ELSEVIER

Annual Reviews in Control 29 (2005) 181–190

Annual Reviews in Control

www.elsevier.com/locate/arcontrol

An overview of approximation methods for large-scale dynamical systems☆

A.C. Antoulas *

Department of Electrical and Computer Engineering, Rice University, Houston, TX 77251, USA

Received 26 October 2004; received in revised form 7 July 2005; accepted 25 August 2005

Abstract

Methods for the approximation of large-scale dynamical systems will be surveyed. There are mainly two families namely, the SVD-based and Krylov-based approximation methods. The former family is based on the singular value decomposition and the second on moment matching. While the former has many desirable properties including an error bound, it cannot be applied to systems of high complexity. The strength of the latter on the other hand, is that it can be implemented iteratively and is thus appropriate for application to high complexity systems. An effort to combine the best attributes of these two families leads to a third class of approximation methods, which will be referred to as SVD/Krylov. Following a survey of these methods we will conclude with a new result concerning model reduction with preservation of passivity which is appropriate for application to large-scale circuits arising in VLSI chip performance verification.

Keywords: Large-scale systems; Model reduction; Approximation; Singular value decomposition; Balanced truncation; Krylov methods; Iterative approximation; Lanczos; Arnoldi; Interpolation; Realization; Passivity; Spectral zeros

1. Introduction

Starting point is a dynamical system which has to be simulated and/or eventually controlled. The first step in this endeavor consists of modeling, that is deriving equations describing its behavior. In our setting these will be assumed to be differential equations, either partial (PDEs) or ordinary (ODEs). To proceed with their solution, the PDEs are often discretized in space which leads to a large number of ODEs. Assume for simplicity that these consist of a set of $n$coupled first order ODEs. Model reduction consists in replacing them with k coupled first order ODEs where $k \ll n$; in addition the reduced set of ODEs must behave as closely as possible to the original one. A pictorial representation of how model reduction fits in the overall picture of simulation/control is shown in Fig. 1.

For details on the material presented in the sequel we refer to the book (Antoulas, 2005a).

1.1. Problem statement

We will consider dynamical systems described by (explicit) state equations

$$\boldsymbol{\Sigma}: \dot{x} = f(x,u), \quad y = h(x,u),$$

with state $x(\cdot)$of dimension n, input $u(\cdot)$ of dimension m, and output $y(\cdot)$ of dimension p, where $n \gg m, p$; we will use the notation $\boldsymbol{\Sigma} = (f,h)$. The approximation or model reduction problem can be formulated as follows.

Problem: Approximate $\boldsymbol{\Sigma} = (f,h)$ with $\hat{\boldsymbol{\Sigma}} = (\hat{f},\hat{h})$, $u(\cdot) \in \mathbb{R}^m, \hat{x}(\cdot) \in \mathbb{R}^k, \hat{y}(\cdot) \in \mathbb{R}^p$, where $k \ll n$, so that as many as possible of the properties below are satisfied:

(1) Approximation error small and existence of an error bound.
(2) Preservation of stability/passivity.
(3) Procedure must be computationally efficient.

1.2. Approximation by projection

The approximation methods to be discussed are obtained by means of *projections*. Let $V, W \in \mathbb{R}^{n\times k}$, be such that $W^*V = I_k$, where the superscript $(\cdot)^*$ denotes transposition. It follows that $\Pi = VW^*$ is a projection. Let $\hat{x} = W^*x \in \mathbb{R}^k$; the state x will

☆ This work was supported in part by the NSF through Grants CCR-0306503 and ACI-0325081.
* Fax: +1 713 348 5686.
E-mail address: aca@rice.edu.

1367-5788/$ – see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.arcontrol.2005.08.002

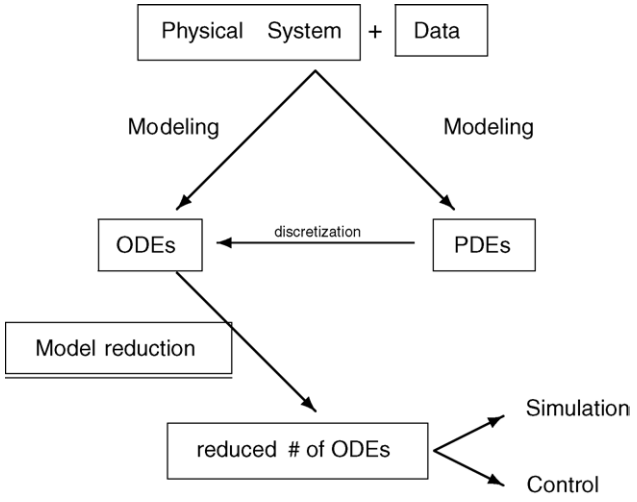

Fig. 1. Model reduction: the big picture.

now be approximated by its projection $\Pi x = V\hat{x}$. This leads to the following system which describes the evolution of $\hat{x}$:

$$\hat{\Sigma} : \dot{\hat{x}} = W^* f(V\hat{x}, u), \qquad \hat{y} = h(V\hat{x}, u).$$

Thus $\hat{\Sigma}$ is "good" approximation of Σ, if $x - \Pi x$ is "small" in some appropriate sense.

1.2.1. Special case: linear dynamical systems

In this case f and h are linear, i.e., $\dot{x} = Ax + Bu, y = Cx + Du$, and the system will be denoted by

$$\Sigma = \left(\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right) \in \mathbb{R}^{(n+p)\times(n+m)}.$$

With V, W as above, a $k < n$ dimensional reduced order model $\hat{\Sigma}$: $\dot{\hat{x}} = \hat{A}\hat{x} + \hat{B}u, \hat{y} = \hat{C}\hat{x} + Du$, of Σ is obtained as follows:

$$\hat{\Sigma} = \left(\begin{array}{c|c} \hat{A} & \hat{B} \\ \hline \hat{C} & D \end{array}\right) = \left(\begin{array}{c|c} W^*AV & W^*B \\ \hline CV & D \end{array}\right) \in \mathbb{R}^{(k+p)\times(k+m)}.$$

The closeness of $\hat{\Sigma}$ to Σ will be measured mostly in terms of the $\mathcal{H}_\infty$ and the $\mathcal{H}_2$ norms. The former is the 2-norm of the worst output error $\|y - \hat{y}\|_2$ for inputs $\|u\|_2 = 1$, while the latter is the 2-norm of the difference between the corresponding impulse responses $\|h - \hat{h}\|_2$.

