

ADAPTIVE RATIONAL KRYLOV SUBSPACES FOR LARGE-SCALE DYNAMICAL SYSTEMS *

V. DRUSKIN [†], V. SIMONCINI[‡], AND M. ZASLAVSKY [§]

Abstract. The rational Krylov space is recognized as a powerful tool within Model Order Reduction techniques for linear dynamical systems. However, its success has been hindered by the lack of a parameter-free procedure, which would effectively generate the sequence of shifts used to build the space. In this paper we propose an adaptive computation of these shifts. The whole procedure only requires to inject some initial rough estimate of the spectral region of the matrix, while further information is automatically generated during the process. The approach is a generalization of the idea first proposed in [15] and it is used for two important problems in control: the approximation of the transfer function and the numerical solution of large Lyapunov equations. The procedure can be naturally extended to other related problems, such as the solution of the Sylvester equation, and parametric or higher order systems. Several numerical experiments are proposed to assess the quality of the rational projection space over its most natural competitors.

1. Introduction. The time-invariant linear system

$$\begin{aligned} \mathbf{x}'(t) &= A\mathbf{x}(t) + B\mathbf{u}(t), & \mathbf{x}(0) &= x_0 \\ y &= C\mathbf{x}(t) \end{aligned}$$

arises in a great variety of application problems such as signal processing, system and control theory. Here we assume that A is negative definite (passive), and that B and C are vectors, although the whole discussion and algorithms could be stated for B and C being matrices, at the price of some extra technical notation. A major task in the analysis of linear dynamical systems consists in reducing the continuous-time system

$$\Sigma = \left(\begin{array}{c|c} A & B \\ \hline C & \end{array} \right), \quad A \in \mathbb{C}^{n \times n}$$

onto the significantly smaller system

$$\hat{\Sigma} = \left(\begin{array}{c|c} \tilde{A} & \tilde{B} \\ \hline \tilde{C} & \end{array} \right)$$

with \tilde{A} of size $m \ll n$, while maintaining the most relevant properties of the original system Σ . We refer to [37], [8] for a discussion of the state of the art methods and for collections of challenging problems.

Various approaches have been explored in the literature, based either on invariant subspaces (see, e.g., [17], [38], [5] and references therein) or on projection spaces (see, e.g., [1] for a general presentation). Within the second class a leading role is played by Krylov subspace type methods, see also [19], [2], [11]. The standard Krylov subspace is defined as $K_m(A, b) = \text{span}\{b, Ab, \dots, A^{m-1}b\}$, with $\dim(K_m(A, b)) \leq m$; by projecting the system matrices onto this space, it is sometimes possible to obtain

*Version of January 5, 2010.

[†]Scientific Advisor, Ph.D., Schlumberger Doll Research, 1 Hampshire st. Cambridge, MA 02139 USA (druskin1@slb.com)

[‡]Dipartimento di Matematica, Università di Bologna, Piazza di Porta S. Donato 5, I-40127 Bologna, Italy and CIRSA, Ravenna, Italy (valeria@dm.unibo.it).

[§]Schlumberger Doll Research, 1 Hampshire st. Cambridge, MA 02139 USA (mzaslavsky@slb.com)

a sufficiently accurate reduced system with a moderate space dimension. Other more effective variants have been analyzed, the most general of which is given by the rational Krylov subspace: Given the possibly complex numbers s_1, s_1, s_2, \dots , this space is defined as

$$K_m(A, v, \mathbf{s}) = \text{span}\{(A - s_1 I)^{-1}v, \dots, \prod_{j=1}^m (A - s_j I)^{-1}v\} \quad (1.1)$$

where $\mathbf{s} = [s_1, s_1, \dots, s_m]$. The rational Krylov method was originally proposed by Ruhe [33] in the context of approximating interior eigenvalues, with which appropriately chosen shifts would accelerate convergence to the sought after spectral region. It is now acknowledged as being one of the most powerful projection approaches for reducing the given linear dynamical system. Major efforts have been devoted to the characterization and parameter evaluations of this space, see, e.g., [21], [22], [41],[31]. However, the success of the rational Krylov subspace has been hindered by the lack of a parameter-free procedure. In most cases information on the problem or a trial and error approach has been used to tune the choice of the shifts s_j . In the context of \mathcal{H}_2 -optimality reduction, an interesting attempt to provide an automatic selection has been recently proposed in [24]. However, the computational and memory costs of this approach have not been fully assessed. We also mention the early efforts due to Grimme [23] for determining a sequence of shifts, which however did not rely on a simple rational function like the one in (2.2).

In this paper we propose an adaptive computation of the shifts for building the rational space. The whole procedure only requires to inject some initial rough estimate of the spectral region of the matrix A , while further information is automatically generated during the process. The overall process is basically parameter-free, although performance can be even improved if accurate initial estimates are at hand. The approach is a generalization of the idea first proposed in [15], where it was used for a specific symmetric application problem, and it is investigated here for two important problems in control: the approximation of the transfer function and the numerical solution of large Lyapunov equations. The procedure is very general and can be naturally extended to other related problems, such as the solution of the Sylvester equation, and parametric or higher order systems [44],[20].

The new procedure is tested on available benchmark problems of medium and large dimensions; our experimental results show that the adaptive approach allows one to obtain a reduced system of very small dimension compared to available projection type techniques.

Throughout the paper the Euclidean norm for vectors and the induced norm for matrices is used, except when using the Frobenius norm, which is defined as $\|A\|_F^2 = \sum_{i,j} |a_{i,j}|^2$, where $a_{i,j}$ are the entries of A . Moreover, A^\top denotes the transpose of a matrix, and A^* its conjugate transpose.

2. Computation of the Rational Krylov subspace. In this section we describe the algorithm for computing an orthonormal basis and a representation matrix for A in the rational Krylov subspace, with a dynamic computation of the shifts. In the next sections the use of the computed space in two related contexts will be described.

