see Fig. 2c). Two classes of methods for the propagation and analysis of probabilistic uncertainties are summarized in this section. The first class of methods uses an approximation of the Liouville equation and its solution using the Fokker-Planck equation. The second class of methods employs polynomial chaos theory, which expands system responses to probabilistic uncertainties in appropriate polynomial basis functions, which are determined by the uncertainty distributions. Approximating system responses by these polynomial chaos expansions (PCEs) has the advantage of decreasing computation time compared to Monte Carlo simulation which requires the solution of the system equations for each sample. The two classes of methods are described below.

6.1. Approximate analysis using the Fokker-Planck equation

Analysis of the effects of randomness (i.e., probabilistic uncertainty) of the initial concentrations and kinetic reaction parameter variations can be studied by solving a partial differential equation (PDE) known as the *Fokker-Planck* or *Liouville equation* for a given probability distribution of initial concentration profiles and kinetic parameters. It is usually very computationally expensive to compute an exact solution for such PDEs, and there have been made many efforts to either avoid solution of the PDE or compute an approximate solution instead. Such methods for stochastic stability analysis and sensitivity analysis are described next.

6.1.1. Stochastic stability analysis

Stochastic stability of two gene regulatory networks that were built and analyzed in the biological literature were analyzed in [26]. In this work, standard theorems in the literature [76] were specialized and enabled the analysis of stochastic stability using Lyapunov functions, which is much simpler than direct analysis of the Fokker-Planck equation for the probability distribution. [64] characterized the dynamics of a motif consisting of interlinked fast and slow positive feedback loops, which regulate polarization of budding yeast, calcium signaling, Xenopus oocyte maturation, and other processes [18]. Interest in this motif as a component in synthetic genetic networks is that it provides a dual-time switch that can be rapidly and reliably induced while being relatively insensitive to noise in the stimulus [18]. [64] discussed how the expressions derived from this approach can be used to *design* robust biological gene switch circuits that perform programmed desired behaviors in the presence of intrinsic and extrinsic perturbations.

6.1.2. Sensitivity analysis

The *Trapezoid Rule for Adaptive Integration of Liouville dynamics* (TRAIL) to analyze the propagation of randomness (distribution) in nonlinear dynamical systems was developed and applied in [43]. This technique consists of two steps: (a) *prediction*, in which linear ordinary differential equations (ODEs) corresponding to reaction dynamics and the associated Fokker-Planck equations are used, and (b) *correction*, in which nonlinear effects are treated to refine the accuracy.

The reaction dynamics (called macroscopic dynamics) was linearized with respect to a steady-state in [27]. Also the covariance of a solution of the corresponding Fokker-Planck equation in terms of a solution of Lyapunov equations was computed. The method was shown to efficiently estimate sensitivity criteria and analyze the effects of the elimination of fast variables (that is, removing fast subsystems, which corresponds to unmodeled dynamics).

6.2. Polynomial chaos theory

This section considers a stochastic spectral method for propagation and quantification of probabilistic uncertainties, called *polynomial chaos theory*. Introductory tutorials are available on the use of polynomial chaos theory that emphasize the application to systems and control problems [66,93,92] or focus on sensitivity analysis [126]. This section provides a tutorial introduction to the ideas of polynomial chaos theory with some simple examples of biochemical reaction network systems.

6.2.1. Propagation of probabilistic uncertainty

The polynomial chaos method was first introduced by [139] for turbulence modeling for uncertainties that are Gaussian random variables, which was later extended to other random variables [8,143]. The underlying idea is to use a spectral decomposition for which solutions for differential equations with stochastic uncertainties in the initial conditions and parameters in an infinite-dimensional probability space are projected onto a finite-dimensional subspace spanned by a set of certain polynomial basis functions. Selection of the types of basis functions depends on the types of random variables whose stochasticity is to be propagated, and the methods of projections determine ways of computing the associated coefficients of basis functions that minimize an approximation error. The optimal choice of basis functions have been derived for all of the commonly used probability density functions [143].

Example *(Uncertainty propagation using PCE).* To illustrate the use of PCE methods for uncertainty propagation, consider the biochemical reaction network model underlying cAMP oscillations observed in chemotactic *Dictyostelium discoideum* cells [80]:

$$\frac{dx}{dt} = \begin{bmatrix} k_1x_7 - k_2x_1x_2 \\ k_3x_5 - k_4x_2 \\ k_5x_7 - k_6x_2x_3 \\ k_7 - k_8x_3x_4 \\ k_9x_1 - k_{10}x_4x_5 \\ k_{11}x_1 - k_{12}x_6 \\ k_{13}x_6 - k_{14}x_7 \end{bmatrix} \tag{44}$$

where

$$p = [k_1, k_2, \ldots, k_{14}]^\top$$

is the vector of parameters, and the state vector

$$x = [x_1, \ldots, x_7]^\top$$

contains the concentrations of seven proteins: [ACA], [PKA], [ERK2], [REG A], [Internal cAMP], [External cAMP], and [CAR1], respectively. Suppose that the only source of uncertainty is in the kinetic parameters[11] described by a normal distribution with mean 0.8 and standard deviation 0.1:

$$k_6 \sim \text{norm}(0.8, 0.1).$$

The states $x_i(t\,; k_6)$ were approximated by PCEs of the form

$$\hat{x}_i(t;\theta) = \sum_{j=1}^{m_i} \xi_{ij}(t)\phi_j(\theta), \tag{45}$$

where

$$\theta \sim \text{norm}(0,\, 1),$$

i.e., $k_6 = 0.1\theta + 0.8$, and ϕ_j are Hermite polynomial basis functions, which are optimal for normal distributions for the parameters. This example uses a collocation method to determine the coefficients in the PCE. A least-squares fit was used to determine $\xi_{ij}(t)$, which minimizes

$$[x_i(t; k_6) - \hat{x}_i(t;\theta)]^2 \tag{46}$$

for m_i distinct values of θ that are roots of ϕ_{m_i+1}. Fig. 8 shows the average and variance of [ACA] computed from an 8*th*-order PCE ($m_i = 9$) and from Monte Carlo simulation (10^4 samples). Compared to Monte Carlo simulation, which takes 342.7 seconds, PCE decreased the computational time by a factor of more than 500, to 0.644 seconds. As shown by Fig. 8, PCE agrees with Monte Carlo for short-time behavior; however, as time progresses, disagreement in the variance increases. This result suggests that PCE should be applied to oscillatory systems with caution, which has motivated the development of modified PCE methods that are more accurate for such systems [78].