2. Motivating examples

There is a great variety of examples which motivate the need for model reduction of large-scale systems, both for simulation and control. A partial list is given in Table 1.

In the sequel we will summarize a few of these applications; for a complete discussion we refer to the book (Antoulas, 2005a).

2.1. Passive devices: VLSI circuits

Evolution of VLSI design. The integrated circuit (IC) was invented in the 1960's. In 1971 the Intel 4004 processor was introduced. It had 2300 components of size approximately 10μm and an operating frequency of 64 KHz. Thirty years later, in 2001, the Intel Pentium IV was introduced, having 42 million components with size of the order of 0.18μm and 2 GHz operating frequency; as a result the *interconnect length*(length of all interconnections between the components) is about 2 km, and the chip has seven layers. In this case the interconnections must be modeled as transmission lines and simulations are required to verify that internal electromagnetic fields do not significantly delay or distort circuit signals. This leads to the need for electromagnetic modeling of interconnects (and packages). The resulting models are very complex; using PEEC (Partial Element Equivalent Circuit) methods to discretize Maxwell's equations in three-dimensions, we obtain systems of complexity $n \approx 10^5 \cdots 10^6$; thus model reduction methods are necessary to verify the performance of the underlying chip.

For details see e.g. van der Meijs (2000).

Table 1
Motivating examples of large-scale systems

1. *Passive devices*:
   VLSI circuits
2. *Weather prediction, data assimilation*:
   North sea forecast
   Air quality forecast
   America's cup
3. *Biological systems*:
   Honey comb vibrations
4. *Molecular systems*:
   Dynamics simulations
   Heat capacity
5. *ISS: International space station*:
   Stabilization
6. *Vibration/acoustic problems*:
   Windscreen vibrations
7. *CVD reactor*:
   Bifurcations
8. *Optimal cooling*:
   Steel profile
9. *MEMS: Micro-electro-mechanical systems*:
   Elf sensor

2.2. Weather: wave surge forecast

This problem concerns the prediction of wave surge at the coast of the Netherlands; in cases of high waves certain dams need to be closed to prevent flooding.

The North Sea is shallow compared, say, to the Atlantic ocean and therefore wave propagation is governed by the *shallow water equations*. A typical discretization results in about 60,000 ODEs.The computational time (on a laptop) is about 2 days which is prohibitively long. Therefore the system needs to be reduced.

Actually, there are eight measurement stations which provide weather data, and hence the resulting problem is one of *data assimilation*. Thus a Kalman Filter is needed, and the reduced filter propagates low-rank factors of the data covariance matrix.

For details see Heemink, Verlaan and Segers (2000).

2.3. Weather: air quality simulations

Here we are concerned with data assimilation for atmospheric chemistry models. This requires modeling of the formation and transport of air pollutants (CO). Chemistry transport models (CTM) are governed by PDEs together with boundary conditions. After spatial discretization with a typical grid of 100 points per 1 km in the $x$and y direction, and 30 points per 10 km in the z direction, we obtain a high number of ODEs which needs to be reduced.

Given pictures of the concentration of pollutants obtained e.g. by the *Terra satellite*, the goal is to derive a model which will predict these and other occurrences of high pollutant concentration (e.g. the propagation of CO across the Pacific).

For details, see Elbern and Schmidt (2001).

2.4. Vibration/acoustic systems

The problem here is the simulation of a car windscreen subject to an acceleration load. We seek to compute the noise generated at points away from the window while the car is in motion.

The model in this case is provided by a PDE which describes the deformation of a structure made of a specific material (glass in this case); In a typical case a finite element discretization yields 7564 nodes (three layers of 60 by 30 elements). The material is glass with Young modulus $7 \times 10^{10}\mathrm{N/m}^2$, density 2490 $\mathrm{kg/m}^3$, and Poisson ratio 0.23. Some coefficients of the finite element model are determined experimentally. The set-up yields second order equations of the type $M\ddot{x} + C\dot{x} + Kx = F$, where $x$has dimension 22,692; therefore the actual complexity of this problem is $n = 45,384$.

For details see Meerbergen (2001).

2.5. CVD reactors

The problem here is to analyze the stability of steady state flows in a CVD (chemical vapor deposition) reactor as a function of flow parameters like the Grashof, Rayleigh, and Reynolds numbers. These flows are solutions of the (full 3D) Navier Stokes equation. Therefore the resulting FEMs (finite element models) may contain a large number of variables. Indeed for this stability analysis the Amatrix of the linearized problem has size in excess of 10 million. To address the stability problem about 6–10 eigenvalues of A of largest real part are sought as a function of the above parameters.

This problem was solved using the software package P_ARPACK on a Sandia-Intel Tflop computer with 1024 processors.

For details, see Lehoucq and Salinger (2001).

2.6. MEMS: Elk sensor

A few years ago, production of the Mercedes A Class cars had to be stopped because they failed the *elk test*. This test consists in forcing the car to take a sharp turn in order to avoid an obstacle which suddenly appears on the road. To remedy this situation the company incorporated a roll-over (elk) sensor which could detect turning movement and apply the breaks to slow down the rotational movement. The first roll-over sensors were mechanical. Subsequently the company Robert Bosch AG developed a micro-electro-mechanical sensor at reduced cost and size. Nowadays the elk sensor can be found in many cars.

Once such a device has been designed, its performance must be tested by simulation. One method is physically oriented modeling, using appropriate simulation packages as described in Schwarz and Schneider (2001). The more detailed the modeling the higher the complexity (i.e. the number of differential equations) of the resulting model. And as the available simulation packages are built to handle low complexities, there is a need for simplification of the model through model reduction.

3. Approximation methods

The issues with large-scale systems are on one hand of numerical nature, to wit *storage*, *computational speed*, and *accuracy*. But in our setting they are also *system theoretic*, like the existence of an error bound the preservation of stability, passivity, etc. Thus large-scale problems have numerical and system components. We classify approximation methods in two basic categories (see also Fig. 4).

3.1. SVD-based methods

- Linear systems
  - Balanced truncation
  - Hankel approximation
- Nonlinear systems
  - POD methods
  - Empirical gramians

3.2. Krylov-based methods

- Realization
- Interpolation
- Lanczos procedure
- Arnoldi procedure

The third class of approximation methods which follows, aims at combining the best attributes of the SVD-based and the Krylov-based approximation methods.

