Automatica 47 (2011) 1165–1174

Contents lists available at ScienceDirect

Automatica

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

Model decomposition and reduction tools for large-scale networks in systems biology☆

James Anderson [a,b], Yo-Cheng Chang [b,c], Antonis Papachristodoulou [a,*]

[a] Department of Engineering Science, University of Oxford, Parks Road, Oxford, OX1 3PJ, United Kingdom

[b] Life Sciences Interface Doctoral Training Centre, University of Oxford, Wolfson Building, Parks Road, Oxford, OX1 3QD, United Kingdom

[c] Department of Biochemistry, University of Oxford, South Parks Road, Oxford, OX1 3QU, United Kingdom

ARTICLE INFO

Article history:
Received 15 February 2010
Received in revised form
11 August 2010
Accepted 21 February 2011
Available online 8 April 2011

Keywords:
Systems biology
Large-scale systems
Model reduction
Model decomposition

ABSTRACT

Biological system models are routinely developed in modern systems biology research following appropriate modelling/experiment design cycles. Frequently these take the form of high-dimensional nonlinear Ordinary Differential Equations that integrate information from several sources; they usually contain multiple time-scales making them difficult even to simulate. These features make systems analysis (understanding of robust functionality) — or redesign (proposing modifications in order to improve or modify existing functionality) a particularly hard problem. In this paper we use concepts from systems theory to develop two complementary tools that can help us understand the complex behaviour of such system models: one based on model decomposition and one on model reduction. Our aim is to algorithmically produce biologically meaningful, simplified models, which can then be used for further analysis and design. The tools presented are applied on a model of the Epidermal Growth Factor signalling pathway.

1. Introduction

Mathematical modelling is undoubtedly an important tool for understanding the behavior of complex biological systems (Lander, 2004) but also for making predictions about system behaviour far from its nominal operating conditions. The field of systems biology has seen a considerable increase in the number of system models, thanks in part to the continual increase in computational power and high throughput experimental techniques. This has allowed modellers to combine information and experimental data from several sources to construct very complicated system descriptions. For example, in cell signalling models, information from numerous experiments taken from different parts of these large pathways are frequently integrated to produce a single large-scale model that aims to describe the behaviour of the entire network (Hildebrandt, Raden, Petzold, Robinson, & Doyle, 2008).

Typically these integrative models are described by sets of coupled, high-dimensional nonlinear differential equations which contain a large number of uncertain parameters. Other frameworks include statistical models, stochastic differential equations and Boolean networks. An overview of these and other approaches is given in Klipp et al. (2009). The aim is to use these models to gain insight into the functionality of biological systems, but their underlying complexity is sometimes overwhelming. Of particular benefit would be a simpler model description which is able to capture key characteristics of the system and which distills the essence of the regulatory mechanisms that the system employs. Such a model would be more amenable for analysis and redesign.

While mathematical and computational biology have provided valuable modelling and simulation tools, systems engineering principles and appropriate algorithmic tools need to be developed to analyse the properties of large signalling pathways (Sontag, 2004). There are several examples of work in this direction—e.g., the structured singular value framework was used to identify the robustly stable volume in the multi-dimensional parameter space of a Fas signalling-induced apoptosis model (Shoemaker & Doyle, 2008). In this example, the system model contained more than 60 parameters and had 31 coupled differential equations based on the original model described in Bagci, Vodovotz, Billiart, and Bahar (2006). Models of this degree of complexity are common in the literature. For more examples of large scale system models (Kim, Rath, Kolch, & Cho, 2007; Lee, Gianchandani, Eddy, & Papin, 2008; Singhand, Jayaraman, & Hahn, 2006; Vilar, Jansen, & Sander,

☆ The material in this paper was partially presented at the 49th IEEE Conference on Decision and Control, December 15–17, 2010, Atlanta, Georgia, USA, and the 2010 American Control Conference, June 30–July 2, 2010, Baltimore, Maryland, USA. This paper was recommended for publication in revised form by Associate Editor Elling Jacobsen under the direction of Guest Editor Francis J. Doyle III. J. Anderson and Y.-C. Chang contributed equally to this research.

* Corresponding author. Tel.: +44 1865 283036.
E-mail addresses: james.anderson@eng.ox.ac.uk (J. Anderson), yo-cheng.chang@univ.ox.ac.uk (Y.-C. Chang), antonis@eng.ox.ac.uk (A. Papachristodoulou).

0005-1098/$ – see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.automatica.2011.03.010

2006) provide a selection that cover various metabolic and signal transduction pathways. In recent work (Huang, Chu, & Hahn, 2010) the 68-state model of the IL-6 signal pathway (Singhand et al., 2006) containing over 100 parameters was reduced using techniques from systems theory. Another example of how systems/control theoretical tools can aid simplification is given in Hornberg et al. (2005) where kinetic parameter perturbations were used to measure the robustness of the MAPK signaling cascade (Waldherr, Allgöwer, & Jacobsen, 2009). Along similar lines the objective of this paper is to introduce new theoretical tools rooted in control theory to provide simplified models that can be used to gain insight into the original system properties.

In this paper two complementary approaches are presented that can be used to simplify large-scale signalling pathways to enable the understanding of their structure and dynamics. The first is a model reduction algorithm that is particularly suited to biological systems analysis. The motivation is to develop a lower order model with very similar dynamic properties as the original model but with a state vector which is a strict subset of the state vector in the original system and avoids a coordinate transformation. The algorithm can be thought of as collapsing species in a signalling pathway according to their importance to the system dynamics. The resulting model can be used to determine functionality of the original one, as a bound on the system error is also obtained. The algorithm is developed for linear systems and then extended for use with systems containing parametric uncertainty in the Linear Parameter Varying (LPV) framework.

Next, a decomposition algorithm is presented that allows for the identification of 'modular' components in the signalling network from the system dynamics. In contrast to the reduction approach, the decomposition algorithm aims to break the original large-scale system into a set of interacting subsystems by minimising their interaction. Finding these subunits can allow us to gain a better understanding of the interacting components and their dynamic behaviour that characterise the full pathway.

The reduction and decomposition techniques are applied on the Epidermal Growth Factor Receptor (EGFR) signalling pathway, involved in various mammalian cellular processes (Herbst, 2004). The system model considered is typical of the type of high-fidelity models mentioned previously making it difficult to analyse using conventional techniques. It consists of 35 coupled differential equations with over 70 parameters, some of which are uncertain, and includes states that evolve over largely differing time scales. We show how our model reduction approach reveals the underlying pathway structure, while the decomposition algorithm determines the biologically significant functional subsystems in the large-scale model.

In Section 2 we present some background material and the notation used in the sequel. Section 3 introduces a new model reduction algorithm while in Section 4 we develop a system decomposition algorithm for complex signalling pathway models. Section 5 illustrates the reduction and decomposition algorithms on an EGFR-MAPK pathway. The paper is concluded in Section 6.

2. Background

In this paper we consider Ordinary Differential Equation (ODE) models of biochemical reaction networks representing signalling pathways, which take the form

$$\dot{x}(t) = f(x(t), p), \qquad x(0) = x_0 \tag{1}$$
$$y(t) = g(x(t))$$

where $x \in \mathbb{R}^n$ is the state vector and $p \in \mathcal{P}$ is a vector of parameters where

$$\mathcal{P} = \left\{ p \in \mathbb{R}^{n_p} \mid p_i \in \left[\underline{p_i}, \overline{p_i}\right],\ i = 1, \ldots, n_p \right\}. \tag{2}$$

It is assumed that the vector field $f : \mathbb{R}^n \times \mathbb{R}^{n_p} \to \mathbb{R}^n$ is sufficiently smooth so as to ensure existence and uniqueness of solutions. The output function $g : \mathbb{R}^n \to \mathbb{R}^m$ is also assumed to be smooth. Such assumptions are justified as the models we consider are derived using the law of mass action. Moreover, we assume throughout that for every parameter choice p there is a stable equilibrium point x^* located at the origin; we are interested in the properties of the system about this equilibrium point.

