BALANCED TRUNCATION MODEL REDUCTION OF LARGE AND SPARSE DESCRIPTOR LINEAR SYSTEMS

J. M. BADíA*, P. BENNER†, AND R. MAYO*, E. S. QUINTANA-ORTí*, G. QUINTANA-ORTí*, A. REMÓN*

Abstract. In this paper we describe our parallel package SpaRed for model reduction of large-scale (sparse) systems via SVD-based methods. Numerical preliminary experiments on a parallel platform illustrate the numerical and parallel performances of this approach.

Key words. Model reduction, dynamical linear systems, parallel dense and sparse linear algebra.

1. Introduction. We consider linear time-invariant (LTI) systems in generalized state-space form defined by

$$\begin{array}{rcll} E\dot{x}(t) & = & Ax(t)+Bu(t), & t>0, \quad x(0)=x^0, \\ y(t) & = & Cx(t)+Du(t), & t\geq 0, \end{array} \tag{1.1}$$

where $E, A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$, $C \in \mathbb{R}^{p\times n}$, $D \in \mathbb{R}^{p\times m}$, and $x^0 \in \mathbb{R}^n$ is the initial state of the system. The number of states n is known as the order of the system. In *model order reduction*, we are interested in finding a second system

$$\begin{array}{rcll} E_r\dot{x}_r(t) & = & A_rx_r(t)+B_ru(t), & t>0, \quad x_r(0)=x_r^0, \\ y_r(t) & = & C_rx_r(t)+D_ru(t), & t\geq 0, \end{array} \tag{1.2}$$

of order r, so that the distance between $y(t)$ and $y_r(t)$ is "small". Model reduction of large LTI systems, with n of $\mathcal{O}(10^5)$-$\mathcal{O}(10^6)$ and sparse (A, E), is necessary in weather forecast, circuit simulation, VLSI chip design, and air quality simulation, among others [1]. In these applications, usually $m, p \ll n$.

The efficacy of model reduction methods strongly relies on the problem so that no technique can be considered as globally optimal. SVD-based methods are appealing in that they usually preserve system properties and provide bounds on the approximation error. All SVD-based methods require, as the major computational stage, the solution of two Lyapunov equations. This class of methods has regained interest in the last years with the formulation of the low-rank alternating direction implicit (LR-ADI) iteration for the Lyapunov equation [2], which exploits the sparse structure of the coefficient matrices of the equation.

In this paper we describe our package for model reduction of large LTI systems, Spared (`http://www.pscom.act.uji.es/~plicmr/SpaRedW3/SpaRed.html`), which has been built using existing efficient parallel dense and sparse linear algebra libraries. Experimental results on a parallel platform illustrate the accuracy and efficacy of the methods in SpaRed, which can be used to reduce systems with n as large as $\mathcal{O}(10^5)$.

2. Numerical Methods for Model Reduction. Balanced Truncation (BT) methods are strongly related to the controllability and observability Gramians of the system, W_c and W_o respectively, given by the solutions of the *generalized Lyapunov equations*

$$\mathcal{R}_{W_c} := AW_cE^T + EW_cA^T + BB^T = 0, \tag{2.1}$$

$$\mathcal{R}_{W_o} := A^T\hat{W}_oE + E^T\hat{W}_oA + C^TC = 0, \tag{2.2}$$

and $W_o = E^T\hat{W}_oE$. Let $W_c = S^TS$ and $W_o = R^TR$ be Cholesky factorizations of the Gramians, and consider now the singular value decomposition (SVD)

$$SR^T = U\Sigma V^T = [U_1\ U_2]\begin{bmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{bmatrix}\begin{bmatrix} V_1^T \\ V_2^T \end{bmatrix}, \tag{2.3}$$

*Depto. de Ingeniería y Ciencia de Computadores, Universidad Jaume I, 12.071–Castellón, Spain ({badia,mayo,quintana,gquintan,remon}@icc.uji.es).

