Model reduction using the AISIAD algorithm

2^{nd} Young ERCOFTAC Workshop

Gary Chandler*, Christoph Mack† and Shervin Bagheri‡

*University of Cambridge, UK
† UniBw München, Germany
‡Linné Flow Centre, Dept. of Mechanics, Stockholm, Sweden

1 Introduction

Many powerful linear systems and control theory tools have been out of the reach of the fluids community due to the complexity of the Navier-Stokes equations. Flow control based on systematic methods adopted from control theory is becoming a fairly mature field, with both computational and experimental advances. In this sense, model reduction plays an important role in developing effective control strategies for practical applications, since the dynamical systems which describe most flows are discretized partial differential equations with many degrees of freedom. Currently, balanced truncation represents the standard method of model reduction in systems and control theory which preserves the main input-output characteristics of the system. In this article we investigate an approximate balanced truncation by applying a modified version of the AISIAD algorithm [4]. The method is demonstrated on the Ginzburg-Landau equation and the linearized flow in a plane channel.

2 Background

In this section we present the concept of model reduction as well as the main steps of the AISIAD algorithm.

2.1 Model reduction

Consider a stable linear system in state-space form

$$\begin{aligned} \dot{x} &= Ax + Bu \\ y &= Cx \ , \end{aligned} \tag{1}$$

where A denotes the $n \times n$ *system matrix*, B is the $n \times m$ *input matrix* and C represents the $l \times n$ *output matrix*, with n as the dimension of the full system. x, u and y denote the n-dimensional state vector, the m-dimensional control vector and the l-dimensional measurement vector, respectively. For fluid-dynamical systems, the system matrix A represents the linearized and discretized Navier-Stokes equations including the corresponding boundary conditions. For two- or three-dimensional flows the size of A is often above 10^5.

The idea of model reduction is now to construct a reduced system of order $q \ll n$

$$\begin{aligned} \dot{x}_r &= A_r x_r + B_r u \\ y &= C_r x_r \ , \end{aligned} \tag{2}$$

so that the input-output behavior (frequency response) of the original system (1) is preserved. To achieve this the reduced system is projected onto a bi-orthogonal set of functions, called the balanced modes and defined as the eigenfunctions of the cross Gramian PQ, where P and Q are the *controllability* Gramian and the *observability* Gramian, respectively. These two $n \times n$ matrices are solutions of the following Lyapunov equations:

$$AP + PA^H = -BB^H \tag{3a}$$
$$A^H Q + QA = -C^H C \ . \tag{3b}$$

In the method of balanced truncation, the left and the right dominant eigenvectors of the resulting cross Gramian PQ can then be used to reduce the full system by projecting the system matrices via (see [4] algorithm 1)

$$A_r = S_o^H A S_c \ , \quad B_r = S_o^H B \ , \quad C_r = C S_c \ , \tag{4}$$

where S_c and S_o denote the projection matrices, the balanced modes and their associated adjoint modes, respectively.

Solving the Lyapunov equations is, however, prohibitively expensive for large systems. When $m \ll n$ and outputs $l \ll n$ the right-hand side of (3) are of low rank, which in turn indicates that the Lyapunov equations have low rank approximations. Two approaches exist: the snapshot-based balanced truncation [5] and iterative methods based on power iterations or Krylov subspace methods. In what follows, we will focus on iterative methods.

It is worth mentioning, that the outlined concept of model reduction only produces good results if the dominant controllability and observability eigenspaces coincide. In other words, the states need to be both *controllable* and *observable*.

2.2 The modified AISIAD algorithm

The AISIAD (Approximate Implicit Subspace Iteration with Alternating Directions) algorithm, as described in [4], directly approximates the dominant left and right eigenspaces of the cross Gramian PQ. It is based on a block power iteration method and consists of the following steps:

1. Choose an orthogonal matrix V_1 of size $n \times q$ as an initial guess
2. Iterate until convergence:
   (a) Solve the projected Lyapunov equation for X_i,