Traditional control theoretical analysis approaches cannot be applied directly to a system of the form (1). Instead, it is customary to develop a Linear Parameter Varying (LPV) model for system (1), which can be analysed more efficiently,

$$\dot{x} = F(\theta)x, \qquad y = Gx, \qquad x(0) = x_0. \tag{3}$$

The state matrix $F(\theta)$ depends affinely on $\theta \in \Theta$

$$F(\theta) = F_0 + \sum_{i=1}^{k} \theta_i F_i, \tag{4}$$

where F_i are known and fixed. Here,

$$\Theta = \{\theta \in \mathbb{R}^k \mid \theta_i \in [\underline{\theta_i}, \overline{\theta_i}],\ i = 1, \ldots, k\}. \tag{5}$$

There are numerous methods for developing an LPV model of a system; these include classical symbolic linearisation approaches and fitting. The method used in this paper determines the parameters which the nonlinear system is most sensitive to by carrying out a preliminary sensitivity analysis as described in Varga, Moormannand, and Grubel (1998). The most sensitive parameters are then chosen to be the uncertain variables, θ_i, in the LPV model. The parameter space is then discretised at R points and *pointwise* state space matrices $(\bar{F}_r, \bar{G}_r)$, $r = 1, \ldots, R$ are created for all parameter combinations through linearisation. A least squares error approach is subsequently used to fit a plane to each of the elements in the pointwise state space matrices. An extra variable can be used to cover for the error in the fit (Mannchen & Well, 2002).

The LPV modelling framework is attractive as several system analysis tools have been developed in the past decade based on Linear Matrix Inequalities and Semidefinite Programming, see Apkarian and Tuan (2000) and Boyd, El Ghaoui, Feron, and Balakrishnan (1994). These algorithms have a worst-case polynomial time complexity; the methods developed in this paper are designed to scale well with system size and complexity, but also to respect the original system's structure in order to facilitate analysis.

3. Model reduction

The objective of a model reduction procedure is to construct a new system model with state $\tilde{x} \in \mathbb{R}^{\tilde{n}}$ where $\tilde{n} < n$, so that the error between the system output of the original system, y, and the system output of the reduced system, $\tilde{y}$, is as small as possible, in some norm.

In the control literature various metrics, usually based on system norms, have been used to measure how close the reduced order model is to the original system. The balanced truncation method (Glover, 1984; Moore, 1981) is frequently used for linear model reduction and can cope with structured uncertainty when the system is represented as a Linear Fractional Transform (LFT) (Beck, Doyle, & Glover, 1996). A strong argument for using balanced truncation is that a bound on the size of the error between the models is provided; some recent extensions to balanced truncation allow for the reduction of nonlinear models (Lall, Marsden, & Glavaški, 2002) and preserve structure (Sandberg & Murray, 2009). Unfortunately balanced truncation is of limited use when trying to infer functionality of a reduced order systems biology model

as it introduces a coordinate transformation that may not make sense biologically, and which destroys the signalling network structure. Okino and Mavrovouniotis (1998) provide an overview of reduction techniques that are specifically designed for chemical reaction systems.

A well-utilised technique for model reduction of biochemical reaction networks is *lumping* (Huang, Fairweather, Griffiths, Tomlin, & Brad, 2005; Kuo & Wei, 1969). The term refers to reducing the system dimensionality by replacing a number of system states by affinely (and in some cases nonlinearly) combining them into one 'supernode'. While the lumping method produces lower order models it is not easily automated nor particularly useful for detecting the underlying structure of the system of interest. It is also likely that the states retained, which now contain an aggregation of numerous other states, will not be biologically meaningful thus rendering the model of limited practical use.

A method that maintains the coordinate system of the original model and which is often used in mathematical biology is based on singular perturbation. This takes advantage of the fact that frequently, biological systems contain states that evolve over differing time scales: simulating or analysing such systems is both numerically and computationally demanding due to the stiffness in their dynamics. Time scale separation can be used to reduce the system's size, thus alleviating the high dimensionality and ill-conditioning caused by the interaction of fast and slow acting systems (Gerdtzen, Daoutidis, & Hu, 2004; Kokotović, Khalil, & O'Reilly, 1999; Lee & Othmer, 2010; Naidu, 2002; Saksena, O'Reilly, & Kokotovic, 1984). Typically the solution involves identifying which states in the system evolve faster than others. The assumption that the fast dynamics tend to a quasi-steady state quickly, leads to a set of algebraic equations which are used in place of the fast differential equations. Such an approach is often used in the analysis of biochemical reaction networks and is known as the *quasi-steady state approximation* (Murray, 1993; Zagaris, Kaper, & Kaper, 2004). Singular perturbation methods have proven very effective for analysing large-scale systems by providing reduced order models, however, it is often difficult and requires intuition to recognise by inspection when such an analysis can be applied. Furthermore, the process of carrying out singular perturbation is not particularly suitable for automation and can in some instances only work with particular parameter values (Ciliberto, Capuani, & Tyson, 2007). It is also not straightforward to derive an error bound between the original and reduced order model using the singular perturbation framework.

In the next section we present a model reduction algorithm which can be used to uncover the structure of the underlying biological system while avoiding any co-ordinate transformations and ensuring that the state vector in the reduced model is a strict subset of the one in the full model. The approach is similar to singular perturbation but does not try to identify fast or slow states; in fact, it collapses states based on the worst-case error that could result between the original and the reduced order model if a particular set of states were collapsed. In the sequel the algorithm will be extended to the Linear Parameter Varying (LPV) case.

3.1. Linear model reduction algorithm

Consider a linear system to be reduced,

$$\dot{x} = Fx, \qquad y = Gx, \qquad x(0) = x_0, \quad x \in \mathbb{R}^n,\ y \in \mathbb{R}^m, \tag{6}$$

and suppose the reduced system is of the form

$$\dot{\tilde{x}} = \tilde{F}\tilde{x}, \qquad \tilde{y} = \tilde{G}\tilde{x}, \qquad \tilde{x}(0) = \tilde{x}_0, \quad \tilde{x} \in \mathbb{R}^{\tilde{n}},\ \tilde{y} \in \mathbb{R}^m \tag{7}$$

where $\tilde{n} < n$. It is assumed that F is Hurwitz. We want to minimise $\|y - \tilde{y}\|_2^2$ from initial conditions with $\|x_0\|_2 = 1$, while ensuring that $\tilde{x}$ is a strict subset of x. Assume for now that the states to be collapsed to algebraic relations, $\hat{x}$, have been identified and that the original state vector x is reordered such that

$$x = \begin{bmatrix} \tilde{x} \\ \hat{x} \end{bmatrix}. \tag{8}$$

Construct the augmented state vector

$$e = \begin{bmatrix} x \\ \tilde{x} \end{bmatrix}, \quad e \in \mathbb{R}^{n+\tilde{n}}$$

and the corresponding dynamical system:

$$\dot{e} = \underbrace{\begin{bmatrix} F & 0 \\ 0 & \tilde{F} \end{bmatrix}}_{F_e} e, \qquad z = \underbrace{\begin{bmatrix} G & -\tilde{G} \end{bmatrix}}_{G_e} e, \qquad e_0 = \begin{bmatrix} x_0 \\ \tilde{x}_0 \end{bmatrix}. \tag{9}$$

Note that the second element of G_e in (9) is negated, thus the output is given by $z = y - \tilde{y}$. The following result gives the worst case error between the original and reduced systems, for all initial conditions x_0 of unit size.

Theorem 1. *Consider systems (6) and (7). Suppose that there is a* $\Psi \succeq 0$ *such that*

$$F_e^T\Psi + \Psi F_e + G_e^T G_e = 0.$$

Partition Ψ *appropriately into*

$$\Psi = \begin{bmatrix} \Psi_{11} & \Psi_{12} & \Psi_{13} \\ \Psi_{12}^T & \Psi_{22} & \Psi_{23} \\ \Psi_{13}^T & \Psi_{23}^T & \Psi_{33} \end{bmatrix},$$

where $\Psi_{11},\ \Psi_{33} \in \mathbb{R}^{\tilde{n}\times\tilde{n}}$ *and* $\Psi_{22} \in \mathbb{R}^{(n-\tilde{n})\times(n-\tilde{n})}$. *Then* $\|y - \tilde{y}\|_2^2 = x_0^T Q x_0$, *where*

$$Q = \begin{bmatrix} \Psi_{11} + \Psi_{13} + \Psi_{13}^T + \Psi_{33} & \Psi_{12} + \Psi_{23}^T \\ \Psi_{12}^T + \Psi_{23} & \Psi_{22} \end{bmatrix},$$

and the worst case error results from the initial condition x_0 *aligned with the eigenvector direction corresponding to the largest eigenvalue of* Q, $\bar{\lambda}_Q$. *In fact we have*

