Nonlinear lumping of cascade activation reactions

Dávid Csercsik & Katalin M. Hangos

Research Report SCL-003/2007

Nonlinear lumping of cascade activation reactions

Dávid Csercsik, Katalin M. Hangos

November 5, 2007

[1]Process Control Research Group, Systems and Control Laboratory
Computer and Automation Research Institute,
Hungarian Academy of Sciences
H-1518, P.O. Box 63, Budapest, Hungary
Tel: +36 1 279 6000
Fax: +36 1 466 7503
e-mail: {csercsik, hangos}@scl.sztaki.hu

Abstract

This paper describes the extension of the variable lumping transformation, being a well-known engineering-type model simplification method, to a special class of nonlinear reaction kinetic systems, to the cascade activation reactions. Two methods of cascade lumping is investigated: (i) variable lumping with changing the stoichiometric coefficients, and (ii) variable lumping with output variable transformation.

It is shown that the simplified model remains in the class of reaction kinetic systems in both cases. The methods are illustrated on one and two step cascade activation reactions. A more complex case study of a feedback regulated cascade structure is also presented, which is related to the MAPK intracellular signaling cascade.

Keywords:

Reaction kinetic networks, model simplification, variable lumping transformation, systems biology, minimal model, MAPK cascade

Contents

1 Introduction

In biology there are many phenomena that cannot be explained with relevant properties of it's physical elements (e.g. molecules taking part in the phenomena), but with the dynamics of the interactions between its elements. Systems biology is the science that aims at studying these interaction networks in biological systems. The spectrum of systems biology is very wide, ranging from population dynamics to biochemical networks.

For the appropriate description of complex dynamical phenomena in bio-chemical networks in particular, one needs solid system theoretical methods, which are yet practically feasible. In the general case, the theory of reaction kinetic systems [6, 7, 10] provides tools for model synthesis, analysis and model simplification of biochemical reaction systems (e.g. signaling pathways). The resulting mathematical models, however, are usually very complex, therefore they are usually not suitable for model verification, validation and model parameter estimation using experimental data. Thus the need to simplify these models naturally arises.

In the case of reaction kinetic systems the most common ways of model simplification are the utilization of conservation equations and quasi steady state approximations [21, 26]. In the case of this system class, there are further promising model reduction methods based on invariant manifolds and optimization [12, 22].

Some other results of nonlinear model reduction, based on singular perturbation or other types of rigorous assumptions can also be found in the literature [11, 14]. These results are applied dominantly in the case of metabolic reaction networks. In fact the application of such techniques for enzyme-type cascade activation reactions can be problematic, because of the strict requirements corresponding to the reaction kinetic system's structure and equations, for which these methods can be applied.

At the same time, there are general and reaction kinetic systems-related model simplification and reduction methods in the engineering literature [19, 20, 2] that can possibly be applied to biochemical kinetic models, as well. In the case of engineering type model reduction techniques, like quasi steady-state assumptions or variable lumping, the main aim is the reduction of the model in order to obtain a system, which can describe the input-output relationships of the original system with a minimal number of state variables and parameters.

The aim of this paper is to provide new model simplification methods for a special class of reaction kinetic systems-based nonlinear ODE representation of biochemical systems, which utilizes the special structural properties of the system class. The method is a nonlinear extension of the engineering type model reduction method called variable lumping applied to cascade activation structures. This nonlinear cascade lumping is designed in a such a way that the resulting reduced system remains in the class of reaction kinetic networks.

2 Basic notions: reaction kinetic systems and their model simplification approaches

The mathematical description of reaction kinetic systems in the form of ordinary differential equations (ODEs) is described here together with the most widely used model simplification approaches: the method of conservation equations, steady-state assumptions and variable lumping.

2.1 Reaction kinetic systems with mass action law kinetics

We consider a system of n chemical species participating in an r reversible steps reaction network **in a closed system under isothermal and isobaric conditions**:

$$\sum_{i=1}^{n} \alpha_{ij}\chi_i \leftrightarrows \sum_{i=1}^{n} \beta_{ij}\chi_i \quad for\ j = 1, ...r \tag{1}$$

where the integers $\alpha_{ij}, \beta_{ij} \in \mathbb{N}$ are the stoichiometric coefficients for specie χ_i in the reaction step j. The r stoichiometric vectors are defined as $\nu_{ij} = \beta_{ij} - \alpha_{ij}$. The reaction rate in each reversible step is assumed to obey the Mass Action Law [10]:

$$W_j = k_j^+ \prod_{i=1}^{n} x_i^{\alpha_{ij}} - k_j^- \prod_{i=1}^{n} x_i^{\beta_{ij}} \tag{2}$$

where k_j^+ and k_j^- are the constants of the direct and of the inverse reaction rates of the j-th reaction step. The concentration vector is represented by x where the component $x_i \geq 0$ is the concentration of the specie χ_i.

Reaction kinetics equations that describe the evolution of the states in time can be expressed in matrix notation ([13]):

$$\dot{x} = \mathcal{N} \cdot W \tag{3}$$

Where $\mathcal{N} \in \mathbb{R}^{n \times r}$ and $W \in \mathbb{R}^{r \times 1}$ are the full rank matrix of stoichiometric vectors and the vector of reaction rates, respectively.

The linear combination of species defined by the stoichiometric vectors are called complexes.

2.2 Graphical description of reaction kinetic systems

Reaction schemes Reaction schemes [17, 5, 25, 23] can be depicted using **hypergraphs** in mathematical terms, where the edges may be adjacent to more than two vertices. The vertices of a reaction scheme correspond to the non enzymatic complex type components, while the hyper-edges describe chemical reactions (not necessarily reaction steps!). An enzyme-catalytic reaction corresponds to a pair of hyper-edges with different directions both adjacent to three

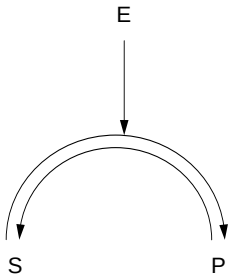

Figure 1: A simple reaction scheme

components S, P and E (being the substrate, product and enzyme, respectively) as shown in Fig. 1. The scheme in Fig. 1 corresponds to the reaction steps

$$E + S \underset{k_1^-}{\overset{k_1^+}{\rightleftharpoons}} ES \underset{k_2^-}{\overset{k_2^+}{\rightleftharpoons}} E + P \tag{4}$$

with ES being an enzymatic complex type component.

Reaction graph The reaction graph of a system is defined by the vertices, which are related to complexes (linear combination of species), and edges, related to reactions between complexes. Edges are weighted with reaction rates. Linkage classes are defined as connected components of the reaction graph. The reaction graph of the enzymatic reaction scheme depicted in Fig. 1 can be seen in Fig. 2. This reaction graph contains only one linkage class.

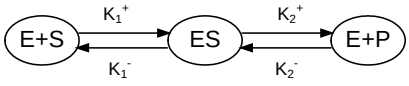

Figure 2: A simple reaction graph

Reaction simplex In a system with m species, we call the vector space $V = \mathbb{R}^m$ the species space, with each coordinate representing a different specie. A complex can be represented as a vector in species space as the sum of its constituent species multiplied by the appropriate coefficients, defined by the stoichiometric vectors. Each reaction can be represented as the vector associated with the reactant complex subtracted from the vector associated with the product complex. This is called the **reaction vector** for the corresponding reaction. The matrix $\mathcal{N}$ is composed of columns of the reaction vectors.

The subspace of reaction space defined by the linear span of the reaction vectors is called S, the **stoichiometric subspace**.

For a vector $v \in V$, the coset $v + S$ intersected with the positive subset of V (V^+) is called a reaction simplex.

Structure graphs A structure graph [4] is a signed directed graph that depicts the direct influences between the state, input and output variables of a system. For closed reaction kinetic systems the state variables are the concentrations of the components that correspond to the vertices of the structure graph. There is a directed edge from vertex v_i to v_j, if the time variation of v_j depends on v_i, i.e. they are both present in the same reaction step r_ℓ.

The structure graph of the scheme shown in Fig. 1 is depicted in Fig. 3

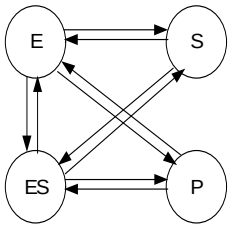

Figure 3: A simple structure graph

2.3 Steady states

Formally speaking, one has to solve the static (algebraic) counterpart of the reaction kinetic equations (3)

$$0 = N \cdot W(x)$$

with the reaction rates $W(x)$ given in Eq. (2) for x. Because the reaction rates depend on the concentrations x in a non-linear way, this set of equations may not have solution, or may have multiple solutions depending on the reaction kinetic mechanism and on its parameters.

2.4 Conservation equations

The reaction simplex is in every case a positively invariant set, i.e. the trajectories remain in an s dimensional translated linear subspace, where s is the dimension of S.

Therefore, $n - s$ linearly independent linear conservation equations appear (where n is the number of species) in the form:

$$b_i^T x = b_i^T x(0), \qquad b_i \in \mathbb{R}^n,\ i = 1, \ldots, n.$$

Moreover, $b_i \in \ker(\mathcal{N}^T)$.

For example, in the case of reactions described in Eq. (4) we obtain

$$\mathcal{N} = \begin{pmatrix} -1 & 1 \\ -1 & 0 \\ 1 & -1 \\ 0 & 1 \end{pmatrix}$$

The first row of the above matrix corresponds to E, the second to S, the third to ES and the fourth to P.

$$\ker(\mathcal{N}^T) = \begin{pmatrix} 1 & -1 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}$$

If we take the first column, we get the conservation equation for the enzyme: $E+ES = E^{tot}$. The sum of the first and the second column gives the conservation for the substrate and the product: $S + ES + P = S^{tot}$

The utilization of conservation equations, i.e. their substitution into the differential equations decreases the number of state-space variables in the ODE (ordinary differential equation) description of a reaction kinetic system providing model, which exhibits the same dynamics as the original. The model obtained with this method is minimal in the sense, that the degree of freedom of the system is equal to the number of ODE's describing the model - the original differential-algebraic system is reduced to a differential system. In fact the model will not be necessary minimal from the point of view of input-output behavior.

The limitation of this method is, that the conservation assumptions are not always valid (e.g. in the case of open systems which interact with their environment).

2.5 Quasi-steady state assumption

The quasi-steady state (QSS) approximation can be applied to decrease the number of ODE's and the number of differential (state) variables at the same time, if certain conditions are fulfilled. The most widespread of those conditions in reaction kinetics when some reactions are considered to be much faster or much slower than the others, and are assumed to be in quasi-equilibria.

Formally, this assumption imposed on the variable x_i implies

$$\frac{dx_i}{dt} = 0 \quad (x_i = const) \quad \text{and} \quad F_i(x) = 0$$

where F_i is the right-hand side expression of the ODE with the left-hand side $\frac{dx_i}{dt}$.

2.6 Variable lumping

The main aim of variable lumping is to approximate the original model's input-output (or output) behavior with less state variables. It is achieved by finding two (or more) state variables with similar dynamics (say x_i and x_j) and describe them with a single pseudo-variable $\hat{x}_{i+j}$. More formally, the two variables x_i and x_j are replaced with a single state variable $\hat{x}_{i+j}$, that "inherits" all the effects acting on both x_i and x_j.

Note, that the elementary step of variable lumping is to lump two variables together, and this will be the case that is considered for reaction kinetic systems.

In the case of linear systems, the variable lumping transformation can be applied to state variables, which are strongly correlated. This technique can be extended to less correlated state variables in the case of cascade systems [20].

The method of variable lumping may be a step towards the synthesis of a minimal model, which produces the input-output dynamics of the original system with the minimal number of state variables and parameters.

3 Nonlinear lumping of cascade activation reactions

In this section the class of cascade activation reactions is defined, and the basic properties of the variable lumping transformation are detailed.

3.1 Cascade activation reactions

In the following a set of enzymatic reaction levels, where the product of a higher level acts as an enzyme in the subsequent reaction level will be called **cascade activation reaction**. The most simple reaction scheme of cascade activation can be seen in figure 4.

The original system is described by the following reactions:

$$\begin{aligned} E + S_1 &\underset{k_1^-}{\overset{k_1^+}{\rightleftharpoons}} ES_1 \underset{k_2^-}{\overset{k_2^+}{\rightleftharpoons}} E + P_1 \\ P_1 + S_2 &\underset{k_3^-}{\overset{k_3^+}{\rightleftharpoons}} P_1S_2 \underset{k_4^-}{\overset{k_4^+}{\rightleftharpoons}} P_1 + P_2 \end{aligned} \tag{5}$$

The first linkage class refers to the first level of the activation cascade, and the second one refers to second level.