†Fakultät für Mathematik, Technische Universität Chemnitz, D-09107 Chemnitz, Germany (benner@mathematik.tu-chemnitz.de).

where the matrices Σ, U, and V are conformally partitioned at a given dimension r. In the *Square-Root* (SR) BT method a reduced model of order r is then obtained as

$$(E_r, A_r, B_r, C_r, D_r) := (LET, LAT, LB, CT, D), \tag{2.4}$$

where

$$L := \Sigma_1^{-1/2} V_1^T R E^{-1},\ T := S^T U_1 \Sigma_1^{-1/2}. \tag{2.5}$$

We have adapted the LR-ADI iteration in [2] for the solution of generalized Lyapunov equations. The numerical method resulting from it basically requires, at each iteration, the solution of a sparse linear system of the form $(A - \tau_j E)x = b$, where τ_j is a "cyclic" (complex) shift parameter with $\tau_j = \tau_{j+t_s}$. We have also generalized the procedure for selecting the shifts and proposed a practical convergence criterion for the iteration. The method thus obtained can exploit the sparsity of the coefficient matrices of the generalized Lyapunov equation. Furthermore, after l_c iterations, the solutions W_o, W_c are provided in the form of low-rank approximations of the Cholesky factors S, R, of dimensions $l_c m$-by-n, $l_c p$-by-n, respectively, so that the computation of the SVD in (2.3) becomes feasible.

3. Implementation and parallelization. The LR-ADI iteration is composed of linear algebra operations such as sparse matrix-vector products and the solution of sparse linear systems. Direct solvers are to be prefered because of the possibility of factorizing $(A - \tau_j E)$ once and reusing the factors in iterations $j + kt_s$, $k = 1, 2, \ldots$. Once the generalized Lyapunov equations are solved, the final stages of the BT methods require the computation of an SVD of a small dense matrix and, for the SRBT method, a few dense matrix products. Our hybrid approach for dealing with these operations is based on the combination of a coarse-grain parallel implementation for certain stages of the procedure, and the use of parallel linear algebra routines in MUMPS, ScaLAPACK, and multithreaded implementations of BLAS.

4. Preliminary Experimental Results. The following results were obtained on a cluster of 16 nodes connected via a *Myrinet* network. Each node consists of an Intel Pentium Xeon processor@2.4 GHz with 1 GByte of RAM. We employ four examples from the Oberwolfach model reduction collection corresponding to models of a steel profile (**STEEL_I** and **STEEL_II**), a microthruster array (**T3DL**), and a butterfly gyroscope (**GYRO**). The order n of these systems ranges from 20,360 to 79,841 with a sparsity degree well below 1% in all cases. The values of m, p are below 8 in all four examples.

Table 4.1 reports the number of shifts used in the LR-ADI iteration, t_s; the number of iterations required for convergence, l_c; the relative residuals for the approximations to the solutions of the Lyapunov equations; the dimension of the reduced-order system, r; the distance between the transfer function matrices of the original and reduced systems, G and G_r, respectively (a magnitude related to the distance between y and y_r); and the execution time of the parallel model reduction routine executed using 16 nodes. Although a large residual is found for one of the solutions for the equations, the distance $\|G - G_r\|_\infty$ is still satisfactory in that case. The amount of time required for model reduction of these examples is also reasonable.

Example	t_s	l_c	Resid. $\mathcal{R}_{W_c}$	Resid. $\mathcal{R}_{W_o}$	r	$\|G - G_r\|_\infty$	Time (sec.)
STEEL_I	25	76	2.94e−09	4.89e−20	45	1.54e−5	135
STEEL_II	26	78	1.80e−08	4.54e−21	60	5.29e−6	473
T3DL	29	140	4.04e+00	1.32e−12	30	6.74e−4	805
GYRO	8	50	2.27e−04	1.87e−14	30	2.33e−4	657

TABLE 4.1
Numerical performance of the parallel SRBT model reduction routine in SpaRed.

REFERENCES

[1] A. Antoulas, *Approximation of Large-Scale Dynamical Systems*, SIAM Pub., 2005.
[2] T. Penzl, *A cyclic low rank Smith method for large sparse Lyapunov equations*, SIAM J. Sci. Comput., 21 (2000), pp. 1401–1418.