Laplacian Dynamics and Kron Reduction in Species-Reaction Graphs of Chemical Reaction Networks

Manvel Gasparyan[†,1], Upinder S. Bhalla[2], Ovidiu Radulescu[1,*], Shodhan Rao[3,4,*]

[1]LPHI, University of Montpellier, CNRS and INSERM, Montpellier, France.
[2]National Centre for Biological Sciences, Tata Institute of Fundamental Research, Bangalore, India.
[3]Center for Biosystems and Biotech Data Science, Ghent University Global Campus, Incheon, Republic of Korea.
[4]Department of Data Analysis and Mathematical Modeling, Ghent University, Ghent, Belgium.

*Corresponding author(s). E-mail(s): ovidiu.radulescu@umontpellier.fr; shodhan.rao@ghent.ac.kr;
Contributing authors: manvel.gasparyan@ugent.be; bhalla@ncbs.res.in;

Abstract

We present a new method for deriving the dynamics of chemical reaction networks using the Laplacian matrix of the corresponding species–reaction graph, in contrast to previous works that use the Laplacian of the graph of complexes. Species-reaction graphs are bipartite graphs that contain two sets of vertices, one representing species and the other representing reactions, connected by directed edges that indicate relationships between them. Our approach starts by assigning appropriate edge weights to this bipartite graph, which are then used to compute the weighted graph Laplacian. This Laplacian reformulation of the system of differential equations governing the network dynamics emphasizes the flow of information throughout the chemical reaction network considered as causal network. As an application of this framework, we introduce a novel model reduction technique based on the Kron reduction of the weighted Laplacian matrix associated with the species-reaction graphs. Our systematic approach involves identifying nodes for deletion while preserving the bipartite structure, followed by constructing the Kron-reduced model. To demonstrate the effectiveness of our method, we apply it to a complex biochemical network, showing how model simplification facilitates analysis and interpretation of these systems.

Keywords: systems biology, deterministic models, bipartite graphs, graph Laplacian, Schur complement, model reduction, Glutamate receptor trafficking cycle.

1 Introduction

Chemical Reaction Networks (CRNs) serve as models for complex biochemical processes in cells and tissues. Studying these models is crucial for understanding diseases, developing new therapies, controlling bioengineering processes, and gaining insights into fundamental aspects of living systems [1, 2].

Models of CRNs are typically categorized as deterministic, stochastic, or hybrid. Deterministic models, commonly Ordinary Differential Equation (ODE) models [3] are valued for their simplicity, precision, and efficiency, particularly in systems with high reactant concentrations (large molecular counts, e.g., thousands per well-stirred volume), where random fluctuations in molecule numbers due to the inherent probabilistic nature of chemical reactions become insignificant compared to the average behavior. Stochastic models, such as Markov jump processes [4, 5], incorporate randomness to capture stochastic fluctuations in the species amounts. Though crucial for small-scale systems, these models are computationally intensive. Hybrid models [6, 7] combine deterministic and stochastic approaches, using deterministic methods for abundant species and stochastic methods for rare species, balancing accuracy and computational cost. This strategy reduces computational load while preserving the fluctuations of the species amounts. In this paper, we will focus solely on deterministic ODE models.

[†]Now affiliated with Department of Data Analysis and Mathematical Modeling, Ghent University, Ghent, Belgium, and Center for Biosystems and Biotech Data Science, Ghent University Global Campus, Incheon, Republic of Korea.

Graph theory offers powerful tools for analyzing CRNs. Network structures, such as (un)directed weighted graphs [8], multi-layer networks [9], temporal networks [10], and hierarchical graphs [11], facilitate the study of interactions in biochemical processes by highlighting structural properties such as feedback loops, key pathways, and their impact on qualitative dynamics, stability and robustness of the networks.

There is a growing need of mathematical tools for analyzing complex models of CRNs and for obtaining biological insight from the resulting models. The current state-of-the-art numerical tools for dynamical analysis of bioCRNs are known to suffer from a so-called curse-of-dimensionality. Therefore it is essential to develop mathematical techniques that can reduce a given kinetic model of a bioCRN to a simplified version that more or less mimics the behaviour of the original model, but has lesser number of parameters and differential equations.

For a comprehensive review of key model reduction strategies for CRNs, we refer the reader to [12] and [13]. Some of these model reduction methods have a graphical interpretation and can be described using graph rewriting operations, such as pruning (node elimination) and pooling (node aggregation) . Among them, are the popular Quasi-Steady-State (QSS) and Quasi-Equilibrium (QE) approximations [14, 15], the exact lumping methods [16–19] and the Kron reduction method [20–28]. In this paper, we revisit the latter in the context of a new formulation of CRN dynamics based on the Laplacian matrix of the Species-Reaction Graph (SR-graph).

Kron reduction [29] is a technique developed to simplify electrical grid models represented by directed graphs, by reducing network dimensionality while preserving voltage and current behavior at key nodes. It eliminates internal nodes, generating an equivalent network. The method was adapted for CRNs [20–28], simplifying CRNs represented by Graph of Complexess (GCs) (formal combinations of chemical species that appear together as reactants or products in a reaction) by eliminating some complexes . Following this adaptation, Kron reduction has also been employed as a tool for parameter estimation [23, 30] in CRN models.

The Kron reduction of the GC has limitations, as eliminating a complex may remove multiple species and reactions, potentially distorting the dynamics. For example, the CRN with the reactions $\mathcal{S}_1 \rightleftarrows \mathcal{S}_2$ and $\mathcal{S}_2+\mathcal{S}_3 \longrightarrow \mathcal{S}_4$ has four distinct complexes, namely $\{\mathcal{S}_1, \mathcal{S}_2, \mathcal{S}_2+\mathcal{S}_3, \mathcal{S}_4\}$.

Eliminating the complex $\mathcal{S}_2+\mathcal{S}_3$ would remove the second reaction entirely, leaving only the first one, which constitutes a large perturbation of the overall dynamics and is therefore not a good reduction. To address this, we propose a more flexible reduction using the SR-graph, a bipartite representation where species and reactions form disjoint node sets, with edges representing the relationships between them [31, 32].

This paper addresses two main problems: (i) establishing a link between the SR-graph and the system of ODEs derived from the stoichiometry and rate equations of a CRN, demonstrating how SR-graph edge weights based on stoichiometry and the state-dependent reaction rates can be used to reconstruct the ODE model; and (ii) proposing a novel model reduction technique that applies Kron reduction to the SR-graph of a CRN, rather than to the GC as done in prior work. Earlier, similar results were obtained for CRNs governed by Mass-Action Kinetics (MAK) [33]; here we generalize them to CRNs governed by general enzyme kinetics, which include MAK as a special case.

The structure of the paper is as follows. In Section 2, we introduce the necessary techniques for deriving the main results of this work, including the standard framework for modeling CRNs – a system of ODEs based on stoichiometry. In Section 3, we present the Laplacian of a SR-graph and examine its relationship with the stoichiometry-based ODE model. In Section 4, we propose a novel model reduction technique that applies Kron reduction to the SR-graph of CRNs. Finally, in Section 5, we validate our reduction approach using a real-world CRN example. We apply it to an Activity-dependent Movement of a Glutamate Receptor (AMPAR) signaling network, used to model changes of synaptic strength in neuroscience [34]. Notably, our model reduction preserves the property of bistability which is a crucial property of this model. Our reduction method eliminates 35.71% of the species and 48% of the reactions in this model, while preserving accuracies of 1.73% and 0.34% for trajectories leading to each of the model's two attractors.

Notations: $\mathbb{R}^n$ and $\mathbb{R}^n_{\geq 0}$ represents the n-dimensional Euclidean space and its non-negative orthant. For a vector $\mathbf{v} \in \mathbb{R}^n$, $v_i \in \mathbb{R}$ refers to its ith entry, i.e., the vector $\mathbf{v}$ can be represented as $\mathbf{v} = [v_i]_{i=1}^n$. Vector of length m composed entirely of ones is denoted as $\mathbf{1}_m$. Additionally, $\mathbf{0}_m$ denotes the vector of length m all of whose entries are zeros, and $\mathbf{0}_{m\times n}$ denotes the $m \times n$ matrix with each entry equal to zero. To avoid confusion, if there is no risk of ambiguity, we will use $\mathbf{0}$ in place of these notations. A_{ij} represents the entry of the matrix $\mathbf{A}$ that is located in the ith row and the jth column. For a finite set I, the number of elements in I is denoted by $|I|$, which represents the cardinality of the set. For a mutlivariable, real-valued function $\varphi : \mathbb{R}^n \to \mathbb{R}$, $\partial_i \varphi$ denotes the partial derivative of φ with respect to its ith variable.

2 Background

In this section, we follow the standard CRN framework to derive its mathematical formulation. We also introduce a technique for comparing the dynamics of the original and reduced CRN models, forming the basis of our modeling and reduction approach.

2.1 Stoichiometric mathematical models

We consider CRNs with only non-autocatalytic reactions; that is, no species appears simultaneously as both a substrate and a product of the same reaction. Autocatalysis is a rare phenomenon in biochemistry, and this phenomenon occurs only under special circumstances. Later on, in Section 4 of this manuscript, we will show that our reduction procedure leads to new reactions in which the complexes, i.e. the combinations of species of the left and right hand sides of the reactions of the original network are not retained in most cases. Hence our reduction procedure may lead to reactions that are not contained in the original CRN. In order to make the reactions of our reduced CRN acceptable to biochemists, we ensure that the reduced network arising from a given non-autocatalytic CRN in also non-autocatalytic.

Assume that n_s distinct chemical species $\mathcal{S}_i$, $i = 1, \ldots, n_s$, are involved in n_r irreversible reactions $\mathcal{R}_i$, $i = 1, \ldots, n_r$. We do not exclude the possibility that reactions can be reversible. In this case, we consider each direction of a reversible reaction as a single separate reaction. Such a CRN can generally be written in the following form:

$$\mathcal{R}_j : \sum_{i=1}^{n_s} \alpha_{ij}\mathcal{S}_i \xrightarrow{v_j} \sum_{i=1}^{n_s} \beta_{ij}\mathcal{S}_i, \quad j = 1, \ldots, n_r, \tag{1}$$

where α_{ij} and β_{ij} are non-negative constants representing the number of molecules of the ith species $\mathcal{S}_i$ in the substrates and the products of the jth reaction $\mathcal{R}_j$. Here, v_j denotes the rate of the jth reaction $\mathcal{R}_j$, which is a multivariable, real-valued, and non-negative function of the species' concentration vector. Note that we do not exclude the possibility of having a reaction in which all the constants α_{ij} are zero simultaneously. For example, if there exists a reaction $\mathcal{R}_k$ such that $\alpha_{ik} = 0$ for all i, then the corresponding reaction is of the form $\emptyset \longrightarrow \sum_{i=1}^{n_s} \beta_{ik}\mathcal{S}_i$, where $\emptyset$ indicates the absence of any reactant species. Such reactions represent inflows into the CRN, modeling the spontaneous creation or external supply of the product species involved. Similarly, it is also possible to have a reaction in which all the constants β_{ij} are zero. In this case, the reaction represents an outflow from the CRN and modeling the degradation or removal of the substrate species involved.

The connection between species and reactions is established through an $n_s \times n_r$ stoichiometric matrix, denoted by $\mathbf{S}$. The elements of this matrix are defined as $S_{ij} = \beta_{ij} - \alpha_{ij}$. Let $\mathbf{s} = [s_i]_{i=1}^{n_s}$ and $\mathbf{v} = [v_j]_{j=1}^{n_r}$ represent the vector of species concentrations and the vector of reaction rates, respectively. The fundamental framework characterizing the evolution of the species concentration vector is described by the following system of ODEs:

$$\frac{d\mathbf{s}}{dt} = \mathbf{S}\mathbf{v}(\mathbf{s}) = \left[\sum_{j=1}^{n_r} (\beta_{ij} - \alpha_{ij})\, v_j(\mathbf{s})\right]_{i=1}^{n_s}. \tag{2}$$

Buffered species—i.e., those that remain constant, do not participate as substrates or products, but act as modifiers—are not represented by an ODE of the form given in Eq. (2), as their concentration derivatives are zeros. However, they appear in the model through the rate expressions as constants, influencing reaction kinetics and potentially being treated as parameters. We further assume, that rates of zero order reactions are constant inputs to the system, i.e.

$$v_j(\mathbf{s}) = k_j, \text{ if } \alpha_{ij} = 0 \text{ for all } j, \tag{3}$$

where k_j is a constant.