3.3. Krylov/SVD-based methods

In the sequel we will briefly discuss these three families of approximation methods.

4. The SVD and a prototype problem

Given a matrix $A \in \mathbb{R}^{n\times m}$, its *singular value decomposition* (*SVD*) is a factorization $A = U\Sigma V^*$, where $U \in \mathbb{R}^{n\times n}$ and

$V \in \mathbb{R}^{m\times m}$ are unitary (i.e. $UU^* = I_n$, $VV^* = I_m$) and $\Sigma \in \mathbb{R}^{n\times m}$ is diagonal:

- The singular values are listed in decreasing order $\sigma_i = (\Sigma)_{ii} \geq \sigma_{i+1} = (\Sigma)_{i+1,i+1}$, and it turns out that $\sigma_i = \sqrt{\lambda_i(A^*A)}$, while σ_1 is the 2-induced norm of A.
- A consequence is the *dyadic decomposition*:

$$A = \sigma_1 u_1 v_1^* + \sigma_2 u_2 v_2^* + \cdots + \sigma_k u_k v_k^* + \sigma_{k+1} u_{k+1} v_{k+1}^* + \cdots + \sigma_n u_n v_n^*,$$

where u_i, v_j are columns of U, V.

4.1. Prototype problem

Optimal approximation in the 2-norm. Given the matrix $A \in \mathbb{R}^{n\times m}$, find $X \in \mathbb{R}^{n\times m}$, rank $X = k <$ rank A, such that the 2-norm of the error $E = A - X$, is minimized.

Solution (Schmidt–Mirsky, Eckart–Young, see Stewart and Sun (1990)).

$$\min_{\text{rank}X \leq k} \|A - X\|_2 = \sigma_{k+1}(A).$$

A minimizer $X_\#$ is obtained by truncating the dyadic decomposition of A

$$X_\# = \sigma_1 u_1 v_1^* + \sigma_2 u_2 v_2^* + \cdots + \sigma_k u_k v_k^*.$$

The SVD can be used to approximate images (which can be considered as static systems). In Fig. 2 this method is applied to the approximation of a supernova explosion taken by the Hubble telescope and that of a clown which can be found in *MATLAB*. The former has size 406×406 while the latter 200×320 pixels. The upper left-hand side pane of Fig. 2, depicts the 406 singular values of the supernova and the 200 singular values of the clown. For comparison the largest singular value in each case has been normalized to 1. The plot is in double logarithmic scale.

We thus notice that for approximants of complexity (rank) up to 10 the clown is easier to approximate as its first 10 singular values lie below those of the supernova. For complexities higher than 10 however, the opposite is true. For instance if the (largest) relative error allowed is 1%, the resulting approximant of the supernova has rank 31 while that of the clown 100. In conclusion, *the singular values in this approximation problem provide the trade-off between accuracy and complexity, when errors are measured in the 2-norm*. One of our goals in the sequel would be to come-up with a similar result for *dynamical systems*.

4.2. POD: Proper orthogonal decomposition

A straightforward application of approximation based on the SVD is the *proper orthogonal decomposition* (*POD*). Consider the nonlinear system $\dot{x} = f(x,u), y = h(x,u)$, where the state lives in $\mathbb{R}^n$. Assume that a preferred input and/or initial condition have been determined. The question then becomes as

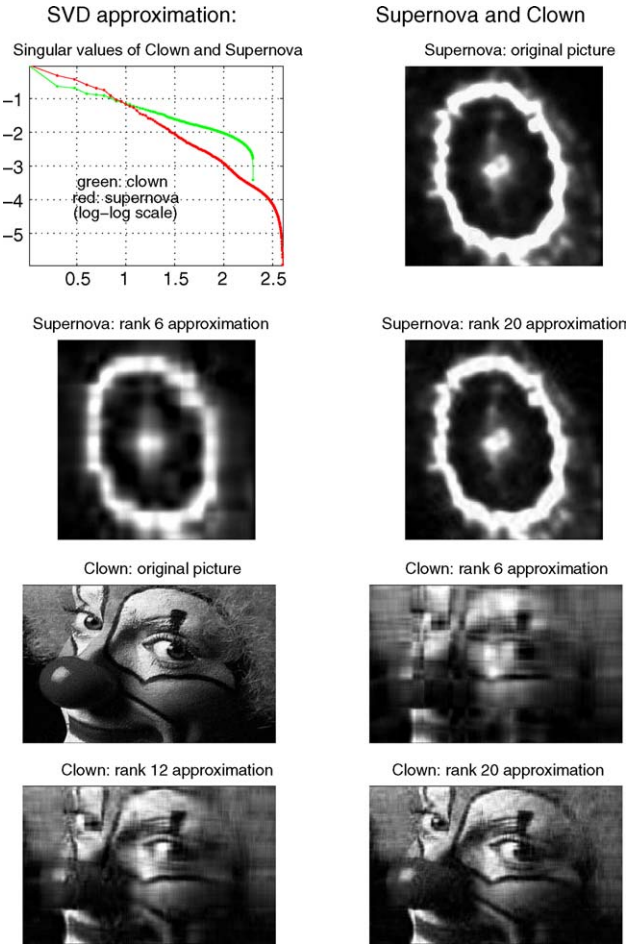

Fig. 2. Trade-off between accuracy and complexity.

to whether the evolution of the resulting state x lies mostly in a lower dimensional space $\mathbb{R}^k, k < n$. This question can be approached by measuring *snapshots* of the state $x(t_i)$, $i = 1, \ldots, N$. These are collected in a matrix

$$\mathcal{X} = [x(t_1)\, x(t_2) \cdots x(t_N)] \in \mathbb{R}^{n\times N},$$

and the SVD of $\mathcal{X}$ is computed: $\mathcal{X} = U\Sigma V^*$. Assuming that the singular values $\sigma_i, i > k$ are "small" with respect to the preceding ones, $\mathcal{X}$ can be approximated by means of $U_k \Sigma_k V_k^*$, where U_k, V_k denotes the submatrices of U, V respectively, consisting of the leading k columns. Thus with the (orthogonal projection) $\Pi = U_k U_k^*$, the state can be approximated as

$$\hat{x}(\cdot) = U_k^* x(\cdot) \in \mathbb{R}^k \Rightarrow x \approx U_k \hat{x}.$$

The projected state and output equations yield

$$\dot{\hat{x}} = U_k^* f(U_k \hat{x}, u), \qquad y = h(U_k \hat{x}, u).$$

Consequently $\hat{x}$ evolves in a *lower-dimensional* space.