Assume first that a sequence of parameters $\{s_2, s_3 \dots, s_{m+1}\}$ is given, together with two starting values $s_1 = s_0^{(1)}, s_0^{(2)}$. Let $D_m = \text{diag}(s_2, s_3 \dots, s_{m+1})$ be the diagonal matrix collecting these parameters. An orthonormal basis can be generated

using a standard Arnoldi process with a modified Gram-Schmidt (re)orthogonalization procedure, starting with $(A - s_1 I)^{-1}b$ and then continuing with a new shift at each iteration. We will use V_m to denote the matrix whose columns are the computed orthonormal basis vectors and $H_{m+1,m} = [H_m; h_{m+1,m}e_m^\top]$ to denote the matrix of orthogonalization coefficients.

Since products by A are not performed during the generation of the basis, the computation of the projected matrix $V_m^* A V_m$ requires some care, to avoid the computation of $O(m)$ extra matrix-vector multiplications per iteration. Different strategies have been employed, which all require an extra matrix-vector multiply. In [4] for instance, it was recently suggested to generate the basis by means of $(A - s_j I)^{-1} A v$, $j = 2, 3, \dots$. Here we follow the procedure suggested by Ruhe, see e.g. [34], where a multiplication by A is only performed to build the projected matrix.

PROPOSITION 2.1 ([34]). *Let the columns of V_m be an orthonormal basis of the rational Krylov subspace. With the notation above, it holds that*

$$T_m := V_m^* A V_m = (I_m + H_m D_m - V_m^* A v_{m+1} h_{m+1,m} e_m^\top) H_m^{-1}.$$

We are thus left with the generation of the shift sequence. Here we give the general procedure, while we postpone the details to after the algorithm description. At iteration m , a new basis vector is computed as $\tilde{v}_{m+1} = (A - s_{m+1} I)^{-1} v_m$, which requires the selection of the shift s_{m+1} at the previous iteration. Clearly, any vector $u \in K_m(A, b, \mathbf{s})$ can be written as

$$u = p_{m-1}(A) q_m(A)^{-1} b, \quad (2.1)$$

where p_{m-1} is a polynomial of degree at most $m-1$, while q_m is a fixed polynomial of degree m , whose roots are the components of \mathbf{s} . In the context of solving shifted systems of the type $(A - sI)x = b$, it was shown in [15] that for A symmetric the residual can be written as

$$b - (A - sI) V_m (T_m - sI)^{-1} V_m^* b = \frac{r_m(A)b}{r_m(s)}, \quad r_m(z) = \prod_{j=1}^m \frac{z - \lambda_j}{z - s_j}, \quad (2.2)$$

for λ_j , $j = 1, \dots, m$ being the eigenvalues of T_m . Since the characteristic polynomial of T_m minimizes $\|p(A)v\|$ among all monic polynomials of degree m , for our numerator in (2.2) we can write

$$\|r_m(A)b\| = \min_{\theta_1, \dots, \theta_m} \left\| \prod_{j=1}^m (A - \theta_j I)(A - s_j I)^{-1} b \right\|. \quad (2.3)$$

With this result in mind, it was proposed in [15] to select the next shift s_{m+2} as $s_{m+2} = \arg(\max_{s \in [\lambda_{\min}, \lambda_{\max}]} |r_m(s)|^{-1})$. Reasoning behind selection of real interpolation frequencies s_i for the approximation of the resolvent on the imaginary axis will be given later.

When working with nonsymmetric matrices, in principle we may allow both sets of interpolatory points $\{\lambda_j\}$ and $\{s_j\}$ to be complex. Obviously, (2.2) and (2.3) are also valid for nonsymmetric A , provided A and T_m do not have Jordan blocks of size larger than 1. Thus, the next shift s_{m+2} is selected as

$$s_{m+2} = \arg \left(\max_{s \in \mathcal{S}} \frac{1}{|r_m(s)|} \right)$$

for a set \mathcal{S} whose choice will be discussed in section 2.1, while the values $\{\lambda_k\}$, $k = 1, \dots, m$, are the eigenvalues of T_m .

The algorithm for the computation of the basis and representation matrix of the rational Krylov subspace can be thus summarized as follows.

Algorithm 1. Given $A, b, m_{\max}, s_0^{(1)}, s_0^{(2)}$

0. Set $s_1 = s_0^{(1)}$
1. Compute $\tilde{v}_1 = (A - s_1 I)^{-1} b$, $v_1 = \tilde{v}_1 / \|\tilde{v}_1\|$
2. Compute $t = v_1^* A v_1$
3. Select $s_2 = \text{newpole}(t, s_0^{(1)}, \{s_0^{(1)}, s_0^{(2)}, t\})$
4. For $m = 1, \dots, m_{\max}$
5. $\tilde{v}_{m+1} = (A - s_{m+1} I)^{-1} v_m$
6. Orthogonalize \tilde{v}_{m+1} against v_1, \dots, v_m ($H_{m+1,m}$ matrix of orth. coefficients)
7. Scale $v_{m+1} = \tilde{v}_{m+1} / \|\tilde{v}_{m+1}\|$, $V_{m+1} = [V_m, v_{m+1}]$
8. Compute $g = V_m^* A v_{m+1}$ and set $D_m = \text{diag}(s_2, \dots, s_{m+1})$
9. Compute $T_m = (I_m + H_m D_m - g h_{m+1,m} e_m^T) H_m^{-1}$.
10. Compute eigenvalues $\{\lambda_k\}$, $k = 1, \dots, m$ of T_m
11. Select $s_{m+2} = \text{newpole}(\{\lambda_k\}_{k=1:m}, \{s_k\}_{k=1:m}, \{s_0^{(1)}, s_0^{(2)}, s_1, \dots, s_{m+1}\})$
12. end

The function `newpole` determines the next pole for the expansion of the rational Krylov space, which will also be the extra pole of the function r_{m+1} , compared to r_m . Therefore, at each iteration the $[m/m]$ rational function r_m changes with respect to r_{m-1} as follows: the numerator polynomial may change completely, since its roots are fully recomputed, whereas the denominator polynomial only gains an extra root. The process is performed as follows:

Algorithm 2 (newpole).

