A combined model reduction algorithm for controlled biochemical systems

Article

Published Version

It is advisable to refer to the publisher's version if you intend to cite from the work. See Guidance on citing.
Published version at: http://dx.doi.org/10.1186/s12918-017-0397-1
To link to this article DOI: http://dx.doi.org/10.1186/s12918-017-0397-1

www.reading.ac.uk/centaur

CentAUR

Central Archive at the University of Reading

METHODOLOGY ARTICLE

Open Access

CrossMark

A combined model reduction algorithm for controlled biochemical systems

Thomas J. Snowden[1,2] , Piet H. van der Graaf[3,2] and Marcus J. Tindall[1,4*]

Abstract

Background: Systems Biology continues to produce increasingly large models of complex biochemical reaction networks. In applications requiring, for example, parameter estimation, the use of agent-based modelling approaches, or real-time simulation, this growing model complexity can present a significant hurdle. Often, however, not all portions of a model are of equal interest in a given setting. In such situations methods of model reduction offer one possible approach for addressing the issue of complexity by seeking to eliminate those portions of a pathway that can be shown to have the least effect upon the properties of interest.

Methods: In this paper a model reduction algorithm bringing together the complementary aspects of proper lumping and empirical balanced truncation is presented. Additional contributions include the development of a criterion for the selection of state-variable elimination via conservation analysis and use of an 'averaged' lumping inverse. This combined algorithm is highly automatable and of particular applicability in the context of 'controlled' biochemical networks.

Results: The algorithm is demonstrated here via application to two examples; an 11 dimensional model of bacterial chemotaxis in *Escherichia coli* and a 99 dimensional model of extracellular regulatory kinase activation (ERK) mediated via the epidermal growth factor (EGF) and nerve growth factor (NGF) receptor pathways. In the case of the chemotaxis model the algorithm was able to reduce the model to 2 state-variables producing a maximal relative error between the dynamics of the original and reduced models of only 2.8% whilst yielding a 26 fold speed up in simulation time. For the ERK activation model the algorithm was able to reduce the system to 7 state-variables, incurring a maximal relative error of 4.8%, and producing an approximately 10 fold speed up in the rate of simulation. Indices of controllability and observability are additionally developed and demonstrated throughout the paper. These provide insight into the relative importance of individual reactants in mediating a biochemical system's input-output response even for highly complex networks.

Conclusions: Through application, this paper demonstrates that combined model reduction methods can produce a significant simplification of complex Systems Biology models whilst retaining a high degree of predictive accuracy. In particular, it is shown that by combining the methods of proper lumping and empirical balanced truncation it is often possible to produce more accurate reductions than can be obtained by the use of either method in isolation.

Keywords: Model reduction, Lumping, Empirical balanced truncation, Controlled systems

*Correspondence: m.tindall@reading.ac.uk
[1] Department of Mathematics and Statistics, University of Reading, RG6 6AX, Reading, UK
[4] The Institute for Cardiovascular and Metabolic Research (ICMR), University of Reading, RG6 6AX Reading, UK
Full list of author information is available at the end of the article

Snowden *et al. BMC Systems Biology* (2017) 11:17

Page 2 of 18

Background

The field of Systems Biology has seen a considerable increase in both the number of models created and their complexity across the past decade. The BioModels Database, which acts as an open repository for Systems Biology models, saw the number of models it stores increase approximately ten-fold between 2005 and 2010, with the average number of reactions per model having nearly tripled in the same period [1]. This increase in complexity, specifically in the number of species or reactions modelled by each system, has become a defining characteristic of research in this area.

Such systems are typically developed by bringing together biochemical and physiological knowledge to inform highly detailed mechanistic models of biological networks (e.g. signalling pathways, protein-protein interactions, and genetic cascades). Mathematically, these networks are typically modelled via high-dimensional systems of stiff, nonlinear ordinary differential equations (ODEs).

This model complexity, however, can present a number of issues with regards to their use and analysis. For example parameter estimation techniques and real-time numerical simulation can both be difficult to perform for high dimensional or overparameterised systems. Even the basic intuitiveness of a system can potentially be obscured by its complexity. Additionally, complexity of this form is often associated with the 'curse of dimensionality', whereby the data that can be obtained for such systems in practice are sparse relative to the volume of the state and parameter spaces.

Model reduction techniques [2] offer one possible approach to easing complexity. A method of model reduction here refers to any method designed to construct a lower order representation (either in terms of the number of state variables or parameters) of a model with which some set of the original dynamical behaviour can be satisfactorily approximated.

A range of model reduction methods exist in the literature, many of which have previously been applied to models of biochemical reaction networks. The most commonly applied methods are based upon time-scale separation. These simplify a system by exploiting the wide ranges in reaction rates and equilibration speeds typical of biochemical reaction networks. These approaches include variants of the quasi-steady state approximation [3–9], variants of the rapid equilibrium approximation [10–15], the intrinsic low dimensional manifold method [16–20], and computational singular perturbation [21–24]. Beyond time-scale exploitation, sensitivity analysis can also be used to guide model reduction by identifying and eliminating those portions of a network least responsible for the dynamical behaviour of interest [25–29]. Optimisation based methods [30–35] seek to

evaluate a range of possible reduced models under a given metric of reduction accuracy before returning the best available option. Lumping methods, meanwhile, reduce a network by treating groups of state-variables as a single dynamical, 'lumped' variable [36–41]. Additionally, there exist a range of singular value decomposition (SVD) reduction methods based upon the matrix decomposition of the same name. These exploit the property that SVD can be used to give lower rank approximations of matrices. Such methods have seen limited application to biochemical systems, but a number of publications employing a particular variant known as balanced truncation can be found in the literature [42–44].

Each model reduction approach has advantages and disadvantages in the reduction of large scale models of biochemical reaction networks. There is no 'one size fits all' method of reduction; the best method available depends inextricably on the properties of the model and the aims of reduction.

Systems Biology models, such as those this paper seeks to reduce, usually possess a number of mathematical properties that can influence the suitability of specific reduction methodologies: most notably, such models are often stiff, nonlinear and of extremely high dimensionality (e.g. containing 10s or even 100s of state-variables and parameters). Stiffness, in this context, refers to a parameter dependent property of a system of differential equations whereby their numerical integration can require taking a step-length that is excessively short relative to the exact solution's smoothness in a given interval. This has relatively severe implications for simulating such models, as stiffness is associated with issues of numerical stability and computational processing time. In the case of large-scale Systems Biology models, stiffness is typically a consequence of reactions in the system evolving and equilibrating at greatly different timescales as compared with one another. Meanwhile, the nonlinearity of these models implies that a number of analytical methodologies will not be applicable. Often linearisation is used in this context, but in many such cases this incurs a prohibitively large error. Finally, the high dimensionality of such systems can also present issues of combinatorial complexity or an excessive computational burden for some methods of mathematical analysis.

This paper specifically seeks to address the topic of controlled biochemical reaction networks. Here a controlled network refers to any model for which the dynamics are influenced by the concentration of a particular reactant that can be considered as an input into the network, within which some given combinations of the reactants can be considered as outputs of the system. In the context of Systems Biology this includes, for example, models of receptor signalling pathways where the concentration of an extracellular ligand may be seen as an input controlling

Snowden *et al. BMC Systems Biology*   (2017) 11:17

Page 3 of 18

the pathway. The concentration of some subset of the intracellular signalling species may also be considered an output that is directly observed or inferred from some measure of the cellular response. The recently emerging field of Quantitative Systems Pharmacology [45, 46], which proposes to mechanistically model the effects of drug action from the genetic scale upwards is particularly amenable to such a formulation.

Here we develop a model reduction algorithm focused on maintaining the input-output relationship of a controlled biochemical reaction network. The algorithm combines several approaches including conservation analysis, proper lumping and empirical balanced truncation. For controlled systems with a specified output, empirical balanced truncation is designed to give a reduction that accurately maintains the input-output relationship. Unfortunately, due to the need to repeatedly simulate the system under a range of perturbed conditions, empirical balanced truncation can be highly sensitive to model stiffness. Hence, given the stiffness that is commonly associated with such systems, in the combined algorithm proper lumping is employed as a preconditioner to enable the application of empirical balanced truncation whilst retaining an accurate reduced model.

Our paper is structured as follows: we will first outline the general model reduction problem and then proceed to provide an overview of conservation analysis, empirical truncation, and proper lumping. A detailed account of how these methods can be brought together to obtain more accurate reductions than can be obtained via application of any single method in isolation will then be given. Finally, we demonstrate the algorithm via application to two examples: an 11 dimensional model of bacterial chemotactic signalling in *Escherichia coli* [47] and a 99 dimensional model of extracellular signal-regulated kinase (ERK) phosphorylation mediated via the epidermal growth factor (EGF) and nerve growth factor (NGF) receptor pathways [48]. Results are compared and a number of enhancements to the core methods are discussed.

Problem outline

The models we seek to address are comprised of systems of coupled, nonlinear ODEs. These are formulated as initial value problems and can be expressed by a control affine, state-space representation such that

$$\dot{\boldsymbol{x}}(t) = \boldsymbol{f}(\boldsymbol{x}(t)) + \sum_{i=1}^{l} \boldsymbol{g}_i(\boldsymbol{x}(t))u_i(t), \tag{1a}$$

$$\boldsymbol{y}(t) = \boldsymbol{h}(\boldsymbol{x}(t)), \tag{1b}$$

with initial conditions $\boldsymbol{x}(0) = \boldsymbol{x}_0$ and where the over-dot represents the time derivative, such that $\dot{\boldsymbol{x}} = \frac{d\boldsymbol{x}}{dt}$. Here the state variables $\boldsymbol{x}(t) \in \mathbb{R}^n$ typically represent the time-varying concentrations of the modelled species, $\boldsymbol{u}(t) \in \mathbb{R}^l$

(such that $u_i(t) \in \boldsymbol{u}(t)$) represent the input variables, and $\boldsymbol{y} \in \mathbb{R}^p$ represent the output variables. Here $\boldsymbol{f}(\boldsymbol{x}(t))$ is the set of functions describing the dynamical interaction between the individual reactants, each set of functions $\boldsymbol{g}_i(\boldsymbol{x}(t))$ describes the dynamical behaviour of the reactants interacting with each of the inputs, and $\boldsymbol{h}(\boldsymbol{x}(t))$ describes the combinations of the reactant concentrations corresponding to each of the outputs.