Difficulties with POD arise because, for instance, the choice of snapshots is not an automated process, and the singular values of the state snapshots are not I/O invariants of the original system; thus they do not provide an error bound for the deviation of the state trajectories for different inputs or initial conditions.

4.3. The Hankel singular values

We now wish to briefly discuss the issue of extending the Schmidt–Mirsky, Eckard–Young result to linear dynamical systems. Consider the reachable, observable and stable linear system $\boldsymbol{\Sigma} = \left(\begin{array}{c|c} A & B \\ \hline C & \end{array}\right)$.. We would like to search for an operator associated with $\boldsymbol{\Sigma}$ whose singular values provide a trade-off between accuracy and complexity in the approximation of linear dynamical systems.

The most natural operator one can think of is the *convolution operator* $\mathcal{S}$, which maps inputs into outputs: $u \mapsto y = \mathcal{S}(u)$ where $y(t) = (h*u)(t) = \int_{-\infty}^{\infty} h(t-\tau)u(\tau)\mathrm{d}\tau$, and $h(t) = D\delta(t) + Ce^{At}B$, $t \geq 0$, is the *impulse response* of $\boldsymbol{\Sigma}$. Unfortunately, the singular values of $\mathcal{S}$ form a continuum and consequently they are *not* appropriate for our purpose.

An operator which is closely related to $\mathcal{S}$ is the *Hankel operator* $\mathcal{H}$ which maps past inputs into future outputs: $u_- \mapsto y_+ = \mathcal{H}(u_-)$ where $y_+(t) = \int_{-\infty}^{0} h(t-\tau)u_-(\tau)\mathrm{d}\tau$, $t \geq 0$. The (non-zero) singular values of $\mathcal{H}$, referred to as the *Hankel singular values* of $\boldsymbol{\Sigma}$, are discrete and finitely many: $\sigma_1 \geq \cdots \geq \sigma_n > 0$, and it can be shown that they are appropriate for model reduction and in particular for providing a trade-off between accuracy and complexity.

4.3.1. Computation of the Hankel singular values

The Hankel singular values are closely related with the system *gramians* which are defined as follows:

$$\mathcal{P} = \int_0^{\infty} e^{At}BB^*e^{A^*t}\,\mathrm{d}t, \qquad \mathcal{Q} = \int_0^{\infty} e^{A^*t}C^*Ce^{At}\,\mathrm{d}t.$$

The computation of these gramians involves the solution of two Lyapunov equations

$$A\mathcal{P} + \mathcal{P}A^* + BB^* = 0, \quad \mathcal{P} > 0,$$
$$A^*\mathcal{Q} + \mathcal{Q}A + C^*C = 0, \quad \mathcal{Q} > 0.$$

The Hankel singular values σ_i are then obtained by means of the eigenvalues of the product $\mathcal{PQ}$: $\sigma_i = \sqrt{\lambda_i(\mathcal{PQ})}$.

4.3.2. Approximation by balanced truncation

There exists a basis in the state space so that $\mathcal{P} = \mathcal{Q} = S = \mathrm{diag}\,(\sigma_1, \cdots, \sigma_n)$. This is the balanced basis of the system. In this basis partition

$$A = \left(\begin{array}{c|c} A_{11} & A_{12} \\ \hline A_{21} & A_{22} \end{array}\right), \qquad B = \left(\begin{array}{c} B_1 \\ \hline B_2 \end{array}\right), \qquad C = (\,C_1 \mid C_2,),$$
$$S = \left(\begin{array}{c|c} \Sigma_1 & 0 \\ \hline 0 & \Sigma_1 \end{array}\right),$$

where Σ_2 contains the small Hankel singular values.

The reduced order model obtained by *balanced truncation* is $\hat{\boldsymbol{\Sigma}} = \left(\begin{array}{c|c} A_{11} & B_1 \\ \hline C_1 & \end{array}\right)$; the associated *projector* is $\Pi = VW^*$ where $V = W = [I_k, 0]^*$.

4.3.3. Properties of balanced reduction

The advantages of this approximation methods are:

- the preservation of stability, and
- the existence of a *global error bound*, namely,

$$\sigma_{k+1} \leq \|\boldsymbol{\Sigma} - \hat{\boldsymbol{\Sigma}}\|_{\infty} \leq 2(\sigma_{k+1} + \cdots + \sigma_n),$$

which yields the desired *trade-off between accuracy and complexity* for linear dynamical systems.

Its *drawbacks* are that:

- *dense computations*, matrix factorizations and inversions of order $\mathcal{O}(n^3)$ are required, which may lead to ill-conditioning.
- The *whole* transformed system is needed in order to compute the truncated system, in other words there is no iterative way of computing the resulting k th order system.

5. Krylov approximation methods

This family of approximation methods is based on *moment matching*. Given

$$\boldsymbol{\Sigma} = \left(\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right) \in \mathbb{R}^{(n+p)\times(n+m)},$$

we expand its transfer function around $s_0 \in \mathbb{C}$:

$$G(s) = \eta_0 + \eta_1(s-s_0) + \eta_2(s-s_0)^2 + \eta_3(s-s_0)^3 + \cdots$$

The *moments* of $\boldsymbol{\Sigma}$ at s_0 are the coefficients η_j in the above expansion. The approximation problem now consists in finding

$$\hat{\boldsymbol{\Sigma}} = \left(\begin{array}{c|c} \hat{A} & \hat{B} \\ \hline \hat{C} & \hat{D} \end{array}\right) \in \mathbb{R}^{(k+p)\times(k+m)}, \qquad k < n,$$

such that

$$\hat{G}(s) = \hat{\eta}_0 + \hat{\eta}_1(s-s_0) + \hat{\eta}_2(s-s_0)^2 + \hat{\eta}_3(s-s_0)^3 + \cdots$$

where for appropriate ℓ:

$$\eta_j = \hat{\eta}_j, \qquad j = 1, 2, \ldots, \ell.$$

The strength of *moment matching* methods lies in the fact that they can be implemented in a *numerically efficient way*, namely:

- there is an *iterative* implementation which leads to the *Arnoldi* and *Lanczos* procedures;
- thus moments can be matched *without* computing them; this is important as the computation of moments is numerically problematic.