To illustrate the aforementioned approach for deriving the system of ODEs, we present a simple example of a CRN and demonstrate the modelling process.

Example 1. *Consider the following example of a CRN of five distinct chemical species and three unidirectional reactions:*

$$\mathcal{S}_1 + 2\mathcal{S}_2 \underset{v_2}{\overset{v_1}{\rightleftharpoons}} 3\mathcal{S}_3 + \mathcal{S}_4, \qquad 4\mathcal{S}_4 \xrightarrow{v_3} 5\mathcal{S}_5 + 2\mathcal{S}_1. \tag{4}$$

The 5×3 stoichiometric matrix $\mathbf{S}$ of the CRN (4) is:

$$\mathbf{S} = \begin{bmatrix} -1 & 1 & 2 \\ -2 & 2 & 0 \\ 3 & -3 & 0 \\ 1 & -1 & -4 \\ 0 & 0 & 5 \end{bmatrix},$$

and the corresponding system of ODEs can be expanded to:

$$\frac{ds_1}{dt} = -v_1 + v_2 + 2v_3, \qquad \frac{ds_3}{dt} = 3v_1 - 3v_2, \qquad \frac{ds_5}{dt} = 5v_3.$$
$$\frac{ds_2}{dt} = -2v_1 + 2v_2, \qquad \frac{ds_4}{dt} = v_1 - v_2 - 4v_3,$$

2.2 Enzyme kinetics law

Every CRN can be uniquely represented by a set of ODEs, as shown in Eq. (2), regardless of its rate governing laws, i.e. the governing principles for constructing the reaction rates as a function of the species concentrations. In this paper, we consider a set of rate governing laws known as General Enzyme Kinetics Law (GEKL) [20], that cover the most prevalent rate laws globally, governing numerous real-life bio-CRNs. According to this law, the rate v_j of the jth reaction is given by:

$$v_j(\mathbf{s}) = g_j(\mathbf{s})\varphi_j(\mathbf{s}), \tag{5}$$

where φ_j is a multivariable monomial of the species concentration vector $\mathbf{s}$ defined by:

$$\varphi_j(\mathbf{s}) = \prod_{i=1}^{n_s} s_i^{\alpha_{ij}}, \tag{6}$$

and g_j is a multivariable, real-valued, and non-negative rational function of its arguments.

MAK is an example of GEKL, where the rational terms are constant functions representing the rate constants of the corresponding reactions. A comprehensive list of well-known GEKLs, such as Michaelis-Menten kinetics [35] and Hill kinetics [36], can be found in the appendix of [37]. As a matter of fact, GEKL is very general, as it encompasses all kinetic laws that depend rationally on species concentrations.

Note that the reaction rate vector $\mathbf{v}$, of a CRN governed by GEKL can be written as $\mathbf{v} = \mathbf{\Gamma_g}\boldsymbol{\varphi}$, where $\boldsymbol{\varphi} = [\varphi_j]_{j=1}^{n_s}$, $\mathbf{g} = [g_j]_{j=1}^{n_s}$, and $\mathbf{\Gamma_g} = \mathrm{diag}(\mathbf{g})$. This representation is especially useful for improving computational efficiency and optimizing automation.

Example 2. *With reference to the CRN depicted in Eq. (4), assuming its reactions are governed by GEKL, one possible formulation for the expressions of the reaction rates is as follows:*

$$v_1 = \frac{k_1 s_1 s_2^2}{1 + s_1 s_2^2}, \qquad v_2 = \frac{k_2 s_3^3 s_4}{1 + s_3^3 + s_4}, \qquad v_3 = \frac{k_3 s_4^4}{1 + s_4}, \tag{7}$$

where k_i are positive constants. In this case, the monomials corresponding to the reactions are given by:

$$\varphi_1 = s_1 s_2^2, \qquad \varphi_2 = s_3^3 s_4, \qquad \varphi_3 = s_4^4,$$

and the rational terms within the expressions of the reaction rates are as follows:

$$g_1 = \frac{k_1}{1 + s_1 s_2^2}, \qquad g_2 = \frac{k_2}{1 + s_3^3 + s_4}, \qquad g_3 = \frac{k_3}{1 + s_4}.$$

2.3 Species-reaction graphs

A SR-graph [31, 32] is a graphical framework for representing the dynamic relationships between species and reactions in a CRN. It provides insights into chemical transformations by illustrating the CRN structure.

The SR-graph of a CRN is a weighted directed bipartite graph $(\mathcal{V}_s, \mathcal{V}_r, \mathcal{E})$, where $\mathcal{V}_s$ and $\mathcal{V}_r$ represent species and reaction nodes, respectively. Reversible reactions are modeled as two irreversible ones. The edges $\mathcal{E}$ follow these rules: (i) if $\mathcal{S}_i$ is a reactant in reaction $\mathcal{R}_j$, the edge goes from species $\mathcal{S}_i$ to reaction $\mathcal{R}_j$; (ii) if $\mathcal{S}_i$ is a product of reaction $\mathcal{R}_j$, the edge goes from reaction $\mathcal{R}_j$ to species $\mathcal{S}_i$. Later, we will demonstrate how to appropriately assign the edge weights and explain their role in our model reduction procedure.

Example 3. *To clarify the construction of the SR-graph for a given CRN, we use the example in Eq. (4). The CRN consists of five species and three reactions, leading to a node set $\mathcal{V}_s = \{\mathcal{S}_1, \ldots, \mathcal{S}_5\}$ for species and $\mathcal{V}_r = \{\mathcal{R}_1, \ldots, \mathcal{R}_3\}$ for reactions. Since $\mathcal{S}_1$ is a reactant in $\mathcal{R}_1$, there is a directed edge from $\mathcal{S}_1$ to $\mathcal{R}_1$ in the SR-graph. Similarly, as $\mathcal{S}_3$ is a product in $\mathcal{R}_1$, there is a directed edge from $\mathcal{R}_1$ to $\mathcal{S}_3$. The remaining edges are constructed in the same manner. The resulting graph is shown in Fig. 1, where species nodes are represented as blue circles, reaction nodes as red squares, edges from species to reactions in orange, and edges from reactions to species in green.*

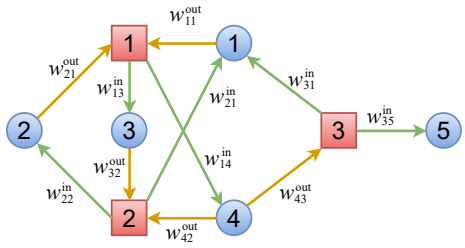


Fig. 1: The SR-graph of the enzymatic CRN in Eq. (4), with species nodes (blue circles) and reaction nodes (red squares). Edges from species to reactions are orange, while those from reactions to species are green.

2.4 Symmetrized error integral

We present a metric designed to quantify the differences in the dynamical behavior between the original model and its corresponding reduced model. To achieve this, we utilize the error integral.

Definition 1 (Error integral)**.** *Let $\widehat{s}_i$ and s_i represent the concentrations of the ith species in the reduced model and the original model, respectively. Further let T_i denote the duration of the time interval during which we examine the differences in the concentration behavior of the ith species between a base trajectory of the original model and the corresponding trajectory of the reduced model. The error integral between the original model and the reduced model is defined as:*

$$\theta(\mathbf{s}, \widehat{\mathbf{s}}) = \frac{1}{|J_s^c|} \sum_{i \in J_s^c} \frac{1}{T_i} \int_0^{T_i} \left| 1 - \frac{\widehat{s}_i(t) + \varepsilon_i}{s_i(t) + \varepsilon_i} \right| dt, \tag{8}$$

where J_s^c is the set of indices in the original model corresponding to the species that are preserved in the reduced model and ε_i is a small positive value used to prevent division by zero and large errors that may arise when $\widehat{s}_i(t) \to 0$, $s_i(t) \to 0$.

The concept of the error integral is adapted from [20], where all T_i are assigned the same value. In contrast, this study allows each species to be evaluated over its own time interval. For zero-deficiency CRNs, a suitable choice for T_i is the settling time of the ith species—defined as the time it takes for its concentration to enter and remain within a small, predefined neighborhood of its steady state s_i^*. An automated method for computing these settling times is presented in [22, 23], where T_i is determined as the earliest time the concentration s_i enters this neighborhood and stays within it for the remainder of the simulation. The overall network settling time is then given by $\max_i T_i$, which is used uniformly for all species in [20].

It is important to note that the error integral, as defined in Eq. (8), is not symmetric. This lack of symmetry arises because the measure is dependent on the direction of comparison between the full and reduced models; the error quantified when comparing the full model to the reduced one is not necessarily the same as when the comparison is reversed. A symmetric error integral, on the other hand, is advantageous because it treats the discrepancies between the full and reduced models equally, regardless of the direction of comparison. For this purpose, we use the symmetrized version of the error integral, which was first introduced and applied in [38].

Definition 2 (Symmetrized error integral)**.** *Let θ be the error integral between the original model and the reduced model, as defined in Eq. (8). The symmetrized error integral between the original model and the reduced model is defined as:*

$$\Theta(\mathbf{s}, \widehat{\mathbf{s}}) = \frac{1}{2}\Big(\theta(\mathbf{s}, \widehat{\mathbf{s}}) + \theta(\widehat{\mathbf{s}}, \mathbf{s})\Big). \tag{9}$$

This symmetry provides a more balanced and comprehensive assessment of the differences in dynamical behavior between the two models, ensuring that both overestimations and underestimations are accounted for equivalently.

2.5 Schur complement

The Schur complement [39] plays a central role in Kron reduction, as it naturally arises from the Gaussian elimination of variables in linear systems of equations. To illustrate this, let us consider a block-partitioned linear system and show how one set of variables can be expressed in terms of the others, leading to the Schur complement. Consider the linear system:

$$\mathbf{L}_{11}\mathbf{x}_1 + \mathbf{L}_{12}\mathbf{x}_2 = 0,$$

$$\mathbf{L}_{21}\mathbf{x}_1 + \mathbf{L}_{22}\mathbf{x}_2 = 0,$$

where $\mathbf{L}_{11} \in \mathbb{R}^{n\times n}$, $\mathbf{L}_{12} \in \mathbb{R}^{n\times m}$, $\mathbf{L}_{21} \in \mathbb{R}^{m\times n}$, and $\mathbf{L}_{22} \in \mathbb{R}^{m\times m}$. Assuming that $\mathbf{L}_{22}$ is invertible, we can express $\mathbf{x}_2$ in terms of $\mathbf{x}_1$ as $\mathbf{x}_2 = -\mathbf{L}_{22}^{-1}\mathbf{L}_{21}\mathbf{x}_1$. Substituting this relation into Eq. (10) yields the reduced equation:

$$\left(\mathbf{L}_{11} - \mathbf{L}_{12}\mathbf{L}_{22}^{-1}\mathbf{L}_{21}\right)\mathbf{x}_1 = 0,$$

which illustrates the construction of the Schur complement.

Definition 3 (Schur complement)**.** *Let $\mathbf{L}_{11} \in \mathbb{R}^{n\times n}$, $\mathbf{L}_{12} \in \mathbb{R}^{n\times m}$, $\mathbf{L}_{21} \in \mathbb{R}^{m\times n}$, and $\mathbf{L}_{22} \in \mathbb{R}^{m\times m}$ be constant matrices such that $\mathbf{L}_{22}$ is invertible. Consider the following $(n+m)\times(n+m)$ block matrix:*

$$\mathbf{L} = \begin{bmatrix} \mathbf{L}_{11} & \mathbf{L}_{12} \\ \mathbf{L}_{21} & \mathbf{L}_{22} \end{bmatrix}.$$

The Schur complement of the block $\mathbf{L}_{22}$ of the matrix $\mathbf{L}$ is the $n\times n$ matrix $\widehat{\mathbf{L}}$ defined as:

$$\widehat{\mathbf{L}} = \mathbf{L}_{11} - \mathbf{L}_{12}\mathbf{L}_{22}^{-1}\mathbf{L}_{21}.$$

Remark 1. *The definition of the Schur complement requires the block $\mathbf{L}_{22}$ to be invertible. If $\mathbf{L}_{22}$ is singular, the classical Schur complement is not defined. In such cases, a common generalization replaces the inverse $\mathbf{L}_{22}^{-1}$ with the Moore–Penrose pseudoinverse $\mathbf{L}_{22}^{\dagger}$, i.e $\widehat{\mathbf{L}} = \mathbf{L}_{11} - \mathbf{L}_{12}\mathbf{L}_{22}^{\dagger}\mathbf{L}_{21}$, leading to the generalized Schur complement.*

3 Dynamics on a species-reaction graph

In this section we present one of our study's key findings: we show how to derive a new representation of the enzymatic CRNs using their SR-graphs. By defining edge weights, we utilize the graph Laplacian of the SR-graph to construct a system of ODEs that accurately reflect the CRN dynamics.