Given $\{\lambda_k\}_{k=1:m}, \{s_k\}_{k=1:m}$, and the set of values $\{\eta_1, \eta_2, \dots, \eta_\ell\}$;

1. For $j = 1, \dots, \ell - 1$
2. $\sigma_j = \text{argmax}_{\sigma \in [\eta_j, \eta_{j+1}]} \frac{1}{|r_m(\sigma)|}$, $r_m(z) = \prod_{k=1}^m \frac{z - \lambda_k}{z - s_k}$
3. end
4. $s_{m+2} = \text{argmax}_{j=1, \dots, \ell-1} \frac{1}{|r_m(\sigma_j)|}$

2.1. Selection of the shifts. There are two issues that require careful consideration: the iterative selection of the shifts, and the choice of their initial values. A careful selection of the shifts, that is the poles of r_m , is what makes the whole procedure work efficiently.

It is known that standard Krylov subspace - based methods may be accurate at the local level, but are far less so in the \mathcal{H}_∞ or \mathcal{H}_2 norms, compared to methods such as balanced truncation; see, e.g., [1]. On the other hand, it was recently shown in [14] that the Rational Krylov subspace using real shifts obtained via the classical Zolotaryov solution, yields asymptotically optimal linear convergence rate for the approximation of the resolvent on the whole imaginary axis for bounded positive-definite self-adjoint operators with continuous spectrum. In particular, the following optimal convergence rate for the L_∞ error of the approximation of the resolvent on $i\mathbb{R}$ was derived:

$$O \left[\exp \left(- \frac{\pi^2 m (1 + o(1))}{2 \log \frac{4\lambda_{\max}}{\lambda_{\min}}} \right) \right], \quad \text{for } \frac{\lambda_{\max}}{\lambda_{\min}} \gg 1. \quad (2.4)$$

In [14] the authors also proposed the computation of a nested sequence of shifts (so-called equidistributed sequences or EDS), so that the asymptotically optimal rational

space with convergence given by (2.4) could be computed iteratively.

The EDS is generated as a pseudo-random sequence distributed with some density $\rho[a, b](x)$ on $[\lambda_{\min}, \lambda_{\max}]$. This density corresponds to the equilibrium charge distribution (for $x < 0$) of the condenser with positive and negative plates $[-\lambda_{\max}, -\lambda_{\min}]$ and $[\lambda_{\min}, \lambda_{\max}]$, respectively. Bound (2.4) corresponds to the worst possible scenario for a given spectral interval, since it assumes there is no possibility for essential spectral adaptation. It is exact for the cases when the spectral measures are proportional to the limiting measure of the EDS, i.e., to the above mentioned $\rho[\lambda_{\min}, \lambda_{\max}](x)$. In fact, convergence can be better for essentially different spectral measures.

The iterative approach proposed in [24] generates a sequence of Krylov subspaces with the shifts converging to those satisfying the first-order \mathcal{H}_2 -optimality conditions

$$s_i = -\lambda_i, \quad i = 1, \dots, m, \quad (2.5)$$

where λ_i are the eigenvalues of T_m . The main disadvantage of this method is that it requires the construction of many Krylov subspaces which will not be utilized in the final model (only the last subspace is used).

An adaptive strategy allowing to beat (2.4) when possible was proposed in [15], again for A symmetric. In there, the next (real) pole s_{m+2} is chosen to be the value that yields the maximum of $|r_m|^{-1}$ over the mirrored image of the real spectral interval of A ; this way, the next function will have damped values around the new pole. The discussion above justifies the use of real shifts for A symmetric; see also the discussion at the beginning of section 3. It was experimentally shown that such an approach always gives linear convergence rate that is at least as good as the one using the EDS points of [14], although significant speedups were observed for nonuniform spectral distributions.

The problem remains open as of which shifts should be used for A nonsymmetric with a complex spectrum. As a generalization to the real case, we can either consider real shifts, taken over the mirrored interval of the eigenvalue real parts, or complex shifts taken, say, over the mirrored convex hull of the complex eigenvalues. Algorithm 1 is stated in terms of real shifts, therefore we take η_j to be the (increasingly ordered) previously chosen shifts, although Algorithm 2 also holds for complex intervals. Our decision is mainly based on computational ground, since solving a complex shifted system may be significantly more expensive than solving a real one. Nonetheless, on certain problems the use of complex shifts may be beneficial, at least in terms of number of iterations, as the left plot of Figure 2.1 shows (data from the FOM set of Example 3.1). The plot reports the \mathcal{H}_∞ norm of the error for approximating the transfer function as the space dimension increases, as discussed in the next section.

We also recall that whenever data are real and shifts are complex *conjugate*, it is possible to maintain a real basis and projection matrix T_m , although the expensive solution of the linear system still requires complex arithmetic; we refer to [34] for a possible implementation.