We seek a reduced model of the form

$$\dot{\tilde{\boldsymbol{x}}}(t) = \tilde{\boldsymbol{f}}(\tilde{\boldsymbol{x}}(t)) + \sum_{i=1}^{l} \tilde{\boldsymbol{g}}_i(\tilde{\boldsymbol{x}}(t))u_i(t), \tag{2a}$$

$$\bar{\boldsymbol{y}}(t) = \tilde{\boldsymbol{h}}(\tilde{\boldsymbol{x}}(t)), \tag{2b}$$

where $\tilde{\boldsymbol{x}} \in \mathbb{R}^{\tilde{n}}$ represents a reduced set of state variables (such that $\tilde{n} < n$) and for which, given a set of inputs $\boldsymbol{u}(t)$, the reduced set of outputs $\bar{\boldsymbol{y}}(t)$ represents an approximation of the original set $\boldsymbol{y}(t)$. Similarly to $\boldsymbol{f}$, $\boldsymbol{g}$, and $\boldsymbol{h}$ in the unreduced model, $\tilde{\boldsymbol{f}}(\tilde{\boldsymbol{x}}(t))$ and $\tilde{\boldsymbol{g}}_i(\tilde{\boldsymbol{x}}(t))$ are sets of functions describing the dynamical effects of interactions between the reduced state-variables and inputs. Meanwhile $\tilde{\boldsymbol{h}}(\tilde{\boldsymbol{x}}(t))$ approximately maps the reduced state-variables to the outputs.

The accuracy of a reduced model in capturing the dynamics of the original can be defined in a number of ways to suit the needs of the modeller. The most common approaches, however, are based upon the instantaneous error between the outputs of the two systems

$$\boldsymbol{\epsilon}(t) = \left| \boldsymbol{y}(t) - \bar{\boldsymbol{y}}(t) \right|. \tag{3}$$

The most common metrics are the L^2-norm $\|\boldsymbol{\epsilon}(t)\|_2 = \left(\int \boldsymbol{\epsilon}(t)^2 \, dt \right)^{1/2}$ or the ∞-norm $\|\boldsymbol{\epsilon}(t)\|_\infty = \sup \{\boldsymbol{\epsilon}(t)\}$. Throughout this paper we employ a form of maximal relative error E as the metric that we aim to minimise, such that

$$\epsilon_i \in \boldsymbol{\epsilon}: \ \epsilon_i(t) = \frac{\left\| y_i(t) - \bar{y}_i(t) \right\|}{y_i(t)} \ \text{ and } \ E = \|\boldsymbol{\epsilon}(t)\|_\infty. \tag{4}$$

Here, the relative error is selected such reduced models can be compared fairly for a range of different perturbation magnitudes applied both to the inputs and the initial condition of the state-variables.

Methods

Our combined model reduction algorithm is designed to bring together the methods of nondimensionalisation, conservation analysis, proper lumping and empirical balanced truncation. At its core the method employs proper lumping as a preconditioner (to reduce model stiffness) prior to the application of empirical balanced truncation. In this section we briefly review the variants of the methods employed before providing a more detailed account of the algorithm.

Snowden *et al. BMC Systems Biology* (2017) 11:17

Page 4 of 18

Nondimensionalisation

Nondimensionalisation refers to the process of rescaling the variables in a system such that the physical units (typically units of concentration and time) are removed from the model [49]. There are number of purposes for nondimensionalisation in the analysis of biochemical systems — primary amongst these is its use in accessing characteristic or intrinsic properties of the reaction network. Usually these are ratios of kinetic rate constants and conserved values that enable greater intuition into how the parameterisation of a model governs its behaviour.

This yields a nondimensionalised parameter set $\tilde{p}$ with entries representing ratios of the original parameters p. Often, nondimensionalisation can result in a reduction in the dimension of the new parameter set $\tilde{p}$ by finding ratios that are fixed to 1 irrespective of the original parameterisation. This does not, however, result in a reduction in the number of modelled reactions or reactants and hence does not reduce complexity as previously defined. Additionally, the dimensionless parameters may lose their innate biological meaning as the ratios they represent may not always hold particular biological significance.

There is usually a large possible number of combinations of the original parameters that could be used to yield these nondimensional ratios - for example, it is common to rescale time relative to a single kinetic rate parameter, amongst which a great number of choices may exist. Whilst there is no single method to determine a 'best' or 'optimal' nondimensionalisation, in the case where the system is fully parameterised and parameter values are fixed, selecting a nondimensionalised parameter set $\tilde{p}$ with entries all of a similar order will typically improve the numerical properties of the model for computational issues such as rounding error. To achieve this, the combined model reduction algorithm randomly samples 50 possible parameter combinations as nondimensionalsations and selects the one that minimises σ, where

$$\sigma = \log_{10}\left(\frac{\max(\tilde{p})}{\min(\tilde{p})}\right). \tag{5}$$

Conservation analysis

It is common in models of biochemical reaction networks for the total concentration of certain subsets of the species to remain constant at all times independent of the model's specific parameterisation [50]. Such subsets are commonly referred to as conserved moieties. Replacement of state-variables via the algebraic exploitation of conservation relations is a common first step in the analysis of biochemical reaction networks. Eliminating one of the conserved state-variables for each of the conservation relations can be used to yield a simplified realisation of the system.

In principle, all such conservation relations for a given biochemical reaction network can be determined by finding the linear dependencies in the associated stoichiometry matrix of the system. Mathematically this relies upon computing the left null space of the stoichiometry matrix for which a number of well-established methods exist. A review of a range of such techniques, including Gaussian elimination and singular value decomposition, can be found in Sauro and Ingalls [51]. For smaller scale systems such methods will usually find all available conservation relations, but for higher dimensional systems difficulties may occur. In particular, it is often necessary to select more stable computational methods to avoid missing conservation relations or finding false ones due to numerical error. A particularly stable method based upon the construction of a QR decomposition via Householder reflections has been developed by Vallbhajosyula et al. [52]. We therefore employ a form of this approach here, a more detailed mathematical treatment of which can be found in Section 1.2 of the Additional file 1.

This form of analysis leads to a simplified realisation of the system under which it is possible to obtain nonsingular Jacobians for given states of the model. As such it can be seen as a first step in model reduction schemes involving numerical methods.

Proper lumping

A lumping can potentially refer to any direct mapping $L : \mathbb{R}^n \to \mathbb{R}^{\tilde{n}}$ of the original state variables $x(t) \in \mathbb{R}^n$ to a reduced set $\tilde{x}(t) \in \mathbb{R}^{\tilde{n}}$ where $\tilde{n} < n$. Here we limit ourselves to linear lumping, such that we have a projection of the form
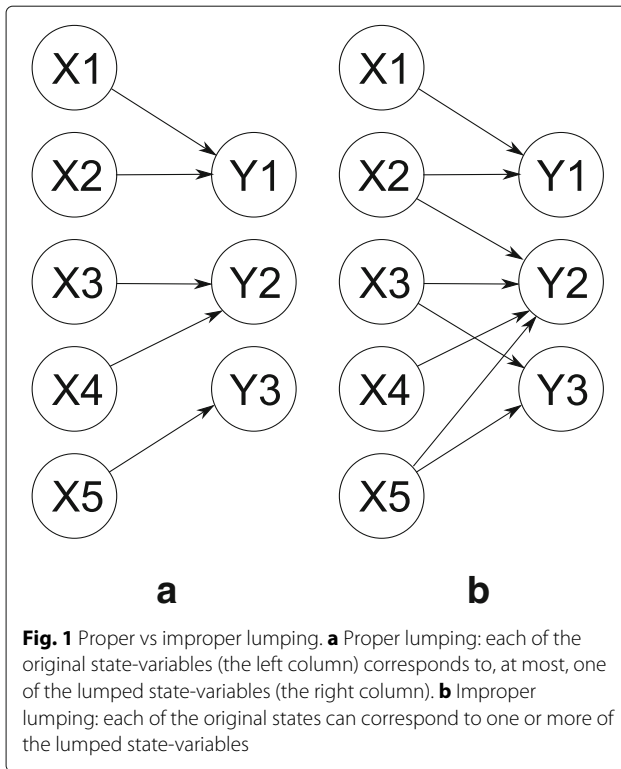
$$\tilde{x}(t) = Lx(t), \tag{6}$$

and, additionally, proper lumping such that the projection L becomes a matrix $L \in \{0, 1\}^{\tilde{n} \times n}$ where each column is pairwise orthogonal. This implies that each of the original state-variables corresponds to, at most, one of the lumped state-variables in the reduced model as is depicted in Fig. 1(a).

The dynamics of the reduced variables $\tilde{x}(t)$ are obtained via application of the Galerkin projection [53] (a detailed account of which can be found in the Section 1.1 of the Additional file 1) to the original system (2), such that

$$\dot{\tilde{x}}(t) = Lf(\bar{L}\tilde{x}(t)) + \sum_{i=1}^{l} Lg_i(\bar{L}\tilde{x}(t))u_i(t),$$

$$\text{with } \tilde{x}(0) = \tilde{x}_0 = Lx_0, \tag{7a}$$

$$\tilde{y}(t) = h(\bar{L}\tilde{x}(t)), \tag{7b}$$

where $\bar{L} \in \mathbb{R}^{n \times \tilde{n}}$ represents a generalised inverse of L such that $L\bar{L} = I_{\tilde{n}}$ (the $\tilde{n} \times \tilde{n}$ identity matrix). An approximation for the original state variables from the reduced variables can therefore be computed as

Snowden *et al. BMC Systems Biology* (2017) 11:17

Page 5 of 18

Fig. 1 Proper vs improper lumping. **a** Proper lumping: each of the original state-variables (the left column) corresponds to, at most, one of the lumped state-variables (the right column). **b** Improper lumping: each of the original states can correspond to one or more of the lumped state-variables

$$\boldsymbol{x}(t) \approx \bar{L}\tilde{\boldsymbol{x}}(t). \tag{8}$$

Finding an optimal lumping

There exist a number approaches for selecting an appropriate lumping matrix L for producing a system of given reduced dimensionality $\tilde{n}$. Here we choose to employ the scheme described by Dokoumetzidis and Aarons [37]. This algorithm runs an exhaustive search of possible lumping matrices to determine which produces the lowest error between simulation of the original model and the reduced model. To speed up this process, it is assumed (from justifications given in the original paper) that the lowest error k dimensional reduction obtained via lumping of an n dimensional system can also be found as the optimal lumping of two states in the $k + 1$ dimensional reduction. This yields a 'forward selection' strategy, where 2 of the state-variables are lumped at each step, which greatly decreases the combinatorial burden of possible lumping matrices that must be evaluated.

Lumping and stiffness

Stiffness here is defined as the ratio between the largest and smallest eigenvalues of the Jacobian matrix of the system evaluated at its unperturbed initial condition, such that

$$\chi = \frac{\lambda\left(J|_{\boldsymbol{x}=\boldsymbol{x}_0}\right)_{\max}}{\lambda\left(J|_{\boldsymbol{x}=\boldsymbol{x}_0}\right)_{\min}}, \quad \boldsymbol{x}_0 \in \mathbb{R}^n. \tag{9}$$

A high stiffness coefficient implies that parts of the system are evolving at significantly greater rates than others which can lead to traditional methods of simulation being numerically unstable.