$$\max_{\|x_0\|_2=1} \|y - \tilde{y}\|_2 = \sqrt{\bar{\lambda}_Q}.$$

Proof. The $\mathcal{L}_2$ norm of the output signal z, $\|z\|_2^2 = \int_0^\infty z^T(t)z(t)\mathrm{d}t$ is the reduction error from a given initial condition e_0, and can be computed using

$$\|z\|_2^2 = e_0^T\Psi e_0, \tag{10}$$

where $\Psi \succeq 0$ solves the Lyapunov equation (Dullerud & Paganini, 2000)

$$F_e^T\Psi + \Psi F_e + G_e^T G_e = 0. \tag{11}$$

The objective now is to find the initial condition e_0 that produces the largest error (10). There is however a special structure that needs to be imposed on e_0: the states in the reduced order system $\tilde{x}$, must be released from the same initial conditions $\tilde{x}_0$ as those in the original system, x_0. Using (8) and (10) we get:

$$\|z\|_2^2 = \begin{bmatrix} \tilde{x}_0 \\ \hat{x}_0 \\ \tilde{x}_0 \end{bmatrix}^T \begin{bmatrix} \Psi_{11} & \Psi_{12} & \Psi_{13} \\ \Psi_{12}^T & \Psi_{22} & \Psi_{23} \\ \Psi_{13}^T & \Psi_{23}^T & \Psi_{33} \end{bmatrix} \begin{bmatrix} \tilde{x}_0 \\ \hat{x}_0 \\ \tilde{x}_0 \end{bmatrix} \triangleq x_0^T Q x_0, \tag{12}$$

where Q is as defined in the statement of the theorem. The direction of a unit-norm input x_0 that produces the worst-case error is in the direction of the eigenvector associated with the largest eigenvalue of Q, with the size of the error squared clearly being equal to said eigenvalue. □

Table 1
Greedy linear reduction algorithm.

```
1 x̃ ← x, x̂ ← []
2 For j = 1 : n
3     For i = 1: length (x̃)
4         x̂[j] ← x̃[i]
5         Compute F̃, G̃
6         Construct F_e and G_e from F̃ and G̃
7         Solve (11) to obtain Ψ
8         Construct Q
9         errornorm [i] ← sqrt(λ̄_Q)
10    End For
11    minerrornorm [j] ← min (errornorm)
12    x̂[j] ←species corresponding to minerrornorm[j]
13    Clear errornorm
14    Remove x̂[j] from x̃
15 End For
```

Remark 1. As states are removed from the original system, the norm of the initial condition vector in the reduced system may decrease. To ensure a fair comparison, the programme below maintains a proportional balance between the size of $\tilde{x}_0$ and $\hat{x}_0$, depending on their dimension:

$$\begin{aligned} \max \quad & x_0^T Q x_0 \\ \text{subject to} \quad & \|\tilde{x}_0\|_2^2 = \frac{\tilde{n}}{n} \\ & \|\hat{x}_0\|_2^2 = \frac{\hat{n}}{n}. \end{aligned} \tag{13}$$

This is a Quadratically Constrained Quadratic Program (QCQP) (Lobo, Vandenberghe, Boyd, & Lebret, 1998) and can be addressed using Semidefinite Programming (Vandenberghe & Boyd, 1996). The constraints ensure that both norms, $\tilde{x}_0$ and $\hat{x}_0$ sum to one, i.e. $\|\tilde{x}_0\|_2^2 + \|\hat{x}_0\|_2^2 = \|x_0\|_2^2 = 1$.

The full reduction algorithm (Table 1) embeds Theorem 1 in a greedy scheme that systematically removes states which minimise the worst case output of the error system (9). Insight into when a greedy algorithm provides near optimal solutions can be found in Tropp (2004). A greedy algorithm is desirable because the alternatives, optimising over the space of permutation matrices, is a non-convex combinatorial problem and an exhaustive search would incur a computational cost of order $\mathcal{O}(2^n)$. In contrast, greedy methods are simple to implement and their performance can be easily kept track of; the method implemented here scales according to $\mathcal{O}(n^2)$. The input to the reduction algorithm is the system model to be reduced in the form of (6). The output is an ordered list of states and the associated worst-case $\mathcal{L}_2$ error incurred by the removal of these states.

3.2. Uncertain model reduction

The model reduction algorithm discussed above can be extended to deal with uncertain system models. The objective is to produce an uncertain, low order model such that the worst-case error between the original and reduced models is minimised for all uncertainties in the original model. For system models that allow for structured uncertainty, the LFT framework can be used in conjunction with a balanced truncation approach (Beck et al., 1996; Lall & Beck, 1999). For systems that can be described in the Linear Parameter Varying (LPV) framework, specifically systems that contain polytopic uncertainty, methods exist that give a guaranteed induced $\mathcal{L}_2$ norm on the error (Wu., 1996), however this approach leads to a non-convex optimisation problem.

We assume that the biological system of interest is in LPV form. The objective is to be able to calculate the error norm using LMIs. For that reason, it would be convenient if the reduced system was in LPV form too. However, this requires that we do not back-substitute the collapsed states but rather keep the reduced system in descriptor form.

The error system analogous to (9) takes the form

$$\begin{bmatrix} \dot{\tilde{x}} \\ \dot{\hat{x}} \\ \dot{\tilde{x}} \\ 0 \end{bmatrix} = \overbrace{\begin{bmatrix} F_{11}(\theta) & F_{12}(\theta) & 0 & 0 \\ F_{21}(\theta) & F_{22}(\theta) & 0 & 0 \\ 0 & 0 & F_{11}(\theta) & F_{12}(\theta) \\ 0 & 0 & F_{21}(\theta) & F_{22}(\theta) \end{bmatrix}}^{F_e} \overbrace{\begin{bmatrix} \tilde{x} \\ \hat{x} \\ \tilde{x} \\ \hat{x} \end{bmatrix}}^{e} \tag{14}$$

$$z = \overbrace{\begin{bmatrix} G_1 & G_2 & -G_1 & -G_2 \end{bmatrix}}^{G_e} e.$$

The objective of the reduction algorithm is then to minimise $\|z\|_2$, but this time without back-substituting variables, which would most probably disturb the affine structure within the F matrices. Theorem 1 needs to be modified accordingly in this case.

Theorem 2. *Consider system (14). Suppose that there exists a $\Psi \succeq 0$, $\Psi \in \mathbb{R}^{(n+\tilde{n})\times(n+\tilde{n})}$ such that for all $\theta \in \Theta$*

$$F_e^T \begin{bmatrix} \Psi & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} \Psi & 0 \\ 0 & 0 \end{bmatrix} F_e + \begin{bmatrix} 0 \\ 0 \\ \Lambda_1^T \\ \Lambda_2^T \end{bmatrix} \begin{bmatrix} 0 & 0 & F_{21}(\theta) & F_{22}(\theta) \end{bmatrix}$$

$$+ G_e^T G_e + \begin{bmatrix} 0 \\ 0 \\ F_{21}^T(\theta) \\ F_{22}^T(\theta) \end{bmatrix} \begin{bmatrix} 0 & 0 & \Lambda_1 & \Lambda_2 \end{bmatrix} \preceq 0,$$

where $\Lambda = [\Lambda_1^T, \Lambda_2^T]^T$ with $\Lambda_1 \in \mathbb{R}^{(n-\tilde{n})\times\tilde{n}}$ and $\Lambda_2 \in \mathbb{R}^{(n-\tilde{n})\times(n-\tilde{n})}$ act as Lagrange multipliers. Partition Ψ as in Theorem 1 into

$$\Psi = \begin{bmatrix} \Psi_{11} & \Psi_{12} & \Psi_{13} \\ \Psi_{12}^T & \Psi_{22} & \Psi_{23} \\ \Psi_{13}^T & \Psi_{23}^T & \Psi_{33} \end{bmatrix}.$$

Then $\|y - \tilde{y}\|_2^2 \le x_0^T Q x_0$, where

$$Q = \begin{bmatrix} \Psi_{11} + \Psi_{13} + \Psi_{13}^T + \Psi_{33} & \Psi_{12} + \Psi_{23}^T \\ \Psi_{12}^T + \Psi_{23} & \Psi_{22} \end{bmatrix},$$

and the worst case error results from the initial condition x_0 aligned with the eigenvector direction corresponding to the largest eigenvalue of Q, $\bar{\lambda}_Q$. In fact we have

$$\max_{\|x_0\|_2=1} \|y - \tilde{y}\|_2 \le \sqrt{\bar{\lambda}_Q}.$$

Proof. This follows directly from the proof of Theorem 1, while the equality constraint in the descriptor system is adjoined on the grammian computation using free Lagrange multipliers Λ_1 and Λ_2 of appropriate dimension using Finsler's lemma (de Oliveira & Skelton, 2001; Skelton, Iwasaki, & Grigoriadis, 1998). □

The advantage of Theorem 2 is that it retains an LMI computation which is affine in the uncertain parameters, and hence its validity in the interior of the domain Θ can be established by ensuring that the LMI is feasible on the corners of Θ (Gahinet, Apkarian, & Chilali, 1996). An illustrative example is presented in Section 5.

