Dimension reduction by balanced truncation applied to a model of glycolysis

Wolfram Liebermeister
lieberme@molgen.mpg.de
MPI for Molecular Genetics
Ihnestr. 73, D-14195 Berlin

Abstract

Balanced truncation is an established method for reducing large linear dynamical systems. Metabolic pathways are usually surrounded by a large metabolic network, the "environment". In order to simplify it, we employ linearisation and balanced truncation and construct a small black-box model that mimicks its behaviour - as seen by the subsystem of interest. To illustrate this procedure, we apply it to a mathematical model of glycolysis in yeast.

1 Introduction

As mathematical models in cell biology grow to larger sizes, complexity reduction becomes an important issue: models - or parts of them - are replaced by effective models that are easier to solve and that are supposed to reproduce, with sufficient precision, the results of the original model. In the context of control theory, the reduction of linear models has been studied for a long time, and various methods have been proposed [1]. One of them is balanced truncation [2], which is numerically demanding, but yields a stable reduced system with a bounded approximation error. Moreover, by tuning the dimensionality, one can choose a compromise between approximation accuracy and numerical efficiency. Balanced truncation has successfully been applied to linear control systems of high state-space dimensions (see [3] and the examples therein). Importantly, balanced truncation does not rely on a

time scale separation of fast and slow processes (see, for instance, [4] [5]), but removes parts of the dynamics that contribute little to the input-output relation of the system.

Here we apply balanced truncation to metabolic models. As metabolic systems are usually nonlinear, and as linearising may change their behaviour considerably, we restrict ourselves to a special application: a small system of interest (a pathway) is embedded in a larger environment (a metabolic network). Usually, this environment is described by fixed concentrations. In order to construct a model that is more accurate, but still easy to solve, we start with a model of the entire system and use model reduction to obtain a dynamic black-box model of the environment. This effective model can provide the system of interest with dynamical boundary conditions, and the reduced variables describe the dynamical modes in the environment that dominate its interaction with the subsystem.

This paper is a shortened version of the article [6], in which the mathematical details are treated at length. As an illustrative example, we added the application to the glycolysis model of Hynne et al. [7]. The author likes to thank U. Baur and E. Klipp who participated in the project. This work was funded by the European commission, grant 503269.

2 Model reduction by balanced truncation

2.1 Linear models

Let us first outline the idea of model reduction: we consider a linear dynamical system of the standard (vectorial) form

$$\begin{aligned} \dot{\mathbf{x}}(t) &= A\,\mathbf{x}(t) + B\,\mathbf{u}(t), \quad t > 0, \quad \mathbf{x}(0) = \mathbf{x}_0 \\ \mathbf{y}(t) &= C\,\mathbf{x}(t) + D\,\mathbf{u}(t), \quad t \geq 0. \end{aligned} \tag{1}$$

The system comprises n state variables x_i, which are controlled by m input variables u_k and can be observed via the p output variables y_l. We postulate that for $\mathbf{u} = 0$, the system has a steady state at $\mathbf{x} = 0$. For fixed initial conditions, any time course $\mathbf{u}(\cdot)$ of the controlling variables leads to a time course $\mathbf{y}(\cdot)$ of the observables.

The same input-output relation can be exactly represented by a system with transformed variables $\hat{\mathbf{x}}$. If T is an invertible $n \times n$ matrix, we can apply

the transformation

$$\begin{aligned} \mathbf{x} &\rightarrow \hat{\mathbf{x}} = T\mathbf{x} \\ A &\rightarrow \hat{A} = TAT^{-1} \\ B &\rightarrow \hat{B} = TB \\ C &\rightarrow \hat{C} = CT^{-1} \end{aligned}$$

without changing the input-output relation between $\mathbf{u}(\cdot)$ and $\mathbf{y}(\cdot)$.