The lumping algorithm of Dokoumetzidis and Aarons [37] will tend to sum together those state-variables that interact on faster timescales than their neighbours, and hence rapidly reach a form of proportional equilibrium. As a result the reduced model will tend to contain a lower range of timescales and a lower stiffness coefficient with every additional dimension eliminated. Furthermore, proper lumping, which creates reduced state-variables as straightforward sums of the originals, retains a degree of a biological interpretability that may not hold for alternative, coordinate transforming methods of model reduction.

Empirical balanced truncation

Our approach to empirical balanced truncation is based upon the procedure developed by Hahn and Edgar [54]. The aim is to construct two covariance matrices, known as the empirical controllability and observability Gramians, via repeated simulations of the system under perturbations of the input and the initial conditions respectively. The controllability Gramian provides information on how changes in the input will alter the state of the system. The observability Gramian provides information on the magnitude of the output any given initial condition of the system can produce. As these Gramians are positive semi-definite matrices, one good way to interpret them is as ellipsoids in the phase-space of the model [55]. The ellipsoid of the controllability Gramian represents the set of states that can be reached for a given magnitude of input, the ellipsoid of the observability Gramian describes the set of initial states that can produce an output of a given magnitude. A detailed mathematical account of the computation of empirical Gramians can be found in Section 1.3 of the Additional file 1.

We have modified the empirical balanced truncation procedure of Hahn and Edgar to deal with a system where state-variables have been eliminated via conservation analysis. In particular the conserved totals should be treated as functions of the initial conditions of the system and be altered in accordance with the perturbations used to create the observability Gramian. This allows us to perturb the initial conditions of the system without risking violation of conservation.

Once the Gramians have been computed the aim is to construct a balancing transformation of the state variables which also acts to equalise and diagonalise the Gramians. As the Gramians are now diagonal, each of the associated state-variables is therefore orthogonal in terms of their contribution to the input-output relationship. Hence, under this transformation, the state-variables

Snowden *et al. BMC Systems Biology*   (2017) 11:17

Page 6 of 18

which contribute least can be truncated without influencing the remaining terms.

The balancing transformation can be obtained in a numerically stable way following the approach of Laub et al. [56]. Firstly, perform a Cholesky factorisation of both the controllability Gramian $\mathcal{P}$ and the observability Gramian $\mathcal{Q}$ to obtain

$$\mathcal{P} = L^{\mathsf{T}}L \quad \text{and} \quad \mathcal{Q} = R^{\mathsf{T}}R,$$

where L and R represent the upper triangular factors of the Gramians. Now, take a singular value decomposition of the newly formed matrix LR^T and select a reduced dimensionality $\tilde{n}$ of the new model to obtain

$$LR^{\mathsf{T}} = (U_1 \; U_2) \begin{pmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{pmatrix} \begin{pmatrix} V_1^{\mathsf{T}} \\ V_2^{\mathsf{T}} \end{pmatrix},$$

where U_1 is an $n \times \tilde{n}$ matrix, Σ_1 is an $\tilde{n} \times \tilde{n}$ matrix (of the form $\Sigma_1 = \mathrm{diag}(\sigma_1^2, \ldots, \sigma_{\tilde{n}}^2)$) and V_1^{T} is a $\tilde{n} \times n$ matrix. Finally, set

$$T_1 = \Sigma_1^{-\frac{1}{2}} V_1^{\mathsf{T}} R \quad \text{and} \quad S_1 = L^{\mathsf{T}} U_1 \Sigma_1^{-\frac{1}{2}}. \tag{10}$$

Given the state-variable projection $\tilde{x} = T_1 x$ and the particular generalised inverse leading to the approximation $x \approx S_1 \tilde{x}$, we construct the reduced dynamics for this system again via application of the Galerkin projection to the original system (2), such that

$$\dot{\tilde{x}}(t) = T_1 f(S_1 \tilde{x}(t)) + \sum_{i=1}^{l} T_1 g_i(S_1 \tilde{x}(t)) u_i(t), \quad T_1 x(0) = \tilde{x}_0, \tag{11a}$$

$$\tilde{y}(t) = h(S_1 \tilde{x}(t)). \tag{11b}$$

This reduced system will not only feature fewer state-variables, but will also typically be faster to simulate and contain fewer parameters.

Stiffness and empirical balanced truncation

The construction of empirical Gramians via repeated simulation of the system under perturbations of input and initial conditions can be sensitive to numerical error. Where balanced truncation is applied using Gramians with a high degree of associated error, blow-up problems can occur for simulations of the reduced system. This accumulation of numerical error is particularly likely to occur for systems with a high-stiffness coefficient ($\chi \gg 1$).

Given the stiffness reducing property of lumping, however, and its retention of some biological meaning it can potentially be treated as a preconditioning step - enabling the application of more numerically sensitive methods

(such as empirical balanced truncation) to previously lumped systems.

The combined model reduction algorithm

Here a combined, automatable algorithm for model reduction bringing together the methods of nondimensionalisation, conservation analysis, proper lumping and empirical balanced truncation is introduced. A high-level overview of the algorithm's steps is shown in Fig. 2.
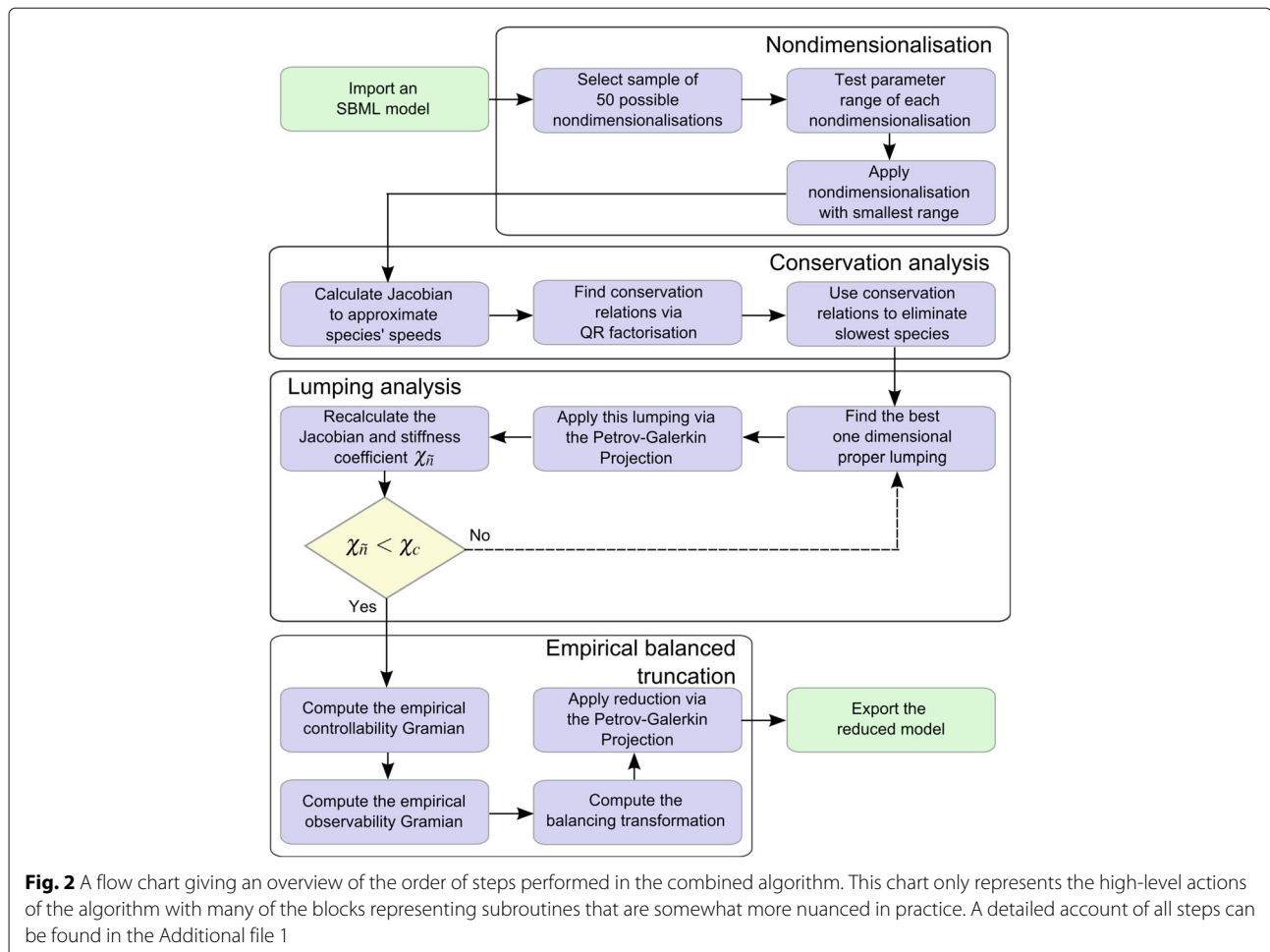
The algorithm is designed to be automatically applicable to models given in Systems Biology Markup Language (SBML) form — a standardised format for the representation, storage and easy communication of Systems Biology models. Many such models of this form can commonly be obtained from online repositories. Publicly open databases, such as the previously mentioned BioModels Database, contain thousands of such models enabling researchers to share their work in a more accessible way.

As preliminary steps, nondimensionalisation and conservation analysis are applied to the model. Nondimensionalisation is applied to improve numerical accuracy by reducing the range of parameters accounted for in the model. Conservation analysis is then applied in order to obtain a simplified realisation of the system and remove any associated singularities from the system's associated Jacobian.

At the core of the algorithm, however, are the methods of proper lumping and empirical balanced truncation. Theoretically, empirical balanced truncation should produce lower output error reductions across a range of inputs than lumping. However, due to the need to construct accurate covariance matrices of data from repeated numerical simulations, empirical balanced truncation can fail when applied to highly stiff systems. Conversely, lumping strips the model of some of its stiffness for each reduced dimension. In the combined algorithm, therefore, the complimentary aspects of proper lumping and empirical balanced truncation are exploited. Lumping is used to reduce the system until the stiffness coefficient $\chi_{\tilde{n}}$ of the reduced model is within a satisfactory range $\chi_{\tilde{n}} < \chi_c$, for χ_c some pre-chosen critical stiffness value (from numerical experimentation with example models we set this to be 250 in the automated algorithm). Empirical balanced truncation is then employed to obtain further model reduction whilst maintaining a good error bound between the outputs of the original and reduced models.

The algorithm as presented will proceed until the reduced system exceeds the maximum tolerated error, here defined to be 5%. The algorithm then returns the lowest dimensional reduced system that meets this constraint. It is not possible to know a priori what degree of reduction will be attainable by the algorithm, and the actual reduction achieved is both model structure and parameterisation dependent. Application of the combined

Snowden *et al. BMC Systems Biology* (2017) 11:17

Page 7 of 18

Fig. 2 A flow chart giving an overview of the order of steps performed in the combined algorithm. This chart only represents the high-level actions of the algorithm with many of the blocks representing subroutines that are somewhat more nuanced in practice. A detailed account of all steps can be found in the Additional file 1

reduction algorithm does require that the model it is applied to is fully parameterised prior to reduction. As such, the reduced model is only guaranteed to be accurate locally to a point in parameter space. For small perturbations to the initial parameterisation it is likely that the reduced model will remain accurate, however for larger deviations it may not be suitable for describing the overall input-output behaviour of the system.