Some special cases of this problem are worth mentioning. (a) If the expansion is around infinity, the resulting problem is known as *partial realization*, and η_j are the *Markov parameters* of $\boldsymbol{\Sigma}$ ($\eta_0 = D$, $\eta_j = CA^{j-1}B$, $j > 0$). (b) If the expansion is around zero, the resulting problem is known as *Padé approximation*, and the coefficients η_j are known as *moments*

of $\boldsymbol{\Sigma}$. (c) In general, for arbitrary $s_0 \in \mathbb{C}$, the resulting problem is *rational interpolation*.

5.1. The two-sided Lanczos procedure

Given $\boldsymbol{\Sigma}$, with $m = p = 1$ for simplicity, let $\mathcal{R}_k = [B, AB, \ldots, A^{k-1}B]$, $\mathcal{O}_k = [C^*, A^*C^*, \ldots, (A^*)^{k-1}C^*]^*$, $\mathcal{H}_k = \mathcal{O}_k\mathcal{R}_k$, be the partial reachability, observability, and Hankel matrices. The *key* step in order to obtain the projector $\Pi = VW^*$, consists in computing the LU (lower–upper triangular) factorization (if it exists) of the Hankel matrix $\mathcal{H}_k = LU = \mathcal{O}_k\mathcal{R}_k$. The projector is defined as $W^* = L^{-1}\mathcal{O}_k$, $V = \mathcal{R}_k U^{-1}$, and the resulting *reduced order system*:

$$\hat{\boldsymbol{\Sigma}} : \hat{A} = W^*AV, \qquad \hat{B} = W^*B, \qquad \hat{C} = CV, \qquad \hat{D} = D,$$

has the following properties:

(a) $\hat{\boldsymbol{\Sigma}}$ matches $2k$ Markov parameters of $\boldsymbol{\Sigma}$;
(b) $\hat{A}$ is tridiagonal;
(c) $\hat{B}, \hat{C}^*$ are multiples of the unit vector e_1.

Next we state the Lanczos algorithm; this is intended to stress the fact that its *iterative implementation* requires vector-matrix multiplications exclusively.

- $\beta_1 = \sqrt{|B^*C^*|}$, $\gamma_1 = \operatorname{sgn}(B^*C^*)\beta_1$, $v_1 = B/\beta_1$, $w_1 = C^*/\gamma_1$.
- For $j = 1, \ldots, k$, set

$$\alpha_j = w_j^* A v_j, \qquad r_j = Av_j - \alpha_j v_j - \gamma_j v_{j-1},$$

$$q_j = A^* w_j - \alpha_j w_j - \beta_j w_{j-1}, \qquad \beta_{j+1} = \sqrt{|r_j^* q_j|},$$

$$\gamma_{j+1} = \operatorname{sgn}(r_j^* q_j)\beta_{j+1}, \qquad v_{j+1} = r_j/\beta_{j+1},$$

$$w_{j+1} = q_j/\gamma_{j+1}.$$

5.1.1. Properties of Krylov methods

(a) *Numerical efficiency*, since the number of operations is $\mathcal{O}(kn^2)$ or $\mathcal{O}(k^2n)$ versus $\mathcal{O}(n^3)$ for balanced truncation.
(b) Only *matrix-vector* multiplications are required; no matrix factorizations and/or inversions; there is no need to compute the whole transformed model and then truncate.
(c) Its *drawbacks* are
  - no apriori computable error bound.
  - $\hat{\boldsymbol{\Sigma}}$ may not be stable.
  - Lanczos breaks down if $det\, \mathcal{H}_i = 0$ (remedy: look-ahead methods).
  - $\hat{\boldsymbol{\Sigma}}$ tends to approximate the high frequency behavior of $\boldsymbol{\Sigma}$ (remedy: match expansions around other frequencies, i.e. use rational Lanczos methods).

6. SVD/Krylov approximation methods

As we argued above, SVD-based and Krylov-based approximation methods have distinct sets of advantages (see Fig. 4). Therefore a lot of effort currently goes into devising approximation methods which combine the best attributes of these two approximation methods. We shall refer to these as SVD/Krylov approximation methods and will only point out some of their salient features; for details we refer to the book (Antoulas, 2005a).

- By appropriate choice of weightings, weighted balanced truncation methods can be reduced to Krylov methods. Thus weighted SVD methods have a direct connection with Krylov methods.
- Projectors for rational Krylov methods (generalized reachability/observability matrices) can be obtained as solutions of Sylvester equations.
- There are methods which combine attributes of both approaches. One of them is *least squares approximation*.
- The bottleneck in applying SVD methods (balanced truncation) to large-scale systems is the fact that the solution of Lyapunov equations requires $\mathcal{O}(n^3)$operations where $n$is the dimension of the original system.

To remedy this situation *iterative methods* for solving Lyapunov equations *approximately* have been proposed. This leads to approximately balancing transformations and approximately balanced and truncated systems.

6.1. Example

We consider the simulation of the structural model of an *aluminium plate*; its size is 0.5 m by 0.5 m, Young modulus 7.0×10^{10} N/m^2, thickness 0.001 m, Poisson ratio 0.33, and density 2700 kg/m^3 with no structural damping. It was discretized by a grid of 16×16 solid shell elements. The plate is subjected to a unit point force in the coordinate (0.125, 0.125 m). The goal is to compute the amplitude of the vertical displacement in the same position for the frequency range $\omega \in [10, 110]$ Hz. The resulting complexity is $n = 3468$.