The PCE methods require the determination of the associated coefficients $\xi_{ij}(t)$. In addition to the least-squares fit, there are several methods available for the determination of the coefficients of a PCE. For the details of various methods of PCE coefficient determination, the readers are referred to the research monographs [77,142].

In addition to quantifying uncertainty propagation in dynamical systems, PCE methods have been used to facilitate solving Bayesian inference problems [81,86]. Direct Monte Carlo simulations are replaced by PCEs to efficiently approximate the likelihood functions associated with the Bayesian rule of computing posterior distributions. Computational efficiency and approximation accuracy of using the PCE methods for solving Bayesian inference problems have been demonstrated with genetic circuit models [85] and a diffusion process [86].

6.2.2. Steady-state probability distributions

In the presence of randomness of kinetic parameters, the steady-state of a locally asymptotically stable biochemical reaction network has a probability distribution in which the probability density function of the state or the output converges as time goes to infinity. For analysis of the steady-state of the output y_{ss}, a goal is to compute or estimate such a converging probability distribution. One approach to estimate the steady-state distribution is to use a sampling-based method such as Monte Carlo simulation, which is computationally inefficient, especially for analysis of a large-scale biochemical reaction network. A spectral method based on PCE can

[11] PCE methods can be also used for uncertainty propagation and quantification of uncertain initial conditions and/or inputs in straightforward ways, see [60–62] for technical details.

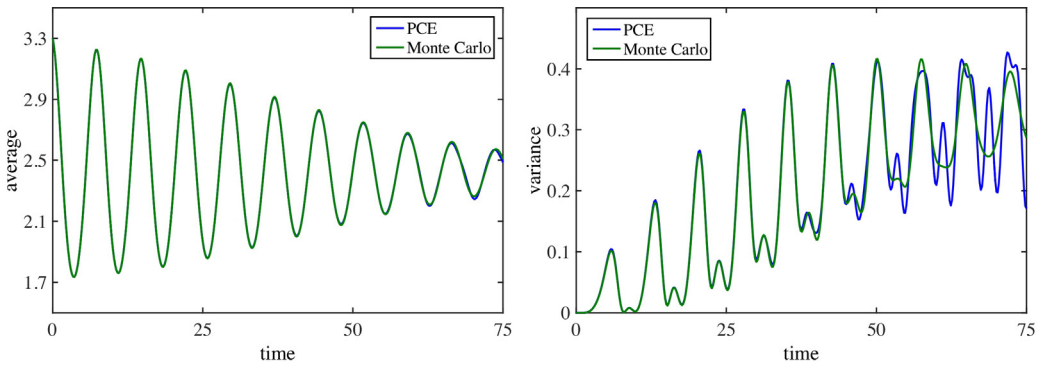

Fig. 8. Average and variance of the state x_1 computed from an 8th-order PCE and from Monte Carlo simulation for 10^4 samples (computation time: 0.644 seconds for PCE and 342.7 seconds for Monte Carlo simulation).

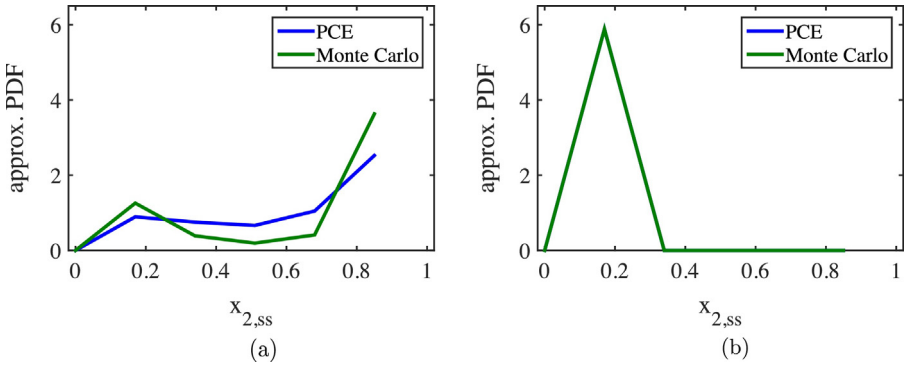

Fig. 9. PCE approximation of 3rd-order for the distribution of steady-state $x_{ss,2} = [A^*]$ if the uniform uncertainty in p_5 is (a) from 10^{-6} to 10^3, or (b) from 10^2 to 10^3.

be used as a computationally efficient alternative to analyze the probability distribution of the steady-state.

Example *(Approximation of the probability density function (PDF) of the steady-state behavior using PCE).* Using the initial example, a PCE based on the Legendre polynomials was used to approximate the probability density function of steady-state $[A^*]$ as a result of uniformly distributed uncertainty in p_5. Fig. 9a shows that the PDF approximated with a 3rd-order PCE deviates significantly from that approximated with 10^3 Monte Carlo simulations for $10^{-6} \leq p_5 \leq 10^3$. This is due to the nonlinear dependence of steady-state $[A^*]$ on the value of p_5, see Fig. 4. As the uncertainty range decreases, PCE of the same order becomes a better approximation for the actual PDF. When $10^2 \leq p_5 \leq 10^3$, the PDF approximated using a 3rd-order PCE is very close to the Monte-Carlo-simulation based approximation (Fig. 9b).