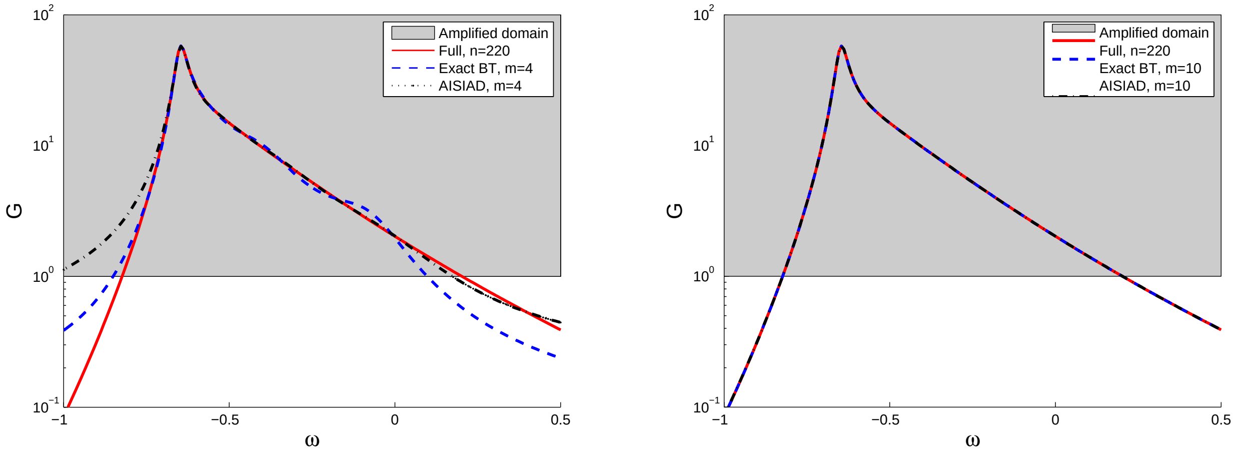

Figure 1: A comparison of the frequency response of the full model (red), the exact balanced truncation model (blue dashed) and the AISIAD-based model (black dashed dotted). The reduced system consists of a dimension of $m = 4$ (left) and $m = 10$ (right). The gray region marks the amplified frequencies.

$$AX_i + X_i H_i^H + BB^H V_i = 0, \qquad (5)$$

where $H_i^H = V_i^H A^H V_i$ and $X_i = PV_i$. Note that H_i should be stable.

(b) Obtain an orthogonal basis which spans the same subspace as V_i

$$[W_i, S_i] = qr(X_i, 0). \qquad (6)$$

(c) Solve the projected Lyapunov equation for Y_i

$$A^H Y_i + Y_i F_i + C^H C V_i = 0, \qquad (7)$$

where $F_i = W_i^H A W_i$ and $Y_i = QW_i$.

(d) Obtain an orthogonal basis which spans the same subspace as W_i

$$[V_{i+1}, R_i] = qr(Y_i, 0). \qquad (8)$$

(e) Check tolerance

$$\|\mathrm{diag}\{R_i S_i\} - \mathrm{diag}\{R_{i-1} S_{i-1}\}\| \leq tol, \quad (9)$$

where the diagonal elements of RS are the eigenvalues of PQ at convergence. The square root of these eigenvalues are called the Hankel singular values (HSV).

3. Normalize $V_L = V_{i+1}$ and $W_R = W_i$,

$$[U, \Sigma, V] = svd(V_L^H W_R), \qquad (10)$$

and recover the balanced modes and their associated adjoint modes from

$$S_c = W_R V \Sigma^{-1/2} \;, \qquad S_o = V_L U \Sigma^{-1/2} \;. \quad (11)$$

At convergence S_o and S_c are the left and the right eigenmodes of the cross Gramian PQ, respectively. For further details, the reader is referred to [4] algorithm 2. Finally, the projection matrices S_o and S_c can be applied to reduce the full system following (4).

To improve the performance of the (original) AISIAD algorithm, Vasilyev and White [4] proposed an efficient solution of the Sylvester equations (5) and (7). By transforming (5) via $\widetilde{X} = XU$, with U resulting from a Schur decomposition of H, the modified equation yields:

$$A\widetilde{X} + \widetilde{X}S = -\widetilde{M}. \qquad (12)$$

This linear system is then solved backwardly, x_j representing the j^{th} column of $\widetilde{X}$

$$(A + s_{jj} I_n)\, \tilde{x}_j = -\tilde{m}_j \;, \qquad (13)$$

and direct or sparse linear solvers can be applied. In addition to that, an iterative Krylov-subspace solver such as BiCGStab or GMRES can be employed too. Even more, since the latter methods are based on matrix-vector products they permit a matrix-free implementation, e.g., based on direct numerical simulations (DNS).