Given this relatively high complexity, we apply a rational Krylov method, in other words moment matching, at the frequency $s = 0$. The results are shown in Fig. 3. The upper pane shows the amplitude Bode plot of the original system with

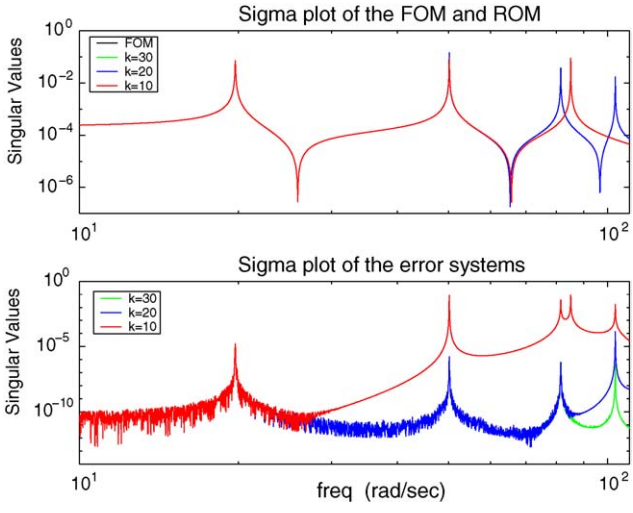

Fig. 3. Amplitude Bode plots for the aluminum plate example.

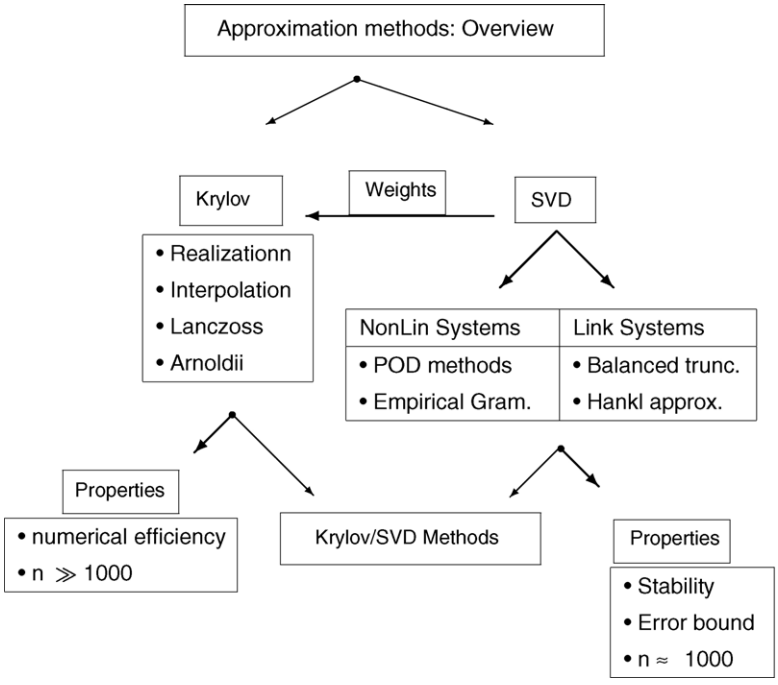

Fig. 4. There are two basic categories of approximation methods, namely the Krylov- and SVD-based methods, each with a different set of attributes. The attempt to combine the best attributes of each of these methods leads to a third category, that of Krylov/SVD approximation methods. Finally, weightings provide a link between SVD-based and Krylov-based approximation methods.

$n = 3468$, and three reduced systems of order $k = 10, 20, 30$. The lower pane depicts the error Bode plots; moments are matched around $s = 0$. We conclude that the system is rather easy to approximate as dimension reduction of 1:1000 yields a Bode plot which is practically indistinguishable from that of the original system.

7. A new result on passive model reduction

We will now briefly describe a new result which concerns model reduction with preservation of stability and passivity. This is important in model reduction as applied for the verification of VLSI chip performance.

*Passive systems*are those which do not generate energy. More precisely, in terms of inputs and outputs the following relationship must be satisfied: $\mathrm{Re}\int_{-\infty}^{t} u(\tau)^* y(\tau)\mathrm{d}\tau \geq 0$, for all $t \in \mathbb{R}$, and all inputs u which are square integrable on $\mathbb{R}$.

A function G of the complex variable s is *positive real*, provided that (1) G is analytic for $\mathrm{Re}(s) > 0$, and (2) $\mathrm{Re}G(s) \geq 0$ for $\mathrm{Re}(s) \geq 0$, where s is not a pole of G.

A well known result asserts that the system $\boldsymbol{\Sigma} = \left(\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right)$ is passive if, and only if, its transfer function $G(s) = D + C(sI - A)^{-1}B$ is positive real. The new approach makes use of the following results.

7.1. Preliminaries

- Positive realness of $G(s)$ implies the existence of a *spectral factorization* $G(s) + G^*(-s) = \Phi(s)\Phi^*(-s)$, where $\Phi(s)$ and $\Phi(s)^{-1}$ are both stable. The *spectral zeros* of the system λ_i, are the zeros of the spectral factor, i.e. $\Phi(\lambda_i) = 0, i = 1, \ldots, n$.
- *Rational Lanczos*. Skelton-de Villemagne and Grimme-Van Dooren (see e.g. Antoulas, 2005a, b) have shown (assuming for simplicity $m = p = 1$) the following. Let

$$V = [(\lambda_1 I - A)^{-1}B \ldots (\lambda_k I - A)^{-1}B] \in \mathbb{R}^{n\times k},$$

$$\bar{W}^* = \begin{bmatrix} C(\lambda_{k+1}I - A)^{-1} \\ \vdots \\ C(\lambda_{2k}I - A)^{-1} \end{bmatrix} \in \mathbb{R}^{k\times n},$$

where $\lambda_i \neq \lambda_j$, $\lambda_{i+k} \neq \lambda_{j+k}$, $i, j = 1, \cdots, k$; if $\Gamma = \bar{W}^* V \in \mathbb{R}^{k\times k}$ is non-singular $\det \Gamma \neq 0$, define $W^* = \Gamma^{-1}\bar{W}^*$. The transfer function of the projected system $\hat{\boldsymbol{\Sigma}}$ where $\Pi = VW^*$

$$\hat{A} = W^*AV, \qquad \hat{B} = W^*B, \qquad \hat{C} = CV, \qquad \hat{D} = D,$$

interpolates the transfer function of the original system $\boldsymbol{\Sigma}$ at the interpolation points $\lambda_i : G(\lambda_i) = \hat{G}(\lambda_i), i = 1, \ldots, 2k$.

7.1.1. Positive real interpolation

Problem. Given the pairs of points (λ_i, h_i), find a positive real rational function $G(s)$ such that $G(\lambda_i) = h_i, i = 1, \ldots, n$. It is well known that this problem is solvable if and only if the Pick matrix $P = \left[\frac{h_i + h_j^*}{\lambda_i + \lambda_j^*}\right]$, is positive semi-definite. The following result proved in (Antoulas, 2005b), forms the basis of the new approach.