3.1 Graph Laplacian

We now demonstrate how to appropriately define the weights of the edges in the SR-graph and subsequently explain the construction of the weighted directed Laplacian matrix of the SR-graph associated with the given CRN using the weighted directed adjacency matrix. We define the weights of the edges as follows:

1. For an edge with the node corresponding to species $\mathcal{S}_i$ as the source and the node corresponding to reaction $\mathcal{R}_j$ as the target, the weight of this edge is $w_{ij}^{\text{out}} = g_j\partial_i\varphi_j = \alpha_{ij}v_j/s_i$. Here, as previously described, g_j and φ_j denote the rational and the monomial terms, respectively, in the expression for the rate v_j of reaction $\mathcal{R}_j$.
2. For an edge with the node corresponding to reaction $\mathcal{R}_j$ as the source and the node corresponding to species $\mathcal{S}_i$ as the target, the weight w_{ji}^{in} of this edge is the number of molecules of species $\mathcal{S}_i$ resulting from the reaction $\mathcal{R}_j$, namely $w_{ji}^{\text{in}} = \beta_{ij}$.

Let us now re-index the n_s species and n_r reactions as nodes of the SR-graph using the following substitution: $i \mapsto i$, for $i \in \mathcal{S}$, and $j \mapsto j + n_s$, for $j \in \mathcal{R}$. In the new indexation, the species use the indices $1, \ldots, n_s$ and reactions the indices $n_s + 1, \ldots, n_s + n_r$.

The adjacency matrix $\mathbf{A}(\mathbf{s})$ of the SR-graph associated with a CRN is a $(n_s+n_r)\times(n_s+n_r)$ matrix, such that its entry A_{mk} is equal to the weight of the edge having the kth node as the source node and mth node as the target node. In the case when there is no such edge, we have $A_{mk} = 0$. Note that the definition of the adjacency matrix we use is the transpose of the adjacency matrix typically defined in the literature. To minimize confusion and avoid excessive notation, we will adopt our definition throughout, without any loss of generality. More precisely, one has:

$$A_{mk} = \begin{cases} w_{k,m-n_s}^{\text{out}}, & 1 \le k \le n_s \text{ and } n_s + 1 \le m \le n_s + n_r, \\ w_{k-n_s,m}^{\text{in}}, & n_s + 1 \le k \le n_s + n_r \text{ and } 1 \le m \le n_s, \\ 0, & \text{elsewhere.} \end{cases}$$

The weighted directed degree matrix $\mathbf{D}(\mathbf{s})$ of the SR-graph is a $(n_s + n_r)\times(n_s + n_r)$ diagonal matrix such that its mth diagonal entry is equal to the sum of the elements of the mth column of the weighted adjacency matrix $\mathbf{A}(\mathbf{s})$, namely:

$$D_{mk} = \delta_{mk}\sum_{p=1}^{n_s+n_r} A_{pk},$$

where δ_{mk} is the Kronecker delta.

The Laplacian matrix is an $(n_s + n_r) \times (n_s + n_r)$ square matrix defined as:

$$\mathbf{L}(\mathbf{s}) = \mathbf{D}(\mathbf{s}) - \mathbf{A}(\mathbf{s}). \tag{10}$$

The Laplacian matrix can be described as follows: for distinct nodes $m \neq k$, if there is an edge from the kth node to the mth node, L_{mk} is the negative of the weight of that edge. Otherwise, $L_{mk} = 0$. For the diagonal entries $m = k$, we set $L_{kk} = -\sum_{m \neq k} L_{mk}$, such that the column-wise sum of $\mathbf{L}(\mathbf{s})$ is the zero vector. Note that the main $n_s \times n_s$ block matrix of $\mathbf{L}(\mathbf{s})$ is a diagonal matrix. This is because in the SR-graph there are no edges connecting two species nodes. Similarly, the right-lower $n_r \times n_r$ block matrix is also diagonal because there are no edges connecting two reaction nodes. The other entries of the Laplacian matrix can be described as follows. If there is a reaction $\mathcal{R}_j$ for which the species $\mathcal{S}_i$ is a substrate, then the off-diagonal element $L_{(n_s+j)i}$ is equal to the negative product of the rational term in the rate expression for the jth reaction $\mathcal{R}_j$ and the partial derivative of the corresponding monomial with respect to the concentration of the ith species $\mathcal{S}_i$, i.e. $L_{(n_s+j)i} = -g_j \partial_i \varphi_j$. If there is a reaction $\mathcal{R}_j$ for which the ith species $\mathcal{S}_i$ is a product, then the off diagonal element $L_{i(n_s+j)}$ is the number of molecules of the ith species $\mathcal{S}_i$ in the expression of the jth reaction $\mathcal{R}_j$, i.e. $L_{i(n_s+j)} = -\beta_{ij}$.

Any directed graph is defined by an incidence matrix, which represents the connections between its vertices and edges. In the case of a SR-graph, the $(n_s + n_r) \times n_e$ incidence matrix $\mathbf{B}$, with n_e being the number of edges, is defined as follows:

$$B_{ij} = \begin{cases} -1, & \text{if the } i\text{th node is the source of the } j\text{th edge,} \\ 1, & \text{if the } i\text{th node is the target of the } j\text{th edge,} \\ 0, & \text{otherwise.} \end{cases} \tag{11}$$

Define the $(n_s + n_r) \times n_e$ outgoing matrix $\mathbf{\Delta}$ of the SR-graph by:

$$\Delta_{ij} = \begin{cases} 0, & \text{if } B_{ij} = 1, \\ B_{ij}, & \text{otherwise.} \end{cases}$$

It can be shown that the Laplacian matrix is given in terms of matrix multiplication by:

$$\mathbf{L}(\mathbf{s}) = \mathbf{B}\mathbf{\Gamma}_w \mathbf{\Delta}^\top, \tag{12}$$

where $\mathbf{\Gamma_w} = \text{diag}(\mathbf{w})$ is the conductance matrix with $\mathbf{w}$ being the vector of weights.

Example 4. *Consider again the CRN example presented in Eq. (4) and its corresponding SR-graph visualized in Fig. 1. By assigning weights to the edges according to the prescribed rule, the resulting edge weights are as follows:*

$$\begin{aligned}
&w_{11}^{\text{out}} = g_1 s_2^2, && w_{21}^{\text{in}} = 1, && w_{31}^{\text{in}} = 2, \\
&w_{21}^{\text{out}} = 2g_1 s_1 s_2, && w_{22}^{\text{in}} = 2, && w_{43}^{\text{out}} = 4g_3 s_4^3, \\
&w_{13}^{\text{in}} = 3, && w_{32}^{\text{out}} = 3g_2 s_3^2 s_4, && w_{35}^{\text{in}} = 5. \\
&w_{14}^{\text{in}} = 1, && w_{42}^{\text{out}} = g_2 s_3^3,
\end{aligned}$$

The weighted adjacency matrix $\mathbf{A}(\mathbf{s})$ and the weighted degree matrix $\mathbf{D}(\mathbf{s})$, both of size 8×8, are as follows:

$$\mathbf{A}(\mathbf{s}) = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 5 \\ g_1 s_2^2 & 2g_1 s_1 s_2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 3g_2 s_3^2 s_4 & g_2 s_3^3 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 4g_3 s_4^3 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad \mathbf{D}(\mathbf{s}) = \begin{bmatrix} g_1 s_2^2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2g_1 s_1 s_2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 3g_2 s_3^2 s_4 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & g_2 s_3^3 + 4g_3 s_4^3 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 4 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 7 \end{bmatrix}$$

and the 8×8 weighted directed Laplacian matrix $\mathbf{L}(\mathbf{s})$ is:

$$\mathbf{L}(\mathbf{s}) = \begin{bmatrix} g_1 s_2^2 & 0 & 0 & 0 & 0 & 0 & -1 & -2 \\ 0 & 2g_1 s_1 s_2 & 0 & 0 & 0 & 0 & -2 & 0 \\ 0 & 0 & 3g_2 s_3^2 s_4 & 0 & 0 & -3 & 0 & 0 \\ 0 & 0 & 0 & g_2 s_3^3 + 4g_3 s_4^3 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -5 \\ -g_1 s_2^2 & -2g_1 s_1 s_2 & 0 & 0 & 0 & 4 & 0 & 0 \\ 0 & 0 & -3g_2 s_3^2 s_4 & -g_2 s_3^3 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & -4g_3 s_4^3 & 0 & 0 & 0 & 7 \end{bmatrix}.$$

Note that the Laplacian matrix can also be represented in the matrix multiplication form given in Eq. (12) with the 8×11 incidence matrix $\mathbf{B}$ and the 8×11 outgoing matrix $\mathbf{\Delta}$ given as follows:

$$\mathbf{B} = \begin{bmatrix} -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & -1 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & -1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & -1 \end{bmatrix}, \quad \mathbf{\Delta} = \begin{bmatrix} -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & -1 \end{bmatrix}.$$

3.2 Properties of Laplacian matrices of SR-graphs

Beyond the standard properties common to all Laplacian matrices, the Laplacians of SR-graphs constructed using our weight assignment scheme exhibit additional structural features. In particular, certain block structures within these matrices take on specific forms, owing to the bipartite nature of the SR-graphs and the characteristics of the assigned weights. We summarize these additional properties in the following. The $(n_s + n_r) \times (n_s + n_r)$ Laplacian matrix has the following block structure:

$$\mathbf{L}(\mathbf{s}) = \begin{bmatrix} \mathbf{L}_{11}(\mathbf{s}) & \mathbf{L}_{12} \\ \mathbf{L}_{21}(\mathbf{s}) & \mathbf{L}_{22} \end{bmatrix}, \tag{13}$$

where:

P1) $\mathbf{L}(\mathbf{s})$ possesses the standard properties of Laplacian matrices, that is, for every $\mathbf{s} \in \mathbb{R}^{n_s}_{\geq 0}$, $\mathbf{L}(\mathbf{s})$ has non-negative diagonal entries, non-positive off-diagonal entries, and zero column sums.
P2) $\mathbf{L}_{11}(\mathbf{s})$ is an $n_s \times n_s$ diagonal matrix with entries depending on the vector of species concentrations $\mathbf{s}$. The matrix is diagonal because, due to the bipartite structure of the SR-graph, there are no edges connecting species nodes to one another.

P3) $\mathbf{L}_{12}$ is an $n_s \times n_r$ constant matrix with integer entries. This is due to the fact that the weights of the edges from reaction nodes to species nodes in the SR-graph are constants. Specifically, $[\mathbf{L}_{12}]_{ij} = -\beta_{ij}$, where β_{ij} is the stoichiometric coefficient of the ith species in the product of the jth reaction.

P4) $\mathbf{L}_{22}$ is an $n_r \times n_r$ constant diagonal matrix. The matrix is diagonal because, again, the bipartite structure of the SR-graph prevents any edges between reaction nodes. Specifically $[\mathbf{L}_{22}]_{jj} = \sum_{i=1}^{n} \beta_{ij}$.

P5) $\mathbf{L}_{21}(\mathbf{s})$ is an $n_r \times n_s$ matrix with entries depending on $\mathbf{s}$. For each row j of $\mathbf{L}_{21}(\mathbf{s})$, there is a rational function $g_j(\mathbf{s})$ and a monomial $\varphi_j(\mathbf{s}) = s_1^{\alpha_{1j}} \dots s_{n_s}^{\alpha_{n_s,j}}$ such that

$$[\mathbf{L}_{21}(\mathbf{s})]_{ji} = -g_j(\mathbf{s})\partial_i \varphi_j = -\alpha_{ij} v_j(\mathbf{s})/s_i, \tag{14}$$

where $v_j(\mathbf{s}) = g_j(\mathbf{s})\varphi_j(\mathbf{s})$. This results from (5). In other words, each row of $\mathbf{L}_{21}(\mathbf{s})$ is the product of a rational integrating factor and the gradient of a monomial with respect to $\mathbf{s}$.

P6) In the Laplacian matrix $\mathbf{L}(\mathbf{s})$, there is no $j \in \{1, \dots, n_s + n_r\}$ such that the jth row and the jth column are entirely filled with zeros. This is because we consider CRNs with connected SR-graphs.