3.2 The basis of cascade reaction lumping

If we analyze this structure from the point of view of information flow taking place in cell-signaling with zero initial conditions of P_1 and P_2, we can observe that first the concentration

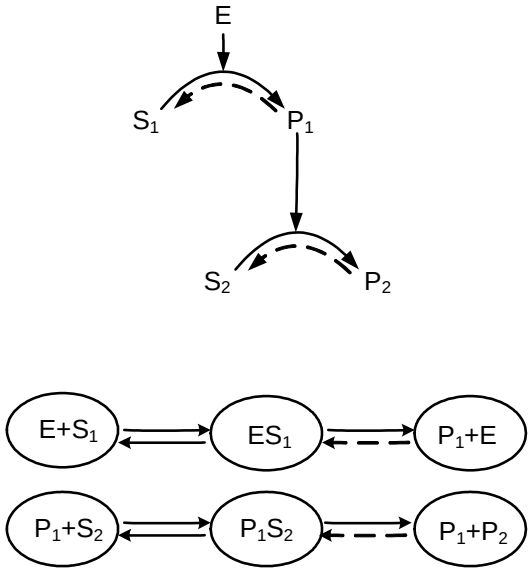

Figure 4: The activation scheme and reaction graph of the basic cascade activation reaction. The dashed lines denote the reversible case

of P_1 increases caused by the effect of E, and then due to the effect of P_1 the concentration of P_2 rises, too.

The basic idea is to shorten the way between the beginning and the end of the signaling path ($E \rightarrow P_1 \rightarrow P_2$), and eliminate the intermediate component in a way, that approximately preserves the dynamic properties of the system at least from the input-output point of view. In this case the intermediate product P_1, which acts as an enzyme in the second reaction is lumped together with the final product P_2 of the cascade into a new variable $\hat{P}$, as seen in the right sub-figure of figure 5.

The new variable $\hat{P}$ is a result of the lumping of a product, which acts as an enzyme (P_1) and product (P_2), therefore the enzymatic property of the component is preserved in the second linkage class of reactions.

This new reaction scheme implies the following reactions:

$$\begin{aligned} E + S_1 &\underset{k_1^-}{\overset{k_1^+}{\rightleftharpoons}} ES_1 \underset{k_2^-}{\overset{k_2^+}{\rightleftharpoons}} E + \hat{P} \\ \hat{P} + S_2 &\underset{k_3^-}{\overset{k_3^+}{\rightleftharpoons}} \hat{P}S_2 \underset{k_4^-}{\overset{k_4^+}{\rightleftharpoons}} \hat{P} \end{aligned} \tag{6}$$

As mentioned before in sub-section 2.6, the variable $\hat{P}$ inherits all of the connections of the lumped variables, as it can be seen in the activation scheme (Fig. 5), and also in the structure

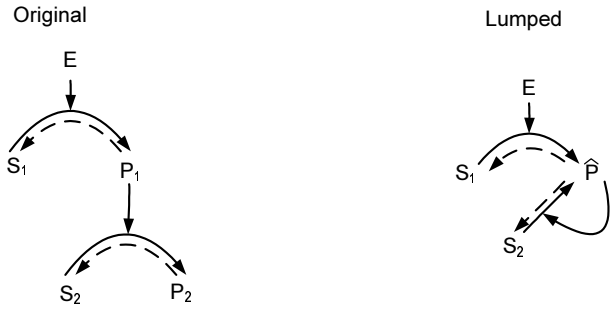

Figure 5: The reaction scheme of the original and the lumped reaction. The dashed lines denote the reversible case

graph (Fig. 6 where the loop edges are not included, because their existence is trivial in the case of reaction kinetic systems).

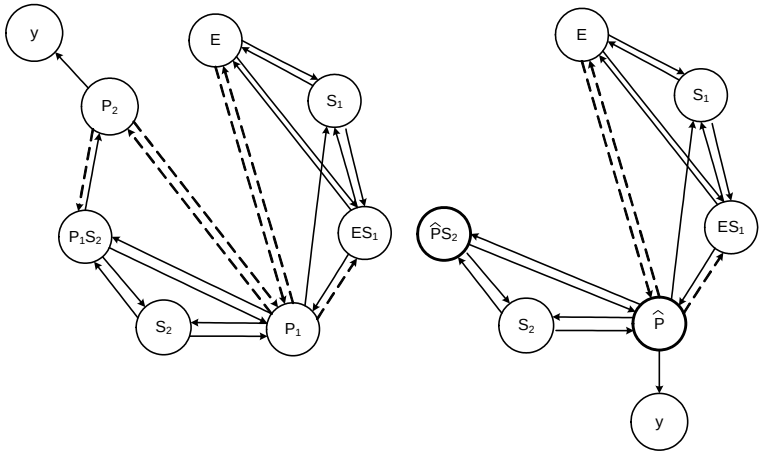

Figure 6: The structure graph of the original and the lumped reaction. The dashed lines denote the reversible case

3.3 The variable lumping as graph transformation

The lumping transformation can be considered as a transformation of the corresponding graphs, the reaction graph and the structure graph.

The transformation of the reaction graph occurs only on the level of complexes, the structure of the graph remains unchanged. To see this, we recall, that the reaction graph of the original activation scheme has essentially the same structure as Eq. (5) while the new reaction graph is described in Eq. (6). This means that the matrix representation of the graph is the same, only the rows and columns correspond to other complexes. The number of components is decreased by 1. This means that, of course, the ODE-s implied by the reactions also change, as described

later.

The structure graph of the lumped system can be obtained by dragging together the corresponding vertices P_1 and P_2 into $\hat{P}$.

3.4 The transformation of the model equations

As described in [13] the differential equations of the model can be written in the form of

$$\frac{dx}{dt} = \mathcal{N}W$$

where in the original case

$$\mathcal{N} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

with the rows corresponding to S_1, E, ES_1, P_1, S_2, P_1S_2, P_2.

In the transformed case the matrix takes the following form:

$$\mathcal{N}_L = \begin{pmatrix} -1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 1 & -1 \end{pmatrix}$$

with the rows corresponding to S_1, E, ES_1, $\hat{P}$, S_2, $\hat{P}S_2$.

The transformation of the matrix $\mathcal{N}$ into $\mathcal{N}_L$ can be described by the **non-invertible** linear transformation $\mathcal{N}_L = T\,\mathcal{N}$ where

$$T = \begin{pmatrix} 0 & 1 & 0 & -1 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

3.5 The constraints related to the conservation laws

It is a primary requirement that any simplified model should obey the conservation laws underlying the original mechanism. This poses additional requirements or constraints on the structure and/or parameters of the simplified model.

For our cascade activation reaction scheme one can observe that in both reaction steps the total normalized concentration of the substrate and product (the concentration of S_1 and P_1, or S_2, P_2, respectively together with the intermediate complex) is constant, therefore these total normalized concentrations can be regarded as 1. Furthermore, we can observe that S_1 and S_2 is transformed into P_1 and P_2, therefore the sum of their concentrations will reach the concentration value of 1 in the irreversible case.

In the original reaction scheme P_2 was considered as the output of the system. If we, however, consider the lumped reactions (6), S_1 is transformed into $\hat{P}$, and S_2 is eliminated via $\hat{P}$. $\hat{P}$ acts as an "eliminating enzyme" of S_2.

In order to preserve the conservation laws, we can, for example, modify the second reaction as follows:

$$\hat{P} + S_2 \underset{k_3^-}{\overset{k_3^+}{\rightleftharpoons}} \hat{P}S_2 \underset{k_4^-}{\overset{k_4^+}{\rightleftharpoons}} 2\hat{P}$$

This way we have modified the stoichiometric coefficient of $\hat{P}$. In fact, both S_1 and S_2 are transformed into $\hat{P}$ in this case, that will reach the concentration value of 2 at the end of the reaction.

Motivated by the above reasoning, two possible approaches will be investigated in the following to modify the result of the raw variable lumping transformation in order to satisfy the conservation laws.

- the suitable **adjustment of stoichiometric coefficients** in order to achieve a good approximation of the original output behavior,
- the **output transformation method**, which will be discussed later in section 5.

4 Cascade lumping with the adjustment of stoichiometric coefficients

In this case the stoichiometric coefficients are adjusted in such a way, that the final concentration of the final product of the original reaction should be the same as in the original case. We consider the case when one unit of both substrates are transformed into half unit of the product.

$$\begin{aligned} E + S_1 &\underset{k_1^-}{\overset{k_1^+}{\rightleftharpoons}} ES_1 \underset{k_2^-}{\overset{k_2^+}{\rightleftharpoons}} E + 0.5\hat{P} \\ \hat{P} + S_2 &\underset{k_3^-}{\overset{k_3^+}{\rightleftharpoons}} \hat{P}S_2 \underset{k_4^-}{\overset{k_4^+}{\rightleftharpoons}} 1.5\hat{P} \end{aligned} \tag{7}$$

This reactions will imply the final concentration of 1 for $\hat{P}$ in the irreversible case, as later confirmed by the simulations.

4.1 State-space model

The original reactions (5) imply the following differential equations if they obey the mass-action law:

$$\begin{aligned}
\frac{d[S_1]}{dt} &= -k_1^+[E][S_1] + k_1^-[ES_1] \\
\frac{d[E]}{dt} &= -k_1^+[E][S_1] + k_1^-[ES_1] + k_2^+[ES_1] - k_2^-[E][P_1] \\
\frac{d[ES_1]}{dt} &= k_1^+[E][S_1] - k_1^-[ES_1] - k_2^+[ES_1] + k_2^-[E][P_1] \\
\frac{d[P_1]}{dt} &= k_2^+[ES_1] - k_2^-[E][P_1] - k_3^+[P_1][S_2] + k_3^-[P_1S_2] + k_4^+[P_1S_2] - k_4^-[P_1][P_2] \\
\frac{d[S_2]}{dt} &= -k_3^+[S_2][P_1] + k_3^-[P_1S_2] \\
\frac{d[P_1S_2]}{dt} &= k_3^+[S_2][P_1] - k_3^-[P_1S_2] - k_4^+[P_1S_2] + k_4^-[P_1][P_2] \\
\frac{d[P_2]}{dt} &= k_4^+[P_1S_2] - k_4^-[P_1][P_2]
\end{aligned} \tag{8}$$

The **output of the model** is P_2.

As detailed before, the model can be written in the form of

$$\frac{dx}{dt} = \mathcal{N}W$$

where

$$\mathcal{N} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \qquad W = \begin{pmatrix} k_1^+[E][S_1] - k_1^-[ES_1] \\ k_2^+[ES_1] - k_2^-[E][P_1] \\ k_3^+[S_2][P_1] - k_3^-[P_1S_2] \\ k_4^+[P_1S_2] - k_4^-[P_1][P_2] \end{pmatrix}$$

The lumped reactions 5 imply the following differential equations:

$$\begin{aligned}
\frac{d[S_1]}{dt} &= -k_1^+[E][S_1] + k_1^-[ES_1] \\
\frac{d[E]}{dt} &= -k_1^+[E][S_1] + k_1^-[ES_1] + k_2^+[ES_1] - k_2^-[E][\hat{P}]^{0.5} \\
\frac{d[ES_1]}{dt} &= k_1^+[E][S_1] - k_1^-[ES_1] - k_2^+[ES_1] + k_2^-[E][\hat{P}]^{0.5} \\
\frac{d[\hat{P}]}{dt} &= 0.5k_2^+[ES_1] - 0.5k_2^-[E][\hat{P}]^{0.5} - k_3^+[\hat{P}][S_2] + k_3^-[\hat{P}S_2] + 1.5k_4^+[\hat{P}S_2] - 1.5k_4^-[\hat{P}]^{1.5} \\
\frac{d[S_2]}{dt} &= -k_3^+[\hat{P}][S_2] + k_3^-[\hat{P}S_2] \\
\frac{d[\hat{P}S_2]}{dt} &= k_3^+[\hat{P}][S_2] - k_3^-[\hat{P}S_2] - k_4^+[\hat{P}S_2] + k_4^-[\hat{P}]^{1.5}
\end{aligned} \tag{9}$$

The **output of the simplified model** is $\hat{P}$.

In this case the algebraic description will be the following:

$$\mathcal{N} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 0.5 & -1 & 1.5 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 1 & -1 \end{pmatrix} \qquad W = \begin{pmatrix} k_1^+[E][S_1] - k_1^-[ES_1] \\ k_2^+[ES_1] - k_2^-[E][\hat{P}]^{0.5} \\ k_3^+[\hat{P}][S_2] - k_3^-[\hat{P}S_2] \\ k_4^+[\hat{P}S_2] - k_4^-[\hat{P}]^{1.5} \end{pmatrix}$$

Effect on the deficiency With the variable lumping transformation, we reduced the number of species, but the number of complexes and linkage classes remain unchanged. Furthermore the rank of the matrix $\mathcal{N}$ is also conserved, so the **deficiency is invariant under the variable lumping transformation**.

4.2 Steady-states of the original and the simplified model

In this part, we examine the equations required for steady-state of the original and the lumped model.