6.2.3. Stochastic stability

Section 6.1 discusses methods to rigorously analyze stochastic stability without approximation by employing Lyapunov theory proven to be rigorous using the Fokker-Planck equation for the probability distribution of the states. Alternative approaches for analysis of stochastic stability have been proposed based on the PCE approximation of the probability distribution. For example, [32,33] presented stability tests for linear and polynomial stochastic systems that contain random variables. The tests were based on PCEs for which the intrusive Galerkin projection method was used to obtain deterministic dynamical system equations for the associated coefficients, and the convergence of the coefficients was used to define notions of stochastic stability, for which Lyapunov methods were applied. As another example, [45] applied PCEs to estimate short-term statistics and stability of a solution trajectory in the presence of random parameters and initial conditions with known probability distributions.

However, the stability analysis using PCE methods presented by [32,33] is very limited. Most importantly, the notion of stability was very restrictive, requiring the probability distribution of a state solution to converge to a Dirac delta function that peaks at the origin in some probabilistic convergence sense. However, the main purpose of stability analysis of a stochastic dynamical system is to determine *whether there exists* a stationary probability distribution to which distribution of a state solution converges [26,76]. This issue can be corrected by employing the three steps:

1 compute/check the existence of equilibrium points of an extended state-space model corresponding to an ODE for the coefficients of the polynomial chaos expansion,
2 rewrite the ODE in terms of the deviations from the equilibrium points, and
3 test stability of the equilibrium points by using Lyapunov analysis.

Checking the existence of an equilibrium point in Step 1 should be considered as *approximate*[12] existence of a stationary probability distribution in the sense of convergence in distribution, by implying there exists a random vector with well-defined probability distribution such that its moments have a convergent stationary point for the associated stochastic differential equation.

[12] PCE methods provide an approximation, not an exact solution, for the associated stochastic differential equation.

6.2.4. Scalable uncertainty quantification

The exact and ϵ-approximate computation of worst-case bounds on parameters or variables for polynomial or rational systems is NP-hard [17,16,63], which implies that it is very unlikely that an algorithm exists that can perform such computations with a computation time that is a polynomial function of the number of parameters. The exact computation of probability distribution functions for parameters or variables for polynomial or rational systems would also be expected to be NP-hard. Since biochemical reaction networks are often large-scale interconnected dynamical systems, the associated sensitivity analysis and uncertainty quantification can have very high computational cost as a function of the number of parameters.

Such computational complexity results only indicate, however, that *there exist* polynomial or rational systems whose robustness analysis is computationally expensive, and do not indicate that the robustness analysis of *all* polynomial or rational systems is difficult. A natural approach to development of scalable robustness analysis is to exploit the interconnection (or reaction network) structure and to decompose the whole system into many several subsystems. Due to the importance of developing scalable robustness analysis for large-scale systems, there have been some algorithmic efforts for model decomposition that facilitate the use of PCE by parallelizing the associated computations, for which the underlying interconnection structure is used. [127,9,128] used a graph theory-based method of spectral clustering for decomposition. For an indirected graph, each state variable is associated with a node, and an edge connecting the node pair (i, j) is weighted by the absolute sum of time average (i, j) and (j, i) elements of the Jacobian. The eigenvalues and eigenvectors of the associated graph Laplacian can be computed in a distributed manner [55] and groups of distinct eigenvalues determine clusters of decomposition [134]. Although such algorithmic approaches have been demonstrated with some examples, the mechanism is still opaque in the sense that the performance bounds are not obtainable and algorithmic stability cannot guaranteed in general. Furthermore, the aforementioned approaches are for general large-scale systems, not restricted to biochemical reaction networks. Biochemical reaction networks have some useful properties such as positivity and boundedness of state variables. Since (at least some of) biochemical reaction networks share many similarities with Markov chains, model reduction techniques for Markov chains can be used for model reduction of biochemical reaction kinetics, e.g., aggregation-based model reduction hidden Markov chains (see [3,75,133] and references therein for details).

7. Discussion and conclusions

This article presents an overview of various approaches to model and analyze robustness in biochemical reaction networks. Usually, *robustness analysis* aims to quantify the perturbations for which a network loses or gains some qualitative behavior. This challenge can be addressed by prediction and estimation methods as illustrated in Fig. 1.

When constructing or working with mathematical models of biochemical networks, there are not only a large amount of uncertainties, but also significantly different classes of uncertainties such as set-based, qualitative, or probabilistic uncertainties. These uncertainties typically result from limited or expensive measurement techniques, precision, and sampling frequency. Moreover, the descriptions of the network functions or characteristics analyzed for robustness are also inherently uncertain. Robustness analysis is hence inevitably linked to uncertainty analysis.

Due to the different types of encountered uncertainties, the analysis method must be chosen carefully. Moreover, it is important to examine the influence of the uncertainties on the model response and consequently to assess the validity of the analysis conclusions. We feel that the different approaches have not been reviewed systematically so far, especially within the context of biochemical reaction networks. Therefore, this article first defines the different uncertainty classes and, secondly, formally states the different robustness analysis questions of interest. Finally, recent developments for the uncertainty and robustness analysis are presented.

Besides a quantification of robustness, we expect that in the future quantitative predictions of responses to perturbations and uncertainty propagation will become increasingly important, for instance, in therapy design and synthetic biology. This article presents methods that are—from our perspective—especially suited for robustness analysis, prediction, and estimation. For consistency, the focus was on a particular class of methods and models that we feel can be used to deal with the mentioned challenges. The presentation is weighted towards set-based deterministic methods and stochastic methods based on the Fokker-Planck equation and polynomial chaos theory, with minimal discussion of Monte Carlo sampling methods or local analyses.

This perspective paper also touches on advances being made in robust estimation and prediction and the design of biochemical reaction networks, also known as *synthetic biology*. Substantial research is still needed in the development of practically implementable algorithms, especially for larger networks. A natural approach to the development of scalable methods is to exploit the interconnection (or reaction) structure and to decompose the whole system into many several subsystems that are analyzed in isolation, for example, see [25].

Acknowledgements

Financial support is acknowledged from the Edwin R. Gilliland Professorship at the Massachusetts Institute of Technology.