P7) Since we consider CRNs with non-autocatalytic reactions, the SR-graphs cannot contain both an edge from a species to a reaction and the reverse edge simultaneously. Consequently, the entries of the Laplacian matrix satisfy the following property: $L_{i,j+n_s} L_{j+n_s,i} = 0$, for all $i = 1, \dots, n_s$ and $j = 1, \dots, n_r$.

Remark 2. *Note that, a matrix is the Laplacian matrix of the SR-graph of a CRN if it satisfies properties P1–P5 listed above. Also note that a CRN with non-integer rational stoichiometric coefficients may be modified so as to satisfy condition* P3 *without altering its dynamics. Assuming that the jth reaction has non-integer rational coefficients, let q denote the least common multiple of the denominators of the coefficients α_{ij} and β_{ij} for $i = 1, \ldots, n_s$. The transformation $\alpha_{ij} \mapsto q\alpha_{ij}$, $\beta_{ij} \mapsto q\beta_{ij}$ and $v_j(\mathbf{s}) \mapsto \frac{v_j(\mathbf{s})}{q}$ produces an equivalent reaction (i.e., a reaction with the same dynamics) with integer stoichiometric coefficients (see [23, Lemma 1]). The properties P6 and P7 have been included because for our reduction, we only consider autocatalytic CRNs with connected SR-graphs.*

3.3 Laplacian-based dynamical model

Here, we demonstrate how to derive the dynamical system given in Eq. (2) and the corresponding enzymatic reaction rates from the SR-graphs by utilizing its Laplacian matrix. We first introduce the vector of *net stoichiometric imbalances* of reactions. It is a vector $\mathbf{d} \in \mathbb{R}^{n_r}$, where the net stoichiometric imbalance of the reaction $\mathcal{R}_i$, i.e., $d_i := \sum_{j=1}^{n_s}(\alpha_{ji} - \beta_{ji})$, represents the difference between the total number of molecules consumed and the total number produced in the reaction $\mathcal{R}_i$. Moreover, $d_i v_i$ represents the net rate of change of the total species concentration due to the reaction $\mathcal{R}_i$. Therefore, reactions with zero net stoichiometric imbalance conserve the total species concentration.

Theorem 1. *Let $\mathbf{s} = [s_i]_{i=1}^{n_s}$ and $\mathbf{v} = [v_i]_{i=1}^{n_r}$ denote the vector of species concentrations and the vector of reaction rates. Further let $\mathbf{\Gamma_d} = \mathrm{diag}(\mathbf{d})$, with $\mathbf{d} = [d_i]_{i=1}^{n_r}$ being the vector of net stoichiometric imbalances of the reactions. Then the following identity holds:*

$$\begin{bmatrix} \dfrac{d\mathbf{s}}{dt} \\ \mathbf{\Gamma_d}\mathbf{v} \end{bmatrix} = -\mathbf{L}(\mathbf{s}) \begin{bmatrix} \mathbf{s} \\ \mathbf{v} \end{bmatrix}. \tag{15}$$

Proof. Let $\mathbf{\Lambda}$ denote the right-hand side of Eq. (15). For $i = 1, \ldots, n_s$, the ith component of $\mathbf{\Lambda}$ can be rewritten as:

$$\begin{aligned}
[\mathbf{\Lambda}]_i &= -\left(\sum_{j=1}^{n_r} g_j \frac{\partial \varphi_j}{\partial s_i}\right) s_i + \sum_{j=1}^{n_r} \beta_{ij} v_j = -\left(\sum_{j=1}^{n_r} g_j \alpha_{ij} s_i^{\alpha_{ij}-1} \left[\prod_{\substack{p=1,\\ p\neq i}}^{n_s} s_p^{\alpha_{pj}}\right]\right) s_i \\
&+ \sum_{j=1}^{n_r} \beta_{ij} v_j = -\sum_{j=1}^{n_r} \alpha_{ij} v_j + \sum_{j=1}^{n_r} \beta_{ij} v_j = \sum_{j=1}^{n_r} (\beta_{ij} - \alpha_{ij}) v_j = \frac{ds_i}{dt}.
\end{aligned}$$

The last identity follows form Eq. (2). Concerning the remaining components of $\mathbf{\Lambda}$, for $j = 1, \ldots, n_r$, we have:

$$\begin{aligned}
[\mathbf{\Lambda}]_{j+n_s} &= \sum_{i=1}^{n_s} \left(g_j \frac{\partial \varphi_j}{\partial s_i}\right) s_i - \left(\sum_{i=1}^{n_s} \beta_{ij}\right) v_j \\
&= \sum_{i=1}^{n_s} \left(g_j \alpha_{ij} s_i^{\alpha_{ij}-1} \left[\prod_{\substack{p=1,\\ p\neq i}}^{n_s} s_p^{\alpha_{pj}}\right] s_i\right) - \left(\sum_{i=1}^{n_s} \beta_{ij}\right) v_j \\
&= \left(\sum_{i=1}^{n_s} \alpha_{ij}\right) v_j - \left(\sum_{i=1}^{n_s} \beta_{ij}\right) v_j = \left(\sum_{i=1}^{n_s} \alpha_{ij} - \sum_{i=1}^{n_s} \beta_{ij}\right) v_j = d_j v_j.
\end{aligned}$$

This concludes the proof. □

Given a model of the form presented in Eq. (15), we can derive the expressions for the reaction rates $\mathbf{v}$ as functions of the species concentrations $\mathbf{s}$, as well as the corresponding system of ODEs provided in Eq. (2), as detailed in the following remark.

Remark 3. *Consider the block decomposition of the Laplacian matrix $\mathbf{L}(\mathbf{s})$ given in EQ. (13). The key argument is that the last equations of Eq. (15), which correspond to $\mathbf{v}$, depend linearly on $\mathbf{v}$, allowing $\mathbf{v}$ to be determined using Gaussian elimination. Consequently, Eq. (15) implies that:*

$$\begin{aligned}
\frac{d\mathbf{s}}{dt} &= -\Big(\mathbf{L}_{11}(\mathbf{s}) - \mathbf{L}_{12}(\mathbf{L}_{22} + \mathbf{\Gamma_d})^{-1}\mathbf{L}_{21}(\mathbf{s})\Big)\mathbf{s}, \\
\mathbf{v}(\mathbf{s}) &= -\Big((\mathbf{L}_{22} + \mathbf{\Gamma_d})^{-1}\mathbf{L}_{21}(\mathbf{s})\Big)\mathbf{s}.
\end{aligned} \tag{16}$$

Therefore, given the expression on the right-hand side of Eq. (15), simple algebraic manipulations can be employed to derive both the reaction rate vector and the system of ODEs outlined in Eq. (2). One can recognize in the expression of $\frac{d\mathbf{s}}{dt}$ *the Schur complement of:*

$$\begin{bmatrix} \mathbf{L}_{11}(\mathbf{s}) & \mathbf{L}_{12} \\ \mathbf{L}_{21}(\mathbf{s}) & \mathbf{L}_{22} + \mathbf{\Gamma_d} \end{bmatrix}$$

corresponding to the block $\mathbf{L}_{22}+\mathbf{\Gamma_d}$*. Note that the matrix* $\mathbf{L}_{22}+\mathbf{\Gamma_d}$ *is diagonal, i.e* $\mathbf{L}_{22}+\mathbf{\Gamma_d} = \mathrm{diag}\left([\sum_{i=1}^{n_s} \alpha_{ij}]_{j=1}^{n_r}\right)$*. This implies that all reaction rates can be computed using equation (16), except for zero-order reactions, which involve no reactants. However, this has no impact if the rates of zero-order reactions are treated as input parameters, see Eq. (3).*

This novel Laplacian-based and bipartite graph representation of CRNs provides a structured framework for analyzing species-reaction interactions. By explicitly capturing the network's topology, it offers valuable insights into both its structural and dynamic properties. The Laplacian matrix governs the diffusion of information in complex networks, with its eigenvalues and eigenvectors determining how perturbations spread and at what rate [40, 41]. In the context of CRNs, this corresponds to the propagation of concentration changes through reaction-mediated interactions [42]. This perspective can inform strategies for assessing system stability, identifying key pathways, and optimizing network behavior, with applications in synthetic biology, systems biology, and biochemical engineering. While further theoretical developments are left for future work, we demonstrate the practical utility of this approach for model reduction.

4 Model reduction

A real-life CRN model is often complex, involving many variables and parameters that require substantial computational resources for analysis [43]. Model reduction addresses this by simplifying the model, retaining key dynamical features while reducing the number of variables and parameters for easier analysis.

We apply our novel modeling approach to reduce CRNs, using the SR-graph and its Laplacian matrix to simplify the system of ODEs. Our reduction method for CRNs governed by GEKL, draws inspiration from the Kron reduction technique for resistive electrical networks [29], and is supported by insights from related CRN studies [20–22].

4.1 Methodology Outline

The goal of our model reduction is to remove a specific set of species and reactions from the SR-graph, ensuring that the resulting directed graph remains a valid SR-graph. Removing a reaction node j is equivalent to enforcing that $d_j v_j = 0$. The application of these balance constraints lead to a reduced model. Because the constraints are applied formally, their (approximate) validity must be tested. This is done numerically using the symmetrized error integral introduced in Section 2.4. We also require that the reduced model retains a SR-graph structure. To ensure this, we verify symbolically that the Schur complement Laplacian satisfies the properties listed in Section 3.2, as outlined in Algorithm 2.

Consider a CRN with dynamics given by Eq. (15). Let I_s and I_r represent the set of species and reaction indices, respectively. Further, let J_s and J_r denote the sets of indices of species and reactions to be deleted from the SR-graph, with $J_s^c = I_s \setminus J_s$ and $J_r^c = I_r \setminus J_r$. Define the following partitions for the concentration and reaction rate vectors.

$$\mathbf{s} = \begin{bmatrix} \mathbf{s}_1 \\ \mathbf{s}_2 \end{bmatrix}, \quad \mathbf{v} = \begin{bmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \end{bmatrix}.$$

Here, $\mathbf{s}_1$ is the vector of length $|J_s^c|$ representing the concentrations of the species remaining in the model, and $\mathbf{s}_2$ is the vector of length $|J_s|$ representing the concentrations of the species being removed from the model. Similarly, $\mathbf{v}_1$ is the vector of length $|J_r^c|$ representing the rates of the reactions remaining in the model, and $\mathbf{v}_2$ is the vector of length $|J_r|$ representing the rates of the reactions being removed from the model. In this case, the model given in Eq. (15) can be rewritten as follows:

$$\begin{bmatrix} \frac{d\mathbf{s_1}}{dt} \\ \mathbf{\Gamma_{d_1} v_1} \\ \frac{d\mathbf{s}_2}{dt} \\ \mathbf{\Gamma_{d_2} v_2} \end{bmatrix} = - \begin{bmatrix} \mathbf{L}_{11}(\mathbf{s}) & \mathbf{L}_{12}(\mathbf{s}) \\ \mathbf{L}_{21}(\mathbf{s}) & \mathbf{L}_{22}(\mathbf{s}) \end{bmatrix} \begin{bmatrix} \mathbf{s_1} \\ \mathbf{v_1} \\ \mathbf{s_2} \\ \mathbf{v_2} \end{bmatrix},$$

where $\mathbf{d}_i$ is the vector of stoichiometric imbalances corresponding to the reactions with rates given by $\mathbf{v}_i$. Here, the blocks $\mathbf{L}_{11}(\mathbf{s})$, $\mathbf{L}_{12}(\mathbf{s})$, $\mathbf{L}_{21}(\mathbf{s})$, and $\mathbf{L}_{22}(\mathbf{s})$ of the Laplacian matrix $\mathbf{L}(\mathbf{s})$ are matrices of sizes $(|J_s^c|+|J_r^c|)\times(|J_s^c|+|J_r^c|)$, $(|J_s^c|+|J_r^c|)\times(|J_s|+|J_r|)$, $(|J_s|+|J_r|)\times(|J_s^c|+|J_r^c|)$, and $(|J_s|+|J_r|)\times(|J_s|+|J_r|)$, respectively.