4.2.1 Original model

In the case of the original model we can derive the steady-state equations by the following way:

$$\begin{aligned} \frac{dS_1}{dt} &= 0 \quad \Rightarrow \quad k_1^+[E][S_1] = k_1^-[ES_1] \\ \frac{dE}{dt} &= 0 \quad \Rightarrow \quad k_2^+[ES_1] = k_2^-[E][P_1] \\ \frac{dS_2}{dt} &= 0 \quad \Rightarrow \quad k_3^+[S_2][P_1] = k_3^-[P_1S_2] \\ \frac{dP_1S_2}{dt} &= 0 \quad \Rightarrow \quad k_4^+[P_1S_2] = k_4^-[P_1][P_2] \end{aligned} \tag{10}$$

These equations imply the steady-state of the remaining two equations in (8) for $[P_2]$ and $[P_1]$.

In the original model we have 3 conservation laws (for $[S_1]$,$[E]$ and $[S_2]$) applied for 7 equations. This means that the reduced minimal model is 4 dimensional, thus we need 4 equations describe the steady state.

Note that there is a **unique steady-state** if the deficiency of the model is zero.

4.2.2 Lumped model

In the case of the lumped-variable model, it follows from the steady state of (9) that:

$$\begin{aligned}
\frac{dS_1}{dt} &= 0 \quad \Rightarrow \quad k_1^+[E][S_1] = k_1^-[ES_1] \\
\frac{dE}{dt} &= 0 \quad \Rightarrow \quad k_2^+[ES_1] = k_2^-[E][\hat{P}]^{0.5} \\
\frac{dS_2}{dt} &= 0 \quad \Rightarrow \quad k_3^+[S_2][\hat{P}] = k_3^-[\hat{P}S_2] \\
\frac{d\hat{P}}{dt} &= 0 \quad \Rightarrow \quad 1.5k_4^+[\hat{P}S_2] = 1.5k_4^-[\hat{P}]^{1.5}
\end{aligned} \tag{11}$$

These equations imply the steady state of the remaining two equations in (9) for $[ES_1]$ and $[\hat{P}S_2]$. This means that the reduced minimal model has to be 4 dimensional, so we need to have 2 conservation equations. The first one is

$$[E] + [ES] = [E]^{tot}$$

and the second one is a pseudo-conservation law which reads:

$$[S_1] + [ES_1] + [S_2] + 2[\hat{P}] + 3[\hat{P}S_2] = [S]^{tot} \tag{12}$$

This pseudo-conservation can be derived from the sum of the conservations

$$[S_1] + [ES_1] + 0.5[\hat{P}] = [S_1^{tot}]$$

$$[S_2] + 3[\hat{P}S_2] + 1.5[\hat{P}] = [S_2^{tot}]$$

where the multiplier 3 in the second equations is derived from the intuition that $\hat{P}$ stands double-weighted and S_2 stands single weighted.

It is important to note that - here again - there is a **unique steady-state** if the deficiency of the model is zero. As we have seen before, the deficiency does not change if one applies variable lumping, therefore the uniqueness of the steady-state of the original model implies that of the lumped one.

In the irreversible case, this steady state means, that all of the substrates are transformed into products (into the lumped variable, describing the products in the lumped case), so the invariant coordinates of the steady-state (eg. substrate concentrations) are equal in the case of the original and the lumped system.

As we will see later, in the case of simulations, in the irreversible case the steady-state is not invariant in any sense.

4.3 Controllability

A called minimal model of a system should be controllable. To show that this is the case, we examine the controllability of the lumped model.

We consider an **additive term in the differential equation of the enzyme as input**. This choice is related to the general structure of biochemical signaling pathways, where the receptor related G-protein, and also the remaining free receptor, which can be phosphorylated, has enzymatic activity [3].

$$\begin{aligned}
\frac{d[S_1]}{dt} &= -k_1^+[E][S_1] + k_1^-[ES_1] \\
\frac{d[E]}{dt} &= -k_1^+[E][S_1] + k_1^-[ES_1] + k_2^+[ES_1] - k_2^-[E][\hat{P}]^{0.5} + u \\
\frac{d[ES_1]}{dt} &= k_1^+[E][S_1] - k_1^-[ES_1] - k_2^+[ES_1] + k_2^-[E][\hat{P}]^{0.5} \\
\frac{d[\hat{P}]}{dt} &= 0.5k_2^+[ES_1] - 0.5k_2^-[E][\hat{P}]^{0.5} - k_3^+[\hat{P}][S_2] + k_3^-[\hat{P}S_2] + 1.5k_4^+[\hat{P}S_2] - 1.5k_4^-[\hat{P}]^{1.5} \\
\frac{d[S_2]}{dt} &= -k_3^+[\hat{P}][S_2] + k_3^-[\hat{P}S_2] \\
\frac{d[\hat{P}S_2]}{dt} &= k_3^+[\hat{P}][S_2] - k_3^-[\hat{P}S_2] - k_4^+[\hat{P}S_2] + k_4^-[\hat{P}]^2
\end{aligned} \tag{13}$$

Further we consider the irreversible case ($k_2^- = k_4^- = 0$) when the equations read

$$\begin{aligned}
\frac{d[S_1]}{dt} &= -k_1^+[E][S_1] + k_1^-[ES_1] \\
\frac{d[E]}{dt} &= -k_1^+[E][S_1] + k_1^-[ES_1] + k_2^+[ES_1] + u \\
\frac{d[ES_1]}{dt} &= k_1^+[E][S_1] - k_1^-[ES_1] - k_2^+[ES_1] \\
\frac{d[\hat{P}]}{dt} &= 0.5k_2^+[ES_1] - k_3^+[\hat{P}][S_2] + k_3^-[\hat{P}S_2] + 1.5k_4^+[\hat{P}S_2] \\
\frac{d[S_2]}{dt} &= -k_3^+[\hat{P}][S_2] + k_3^-[\hat{P}S_2] \\
\frac{d[\hat{P}S_2]}{dt} &= k_3^+[\hat{P}][S_2] - k_3^-[\hat{P}S_2] - k_4^+[\hat{P}S_2]
\end{aligned} \tag{14}$$

If we define the notations $[S_1] = x_1$ $[E] = x_2$ $[ES_1] = x_3$ $[\hat{P}S_2] = x_4$ and include the conservation based expressions of $[\hat{P}] = 2([S_1^{tot}] - [S_1] - [ES_1])$ $[S_2] = ([S_2^{tot}] - 3[\hat{P}S_2] - 1.5[\hat{P}]) =$

$([S_2^{tot}] - 3[\hat{P}S_2] - 3([S_1^{tot}] - [S_1] - [ES_1]))$ we get the reduced lumped model in the form

$$\begin{aligned}
\frac{dx_1}{dt} &= -k_1^+ x_2 x_1 + k_1^- x_3 \\
\frac{dx_2}{dt} &= -k_1^+ x_2 x_1 + k_1^- x_3 + k_2^+ x_3 + u \\
\frac{dx_3}{dt} &= k_1^+ x_2 x_1 - k_1^- x_3 - k_2^+ x_3 \\
\frac{dx_4}{dt} &= k_3^+ 2([S_1^{tot}] - x_1 - x_3)([S_2^{tot}] - 3[\hat{P}S_2] - 3([S_1^{tot}] - x_1 - x_3)) \\
&\quad -k_3^- x_4 - k_4^+ x_4
\end{aligned} \tag{15}$$

In order to obtain the controllability distribution, we compute $\phi_2 = [f, g] = \frac{\partial g(x)}{\partial x} f(x) - \frac{\partial f(x)}{\partial x} g(x) = -\frac{\partial f(x)}{\partial x} g(x)$

$$= \begin{pmatrix} k_1^+ x_1 \\ k_1^+ x_1 \\ -k_1^+ x_1 \\ 0 \end{pmatrix}$$

$\phi_3 = [f, \phi_2] = \frac{\partial \phi_2}{\partial x} f(x) - \frac{\partial f(x)}{\partial x} \phi_2 =$

$$= \begin{pmatrix} k_1^+ k_1^- x_3 + k_1^{+2} x_1^2 + k_1^- k_1^+ x_1 \\ k_1^+ k_1^- x_3 + k_1^{+2} x_1^2 + k_1^- k_1^+ x_1 + k_1^+ x_1 k_2^+ \\ -k_1^+ k_1^- x_3 - k_1^{+2} x_1^2 - k_1^- k_1^+ x_1 - k_1^+ x_1 k_2^+ \\ 0 \end{pmatrix}$$

$\phi_4 = [f, \phi_3] = \frac{\partial \phi_3}{\partial x} f(x) - \frac{\partial f(x)}{\partial x} \phi_3 =$

$$= \begin{pmatrix} -k_1^{+3} x_1^2 x_2 + 3k_1^{+2} x_1 k_1^- x_3 + k_1^- k_1^{+2} x_1 x_2 + k_1^{-2} k_1^+ x_3 - k_1^+ k_2^+ k_1^- x_3 + k_1^{+2} x_2 k_1^- x_3 + \\ k_1^{+3} x_1^3 + 2k_1^{+2} x_1^2 k_1^- + k_1^{+2} x_1^2 k_2^+ + k_1^{-2} k_1^+ x_1 + k_1^- k_1^+ x_1 k_2^+ \\ \\ -k_1^{+3} x_1^2 x_2 + 3k_1^{+2} x_1 k_1^- x_3 + k_1^- k_1^{+2} x_1 x_2 + k_1^{-2} k_1^+ x_3 - k_1^{+2} k_2^+ x_1 x_2 + k_1^+ k_2^+ k_1^- x_3 + \\ k_1^{+2} x_2 k_1^- x_3 + k_1^{+3} x_1^3 + 2k_1^{+2} x_1^2 k_1^- + 2k_1^{+2} x_1^2 k_2^+ + k_1^{-2} k_1^+ x_1 + 2k_1^- k_1^+ x_1 k_2^+ + k_1^+ x_1 k_2^{+2} \\ \\ k_1^{+3} x_1^2 x_2 - 3k_1^{+2} x_1 k_1^- x_3 - k_1^- k_1^{+2} x_1 x_2 - k_1^{-2} k_1^+ x_3 + k_1^{+2} k_2^+ x_1 x_2 - k_1^+ k_2^+ k_1^- x_3 - \\ k_1^{+2} x_2 k_1^- x_3 - k_1^{+3} x_1^3 - 2k_1^{+2} x_1^2 k_1^- - 2k_1^{+2} x_1^2 k_2^+ - k_1^{-2} k_1^+ x_1 - 2k_1^- k_1^+ x_1 k_2^+ - k_1^+ x_1 k_2^{+2} \\ \\ -12k_3^+ x_1^2 k_1^+ k_2^+ - 2k_3^+ S_2^{tot} k_1^+ x_1 k_2^+ + 6k_3^+ x_4 k_1^+ x_1 k_2^+ + 12k_3^+ S_1^{tot} k_1^+ x_1 k_2^+ - 12k_3^+ x_3 k_1^+ x_1 k_2^+ \end{pmatrix}$$

which is of full rank in almost every point of the state space, so our **lumped reduced model is controllable**.

4.4 Simulation results

The validity range of the simplification method, the variable lumping with the adjustment of the stoichiometric coefficients was investigated by simulation both in the case of irreversible and reversible reaction schemes.

Basic reversible case In figure 7 a simulation of an irreversible reaction is shown, where the parameters were set as follows:

$$k_1^+ = 0.6 \quad k_1^- = 0.6 \quad k_2^+ = 0.4 \quad k_2^- = 0 \quad k_3^+ = 0.6 \quad k_3^- = 0.6 \quad k_4^+ = 0.4 \quad k_4^- = 0$$

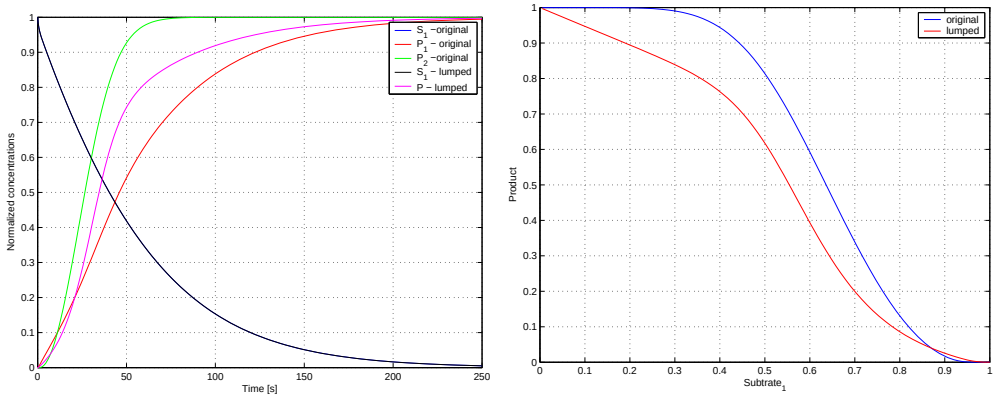

Figure 7: The normalized concentrations of the substrates and the products in the case of the original and the lumped reaction as functions of time, and as they evolve in the state space

In the figure P - *lumped* refers to $\hat{P}$. As it can be seen in Fig. 7, the lumped output $\hat{P}$ approximately corresponds to the average of the two products in this basic case, when the reaction rates of the reactions are of the same order of magnitude. This is in good agreement of the original engineering intuition behind variable lumping [20]: in the linear case the lumped variable is introduced formally as exactly the average of the two original ones.