Our final considerations of this section concern the choice of the initial values $s_0^{(1)}, s_0^{(2)}$. These values are used in the determination of the spectral region of A to generate the set of points used to compute the next shift in `newpole`. The smallest value is also used as initial pole, $s_1 = s_0^{(1)}$. In our experiments, we selected $s_0^{(1)}, s_0^{(2)}$ to be a rough estimate of the eigenvalues of A closest to and farthest from the origin, respectively, with $s_0^{(1)}$ usually being real. If $s_0^{(2)}$ is complex, then either its real part is retained, or in case of complex shifts its complex conjugate may also be added. Some problems appeared to be sensitive to the accuracy of the initial values as estimates

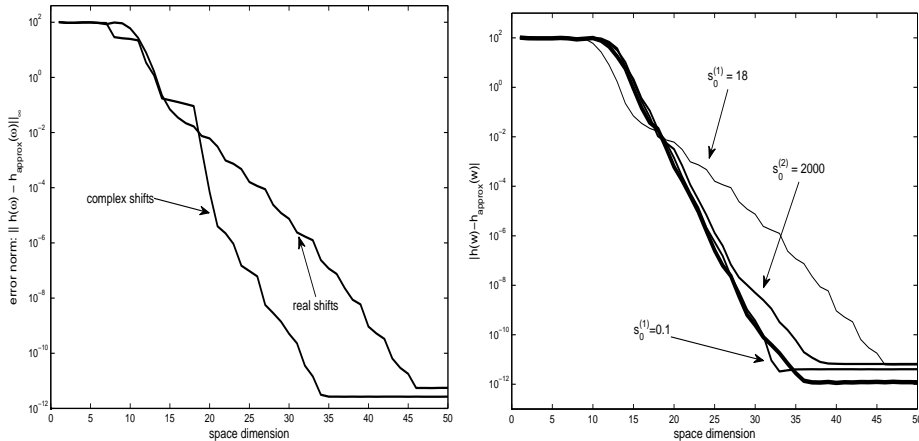


FIG. 2.1. FOM data set. Matrix A is normal, with $\Re(\lambda) \in [1, 1000]$. Left: Error in the \mathcal{H}_∞ -norm for real and complex shifts. Right: Dependence of \mathcal{H}_∞ on the choice of starting parameters $s_0^{(1)} \approx 1$, $s_0^{(2)} \approx 1000$ when any of them is moved away from the exact value. The thickest curve is for both parameters equal to the exact values $s_0^{(1)} = 1$, $s_0^{(2)} = 1000$.

of the spectral region, therefore any relevant a-priori knowledge could be exploited here. As an example, the right plot of Figure 2.1 reports the \mathcal{H}_∞ error norm as a function of the space dimension, for different starting values $s_0^{(1)}$, $s_0^{(2)}$ for the FOM set. The matrix A there is normal, with $\Re(\lambda) \in [1, 1000]$, and the thickest curve corresponds to the selection $s_0^{(1)} = 1$, $s_0^{(2)} = 1000$. Note that the value $s_0^{(1)} = 18$ was obtained with the following rough estimate $s_0^{(1)} = \text{cond}(A)/\|A\|_F$, which is the approximation used in all our tests. The method seems to perform less well if the real part of the spectral interval is not captured. It is also interesting that the use of complex shifts would allow one to recover the best observed convergence rate (cf. left plot of Figure 2.1). More details on the actual computation of $s_0^{(1)}$, $s_0^{(2)}$ may be found in the experiment sections.

In the next sections we analyze two specific problems that are tightly related to the approximation of the resolvent of the given linear dynamical system. We report on the performance of Galerkin-type projection onto the Rational Krylov subspace with adaptive choice of the shifts.

3. Evaluation of the transfer function. We consider the approximation of the transfer function

$$h(\omega) = c(A - \omega iE)^{-1}b, \quad \omega \in \mathbb{R}. \quad (3.1)$$

with A negative (semi-)definite and E symmetric positive definite. Let h_{approx} be a reduced transfer function, the \mathcal{H}_∞ -norm of the error is given as (cf., e.g., [1, sec.5.3])

$$\|h - h_{approx}\|_{\mathcal{H}_\infty} = \sup_{\omega \in \mathbb{R}} |h(\omega) - h_{approx}(\omega)|.$$

The adaptive procedure in Algorithm 1 appears to numerically maintain the high convergence rate obtained by the theoretical Rational Krylov space with A symmetric and a-priori computed weights, as described in [14]; see also [15]. As an example, we

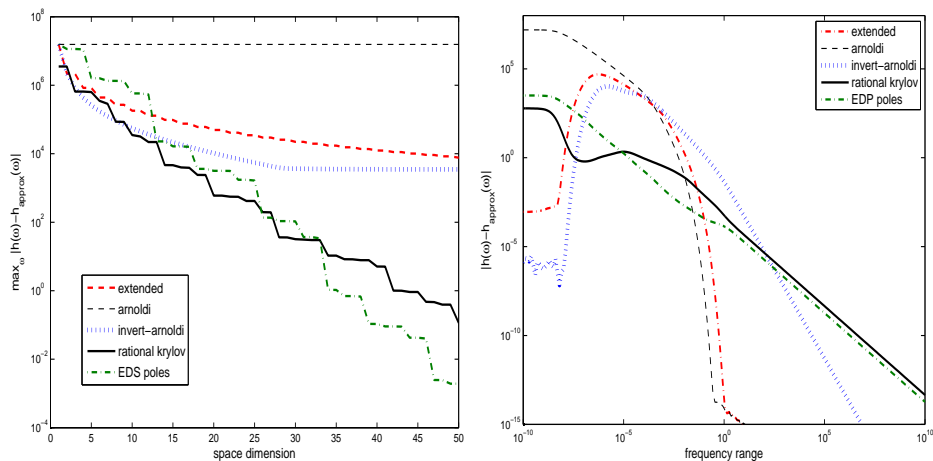


FIG. 2.2. Diagonal matrix with log-linearly distributed values in spectral interval. Error \mathcal{H}_∞ -norm (left), and $|h(\omega) - h_{\text{approx}}(\omega)|$. (right) for $m = 20$. For Rational Krylov, the initial parameters were estimated as $s_0^{(1)} = 1.6086 \cdot 10^{-8}$, $s_0^{(2)} = 1$.

consider a diagonal matrix of size 900 with log-uniformly distributed (eigen)values in $[-1, -3.3164 \cdot 10^{-9}]$, $b = c^\top = \mathbf{1}$. The left plot of Figure 2.2 shows the error \mathcal{H}_∞ -norm in (3.1) for various projection methods as the space dimension increases: Arnoldi (using $K_m(A, b)$) Inverted-Arnoldi (using $K_m(A^{-1}, b)$), see, e.g., [1] for their description in context), Extended Krylov (using $\mathbf{K}_m(A, b) = K_m(A, b) + K_m(A^{-1}, A^{-1}b)$) [13], [39],[40], Rational Krylov space method with a-priori EquiDistributed Sequences of shifts (dash-dotted, only for theoretical purposes) and adaptive (solid) shifts.