Our model reduction procedure is based on the following heuristics. Deleting the species nodes J_s is equivalent to assuming that the concentration vector $\mathbf{s}_2$ satisfy the equation $\frac{d\mathbf{s}_2}{dt}=0$. We therefore assume that the net effect of all the reactions on the species that have been removed is balanced, i.e., that the removed species are in a quasi-steady state. Similarly, deleting reaction nodes J_r is equivalent to assuming that the eliminated reactions $\mathbf{v}_2$ satisfy the equation $\mathbf{\Gamma}_{\mathbf{d}_2}\mathbf{v}_2=0$. Our model reduction technique has similarities with both the QSS and the QE approaches, which are elaborated in the Supplementary Material.
The reduced model reads:

$$\begin{bmatrix}\frac{d\mathbf{s_1}}{dt}\\ \mathbf{\Gamma}_{\mathbf{d}_1}\mathbf{v_1}\end{bmatrix}=-\mathbf{L}_{11}(\mathbf{s})\begin{bmatrix}\mathbf{s}_1\\ \mathbf{v}_1\end{bmatrix}-\mathbf{L}_{12}(\mathbf{s})\begin{bmatrix}\mathbf{s}_2\\ \mathbf{v}_2\end{bmatrix},\qquad -\mathbf{L}_{21}(\mathbf{s})\begin{bmatrix}\mathbf{s}_1\\ \mathbf{v}_1\end{bmatrix}-\mathbf{L}_{22}(\mathbf{s})\begin{bmatrix}\mathbf{s}_2\\ \mathbf{v}_2\end{bmatrix}=\mathbf{0}.$$

From the second equation we find that:

$$\begin{bmatrix}\mathbf{s}_2\\ \mathbf{v}_2\end{bmatrix}=-\mathbf{L}_{22}^{-1}(\mathbf{s})\mathbf{L}_{21}(\mathbf{s})\begin{bmatrix}\mathbf{s}_1\\ \mathbf{v}_1\end{bmatrix}.$$

Substituting this result into the first equation yields:

$$\begin{bmatrix}\frac{d\mathbf{s_1}}{dt}\\ \mathbf{\Gamma}_{\mathbf{d}_1}\mathbf{v_1}\end{bmatrix}=-\Big(\mathbf{L}_{11}(\mathbf{s})-\mathbf{L}_{12}(\mathbf{s})\mathbf{L}_{22}^{-1}(\mathbf{s})\mathbf{L}_{21}(\mathbf{s})\Big)\begin{bmatrix}\mathbf{s}_1\\ \mathbf{v}_1\end{bmatrix}=-\widehat{\mathbf{L}}(\mathbf{s})\begin{bmatrix}\mathbf{s}_1\\ \mathbf{v}_1\end{bmatrix},\tag{17}$$

where $\widehat{\mathbf{L}}(\mathbf{s})=\mathbf{L}_{11}(\mathbf{s})-\mathbf{L}_{12}(\mathbf{s})\mathbf{L}_{22}^{-1}(\mathbf{s})\mathbf{L}_{21}(\mathbf{s})$. Note that $\widehat{\mathbf{L}}(\mathbf{s})$ is the Schur complement of $\mathbf{L}(\mathbf{s})$ with respect to the indices corresponding to $J_s\cup J_r$. The reduction procedure does not apply when the matrix $\mathbf{L}_{22}(\mathbf{s})$ is singular.

It is important to note that the reduced model remains independent of the order in which species and reactions are deleted. The matrix $\widehat{\mathbf{L}}(\mathbf{s})$ corresponds to a SR-graph where the species nodes correspond to the species with indices J_s^c and the reaction nodes correspond to the reactions with indices J_r^c. The theorem that follows establishes that $\widehat{\mathbf{L}}(\mathbf{s})$ adheres to the first property mentioned in Section 3.2, expected from a weighted Laplacian matrix within the context of a SR-graph.

Theorem 2. *Let $\mathbf{L}(\mathbf{s})$ be the Laplacian matrix of the SR-graph associated with the CRN. Let $I\subset\{1,\dots,n_s+n_r\}$ and $\widehat{\mathbf{L}}(\mathbf{s})$ be the Schur complement of $\mathbf{L}(\mathbf{s})$ with respect to the set of indices I. Then $\widehat{\mathbf{L}}(\mathbf{s})$ is again a Laplacian matrix, i.e., the following properties hold:*

1. *All non-diagonal entries of $\widehat{\mathbf{L}}(\mathbf{s})$ are non-positive for $s\in\mathbb{R}_{\geq 0}^{n_s}$.*
2. *All diagonal entries of $\widehat{\mathbf{L}}(\mathbf{s})$ are non-negative for $s\in\mathbb{R}_{\geq 0}^{n_s}$.*
3. *$\widehat{\mathbf{L}}(\mathbf{s})$ has zero column sums.*

Proof. The proof of the theorem directly follows from the proof provided in [44]. □

Remark 4. *Theorem 2 establishes that the Schur complement of the Laplacian matrix is again a Laplacian matrix of an ordinary graph, implying that it need not correspond to a bipartite SR-graph. In the reduction algorithm described in the next section, we therefore additionally verify that the Schur complement satisfies the structural properties listed in Section 3.2, which ensure that the reduced matrix corresponds to a valid connected SR-graph with non-autocatalytic reactions.*

Remark 5. *In Eq. (17), the Schur complement eliminates the reaction rates $\mathbf{v}_2$, but the remaining Laplacian blocks, and thus $\widehat{\mathbf{L}}(\mathbf{s})$, still depend on both $\mathbf{s}_1$ and $\mathbf{s}_2$. To reduce the model by eliminating $\mathbf{s}_2$, one must solve the quasi-stationarity equation:*

$$\mathbf{s}_2=-\Pi_{\mathbf{s}_2}\mathbf{L}_{22}^{-1}(\mathbf{s}_1,\mathbf{s}_2)\mathbf{L}_{21}(\mathbf{s}_1,\mathbf{s}_2)\begin{bmatrix}\mathbf{s}_1\\ \mathbf{v}_1\end{bmatrix},\tag{18}$$

where $\Pi_{\mathbf{s}_2}$ projects onto $\mathbf{s}_2$. Substituting this into $\widehat{\mathbf{L}}(\mathbf{s})$, the reaction rates $\mathbf{v}_1$ are obtained from Eq. (17) and expressed in terms of $\mathbf{s}_1$ via Gaussian elimination, as in Remark 3.

First note that, in some cases, Eq. (18) can be solved analytically to give an explicit expression for $\mathbf{s}_2$ in terms of $\mathbf{s}_1$ and $\mathbf{v}_1$. In such cases, we can obtain a reduced model in an analytical form. This corresponds to the cases where the vector of reaction rates $\mathbf{v}_1$ of the reduced model can be expressed as an analytical function of $\mathbf{s}_1$. However, in other cases where Eq. (18) cannot be analytically solved, it still can be solved numerically. In such cases, we

Algorithm 1 CandidateCombinations

Require: The Laplacian matrix $\mathbf{L}$ as described in Eq.(10) and an integer N.

Ensure: Collection $\mathcal{I}$ of pairs $\{J_s, J_r\}$, where J_s is a subset of species indices and J_r is a subset of reaction indices, such that their removal results in a meaningful reduced model.

```
Initialize the output collection I as an empty set.
Find all subsets C_i of the species nodes containing exactly N elements.
for Every C_i do
  Find all possible combinations D_j of neighboring reaction nodes in the SR-graph.
  for For every D_j do
    Compute the Schur complement L_s of L corresponding to the set of indices C_i ∪ D_j.
    if L_s satisfies the properties described in Section 3.2, then
      Add the pair {C_i, D_j} to the output collection I.
    end if
  end for
end for

return The output collection I.
```

cannot obtain analytical expressions for vector of the reaction rates $\mathbf{v}_1$ *of the reduced model in terms of* $\mathbf{s}_1$*. The reduced model will be a computational model which is capable of producing the output trajectories for given initial conditions of species concentrations. It is important to note that other model reduction methods involving species elimination, such as the QSS and QE approximations, face similar challenges.*

4.2 Reduction algorithm

In this section, we present the reduction algorithm, outlining the key steps and principles. We demonstrate how to identify the optimal subset of species and reactions for removal, ensuring that the resulting bipartite graph remains a valid SR-graph. A step-by-step reduction procedure applied to a simple CRN example is provided in the supplementary material.

Combination of candidate species and reactions for effective removal: We now outline the methodology for selecting candidate species and reactions for removal from the SR-graph via Kron reduction.

The core idea is to select a combination of species and reaction subsets such that the corresponding Schur complement of the Laplacian, associated with these subsets, remains a valid reduced Laplacian matrix – meaning that it satisfies the conditions stated in Section 3.2. Ideally, we would examine all possible combinations of species and reactions. However, this is computationally expensive. With n_s species and n_r reactions, the number of combinations would be $2^{n_s+n_r} - 2^{n_s} - 2^{n_r} + 1$, which is prohibitive for large values of n_s and n_r. Thus, we need a more feasible approach to identify candidate subsets for removal.

To implement this procedure, we begin by selecting an integer N, which specifies the number of species nodes in each candidate subset. We then enumerate all distinct subsets $\mathcal{C}_i$ of species nodes containing exactly N elements. For each subset $\mathcal{C}_i$, we identify all possible combinations $\mathcal{D}_j$ of neighboring reaction nodes. A subset $\mathcal{C}_i \cup \mathcal{D}_j$ is considered a candidate for removal if its elimination results in a Schur complement satisfying the criteria of a valid Laplacian matrix, as presented in the Algorithm 1. This process is repeated for increasing values of N, starting with $N = 1$, systematically identifying increasingly larger candidate subsets. Since the symmetrized error integral tends to increase with N, the procedure is terminated when the error exceeds a predefined threshold.

Main reduction algorithm: We now present the main reduction algorithm, which is also schematized in Algorithm 2. Consider a given enzymatic CRN described by its stoichiometric matrix $\mathbf{S} \in \mathbb{R}^{n_s \times n_r}$ and the vector of reaction rates $\mathbf{v}$ as a function of species concentrations. The first step is to determine the SR-graph, which can be derived from the stoichiometric matrix as follows: the species nodes correspond to the rows, and the reaction nodes correspond to the columns of $\mathbf{S}$. The edges are constructed based on the entries of S_{ij}, where i is the row index and j is the column index. If the entry S_{ij} is negative, there is a directed edge from the ith species node to the jth reaction node in the SR-graph. If S_{ij} is positive, there is a directed edge from the jth reaction node to the ith species node. The weights of these edges are determined from the stoichiometric matrix and the vector of reaction rates, as explained at the beginning of Section 3.1.

Starting with the original model, we first generate candidate subsets of species and reactions for removal, as outlined in Algorithm 1. Among these, we select the subset whose deletion minimizes the value of the symmetrized error integral given in Eq. (9). We then apply Kron reduction to obtain the reduced model. This procedure is repeated iteratively on the reduced model: in each iteration, candidate subsets are determined, ranked, and the

Algorithm 2 KronReduction

Require: A CRN given by its stoichiometry matrix and its vector $\mathbf{v}$ of reaction rates as a function of species concentrations; and the maximum admissible value ε of the symmetrized error integral given in Eq. (9).

Ensure: Final reduced model with corresponding system of ODEs and reaction rates; final value ε_{F} of the symmetrized error integral.

Construct the Laplacian matrix of the SR-graph of the original model.
Set the final symmetrized error $\varepsilon_F = 0$.

repeat
  Select the Laplacian matrix $\mathbf{L}$.
  Set $\mathbf{L}_{\mathrm{F}} = \mathbf{L}$.
  Set $N = 0$.
  Initialize the set of candidate subsets $\mathcal{I}$ for deletion as the empty set.

  while $\mathcal{I}$ is empty, **do**
    $N = N + 1$.
    Determine the set of candidate subsets $\mathcal{I}$ of species and reaction indices for N, as described in Algorithm 1.
  end while

  Rank all candidate subsets in $\mathcal{I}$ based on the symmetrized error integral values Θ (compared to the initial model).
  Select the candidate subset from $\mathcal{I}$ whose removal leads to the smallest value of the symmetrized error integral Θ.
  Determine the Laplacian matrix $\widehat{\mathbf{L}}$ of the reduced model.
  Set $\mathbf{L} = \widehat{\mathbf{L}}$.
  Update ε_F to the corresponding Θ of the new reduced model.
until ε_{F} exceeds the cutoff threshold ε

return The final reduced model, which is the model corresponding to the Laplacian matrix $\mathbf{L}_{\mathrm{F}}$.

subset yielding the smallest symmetrized error integral is removed. The iteration is stopped when the symmetrized error integral of the reduced model (compared to the initial model) obtained at the end of an iteration exceeds a certain cutoff threshold. The reduced model obtained at the end of the previous iteration is then considered the final reduced model.