Basic irreversible case Here the parameters were the following:

$$k_1^+ = 0.6 \quad k_1^- = 0.6 \quad k_2^+ = 0.4 \quad k_2^- = 0.1 \quad k_3^+ = 0.6 \quad k_3^- = 0.6 \quad k_4^+ = 0.4 \quad k_4^- = 0.1$$

As it can be seen in Fig. 8, in this case the final concentrations do not converge to the same value, yet the qualitative behavior is similar.

In the **irreversible case of higher reaction rates for the second reaction**, the results can be seen in figure 9. Here the parameters were:

$$k_1^+ = 0.6 \quad k_1^- = 0.6 \quad k_2^+ = 0.4 \quad k_2^- = 0 \quad k_3^+ = 6 \quad k_3^- = 6 \quad k_4^+ = 4 \quad k_4^- = 0$$

The breakpoint in the trajectory of $\hat{P}$ can be related to the total transformation of S_2.

In the **reversible case with higher reaction rates for the second reaction**, the results are depicted in Fig. 10 with the parameters:

$$k_1^+ = 0.6 \quad k_1^- = 0.6 \quad k_2^+ = 0.4 \quad k_2^- = 0.1 \quad k_3^+ = 6 \quad k_3^- = 6 \quad k_4^+ = 4 \quad k_4^- = 1$$

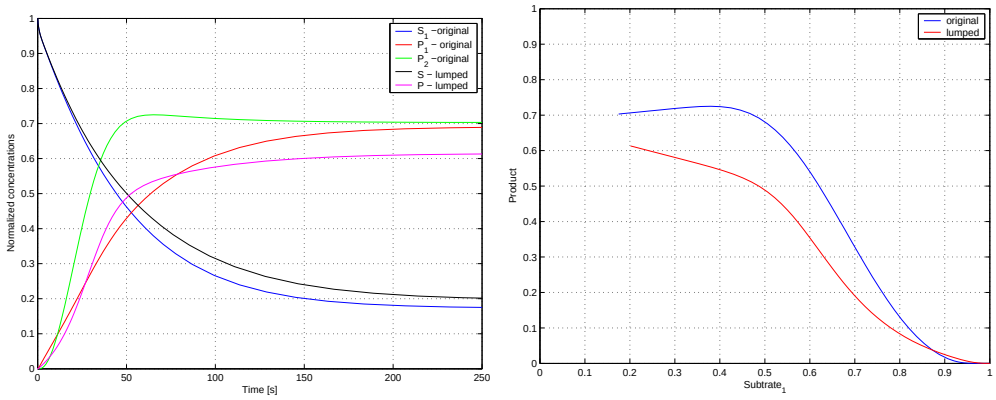

Figure 8: The normalized concentrations of the substrates and the products in the case of the original and the lumped reaction as functions of time, and as they evolve in the state space

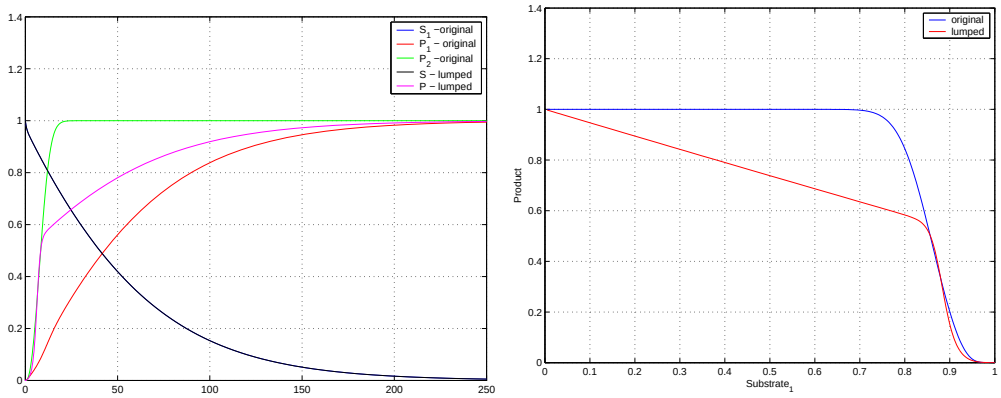

Figure 9: The normalized concentrations of the substrates and the products in the case of the original and the lumped reaction as functions of time, and as they evolve in the state space

In the **irreversible case of higher reaction rates for the first reaction** the results can be seen in figure 11, and the parameters were:

$$k_1^+ = 6 \quad k_1^- = 6 \quad k_2^+ = 4 \quad k_2^- = 0 \quad k_3^+ = 0.6 \quad k_3^- = 0.6 \quad k_4^+ = 0.4 \quad k_4^- = 0$$

In the **reversible case with higher reaction rates for the first reaction**, the results are shown in Fig. 12 with the parameters:

$$k_1^+ = 6 \quad k_1^- = 6 \quad k_2^+ = 4 \quad k_2^- = 1 \quad k_3^+ = 0.6 \quad k_3^- = 0.6 \quad k_4^+ = 0.4 \quad k_4^- = 0.1$$

4.5 Discussion: the validity range of the lumping

The simulation results show that the method of variable lumping with the adjustment of the stoichiometric coefficients can give acceptable results in the cases, when the second enzymatic activation reaction has a higher reaction rate than the first.

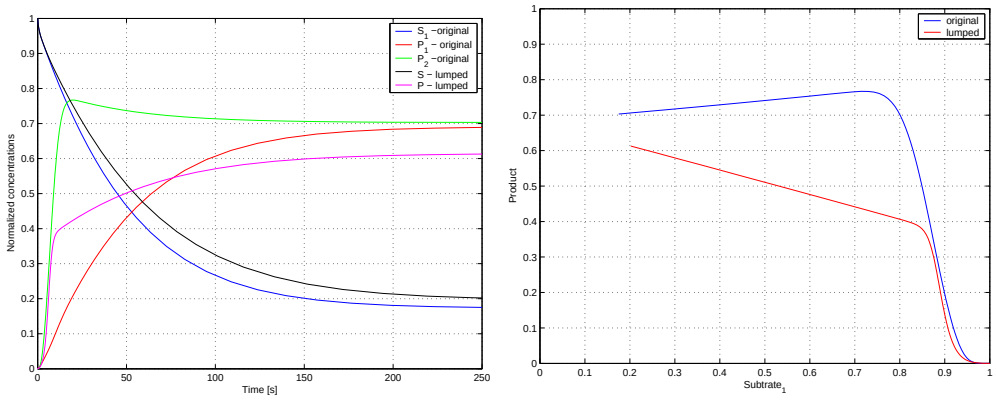

Figure 10: The normalized concentrations of the substrates and the products in the case of the original and the lumped reaction as functions of time, and as they evolve in the state space

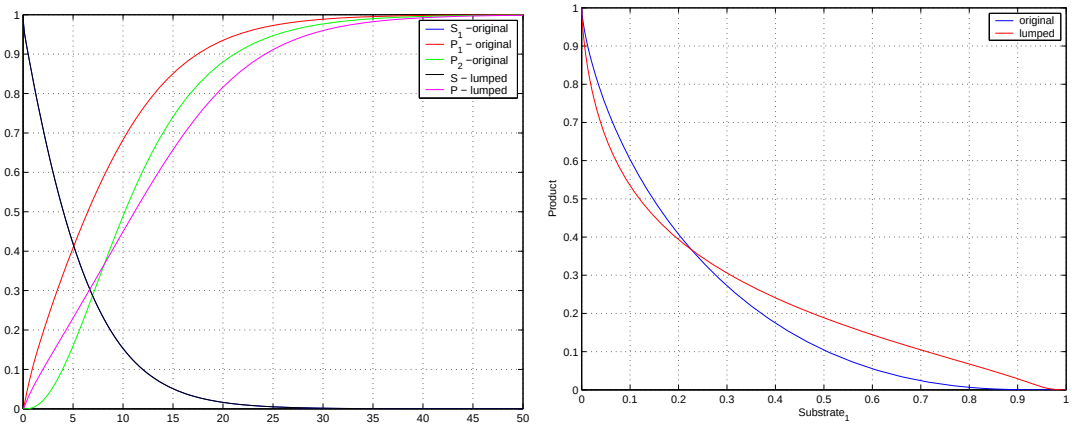

Figure 11: The normalized concentrations of the substrates and the products in the case of the original and the lumped reaction as functions of time, and as they evolve in the state space

In the reversible case, the results are worse, and the approximation is more acceptable in the case of reaction cascades with higher rates at the first level.

Note, that the parameters (the reaction rates) of the original and the lumped model were the same during the simulations. A proper tuning of the parameters of the lumped system would probably lead to the further improvement of the approximation. This could be a subject of future studies.

5 Cascade lumping with output transformation

Another way for performing the variable lumping transformation is to transform the output of the system. As mentioned before, the output of the original system is P_2. The behavior of

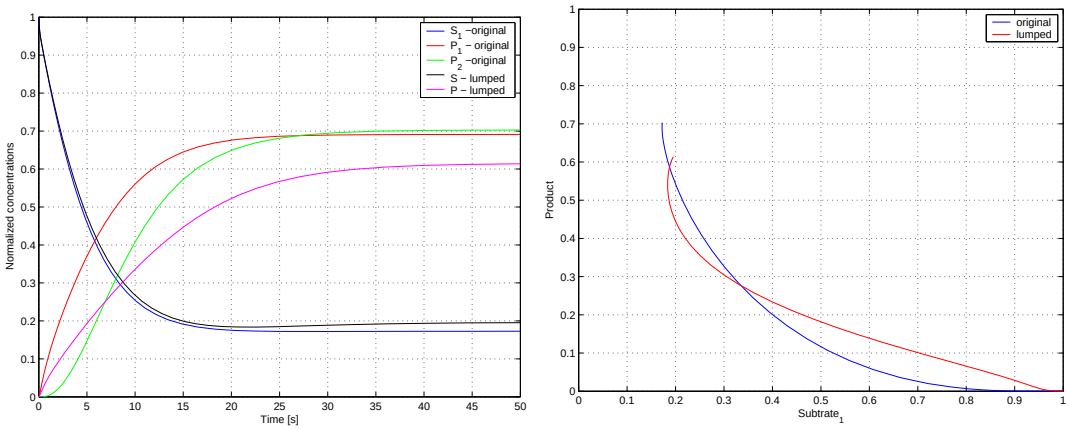

Figure 12: The normalized concentrations of the substrates and the products in the case of the original and the lumped reaction as functions of time, and as they evolve in the state space

$\hat{P}$ in the basic lumped case (without the variation of stoichiometric coefficients) approximates potentially P_1, so we can not use it directly as output. Therefore, the basic idea is that the transformed (eliminated) $\hat{S}_2$ ($=S_2^{tot} - S_2$) in the second level of reactions is considered as an estimate of the product P_2. It is important to note that **this approximation is only valid in the irreversible case, thus we assume irreversible reactions in this section**.

If we only change the output but not the stoichiometric coefficients, then we can use the following pseudo-reactions, which are identical to the ones which were described at the definition of cascade lumping (in Eq. (5), no adjustment of the stoichiometric coefficients):

$$E + S_1 \underset{k_1^-}{\overset{k_1^+}{\rightleftharpoons}} ES_1 \underset{k_2^-}{\overset{k_2^+}{\rightleftharpoons}} E + \hat{P}$$

$$\hat{P} + S_2 \underset{k_3^-}{\overset{k_3^+}{\rightleftharpoons}} \hat{P}S_2 \underset{k_4^-}{\overset{k_4^+}{\rightleftharpoons}} \hat{P}$$

The lumped reactions imply the following differential equations:

$$\begin{aligned}
\frac{d[S_1]}{dt} &= -k_1^+[E][S_1] + k_1^-[ES_1] \\
\frac{d[E]}{dt} &= -k_1^+[E][S_1] + k_1^-[ES_1] + k_2^+[ES_1] - k_2^-[E][\hat{P}] \\
\frac{d[ES_1]}{dt} &= k_1^+[E][S_1] - k_1^-[ES_1] - k_2^+[ES_1] + k_2^-[E][\hat{P}] \\
\frac{d[\hat{P}]}{dt} &= k_2^+[ES_1] - k_2^-[E][\hat{P}] - k_3^+[\hat{P}][S_2] + k_3^-[\hat{P}S_2] + k_4^+[\hat{P}S_2] - k_4^-[\hat{P}] \\
\frac{d[S_2]}{dt} &= -k_3^+[\hat{P}][S_2] + k_3^-[\hat{P}S_2] \\
\frac{d[\hat{P}S_2]}{dt} &= k_3^+[\hat{P}][S_2] - k_3^-[\hat{P}S_2] - k_4^+[\hat{P}S_2] + k_4^-[\hat{P}]
\end{aligned} \tag{16}$$

But now **the output of the model is** $y = [\hat{S}_2] = [S_2^{tot} - S_2]$**.**

The matrices of the algebraic description are:

$$\mathcal{N} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 1 & -1 \end{pmatrix} \qquad W = \begin{pmatrix} k_1^+[E][S_1] - k_1^-[ES_1] \\ k_2^+[ES_1] - k_2^-[E][\hat{P}] \\ k_3^+[\hat{P}][S_2] - k_3^-[\hat{P}S_2] \\ k_4^+[\hat{P}S_2] - k_4^-[\hat{P}] \end{pmatrix}$$

The steady-states, controllability and further properties can be derived in a way similar to the case of variable lumping with adjusting the stoichiometric coefficients. The only difference will be the more simple form of the expressions, which is a result of the more simple stoichiometric coefficients.