Combined algorithm implementation
The main steps of the combined model reduction algorithm are as follows.

1. Import the model of interest with a set of predefined state-variables ($\boldsymbol{x}(t) \in \mathbb{R}^n$), reactions, and rate parameters into the algorithm.
2. Request the user to define the set of inputs ($\boldsymbol{u}$) and the set of outputs ($\boldsymbol{y} = h(\boldsymbol{x}(t))$) in which they are interested.
3. Randomly select a sample of 50 possible nondimensionalisations under the original

parameterisation of the model.
4. Calculate the range of the new parameter set for each nondimensionalisation.
5. Select and apply the nondimensionalisation sampled with the smallest range of orders of magnitude of parameters to help avoid rounding-off and truncation errors.
6. Compute the conservation relationships using a standard QR factorization method via Householder reflections (outlined, for example, in [51]).
7. Prioritise which state-variables to replace in the dynamical system via exploitation of the conservation relations.
8. Compute the linear, proper lumping matrix for all lumpable pairs of state-variables remaining in the system. Expressed alternatively, compute the set of all possible proper lumping matrices that will reduce the number of state-variables by one.
9. Compute the lumping inverse.
10. Simulate each lumped system and compute the associated maximal relative output error.

Snowden *et al. BMC Systems Biology*   (2017) 11:17

Page 8 of 18

11. Select the lumping matrix L and associated inverse that produces the lowest error and apply this reduction.
12. Calculate the stiffness coefficient $\chi_{\tilde{n}}$ for this reduced system.
13. Return to step (8). Exit the loop of steps (8)-(13) when the lumping either violates the maximum tolerated error (typically set to 5%) or reduces the stiffness coefficient to within an acceptable range $\chi_{\tilde{n}} < \chi_c$ or when the reduced dimensionality $\tilde{n} = 1$.
14. Compute the empirical Gramians for the lumped system.
15. Compute the balancing transformation for the given Gramians.
16. Apply the transformation to the model via the Galerkin projection and truncate state-variables until the maximum acceptable error is reached.
17. Produce the symbolic form of the reduced model and terminate.

The main algorithm was implemented in a commercial software package (Matlab 2013b, MathWorks Inc., Natick, MA). To enable the importing and manipulation of models stored in SBML, additional use was made of two open-source toolboxes - libSBML [57] and SBtoolbox2 [58]. Note that a more detailed account of the steps and implementation of the combined algorithm can be found in Section 3 of the Additional file 1.

Algorithm enhancements

The algorithm additionally features two specific enhancements to the combined methodologies that will be outlined in the following section. The first of these addresses the question of how to select which state-variables should be explicitly eliminated from the system of ODEs via the exploitation of conservation relations. The second addresses the question of which of the possible generalised inverses should be used to reduce the system via lumping with the Galerkin projection.

Selection of state-variables for elimination via conservation analysis: Typically, the most accurate possible proper lumping of two state-variables in a model will occur between those two that most rapidly equilibrate with respect to the remainder of the system. In short, if two variables quickly reach a point where their proportional concentration with respect to each other is approximately constant, they can be lumped accurately. If a state-variable from this 'best' lumped pair has been replaced by an algebraic equation from the conservation relations, then it will necessarily be unavailable for lumping in the combined algorithm. Another concern is that the choice of state-variables eliminated will have an effect upon the reduction of stiffness attainable via

the application of lumping; if the fastest interacting state-variables are not represented explicitly the stiffness will require many more steps of lumping in order to be reduced.

To avoid this issue the combined algorithm initially 'speed' ranks the state-variables in the model such that the slowest variables can then be selected for elimination. This speed ranking is achieved via numerical calculation of the Jacobian evaluated at the steady-state $x = x^*_{u_0}$ of the system attained under an unperturbed input $u = u_0$. The absolute value of the diagonal entries of the Jacobian $J|_{x=x^*_{u_0}}$ are then used as an approximate metric of the initial time-scale or speed of each of the state-variables. This is equivalent to the outgoing rate of concentration from each state-variable at the unperturbed steady-state of the network. Those state-variables corresponding to the largest values are deemed fast, whilst those corresponding to the smallest values are deemed slow and prioritised for elimination via application of conservation analysis.

Alternative lumping inverses: In order to apply a lumping L under the Galerkin projection, some generalised right-inverse $\bar{L}$ of L such that $L\bar{L} = I_{\tilde{n}}$ is required. Note that, as $\bar{L}$ could potentially be any generalised right-inverse of L, and there exists an infinite number of ways to construct such a matrix. However, the particular choice made will have an effect on the maximal relative output error incurred by the reduced model as defined by Eq. (4). Some of the alternatives are evaluated here.

In the original Wei and Kuo papers [59, 60] they suggest selecting the $\bar{L}$ that reconstructs the steady-state of the system such that $x(t) = \bar{L}\tilde{x}(t)$ for $t \to \infty$. This can be constructed as follows. Let

$$X := \text{diag}\left\{x^*\right\},$$

where x^* represents the steady-state of the system such that $\lim_{t \to +\infty} x(t) = x^*$, then

$$\bar{L} = XL^\mathsf{T}\left(LXL^\mathsf{T}\right)^{-1}. \tag{12}$$

As, in the case of a controlled biochemical network, the steady-state of the system is necessarily dependent on the value of the inputs $u(t)$. We here seek to maintain the generality of the reduced model by employing the steady-state attained under an unperturbed input $u = u_0$, such that $x = x^*_{u_0}$.

The Dokoumetzidis and Aarons [37] paper, which introduces the proper lumping algorithm that the scheme presented here is based upon, follows the work of Li and Rabbitz [61]. They suggested the somewhat simpler approach of using the Moore-Penrose inverse of L such that $\bar{L} = L^+$, which can similarly be computed as

$$L^+ = L^\mathsf{T}\left(LL^\mathsf{T}\right)^{-1}. \tag{13}$$

Snowden *et al. BMC Systems Biology* (2017) 11:17

Page 9 of 18

Typically the approach of reconstructing the steady-state values will lead to a better approximation of the original dynamics. In the case where a lumping sums a group of species all with a steady-state value of zero, however, this will lead to a singularity and thus the Moore-Penrose approach is to be preferred.

We thus consider an alternative to the above methodologies; we design an inverse $\hat{L}$ designed to reconstruct the proportionally average value of the states-variables for all time-points with non-zero values. To understand this alternative lumping inverse, first observe that all lumping matrices L (and inverses $\bar{L}$) can be expressed as the composition of sequential lumping matrices of two state-variables. For example

$$\begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

which demonstrates that a lumping of 3 state-variables can be achieved via two sequential lumpings of two state-variables. The novel inverse $\hat{L}$ is constructed for the two state-variable lumping case and can be generalised to any lumping via this process of sequential composition. Hence consider the general case seeking to lump the state-variables $x_h(t)$ and $x_k(t)$ (with $h < k$), where the corresponding lumping matrix $L = \{l_{ij}\}$ has the entries

$$l_{ij} = \begin{cases} 1 & \text{for } i = j \text{ with } i < k, \\ 1 & \text{for } i = j - 1 \text{ with } i \geq k, \\ 1 & \text{for } i = h, j = k, \\ 0 & \text{otherwise.} \end{cases} \tag{14}$$

The lumping inverse $\hat{L} = \left\{ \hat{l}_{ij} \right\}$ is then based upon calculating the average proportion of each of the lumped state-variables $x_h(t)$ and $x_k(t)$ during the interval $0 \leq t \leq T$; here, T is defined as some time point where the system can be assumed to have approximately reached steady-state, such that

$$\left\| \boldsymbol{x}(T) - \boldsymbol{x}^* \right\| \ll 1.$$

Hence each element of the inverse is constructed as follows

$$\hat{l}_{ij} = \begin{cases} 1 & \text{for } i = j \text{ with } i < h, \\ 1 & \text{for } i = j \text{ with } i > h \text{ and } j < k, \\ 1 & \text{for } i = j + 1 \text{ with } i \geq k, \\ \frac{1}{T} \int_0^T \frac{x_i(t)}{x_i(t) + x_j(t)} dt & \text{for } i = h, j = h, \\ \frac{1}{T} \int_0^T \frac{x_j(t)}{x_i(t) + x_j(t)} dt & \text{for } i = k, j = h, \\ 0 & \text{otherwise.} \end{cases} \tag{15}$$

In the case where the time $T \to \infty$ and either or both state-variables have a non-zero steady-state this reduces to the steady-state reconstructing lumping inverse defined

by Eq. (12). For the zero steady-state situation, however, it avoids the issue of numerical singularities.

Indices of controllability and observability

One of the benefits of the application of empirical balanced truncation is its potential use in obtaining metrics of observability and controllability for the individual state-variables of nonlinear systems. Given the typically nonlinear nature of cell signalling models, this potentially allows a framework for accessing indices of the controllability of each state-variable via receptor activation or suspension, the observability of each state-variable in influencing the output of interest, and the contribution of each state-variable to the overall input-output relationship of the model.

The interpretation of standard Gramians as ellipsoids describing the controllability and observability of the directions in phase-space [55] can also be extended to the interpretation of empirical Gramians. Hence it is possible to employ these matrices to obtain indices of controllability and observability for individual state-variables in nonlinear systems. Given this, we define the following indices.

Observability index: The observability index of the i-th state-variable $v_{oi} \in \boldsymbol{v}_o$ is defined as

$$v_{oi} := \frac{\sqrt{\boldsymbol{e}_i^\mathsf{T} Q \boldsymbol{e}_i}}{\max\left\{ \sqrt{\boldsymbol{e}_i^\mathsf{T} Q \boldsymbol{e}_i} \right\}},$$

where $0 \leq v_{oi} \leq 1$ and $\boldsymbol{e}_i$ represents the ith unit vector. Note the index has been rescaled relative to the maximally observable state-variable, such that each represents the comparable influence of the associated state-variable. The greater the value the more influence the state-variable has on the observed output. A value of zero indicates the state-variable has no effect on the output of interest.

Controllability index: The controllability index of the i-th state-variable $v_{ci} \in \boldsymbol{v}_c$ is defined as

$$v_{ci} := \frac{\sqrt{\boldsymbol{e}_i^\mathsf{T} \mathcal{P} \boldsymbol{e}_i}}{\max\left\{ \sqrt{\boldsymbol{e}_i^\mathsf{T} \mathcal{P} \boldsymbol{e}_i} \right\}},$$

with $0 \leq v_{ci} \leq 1$. Again, the indices here are rescaled relative to their maximal value such that they are more comparable to each other and to other indexes. The greater the value the more controllable the state-variable is via the input. A value of zero indicates that the input has no effect on the corresponding state-variable.