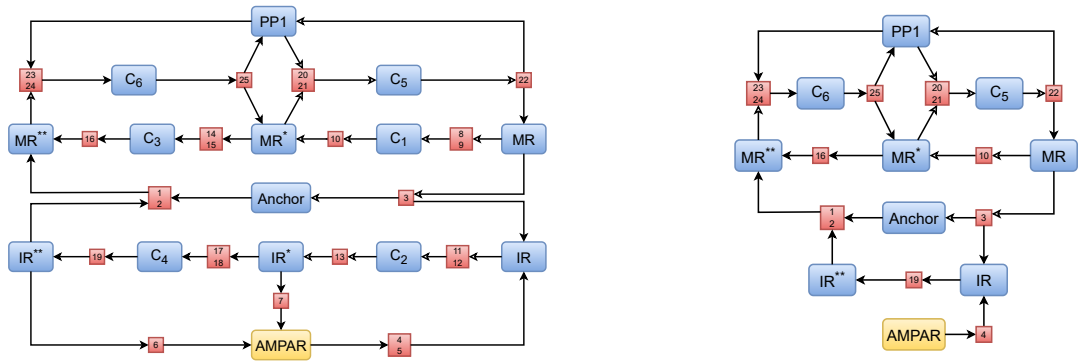
Automated Model Reduction Tools: We have developed Python and MATLAB packages for the automatic reduction of CRNs. These packages are provided as supplementary material and are publicly accessible via the following GitHub repository: https://github.com/manvelgasparyan/LK_red. The tools require the biochemical model information, including the stoichiometric matrix and the expressions of the rate functions with available parameter values. In addition, a user-defined parameter must be specified, which is the maximum admissible threshold for the symmetrized error integral (15% in our examples, as detailed in the next section).

After submission, the tool processes the model and produces several outputs, including a list of removed species and reactions, the symmetrized error integral – which quantifies the discrepancy between the original and reduced models – the SR-graphs of both the models, and a visual comparison of their base trajectories. This suite of features ensures efficient model reduction while preserving key dynamical properties within a pre-specified tolerance.

5 Application of model reduction to a real-life example

In this section, we demonstrate the practical utility of our modeling and model reduction methods through a real-life CRN example. These techniques simplify complex biological systems while preserving model accuracy. The case study showcases how our approach can make computational models more manageable without losing essential dynamics, validating the applicability of our methods for analyzing and optimizing biological systems. We examine the AMPAR trafficking cycle model, originally derived from [34]. In addition, three further real-world examples, namely a model of N-(1-deoxy-D-fructose-1-yl)-glycine (DFG) thermal decomposition [45], a model of Hedgehog Signaling Pathway (HSP) [46], and a model of Cell Division Cycle (CDC) [47], are provided in the Supplementary Material. These models are available from the BioModels database at https://www.ebi.ac.uk/biomodels.

The considered AMPAR trafficking cycle model forms a bistable system in which the low state has few (~ 10) membrane bound AMPA receptors (MR, MR* and MR** in the equations below), and the high state has large



(a) Original model.

(b) Kron-reduced model.

Fig. 2: Schematic visualization of the original and Kron-reduced model of AMPAR trafficking cycle model with bistability. The left-hand panel shows the original model, as developed in [34]. The right-hand panel displays the Kron-reduced model, which is obtained by removing the species nodes $\{C_1, C_2, C_3, C_4, IR^*\}$ and reaction nodes $\{\mathcal{R}_i : i = 5, \ldots 9, 11 \ldots, 15, 17, 18\}$ from the original graph. Nodes for non-buffered species are highlighted in blue, while the AMPAR node, as a buffered species, is in yellow, and reaction nodes appear in red.

Parameter	k_{1f}	k_{1b}	k_2	k_{3f}	k_{3b}	k_4	k_5
Value	$2 \cdot 10^{-4}$	$8 \cdot 10^{-3}$	$8 \cdot 10^{-4}$	$1.8 \cdot 10^{-2}$	1	$2 \cdot 10^{-5}$	$2 \cdot 10^{-5}$
Unit	$\#^{-1}s^{-1}$	s^{-1}	s^{-1}	s^{-1}	s^{-1}	s^{-1}	s^{-1}

Parameter	k_6	k_7	k_8	k_9	k_{10}	k_{11}	k_{12}	k_{13}	k_{14}
Value	0.14815	24	6	0.14815	24	6	0.14815	24	6
Unit	$\#^{-1}s^{-1}$	s^{-1}	s^{-1}	$\#^{-1}s^{-1}$	s^{-1}	s^{-1}	$\#^{-1}s^{-1}$	s^{-1}	s^{-1}

Parameter	k_{15}	k_{16}	k_{17}	k_{18}	k_{19}	k_{20}	k_{21}	k_{22}	k_{23}
Value	0.14815	24	6	0.29167	1.4	0.35	0.29167	1.4	0.35
Unit	$\#^{-1}s^{-1}$	s^{-1}	s^{-1}	$\#^{-1}s^{-1}$	s^{-1}	s^{-1}	$\#^{-1}s^{-1}$	s^{-1}	s^{-1}

Table 1: Parameter properties of the AMPAR trafficking cycle model.

numbers ($\sim$ 164 receptors). A brief increase in Protein Kinase A (PKA) promotes the model to the high state, representing synaptic potentiation, and a temporary reduction of PKA to near zero pulls the model from high to low state.

The original model comprises 2 buffered species and 14 non-buffered species. We decompose each reversible reaction into two unidirectional reactions, resulting in a total of 25 unidirectional reactions. A schematic representation of the CRN is shown in Fig. 2, while the complete list of original reactions is provided below.

$$\mathcal{R}_{1,2}: \ \text{Anchor} + \text{IR}^{**} \underset{k_{1b}}{\overset{k_{1f}}{\rightleftharpoons}} \text{MR}^{**},$$

$$\mathcal{R}_{3}: \ \text{MR} \xrightarrow{k_2} \text{Anchor} + \text{IR},$$

$$\mathcal{R}_{4,5}: \ \emptyset \underset{k_{3b}}{\overset{k_{3f}}{\rightleftharpoons}} \text{IR},$$

$$\mathcal{R}_{6}: \ \text{IR}^{**} \xrightarrow{k_4} \emptyset,$$

$$\mathcal{R}_{7}: \ \text{IR}^{*} \xrightarrow{k_5} \emptyset,$$

$$\mathcal{R}_{8,9,10}: \ \text{MR} \underset{k_7}{\overset{k_6}{\rightleftharpoons}} \text{C}_1 \xrightarrow{k_8} \text{MR}^{*}$$

$$\mathcal{R}_{11,12,13}: \ \text{IR} \underset{k_{10}}{\overset{k_9}{\rightleftharpoons}} \text{C}_2 \xrightarrow{k_{11}} \text{IR}^{*}$$

$$\mathcal{R}_{14,15,16}: \ \text{MR}^{*} \underset{k_{13}}{\overset{k_{12}}{\rightleftharpoons}} \text{C}_3 \xrightarrow{k_{14}} \text{MR}^{**}$$

$$\mathcal{R}_{17,18,19}: \ \text{IR}^{*} \underset{k_{16}}{\overset{k_{15}}{\rightleftharpoons}} \text{C}_4 \xrightarrow{k_{17}} \text{IR}^{**}$$

$$\mathcal{R}_{20,21,22}: \ \text{PP1} + \text{MR}^{*} \underset{k_{19}}{\overset{k_{18}}{\rightleftharpoons}} \text{C}_5 \xrightarrow{k_{20}} \text{MR} + \text{PP1}$$

$$\mathcal{R}_{23,24,25}: \ \text{PP1} + \text{MR}^{**} \underset{k_{22}}{\overset{k_{21}}{\rightleftharpoons}} \text{C}_6 \xrightarrow{k_{23}} \text{MR}^{*} + \text{PP1}$$

All reactions follow MAK, with reaction $\mathcal{R}_4$ being influenced by the modifier AMPAR, while reactions $\mathcal{R}_i$, $i = 8, 11, 14, 17$, are influenced by the modifier PKA. The model parameters and their respective values are listed in Table 1.

Through our iterative process, the optimal combination of species and reactions for removal is identified, which includes the subset of species $\{\mathrm{C}_1, \mathrm{C}_2, \mathrm{C}_3, \mathrm{C}_4, \mathrm{IR}^*\}$ and the subset of reactions $\{\mathcal{R}_i : i = 5, \ldots, 9, 11, \ldots, 15, 17, 18\}$. These subsets represent the largest group of nodes that can be removed while keeping the symmetrized error integral below 15%. The reduced model, therefore, consists of 9 species and 13 unidirectional reactions. The reduced CRN, along with its corresponding rates, is provided below.

$$\mathcal{R}_{1,2}: \ \mathrm{Anchor} + \mathrm{IR}^{**} \underset{k_{1b}}{\overset{k_{1f}}{\rightleftharpoons}} \mathrm{MR}^{**},$$

$$\mathcal{R}_3: \ \mathrm{MR} \xrightarrow{k_2} \mathrm{Anchor} + \mathrm{IR},$$

$$\mathcal{R}_4: \ \emptyset \xrightarrow{k_{3f}} \mathrm{IR},$$

$$\mathcal{R}_{10}: \ \mathrm{MR} \xrightarrow{p_1} \mathrm{MR}^*,$$

$$\mathcal{R}_{16}: \ \mathrm{MR}^* \xrightarrow{p_2} \mathrm{MR}^{**},$$

$$\mathcal{R}_{19}: \ \mathrm{IR} \xrightarrow{p_3} \mathrm{IR}^{**}$$

$$\mathcal{R}_{20,21,22}: \ \mathrm{PP1} + \mathrm{MR}^* \underset{k_{19}}{\overset{k_{18}}{\rightleftharpoons}} \mathrm{C}_5 \xrightarrow{k_{20}} \mathrm{MR} + \mathrm{PP1}$$

$$\mathcal{R}_{23,24,25}: \ \mathrm{PP1} + \mathrm{MR}^{**} \underset{k_{22}}{\overset{k_{21}}{\rightleftharpoons}} \mathrm{C}_6 \xrightarrow{k_{23}} \mathrm{MR}^* + \mathrm{PP1}.$$

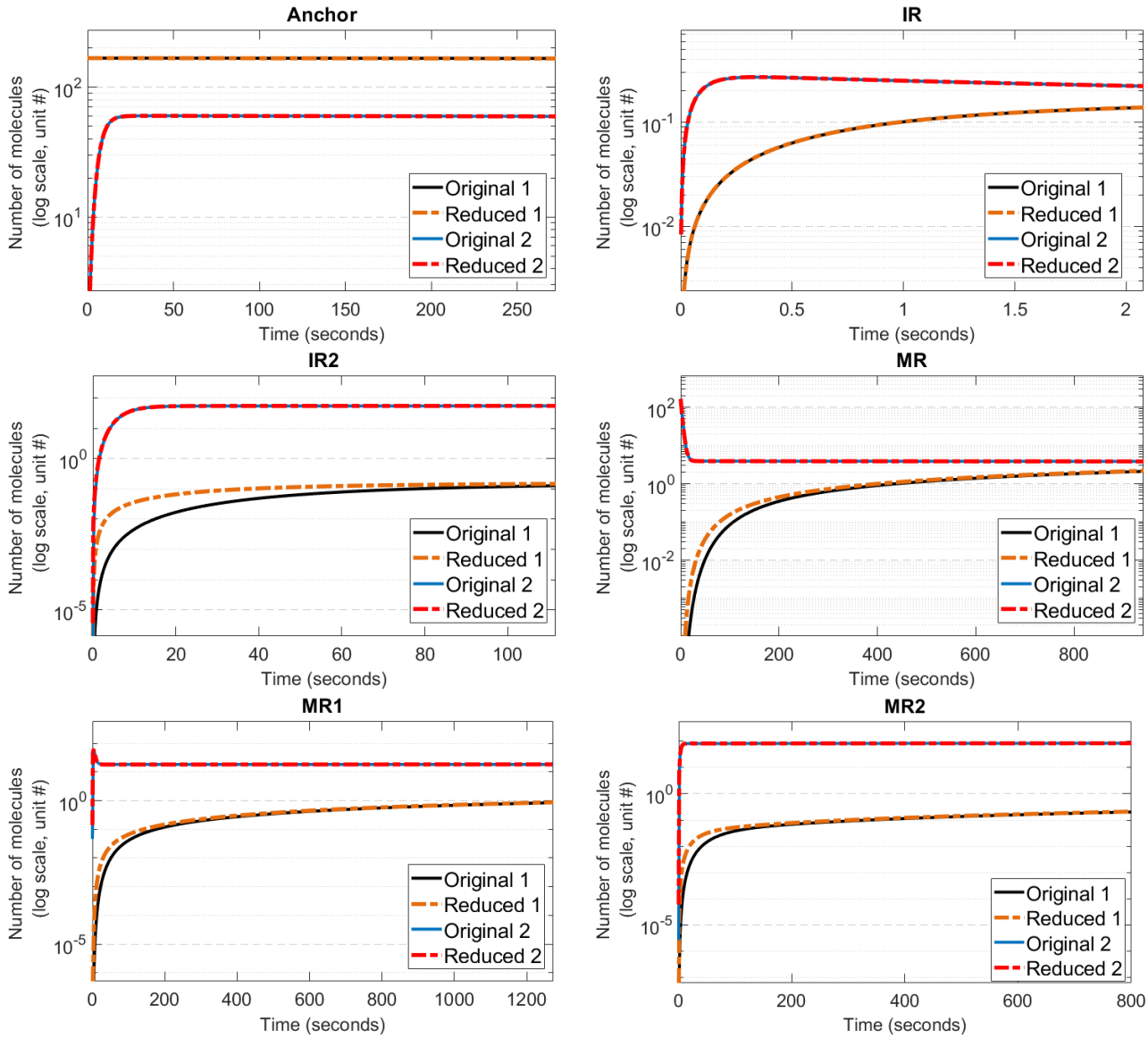


Fig. 3: A comparison of species molecule numbers between the original and reduced AMPAR trafficking cycle models for trajectories converging to different stable equilibrium points.