5.1 Simulation results

First the **basic irreversible case** was investigated with the parameters:

$$k_1^+ = 0.6 \quad k_1^- = 0.6 \quad k_2^+ = 0.4 \quad k_2^- = 0 \quad k_3^+ = 0.6 \quad k_3^- = 0.6 \quad k_4^+ = 0.4 \quad k_4^- = 0$$

the result of which is shown in Fig. 13

In the case of **higher reaction rates for the second reaction**, the results can be seen in figure 14 with the parameters

$$k_1^+ = 0.6 \quad k_1^- = 0.6 \quad k_2^+ = 0.4 \quad k_2^- = 0 \quad k_3^+ = 6 \quad k_3^- = 6 \quad k_4^+ = 4 \quad k_4^- = 0$$

In the case of **higher reaction rates for the first reaction** the results can be seen in figure 15, where the parameters were as follows

$$k_1^+ = 6 \quad k_1^- = 6 \quad k_2^+ = 4 \quad k_2^- = 0 \quad k_3^+ = 0.6 \quad k_3^- = 0.6 \quad k_4^+ = 0.4 \quad k_4^- = 0$$

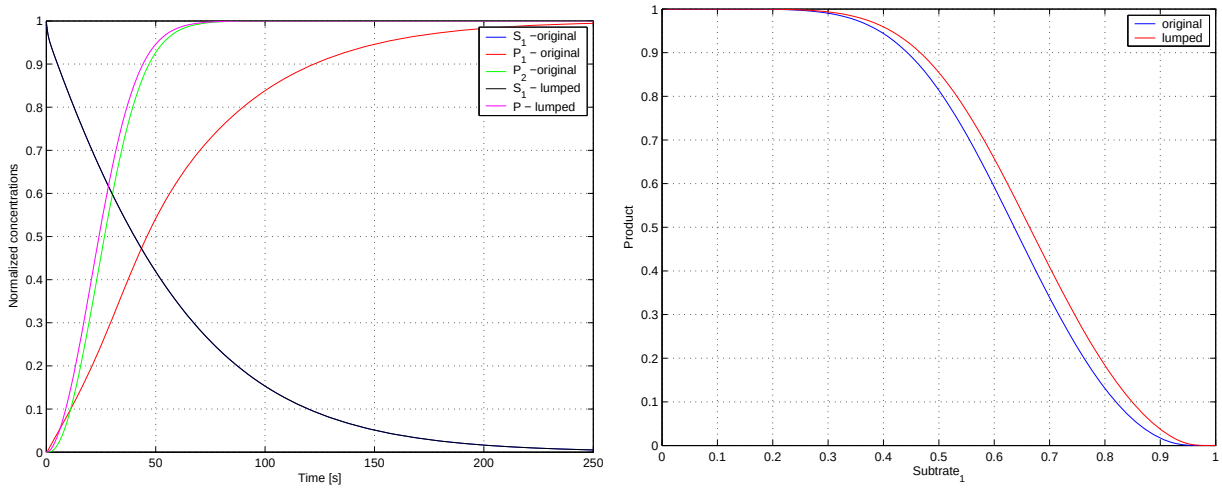

Figure 13: The normalized concentrations of the substrates and the products in the case of the original and the lumped reaction as functions of time, and as they evolve in the state space

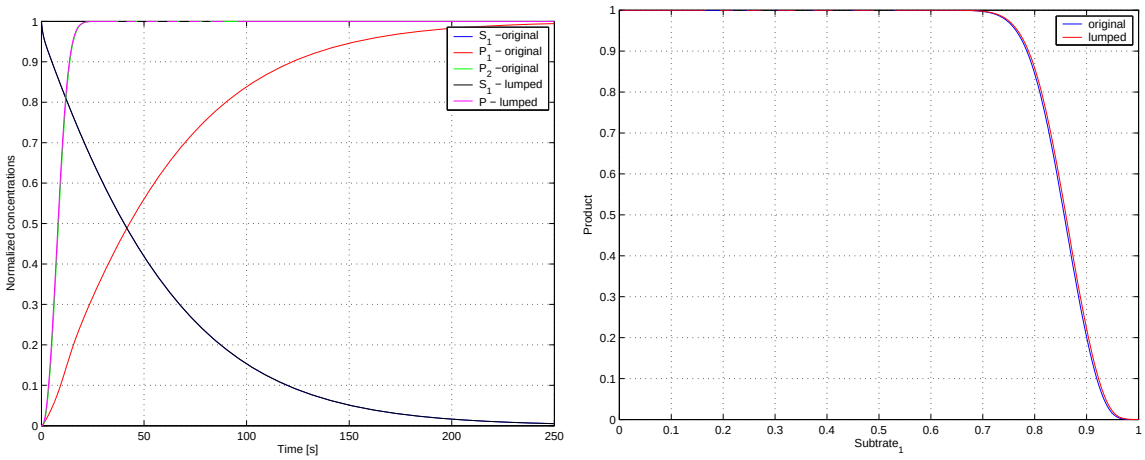

Figure 14: The normalized concentrations of the substrates and the products in the case of the original and the lumped reaction as functions of time, and as they evolve in the state space

5.2 Discussion: the validity range of the lumping

As it can be seen in the figures of the simulation results in the previous section, the output transformation method provides better approximation in all investigated cases of parameters than the one using the adjustment of stoichiometric coefficients. The worst, but sill acceptable result is achieved in the case, when the first activation reaction inhibits reaction constants one order of magnitude higher than the second level.

It should be emphasized again, that this approximation method is strongly related to irreversible reactions. Due to the fact that in this case the dynamics of $\hat{P}$ approximates P_1 in the original system, and the elimination of S_2 corresponds to the concentration of $\hat{P}$, a better

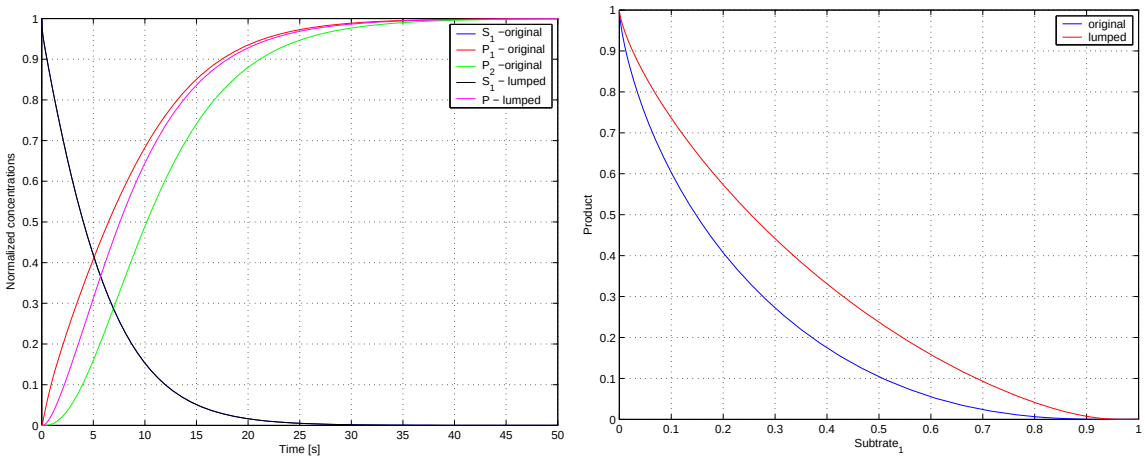

Figure 15: The normalized concentrations of the substrates and the products in the case of the original and the lumped reaction as functions of time, and as state trajectory

approximation of the final product concentration (P_2) is achieved by using the transformed output. The parameters of the reactions (original and lumped) were the same in this case, too.

6 The effect of quasi steady-state assumption on the original and the lumped model

In this section we will examine the effect of the other, widely applied method, the quasi steady-state (QSS) assumption-based model simplification method on the original, and on the lumped model. In all cases the intermediate complexes of the enzymatic reactions (ES_1,P_1S_2 and respectively ES_1 and $\hat{P}S_2$) are assumed to be in steady-state.

6.1 Quasi steady-state simplification of the original model

The assumption of the quasi steady-state of the intermediate complexes of the enzymatic reactions (ES_1,P_1S_2) implies the steady-state of the 3rd and 6th equation in the state space equations of the basic model (8).

$$\begin{aligned} \frac{d[ES_1]}{dt} &= k_1^+[E][S_1] - k_1^-[ES_1] - k_2^+[ES_1] + k_2^-[E][P_1] = 0 \\ \frac{d[P_1S_2]}{dt} &= k_3^+[S_2][P_1] - k_3^-[P_1S_2] - k_4^+[P_1S_2] + k_4^-[P_1][P_2] = 0 \end{aligned} \tag{17}$$

The explicit expression of $[ES_1]$ and $[P_1S_2]$ can be derived from the above equations:

$$
\begin{aligned}
[ES_1] &= \frac{k_1^+[E][S_1] + k_2^-[E][P_1]}{k_1^- + k_2^+} \\
[P_1S_2] &= \frac{k_3^+[S_2][P_1] + k_4^-[P_1][P_2]}{k_3^- + k_4^+}
\end{aligned} \tag{18}
$$

The simplified state space model can be obtained with the substitution of the above expressions into the remaining equations of (8).

$$
\begin{aligned}
\frac{d[S_1]}{dt} &= -k_1^+[E][S_1] + k_1^- \frac{k_1^+[E][S_1] + k_2^-[E][P_1]}{k_1^- + k_2^+} \\
\frac{d[E]}{dt} &= -k_1^+[E][S_1] + k_1^- \frac{k_1^+[E][S_1] + k_2^-[E][P_1]}{k_1^- + k_2^+} + k_2^+ \frac{k_1^+[E][S_1] + k_2^-[E][P_1]}{k_1^- + k_2^+} - k_2^-[E][P_1] \\
\frac{d[P_1]}{dt} &= k_2^+ \frac{k_1^+[E][S_1] + k_2^-[E][P_1]}{k_1^- + k_2^+} - k_2^-[E][P_1] - k_3^+[P_1][S_2] + k_3^- \frac{k_3^+[S_2][P_1] + k_4^-[P_1][P_2]}{k_3^- + k_4^+} \\
&\quad + k_4^+ \frac{k_3^+[S_2][P_1] + k_4^-[P_1][P_2]}{k_3^- + k_4^+} - k_4^-[P_1][P_2] \\
\frac{d[S_2]}{dt} &= -k_3^+[S_2][P_1] + k_3^- \frac{k_3^+[S_2][P_1] + k_4^-[P_1][P_2]}{k_3^- + k_4^+} \\
\frac{d[P_2]}{dt} &= k_4^+ \frac{k_3^+[S_2][P_1] + k_4^-[P_1][P_2]}{k_3^- + k_4^+} - k_4^-[P_1][P_2]
\end{aligned} \tag{19}
$$

If we perform some simplification on the equations, the model will be as follows:

$$
\begin{aligned}
\frac{d[S_1]}{dt} &= -k_1^+[E][S_1] + k_1^- \frac{k_1^+[E][S_1] + k_2^-[E][P_1]}{k_1^- + k_2^+} \\
\frac{d[E]}{dt} &= 0 \\
\frac{d[P_1]}{dt} &= k_2^+ \frac{k_1^+[E][S_1] + k_2^-[E][P_1]}{k_1^- + k_2^+} - k_2^-[E][P_1] \\
\frac{d[S_2]}{dt} &= -k_3^+[S_2][P_1] + k_3^- \frac{k_3^+[S_2][P_1] + k_4^-[P_1][P_2]}{k_3^- + k_4^+} \\
\frac{d[P_2]}{dt} &= k_4^+ \frac{k_3^+[S_2][P_1] + k_4^-[P_1][P_2]}{k_3^- + k_4^+} - k_4^-[P_1][P_2]
\end{aligned} \tag{20}
$$

The equation $\frac{d[E]}{dt} = 0$ is a trivial consequence of the assumption $\frac{d[ES_1]}{dt} = 0$ and the conservation law for the enzyme E.

Furthermore the steady-state of P_1S_2 and the conservation of S_2, P_1S_2 and P_2 implies that $\frac{d[P_2]}{dt} = -\frac{d[S_2]}{dt}$. This can be proved also in algebraic way from the corresponding differential equations. Therefore, the dimension of the resulting simplified model is 3.