Input-output importance index: The input-output index $\boldsymbol{v}$ is then defined as the element-wise product of

Snowden *et al. BMC Systems Biology* (2017) 11:17

Page 10 of 18

the observability and controllability indices rescaled proportionate to the maximal value. Hence the index corresponding to the *i*-th state-variable $v_i \in \boldsymbol{v}$ is given by

$$v_i := \frac{v_{oi} v_{ci}}{\max\{v_{oi} v_{ci}\}},$$

with $0 \leq v_i \leq 1$. This therefore provides an overall metric of the corresponding state-variable's influence on the overall input-output relationship of the system.

Results and discussion

In this section two examples are employed to demonstrate the application of the combined model reduction algorithm, the enhancements made to the base methods, and the calculation of the indices of controllability and observability. The first of these systems is a relatively simple (11 dimensional) model of bacterial chemotaxis in *Escherichia coli* - the modest scope of this model allows the application of our methods to be more easily intuited. The second is a significantly more substantial model (99 dimensional) describing the mediation of ERK activation via both the EGF and NGF receptor pathways. Through application to a model of this size, our methods demonstrate their potential usefulness in analysing models that might be inscrutable under traditional approaches.

A model of bacterial chemotaxis

We have applied our combined model reduction methodology to a model of chemotactic signalling in *E. coli* as detailed in Tindall et al. [47] and summarised in Section 2 of the Additional file 1. This is a modest 11 dimensional model detailing a system of 12 biochemical reactions. This serves as a reasonable starting example as the model is intuitively tractable, but also of a sufficient size and complexity to be meaningfully reduced by the combined algorithm.

E. coli is a common, rod shaped, gram-negative bacterium often used as a model organism in biological studies due to both the large body of existing literature characterising its behaviour as well as the relative ease and inexpensiveness in its growth and experimental manipulation. There are many strains of *E. coli* present in nature, but the model discussed here pertains specifically to those strains that exhibit a chemotactic response. Chemotaxis is the process by which a cell senses an environmental chemical gradient and biases its movement towards those regions most suitable for growth and reproduction. In the model presented here, this process involves the transmembrane receptors on the surface of the bacterium sensing the local concentrations of an attractant or repellent; a decrease in attractant or an increase in repellent will cause the receptors to activate a signalling pathway inside the cell resulting in an increase of the intracellular concentration of the phosphorylated chemotactic CheY

protein, referred to here as CheY$_P$. This concentration, in turn, modulates the flagellum's movement, resulting in a change of direction for the cell either towards attractants or away from repellents.

Chemotaxis represents a good example model to work with as its attractant-receptor behaviour represents a controlled system, and it is highly amenable to the input-output problem formulation that the combined algorithm seeks to address. For our analysis, the external concentration of some chemotactic attractant was treated as the input into the system and the total concentration of CheY$_P$ (in both free and complex forms). Hence,

$$y(t) = [\text{CheY}_P] + [\text{CheA} \cdot \text{CheY}_P] + [\text{CheY}_P \cdot \text{CheZ}].$$

(16)

The model has a stiffness coefficient at the initial condition of the system of 958.3, which is relatively high. Values for these initial conditions can be found in Section 2 of the Additional file 1.

Reduction

We sought to compare the reduction of this example via lumping and empirical balanced truncation alone, and our combined algorithm.

In respect of the combined algorithm, the process of reduction began by nondimensionalising the variables of the system - specifically seeking to rescale the initial conditions and coefficients in the model such that they spanned the fewest orders of magnitude with the aim of avoiding possible issues of numerical truncation. The conservation relations for the system were then calculated; as 4 conservation relations were found, this lead to a system of 7 ODEs that exactly replicated the original system's dynamics.

The algorithm then lumped the state-variables until the stiffness coefficient was less than 250. In our example this occurred at 5 state-variables. Finally, empirical balanced truncation was applied to the 5 dimensional reduced model. In this case, the empirical Gramians were constructed using data from 100 distinct simulations covering perturbations to both the models input parameter $u(t)$ and the initial conditions $\boldsymbol{x}_0$. In both cases, perturbations were uniformly sampled from 0.2 to 1.8 times each parameter's original, unperturbed value. Using the balancing transformation and straightforward truncation the associated error for each possible dimensionality of reduction (between four and one) was calculated. A significantly more detailed account of reducing the bacterial chemotaxis model can be found in Section 2.1 of the Additional file 1.

Snowden *et al. BMC Systems Biology* (2017) 11:17

Page 11 of 18

The results in Table 1 along with Figs. 3 and 4 clearly demonstrate the combined algorithm produces more accurate reductions than those of either method in isolation. At the 2-dimensional reduction level, the combined algorithm exhibited an approximately 78% improvement in reduction error in comparison to the use of lumping alone. Additionally, this reduced model produced a significant speed-up in simulation time. In the case of 100 repeated simulations over a 3 second period under the introduction of $10\mu M$ concentration of attractant ligand at $t = 0$, the original 11 dimensional system required 0.2594 seconds on average to be simulated, whilst the 2 dimensional reduced model was solved in 0.0101 seconds.

Selection of species eliminated via conservation analysis
As was previously discussed, the selection of species eliminated via conservation analysis can have a substantial effect on model reduction error. The combined algorithm hence applies a speed-ranking step designed to estimate which of the biochemical reactants are likely to be most readily lumped.

To demonstrate the use of such an approach we have calculated the associated error and stiffness coefficients incurred under lumping for the chemotaxis model where the species eliminated via conservation were selected either naively or via the speed-ranking approach. These results can be found in Table 2. In the naive case the algorithm selected species CheA, $CheA_P \cdot CheY$, CheZ and CheB to be eliminated via conservation. Whilst in the selective case the algorithm selected species CheA, CheA·CheY, CheZ, and $CheB_P$. These results demonstrate that carefully considering which state-variables should be replaced via conservation relations can greatly improve the overall error incurred via lumping and yield a better reduction in model stiffness.

Table 1 Error results for the nonlinear methods of model reduction applied to the *E. coli* chemotaxis model

Number of dimensions	Empirical balanced truncation	Lumping	Combined method
6	1.396%	0.15%	-
5	1.655%	0.51%	-
4	#	0.54%	0.51%
3	#	4.77%	2.00%
2	#	12.88%	2.84%
1	#	75.56%	21.86%

Note: '#' implies Matlab could not numerically simulate this reduction using ode15s due to stiffness. '-' implies the reduction error at this point was equal to the lumping error. The errors stated represent the maximal relative error between the original and reduced systems when simulated to steady-state under the introduction of $10\mu M$ concentration of attractant ligand at $t = 0$

Analysis of alternative lumping inverses
The second enhancement to our combined method concerns the choice of lumping inverse used during model reduction. This topic does not seem to have been well explored in the literature, but as will be demonstrated here this choice can have a sizeable effect on the overall accuracy of a reduced model. In the case of the bacterial chemotaxis model, results comparing the three approaches for selecting a matrix inverse can be found in Table 3. In the case of the average inverse $\hat{L}$ we have set $T = 5$ seconds.

These results show that, whilst the Moore-Penrose inverse performs worse than the others, both the steady-state and averaged inverses can produce very good results. Given that the optimal approach can vary between lumping steps, the combined algorithm has been designed in such a way as to trial both the average and steady-state inverse at each step before selecting the most accurate. The overall inverse is then returned as the composition of the sequentially best inverses for each dimensionality of lumping.

Indices of controllability and observability
In the 5 dimensional lumped model created via the algorithm, only one lumped state-variable was created. This variable represents the sum of species Y_P, $A \cdot Y_P$, and $A_P \cdot Y_P$ which can be thought of as representing the phosphotransfer chain between species i.e. A and Y.

Calculating empirical Gramians for this lumped system yields the indices given in Table 4 (note that indices here can only be explicitly calculated for the retained species that were not eliminated via exploitation of conservation relations). Here the lumped variable, $Y_P + A \cdot Y_P + A_P \cdot Y_P$, is shown to be the most easily controlled, the most easily observed, and in total the most responsible for carrying the input-output signal from change in attractant concentration through to the phosphorylation of species Y and hence chemotaxis. Additionally, this set of species is readily lumpable, having such a low associated error cost and resulting in a large reduction in model stiffness. This suggests that the phosphotransfer process occurs significantly faster than the remainder of the network and hence equilibrates quickly. This finding concurs with the known biology, and demonstrates that $Y_P + A \cdot Y_P + A_P \cdot Y_P$ is more significant in the process of chemotactic signalling than the other species described in this model. Also noteworthy is the importance of species A·Y which agrees with other work (i.e. Tindall et al. [47]). Finally, the extremely low observability and controllability of CheB suggests that the overall concentration of it has an almost negligible response to the change of extracellular attractant and very little effect on the output in comparison to the remainder of the network. This also makes sense biologically, as CheB is functionally involved in the process of adaptation

Snowden *et al. BMC Systems Biology* (2017) 11:17

Page 12 of 18
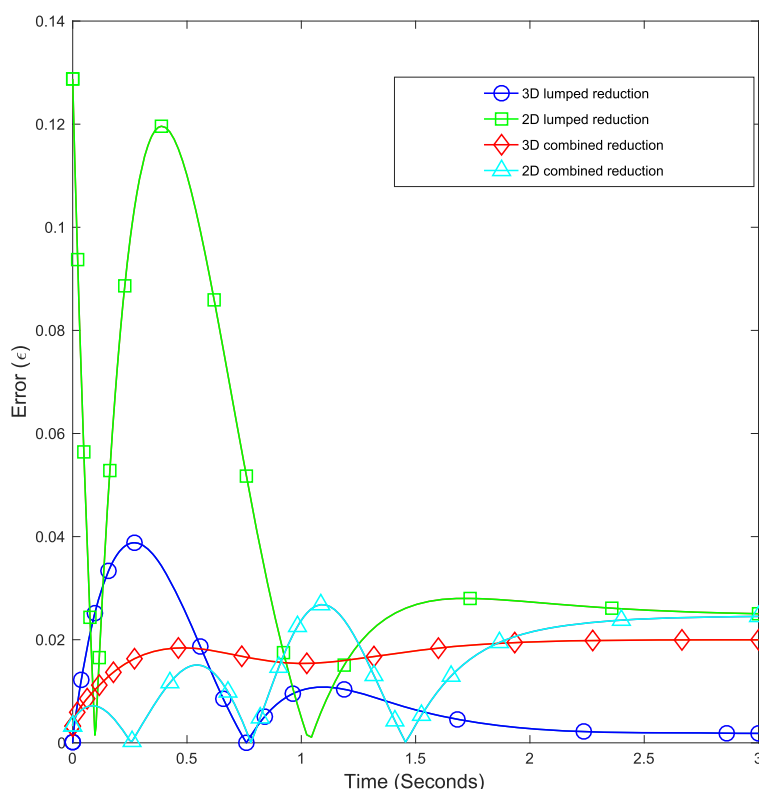
Fig. 3 Relative error between the original and reduced models of *E. coli* chemotaxis. These systems were simulated to steady-state under the introduction of $10\mu M$ concentration of extracellular attractant at $t = 0$. Figure depicts time varying errors incurred for the 3 and 2 dimensional reduced models under lumping applied in isolation and via the combined reduction algorithm

(how the cell steadily adjusts to the concentration of the chemotactic attractant), as opposed to being directly involved in the actual chemotactic flagellar response.