The latter two spaces are the only ones that maintain a good convergence rate, in spite of the large condition number and densely distributed spectrum of the matrix. In particular, we remark that the log-uniform spectral distribution is asymptotically close to $\rho[\lambda_{\min}, \lambda_{\max}](x)$ for large condition numbers [26], i.e., it is asymptotically optimal for the theoretical ‘‘EDS point’’-based method. However the adaptive shifts are so well chosen that no significant difference is observed with respect to the former.

For the sake of completeness, in the right plot of Figure 2.2 we also report the behavior of the methods for $\omega \in [10^{-10}, 10^{10}]$ and $m = 20$, showing that the other projection methods may behave well locally, cf. the frequency range $[10^0, 10^4]$, but perform worse over the whole imaginary axis.

It is also interesting to remark that for spectral measures significantly different from $\rho[\lambda_{\min}, \lambda_{\max}](x)$, the adaptive algorithm is observed to show a faster convergence rate than that with the EDS points. This fact is confirmed by the convergence history in Figure 3.1, where the diagonal matrix A now consists of *linearly* distributed (eigen)values in $[-1, -3.3164 \cdot 10^{-9}]$.

The theory in [14] does not cover the case of nonsymmetric A with complex spectrum. However, we hope that for moderate imaginary parts, the convergence behavior will not be very different. As in Algorithm 2, we consider computing the next shift by maximizing $1/|r_m|$ over certain intervals, where the zeros of r_m are rational Ritz values, while the poles of r_m are the previously computed shifts. In both plots of Figure 3.2 we show the \mathcal{H}_∞ -norm of the error of all methods versus the space dimension, for the matrix A of size 1600×1600 , stemming from the centered

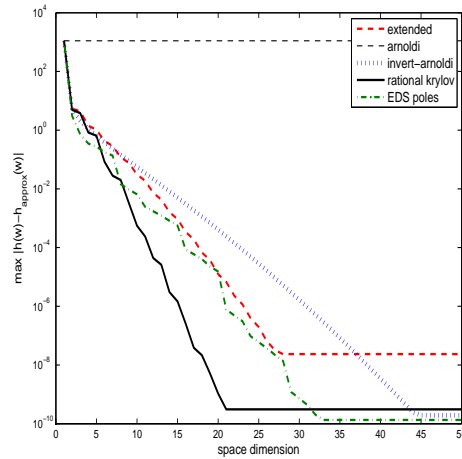


FIG. 3.1. Diagonal matrix with linearly distributed values in spectral interval. Error \mathcal{H}_{∞} -norm.

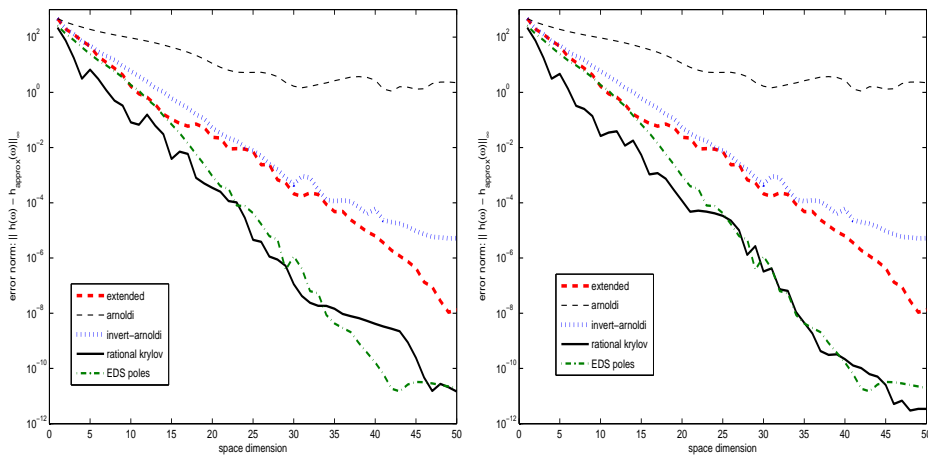


FIG. 3.2. Nonsymmetric matrix. Convergence history versus space dimension. Nonsymmetric matrix. Left: Rational Krylov subspace method with real shifts. Right: Rational Krylov subspace method with complex shifts.

finite difference discretization of the operator

$$\mathcal{L}(u) = (\exp(-xy)u_x)_x + (\exp(xy)u_y)_y - (10(x+y)u)_x$$

on the unit square with Dirichlet homogeneous boundary conditions. The eigenvalues of A are contained in the complex box $[-10.583, -0.023950] \times [-0.075561, 0.075561]$. The solid line reports the accuracy of the Rational Krylov subspace method on the whole imaginary axis as the space dimension increases, and it is the only curve changing in the two plots. The left plot shows the use of *real* shifts, computed in the reference interval of the Ritz value real part, whereas the right plot shows the use of *complex* shifts over the convex hull of the rational Ritz values. For theoretical purposes, we consider generalizing the idea of EDS shifts to the complex case. As it was mentioned already, the EDS for symmetric negative definite problems is a (randomized) approximation of the equilibrium charge (in the left complex half plane) of