2.2 Model reduction

In model reduction, we replace the system (1) by a lower-dimensional system of order r ($r \ll n$) that yields a good approximation of the input-output relation. For a chosen dimensionality r, we can split $T = \binom{T_1}{T_2}$, $T^{-1} = (S_1 S_2)$ with an $r \times n$ matrix T_1 and an $n \times r$ matrix S_1. The transformation

$$\begin{aligned} \mathbf{x} &\rightarrow \tilde{\mathbf{x}} = T_1\mathbf{x} \\ A &\rightarrow \tilde{A} = T_1 A S_1 \\ B &\rightarrow \tilde{B} = T_1 B \\ C &\rightarrow \tilde{C} = C S_1 \end{aligned} \tag{2}$$

yields a reduced model of dimension r

$$\begin{aligned} \dot{\tilde{\mathbf{x}}}(t) &= \tilde{A}\,\tilde{\mathbf{x}}(t) + \tilde{B}\,\mathbf{u}(t), \qquad t > 0, \quad \tilde{\mathbf{x}}(0) = \tilde{x}_0 \\ \tilde{\mathbf{y}}(t) &= \tilde{C}\,\tilde{\mathbf{x}}(t) + D\,\mathbf{u}(t), \qquad t \geq 0 \end{aligned} \tag{3}$$

that approximates the input-output relation. For a broad collection of survey papers on model reduction, see [1], where also a couple of benchmark examples are presented.

2.3 Balanced truncation

Closely connected with the stable, continuous-time system (1) are the two matrices $\mathcal{P}$ and $\mathcal{Q}$, the infinite reachability Gramian and the infinite observability Gramian:

$$\mathcal{P} := \int_0^\infty e^{At} B B^T e^{A^T t}\, \mathrm{d}t, \quad \mathcal{Q} := \int_0^\infty e^{A^T t} C^T C e^{At}\, \mathrm{d}t.$$

The Gramians can be computed by solving the Lyapunov equations

$$\begin{aligned} A\mathcal{P} + \mathcal{P}A^T + BB^T &= 0 \\ A^T\mathcal{Q} + \mathcal{Q}A + C^TC &= 0. \end{aligned} \tag{4}$$

Model reduction by balanced truncation[1] [2] is based on a special transformation into so-called balanced coordinates. The basic concept of balancing is finding a basis in which the two Gramians are equal and diagonal

$$\mathcal{P} = \mathcal{Q} = \mathrm{diag}(\sigma_1, \ldots, \sigma_n),$$

with ordered diagonal entries σ_i, the Hankel singular values of the system. In these new coordinates, states that are difficult to control are also difficult to observe and vice versa. Model reduction by balanced truncation removes these state components.

Model reduction by balanced truncation has some desirable properties: the reduced system (3) remains stable and has a low approximation error with an a priori known upper bound. Therefore, the size of the reduced system can be chosen adaptively depending on the permitted error size.

3 Selective reduction of biochemical models

Metabolic systems have to be linearised before reduction, which may in many cases produce unacceptable approximation errors. Still, balanced truncation can be useful for simplifying parts of a model that are (i) large, (ii) close to a stable steady state, and (iii) uninteresting, except for their input-output relation. This holds for the large parts of the cell that are usually neglected in modelling - because they only surround the metabolic pathways of interest.

Modelling of biochemical systems usually zooms in on certain pathways, while the concentrations of so-called external metabolites are considered fixed. This approximation ignores feedback loops mediated by the environment, that is, via external metabolites and reactions. To achieve a more realistic, dynamical description that is still numerically efficient, we describe the environment by a linear effective model of adjustable dimensionality. In particular, we (i) split the entire model into a subsystem and its environment,

[1] Matlab code for balanced truncation can be found at
`www-user.tu-chemnitz.de/~benner/software.php`

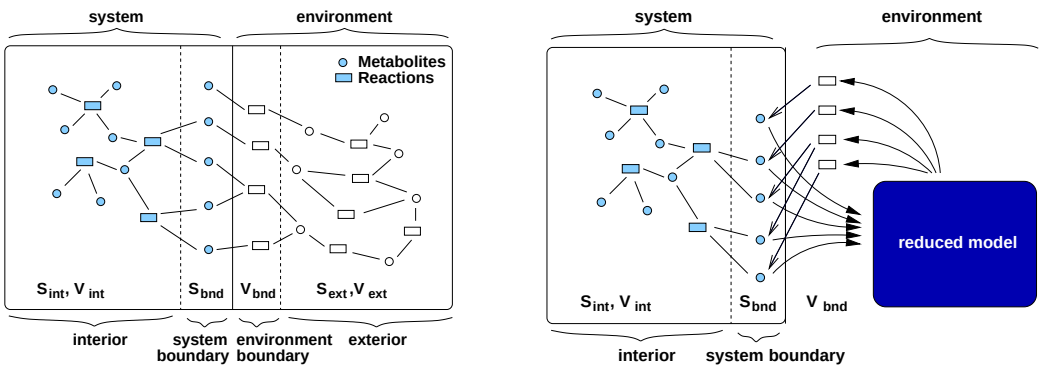

Figure 1: Schematic picture of selective model reduction. Left diagram: a biochemical network comprising metabolites (circles) and reactions (little boxes). The system is divided into a subsystem of interest (left half) and its environment (right half). Both systems are only connected by the communicating metabolites (boundary of the subsystem) and the communicating reactions (boundary of the environment). Right diagram: the environment part of the model has been replaced by an effective (black-box) model that senses the communicating metabolites and computes - approximately - the velocities of the communicating reactions.