3 Application to the Ginzburg-Landau equation

The complex Ginzburg-Landau equation is an amplitude equation, which arises in the context of non-equilibrium systems. It is often used to describe the system dynamics near the onset of linearized instability. A linearized version of this model is used to mimic spatially developing flows, such as boundary-layers, jets and cylinder wakes (see [1]). The equation is of a convection-diffusion type with one additional term to model different types of instabilities,

$$A = -\nu \frac{\partial}{\partial x} + \gamma \frac{\partial^2}{\partial x^2} + \mu(x) \;, \qquad (14a)$$

$$B = \exp\left[-\left(\frac{x - x_{w,i}}{s} \right)^2 \right] \;, \qquad (14b)$$

$$C\boldsymbol{x} = \int_{-\infty}^{\infty} \exp\left[-\left(\frac{x - x_{s,i}}{s} \right)^2 \right]^H \boldsymbol{x}(t) \mathrm{d}\boldsymbol{x}, \quad (14c)$$

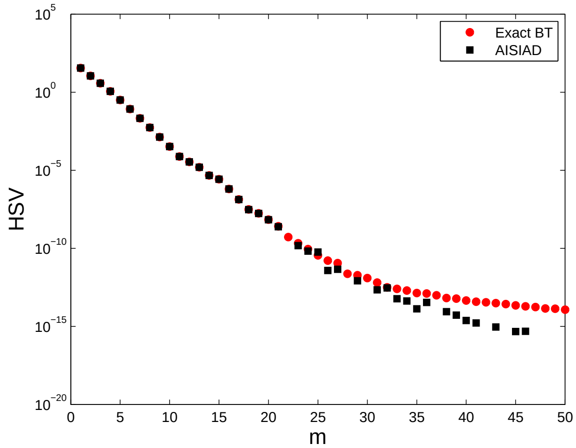

Figure 2: A comparison of the Hankel singular values (HSV) of exact balanced truncation (red circles) and the AISIAD method (black squares).

with $\boldsymbol{x}(t) < \infty$ as $x \to \pm\infty$. Moreover, the convection and diffusion term are complex valued functions in order to model dispersion and frequency selection effects. All eigenvalues of A have negative real part, and, therefore, the system is stable. However, when considering a quadratic instability function

$$\mu(x) \quad = (\mu_0) + \mu_2 \frac{x^2}{2}, \qquad \mu_2 < 0, \quad (15)$$

the flow becomes susceptible to instabilities for $\mu(x) > 0$, which defines a confined unstable region in the x-direction as given by $-\sqrt{-2(\mu_0)/\mu_2} < x < \sqrt{-2(\mu_0)/\mu_2}$.

We now compare the AISIAD method with the exact balanced truncation for computing a reduced-order model. The latter method is computationally feasible since a spectral discretization of the above one-dimensional PDE results in $n = 220$. The exact balanced truncation is computed using the square-root method (see [2]). In Figure 1 the frequency response defined as

$$|G| = |C(i\omega I - A)^{-1}B|, \quad (16)$$

is compared for the full model, the exact balanced truncation and the AISIAD balanced truncation. In Figure 1a the reduced system is of order $m = 4$ and the AISIAD algorithm required 45 iterations for convergence to a tolerance of 10^{-6}. In Figure 1b the reduced system is of order $m = 10$ and the AISIAD algorithm required 7 iterations for convergence to the same tolerance. We observe that both the exact balanced truncation method and the AISIAD method approximate the full frequency response very well for $m = 4$ and almost exactly for $m = 10$. In Figure 2 the Hankel singular values, i.e. the square-root of the eigenvalues of the cross Gramian PQ, are compared for the exact balanced truncation and the AISIAD method. We observe that for $m = 50$ the AISIAD method (which required only 7 iterations for convergence of 10^{-6}) approximates the first 20 Hankel singular values very well, but fails to approximate the very small eigenvalues close to machine epsilon.