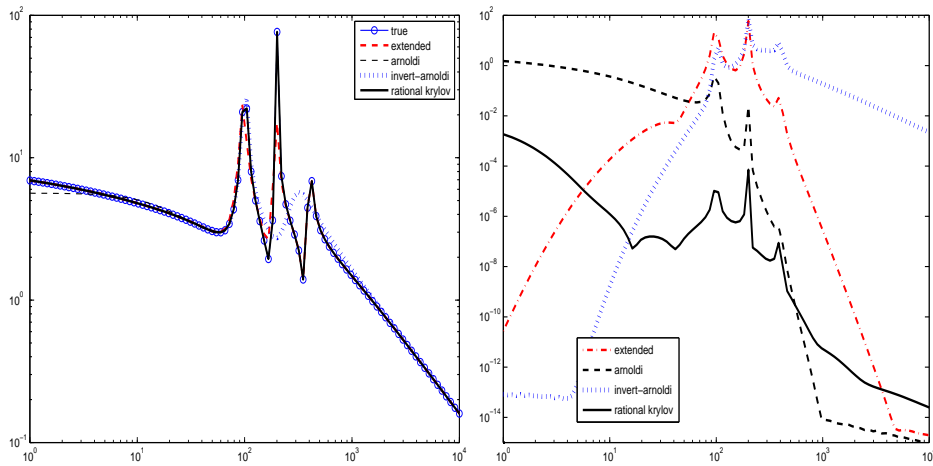
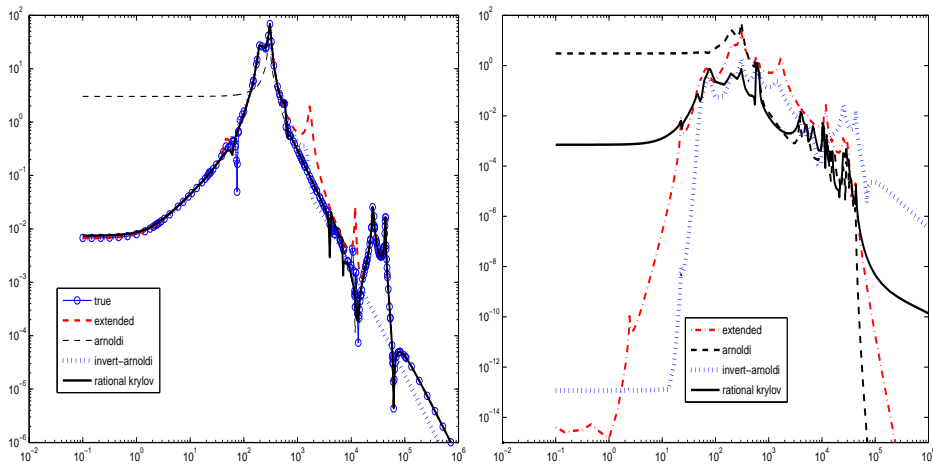
TABLE 3.1
Relevant data information for the test problems

Pb	n	$\ A\ _F$	$\text{condest}(A)$	$\ E\ _F$	$\text{condest}(E)$	A sym
FOM	1006	1.8283e+04	1000	Id	Id	no
CDPlayer	120	2.3095e+05	1.8149e+04	Id	Id	no
FILTER	1668	2.3136e-03	7.2321e+04	4.6926e-09	67.285	yes
EADY	598	1.2660e+02	5.3727e+02	Id	Id	no
FLOW	9669	2.5438e+04	1.6193e+07	6.8072	1.0331e+06	no

the plane condenser with the positive and negative plates being respectively the A 's spectral interval and its mirrored reflection with respect to the imaginary axis. The spectral interval of a symmetric matrix is its numerical range, so a natural extension to a nonsymmetric matrix is to consider its numerical range and its mirrored reflection as condenser's plates. Following this reasoning, a possible rough approximation of the a-priori reference shifts in the numerical range is computed by using 50 points on the spectral convex hull of A : these are taken to be the vertexes of the convex hull of the complex spectrum of A , expanded by a few more points on the hull's edges so as to account for 50 points. The convergence of the rational Krylov method with these precomputed shifts is also reported in the plots of Figure 3.2, as dash-dotted line. The similarity with the adaptive method in both cases is striking. Indeed, finding such a-priori shifts requires the solution of the full eigen-problem for A , whereas computing the adaptive shifts involves negligible additional cost.

3.1. Numerical experiments. In the following we report on our experience with various symmetric and nonsymmetric benchmark problems, whose main properties are reported in Table 3.1. For each test case, we show the Bodeplot of the true transfer function and of its approximations, and also the Bodeplot of the error in the frequency range of interest (in most cases, this range is part of the data set), keeping in mind that this latter plot only reports on the local behavior of the methods. As in the previous test, we compare the performance of the Standard Arnoldi, Inverted-Arnoldi, Extended Krylov subspace method, with that of the new adaptive Rational Krylov subspace method. Test cases include both symmetric and nonsymmetric matrices, some of which are hard to reduce in spite of the small size. To simplify the presentation, we consider approximating the transfer function around zero.

EXAMPLE 3.1. We consider a few benchmark problems, from the NICONET repository ([30]) and the Oberwolfach collection [12]. All relevant information on the problems are reported in Table 3.1. The performance of the methods is shown in the plots of Figure 3.3-3.6, which report the Bode plot of the transfer function (left plots) and the associated errors (right plots) for the reference (small) value $m = 20$. In all cases, the initial values $s_0^{(1)}, s_0^{(2)}$ were estimated as $s_0^{(1)} = \|A\|_F / \text{condest}(A)$ and $s_0^{(2)}$ as the eigenvalue with largest magnitude of $(-A, E)$ (using Matlab's `eigs`), with a very rough tolerance of 10^{-3} . In all cases, the matrices provided in the data set were used; otherwise, we set $c = b^\top$ and $E = I$ where needed. In almost all cases, the frequency range was also provided. For the CDplayer example, the second input and first output were used, otherwise the first input and output were selected. All results consistently confirm the good performance of the rational Krylov subspace method, for a given final reduced space dimension.