(ii) linearise the environment model around a steady state, and (iii) reduce its dimensionality by balanced truncation.

Let us recall some basic definitions for biochemical network models [8] [9]: a metabolic system is described by the differential equation system

$$\dot{\mathbf{s}}(t) = N\mathbf{v}(\mathbf{s}(t), \mathbf{p}) \tag{5}$$

where $\mathbf{s}$ is the vector of metabolite concentrations and $\mathbf{v}$ is the vector of reaction velocities. The vector $\mathbf{p}$ contains the kinetic parameters, and the stoichiometric matrix N contains in its k^{th} column the stoichiometric coefficients for the k^{th} reaction.

We now assume that only a subsystem (e.g., a certain metabolic pathway) is in the focus of interest. As shown in Figure 1, the entire system can be split into subsystem and environment which influence each other only via certain communicating metabolites and reactions, which are located in the boundaries.

We aim at maintaining the subsystem in its original form while replacing the environment by a linear model of lower dimensionality. To obtain such

a reduced, effective model, we first linearise the environment model around some stable steady state (the "reference" state), and introduce the deviations

$$\begin{aligned} \mathbf{x} &= \Delta\mathbf{s}_{\text{ext}} = \mathbf{s}_{\text{ext}} - \bar{\mathbf{s}}_{\text{ext}} \\ \mathbf{u} &= \Delta\mathbf{s}_{\text{bnd}} = \mathbf{s}_{\text{bnd}} - \bar{\mathbf{s}}_{\text{bnd}} \\ \mathbf{y} &= \Delta\mathbf{v}_{\text{bnd}} = \mathbf{v}_{\text{bnd}} - \bar{\mathbf{v}}_{\text{bnd}}. \end{aligned} \tag{6}$$

The bars refer to the reference values, while sub- and superscripts "bnd" and "ext" indicate the boundaries and the the external region. If the model parameters depend on time, then their deviations $\Delta\mathbf{p}(t)$ can also be incorporated into $\mathbf{u}$. Altogether, we can rewrite the entire system as a coupled equation system for internal concentrations $\mathbf{s}$ and external deviations $\mathbf{x}$

$$\begin{aligned} \mathbf{u}(t) &= P\ \mathbf{s}(t) - \bar{\mathbf{s}}_{\text{bnd}} \\ \dot{\mathbf{x}}(t) &= A\ \mathbf{x}(t) + B\ \mathbf{u}(t) \\ \mathbf{y}(t) &= C\ \mathbf{x}(t) + D\ \mathbf{u}(t) \\ \mathbf{v}_{\text{bnd}}(t) &= \bar{\mathbf{v}}_{\text{bnd}} + \mathbf{y}(t) \\ \dot{\mathbf{s}}(t) &= N\ \mathbf{v}(\mathbf{s}(t), \mathbf{p}(t)) + N_{\text{bnd}}\ \mathbf{v}_{\text{bnd}}(t). \end{aligned} \tag{7}$$

The matrix P projects the subsystem metabolites to the boundary metabolites. The matrices A, B, C, D can be computed from the stoichiometric matrix and the reaction elasticities (see [6]). The equation system (7) consists of (i) a model of the subsystem with external fluxes $\mathbf{v}_{\text{bnd}}$ (last equation), (ii) a linear model of the standard form (1), describing the environment (second and third equation), and (iii) instructions on how to match both modules (first and fourth equation).

To reduce the environment part of the model (the equations for $\dot{\mathbf{x}}$ and $\mathbf{y}$), we replace $\mathbf{x}$ by a low-dimensional vector $\tilde{\mathbf{x}}$ as described in section 2.2, and the matrices A, B, C are transformed accordingly. The new, reduced variables do no longer represent individual metabolites, but global dynamical modes that are most strongly involved in the feedback from boundary metabolites to boundary reactions.