References

[1] B. Aguda, A. Friedman, Models of Cellular Regulation, Oxford University Press, Oxford, United Kingdom, 2008.
[2] U. Alon, M.G. Surette, N. Barkai, S. Leibler, Robustness in bacterial chemotaxis, Nature 397 (6715) (1999) 168–171.
[3] B.D.O. Anderson, The realization problem for hidden Markov models, Math. Control Signals Syst. 12 (1) (1999) 80–120.
[4] J. Anderson, A. Papachristodoulou, On validation and invalidation of biological models, BMC Bioinf. 10 (2009), Art. No. 132.
[5] B. Andrews, E.D. Sontag, P. Iglesias, An approximate internal model principle: applications to nonlinear models of biological systems, in: Proceedings of the 17th IFAC World Congress, Seoul, Korea, 2008, pp. 11–60.
[6] D. Angeli, J.E. Ferrell, E.D. Sontag, Detection of multistability, bifurcations, and hysteresis in a large class of biological positive-feedback systems, Proc. Natl. Acad. Sci. U. S. A. 101 (7) (2004) 1822–1827.
[7] D. Angeli, E.D. Sontag, Multi-stability in monotone input/output systems, Syst. Control Lett. 51 (3) (2004) 185–202.
[8] R. Askey, J. Wilson, Some basic hypergeometric orthogonal polynomials that generalize Jacobi polynomials, Mem. Am. Math. Soc. 54 (319) (1985) 1–55.
[9] A. Banaszuk, V.A. Fonoberov, T.A. Frewen, M. Kobilarov, G. Mathew, I. Mezic, A. Pinto, T. Sahai, H. Sane, A. Speranzon, A. Surana, Scalable approach to uncertainty quantification and robust design of interconnected dynamical systems, Annu. Rev. Control 35 (1) (2011) 77–98.
[10] A.-L. Barabási, Z.N. Oltvai, Network biology: understanding the cell's functional organization, Nat. Rev. Genet. 5 (2) (2004) 101–113.
[11] N. Barkai, S. Leibler, Robustness in simple biochemical networks, Nature 387 (6636) (1997) 913–917.
[12] B.R. Barmish, A generalization of Kharitonov's four-polynomial concept for robust stability problems with linearly dependent coefficient perturbations, IEEE Trans. Autom. Control 34 (2) (1989) 157–165.
[13] A. Bemporad, M. Morari, Control of systems integrating logic, dynamics, and constraints, Automatica 35 (3) (1999) 407–427.
[14] A. Ben-Tal, L. El Ghaoui, A. Nemirovski, Robust Optimization, Princeton University Press, Princeton, NJ, 2009.
[15] M. Berz, K. Makino, Verified integration of ODEs and flows using differential algebraic methods on high-order Taylor models, Reliab. Comput. 4 (4) (1998) 361–369.