6.1.1 Simulation results

The acceptability of the steady-state assumptions in the case of the original model was analyzed via simulations. The simulation results are depicted in the following figures.

The first simulation was performed in the case of **basic parameters in the irreversible case**:

$$k_1^+ = 0.6 \quad k_1^- = 0.6 \quad k_2^+ = 0.4 \quad k_2^- = 0 \quad k_3^+ = 0.6 \quad k_3^- = 0.6 \quad k_4^+ = 0.4 \quad k_4^- = 0$$

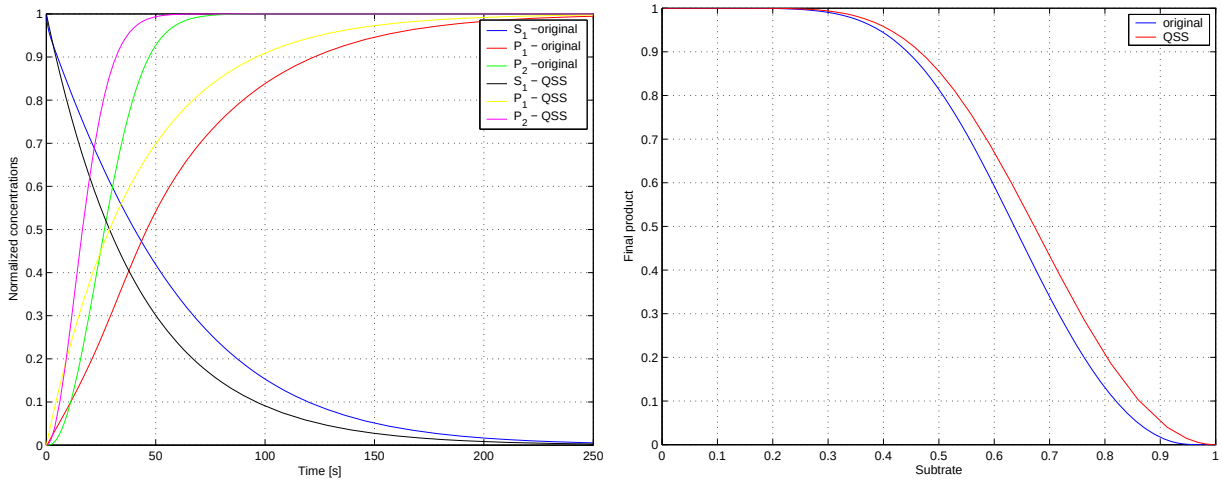

Figure 16: The normalized concentrations of the substrates and the products in the case of the original and the quasi steady-state simplified reaction as functions of time, and as they evolve in the state space

In the case of **higher reaction rates for the second reaction** the results can be seen in figure 17, where the parameters were

$$k_1^+ = 0.6 \quad k_1^- = 0.6 \quad k_2^+ = 0.4 \quad k_2^- = 0 \quad k_3^+ = 6 \quad k_3^- = 6 \quad k_4^+ = 4 \quad k_4^- = 0$$

In the case of **higher reaction rates for the first reaction** the results can be seen in figure 18 with the parameters:

$$k_1^+ = 6 \quad k_1^- = 6 \quad k_2^+ = 4 \quad k_2^- = 0 \quad k_3^+ = 0.6 \quad k_3^- = 0.6 \quad k_4^+ = 0.4 \quad k_4^- = 0$$

In the figures it can be seen that if the second reaction has higher reaction rate, then the steady-state based simplification produces more acceptable input output behavior, regarding the trajectories of P_2.

6.2 Quasi steady-state simplification of the lumped model

The assumption of the quasi steady-state of the intermediate complexes of the enzymatic reactions (ES_1,$\hat{P}S_2$ in the case of the lumped system) implies the steady-state of the 3rd and 6th equation in the state space equations of the basic lumped model (16). In this subsection

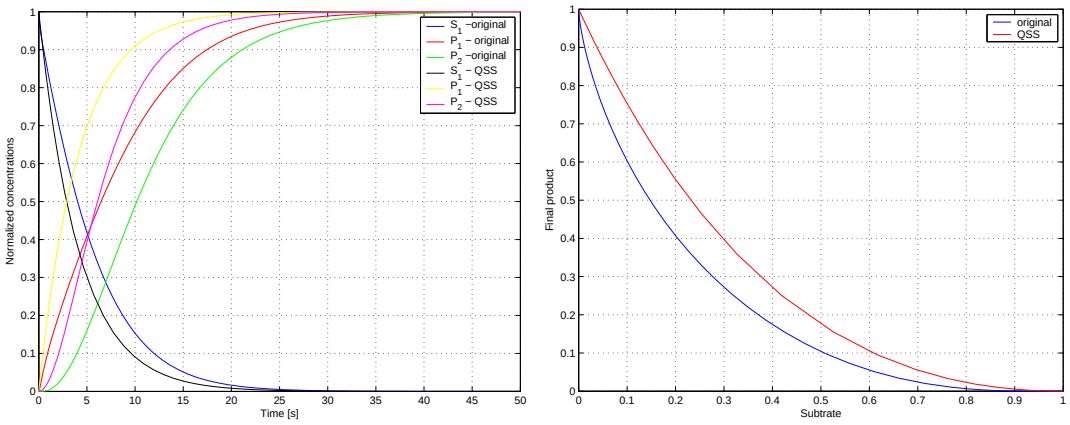

Figure 17: The normalized concentrations of the substrates and the products in the case of the original and the quasi steady-state simplified reaction as functions of time, and as they evolve in the state space

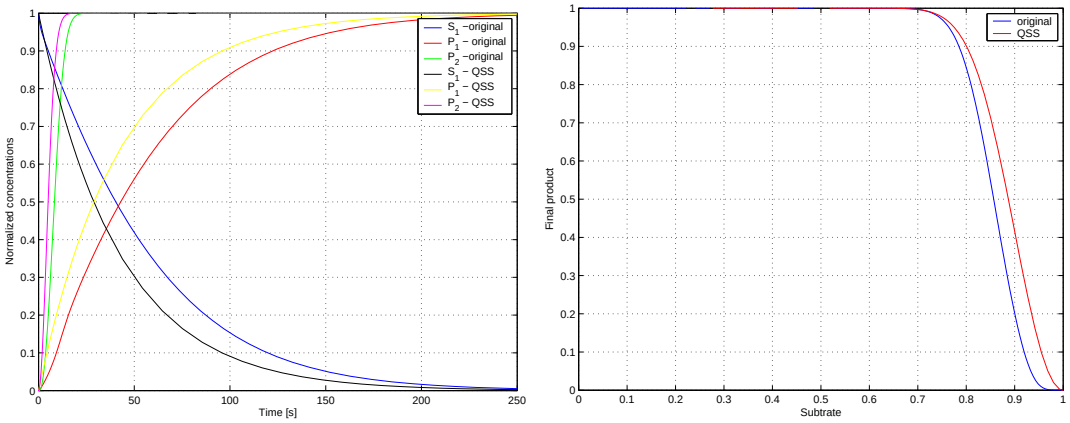

Figure 18: The normalized concentrations of the substrates and the products in the case of the original and the quasi steady-state simplified reaction as functions of time, and as they evolve in the state space

we compare the model derived with QSS-approximation from the original equations, with the QSS model we derive from the lumped model.

In the case of the variable lumped model, the output transformation method is used, because it provided better results as discussed in section 5. The effect of the quasi steady-state assumption is described as

$$\begin{aligned} \frac{d[ES_1]}{dt} &= k_1^+[E][S_1] - k_1^-[ES_1] - k_2^+[ES_1] + k_2^-[E][P_1] = 0 \\ \frac{d[\hat{P}S_2]}{dt} &= k_3^+[\hat{P}][S_2] - k_3^-[\hat{P}S_2] - k_4^+[\hat{P}S_2] + k_4^-[\hat{P}] = 0 \end{aligned} \tag{21}$$

The explicit expression for $[ES_1]$ and $[\hat{P}S_2]$ can be derived from the above equations:

$$\begin{aligned}[ES_1] &= \frac{k_1^+[E][S_1] + k_2^-[E][P_1]}{k_1^- + k_2^+} \\ \left[\hat{P}S_2\right] &= \frac{k_3^+[\hat{P}][S_2] + k_4^-[\hat{P}]}{k_3^- + k_4^+}\end{aligned} \tag{22}$$

The simplified state space model can be obtained with the substitution of the above expressions into the remaining equations of (16).

$$\begin{aligned}\frac{d[S_1]}{dt} &= -k_1^+[E][S_1] + k_1^- \frac{k_1^+[E][S_1] + k_2^-[E][P_1]}{k_1^- + k_2^+} \\ \frac{d[E]}{dt} &= -k_1^+[E][S_1] + k_1^- \frac{k_1^+[E][S_1] + k_2^-[E][P_1]}{k_1^- + k_2^+} + k_2^+ \frac{k_1^+[E][S_1] + k_2^-[E][P_1]}{k_1^- + k_2^+} - k_2^-[E][\hat{P}] \\ \frac{d[\hat{P}]}{dt} &= k_2^+ \frac{k_1^+[E][S_1] + k_2^-[E][P_1]}{k_1^- + k_2^+} - k_2^-[E][\hat{P}] - k_3^+[\hat{P}][S_2] + k_3^- \frac{k_3^+[\hat{P}][S_2] + k_4^-[\hat{P}]}{k_3^- + k_4^+} + \\ &\quad k_4^+ \frac{k_3^+[\hat{P}][S_2] + k_4^-[\hat{P}]}{k_3^- + k_4^+} - k_4^-[\hat{P}] \\ \frac{d[S_2]}{dt} &= -k_3^+[\hat{P}][S_2] + k_3^- \frac{k_3^+[\hat{P}][S_2] + k_4^-[\hat{P}]}{k_3^- + k_4^+}\end{aligned} \tag{23}$$

After rearrangement we obtain

$$\begin{aligned}\frac{d[S_1]}{dt} &= -k_1^+[E][S_1] + k_1^- \frac{k_1^+[E][S_1] + k_2^-[E][P_1]}{k_1^- + k_2^+} \\ \frac{d[E]}{dt} &= 0 \\ \frac{d[\hat{P}]}{dt} &= k_2^+ \frac{k_1^+[E][S_1] + k_2^-[E][P_1]}{k_1^- + k_2^+} - k_2^-[E][\hat{P}] \\ \frac{d[S_2]}{dt} &= -k_3^+[\hat{P}][S_2] + k_3^- \frac{k_3^+[\hat{P}][S_2] + k_4^-[\hat{P}]}{k_3^- + k_4^+}\end{aligned} \tag{24}$$

If we compare this system of ODE-s with Eq. (20), we can find out that the equation of P_1 and $\hat{P}$ and the two equations of S_1 and S_2 are the same. Furthermore the transformed output exhibits the same dynamics as the final product P_2 in the original (QSS-simplified) model (20).

This means that the effect of QSS-simplification implies the same dynamic properties in the case of the original and the lumped system.

6.2.1 Simulation results

In this subsection we will demonstrate that the effects of quasi steady-state assumptions on the output transformed lumped model produces the same approximation of input-output behavior than the QSS model derived from the original equations.

The first simulation was performed in the **irreversible case with basic parameters**:

$$k_1^+ = 0.6 \quad k_1^- = 0.6 \quad k_2^+ = 0.4 \quad k_2^- = 0 \quad k_3^+ = 0.6 \quad k_3^- = 0.6 \quad k_4^+ = 0.4 \quad k_4^- = 0$$

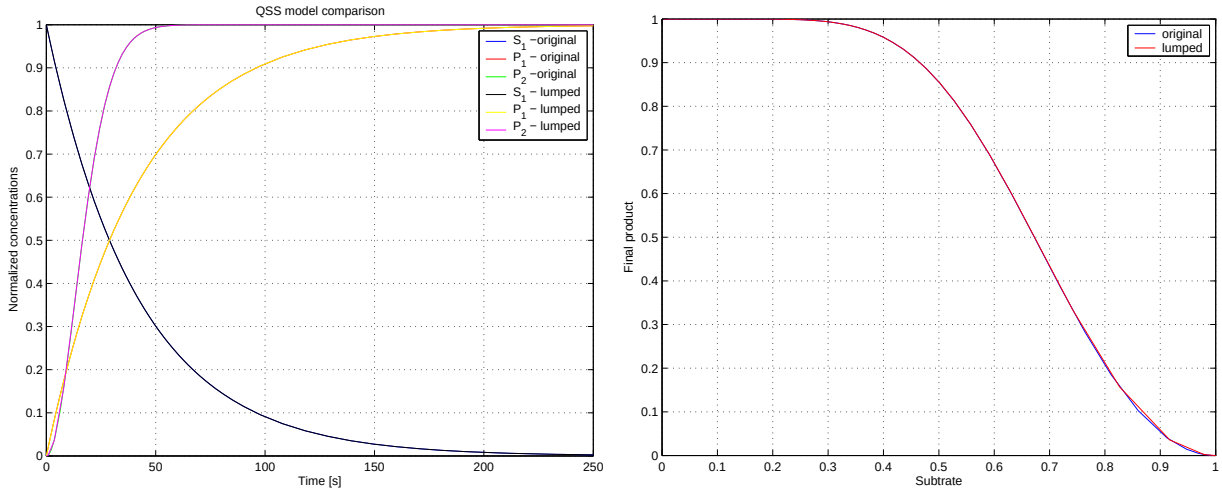

Figure 19: The normalized concentrations of the substrates and the products in the case of the original and the quasi steady-state simplified reaction as functions of time, and as they evolve in the state space

As it can be observed in figure 19, the QSS behavior of the original and lumped model is exactly the same, which is in good agreement with the algebraic results.