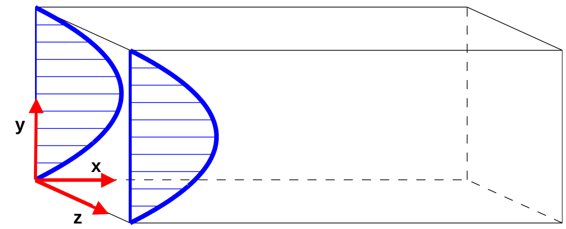

Figure 3: Sketch of plane channel flow

4 Application to the Linearized Channel Flow

The flow configuration under investigation is the laminar Poiseuille flow in a plane channel with the velocity profile $U(y) = 1 - y^2$ as displayed in Figure 3. According to [3], this flow is linearly stable for Reynolds numbers $Re < 5772$ which leads to a stable input-output system of the form (1).

4.1 Equations of motion

The dynamics of small perturbations to the plane channel flow is governed by the following formulation of the incompressible linearized Navier-Stokes equations

$$\frac{\partial u}{\partial t} + U\frac{\partial u}{\partial x} + \frac{\partial U}{\partial y}v = -\frac{\partial p}{\partial x} + \frac{1}{Re}\Delta u \quad (17a)$$
$$\frac{\partial v}{\partial t} + U\frac{\partial v}{\partial x} = -\frac{\partial p}{\partial y} + \frac{1}{Re}\Delta v \quad (17b)$$
$$\frac{\partial w}{\partial t} + U\frac{\partial w}{\partial x} = -\frac{\partial p}{\partial z} + \frac{1}{Re}\Delta w \quad (17c)$$
$$\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z} = 0, \quad (17d)$$

where u, v and w denote the velocity components of the perturbation in the streamwise x-direction, the wall-normal y-direction and the spanwise z-direction, respectively. The Reynolds number is defined as $Re = (UL)/\nu$ and the continuity equation (17d) is applied to constrain the solution to be divergence free.

The flow is assumed to be periodic in the x- and z-direction, with no-slip boundary conditions at the walls, and we assume the following traveling-wave form

$$\phi(x, y, z, t) = \widetilde{\phi}(x, y)\mathrm{e}^{\mathrm{i}(\beta z - \omega t)} \quad (18)$$

with $\phi = (u, v, w, p)^T$. In this expression, $\widetilde{\phi}(x, y)$ denotes the complex amplitude and β the real spanwise wavenumber of the perturbation. The parameter ω characterizes the temporal long-term evolution of this disturbance.

Under these assumptions, the system can be written as $\partial\widetilde{\phi}/\partial t = \mathcal{L}(U)\widetilde{\phi}$, where $\mathcal{L}(U)$ represents the linear stability operator, which, in this case, is the Navier–Stokes equations linearized about the laminar state $U(y)$. This continuous system is descretized in the x- and the y-direction and the discrete system can be rewritten in the standard state-space form

$$\dot{\boldsymbol{x}} = \boldsymbol{A}\boldsymbol{x}. \quad (19)$$

Herein, $\dot{\boldsymbol{x}}$ denotes the time derivative of $\boldsymbol{x}$ and A is the discrete system matrix.

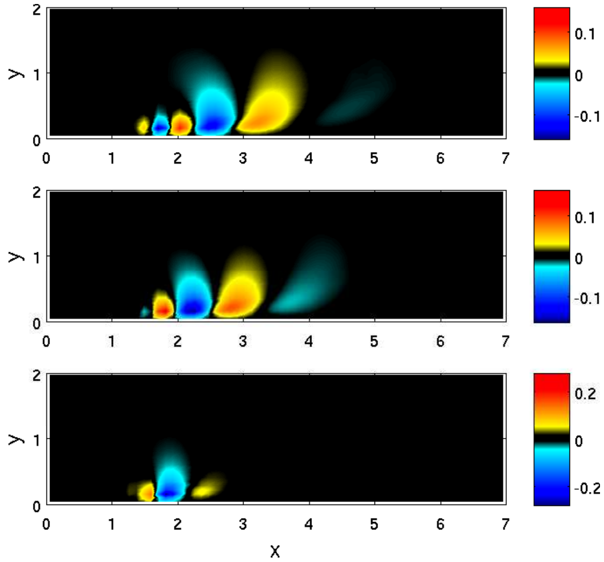

Figure 4: First three computed balanced modes S_c using the AISIAD method ($m = 4$). The $\widetilde{v}$ velocity is displayed.