[16] R.D. Braatz, E.L. Russell, Robustness margin calculation for large scale systems, Comput. Chem. Eng. 23 (1999) 1021–1030.
[17] R.P. Braatz, P.M. Young, J.C. Doyle, M. Morari, Computational complexity of μ calculation, IEEE Trans. Autom. Control 39 (5) (1993) 1000–1002.
[18] O. Brandman, J.E. Ferrett Jr., R. Li, T. Meyer, Interlinked fast and slow positive feedback loops drive reliable cell decisions, Science 310 (5747) (2005) 496–498.
[19] E. Bullinger, R. Findeisen, D. Kalamatianos, P. Wellstead, System and control theory furthers the understanding of biological signal transduction, in: I. Queinnec, S. Tarbouriech, G. Garcia, S.-I. Niculescu (Eds.), Biology and Control Theory: Current Challenges, Springer, Berlin/Heidelberg/New York, 2007, pp. 123–135.
[20] M. Cascante, A. Sorribas, R. Franco, E.I. Canela, Biochemical systems theory: increasing predictive power by using second-order derivatives measurements, J. Theor. Biol. 149 (4) (1991) 521–535.
[21] M. Chaves, A. Sengupta, E.D. Sontag, Geometry and topology of parameter space: investigating measures of robustness in regulatory networks, J. Math. Biol. 59 (3) (2009) 315–358.
[22] G. Craciun, C. Pantea, E.D. Sontag, Graph-theoretic analysis of multistability and monotonicity for biochemical reaction networks, in: H. Koeppl, G. Setti, M. di Bernardo, D. Densmore (Eds.), Design and Analysis of Biomolecular Circuits, Springer, New York, 2011, pp. 63–72.
[23] R. D'Andrea, S. Khatri, Kalman decomposition of linear fractional transformation representations and minimality, in: Proceedings of the American Control Conference, IEEE Press, Piscataway, NJ, 1997, pp. 3557–3561.
[24] H. De Jong, Modeling and simulation of genetic regulatory systems: a literature review, J. Comput. Biol. 9 (1) (2002) 67–103.
[25] D. Del Vecchio, A.J. Ninfa, E.D. Sontag, Modular cell biology: retroactivity and insulation, Mol. Syst. Biol. 4 (1) (2008), Art. No. 161.
[26] H. El Samad, M. Khammash, Stochastic stability and its application to the analysis of gene regulatory networks, in: Proceedings of the IEEE Conference on Control and Decision, IEEE Press, Piscataway, NJ, 2004, pp. 3001–3006.
[27] J. Elf, M. Ehrenberg, Fast evaluation of fluctuations in biochemical networks with the linear noise approximation, Genome Res. 13 (11) (2003) 2475–2484.
[28] G. Enciso, E.D. Sontag, Monotone systems under positive feedback: multistability and a reduction theorem, Syst. Control Lett. 54 (2) (2005) 159–168.
[29] G. Ferreres, A Practical Approach to Robustness Analysis with Aeronautical Applications, Kluwer Academic, New York, 1999.
[30] G. Ferreres, J.-M. Biannic, J.-F. Magni, A skew mu toolbox (SMT) for robustness analysis., in: IEEE International Symposium on Computer Aided Control Systems Design, 2004, pp. 309–314.
[31] D. Fey, E. Bullinger, Limiting the parameter search space for dynamic models with rational kinetics using semi-definite programming, in: Proceedings of the 11th IFAC Symposium on Computer Applications in Biotechnology, Leuven, Belgium, 2010, pp. 150–155.
[32] J. Fisher, R. Bhattacharya, Stability analysis of stochastic systems using polynomial chaos, in: Proceedings of the American Control Conference, IEEE Press, Piscataway, NJ, 2008, pp. 4250–4255.
[33] J.R. Fisher, Stability analysis and control of stochastic dynamic systems using polynomial chaos (Ph.D. dissertation, Texas A&M University, College Station, TX, 2008.
[34] L. Goh, M. Kishida, R.D. Braatz, On the analysis of robust stability of metabolic pathways, IEEE Control Syst. 32 (4) (2012) 92–94.
[35] A. Goldbeter, D.E. Koshland, An amplified sensitivity arising from covalent modification in biological systems, Proc. Natl. Acad. Sci. U. S. A. 78 (11) (1981) 6840–6844.
[36] R. Gunawan, E.L. Russell, R.D. Braatz, Comparison of theoretical and computational characteristics of dimensionality reduction methods for large scale uncertain systems, J. Process Control 11 (2001) 543–552.
[37] E.J. Hancock, A. Papachristodoulou, Generalised absolute stability and sum of squares, Automatica 49 (4) (2013) 960–967.
[38] R. Hartmann, D. Oesterhelt, Bacteriorhodopsin-mediated photophosphorylation in *Halobacterium halobium*, Eur. J. Biochem. 77 (2) (1977) 325–335.
[39] J. Hasenauer, P. Rumschinski, S. Waldherr, S. Borchers, F. Allgöwer, R. Findeisen, Guaranteed steady state bounds for uncertain (bio-)chemical processes using infeasibility certificates, J. Process Control 20 (9) (2010) 1076–1083.
[40] S. Hecker, A. Varga, J.-F. Magni, Enhanced LFR-toolbox for MATLAB, Aerosp. Sci. Technol. 9 (2005) 173–180.
[41] R. Heinrich, S. Schuster, The Regulation of Cellular Systems, vol. 416, Chapman & Hall, New York, 1996.
[42] Y. Hijazi, H. Hagen, C.D. Hansen, K.I. Joy, Why interval arithmetic is so useful, in: H. Hagen, M. Hering-Bertram, C. Garth (Eds.), Visualization of Large and Unstructured Data Sets, Gesellschaft für Informatik (GI), Bonn, Germany, 2007, pp. 148–163.
[43] I. Horenko, S. Lorenz, C. Schütte, W. Huisinga, Adaptive approach for nonlinear sensitivity analysis of reaction kinetics, J. Comput. Chem. 26 (9) (2005) 941–948.
[44] R.A. Horn, C.R. Johnson, Topics in Matrix Analysis, Cambridge University Press, New York, 1991.
[45] F.S. Hover, Application of polynomial chaos in stability and control, Automatica 42 (5) (2006) 789–795.
[46] J.T. Hwang, Sensitivity analysis in chemical kinetics by the method of polynomial approximations, Int. J. Chem. Kinet. 15 (10) (1983) 959–987.
[47] B. Ingalls, Sensitivity analysis: from model parameters to system behaviour, Essays Biochem. 45 (2008) 177–194.