A model of extracellular regulatory kinase activation

The second example provided here is a model of ERK phosphorylation mediated via the EGF and NGF receptor pathways that was originally detailed in Sasagawa et al. [48]. This biological system commonly arises in the study of cancer and pain, and remains an area of ongoing clinical significance. The SBML representation employed here of this model including the parameterisation and initial conditions is available at www.ebi.ac.uk/biomodels-main/BIOMD0000000049.

This is a relatively large biochemical model describing 150 reactions and 99 species. The model additionally integrates two receptor pathways (EGFR and NGFR) allowing exploration as to how they interact. Due to its size and clinical relevance, this model therefore represents a prime candidate for the application of model reduction techniques. Although fully parameterised systems of this scope remain somewhat uncommon, their occurrence in the literature is increasing; primarily a result of increases

in data and knowledge of cellular systems at finer spatial scales. However, even with such data available, approximations may still be required to model parameters particularly in cases where the model acts as a representation of more complex underlying biochemical mechanisms. We have thus employed this system as an example to demonstrate that our methodology remains valid for systems of varying complexity and size, assuming all parameter values are known.

A full description of the model and its parameterisation can be found in Sasagawa et al. [48]. A block schematic diagram of the system is given in Fig. 5; the blocks in this diagram represent various 'submodules' of the system each containing a number of reactants and the reactions that link them. We made one minor modification to the original model; the state-variable representing total concentration of proteasome was altered to have a rate of depletion such that the model was asymptotically stable. Without this proteasome would accumulate indefinitely and cause the system to be unstable.

The analysis performed here treats only one of the pathways as salient, such that the rate of EGF binding represents the only input under consideration. Additionally, the

Snowden *et al. BMC Systems Biology* (2017) 11:17

Page 13 of 18

Fig. 4 Timecourse of output from the original 11-dimensional and reduced, 2-dimensional models of *E. coli* chemotaxis. Here, the systems were simulated to steady-state under the introduction of $10\mu M$ concentration of extracellular attractant at $t = 0$

Table 3 Comparison of maximal relative error under differing lumping inverses for the reduction of the *E coli* model via lumping in isolation

Dimensions	Moore-Penrose	Steady-state	Average
6	6.67%	0.15%	0.18%
5	10.58%	0.55%	0.51%
4	13.99%	0.54%	1.26%
3	26.93%	4.78%	4.77%
2	78.36%	13.45%	12.88%
1	70.92%	75.56%	82.34%

Each row represents a further level of dimensional reduction for the model, whilst the columns represent the different methods of lumping inverse. The 'Moore-Penrose' column contains values where the lumping inverse $\bar{L}$ is the Moore-Penrose or pseudoinverse of the lumping matrix L. The 'steady-state' column contains values where the lumping inverse $\bar{L}$ is selected to reconstruct the unperturbed steady-state values of the original system such that $\boldsymbol{x}_{\boldsymbol{u}_0}^* = \bar{L}\bar{\boldsymbol{x}}_{\boldsymbol{u}_0}^*$. The errors stated represent the maximal relative error between the original and reduced systems when simulated to steady-state under the introduction of $10\mu M$ concentration of attractant ligand at $t = 0$

total concentration of phosphorylated ERK (in complex with other species or in isolation) is regarded as the output, such that

$$y = [\text{ppERK}] + [\text{dppERK}] + [\text{ppERK.MKP3}] + [\text{dppERK.MKP3}].$$

Additionally, the system employs initial conditions such that it begins at a non-zero steady-state with no input into the model (i.e. at the natural rate of EGF binding). Note that we chose to look at the EGF receptor - ERK activation pathway in particular as its dynamical behaviour exhibits an adaptive response [48]. The retention of this nonlinear behaviour in a reduced model serves as a good demonstration of the combined algorithm's strengths and particular applicability in the field.

Conservation analysis was performed via QR factorisation and from which 23 states could be eliminated via the speed-ranking method. Exploitation of these conservation relations resulted in a 76 dimensional model which was then further reduced via the combined algorithm, results of which can be seen in Table 5 and Fig. 6.

The original model had an extremely high stiffness coefficient at the initial condition of the system of 42658. Via lumping this could be reduced to 235.7 for the 11 dimensional lumped reduction enabling application of empirical balanced truncation. In this case, the empirical Gramians were constructed using data from 50 distinct simulations covering perturbations to both the models input parameter $u(t)$ and the initial conditions $\boldsymbol{x}_0$. In both cases, perturbations were uniformly sampled from 0.4 to 1.6 times each parameter's original, unperturbed value.

The results shown in Table 5 highlight the extent of the reductions that can be obtained via the combined method

Table 2 Comparison of reduction error and stiffness coefficients at each level of dimensional reduction for the *E coli* signalling model under different approaches to conservation analysis using the combined algorithm

Dimensions	Naive elimination		Selective elimination	
	Lumping error (ε)	Stiffness	Lumping error (ε)	Stiffness
6	1.37%	876.73	0.15%	794.91
5	2.45%	904.80	0.51%	56.22
4	4.8%	907.59	0.54%	60.16
3	74.18%	542.24	4.77%	58.95
2	339.99%	99.32	12.88%	25.95
1	424.33%	1	75.56%	1

This table compares the effect of either naively or rigorously selecting which species to eliminate via the application of conservation relations. The errors stated represent the maximal relative error between the original and reduced systems when simulated to steady-state under the introduction of $10\mu M$ concentration of attractant ligand at $t = 0$

Snowden *et al. BMC Systems Biology* (2017) 11:17

Page 14 of 18

Table 4 Controllability and observability index values for the model of chemotactic signalling in *E. Coli*

Species	Controllability	Observability	Input-output
CheA · CheY	0.865	0.881	0.762
CheA$_P$	0.846	0.497	0.421
CheY$_P$ + CheA · CheY$_P$ + CheA$_P$ · CheY	1	1	1
CheY$_P$ · CheZ	0.343	0.703	0.241
CheB$_P$	0.004	0.006	2×10^{-5}

with very little associated error. In particular the combined method has again demonstrated that it can produce better reductions than either method in isolation. The 7 dimensional reduced model had only a 4.77% associated maximal relative error as compared to the original model when simulated from steady-state under a 50% inhibition of EGF receptor binding. This is equivalent to an approximately 68% improvement in error over the 7 dimensional reduction achieved by lumping in isolation. Repeated simulation revealed that the original model had an average simulation time of 1.357 seconds, whilst the reduced 7 dimensional model required only 0.144 seconds.

Note that Matlab files regarding the reduction of the ERK activation model can be found online [62].

Indices of controllability and observability

Computing the controllability and observability indices for the 11-dimensional lumped version of the ERK activation model yields the results given in Table 6. In comparison to the results for the *E. coli* chemotactic signalling model, the results here are significantly more difficult to intuit. This is primarily due to the fact that the lumped variables often include species from highly disconnected areas of the original network, making it difficult to interpret their biological significance.

Despite this, however, it is still possible to obtain some insight from the indices calculated. The most controllable

Fig. 5 Reproduction of the block schematic depiction of the ERK activation model due to Sasagawa et al. [48]

Table 5 Error results comparing the application of empirical balanced truncation, lumping, and the combined method of model reduction to the ERK activation model

Dimension	EBT error	Lumping error	Stiffness	Combined error
75	0.76%	$\approx 0\%^*$	42658	–
50	#	0.01%	42633	–
25	#	0.52%	10664	–
15	#	1.26%	7934	–
14	#	2.21%	7934	–
13	#	2.29%	7934	–
12	#	1.21%	1591	–
11	#	3.07%	236	–
10	#	6.02%	264	2.84%
9	#	10.96%	211	4.02%
8	#	13.12%	43	4.32%
7	#	14.18%	42	4.77%
6	#	29.53%	44	13.08%
5	#	39.03%	45	20.81%
4	#	46.47%	212	31.09%
3	#	54.67%	42	34.58%
2	#	53.52%	18	41.10%
1	#	55.73%	1	50.46%

Note: '#' implies Matlab could not numerically simulate this reduction using ode15s due to issues associated with stiffness. '-' implies the reduction error at this point was equal to the lumping error. The errors stated represent the maximal relative error between the original and reduced systems when simulated to steady-state under the introduction of agonist increasing the rate of EGF receptor binding by 50% at $t = 0$. * implies that the error was within the numerical tolerance of the simulator

Snowden *et al. BMC Systems Biology* (2017) 11:17

Page 15 of 18
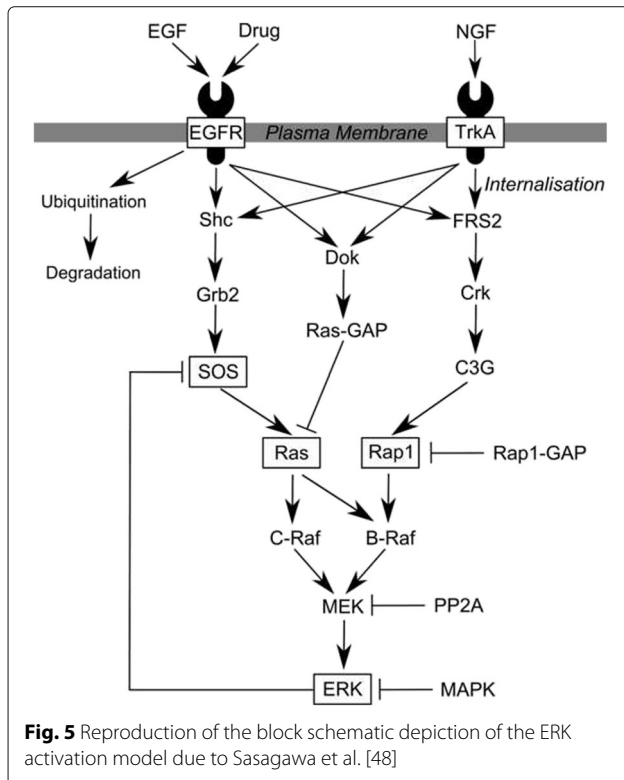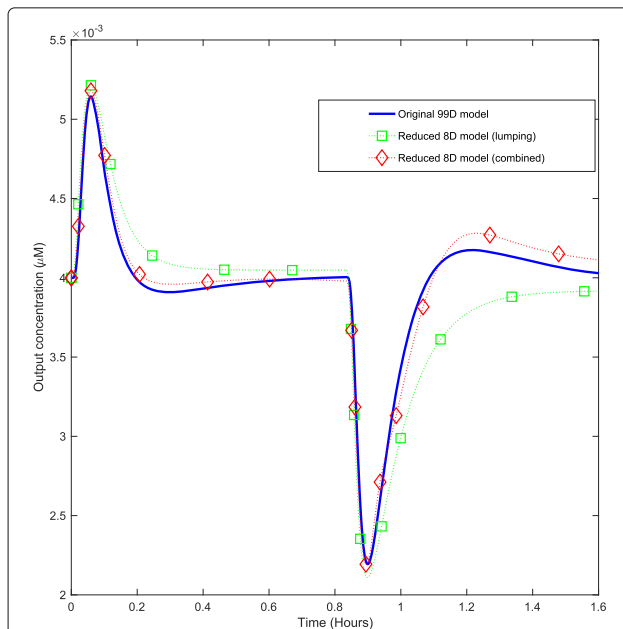
Fig. 6 Timecourses of the output from the original 99-dimensional and the reduced 8-dimensional systems. This plot emphasises the fact that the reduced model is designed to remain valid for any reasonable change in input. The system starts by being affected by an agonist that increases the rate of EGF binding by 25% for 50 minutes, at this point the input flips to an antagonist decreasing the rate of EGF binding by 50% and runs for the same time period. At any given time point the error between the original and reduced model exceeds no more than 5%