4. Model decomposition

The aim of system decomposition is to break complex, large-scale models of signalling pathways into a set of interacting lower order subsystems. The motivation for such a decomposition is to

gain an understanding of the complete pathway in a systematic component-wise manner based on the system dynamics as opposed to traditional methods that only use topological information. To try and infer the functionality of the complete pathway by directly examining the full system is very difficult due to the coupling present in the system equations. A decomposition algorithm is proposed which automatically partitions the system into disjoint sets of weakly interacting subsystems by minimising the 'energy' flow between them. The motivation for such an approach is that states that interact strongly with each other are more likely to be part of a *functional module* than those that do not. For systems that do not have any obvious modular structure such an approach will facilitate analysis, while for systems that are suspected of having modular structure the decomposition will at worst provide a first guess solution to the underlying modularity.

Community detection algorithms have gained popularity recently for their ability to determine modularity in networks (Newman, 2006). However, these methods neglect the system dynamics in favour of understanding the properties of the connection topology. Related work in Del Vecchio, Ninfa, and Sontag (2007) and Saez-Rodriguez, Gayer, Ginkel, and Gilles (2008) focusses on looking at how the dynamics on a signalling network graph can be used to define functional subunits of a complex system. Arcak and Sontag (2006) derive stability conditions for networks that can be formulated as cyclic cascade networks using diagonal stability and passivity arguments. In Buzi, Topcu, and Doyle (2010) a manual decomposition of an autocatalytic network is used to separate the *complicating nonlinear dynamics* from the simple system dynamics. The above methods rely on an assumed structure of the network or obvious system partitions. In general, such assumptions may prove restrictive for some networks and as the state dimension increases, determining where to partition a system is likely to become difficult. In contrast, the method presented here can be applied to a more general class of nonlinear systems and requires no prior knowledge of the system.

Formally the decomposition method partitions (and reorders) the state vector so that $x = \left[x_1^T, x_2^T\right]^T$ where $x_1 \in \mathbb{R}^{n_1}, x_2 \in \mathbb{R}^{n_2}$ and $n_1 + n_2 = n$. System (1) with parameters fixed, can then be written as

$$\dot{x}_1 = f_1(x_1) + g_1(x_1, u_1), \qquad y_1 = x_1 \tag{15a}$$

$$\dot{x}_2 = f_2(x_2) + g_2(x_2, u_2), \qquad y_2 = x_2 \tag{15b}$$

with the system connected so as $y_1 = u_2$ and vice versa. The functions $g_i(x_i, 0) = 0$ represent the sub-system interactions. For the sake of clarity we restrict our attention to the case where the system is partitioned into two subsystems with no shared states. The extension to multiple partitions is straightforward and illustrated in Section 5.

4.1. Algebraic graph theory

A directed graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ consists of a set of vertices $\mathcal{V} = \{v_i\},\ i = 1, \ldots, N$ and a set of edges (links) $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$. If $v_i, v_j \in \mathcal{V}$ and $e_{ij} = (v_i, v_j) \in \mathcal{E}$, then there is an edge (a directed arrow) from node v_i to node v_j.

The *weighted adjacency matrix*, $A \in \mathbb{R}^{N\times N}$, is such that $A_{ij} > 0$ if there is an edge going from v_i to v_j, otherwise $A_{ij} = 0$. For weighted directed graphs define the out-degree matrix $D_{\text{out}} = \text{diag}\{A\mathbf{1}\}$, where $\mathbf{1}$ is the vector of ones. The *Laplacian matrix* of $\mathcal{G}$ is given by $L = D_{\text{out}} - A$. The Laplacian satisfies $L\mathbf{1} = 0$ and for undirected graphs is positive semidefinite.

For a graph with N vertices and M edges, the incidence matrix $C \in \mathbb{R}^{N\times M}$ is given by $C_{ij} = -1$ if edge j exits v_i, 1 if edge j enters v_i and 0 otherwise. The reader is directed to Godsil and Royle (2000) for more details regarding algebraic graph theory.

4.1.1. Spectral graph partitioning

Given a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ the partitioning problem requires one to construct K subgraphs, $\mathcal{G}_k(\mathcal{V}_k, \mathcal{E}_k),\ k = 1, \ldots, K$ such that $\bigcup_{k=1}^{K} \mathcal{V}_k = \mathcal{V}$ and $\mathcal{V}_k \bigcap \mathcal{V}_l = \emptyset$ for all k, l and $\mathcal{E}_k = \{(v_i, v_j) \in \mathcal{E} \mid v_i, v_j \in \mathcal{V}_k\}$. Our aim is to minimise the sum of the edge weights connecting subgraphs, while each partition should contain an approximately equal number of vertices. Such problems are known to be $\mathcal{NP}$-hard.

Consider the problem of partitioning a weighted, undirected graph into two subsets. Define a vector $z \in \mathbb{R}^N$ such that $z_i = \pm 1$ determines which partition $v_i \in \mathcal{V}$ belongs to. Then an approximate solution can be obtained by aligning z with the eigenvector corresponding to the smallest non-zero eigenvalue of the weighted Laplacian matrix for this graph (Karypis & Kumar, 1998), called the Fiedler eigenvector. To obtain multiple partitions the algorithm is recursively called. In some instances it may not be desirable to search for partitions of an equal size, in which case the partition should be made at the point where the largest change in magnitude between elements of the sorted eigenvector occurs. An SDP relaxation of this problem is described in Goemans and Williamson (1995).

4.2. System decomposition algorithm

The decomposition algorithm takes a two step approach: (i) the dynamical system is represented as a graph, (ii) the graph is partitioned using the spectral methods described in the previous section.[1] We note that in Šiljak (1978) the notion of decomposing large scale dynamical systems and a graphical representation of such systems is presented, however no method for finding an appropriate decomposition is given.

First, a weighted graph is constructed where vertices denote states and the weighted edges represent the 'energy' flow between states. A naive approach for decomposing the system would be to perform a Jacobian linearisation of (1) around $x = x^*$, $F \triangleq \frac{\partial f}{\partial x}\Big|_{x=x^*}$, and use the system matrix, F, as a weighted graph adjacency matrix to construct the Laplacian and hence apply a partitioning algorithm. However, this is a poor choice of graph weighting as the size of the elements in F does not give a good indication of the strength of state interactions.

We formulate the dynamical system decomposition problem as an optimisation programme which seeks to find a state partition which minimises the worst-case (maximum) amount of energy flowing between states in different partitions. The 'energy' we refer to is the $\mathcal{L}_2$ norm of the output signal.[2] The following min–max optimisation problem captures these requirements:

$$\begin{aligned} \min_{z}\ \max_{x_0}\quad & \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\left(1 - z_i z_j\right) A_{ij} \int_0^T \left(x_i(\tau) - x_j(\tau)\right)^2 \mathrm{d}\tau \\ \text{s.t.}\quad & \dot{x}(t) = Fx(t), \quad x(0) = x_0 \\ & \|x_0\|_2^2 = 1 \\ & z_i^2 = 1 \quad i = 1, \ldots, n \\ & z^T\mathbf{1} \neq \pm n \end{aligned} \tag{16}$$

where z is the partitioning vector as described in Section 4.1.1 and A is a binary adjacency matrix corresponding to the graph of the dynamical system given by

$$A_{ij} = \begin{cases} 1 & \text{if } F_{ji} \neq 0 \\ 0 & \text{if } i = j \text{ or } F_{ij} = 0. \end{cases} \tag{17}$$

[1] Alternative partitioning methods may be used, however in this work spectral partitioning was chosen.

[2] Note that the $\mathcal{L}_2$ energy of a signal is not equivalent to potential energy of a physical system (Willems, 2007).

Table 2
Energy based system decomposition algorithm.