[48] B.P. Ingalls, H.M. Sauro, Sensitivity analysis of stoichiometric networks: an extension of metabolic control analysis to non-steady state trajectories, J. Theor. Biol. 222 (1) (2003) 23–36.
[49] E.W. Jacobsen, G. Cedersund, On parametric sensitivity and structural robustness of cellular functions – the oscillatory metabolism of activated neutrophils, in: Proceedings of the IEEE Conference on Decision and Control and European Control Conference, IEEE Press, Piscataway, NJ, 2005, pp. 3681–3686.
[50] E.W. Jacobsen, G. Cedersund, Structural robustness of biochemical network models – with application to the oscillatory metabolism of activated neutrophils, IET Syst. Biol. 2 (1) (2008) 39–47.
[51] E.W. Jacobsen, V. Nenchev, Robustification as a tool in modeling biochemical reaction networks, J. Process Control 21 (10) (2011) 1517–1525.
[52] L. Jaulin, M. Kieffer, O. Didrit, E. Walter, Applied Interval Analysis: With Examples in Parameter and State Estimation, Robust Control and Robotics, Springer Verlag, Berlin/Heidelberg/New York, 2001.
[53] H. Kacser, J.A. Burns, The molecular basis of dominance, Genetics 97 (3–4) (1981) 639–666.
[54] S. Karaman, R. Sanfelice, E. Frazzoli, Optimal control of mixed logical dynamical systems with linear temporal logic specifications, in: Proceedings of the IEEE Conference on Decision and Control, IEEE Press, Piscataway, NJ, 2008, pp. 2117–2122.
[55] D. Kempe, F. McSherry, A decentralized algorithm for spectral analysis, J. Comput. Syst. Sci. 74 (1) (2008) 70–83.
[56] H.K. Khalil, Nonlinear Systems, 3rd ed., Prentice Hall, Upper Saddle River, NJ, 2002.
[57] V.L. Kharitonov, Asymptotic stability of an equilibrium position of a family of systems of differential equations, Differentsialnye Uravneniya 14 (1978) 2086–2088.
[58] J. Kim, D.G. Bates, I. Postlethwaite, L. Ma, P.A. Iglesias, Robustness analysis of biochemical network models, IEE Proc. Syst. Biol. 153 (3) (2006) 96–104.
[59] J.-S. Kim, N.V. Valeyev, I. Postlethwaite, P. Heslop-Harrison, K.-H. Cho, D.G. Bates, Analysis and extension of a biochemical network model using robust control theory, Int. J. Robust Nonlinear Control 20 (9) (2010) 1017–1026.
[60] K.-K.K. Kim, Model-based robust and stochastic control, and statistical inference for uncertain dynamical systems (Ph.D. thesis), University of Illinois at Urbana-Champaign, Illinois, 2013.
[61] K.-K.K. Kim, R.D. Braatz, Probabilistic analysis and control of uncertain dynamic systems: generalized polynomial chaos expansion approaches, in: Proceedings of the American Control Conference, IEEE Press, Piscataway, NJ, 2012, pp. 44–49.
[62] K.-K.K. Kim, R.D. Braatz, Generalised polynomial chaos expansion approaches to approximate stochastic model predictive control, Int. J. Control 86 (8) (2013) 1324–1337.
[63] K.-K.K. Kim, R.D. Braatz, Computational complexity and related topics of robustness margin calculation using μ theory: a review of theoretical developments, Comput. Chem. Eng. 70 (5) (2014) 122–132.
[64] K.-K.K. Kim, K.S. Cheong, K. Chen, R.D. Braatz, Parameter estimation, analysis, and design of synthetic gene switching models: system behavior-and performance-based approaches, in: Proceedings of the International Symposium on Advanced Control of Chemical Processes, Singapore, 2012, pp. 946–951.
[65] K.-K.K. Kim, K.S. Cheong, K. Chen, R.D. Braatz, Analysis of a synthetic gene switching motif: systems and control approaches, J. Process Control 24 (2) (2014) 341–347.
[66] K.-K.K. Kim, D.E. Shen, Z.K. Nagy, R.D. Braatz, Wiener's polynomial chaos for the analysis and control of nonlinear dynamical systems with probabilistic uncertainties, IEEE Control Syst. 33 (5) (2013) 58–67.
[67] M. Kishida, R.D. Braatz, Ellipsoid bounds on state trajectories for discrete-time systems with time-invariant and time-varying linear fractional uncertainties, in: Proceedings of the 50th IEEE Conference on Decision and Control and European Control Conference, IEEE Press, Piscataway, NJ, 2011, pp. 5671–5676.
[68] M. Kishida, R.D. Braatz, A model-based approach for the construction of design spaces in quality-by-design, in: Proceedings of the American Control Conference, IEEE Press, Piscataway, NJ, 2012, pp. 1513–1518.
[69] M. Kishida, R.D. Braatz, Non-existence conditions of local bifurcations for rational systems with structured uncertainties, in: Proceedings of the American Control Conference, IEEE Press, Piscataway, NJ, 2014, pp. 5085–5090.
[70] M. Kishida, R.D. Braatz, A skewed structured singular value based approach for the construction of design spaces: theory and applications, IET Control Theory Appl. 8 (14) (2014) 1321–1327.
[71] M. Kishida, R.D. Braatz, Volume maximization of consistent parameter sets for linear fractional models, in: Proceedings of the 50th IEEE Conference on Decision and Control, IEEE Press, Piscataway, NJ, 2014, pp. 1905–1910.
[72] M. Kishida, P. Rumschinski, R. Findeisen, R.D. Braatz, Efficient polynomial-time outer bounds on state trajectories for uncertain polynomial systems using skewed structured singular values, IEEE Trans. Autom. Control 59 (11) (2014) 3063–3068.
[73] H. Kitano, Computational systems biology, Nature 420 (6912) (2002) 206–210.
[74] H. Kitano, Biological robustness, Nat. Rev. Genet. 5 (11) (2004) 826–837.
[75] G. Kotsalis, A. Megretski, M.A. Dahleh, A model reduction algorithm for hidden Markov models, in: Proceedings of the IEEE Conference on Decision and Control, IEEE Press, Piscataway, NJ, 2006, pp. 3424–3429.
[76] A. Lasota, M.C. Mackey, Chaos, Fractals and Noise: Stochastic Aspects of Dynamics, Springer, New York, 1994.