and observable lumped state-variable, $x_6(t)$, for example, primarily contains the concentration of singly phosphorylated mitogen/extracellular signal-regulated kinase (MEK), in isolation and in its various complex forms. This concurs with the known biology as phosphorylated MEK represents the point in the pathway that all possible

Table 6 Controllability and observability index values for the model of ERK activation controlled via the EGFR pathway

Lumped state-variable	Controllability index	Observability index	Input-output index
$x_1(t)$	0.3022	0.0176	0.0053
$x_2(t)$	0.0201	0.0153	0.0003
$x_3(t)$	0.0189	0.0009	1.681×10^{-5}
$x_4(t)$	0.8547	0.2752	0.2352
$x_5(t)$	0.6194	0.0079	0.0049
$x_6(t)$	1	1	1
$x_7(t)$	0.1412	0.1043	0.0147
$x_8(t)$	0.1378	0.0018	0.0003
$x_9(t)$	0.1539	0.0164	0.0025
$x_{10}(t)$	0.0473	0.0092	0.0004
$x_{11}(t)$	0.3312	0.0442	0.0146

routes of activation are directed towards before the phosphorylation of ERK occurs. The second most important state-variable, $x_4(t)$, is somewhat more difficult to parse, but is most distinguishable from the other state-variables by its inclusion of the SOS protein and associated complexes and species along this pathway. This, again, concurs with the known biology as SOS represents the most responsive branch of the pathway to EGFR binding that is described by the model. Similarly, the most unimportant state-variable, $x_3(t)$, can be seen as disproportionately representing the FRS2-C3G branch of the pathway, which is significantly more responsive to binding of the NGF receptor than that of EGF.

Overall the indices do seem to automatically provide a degree of model intuition that is often not possible for such large systems. As such they represent one possible solution to a core issue in the study of complex models. Due to the nature of the lumping algorithm, however, they are somewhat limited in their usefulness. An alternative approach, which we would hope to explore in future work, would be to constrain the lumping algorithm so as to obtain more readily interpreted lumped variables prior to the calculation of indices.

Conclusions

In this paper a combined model reduction algorithm incorporating conservation analysis, proper lumping and empirical balanced truncation for the reduction of high dimensional systems of nonlinear ODEs has been presented. This algorithm was designed for specific application to controlled models of biochemical systems. Such models typically have a number of associated properties, including nonlinearity, stiffness and high dimensionality, which any model reduction approach must address. Under the combined algorithm conservation analysis is first used to obtain a simplified realisation of the system, proper lumping is then used to reduce the model whilst also reducing the stiffness coefficient and finally empirical balanced truncation is employed to achieve a more accurate and lower dimensional reduction than could be obtained by further lumping. The algorithm has additionally been designed to be highly automatable; it is implemented so as to take models stored in the SBML format and return highly accurate reduced systems without significant user input.

Whilst the algorithm has broad applicability, crucially, it focuses on the reduction of controlled biochemical systems that are amenable to an input-output formulation. In reducing a model the algorithm seeks primarily to maintain the input-output response profile, such that the reduced model can accurately predict the output of the original model for a wide range of possible inputs. Such an approach may be particularly applicable to the newly emerging field of Quantitative Systems

Snowden *et al. BMC Systems Biology* (2017) 11:17

Page 16 of 18

Pharmacology, which aims to mechanistically describe the effects of drug administration across multiple scales of action. Here the preservation of dose-response behaviour is of primary importance and tallies well with the algorithm's emphasis on maintaining a model's input-output relationship.

One of the difficulties with reduced models obtained under this approach is that the meaning of the dynamical equations is necessarily obfuscated by the coordinate transformations that have been applied. The combined algorithm starts with a fully mechanistic 'white box' model of a signalling pathway, whereby all of the biological reactions and reactants have a one-to-one correspondence with terms in the system of ODEs. The algorithm then produces a reduced 'grey-box' model where this biological interpretation of the governing equations is now hidden by the transformations applied. Computationally, however, the reduced model remains an accurate approximation to the original. The reduced state-variables can still be mapped back to the original biological species, thus allowing mechanistic insight to be gained as opposed to solely empirical observation. Hence, whilst the reduced system may no longer serve as a directly interpretable description of the original biological system, it still retains a high degree of utility for the analysis of the model's overall input-output relationship and its mechanistic underpinnings.

The algorithm in the form presented here assumes that the model in question is initially fully parameterised. It is however possible to extend the algorithm such that the reduced models constructed are guaranteed to remain valid across a range of parameterisations by additionally evaluating perturbations to the parameter values whilst constructing the lumping matrices and the empirical Gramians. Hence, at each step the possible reduced models are tested not only for perturbations to the original input and the initial conditions, but also for perturbations to the parameterisation of the system. To maintain validity across a defined range of values in parameter space, sampling approaches such as latin hypercube sampling would likely work well, but such ideas are not explored in this initial study.

The algorithm was demonstrated via application to two systems; an 11 dimensional model of bacterial chemotactic signalling in *E. coli* and a 99 dimensional model of ERK activation mediated via the EGF and NGF receptor pathways. The results for both models highlight the extent of reduction that can be obtained with very little associated error and therefore the potential usefulness of these model reduction methods in the analysis of Systems Biology models.

From the results it can be seen that both models began with a high stiffness coefficient. As a result, the application of empirical balanced truncation to the unreduced systems yielded higher error than lumping alone and numerically failed after truncating only a small proportion of the state-variables. Crucially, the proper lumping scheme was able to act as a good pre-conditioner and efficiently stripped stiffness from the systems, hence closing eigengaps in the associated Jacobians. Via subsequent application of empirical balanced truncation the combined algorithm was able to produce significantly better reductions than could be obtained by any of the individual methods in isolation. Furthermore, these reduced systems exhibited a significant speed-up in simulation time. In the case of the *E. coli* model we observed a 96.1% speed-up in simulation time whilst only incurring a 2.8% error. The ERK activation model demonstrated a 89.4% speed-up in simulation time in exchange for a 4.77% maximal relative error. For applications such as parameter optimisation or agent based modelling approaches, where an extremely large number of simulations may be required, speed-up is a substantial benefit. This speed-up can be attributed to two main factors - the reduction in the number of state variables present in each system and the reduction in their overall stiffness coefficients.

There are still a number of open questions with regards to this work, in particular to what extent does a model reduction (specifically, under a coordinate transformation of the state variables) remain accurate if the parameters in the original system are altered? Additionally, are there reduced models that remain valid over a larger range of parameterisations than others? If now, instead of considering the rate parameters to represent a set of fixed values, we consider the parameters to represent a distribution of physically acceptable values how does this affect our selection of the best reduction? Understanding these questions is a necessary next step in further growing the practical applicability of model reduction for biochemical systems.

The paper also introduced the concept of empirical indices of controllability and observability for biochemical systems; as was demonstrated these indices can be used to provide substantial insight into models of even a very large size. This usefulness, however, is dependent upon the biological interpretation of the specific lumping used to precondition the system. Developing a lumping algorithm specifically for the calculation of these indices represents a reasonable extension of this work.

The combined model reduction algorithm has demonstrated good results for the reduction of controlled, nonlinear, stiff, high-dimensional models of biochemical reaction networks. Reduced systems produced under this approach maintain a high degree of input-output accuracy and see a significant speed-up in simulation time. These results justify research into the combining of complementary reduction methodologies and highlight the use of empirical balanced truncation, which had not previously seen application in the reduction of biochemical systems.

Snowden *et al. BMC Systems Biology* (2017) 11:17

Page 17 of 18

Additional file

Additional file 1: Provides a mathematical account of the underlying methodologies employed and gives a detailed, stepped proceedure for the application of the combined reduction algorithm. (PDF 631 kb)

Author details
[1] Department of Mathematics and Statistics, University of Reading, RG6 6AX, Reading, UK. [2] Certara QSP, University of Kent Innovation Centre, CT2 7FG Canterbury, UK. [3] Leiden Academic Centre for Drug Research, Universiteit Leiden, NL-2333 CC Leiden, Netherlands. [4] The Institute for Cardiovascular and Metabolic Research (ICMR), University of Reading, RG6 6AX Reading, UK.