1	*Linearize (1) to obtain F and A*
2	*Construct the incidence matrix C from F*
3	*Compute the energy matrix W*
4	*Symmetrise W using* $W \leftarrow \frac{1}{2}(W + W^T)$
5	*Compute* $L = \text{diag}\{W\mathbf{1}\} - W$
6	*Apply spectral partitioning algorithm to L*
7	*Use state partition obtained to construct (15)*

Then, problem (16) has the interpretation of finding the initial condition, x_0 and the state partition which minimises the worst-case norm of the difference between states ('energy flow') in different partitions up to time T.

Solving (16) directly is a very difficult task as it is a non-convex problem and in addition the partitioning element is $\mathcal{NP}$-hard. We approach this problem by first finding the worst-case energy flow on the graph (and the corresponding initial condition) and then the partition that minimises this worst-case energy flow.

To find the worst-case energy flow on the graph for $T = \infty$, we append an output map to produce the dynamical system

$$\dot{x} = Fx, \qquad y = C^T x, \qquad x(0) = x_0 \tag{18}$$

where C is the graph incidence matrix corresponding to the graph with Adjacency matrix (17). The output energy from a given initial condition is given by $\|y\|_2^2 \triangleq \|C^T x\|_2^2 = x_0^T P x_0$ where $P \succeq 0$ solves $F^T P + PF + CC^T = 0$. The unit-norm initial condition which will produce the most output energy is in the direction of the eigenvector $\bar{x}_0$ corresponding to the largest eigenvalue of P, $\bar{\lambda}_P$. We now produce a weighted adjacency matrix indicating the energy flow between nodes:

$$W_{ij} \triangleq \bar{x}_0^T X^{(ij)} \bar{x}_0 \tag{19}$$

where $X^{(ij)} \succeq 0$ solves the Lyapunov equation

$$F^T X^{(ij)} + X^{(ij)} F = -\delta_k \delta_k^T \tag{20}$$

and δ_k is the kth column of C corresponding to edge (i, j). The interpretation of (20) is that the system is released from the initial condition that maximises the energy observed at the output of the system, in this case the output is just a single edge connecting the two states x_i and x_j. Using this method we have approximated the integral term in (16) by a weighted adjacency matrix W, which is then symmetrised.

An approximate solution to (16) is now achieved by solving

$$\min_z \quad \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \left(1 - z_i z_j\right) W_{ij}. \tag{21}$$

The solution to (21) can be obtained by deriving the Laplacian matrix from the energy matrix W and applying the spectral partitioning algorithm. Once the partitioning has been carried out the nonlinear subsystems in (15) can be constructed. The algorithm is summarised in Table 2.

Note that the linearisation of (1) is only used to obtain the state partition; once this is obtained, the subsystem dynamics can inherit the nonlinearities in the original system, as in (15). Also, when it is desirable to decompose the system into multiple partitions the spectral partitioning algorithm should be called recursively on each sub-graph. In the next section an extension to the method is proposed that allows for decomposition in the presence of parametric uncertainty.

4.2.1. Robust decomposition

Here the decomposition algorithm is extended to take uncertainty into account. The approach taken is to represent the system in LPV form as was done in Section 3.2. It is assumed that all parameters can be represented by (5), i.e. upper and lower bounds are known. Thus the system to decompose is given by

$$\dot{x}(t) = F(\theta)x(t), \qquad x(0) = x_0 \tag{22}$$

where $F(\theta)$ is in the form of (4).

Using the same framework as in the previous section the idea is to decompose (22) into

$$\begin{aligned} \dot{x}_1 &= F_{11}(\theta_1)x_1 + F_{12}(\theta_1)u_1, \qquad y_1 = x_1 \\ \dot{x}_2 &= F_{21}(\theta_2)x_1 + F_{22}(\theta_2)u_2, \qquad y_2 = x_2 \\ \theta_1 &\in \Theta_1, \quad \theta_2 \in \Theta_2, \quad \Theta_1, \Theta_2 \subseteq \Theta \end{aligned} \tag{23}$$

where $y_1 = u_2$ and $y_2 = u_1$ and the matrices $F_{ij}(\theta_k)$ are affine in θ_k. Previously the objective was to partition the states to minimise the energy flow between subsystems. In the LPV case it is also desirable to partition the system's parameters. The notable difference between parameter and state partitioning is that the parameter partitions do not necessarily have to be disjoint, in fact such a partition may be impossible. It would be desirable however to minimise the number of parameters in the intersection of the two sets.

The first step of the robust algorithm is to construct an Energy matrix, W, and graph, $\mathcal{G}_0$, for the nominal system given by:

$$\begin{aligned} \dot{x}(t) &= F_0 x(t), \qquad x(0) = x_0 \\ y(t) &= C^T x(t) \end{aligned} \tag{24}$$

where C is the incidence matrix of $\mathcal{G}_0$. This is equivalent to Steps 2–3 of the decomposition algorithm in Table 2 as F_0 is not parameter dependent. At this point, W can be normalised. In the uncertain case there are now k matrices that describe how the uncertain parameters affect the nominal system (24). Typically the matrices F_i are sparse (in comparison to F_0) indicating that not all parameters affect every element in the nominal system's matrix. The motivation is now to include parameter partitioning with the state partitioning algorithm. To do so a method for combining the information from the matrices F_i, $i = 1, \ldots, k$ needs to be incorporated into the graph $\mathcal{G}_0$.

A graph $\mathcal{G}_i$ with associated binary adjacency matrices A_i for each matrix F_i is created in the following manner:

(1) Represent each state as a vertex.
(2) Construct a binary adjacency matrix (not necessarily symmetric) such that $A_{jk}^{(i)} = 1$ if $F_{kj}^{(i)} \neq 0$, where $A_{jk}^{(i)}$ refers to element jk of A_i.
(3) Define a graph $\mathcal{G}_i$ with adjacency matrix A_i.

The total state-parameter dependence for (22) is then captured by the graph $\mathcal{G}_\theta$ constructed from the adjacency matrix

$$W_\theta = W + \gamma A_\theta \tag{25}$$

where $\gamma > 0$ is a tuning parameter and

$$A_\theta = \frac{1}{k} \sum_{i=1}^{k} A_i. \tag{26}$$

The parameter γ is used to weigh the uncertainty dependence versus energy flow in the decomposition to make sure they are of the same magnitude so that the partitioning algorithm is acting on a relevant graph. By setting $\gamma = 0$ the decomposition algorithm in Table 2 is recovered. Once W_θ has been constructed the algorithm follows on from Step 4 in Table 2 where the energy matrix W is replaced with W_θ.

5. Examples

In this section the model reduction and decomposition algorithms are applied on the Epidermal Growth Factor (EGF)

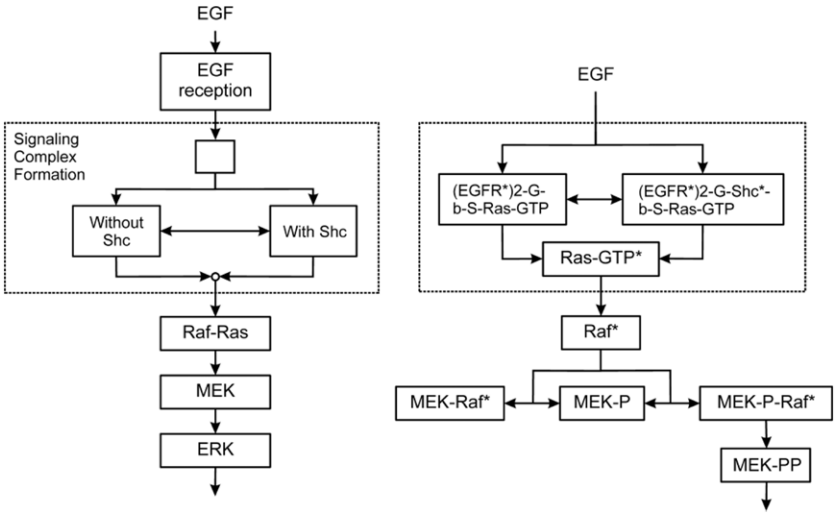

Fig. 1. Reduction of the EGFR-MAPK pathway model. Comparison between the structures of the published reduced order model using prior knowledge (left, reproduced from Saez-Rodriguez et al. (2004)) and the 8-state reduced model generated using the reduction algorithm proposed here (right).