[77] O. Le Maitre, P. Olivier, O.M. Knio, Spectral Methods for Uncertainty Quantification, Springer, Dordrecht, New York, 2010.
[78] J. Le Meitour, D. Lucor, J.-C. Chassaing, Prediction of stochastic limit cycle oscillations using an adaptive polynomial chaos method, ASD J. 2 (1) (2010) 3–22.
[79] Y. Lin, M. Stadtherr, Validated solutions of initial value problems for parametric ODEs, Appl. Numer. Math. 57 (10) (2007) 1145–1162.
[80] L. Ma, P.A. Iglesias, Quantifying robustness of biochemical network models, BMC Bioinf. 3 (2002), Art. No. 38.
[81] X. Ma, N. Zabaras, An efficient Bayesian inference approach to inverse problems based on an adaptive sparse grid collocation method, Inverse Prob. 25 (3) (2009), Art. No. 035013.
[82] J.-F. Magni, Presentation of the linear fractional representation toolbox (LFRT), in: Proceedings of the IEEE International Symposium on Computer Aided Control System Design, IEEE Press, Piscataway, NJ, 2002, pp. 248–253.
[83] A. Marcos, D. Bates, I. Postlethwaite, A symbolic matrix decomposition algorithm for reduced order linear fractional transformation modelling, Automatica 43 (7) (2007) 1211–1218.
[84] S. Marino, I.B. Hogue, C.J. Ray, D.E. Kirschner, A methodology for performing global uncertainty and sensitivity analysis in systems biology, J. Theor. Biol. 254 (1) (2008) 178–196.
[85] Y. Marzouk, D. Xiu, A stochastic collocation approach to Bayesian inference in inverse problems, Commun. Comput. Phys. 6 (4) (2009) 826–847.
[86] Y.M. Marzouk, H.N. Najm, L.A. Rahn, Stochastic spectral methods for efficient Bayesian solution of inverse problems, J. Comput. Phys. 224 (2) (2007) 560–586.
[87] M. Mönnigmann, W. Marquardt, Normal vectors on manifolds of critical points for parametric robustness of equilibrium solutions of ODE systems, J. Nonlinear Sci. 12 (2) (2002) 85–112.
[88] R.E. Moore, Interval Analysis, Prentice-Hall, Englewood Cliffs, NJ, 1966.
[89] R.E. Moore, R. Baker Kearfott, M.J. Cloud, Introduction to Interval Analysis, Society for Industrial and Applied Mathematics, New York, 2009.
[90] M. Morohashi, A.E. Winn, M.T. Borisuk, H. Bolouri, J. Doyle, H. Kitano, Robustness as a measure of plausibility in models of biochemical networks, J. Theor. Biol. 216 (1) (2002) 19–30.
[91] N. Motee, B. Bamieh, M. Khammash, Stability analysis of quasi-polynomial dynamical systems with applications to biological network models, Automatica 48 (11) (2012) 2945–2950.
[92] Z.K. Nagy, R.D. Braatz, Distributional uncertainty analysis using power series and polynomial chaos expansions, J. Process Control 17 (3) (2007) 229–240.
[93] Z.K. Nagy, R.D. Braatz, Distributional uncertainty analysis using polynomial chaos expansions, in: Proceedings of the IEEE International Symposium on Computer-Aided Control System Design, IEEE Press, Piscataway, NJ, 2010, pp. 1103–1108.
[94] N. Nedialkov, K. Jackson, G. Corliss, Validated solutions of initial value problems for ordinary differential equations, Appl. Math. Comput. 105 (1) (1999) 21–68.
[95] T. Ohtsuka, S. Streif, Commutativity of immersion and linearization, IEEE Trans. Autom. Control 54 (4) (2009) 826–829.
[96] S. Prajna, Barrier certificates for nonlinear model validation, Automatica 42 (1) (2006) 117–126.
[97] S. Prajna, A. Rantzer, Convex programs for temporal verification of nonlinear dynamical systems, SIAM J. Control Optim. 46 (3) (2007) 999–1021.
[98] A. Prékopa, Stochastic Programming, Kluwer Academic Publishers, Norwell, MA, 1994.
[99] P.E. Purnick, R. Weiss, The second wave of synthetic biology: from modules to systems, Nat. Rev. Mol. Cell Biol. 10 (6) (2009) 410–422.
[100] A. Rizk, G. Batt, F. Fages, S. Soliman, Continuous valuations of temporal logic specifications with applications to parameter optimization and robustness measures, Theor. Comput. Sci. 412 (26) (2011) 2827–2839.
[101] P. Rumschinski, S. Borchers, S. Bosio, R. Weismantel, R. Findeisen, Set-base dynamical parameter estimation and model invalidation for biochemical reaction networks, BMC Syst. Biol. 4 (2010), Art. No. 69.
[102] P. Rumschinski, D.S. Laila, S. Borchers, R. Findeisen, Influence of discretization errors on set-based parameter estimation, in: Proceedings of the IEEE Conference on Decision and Control, IEEE Press, Piscataway, NJ, 2010, pp. 296–301.
[103] P. Rumschinski, S. Streif, R. Findeisen, Combining qualitative information and semi-quantitative data for guaranteed invalidation of biochemical network models, Int. J. Robust Nonlinear Control 22 (10) (2012) 1157–1173.
[104] E.L. Russell, R.D. Braatz, Model reduction for the robustness margin computation of large scale uncertain systems, Comput. Chem. Eng. 22 (7) (1998) 913–926.
[105] E.L. Russell, C.P.H. Power, R.D. Braatz, Multidimensional realization of large scale uncertain systems for multivariable stability margin computation, Int. J. Robust Nonlinear Control 7 (1997) 113–125.
[106] A. Saltelli, K. Chan, E.M. Scott, Sensitivity Analysis, John Wiley & Sons, Chichester, United Kingdom, 2000.
[107] R. Samaga, S. Klamt, Modeling approaches for qualitative and semi-quantitative analysis of cellular signaling networks, Cell Commun. Signal. 11 (1) (2013), Art. No. 43.
[108] M. Schwarick, M. Heiner, C. Rohr, S. Streif, A comparative study of stochastic analysis techniques, in: Proceedings of the 8th International Conference on Computational Methods in Systems Biology, Trento, Italy, 2010, pp. 96–106.
[109] G. Shinar, M. Feinberg, Structural sources of robustness in biochemical reaction networks, Sci. Signal. 327 (5971) (2010) 1389–1391.
[110] J.E. Shoemaker, F.J. Doyle, Identifying fragilities in biochemical networks: robust performance analysis of FAS signaling-induced apoptosis, Biophys. J. 95 (6) (2008) 2610–2623.
[111] A.K. Singh, J. Hahn, On the use of empirical gramians for controllability and observability analysis, in: Proceedings of the American Control Conference, IEEE Press, Piscataway, NJ, 2005, pp. 140–146.
[112] E.D. Sontag, Adaptation and regulation with signal detection implies internal model, Syst. Control Lett. 50 (2) (2003) 119–126.
[113] E.D. Sontag, Molecular systems biology and control, Eur. J. Control 11 (4) (2005) 1–40.
[114] J. Stelling, S. Klamt, K. Bettenbrock, S. Schuster, E.D. Gilles, Metabolic network structure determines key aspects of functionality and regulation, Nature 420 (6912) (2002) 190–193.
[115] J. Stelling, U. Sauer, Z. Szallasi, F.J. Doyle III, J. Doyle, Robustness of cellular functions, Cell 118 (6) (2004) 675–685.
[116] R. Steuer, S. Waldherr, V. Sourjik, M. Kollmann, Robust signal processing in living cells, PLoS Comput. Biol. 7 (11) (2011), Art. No. e1002218.
[117] S. Streif, R. Findeisen, E. Bullinger, Relating cross gramians and sensitivity analysis in systems biology, in: Proceedings of the Mathematical Theory of Networks and Systems, Kyoto, Japan, 2006, pp. 437–442.
[118] S. Streif, R. Findeisen, E. Bullinger, Sensitivity analysis of biochemical reaction networks by bilinear approximation., in: Proceedings of the Foundations of Systems Biology in Engineering Conference, Stuttgart, Germany, 2007, pp. 521–526.
[119] S. Streif, M. Karl, R. Findeisen, Outlier analysis in set-based estimation for nonlinear systems using convex relaxations, in: Proceedings of the European Control Conference, Zurich, Switzerland, 2013, pp. 2921–2926.
[120] S. Streif, K.-K.K. Kim, P. Rumschinski, M. Kishida, D.E. Shen, R. Findeisen, R.D. Braatz, Robustness analysis, prediction and estimation for uncertain biochemical networks., in: Proceedings of 10th IFAC International Symposium on Dynamics and Control of Process Systems (DyCoPS), Mumbai, India, 2013, pp. 1–20.
[121] S. Streif, D. Oesterhelt, W. Marwan, A predictive computational model of the kinetic mechanism of stimulus-induced transducer methylation and feedback regulation through CheY in archaeal phototaxis and chemotaxis, BMC Syst. Biol. 4 (2010), Art. No. 27.
[122] S. Streif, P. Rumschinski, D. Henrion, R. Findeisen, Estimation of consistent parameter sets of nonlinear continuous-time systems using occupation measures and LMI relaxations, in: Proceedings of the IEEE Conference on Decision and Control, IEEE Press, Piscataway, NJ, 2013, pp. 6379–6384.
[123] S. Streif, A. Savchenko, P. Rumschinski, S. Borchers, R. Findeisen, ADMIT: a toolbox for guaranteed model invalidation, estimation and qualitative-quantitative modeling, Bioinformatics 28 (9) (2012) 1290–1291 http://ifatwww.et.uni-magdeburg.de/syst/ADMIT/.
[124] S. Streif, N. Strobel, R. Findeisen, Inner approximations of consistent parameter sets via constraint inversion and mixed-integer linear programming, in: Proceedings of the 12th IFAC Symposium on Computer Applications in Biotechnology, Mumbai, India, 2013, pp. 326–331.
[125] S. Streif, S. Waldherr, F. Allgöwer, R. Findeisen, Steady state sensitivity analysis of biochemical reaction networks: a brief review and new methods, in: A. Jayaraman, J. Hahn (Eds.), Systems Analysis of Biological Networks, Artech House Press, Boston, MA/London, United Kingdom, 2009, pp. 129–148.
[126] B. Sudret, Global sensitivity analysis using polynomial chaos expansions, Reliab. Eng. Syst. Saf. 93 (7) (2008) 964–979.
[127] A. Surana, A. Banaszuk, Scalable uncertainty quantification in complex dynamic networks, in: Proceedings of the IEEE Conference on Decision and Control, IEEE Press, Piscataway, NJ, 2010, pp. 7278–7285.
[128] A. Surana, T. Sahai, A. Banaszuk, Iterative methods for scalable uncertainty quantification in complex networks, Int. J. Uncertain. Quantif. 2 (4) (2012) 413–439.
[129] S.R. Taylor, R. Gunawan, L.R. Petzold, F.J. Doyle, Sensitivity measures for oscillating systems: application to mammalian circadian gene network, IEEE Trans. Autom. Control 53 (2008) 177–188.
[130] C. Trané, E.W. Jacobsen, Network structure and robustness of intracellular oscillators, in: Proceedings of the IFAC World Congress, Seoul, Korea, 2008, pp. 10989–10994.
[131] N.A.W. van Riel, E.D. Sontag, Parameter estimation in models combining signal transduction and metabolic pathways: the dependent input approach, IEE Proc. Syst. Biol. 153 (4) (2006) 263–274.
[132] K. Venkatesh, S. Bhartiya, A. Ruhela, Multiple feedback loops are key to a robust dynamic performance of tryptophan regulation in *Escherichia coli*, FEBS Lett. 563 (1) (2004) 234–240.
[133] M. Vidyasagar, The complete realization problem for hidden Markov models: a survey and some new results, Math. Control Signals Syst. 23 (1) (2011) 1–65.
[134] U. von Luxburg, A tutorial on spectral clustering, Stat. Comput. 17 (4) (2007) 395–416.
[135] S. Waldherr, Uncertainty and Robustness Analysis of Biochemical Reaction Networks via Convex Optimisation and Robust Control Theory (Ph.D. thesis), Universität Stuttgart, Germany, 2009.
[136] S. Waldherr, F. Allgöwer, Robust stability and instability of biochemical networks with parametric uncertainty, Automatica 47 (6) (2011) 1139–1146.
[137] S. Waldherr, F. Allgöwer, E.W. Jacobsen, Kinetic perturbations as robustness analysis tool for biochemical reaction networks, in: Proceedings of the IEEE