Theorem 1 If the problem is solvable, any minimal rational interpolant of the pairs (λ_i, h_i), together with the Mirror Image pairs $(-\lambda_i^*, -h_i^*)$, is positive real.

We are now ready to state the main result which involves the rational Lanczos procedure and the spectral zeros. For details and proofs see (Antoulas, 2005b).

7.1.2. Main result

Given is the stable and passive system $\boldsymbol{\Sigma}$. If the projector $\Pi = VW^*$is defined as above, where $\lambda_1, \ldots, \lambda_k$ are *stable spectral zeros*, and in addition $\lambda_{k+i} = -\lambda_i^*, i = 1, \ldots, k$, the reduced system $\hat{\boldsymbol{\Sigma}}$ satisfies:

(i) the interpolation constraints,
(ii) it is stable, and
(iii) it is passive.

7.1.3. Spectral zeros as eigenvalues

The *zeros* of the spectral factors Φ can be computed by solving an *eigenvalue problem* involving A, B, C, D. If $D + D^*$ is non-singular this is reduced to the computation of the eigenvalues of the matrix

$$\mathbf{H} = \begin{bmatrix} A & 0 \\ 0 & -A^* \end{bmatrix} - \begin{bmatrix} B \\ -C^* \end{bmatrix} (D + D^*)^{-1} [\, C \qquad B^* \,].$$

Notice that $\mathbf{H}$ is a *Hamiltonian* matrix of size $2n \times 2n$ which means that $(\mathbf{H}J)^* = \mathbf{H}J$, where $J = \begin{pmatrix} 0 & -I_n \\ I_n & 0 \end{pmatrix}$.

If $D + D^*$ is singular, the spectral zeros are the finite generalized eigenvalues $\lambda \in \sigma(\mathcal{A}, \mathcal{E})$ of the pair $(\mathcal{A}, \mathcal{E})$; in other words the spectral zeros are such that rank $(\mathcal{A} - \lambda\mathcal{E}) < 2n + p$, where

$$\mathcal{A} = \begin{pmatrix} A & & B \\ & -A^* & -C^* \\ C & B^* & D + D^*, \end{pmatrix}, \quad \mathcal{E} = \begin{pmatrix} I & & \\ & I & \\ & & \mathbf{0} \end{pmatrix}.$$

7.1.4. Interpolation via invariant subspaces

The above result can be strengthened further. It turns out namely that the projector $\Pi = VW^*$ can be computed by means of appropriate invariant subspaces of the above generalized eigenvalue problem. Let

$$\begin{pmatrix} A & & B \\ & -A^* & -C^* \\ C & B^* & D + D^* \end{pmatrix} \overbrace{\begin{bmatrix} X \\ Y \\ Z \end{bmatrix}}^{k} = \overbrace{\begin{bmatrix} X \\ Y \\ 0 \end{bmatrix}}^{k} R.$$

It can be shown that if all eigenvalues of R are in the right-half of the complex plane, X and Y have both full rank and $X^*Y = Y^*X$. This leads to the following construction of the *projectors* which guarantee passivity and stability. First, compute the SVD

$$X^*Y = Q_x S^2 Q_y^*.$$

Then following (Sorensen, 2005) the projector $\Pi = VW^*$, is defined as follows:

$$V = XQ_x S^{-1}, \qquad W = YQ_y S^{-1}.$$

The resulting reduced order system where $\hat{A} = W^*AV, \hat{B} = W^*B, \hat{C} = CV, \hat{D} = D$, is both *stable* and *passive*.

We conclude this section by pointing out that the problem of approximation with preservation of stability and passivity has thus been reduced to a *structured eigenvalue problem*, which can be solved efficiently for systems of very high complexity.

8. Concluding remarks

8.1. Projectors and complexity

Projectors are a unifying feature of model reduction methods. We project with $\Pi = VW^*, \Pi^2 = \Pi$, to get reduced models of the form $\dot{\hat{x}} = (W^*AV)\hat{x} + (W^*B)u, \hat{y} = (CV)\hat{x} + Du$. The *quality* of the approximation is measured in terms of the frequency response $G(j\omega) = C(j\omega I - A)^{-1}B + D$, and in particular by its *peak*, known as the $\mathcal{H}_\infty$*-norm*, or its *2-norm* (energy) known as the $\mathcal{H}_2$*-norm*.

8.2. Computational complexity

The computational complexity of the various approximation methods is shown in the table below. n denotes the dimension of the *original* system, k denotes the dimension of the *reduced* system and α denotes the average number of non-zero elements per row of A.

	Dense oper.	Approx. sparse oper.
SVD: solve for gramians	$\approx n^3$	$\approx \alpha kn$
SVD: perform balancing	$\approx n^3$	$\approx \alpha kn$
Krylov	$\approx kn^2$	$\approx \alpha kn$

8.3. Software

The goal is to obtain reliable solutions of large dimensional numerical problems encountered in complex control applications. Turning numerically reliable algorithms into high performance, portable and robust software relies on numerical linear algebra, computer science, computer hardware, etc. Next we list some existing software packages.

- General purpose software (including model reduction): MATLAB: http://www.mathworks.com, SCILAB: http://www.scilab.org, MATRIX$_X$: http://www.ni.com/matrixx/
- Model reduction software: SLICOT. http://www.win.tue.nl/niconet
- Parallel model reduction software: PSLICOT.
- Linear algebra package: LAPACK.
- For large-scale eigenvalue problems: ARPACK and its parallel implementation P_ARPACK. These software packages are available at http://www.caam.rice.edu/software/ARPACK
- Some specialized software packages: SUGAR MEMS simulation: http://www-bsac.eecs.berkeley.edu/cadtools/sugar/,

MOZART Ozone propagation:
http://acd.ucar.edu/models/MOZART/,
ANSYS Simulation of system described by PDEs:
http://www.ansys.com/

8.4. Open problems

We conclude with a (partial) list of open problems; again for details see (Antoulas, 2005a).

- Decay rate of the Hankel singular values.