In addition to the forward solution, based on the system matrix A, the solution of the adjoint system, described by the adjoint system matrix A^H, is required to solve (7). The adjoint system is derived in a similar manner as the forward system and, owing to a lack of space, the reader is referred to the literature.

In this article, the spatial discretization in the x- and the y-direction is accomplished using second-order finite difference schemes on a uniform grid; furthermore, the *states* $\boldsymbol{x}$ of the system are the values of $\widetilde{u}$, $\widetilde{v}$ and $\widetilde{w}$ evaluated at the inner grid points.

4.2 Results

The (modified) AISIAD method is now applied to compute the balanced modes S_c and the corresponding adjoint modes S_o of the linearized plane channel flow. For the present numerical experiment we consider the following parameters: $L_x = 7$, $L_y = 2$, $Re = 900$, $\beta = 0$; 66×34 ($n_x \times n_y$) points are used to resolve the flow in the x- and the y-direction, respectively. Furthermore, an actuator and a sensor, both modeled by a Gaussian function, are placed at $n_x/4, n_y/8$ and $3n_x/4, n_y/8$, respectively, to force the system and to measure the corresponding output. Moreover, the linear system (13) is solved iteratively employing GMRES (tol=10^{-3}, maxit=30, no restarts) as implemented in MATLAB, and direct numerical simulations are performed to provide the required matrix-vector products via a matrix-free framework.

As a result, the first three computed balanced modes S_c are shown in Figure 4. The order of the reduced system was chosen as $m = 4$ and 8 block power iterations were performed to converge to a tolerance of $tol = 10^{-3}$.

5 Conclusions

In order to apply control theory to a flow it is often necessary to reduce the number of degrees of freedom of the model, whilst preserving the main input-output characteristics of the system. The way in which the model is reduced as well as the amount of reduction significantly affect a subsequent flow control. The computational cost of the model reduction algorithm must also be taken into account.

Exact balanced truncation is considered to perform the best, by balancing the observability and controllability requirements. For this reason it is used as a benchmark for comparisons of other algorithms, but it is computationally expensive. The AISIAD algorithm performs an approximation to an exact balanced truncation and shows a similar performance at far lower computational cost for the Ginzburg-Landau equation.

Similar performance for the Ginzburg-Landau equation was found using a balanced POD algorithm. This is in contrast to the relatively poor performance given by a regular POD method.

A drawback of the AISIAD algorithm, however, is that the number of degrees of freedom of the reduced model needs to be specified at the start. If the initial guess is too low, the whole process needs to be repeated. This is not the case for other model reduction algorithms such as POD or BPOD where the model can be reduced further by simply taking more snapshots.

Further drawbacks were found when the AISIAD algorithm was implanted for plane channel flow. Due to the extra complexity and greater number of parameters to tune, the AISIAD algorithm is rather difficult to converge. Upon proper tuning, however, the AISIAD algorithm gives good results at reasonable computational cost. The time taken to converge is less dependant on the cost of the DNS than for POD methods and so for DNS expensive problems the AISIAD algorithm offers significant advantages.

6 Acknowledgment

We wish to thank the organizer of the 2nd YOUNG ERCOFTAC Workshop Jörn Sesterhenn as well as the guest speakers Peter Schmid and Francois Gallaire for their thorough and patient guidance during the workshop.

Bibliography

[1] S. Bagheri, J. Hœpffner, P. S. Schmid, D. S. Henningson, *Input-Output Analysis and Control Design Applied to a Linear Model of Spatially Developing Flows*, App. Mech. Rev., In Press, 2008

[2] B. Datta,, *Numerical Methods for Linear Control Systems Design and Analysis*, Elsevier,New York,2003,

[3] S. A. Orszag, *Accurate solution of the Orr-Sommerfeld stability equation*, J. Fluid Mech. 50, p. 689-703, 1971

[4] D. Vasilyev and J. White, *A more reliable reduction algorithm for behavioral model extraction*, ICCAD '05: Proceedings of the 2005 IEEE/ACM International conference on Computer-aided design, IEEE Computer Society, 813–820, 2005

[5] C.W Rowley, *Model reduction for fluids using balanced proper orthogonal decomposition*, Int. J. Bifurc. Chaos, 15: 997–1013, 2005