Conference on Decision and Control, IEEE Press, Piscataway, NJ, 2009, pp. 4572–4577.
[138] S. Waldherr, S. Streif, F. Allgöwer, Design of biomolecular network modifications for adaptation, IET Syst. Biol. 6 (6) (2012) 223–231.
[139] N. Wiener, The homogeneous chaos, Am. J. Math. 60 (1938) 897–936.
[140] T. Wilhelm, J. Behre, S. Schuster, Analysis of structural robustness of metabolic networks, IET Syst. Biol. 1 (1) (2004) 114–120.
[141] A.K. Wilkins, B. Tidor, J. White, P.I. Barton, Sensitivity analysis for oscillating dynamical systems, SIAM J. Sci. Comput. 31 (4) (2009) 2709–2732.
[142] D. Xiu, Numerical Methods for Stochastic Computations: A Spectral Method Approach, Princeton University Press, Princeton, NJ, 2010.
[143] D. Xiu, G.E. Karniadakis, The Wiener-Askey polynomial chaos for stochastic differential equations, SIAM J. Sci. Comput. 24 (2) (2002) 619–644.
[144] T.-M. Yi, Y. Huang, M.I. Simon, J. Doyle, Robust perfect adaptation in bacterial chemotaxis through integral feedback control, Proc. Natl. Acad. Sci. U. S. A. 97 (9) (2000) 4649–4653.
[145] B. Yordanov, C. Belta, A formal verification approach to the design of synthetic gene networks, in: Proceedings of the IEEE Conference on Decision and Control and European Control Conference, IEEE Press, Piscataway, NJ, 2011, pp. 4873–4878.
[146] M. Zettler, J. Garloff, Robustness analysis of polynomials with polynomial parameter dependency using Bernstein expansion, IEEE Trans. Autom. Control 43 (3) (1998) 425–431.
[147] K. Zhou, J.C. Doyle, K. Glover, Robust and Optimal Control, Prentice Hall, Upper Saddle River, NJ, 1995.
[148] Z. Zi, Sensitivity analysis approaches applied to systems biology models, IET Syst. Biol. 5 (6) (2011) 336–346.