4 Example: glycolysis in yeast

We applied model reduction to the glycolysis model of Hynne et al. [7] (as published in the JWS online model database [10]), which describes the production and consumption of energy in a suspension of yeast cells. The

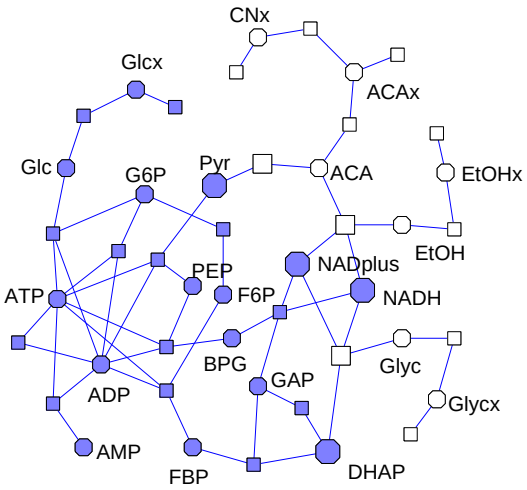

Figure 2: Metabolic model [7] of glycolysis in yeast. Metabolites and reactions are shown as circles and boxes, respectively. Model reduction is applied to the environment (white symbols), while the subsystem of interest (shaded symbols) is maintained in its original form.

model describes the concentrations of 17 metabolite inside the cells and 5 metabolites in the growth medium. The parameter values were determined in [7] at the onset of glycolytic oscillations. For a low Glucose concentration `Glcx0`=5.0 in the inflowing medium, the system has a stable steady state. In order to focus on the actual glycolysis pathway, we regard glycerol, ethanol, acetaldehyde, and external cyanide as external metabolites forming the environment. The metabolic network and its subdivision into subsystem and environment is shown in Figure 2.

Figure 3 shows results from dynamic simulations. Initially, the system is in its steady state, except for the external Glucose concentration, which assumes a higher concentration. Fixing the external concentrations changes the results considerably, while a reduced environment model with two variables yields a much better approximation.

Changes of the reduced variables correspond to changes of the external metabolites: the linear relation between them is given by the transformation weights contained in the matrix S_1. The columns of S_1 are shown in Figure 3: the large white circles (negative values) in the left panel show that x_1 mainly represents a decrease of intracellular acetaldehyde and ethanol. The