Parameter	p_1	p_2	p_3
Value	0.0296	0.0296	0.0296
Unit	s^{-1}	s^{-1}	s^{-1}

Table 2: Parameters introduced in the reduced model that are absent in the original.

	Species	Reactions	Deleted Species	Deleted Reactions	Error 1	Error 2
Original model	14	25	—	—	—	—
Reduced model	9	13	36%	48%	1.73%	0.34%

Table 3: Quantitative comparison of the original AMPAR trafficking cycle model and its reduced model.

All reactions in the reduced model, similar to the original model, follow MAK, with reactions $\mathcal{R}_{10}$, $\mathcal{R}_{16}$, and $\mathcal{R}_{19}$, influenced by the modifier PKA. The parameters p_1, p_2, and p_3 are algebraic expressions of the original model parameters. Due to their complexity and length, explicitly presenting them in terms of the original parameters is impractical. However, their properties are summarized in Table 2.

Fig. 3 presents a comparison of the evolution of molecular count of species between the reduced and original models for two distinct trajectories, both originating from the same stoichiometric compatibility class – that is, the set of all trajectories corresponding to the same conservation laws – but converging to different stable equilibrium points. Despite the removal of 35.71% of species and 48% of reactions, the discrepancy between the original and reduced models, as measured by the symmetrized error integral, remains relatively low. The error is 1.7334% for trajectories converging to one equilibrium state and 0.3377% for those reaching a different equilibrium state. Table 3 provides a quantitative comparison of the original AMPAR trafficking cycle model and the corresponding Kron-reduced model.

6 Conclusion and discussion

High throughput methods for defining protein interaction and reaction maps [48–52] bring the possibility of constructing detailed and complex reaction models for important biological signaling networks such as in the immune system and neurobiology. While such models are necessary to simulate the effect of specific mutations and pharmacological interventions, they are computationally demanding and are harder to parameterize. Simpler but functionally equivalent models are easier to analyze and to parameterize. In this context it is desirable to fluently traverse between complex and functionally similar simple models. Thus, model reduction techniques transforming complex models into simpler models are useful tools in systems biology. Despite the importance of the problem, pipelines for automated model reduction have yet to be developed.

For these reasons, we proposed an automated model reduction method that is both flexible and easy to implement. While it shares some similarities with QSS and QE approximations [12, 14, 15] (as detailed in the supplementary material), there are important differences. QSS and QE methods determine which species and reactions to prune based on timescale and singular perturbation considerations. As discussed in [14, 15], these can be identified statically from the orders of magnitude of CRN parameters, without requiring simulations. However, the quality of the approximation is guaranteed only asymptotically, as timescale ratios (expressed as rational monomials of kinetic parameters) tend to zero.

In contrast, our reduction method simulates a large number of candidate reduced models and selects the best one. Its key advantage is its validity for finite timescale ratios. Another distinction lies in the pooling operations. QSS prunes fast species and replaces the reactions consuming and producing them with elementary modes, i.e., combinations of reactions that conserve the fast species and ensure direct exchanges between non-eliminated species [12]. The choice of these elementary modes is not necessarily unique. In contrast, our reduction method leads to a unique choice.

QE prunes fast reactions and pools species that are consumed or produced by them. The pools correspond to conservation laws of the subsystem of fast reactions [12, 15], and their choice may also be non-unique. In contrast, our reduction method prunes reactions whose effect on the total species concentration is negligible, without pooling species.

Note that our model reduction procedure involves simulations of both the original and reduced models in order to determine the optimal combination of nodes to be deleted and these simulations do pose a computational burden. However, this burden is insignificant compared with the burden of model analysis of the original model including stability/bifurcation analysis and parameter sensitivity analysis as described in the Introduction. Reducing the

size of the model drastically reduces the computational efforts required to analyze the model. Moreover our model reduction procedure provides biological insights on which species or reactions of the original model play a dominant role in the dynamics and which ones have negligible influence since the ones with negligible influence get deleted through our procedure.

In the classical formulation of CRN dynamics based on the GCs [20], the Laplacian matrix is constant only under MAK. For GEKLs, however, the Laplacian may also depend on the system state through the reaction rate functions. In the SR-graph formulation considered here, the Laplacian depends explicitly on the state because the edge weights are defined via the reaction rates.

The original Kron reduction of GCs [20] is particularly effective for CRNs whose complex graphs contain connected components with more than two complexes (nodes). It reduces the number of complexes while retaining the remaining complexes and without adding new ones. This is biologically meaningful because the reduced network retains only complexes that already exist in the original biochemical mechanism, so every state in the reduced model corresponds to a real molecular configuration. Moreover, since no new complexes are introduced, the reduction avoids artificial biochemical states and preserves the interpretability of the CRN. However, when the GC consists of components with only two complexes, the original Kron reduction does not provide a meaningful reduction, as it may remove the entire reaction containing the complex we intend to remove. If such components share common species with other reactions, the extended Kron reduction [21, 22] can be applied. This two-step procedure first eliminates certain species using conservation laws, which can connect previously disconnected components of the GC into larger ones with multiple complexes; the original Kron reduction can then be meaningfully applied. A drawback of this approach is that it introduces additional parameters corresponding to the conservation constants, so the reduced model is valid only on trajectories consistent with these relations. As illustrated in the supplementary material with an example (see Section S1.2), the proposed Kron reduction of the SR-graph can, in certain cases, be dynamically equivalent to the extended Kron reduction of [22] as described above. For the considered example, both approaches lead to dynamically equivalent reduced models, but with different structures. On the other hand, compared with the Kron reduction methods based on the elimination of complexes [20–22], Kron reduction using SR-graphs is more flexible and leads to greater reduction corresponding to the same threshold. This has been illustrated using two examples of biomodels in Section S3 of the Supplementary Material. Indeed, the complex-based method cannot eliminate individual species, only entire complexes, which can be highly restrictive. We would like to remark that while the original complex-based Kron reduction approaches [20, 22] do preserve the kinetics of the original model, the SR-graph based Kron reduction approach introduced in this manuscript, may not do so. While the reduced model is still a valid chemical reaction network in terms of its network topology, it may not preserve the kinetics of the original.

The reduction of CRNs by making use of the graph Laplacian is facilitated by the well-established analogy between CRNs and electrical networks, based on shared structural, dynamical, and thermodynamical properties [42, 53]. Formally, the graph Laplacian is the negative generator of a diffusion process on the SR-graph [54]. In light of the above analogy, this process can be interpreted as the diffusion of information across the SR-graph. The associated diffusion flux respects the direction of causality: from reactions to the species they produce, and from species to the reactions whose rates they influence. Kron reduction enables the elimination of network elements (species and reactions) while preserving the essential pathways of information flow.

In summary, the proposed method offers several advantages, including a straightforward algorithmic implementation and uniquely defined SR-graph rewriting operations determined by the set of nodes to be pruned. Moreover, compared to Laplacian methods based on complexes, it provides greater flexibility by allowing the elimination of individual species or reactions rather than entire complexes. However, the approach has certain drawbacks, notably the need for simulations to validate the results and the risk of combinatorial explosion when identifying the nodes to prune, which may pose computational challenges.

The core procedure of our model reduction method, as discussed earlier, involves the removal of selected species and reaction nodes from the SR-graph associated with the CRN. This reduction is carried out by computing Schur complements of the Laplacian matrix corresponding to the SR-graph. We emphasize this aspect to highlight that the main reduction mechanism is independent of any specific trajectory of the system. However, the selection of species and reaction nodes to be removed is guided by the symmetrized error integral, which does depend on a base trajectory of the CRN.

Looking ahead, it is important to develop a trajectory-independent measure to quantify the discrepancy between the original and reduced models. This would enable a more robust and principled selection of species and reaction subsets for removal. One possible alternative to trajectory-based error measures could be the use of species and/or reaction settling times. Another alternative is to use orders of magnitude of parameters and tropical geometry methods to identify the nodes of the SR-graph to be pruned in the reduction method, along lines similar to those developed in [14, 15]. The proposed reduction method is formulated as a combinatorial optimization problem, in which admissible subsets of nodes are selected for elimination so as to minimize the symmetrized error integral. This problem is NP-hard. To manage its computational complexity, we currently limit the number of nodes that can be eliminated. As future work, we plan to improve our approach by implementing branch-and-bound strategies

to efficiently explore the space of admissible node subsets. We also intend to use the reduction methods introduced in this paper to generate hierarchies of models with varying complexity, which are essential for hierarchical machine learning strategies [11].

Funding. MG, USB and OR acknowledge support for this work from the CEFIPRA grant 68T08-3.

Conflict of interest/Competing interests. The authors have no competing interests to declare that are relevant to the content of this article.

Ethics approval and consent to participate. Not applicable. This study did not involve human participants, animal subjects, or data requiring ethics approval.

Consent for publication. Not applicable. This article does not contain any data, images, or information that requires consent for publication.

Code availability. Python and MATLAB codes developed for this work are available at the following link: https://github.com/manvelgasparyan/model_reduction.

Author contribution. SR formulated the initial concept. MG was responsible for formulating the problem, developing the solution, implementing the algorithm, and designing the automation tools. OR and SR contributed to the development of the solution and the main algorithms. USB and OR provided the models and biological insights, secured funding, and supervised the project. MG wrote the manuscript, with USB, OR and SR providing critical input, revisions, and review throughout the writing process.

Acronyms

AMPAR Activity-dependent Movement of a Glutamate Receptor.

CDC Cell Division Cycle.
CRN Chemical Reaction Network.

DFG N-(1-deoxy-D-fructose-1-yl)-glycine.

GC Graph of Complexes.
GEKL General Enzyme Kinetics Law.

HSP Hedgehog Signaling Pathway.

MAK Mass-Action Kinetics.

ODE Ordinary Differential Equation.

PKA Protein Kinase A.

QE Quasi-Equilibrium.
QSS Quasi-Steady-State.

SR-graph Species-Reaction Graph.