FIG. 3.3. *Example 3.1. FOM data set.*FIG. 3.4. *Example 3.1. CDPlayer data set.*

We next report on the performance of the rational Krylov subspace as a reduction space on a larger problem.

EXAMPLE 3.2. We consider the nonsymmetric linear system stemming from a 2D convective thermal flow problem (set `flow_meter_model_v0.5` in [12]). The involved matrices have size 9669 and we consider the first output vector. The starting shifts $s_0^{(1)}, s_0^{(2)}$ were approximated as before. The results of the approximation for $m = 80$ are shown in the plots of Figure 3.7, except for the Arnoldi method, which did not deliver a useful result because the output vector c appears to be orthogonal to the generated space. The error plot fully confirms our previous findings, showing a consistent error reduction throughout the considered frequency range.

4. The Lyapunov equation. We consider the numerical approximation to the solution of the algebraic equation

$$AX + XA^\top + bb^\top = 0.$$

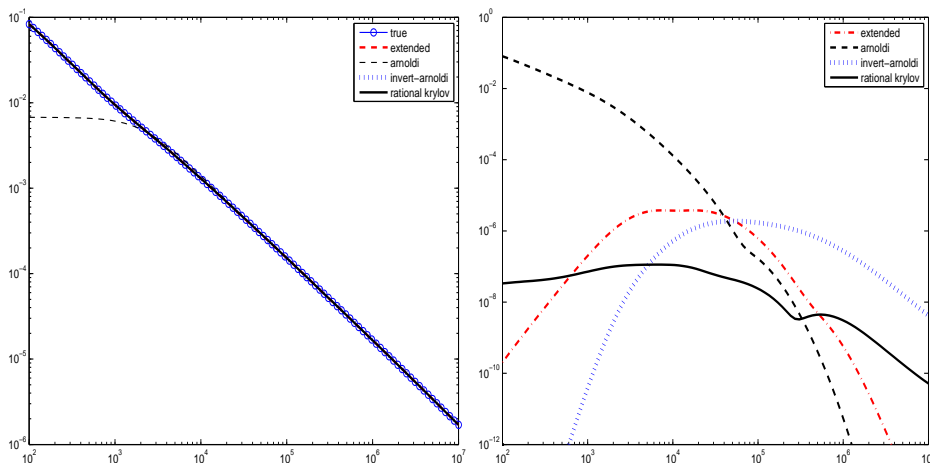


FIG. 3.5. Example 3.1. FILTER data set.

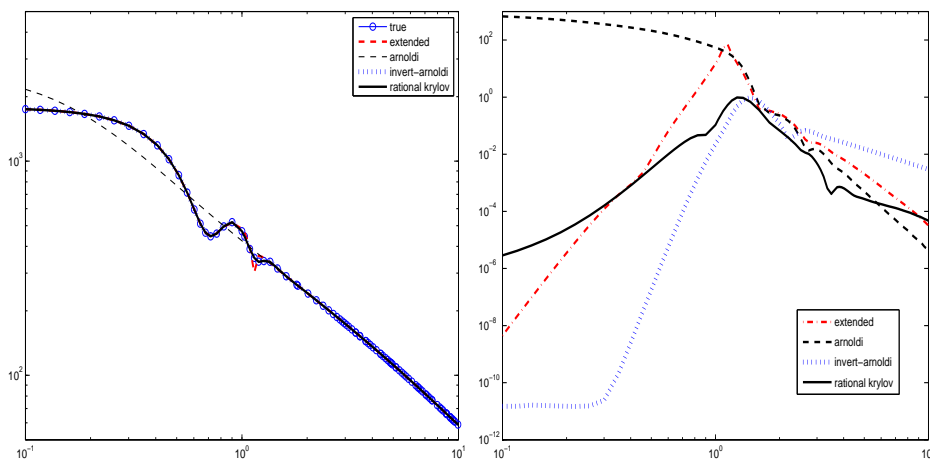
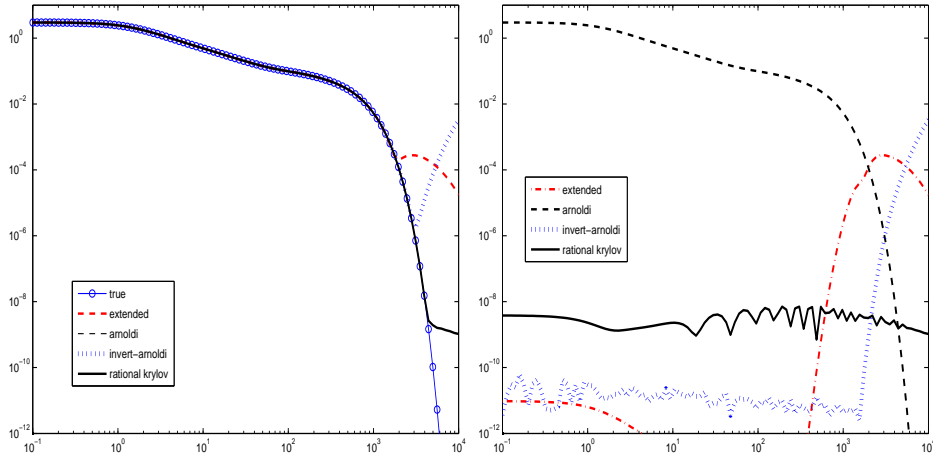


FIG. 3.6. Example 3.1. EADY data set.