References
1. Li C, Donizelli M, Rodriguez N, Dharuri H, Endler L, Chelliah V, Li L, He E, Henry A, Stefan MI, Snoep JL, Hucka M, Le Novere N, Laibe C. BioModels Database: An enhanced, curated and annotated resource for published quantitative kinetic models. BMC Syst Biol. 2010;4:92.
2. Okino M, Mavrovouniotis M. Simplification of mathematical models of chemical reaction systems. Chem Rev. 1998;98(2):391–408.
3. Petrov V, Nikolova E, Wolkenhauer O. Reduction of nonlinear dynamic systems with an application to signal transduction pathways. Syst Biol IET. 2007;1(1):2–9.
4. Schneider KR, Wilhelm T. Model reduction by extended quasi-steady-state approximation. J Math Biol. 2000;40(5):443–50.
5. Vejchodskỳ T, Erban R, Maini PK. Reduction of chemical systems by delayed quasi-steady state assumptions. 2014. arXiv preprint arXiv:1406.4424.
6. Vejchodskỳ T. Accurate reduction of a model of circadian rhythms by delayed quasi steady state assumptions. 2013. arXiv preprint arXiv:1312.2825.
7. Choi J, Yang KW, Lee TY, Lee SY. New time-scale criteria for model simplification of bio-reaction systems. BMC bioinformatics. 2008;9(1):338.
8. West S, Bridge LJ, White MR, Paszek P, Biktashev VN. A method of speed coefficients for biochemical model reduction applied to the NF-κB system. J Math Biol. 2015;70(3):591–620. arXiv preprint arXiv:1403.1610.
9. Härdin HM, Zagaris A, Krab K, Westerhoff HV. Simplified yet highly accurate enzyme kinetics for cases of low substrate concentrations. FEBS J. 2009;276(19):5491–506.
10. Vora N, Daoutidis P. Nonlinear model reduction of chemical reaction systems. AIChE J. 2001;47(10):2320–32.
11. Gerdtzen ZP, Daoutidis P, Hu WS. Nonlinear model reduction for energy metabolism in Saccharomyces cerevisiae. In: American Control Conference, 2002. Proceedings of the 2002; 2002. p. 2867–72. IEEE.
12. Gerdtzen ZP, Daoutidis P, Hu WS. Non-linear reduction for kinetic models of metabolic reaction networks. Metab Eng. 2004;6(2):140–54.
13. Prescott TP, Papachristodoulou A. Layering in networks: The case of biochemical systems. In: American Control Conference (ACC), 2013; 2013. p. 4544–49. IEEE.
14. Prescott TP, Papachristodoulou A. Layered decomposition for the model order reduction of timescale separated biochemical reaction networks. J Theor Biol. 2014;356:113–22.
15. Sivakumar H, Hespanha JP. Towards modularity in biological networks while avoiding retroactivity. In: American Control Conference (ACC), 2013; 2013. p. 4550–556. IEEE.
16. Maas U, Pope SB. Simplifying chemical kinetics: intrinsic low-dimensional manifolds in composition space. Combustion Flame. 1992;88(3):239–64.
17. Vallabhajosyula RR, Sauro HM. Complexity reduction of biochemical networks. In: Simulation Conference, 2006. WSC 06. Proceedings of the Winter; 2006. p. 1690–1697. IEEE.
18. Zobeley J, Lebiedz D, Kammerer J, Ishmurzin A, Kummer U. A new time-dependent complexity reduction method for biochemical systems. In: Transactions on Computational Systems Biology I. Berlin, Heidelberg: Springer; 2005. p. 90–110.
19. Surovtsova I, Zobeley J. Focusing on dynamic dimension reduction for biochemical reaction systems. Understanding Exploiting Syst Biol Biomed Bioprocesses. 2006;31:31–46.
20. Surovtsova I, Simus N, Lorenz T, König A, Sahle S, Kummer U. Accessible methods for the dynamic time-scale decomposition of biochemical systems. Bioinformatics. 2009;25(21):2816–23.
21. Kourdis PD, Goussis DA, Steuer R. Physical understanding via reduction of complex multiscale models: glycolysis in Saccharomyces cerevisiae. In: BioInformatics and BioEngineering, 2008. BIBE 2008. 8th IEEE International Conference On; 2008. p. 1–6. IEEE.
22. Kourdis PD, Steuer R, Goussis DA. Physical understanding of complex multiscale biochemical models via algorithmic simplification: Glycolysis in Saccharomyces cerevisiae. Physica D: Nonlinear Phenomena. 2010;239(18):1798–817.
23. Kourdis PD, Palasantza AG, Goussis DA. Algorithmic asymptotic analysis of the NF-κB signaling system. Comput Math Appl. 2013;65(10):1516–34.
24. Surovtsova I, Simus N, Hübner K, Sahle S, Kummer U. Simplification of biochemical models: a general approach based on the analysis of the impact of individual species and reactions on the systems dynamics. BMC Syst Biol. 2012;6(1):14.
25. Degenring D, Froemel C, Dikta G, Takors R. Sensitivity analysis for the reduction of complex metabolism models. J Process Control. 2004;14(7): 729–45.
26. Liu G, Swihart MT, Neelamegham S. Sensitivity, principal component and flux analysis applied to signal transduction: the case of epidermal growth factor mediated signaling. Bioinformatics. 2005;21(7):1194–202.
27. Smets I, Bernaerts K, Sun J, Marchal K, Vanderleyden J, Van Impe J. Sensitivity function-based model reduction: a bacterial gene expression case study. Biotech Bioeng. 2002;80(2):195–200.
28. Apri M, de Gee M, Molenaar J. Complexity reduction preserving dynamical behavior of biochemical networks. J Theor Biol. 2012;304: 16–26.
29. Maurya M, Bornheimer S, Venkatasubramanian V, Subramaniam S. Reduced-order modelling of biochemical networks: application to the GTPase-cycle signalling module. IEE Proc Syst Biol. 2005;152(4):229–42.

Snowden *et al. BMC Systems Biology* (2017) 11:17

Page 18 of 18

30. Maurya MR, Scott JB, Venkatasubramanian V, Subramaniam S. Model-reduction by simultaneous determination of network topology and parameters: Application to modules in biochemical networks. In: 2005 Annual Meeting AIChE; 2005.

31. Maurya M, Bornheimer S, Venkatasubramanian V, Subramaniam S. Mixed-integer nonlinear optimisation approach to coarse-graining biochemical networks. IET Syst Biol. 2009;3(1):24–39.

32. Hangos KM, Gábor A, Szederkényi G. Model reduction in bio-chemical reaction networks with michaelis-menten kinetics. In: European Control Conference (ECC), July 17-19 2013, Zurich; 2013. p. 4478–483.

33. Taylor SR, Petzold LR, et al. Oscillator model reduction preserving the phase response: application to the circadian clock. Biophys J. 2008;95(4): 1658–73.

34. Anderson J, Chang YC, Papachristodoulou A. Model decomposition and reduction tools for large-scale networks in systems biology. Automatica. 2011;47(6):1165–74.

35. Prescott TP, Papachristodoulou A. Guaranteed error bounds for structured complexity reduction of biochemical networks. J Theor Biol. 2012;304:172–82.

36. Danø S, Madsen MF, Schmidt H, Cedersund G. Reduction of a biochemical model with preservation of its basic dynamic properties. FEBS J. 2006;273(21):4862–77.

37. Dokoumetzidis A, Aarons L. Proper lumping in systems biology models. IET Syst Biol. 2009;3(1):40–51.

38. Gulati A, Isbister G, Duffull S. Scale reduction of a systems coagulation model with an application to modeling pharmacokinetic–pharmacodynamic data. CPT: Pharmacometrics Syst Pharmacol. 2014;3(1):90.

39. Koschorreck M, Conzelmann H, Ebert S, Ederer M, Gilles ED. Reduced modeling of signal transduction–a modular approach. BMC Bioinformatics. 2007;8(1):336.

40. Sunnåker M, Schmidt H, Jirstrand M, Cedersund G. Zooming of states and parameters using a lumping approach including back-translation. BMC Syst Biol. 2010;4(1):28.

41. Sunnåker M, Cedersund G, Jirstrand M. A method for zooming of nonlinear models of biochemical systems. BMC Syst Biol. 2011;5(1):140.

42. Liebermeister W, Baur U, Klipp E. Biochemical network models simplified by balanced truncation. FEBS J. 2005;272(16):4034–43.

43. Härdin H, van Schuppen J. System reduction of nonlinear positive systems by linearization and truncation. Positive Syst. 2006;431–38.

44. Sootla A, Anderson J. On projection-based model reduction of biochemical networks – Part I: The deterministic case. Los Angeles: IEEE 53rd Annual Conference on Decision and Control (CDC); 2014. arXiv preprint arXiv:1403.3579.

45. van der Graaf PH, Benson N. Systems pharmacology: bridging systems biology and pharmacokinetics-pharmacodynamics (PKPD) in drug discovery and development. Pharm Res. 2011;28(7):1460–4.

46. Sorger PK, Allerheiligen SR, Abernethy DR, Altman RB, Brouwer KL, Califano A, D'Argenio DZ, Iyengar R, Jusko WJ, Lalonde R, et al. Quantitative and systems pharmacology in the post-genomic era: new approaches to discovering drugs and understanding therapeutic mechanisms. In: An NIH White Paper by the QSP Workshop Group-October; 2011.

47. Tindall M, Porter S, Wadhams G, Maini P, Armitage J. Spatiotemporal modelling of CheY complexes in *Escherichia coli* chemotaxis. Progress Biophys Mol Biol. 2009;100(1):40–6.

48. Sasagawa S, Ozaki YI, Fujita K, Kuroda S. Prediction and validation of the distinct dynamics of transient and sustained ERK activation. Nat Cell Biol. 2005;7(4):365–73.

49. Murray JD. Mathematical Biology I: An Introduction. New York: Springer; 2002.

50. Klipp E, Liebermeister W, Wierling C, Kowald A, Lehrach H, Herwig R. Systems Biology. Oxford: Wiley-Blackwell; 2013.

51. Sauro HM, Ingalls B. Conservation analysis in biochemical networks: computational issues for software writers. Biophys Chem. 2004;109(1): 1–15.

52. Vallabhajosyula RR, Chickarmane V, Sauro HM. Conservation analysis of large biochemical networks. Bioinformatics. 2006;22(3):346–53.

53. Antoulas AC. Approximation of Large-scale Dynamical Systems. Advances in Design and Control. Philadelphia: Society for Industrial and Applied Mathematics; 2005.

54. Hahn J, Edgar TF. An improved method for nonlinear model reduction using balancing of empirical Gramians. Comput Chem Eng. 2002;26(10): 1379–97.

55. Dullerud GE, Paganini F, Vol. 6. A Course in Robust Control Theory. New York: Springer; 2000.

56. Laub AJ, Heath MT, Paige CC, Ward RC. Computation of system balancing transformations and other applications of simultaneous diagonalization algorithms. Autom Control IEEE Trans. 1987;32(2):115–22.

57. Bornstein BJ, Keating SM, Jouraku A, Hucka M. LibSBML: an API library for SBML. Bioinformatics. 2008;24(6):880–1.

58. Schmidt H, Jirstrand M. Systems biology toolbox for matlab: a computational platform for research in systems biology. Bioinformatics. 2006;22(4):514–5.

59. Wei J, Kuo JC. Lumping analysis in monomolecular reaction systems. analysis of the exactly lumpable system. Ind Eng Chem Fundam. 1969;8(1):114–23.

60. Kuo JC, Wei J. Lumping analysis in monomolecular reaction systems. analysis of approximately lumpable system. Ind Eng Chem Fundam. 1969;8(1):124–33.

61. Li G, Rabitz H. A general analysis of approximate lumping in chemical kinetics. Chem Eng Sci. 1990;45(4):977–1002.

62. Snowden TJ, van der Graaf PH, Tindall MJ. Reduction results for a mathematical model of ERK activation (Matlab Files). 2016. doi:10.5281/zenodo.192503 https://doi.org/10.5281/zenodo.192503. Accessed 12 Apr 2016.