References

[1] Weng, G., Bhalla, U.S., Iyengar, R.: Complexity in biological signaling systems. Science **284**(5411), 92–96 (1999) https://doi.org/10.1126/science.284.5411.92

[2] Shlomi, T., Cabili, M.N., Herrgård, M.J., Palsson, B.Ø., Ruppin, E.: Network-based prediction of human tissue-specific metabolism. Nature Biotechnology **26**(9), 1003–1010 (2008) https://doi.org/10.1038/nbt.1487

[3] Chen, W.W., Niepel, M., Sorger, P.K.: Classic and contemporary approaches to modeling biochemical reactions. Genes & Development **24**(17), 1861–1875 (2010) https://doi.org/10.1101/gad.1945410

[4] Delbrück, M.: Statistical fluctuations in autocatalytic reactions. The Journal of Chemical Physics **8**(1), 120–124 (1940) https://doi.org/10.1063/1.1750549

[5] Gillespie, D.T.: Exact stochastic simulation of coupled chemical reactions. The Journal of Physical Chemistry **81**(25), 2340–2361 (1977) https://doi.org/10.1021/j100540a008

[6] Singh, A., Hespanha, J.P.: Stochastic hybrid systems for studying biochemical processes. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences **368**(1930), 4995–5011 (2010) https://doi.org/10.1098/rsta.2010.0211

[7] Crudu, A., Debussche, A., Radulescu, O.: Hybrid stochastic simplifications for multiscale gene networks. BMC Systems Biology **3**(1), 89 (2009) https://doi.org/10.1186/1752-0509-3-89

[8] Gutman, I.: Chemical graph theory—the mathematical connection. Advances in Quantum Chemistry **51**, 125–138 (2006) https://doi.org/10.1016/S0065-3276(06)51003-2

[9] Smits, J., Melssen, W., Buydens, L., Kateman, G.: Using artificial neural networks for solving chemical problems: Part i. multi-layer feed-forward networks. Chemometrics and Intelligent Laboratory Systems **22**(2), 165–189 (1994) https://doi.org/10.1016/0169-7439(93)E0035-3

[10] Holme, P., Saramäki, J.: Temporal networks. Physics Reports **519**(3), 97–125 (2012) https://doi.org/10.1016/j.physrep.2012.03.001

[11] Viswan, N.A., Tribut, A., Gasparyan, M., Radulescu, O., Bhalla, U.S.: Mathematical basis and toolchain for hierarchical optimization of biochemical networks. PLOS Computational Biology **20**(12), 1012624 (2024) https://doi.org/10.1371/journal.pcbi.1012624

[12] Radulescu, O., Gorban, A.N., Zinovyev, A., Noel, V.: Reduction of dynamical biochemical reactions networks in computational biology. Frontiers in Genetics **3**, 131 (2012) https://doi.org/10.3389/fgene.2012.00131

[13] Snowden, T.J., Graaf, P.H., Tindall, M.J.: Methods of model reduction for large-scale biological systems: a survey of current methods and trends. Bulletin of Mathematical Biology **79**(7), 1449–1486 (2017) https://doi.org/10.1007/s11538-017-0277-2

[14] Kruff, N., Lüders, C., Radulescu, O., Sturm, T., Walcher, S.: Algorithmic reduction of biological networks with multiple time scales. Mathematics in Computer Science **15**(3), 499–534 (2021) https://doi.org/10.1007/s11786-021-00515-2

[15] Desoeuvres, A., Iosif, A., Lüders, C., Radulescu, O., Rahkooy, H., Seiß, M., Sturm, T.: Reduction of chemical reaction networks with approximate conservation laws. SIAM Journal on Applied Dynamical Systems **23**(1), 256–296 (2024) https://doi.org/10.1137/22M1543963

[16] Borisov, N.M., Markevich, N.I., Hoek, J.B., Kholodenko, B.N.: Signaling through receptors and scaffolds: independent interactions reduce combinatorial complexity. Biophysical journal **89**(2), 951–966 (2005) https://doi.org/10.1529/biophysj.105.060533

[17] Conzelmann, H., Saez-Rodriguez, J., Sauter, T., Kholodenko, B.N., Gilles, E.D.: A domain-oriented approach to the reduction of combinatorial complexity in signal transduction networks. BMC Bioinformatics **7**(1), 1–15 (2006) https://doi.org/10.1186/1471-2105-7-34

[18] Feret, J., Danos, V., Krivine, J., Harmer, R., Fontana, W.: Internal coarse-graining of molecular systems. Proceedings of the National Academy of Sciences **106**(16), 6453–6458 (2009) https://doi.org/10.1073/pnas.0809908106

[19] Demin, A., Demitraki, E., Pogudin, G.: Exact hierarchical reductions of dynamical models via linear transformations. Communications in Nonlinear Science and Numerical Simulation **131**, 107816 (2024) https://doi.org/10.1016/j.cnsns.2024.107816

[20] Rao, S., Schaft, A., Eunen, K., Bakker, B.M., Jayawardhana, B.: A model reduction method for biochemical reaction networks. BMC Systems Biology **8**(52), 1–17 (2014) https://doi.org/10.1186/1752-0509-8-52

[21] Gasparyan, M., Van Messem, A., Rao, S.: A novel technique for model reduction of biochemical reaction networks. IFAC-PapersOnLine **51**(19), 28–31 (2018) https://doi.org/10.1016/j.ifacol.2018.09.024

[22] Gasparyan, M., Van Messem, A., Rao, S.: An automated model reduction method for biochemical reaction networks. Symmetry **12**(8), 1321 (2020) https://doi.org/10.3390/sym12081321

[23] Gasparyan, M.: Model reduction and parameter estimation for kinetic models of biochemical reaction networks. PhD thesis, Ghent University (2023). https://biblio.ugent.be/publication/

01GPB70TQ5CYWDB72WABMM7QTG

[24] Jayawardhana, B., Rao, S., Sikkema, W., Bakker, B.M.: Handling biological complexity using Kron reduction. In: Mathematical Control Theory I vol. 461, pp. 73–93. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-20988-3_5

[25] Schaft, A., Rao, S., Jayawardhana, B.: On the mathematical structure of balanced chemical reaction networks governed by mass action kinetics. SIAM Journal on Applied Mathematics **73**(2), 953–973 (2013) https://doi.org/10.1137/11085431X

[26] Rao, S., Schaft, A., Eunen, K., Bakker, B.M., Jayawardhana, B.: Model-order reduction of biochemical reaction networks. In: 2013 European Control Conference (ECC), pp. 4502–4507 (2013). https://doi.org/10.23919/ECC.2013.6669587 . IEEE

[27] Rao, S., Schaft, A., Jayawardhana, B.: A graph-theoretical approach for the analysis and model reduction of complex-balanced chemical reaction networks. Journal of Mathematical Chemistry **51**, 2401–2422 (2013) https://doi.org/10.1007/s10910-013-0218-8

[28] Schaft, A., Rao, S., Jayawardhana, B.: A network dynamics approach to chemical reaction networks. International Journal of Control **89**(4), 731–745 (2016) https://doi.org/10.1080/00207179.2015.1095353

[29] Kron, G.: Tensor Analysis of Networks. New York, Wiley & Sons (1939)

[30] Gasparyan, M., Rao, S.: Parameter estimation for kinetic models of chemical reaction networks from partial experimental data of species' concentrations. Bioengineering **10**(9), 1056 (2023) https://doi.org/10.3390/bioengineering10091056

[31] Angeli, D., De Leenheer, P., Sontag, E.: Graph-theoretic characterizations of monotonicity of chemical networks in reaction coordinates. Journal of Mathematical Biology **61**, 581–616 (2010) https://doi.org/10.1007/s00285-009-0309-0

[32] Feinberg, M.: Foundations of Chemical Reaction Network Theory. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-03858-8

[33] Gasparyan, M., Radulescu, O., Rao, S.: Kron reduction in species-reaction graphs of mass-action chemical reaction networks. In: 2025 13th International Conference on Bioinformatics and Computational Biology (ICBCB), pp. 128–133 (2025). https://doi.org/10.1109/ICBCB64873.2025.11198087 . IEEE

[34] Hayer, A., Bhalla, U.S.: Molecular switches at the synapse emerge from receptor and kinase traffic. PLoS Computational Biology **1**(2), 20 (2005) https://doi.org/10.1371/journal.pcbi.0010020

[35] Michaelis, L., Menten, M.L.: Die kinetik der invertinwirkung. Biochem. z **49**(333-369), 352 (1913)

[36] Hill, A.V.: The possible effects of the aggregation of the molecules of haemoglobin on its dissociation curves. The Journal of Physiology **40**, 4–7 (1910)

[37] Gasparyan, M., Van Messem, A., Rao, S.: Parameter estimation for models of chemical reaction networks from experimental data of reaction rates. International Journal of Control **96**(2), 392–407 (2023) https://doi.org/10.1080/00207179.2021.1998636

[38] Frøysa, H.G., Fallahi, S., Blaser, N.: Evaluating model reduction under parameter uncertainty. BMC Systems Biology **12**(79), 1–10 (2018) https://doi.org/10.1186/s12918-018-0602-x

[39] Zhang, F.: The Schur Complement and Its Applications vol. 4. Springer, New York (2006). https://doi.org/10.1007/b105056

[40] Masuda, N., Porter, M.A., Lambiotte, R.: Random walks and diffusion on networks. Physics Reports **716**, 1–58 (2017) https://doi.org/10.1016/j.physrep.2017.07.007

[41] Newman, M.: Networks. Oxford University Press, Oxford (2018)

[42] Radulescu, O., Lagarrigue, S., Siegel, A., Veber, P., Le Borgne, M.: Topology and static response of interaction networks in molecular biology. Journal of The Royal Society Interface **3**(6), 185–196 (2006) https://doi.org/

10.1098/rsif.2005.0092

[43] Ross, J., Arkin, A.P.: Complex systems: from chemistry to systems biology. Proceedings of the National Academy of Sciences **106**(16), 6433–6434 (2009) https://doi.org/10.1073/pnas.0903406106

[44] Schaft, A.: Characterization and partial synthesis of the behavior of resistive circuits at their terminals. Systems & Control Letters **59**(7), 423–428 (2010) https://doi.org/10.1016/j.sysconle.2010.05.005

[45] Martins, S.I.F.S., Van Boekel, M.A.J.S.: Kinetic modelling of Amadori N-(1-deoxy-D-fructos-1-yl)-glycine degradation pathways. part II—kinetic analysis. Carbohydrate research **338**(16), 1665–1678 (2003) https://doi.org/10.1016/S0008-6215(03)00174-5

[46] Sivakumar, K.C., Dhanesh, S.B., Shobana, S., James, J., Mundayoor, S.: A systems biology approach to model neural stem cell regulation by Notch, Shh, Wnt, and EGF signaling pathways. Omics: a journal of integrative biology **15**(10), 729–737 (2011) https://doi.org/10.1089/omi.2011.00

[47] Tyson, J.J.: Modeling the cell division cycle: cdc2 and cyclin interactions. Proceedings of the National Academy of Sciences **88**(16), 7328–7332 (1991) https://doi.org/10.1073/pnas.88.16.7328

[48] Büchel, F., Rodriguez, N., Swainston, N., Wrzodek, C., Czauderna, T., Keller, R., Mittag, F., Schubert, M., Glont, M., Golebiewski, M., Iersel, M., Keating, S., Rall, M., Wybrow, M., Hermjakob, H., Hucka, M., Kell, D.B., Müller, W., Mendes, P., Zell, A., Chaouiya, C., Saez-Rodriguez, J., Schreiber, F., Laibe, C., Dräger, A., Le Novère, N.: Path2Models: large-scale generation of computational models from biochemical pathway maps. BMC Systems Biology **7**(1), 116 (2013) https://doi.org/10.1186/1752-0509-7-116 . Accessed 2022-01-05

[49] Türei, D., Korcsmáros, T., Saez-Rodriguez, J.: OmniPath: guidelines and gateway for literature-curated signaling pathway resources. Nature methods **13**(12), 966–967 (2016) https://doi.org/10.1038/nmeth.4077

[50] Buffard, M., Naldi, A., Radulescu, O., Coopman, P.J., Larive, R.M., Freiss, G.: Network reconstruction and significant pathway extraction using phosphoproteomic data from cancer cells. Proteomics **19**(21-22), 1800450 (2019) https://doi.org/10.1002/pmic.201800450

[51] Gyori, B.M., Bachman, J.A.: From knowledge to models: Automated modeling in systems and synthetic biology. Current Opinion in Systems Biology **28**, 100362 (2021) https://doi.org/10.1016/j.coisb.2021.100362

[52] Dugourd, A., Kuppe, C., Sciacovelli, M., Gjerga, E., Gabor, A., Emdal, K.B., Vieira, V., Bekker-Jensen, D.B., Kranz, J., Bindels, E.M., *et al.*: Causal integration of multi-omics data with prior knowledge to generate mechanistic hypotheses. Molecular Systems Biology **17**(1), 9730 (2021) https://doi.org/10.15252/msb.20209730

[53] Oster, G.F., Perelson, A.S., Katchalsky, A.: Network thermodynamics: dynamic modelling of biophysical systems. Quarterly Reviews of Biophysics **6**(1), 1–134 (1973) https://doi.org/10.1017/S0033583500000081

[54] Lovász, L.: Random walks on graphs. Combinatorics, Paul erdos is eighty **2**(1-46), 4 (1993)