7 Generalizations

In this section the methods introduced in the previous sections are generalized to multi level cascade activation structures, and feedback regulated cascades. The generalization is illustrated on two level cascades, and on a MAPK signaling based feedback structure.

7.1 Two step cascade activation reactions

A trivial generalization of the cascade activation system is the structure, in which also the second product P_2 acts as an enzyme in the next level of the cascade. In fact such structures can exhibit 4-5 or even higher number of levels in biochemical signaling pathways, which indicates the necessity of the proposed simplification methods in the case of multi level structures.

7.2 Two-step cascade lumping with the adjustment of the stoichiometric coefficients

We can generalize the state-variable lumping method also for a two-step cascade, and define the corresponding reactions based on similar variation of the stochiometric coefficients. Here we lump together the components P_1, P_2 and P_3 into a single lumped component $\hat{P}$ in to elementary (i.e. two-component) variable lumping steps.

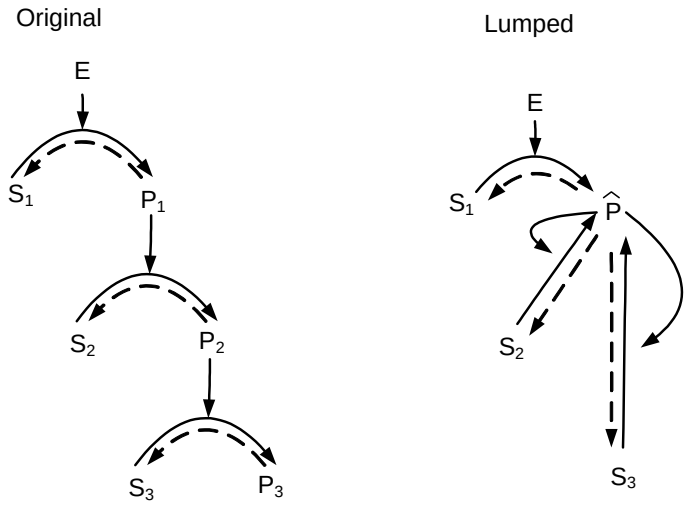

Figure 20: The activation scheme of the original and the lumped reaction. Irreversible case.

The reactions corresponding to the original and the lumped case are the following:

$$E + S_1 \underset{k_1^-}{\overset{k_1^+}{\rightleftharpoons}} ES_1 \underset{k_2^-}{\overset{k_2^+}{\rightleftharpoons}} E + P_1 \qquad E + S_1 \underset{k_1^-}{\overset{k_1^+}{\rightleftharpoons}} ES_1 \underset{k_2^-}{\overset{k_2^+}{\rightleftharpoons}} E + 1/3\hat{P} \tag{25}$$

$$P_1 + S_2 \underset{k_3^-}{\overset{k_3^+}{\rightleftharpoons}} P_1S_2 \underset{k_4^-}{\overset{k_4^+}{\rightleftharpoons}} P_1 + P_2 \qquad \hat{P} + S_2 \underset{k_3^-}{\overset{k_3^+}{\rightleftharpoons}} \hat{P}S_2 \underset{k_4^-}{\overset{k_4^+}{\rightleftharpoons}} 4/3\hat{P} \tag{26}$$

$$P_2 + S_3 \underset{k_5^-}{\overset{k_5^+}{\rightleftharpoons}} P_2S_3 \underset{k_6^-}{\overset{k_6^+}{\rightleftharpoons}} P_2 + P_3 \qquad \hat{P} + S_3 \underset{k_5^-}{\overset{k_5^+}{\rightleftharpoons}} \hat{P}S_3 \underset{k_6^-}{\overset{k_6^+}{\rightleftharpoons}} 4/3\hat{P} \tag{27}$$

The differential equations are derived in a similar way as in section 4.1.

7.2.1 Simulation results

The simulations have been performed on an extended version of the **basic irreversible case** with the parameters:

$$k_1^+ = 0.6 \quad k_1^- = 0.6 \quad k_2^+ = 0.4 \quad k_2^- = 0 \quad k_3^+ = 0.6 \quad k_3^- = 0.6$$

$$k_4^+ = 0.4 \quad k_4^- = 0 \quad k_5^+ = 0.6 \quad k_5^- = 0.6 \quad k_6^+ = 0.2 \quad k_6^- = 0$$

the result of which is seen in Fig. 21. A good qualitative agreement has been observed, but the quantitative agreement is far from being good.

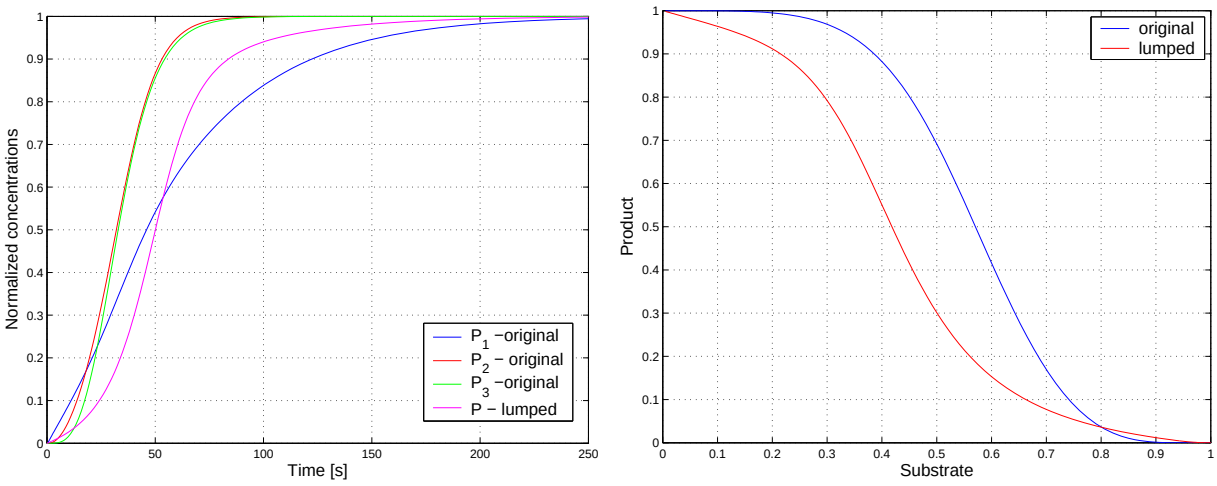

Figure 21: The normalized concentrations of the substrates and the products in the case of the original and the lumped reaction as functions of time, and as they evolve in the state space

7.3 Two-step cascade lumping with output transformation

Similarly to the two level case, the stoichiometric coefficients in the reaction steps remain unchanged to have

$$E + S_1 \underset{k_1^-}{\overset{k_1^+}{\rightleftharpoons}} ES_1 \underset{k_2^-}{\overset{k_2^+}{\rightleftharpoons}} E + \hat{P}$$

$$\hat{P} + S_2 \underset{k_3^-}{\overset{k_3^+}{\rightleftharpoons}} \hat{P}S_2 \underset{k_4^-}{\overset{k_4^+}{\rightleftharpoons}} \hat{P}$$

$$\hat{P} + S_3 \underset{k_5^-}{\overset{k_5^+}{\rightleftharpoons}} \hat{P}S_3 \underset{k_6^-}{\overset{k_6^+}{\rightleftharpoons}} \hat{P}$$

but the **output of the system is the transformed (eliminated) substrate in the 3rd reaction step**, i.e. $y = [\hat{S}_3] = [S_3]^{tot} - [S_3]$.

Effect on deficiency Similarly two the two level case, we reduced the number of species by 2, and the number of complexes by 1 by the variable lumping transformation. The number of linkage classes is also reduced by 1. Furthermore the rank of the matrix $\mathcal{N}$ is also conserved, so the deficiency is invariant under the lumping here, too.

7.3.1 Simulation results

Here again, the simulations have been performed on an extended version of the **basic irreversible case** with the parameters:

$$k_1^+ = 0.6 \quad k_1^- = 0.6 \quad k_2^+ = 0.4 \quad k_2^- = 0 \quad k_3^+ = 0.6 \quad k_3^- = 0.6$$

$$k_4^+ = 0.4 \quad k_4^- = 0 \quad k_5^+ = 0.6 \quad k_5^- = 0.6 \quad k_6^+ = 0.2 \quad k_6^- = 0$$

where the results are depicted in Fig. 22.

As we can see the output transformation based simplification method provides better performance also in the case of multi level cascades.

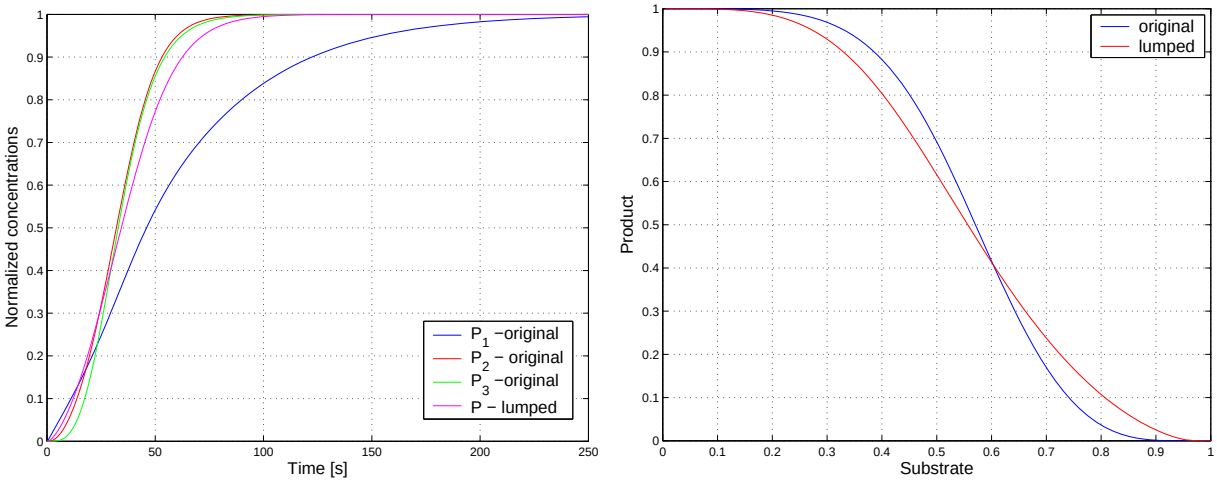

Figure 22: The normalized concentrations of the substrates and the products in the case of the original and the lumped reaction as functions of time, and as they evolve in the state space

7.4 Feedback regulated cascades

In the case of biochemical signaling pathways, a signaling cascade is often (mostly) feedback connected. One important example is the MAPK cascade, which is shortly summarized in the following to underline the biological importance of these reaction kinetic structures.

Four distinct MAPK cascades have been identified, each named after the subgroup of their MAPK components (ERK, JNK ,p38MAPK, and BMK).

The activation of MAPK cascades is initiated either by a small GTP-binding protein (smGP; RAS-family protein) or by an adaptor protein, which transmits the signal either directly or through a mediator kinase to the MAPK kinase kinase (MAP3K) level of the cascades. RAS recruits a RAF kinase from the cytosol to the cell membrane. Here, RAF is activated through a still not completely known process that involves interaction with adaptor proteins and changes in phosphorylation [18]. RAF family isozymes can phosphorylate and activate another kinase, MEK, which in turn phosphorylates and activates ERK. While RAF and MEK have a very restricted set of substrates, ERK features more than 70 substrates including nuclear transcription factors.

As described, the activation pathway of MAPK/ERK is composed by a linear sequence of two other kinases: MEK and RAF, indeed the signal gain is rather modest [8] and if it occurs,

mainly occurs at the MEK-ERK interface [24]. This activation structure has more interfaces available for regulation improving the fine tuning of signal flux through intra-cascade feedback regulation and cross talks with other pathways.

ERK feeds back to MEK activation at several levels. MEK needs to interact with RAF in order to become efficiently phosphorylated on the activating sites S218 and S222. Activated ERK can phosphorylate MEK on the adjacent T292 which precludes phosphorylation of S298 and reduces the formation of RAF-MEK complexes and MEK activation [9].

The properties of the MAPK cascade signaling are detailed in [16, 24, 1, 15].