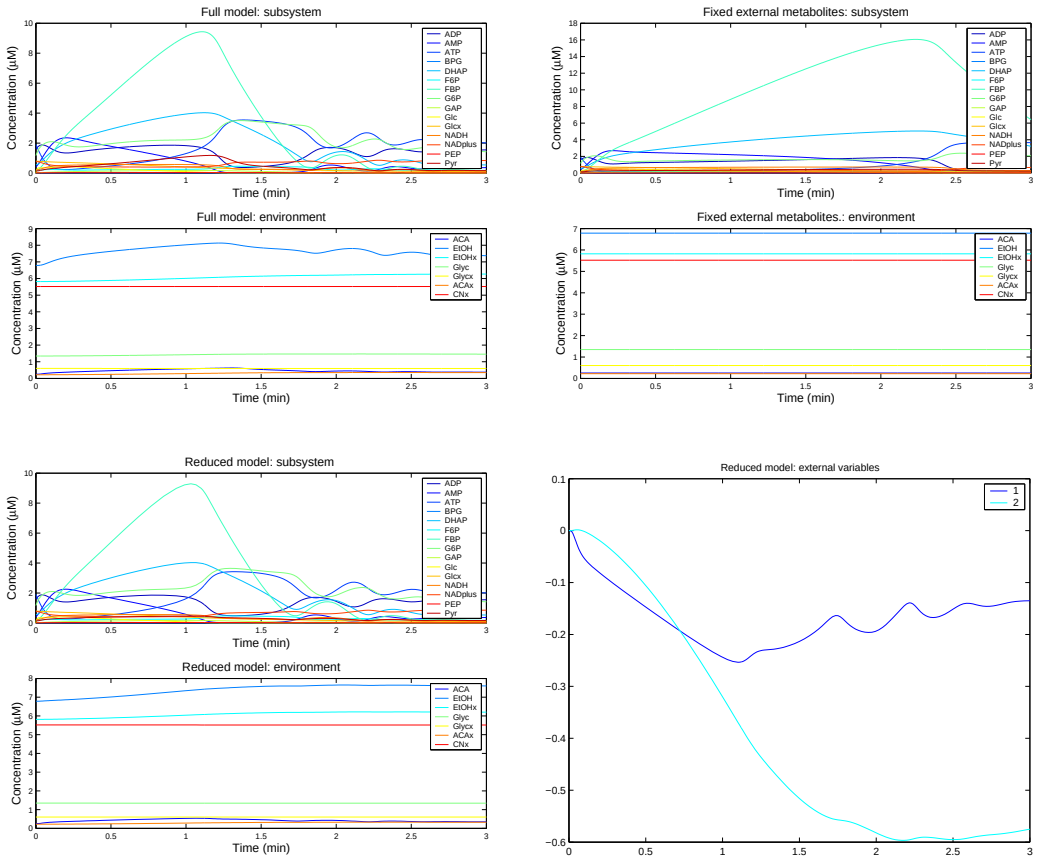

Figure 3: Simulation and reduction of the glycolysis model [7]. Top left: time courses of metabolite concentrations after a pulse of Glucose in the growth medium. The panels at top and bottom show the internal and external metabolites (as indicated in the network in Figure 2). Top right: Fixing the external concentrations at their steady-state values has a strong effect on the simulation results. Bottom left: Mimicking the environment by a two-dimensional effective model almost restores the simulation results of the original model. The external variables (bottom) are approximated by a linear combination of the reduced variables. Bottom right: time courses of the two reduced variables x_1 and x_2.

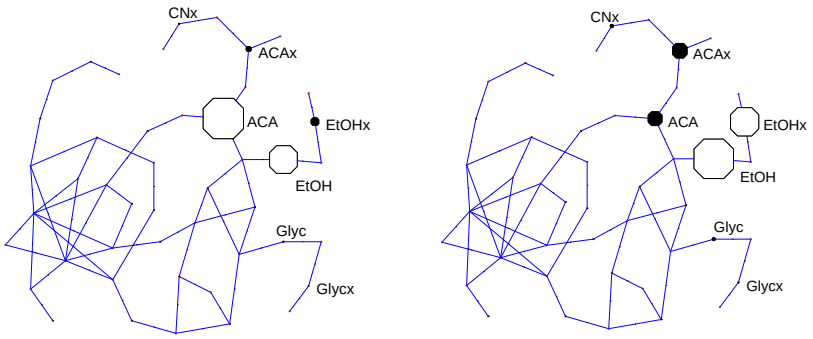

Figure 4: Transformation weights for the reduced variables. The change of a variable approximately corresponds to a change of the original external variables. The weights (elements of the matrix S_1 for the different external variables (compare Figure 2) are shown by circles: their size indicates the absolute value, black and white indicate positive and negative values, respectively.) Left: transformation weights for the variable x_1. Right: transformation weights for x_2.

variable x_2 mainly stands for an increase of acetaldehyde and a decrease of ethanol.

References

[1] P. Benner, V. Mehrmann, and D.C. Sorensen, editors. Lecture Notes in Computational Science and Engineering. Springer-Verlag, Berlin/Heidelberg, Germany, 2005. To appear.

[2] B.C. Moore. Principal component analysis in linear systems: Controllability, observability, and model reduction. *IEEETransAC*, AC-26:17–32, 1981.

[3] A.C. Antoulas and D.C. Sorensen. Approximation of large-scale dynamical systems: An overview. *Int. J. Appl. Math. Comp. Sci.*, 11(5):1093–1121, 2001.

[4] R. Krüger and R. Heinrich. Model reduction and analysis of robustness for the Wnt/β-Catenin signal transduction pathway. *Genome Informatics*, 15(1):138–148, 2004.

[5] J. Zobeley, D. Lebiedz, A. Ishmurzin, and Kummer U. Dynamic complexity reduction in nonlinear biochemical systems. *submitted to Nonlinear Dynamics*, 2003.

[6] W. Liebermeister, U. Baur, and E. Klipp. Biochemical network models simplified by balanced truncation. *to appear in FEBS Journal*, 2005.

[7] F. Hynne, S. Danø, and P.G. Sørensen. Full-scale model of glycolysis in Saccharomyces cerevisiae. *Biophys. Chem.*, 94:121–163, 2001.

[8] R. Heinrich and S. Schuster. *The regulation of cellular systems*. Chapman & Hall, 1996.

[9] J. S. Hofmeyer. Metabolic control analysis in a nutshell. In *ICSB 2001 Online Proceedings, http://www.icsb2001.org/toc.html*, 2001.

[10] B. Olivier and J. Snoep. Web-based kinetic modelling using JWS online. *Bioinformatics*, 20(13):2143–2144, 2004.