  In chapter 9 of (Antoulas, 2005a) upper bounds on the rates of decay as well as approximate rates of decay of the eigenvalues of a single Lyapunov equation are presented. Numerical examples show the usefulness of these results. Computationally, these results require the knowledge of all eigenvalues (poles) of the system. At first one would need to extend the upper bound to the eigenvalues of the product of two gramians. The next step would be to develop bounds on the rates of decay when only partial information about the eigenvalues is available (like an inclusion domain in the complex plane).
- Choice of expansion points for rational Krylov.

  In chapter 11 of (Antoulas, 2005a) it is shown that given two scalar proper rational functions G_1, G_2 of McMillan degree $n_1 > n_2$, respectively, the latter can be obtained almost always, from the former through interpolation which can be implemented iteratively by means of the rational Krylov procedure. Therefore in the generic SISO case *any* reduced order model is attainable through appropriate choice of interpolation points. Thus the question of *how* to choose the interpolation points so as to minimize a desirable error criterion arises. Krylov and rational Krylov methods provide no guidance in that respect. The question thus arises concerning rules for choosing the interpolation points so that a desirable norm of the error system is minimized.
- Choice of spectral zeros in passive reduction.

  In Section 7it was argued that by choosing the interpolation points as spectral zeros of the original system, the reduced system will be automatically passive and stable. As in the previous problem, the issue which arises is to choose those spectral zeros that will reduce or minimize some appropriate norm of the error system.
- Iterative SVD model reduction methods.

  In chapter 12 of (Antoulas, 2005a) iterative methods for computing SVD type approximants are proposed. The open problems are convergence and stability of the first proposed algorithm. The second algorithm is guaranteed to converge to a reduced system which is balanced; however it is not clear how many steps are needed for the reduced system, which is approximately balanced, to be stable. Furthermore iterative methods for weighted balanced reduction are missing.
- Model reduction of second-order systems.

  In many cases, e.g. when mechanical systems modeled as mass-spring-damper are involved, the reduction should respect the second order structure (the state is composed of positions and the corresponding velocities). Converting the system to first order and applying existing methods for model reduction, is an option, but at the expense of destroying the second order structure (the reduced system may no longer be representable as a mass-spring-damper system). Recently (Chahlaoui, Lemonnier, Vandendorpe, & Van Dooren, 2005) have revisited the problem. Issues still remaining are stability, error bounds, the choice of gramians, as well as iterative methods. Weighted reduction for this class of systems remains an open problem.
- Model reduction of structured multi-physics systems.

  The difficulty in analyzing and synthesizing MEMS (Micro-electro-mechanical systems) also known as *Microsystems*, stems from the complexity of the underlying structure as well as its heterogeneity; such a system may contain mechanical, electrical, fluidic, thermal, acoustic, etc., components. This leads to multi-domain and multi-physics problems.

  These issues arising in simulating these devises can be addressed by reduction of the order of the microsystem model. In such a procedure the various physical systems (electrical, mechanical, etc.) must be reduced separately and the reduced models put together by means of the interconnections. Furthermore, the procedure should produce reduced-order systems which have a physical interpretation as electrical, mechanical, etc. The overall goal is to understand as many effects occurring in microsystems as possible through simulation (see e.g. Schwarz and Schneider, 2001).

References

Antoulas, A. C. (2005a). *Approximation of large-scale dynamical systems*. Philadelphia: SIAM.

Antoulas, A. C. (2005b). A new result on passivity preserving model reduction. *Systems and Control Letters, 54*, 361–374.

Chahlaoui, Y., Lemonnier, D., Vandendorpe, A., & Van Dooren, P. M. (2005). *Second order balanced truncation*. Linear Algebra and Its Applications.

Elbern, H., & Schmidt, H. (2001). Ozone episode analysis by four-dimensional variational chemistry data assimilation scheme. *J. Geophysical Research, 106*(D4), 3569–3590.

Heemink, A. W., Verlaan, M., & Segers, A. J. (2000). Variance reduced ensemble Kalman filter. *Technical Report,* 00–03.

Lehoucq, R. B., & Salinger, A. G. (2001). Large-scale eigenvalue calculations for stability analysis of steady flows on massively parallel computers. *Int. J. Numer. Meth. Fluids, 36*, 309–327.

Meerbergen, K. (2001). Model reduction techniques for parametrized linear systems arising from vibro-acoustics. *SIAM Conference on Linear Algebra in Signal Systems and Control*.

Schwarz, P., & Schneider, P. (2001). Model Library and Tool Support for MEMS Simulation in Microelectronics and MEMS Technologies. SPIE Proceedings Series, vol. 4407, Paper 2.

Sorensen, D. C. (2005). Passivity preserving model reduction through interpolation of spectral zeros. *Systems and Control Letters, 54*, 347–360.

Stewart, G. W., & Sun, J. (1990). *Matrix perturbation theory*. Academic Press.

van der Meijs, P. (2000). Model reduction for VLSI physical verification, Technical Report, Department of ITS/EE, Delft University of Technology.

Thanos Antoulas was born in Athens, Greece. He studied at the ETH Zurich, where he obtained the Diploma of Electrical Engineering in 1975, the Diploma of Mathematics in 1975, and the PhD Degree in Mathematics in 1980.Since 1982 he has been with the Department of Electrical and Computer Engineering, Rice University, where he is currently a Professor. Since 2003 he has also been affiliated with the International University Bremen (IUB).Professor Antoulas

was elected Fellow of the IEEE (Institute of Electrical and Electronics Engineers) in 1991 and was awarded the Best Paper Prize of the AIAA (American Institute of Aeronautics and Astronautics) Guidance, Navigation and Control Section, in 1992. In 1995 he was awarded a JSPS (Japan Society for the Promotion of Science) Fellowship. He has served on the Editorial Board of many journals including the IEEE Transactions on Automatic Control, SIAM Journal on Control and Optimization, and Linear Algebra and its Applications. For the past decade he has been serving as one of the Editors-in-Chief of Systems and Control Letters.He has held many visiting appointments, including those at the Australian National University, the University of Groningen, the Catholic University of Leuven and Louvain-la-Neuve, the Tokyo Institute of Technology, Kyoto University, etc. Furthermore he has given invited presentations at numerous conferences and workshops. His research interests are in the broad area of dynamical systems and computation including model reduction of large-scale systems. He is the author of the book ‘‘Approximation of large-scale dynamical systems’’, published by SIAM.