If we analyze the three last step of the cascade, we can summarize the set of interactions in the simplified activation structure shown in Fig. 23: where E denotes Raf, the MAPK-Kinase-

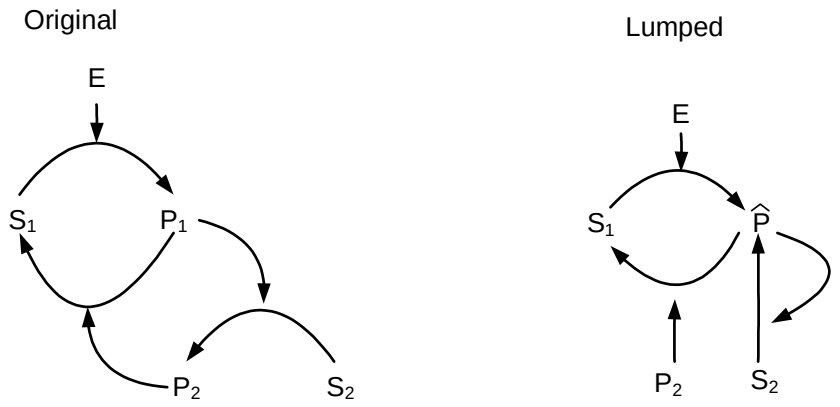

Figure 23: The activation scheme of the original and the lumped reaction. Irreversible case.

Kinase (MAP3K), S_1 and P_1 denotes the deactive/active MEK (the MAPKK), S_2 denotes the deactive ERK and P_2 denotes the activated ERK. Multisite phosphorylation and other interactions are neglected.

On the right side of the figure 23, the activation scheme of the output-transformation based lumped description can be seen.

The **original activation scheme** implies the following reactions describing the cascade activation:

$$E + S_1 \underset{k_1^-}{\overset{k_1^+}{\rightleftharpoons}} ES_1 \underset{k_2^-}{\overset{k_2^+}{\rightleftharpoons}} E + P_1$$

$$P_1 + S_2 \underset{k_3^-}{\overset{k_3^+}{\rightleftharpoons}} P_1S_2 \underset{k_4^-}{\overset{k_4^+}{\rightleftharpoons}} P_1 + P_2$$

The feedback regulation is described by the reaction

$$P_2 + P_1 \underset{k_5^-}{\overset{k_5^+}{\rightleftharpoons}} P_2P_1 \underset{k_6^-}{\overset{k_6^+}{\rightleftharpoons}} P_2 + S_1$$

The **lumped model** implies the following reactions:

$$E + S_1 \underset{k_1^-}{\overset{k_1^+}{\rightleftharpoons}} ES_1 \underset{k_2^-}{\overset{k_2^+}{\rightleftharpoons}} E + \hat{P}$$

$$\hat{P} + S_2 \underset{k_3^-}{\overset{k_3^+}{\rightleftharpoons}} \hat{P}S_2 \underset{k_4^-}{\overset{k_4^+}{\rightleftharpoons}} \hat{P}$$

Here the feedback regulation is described by the reaction

$$P_2 + \hat{P} \underset{k_5^-}{\overset{k_5^+}{\rightleftharpoons}} P_2\hat{P} \underset{k_6^-}{\overset{k_6^+}{\rightleftharpoons}} P_2 + S_1$$

where the output is $y = [P_2] = [S_2]^{tot} - [S_2] - [P_2\hat{P}]$.

7.4.1 Simulation results

Here again, the simulations have been performed on an extended version of the **basic irreversible case** with the parameters:

$$k_1^+ = 0.6 \quad k_1^- = 0.6 \quad k_2^+ = 0.4 \quad k_2^- = 0 \quad k_3^+ = 0.6 \quad k_3^- = 0.6$$

$$k_4^+ = 0.4 \quad k_4^- = 0 \quad k_5^+ = 0.6 \quad k_5^- = 0.6 \quad k_6^+ = 0.2 \quad k_6^- = 0$$

The results are shown in Fig. 24.

Analyzing the figures we can note that the output transformation based variable lumping transformation provides surprisingly good approximation of the feedback regulated cascade's output behavior.

8 Conclusions and Future work

In this paper we have shown that the variable lumping transformation can be generalized to a special class of nonlinear systems - the cascade activation reaction subclass of reaction kinetic systems.

Two ways of performing the lumping have been investigated: (i) the lumping with adjustment of stoichiometric coefficients and (ii) lumping with output transformation for the irreversible case. The two ways have been investigated by using simulation. The simulations were

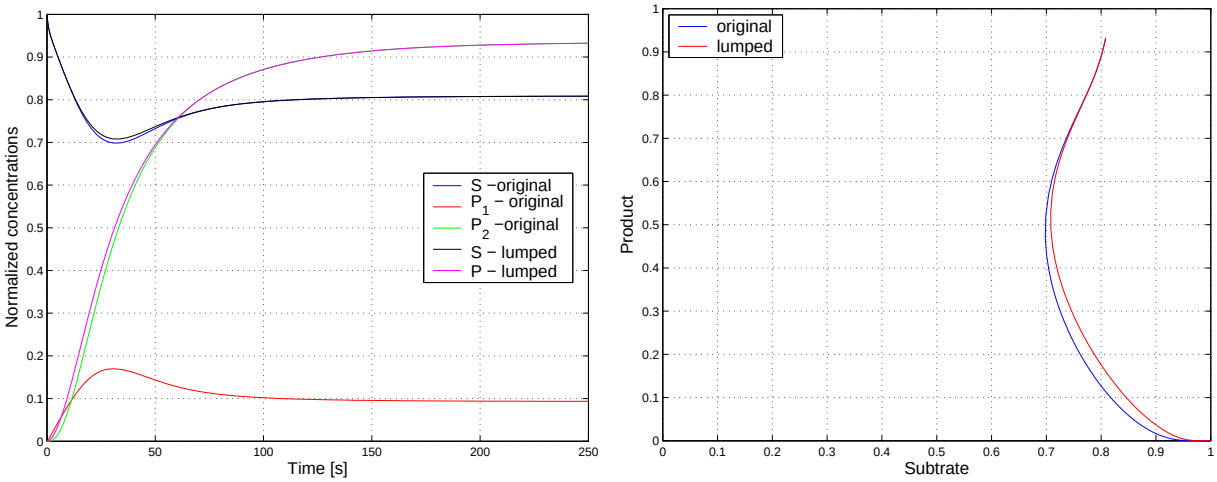

Figure 24: The normalized concentrations of the substrates and the products in the case of the original and the lumped reaction as functions of time, and as they evolve in the state space

performed by using different parameter sets to estimate the validity range of the approximation. All analyzed cases included reversible and irreversible reactions in the case of adjustment of stoichiometric coefficients, and included only irreversible reactions in the case of the output transformation method.

Based on the simulation results, the output transformation based method outperformed the one with adjustment of stoichiometric coefficients, but it is only applicable for irreversible reactions.

The controllability of the lumped model opened for enzyme was analyzed for both methods, and the resulted models were found to be controllable, which is in good agreement with the minimal model requirement.

Furthermore it was shown that the QSS (quasi steady-state) approximation of the original and the lumped reaction kinetic system results in the same model.

The lumping transformation was successfully generalized to multi level and simple feedback regulated cascades.

The results can serve as step on the way of the synthesis of minimal biochemical signaling models.

Future work In the future the real input-output behavior of the original and lumped system will be examined, by using various input signals. The first possibility to apply these stimuli is to consider an additive term in the equation of the enzyme concentrations, that corresponds to the properties of biochemical signaling pathways. An other form of input can be considered as substrate concentration, corresponding to biotechnological applications.

A further challenge is to derive the complete reaction kinetic description of a complex

biological cascade (for example the full MAPK cascade, which is quite often studied in the literature), and apply the described model simplification techniques to determine whether the reduced order models are capable of approximating the input-output behavior or not.

In the case when the reduced model seems capable of describing the qualitative dynamics of the pathway, a methodology of state and parameter estimation of in-vitro or in-vivo will be developed for future biomedical applications.

References

[1] A.R. Asthagiri and D.A. Lauffenburger. A computational study of feedback effects on signal dynamics in a mitogen-activated protein kinase (MAPK) pathway model. *Biotechnol. Prog.*, 17:227, 2001.

[2] E. Palomo Del Barrio, G. Lefebvre, and N. Bailly. Using model size reduction techniques for thermal control applications in buildings. *Energy and Buildings*, 33:1–14, 2000.

[3] J.M. Beaulieu, T.D. Sotinkova, S. Marion, R.J. Lefkowitz, R.R. Gainetdinov, and M.G. Caron. An Akt/β-Arrestin 2/PP2A signaling complex mediates dopaminergic neurotransmission and behavior. *Cell*, 122:261–273, 2005.

[4] I.T. Cameron and K.M. Hangos. *Process Modelling and Model Analysis*. Academic Press, London, 2001.

[5] V. Chickarmane, B.N. Kholodenko, and H.M. Sauro. Oscillatory dynamics arising from competitive inhibition and multisite phosphorylation. *Jornal of Theoretical Biology*, 244:68–76, 2006.

[6] G. Craciun and M. Feinberg. Multiple equilibria in complex chemical reaction networks: II The species-reaction graph. *Siam Journal of Applied Mathematics*, 66:1321–1338, 2006.

[7] G. Craciun and M. Feinberg. Understanding bistability in complex enzyme-driven reaction networks. *Proceedings of the National Academy of Sciences of the United States of America*, 103:8697–8702, 2006.

[8] A.S. Dhillon, C. Pollock, H. Steen, P.E. Shaw, H. Mischak, and W. Kolch. Cyclic AMP-dependent kinase regulates Raf-1 kinase mainly by phosphorylation of serine 259. *Molecular Cell Biology*, 2:3237–3246, 2002.

[9] S.T. Eblen, J.K. Slack-Davis, A. Tarcsafalvi, J.T. Parsons, M.J. Weber, and A.D. Catling. Mitogen-activated protein kinase feedback phosphorylation regulates MEK1 complex formation and activation during cellular adhesion. *Molecular Cell Biology*, 24:2308–2317, 2004.

[10] M. Feinberg. On chemical kinetics of a certain class. *Archive for Rational Mechanics and Analysis*, 46:1–41, 2004.

[11] Z.P. Gerdzten, P. Dautidis, and W.S. Hu. Non-linear reduction for kinetic models of metabolic reaction networks. *Metabolic Engineering*, 6:140–154, 2005.

[12] A.N. Gorban and I.V. Karlin. Method of invariant manifold for chemical kinetics. *Chemical engineering Science*, 58:4751–4768, 2003.

[13] A.N. Gorban, I.V. Karlin, and A.Y.Zinovyev. Invariant grids for reaction kinetics. *Physica A*, 333:106–154, 2004.

[14] J.E. Haag, A.V. Wouwer, and P. Bogaerts. Dynamic modeling of complex biological systems: a link between metabolic and macroscopic description. *Mathematical Biosciences*, 193:25–49, 2005.

[15] C.-Y.F. Huang and J.E. Ferrell Jr. Ultrasensitivity in the mitogen-activated protein kinase cascade. *PNAS*, 93:10078, 1996.

[16] J.E. Ferrel Jr. Tripping the switch fantastic: how a protein kinase cascade can convert graded inputs into switch-like outputs. *TIBS*, 21:460, 1996.

[17] B.N. Kholodenko, J.B. Hoek, and H.V. Westerhoff. Quantification of information transfer via cellular signal transduction pathways. *FEBS Letters*, 414:430–434, 1997.

[18] W. Kolch, F. Calder, and D. Gilbert. When kinases meet mathematics: the systems biology of MAPK signalling. *FEBS Letters*, 579:1891–1895, 2005.

[19] R. Lakner, K.M. Hangos, and I.T. Cameron. On minimal models of process systems. *chemical engineering science*, 60:1127–1142, 2005.

[20] A. Leitold, K.M. Hangos, and Zs. Tuza. Structure simplification of dynamic process models. *Journal of Process Control*, 12:69–83, 2002.

[21] T. Millat, E. Bullinger, J. Rohwer, and O. Wolkenhauer. Approximations and their consequences for dynamic modelling of signal transduction pathways. *Mathematical Biosciences*, 2006.

[22] L. Petzhold and W. Zhu. Model reduction for chemical kinetics: an optimization approach. *A.I.Ch.E. Journal*, 45:869–886, 1999.

[23] H.M. Sauro and B.N. Kholodenko. Quantitative analysis of signaling networks. *Biophysics & Molecular Biology*, 86:5–43, 2004.

[24] B. Schoeberl, C. Eichler-Jonsson, E.D. Gilles, and G. Müller. Computational modeling of the dynamics of the MAP kinase cascade activated by surface and internalized EGF receptors. *Nature Biotech.*, 20:370, 2002.

[25] J.J. Tyson, C.K. Chen, and B. Novák. Sniffers, buzzers, toggles and blinkers: dynamics of regulatory and signaling pathways in the cell. *Current Opinion in Cell Biology*, 14:221–231, 2003.

[26] A.R. Tzafriri and E.R. Edelman. Quasi-steady-state kinetics at enzyme and substrate concentrations in excess of the Michaelis-Menten constant. *Journal of Theoretical Biology*, 245:737–748, 2007.