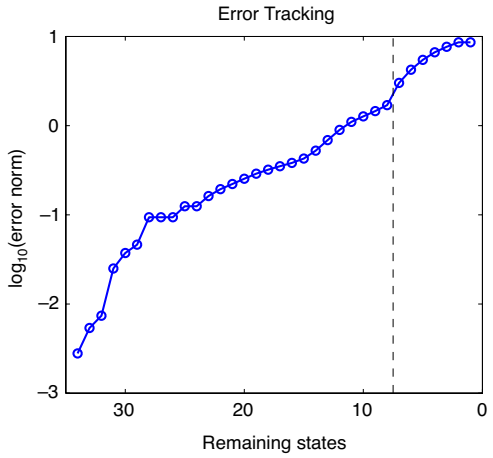

Fig. 2. The worst case error norm incurred as species are systematically removed according to the model reduction algorithm. The dashed line indicates the point at which we stop removing states.

receptor pathway. The model we consider describes the complete signalling cascade of the EGF receptor-induced MAP kinase pathway (Schoeberl, Eichler-Jonsson, Gilles, & Müller, 2002). The EGFR-MAPK pathway is probably the most studied signal transduction pathway in mammalian cells and is responsible for cell proliferation, differentiation and migration. Work in Saez-Rodriguez, Kremling, Conzelmann, Bettenbrock, and Gilles (2004) identifies the following key processes taking place in the pathway: (i) The EGFs bind to the Epidermal Growth Factor Receptors (EGFR) and activate them with every two receptors becoming a dimer. (ii) The cofactors GAP, Shc (optional), Grb2 and Sos bind to the dimer to form a complex. (iii) The complex then activates nearby RAS-GTP on the cell membrane. Raf binds to the activated RAS-GTP and then drives phosphorylation of MEK followed by ERK. (iv) Finally, phosphorylated ERK goes into the nucleus and drives the downstream gene expression.

The original model for this system was presented in Schoeberl et al. (2002). This was a 94-state (species) system of nonlinear differential equations containing over 100 kinetic parameters. Papers (Conzelmann et al., 2004; Saez-Rodriguez et al., 2004) made an assumption to ignore receptor internalisation in the original model. The model without receptor internalisation was used as the starting model for reduction in Conzelmann et al. (2004) and Saez-Rodriguez et al. (2004). This reduced model is shown in Fig. 1 (left). In our study we make the same assumption as in Conzelmann et al. (2004) and Saez-Rodriguez et al. (2004) and obtain a 35-state nonlinear model as the starting model for reduction.

5.1. Model reduction of the EGFR-MAPK pathway

After linearising the 35-state nonlinear model around the equilibrium point of interest the reduction algorithm was applied, giving a list of worst case error norms incurred as each state is removed from the system, as shown in Fig. 2. The states are removed from the system according to the order determined by the list. Removing a state from the system corresponds to removing a species from the reaction graph (network). When a state is removed, the original edges connected to the removed node are deleted and new directed edges that connect all predecessors with successors are generated—these correspond to the algebraic relationship that now relates them from the reduction algorithm. Applying this methodology we obtain a model containing 8 states as shown in Fig. 1(right); the dashed line in Fig. 2 shows the change in error norm at the point we cut the system.

Comparing these results with the previously published reduced order model (Saez-Rodriguez et al., 2004) which used the concept of retroactivity and simulations to produce it, the similarities are instantly recognisable. Our reduced order model and that obtained in Saez-Rodriguez et al. (2004) preserve all essential modular species and additionally our model keeps the fundamental structure of the MAPK cascades for MEK. In Saez-Rodriguez et al. (2004), the MEK module is reduced to a system of integrators and represented in a single block. Such a reduction can approximate the dynamic response of the MEK module in a linear region, but our results indicate that removing more states in the MAPK cascade will result in an increase in the error norm. We note however, that the reduced model from Saez-Rodriguez et al. (2004) preserves the predicted input–output behaviour of each of the subsystems as each module is derived using detailed analysis and simulations. In comparison our approach maintains the structure of the network, retains the steady states, keeps track of the error introduced in the reduction, but does not maintain the fast subsystem transient dynamics.

Next an LPV model for the system was developed with 3 uncertain parameters which were allowed to vary by 5%—these were the parameters to which the system was most sensitive to as determined by a sensitivity analysis using the methods from

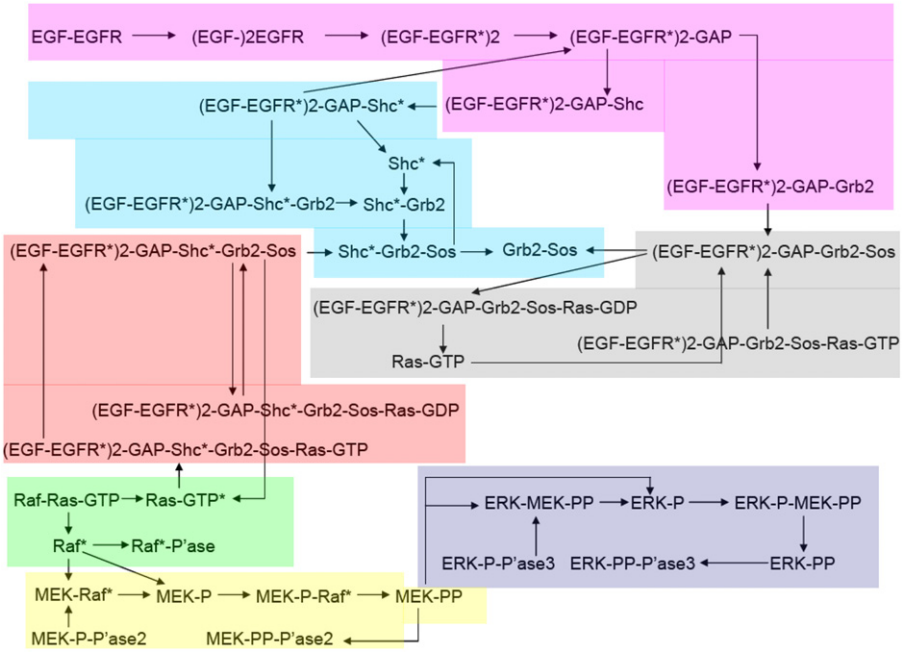

Fig. 3. Decomposition of the EGFR-MAPK pathway model into 7 functional blocks.

Varga et al. (1998). It was also observed from simulation that for larger perturbations the model became unstable. The algorithm from Section 3 was used to obtain the worst case error norm in the uncertain case for the reduced system with 8 states. The bound on the error with no uncertainty was found to be 3.53, while with uncertainty this was 9.61. The algorithm for the uncertain case suffered from numerical error, given the size of the resulting LMI. This advocates for the need of decomposition algorithms to support model reduction approaches.

5.2. Decomposition of the EGFR-MAPK pathway

Here we try to find a modular decomposition of the EGFR-MAPK pathway. The objective is to obtain modular subunits through the decomposition algorithm in Table 2 based on energy flow, that have biological significance. As a benchmark we compare our results to those obtained in Conzelmann et al. (2004). The algorithm was recursively called until the pathway was broken down into 7 subsystems as shown in Fig. 3.

The resulting decomposition is in strong agreement with the modularisation found using retroactivity based methods (Saez-Rodriguez et al., 2008) and analytic methods (Conzelmann et al., 2004). In particular the MAPK cascade is identified and the MEK and ERK blocks are identified and isolated (bottom of Fig. 3). The Raf and RAS activation cycles are identified (middle of Fig. 3) and the EGF binding process is found (top left Fig. 3).

The decomposition of the uncertain (LPV) model was found to be consistent with that of the nominal model. The decomposition shown was achieved initially be selecting the weighting term $\lambda = 1$. However, the same decomposition was achieved for values of $\lambda \in [0.8, 1.95]$.

5.3. Analysis via decomposition

Assessing properties such as stability, robustness and performance of signalling pathways described by coupled nonlinear differential equations such as (1) is a challenging task even when considering pathways of modest sizes. Recent approaches based on Sum of Squares (SOS) programming allow for the computational analysis of nonlinear systems whose dynamics are described by polynomial vector fields (Papachristodoulou & Prajna, 2002); biochemical networks with mass action kinetics fall in this category, but SOS-based methods do not scale well with system size. In related work we have described a set of methods that use the decomposition algorithm and SOS programming to provide subsystem stability certificates which can be integrated allowing conclusions to be made about the original large-scale system (Anderson & Papachristodoulou, 2010a,b). A brief outline of how the system decomposition fits in with large-scale dynamical system analysis is given below.

Assume the state vector x has been appropriately partitioned using the decomposition algorithm, the result is m subsystems connected as in (15). The nature of the decomposition facilitates the construction of subsystem Storage/Lyapunov functions of the form $V_i(x_i) > 0$, for all $x \neq 0$, with $V_i(0) = 0$, so that a Lyapunov function of the form

$$V(x) = \sum_{i=1}^{m} \alpha_i V_i(x_i), \quad \alpha_i > 0 \tag{27}$$

exists. This construction can be formulated as an SOS programme and due to the structure of V it is possible to solve this for systems with higher state dimensions than with a direct method.

The Lyapunov function (27) is referred to as a composite Lyapunov function (Khalil, 2001). Previous work has focussed on how the stability of interacting systems can be inferred from Lyapunov functions such as (27); see Kaszkurewicz and Bhaya (1993) and Moylan and Hill (1978). What sets the approach taken here apart from previous work is that we do not assume that the system exists in a pre-decomposed form.

5.3.1. Computational considerations

Assessing the stability of LPV systems via LMI feasibility tests was first introduced in Gahinet et al. (1996) where affine parameter dependent Lyapunov functions were investigated. It was noted that the number of LMI constraints grows exponentially with the number of parameters, k, of the system. Specifically for time invariant systems there are $2^k + k$ LMI constraints to satisfy. Thus, there are two main benefits of using a decomposition based approach to analyse (22): (i) the size of each of the $2^k + k$ LMI

constraints will potentially be halved for every partition; (ii) if a partition can be found such that the parameters appear only locally, i.e. individual parameters only affect states in one partition, then the number of LMI constraints will exponentially decrease.

6. Conclusion

In this paper a set of computational tools have been developed that facilitate the understanding of the properties of large-scale, complex biochemical signal transduction pathways. We presented a novel model reduction algorithm that keeps track of the worst-case $\mathcal{L}_2$ norm of the output error between the reduced and original system model as states are collapsed. This algorithm is particularly appropriate to biological models as the states of the lower order model are a strict subset of the original model. This property allows for the functionality of the original model to be inferred from the reduced order model. For the example considered, our results are in strong agreement with those obtained by methods that exploit prior knowledge on the system. The functionality of the reduction algorithm is then extended in order to handle uncertain models in the LPV framework.

A systematic method for decomposing signalling networks based on minimising the 'energy' flow between subsystems was also described. The algorithm was shown, through a numerical example, to find biologically significant subsystems. This was validated through comparison with previously published work and current biological knowledge. The decomposition scheme can be used on linear and nonlinear models; an extension to the LPV case was described. The method can be used to analyse complex systems, by analysing the subsystems thus defined first, and integrating the result at the end. This offers the computational advantage of analysing smaller subsystems rather than the full system and can be used to conclude stability, robust stability and robust performance of the full system.

Acknowledgements

The authors gratefully acknowledge Miss N. Inglessis for developing the EGFR-MAPK LPV models and Professor S. Duncan for many helpful discussions related to this work.

References

Anderson, J., & Papachristodoulou, A. (2010a). A network decomposition approach for efficient sum-of-squares programming based analysis. In *Proceedings of the American control conference. Baltimore, MD* (pp. 4492–4497).

Anderson, J., & Papachristodoulou, A. (2010b). Dynamical system decomposition for efficient, sparse analysis. In *Proceedings of the IEEE conference on decision and control. Atlanta, GA* (pp. 6565–6570).

Apkarian, P., & Tuan, H. D. (2000). Parameterized LMIs in control theory. *SIAM Journal on Control and Optimization*, *38*(4), 1241–1264.

Arcak, M., & Sontag, E. D. (2006). Diagonal stability of a class of cyclic systems and its connection with the secant criterion. *Automatica*, *42*, 1531–1537.

Bagci, E. Z., Vodovotz, Y., Billiart, T. R., & Bahar, I. (2006). Bistability in apoptosis: roles of Bax, Bcl-2, and mitochondrial permeability transition pores. *Biophysical Journal*, *90*, 1546–1559.

Beck, C. L., Doyle, J., & Glover, K. (1996). Model reduction of multidimensional and uncertain systems. *IEEE Transactions on Automatic Control*, *41*(10), 1466–1477.

Boyd, S., El Ghaoui, L., Feron, E., & Balakrishnan, V. (1994). *Linear matrix inequalities in system and control theory*. Philadelphia, PA, USA: Society for Industrial and Applied Mathematics.

Buzi, G., Topcu, U., & Doyle, J.C. (2010). Compositional analysis of autocatalytic networks in biology. In *Proceedings of American control conference. Baltimore, MD* (pp. 5929–5935).

Ciliberto, A., Capuani, F., & Tyson, J. J. (2007). Modeling networks of coupled enzymatic reactions using the total quasi-steady state approximation. *PLoS Computational Biology*, *3*(3), e45.

Conzelmann, H., Saez-Rodriguez, J., Sauter, T., Bullinger, E., Allgöwer, F., & Gilles, E. D. (2004). Reduction of mathematical models of signal transduction networks: simulation-based approach applied to EGF receptor signalling. *IEE Systems Biology*, *1*(1), 159–169.

de Oliveira, M. C., & Skelton, R. E. (2001). Stability tests for constrained linear systems. In S. O. Reza Moheimani (Ed.), *Perespectives in robust control* (pp. 241–257). Springer-Verlag.

Del Vecchio, D., Ninfa, A. J., & Sontag, E. D. (2007). Modular cell biology: retroactivity and insulation. *Molecular Systems Biology*, *4*(161).

Dullerud, G., & Paganini, F. (2000). *Texts in applied mathematics*: *Vol. 36. A course in robust control theory: a convex approach*. Springer.

Gahinet, P., Apkarian, P., & Chilali, M. (1996). Affine parameter-dependent Lyapunov functions and real parametric uncertainty. *IEEE Transactions on Automatic Control*, *41*(3), 436–442.

Gerdtzen, Z. P., Daoutidis, P., & Hu, W.-S. (2004). Non-linear reduction for kinetic models of metabolic reaction networks. *Metabolic Engineering*, *6*(2), 140–154.

Glover, K. (1984). All optimal hankel-norm approximations of linear multivariable systems and their l_∞-error bounds. *International Journal of Control*, *39*, 1115–1193.

Godsil, Chris, & Royle, Gordon (2000). *Algebraic graph theory*. Springer-Verlag, Inc..

Goemans, M. X., & Williamson, D. P. (1995). Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. *Journal of the ACM*, *42*, 1115–1145.

Herbst, Roy S. (2004). Review of epidermal growth factor receptor biology. *International Journal of Radiation Oncology Biology Physics*, *59*(Suppl. 1–2), S21–S26.

Hildebrandt, S., Raden, D., Petzold, L., Robinson, A. S., & Doyle, F. J. (2008). A top-down approach to mechanistic biological modeling: application to the single-chain antibody folding pathway. *Biophysical Journal*, *95*, 3535–3558.

Hornberg, J. J., Binder, B., Bruggeman, F. J., Schoeberl, B., Heinrich, R., & Westerhoff, H. V. (2005). Control of MAPK signalling: from complexity to what really matters. *Oncogene*, *24*, 5533–5542.

Huang, Z., Chu, Y., & Hahn, J. (2010). Derivation of simplified signal transduction pathway models: application to IL-6 signaling. In *Proceedings of American control conference. Baltimore, MD* (pp. 5131–5136).

Huang, H., Fairweather, M., Griffiths, J. F., Tomlin, A. S., & Brad, R. B. (2005). A systematic lumping approach for the reduction of comprehensive kinetic models. *Proceedings of the Combustion Institute*, *30*(1), 1309–1316.

Karypis, G., & Kumar, V. (1998). A fast and high quality multilevel scheme for partitioning irregular graphs. *SIAM Journal on Scientific Computing*, *20*(1), 359–392.

Kaszkurewicz, E., & Bhaya, A. (1993). Robust stability and diagonal Lyapunov functions. *SIAM Journal on Matrix Analysis and Applications*, *14*(2), 508–520.

Khalil, H. K. (2001). *Nonlinear systems* (third ed.) Prentice Hall, Inc.

Kim, D., Rath, O., Kolch, W., & Cho, K.-H. (2007). A hidden oncogenic positive feedback loop caused by crosstalk between Wnt and ERK pathways. *Oncogene*, *26*, 4571–4579.

Klipp, E., Liebermeister, W., Wierling, C., Kowland, A., Lehrach, H., & Herwig, R. (2009). *Systems biology: a textbook*. Wiley-VCH Verlag.

Kokotović, P., Khalil, H. K., & O'Reilly, J. (1999). *Singular perturbation methods in control: analysis and design*. Philadelphia, PA, USA: Society for Industrial and Applied Mathematics.

Kuo, J. C. W., & Wei, J. (1969). Lumping analysis in monomolecular reaction systems. Analysis of approximately lumpable system. *Industrial and Engineering Chemistry Fundamentals*, *8*, 124–133.

Lall, S., & Beck, C. (1999). Model reduction of complex systems in the linear-fractional framework. In *Computer aided control system design, 1999. Proceedings of the 1999 IEEE international symposium on* (pp. 34–39).

Lall, S., Marsden, J. E., & Glavaški, S. (2002). A subspace approach to balanced truncation for model reduction of nonlinear control systems. *International Journal of Robust and Nonlinear Control*, *12*(6), 519–535.

Lander, A. D. (2004). A calculus of purpose. *PLoS Biology*, *2*(6), e164.

Lee, J. M., Gianchandani, E. P., Eddy, J. A., & Papin, J. A. (2008). Dynamic analysis of integrated signaling, metabolic, and regulatory networks. *PLoS Computational Biology*, *4*(5), e1000086.

Lee, C. H., & Othmer, H. G. (2010). A multi-time-scale analysis of chemical reaction networks: I. Deterministic systems. *Journal of Mathematical Biology*, *60*(3), 387–450.

Lobo, M. S., Vandenberghe, L., Boyd, S., & Lebret, H. (1998). Applications of second-order cone programming. *Linear Algebra and its Applications*, *284*(1–3), 193–228.

Mannchen, T., & Well, K. H. (2002). Uncertainty bands approach to LFT modelling. In C. Fielding, A. Varga, S. Bennani, & M. Seiler (Eds.), *Lecture notes in control and information sciences*: *Vol. 283. Advanced techniques for clearance of flight control laws* (pp. 211–220). Springer.

Moore, B. (1981). Principal component analysis in linear systems: controllability, observability, and model reduction. *IEEE Transactions on Automatic Control*, *26*(1), 17–32.

Moylan, P. J., & Hill, D. J. (1978). Stability criteria of large-scale systems. *IEEE Transactions on Automatic Control*, *23*(2), 143–149.

Murray, J. D. (1993). *Mathematical biology*. second ed.: Springer-Verlag.

Naidu, D. S. (2002). Singular perturbations and time scales in control theory and applications: overview. In Z. Gajic (Ed.), *Dynamics of continuous, discrete, and impulsive systems*: *Vol. 9. Special issue on singularly perturbed dynamic systems in control technology* (pp. 233–278).

Newman, M. E. J. (2006). Modularity and community structure in networks. *Proceedings of the National Academy of Sciences*, *103*, 8577–8582.

Okino, M. S., & Mavrovouniotis, M. L. (1998). Simplification of mathematical models of chemical reaction systems. *Chemical Reviews*, *98*(2), 391–408.

Papachristodoulou, A., & Prajna, S. (2002). On the construction of Lyapunov functions using the sum of squares decomposition. In *Proceedings IEEE conference on decision and control* (pp. 3482–3487).

Saez-Rodriguez, J., Gayer, S., Ginkel, M., & Gilles, E. D. (2008). Automatic decomposition of kinetic models of signalling networks minimizing the retroactivity among modules. *Bioinformatics*, *24*, i213–i219.
Saez-Rodriguez, J., Kremling, A., Conzelmann, H., Bettenbrock, K., & Gilles, E. D. (2004). Modular analysis of signal transduction networks. *Control Systems Magazine—IEEE*, *24*(4), 35–52.
Saksena, V. R., O'Reilly, J., & Kokotovic, P. V. (1984). Singular perturbations and time-scale methods in control theory: survey 1976–1983. *Automatica*, *20*(3), 273–293.
Sandberg, H., & Murray, R. M. (2009). Model reduction of interconnected linear systems. *Optimal Control Applications and Methods*, *30*(3), 225–245.
Schoeberl, B., Eichler-Jonsson, C., Gilles, E. D., & Müller, G. (2002). Computational modeling of the dynamics of the MAP kinase cascade activated by surface and internalized EGF receptors. *Nature Biotechnology*, *20*, 370–375.
Shoemaker, J. E., & Doyle, F. J. (2008). Identifying fragilities in biochemical networks: Robust performance analysis of Fas signaling-induced apoptosis. *Biophysical Journal*, *95*, 2610–2623.
Šiljak, D. D. (1978). *Large-scale dynamic systems: stability and structure*. New York: North-Holland.
Singhand, A., Jayaraman, A., & Hahn, J. (2006). Modeling regulatory mechanisms in IL-6 signal transduction in hepatocytes. *Biotechnology and Bioengineering*, *95*(5), 850–862.
Skelton, R. E., Iwasaki, T., & Grigoriadis, K. (1998). A unified algebraic approach to linear control design. *Taylor & Francis*.
Sontag, E. D. (2004). Some new directions in control theory inspired by systems biology. *IET Systems Biology*, *1*(1), 9–18.
Tropp, J. A. (2004). Greed is good: algorithmic results for sparse approximation. *IEEE Transactions on Information Theory*, *50*(10), 2231–2242.
Vandenberghe, L., & Boyd, S. (1996). Semidefinite programming. *SIAM Review*, *38*(1), 49–95.
Varga, A., Moormannand, G., & Grubel, G. (1998). Automated generation of LFT–based parametric uncertainty descriptions from generic aircraft models. *Mathematical and Computer Modelling of Dynamical Systems*, *4*(4), 249–274.
Vilar, J. M. G., Jansen, R., & Sander, C. (2006). Signal processing in the TGF-β superfamily ligand-receptor network. *PLoS Computational Biology*, *2*(1), e3.
Waldherr, S., Allgöwer, F., & Jacobsen, E.W. (2009). Kinetic perturbations as robustness analysis tool for biochemical reaction networks. In *Proceedings of the 48th IEEE conference on decision and control* (pp. 4572–4577).
Willems, J. C. (2007). The behavioral approach to open and interconnected systems. *Control Systems Magazine—IEEE*, *27*(6), 46–99.
Wu., F. (1996). Induced L_2 norm model reduction of polytopic uncertain linear systems. *Automatica*, *32*(10), 1417–1426.
Zagaris, A., Kaper, H. G., & Kaper, T. J. (2004). Analysis of the computational singular perturbation reduction method for chemical kinetics. *Journal of Nonlinear Science*, *14*(1), 59–91.

James Anderson received a B.Sc. and M.Sc. degree from the School of Systems Engineering at the University of Reading, UK in 2005 and 2006 respectively. From 2006 to 07 he was a research assistant in the Department of Cybernetics at the University of Reading prior to beginning his D.Phil (Ph.D.) at the University of Oxford's Life Sciences Interface Doctoral Training Centre. He is currently a D.Phil student in the Control Group in the Department of Engineering Science. His research interests include nonlinear dynamical systems, robust control, large-scale networked systems and systems biology.

Yo-Cheng Chang received a B.S. degree in Biomedical Engineering from the Chung-Yuan Christian University, Taiwan, in 1997, the M.S. degree in Biomedical Engineering from the National Yang Ming University, Taiwan, in 1999, and the D.Phil. degree in Biochemistry from the University of Oxford, UK, in 2011. From 2000 to 2005 he was an Associate Researcher working in the Department of Genomics and the Department of Nanobiotechnology, Biomedical Engineering Center, Industrial Technology Research Institute, Taiwan. At the same time he was also an Adjunct Lecturer teaching bioinformatics and computational biology in the Ming Chuan University, Tamkang University, and the National Yang Ming University in Taiwan. He is currently doing postdoctoral research in the University of Oxford, UK. His major research interest is in applying control theory to design and analyze biological networks for systems and synthetic biology.

Antonis Papachristodoulou received an MA/MEng degree in Electrical and Information Sciences from the University of Cambridge in 2000, as a member of Robinson College. In 2005 he received a Ph.D. in Control and Dynamical Systems, with a minor in Aeronautics from the California Institute of Technology. In 2005 he held a David Crighton Fellowship at the University of Cambridge and a postdoctoral research associate position at the California Institute of Technology before joining the Department of Engineering Science at the University of Oxford, Oxford, UK in January 2006, where he is now a University Lecturer in Control Engineering and tutorial fellow at Worcester College. His research interests include scalable analysis of nonlinear systems using convex optimization based on Sum of Squares programming, analysis and design of large-scale networked control systems with communication constraints and Systems and Synthetic Biology.