Consider a subspace \mathcal{K} of \mathbb{R}^n , and let V be a matrix whose orthonormal columns span \mathcal{K} . We project the Lyapunov equation onto this subspace \mathcal{K} : we set $H = V^\top AV$, and seek an approximate solution in \mathcal{K} in the form $\hat{X} = VYV^\top$. The projected problem may be obtained by imposing that the residual $R = A\hat{X} + \hat{X}A^\top + bb^\top$ be orthogonal to the approximation space \mathcal{K} , the so-called Galerkin condition. In matrix terms, this can be written as $V^\top RV = 0$. Since V has orthonormal columns, this yields the following projected Lyapunov equation

$$V^\top AVY + YV^\top A^\top V + V^\top bb^\top V = 0, \quad (4.1)$$

whose solution Y defines \hat{X} . Since we assume that A is dissipative, $V^\top AV$ is stable and the equation (4.1) admits a unique solution. Since Y is symmetric and positive definite, the approximate solution can be written as $\hat{X} = ZZ^\top$ and thus only Z needs to be stored, where the number of columns of Z corresponds to the number of eigenvalues of Y that are larger than, say, the available machine accuracy.

FIG. 3.7. *Example 3.2. Flow problem, size $n = 9669$.*

The projection is effective if a good approximation is attained for a small dimensional subspace, so that (4.1) can be cheaply solved by means of procedures based on matrix factorizations; cf. [3], [25], [28]. In turn, this effectiveness is more likely to take place whenever the exact solution can be accurately approximated by low rank matrices. Projection methods explored in the literature differ in the way the space \mathcal{K} , and in some cases the matrix V , is chosen, and these include the standard Krylov subspace $K_m(A, b)$ first proposed in [35], and the much more recent Extended Krylov subspace $\mathbf{K}_m(A, b) = K_m(A, b) + K_m(A^{-1}, A^{-1}b)$ [39]. A related approach is the cyclic low-rank Smith method (see, e.g., [32]), which is an effective implementation of the classical ADI method [18],[29]. A lot of effort has been recently put in the development of a parameter-free and computationally effective cyclic Smith method; see, e.g., [6] and references therein. It is important to realize that the cyclic Smith method determines the solution in a rational Krylov subspace, but without imposing an orthogonalization condition. The effectiveness of the computation in the Smith method only relies on the closeness of the employed parameters to the optimal ones, for which a fast asymptotic convergence rate can be analytically derived. If such an accuracy is not attained, then a much larger approximation space for the Smith process may be required, also yielding a possibly large approximate solution factor. This weakness in the parameter selection makes the ADI-type approach inferior to the Extended Krylov subspace method, in terms of memory and computational time, both in the symmetric and nonsymmetric cases; see, e.g., the numerical experiments reported in [39].

In this paper we want to show that the adaptive computation of the parameters in Algorithm 2 allows us to generate an effective rational Krylov subspace, which may compete with the Extended Krylov subspace. In particular, our experiments show that the generated approximation space is always (much) smaller than that computed by the Extended Krylov subspace. The final space dimension is probably the most explicit indicator of the performance of the Rational Krylov method compared to the Extended one. The latter requires a system solve at every other iteration, and the coefficient matrix is the same throughout the computation, therefore either the matrix can be factorized once for all, or an effective preconditioner can be precomputed. The

rational Krylov space requires a different system solve at each iteration, since the shifted matrix changes, so that the number of solves equals the space dimension. All other computations are comparable when A is nonsymmetric. As a result, for a fixed space dimension, the generation of the rational space is more expensive than that of the extended space. Therefore, the rational Krylov subspace is competitive only if it can generate a sufficiently good approximation earlier, that is with a smaller space, than with the Extended Krylov space. We have seen in previous sections that this may be the case for the transfer function. In section 4.1 we show that this is so also for some problems in this context.

The use of the rational Krylov subspace has an extra theoretical motivation. Let $\widehat{X}_m = VYV^*$ be an approximate solution, where Y solves the reduced equation (4.1). Let $T = V^*AV$ and $e = V^*b$. Then using the analytic expression for the solution X (see, e.g., [1]), we can write

$$X - \widehat{X}_m = \int_{\mathbb{R}} (x(\omega)x(\omega)^* - \widehat{x}(\omega)\widehat{x}(\omega)^*) d\omega$$

with $x(\omega) = (A - \omega iI)^{-1}b$, $\widehat{x}(\omega) = V(T - \omega iI)^{-1}e$.

This relation emphasizes that the resolvent function $A - \omega iI$ plays a crucial role also in the approximation to the Lyapunov equation, and in particular, we can expect that the behavior of the error $x(\omega) - \widehat{x}(\omega)$ for $\omega \in \mathbb{R}$ influences that of $X - \widehat{X}_m$. A convergence analysis of the rational Krylov method for the Lyapunov equation is currently in progress [27].

In previous sections we defined the rational Krylov subspace as (1.1). However, the inclusion of the vector b may be beneficial, and thus for the computation associated with the Lyapunov equation we use

$$K_m(A, b, s) = \text{span}\{b, (A - s_2I)^{-1}b, \dots, \prod_{j=2}^m (A - s_jI)^{-1}b\}. \quad (4.2)$$

In particular, if $\|b\| = 1$ and b is the vector used to generate the space, then $b = Ve_1$. Thus, $s_1 = s_0^{(1)}$ is not used to build the space. Note however that for $b = (A - s_1I)^{-1}v$ for some v and s_1 , the original space is recovered. In the following we shall consider real shifts, therefore all computation is performed in real arithmetic.

The following proposition shows that the residual norm can be computed at low cost.

PROPOSITION 4.1. *Let \mathcal{K} be the rational Krylov subspace in (4.2) of dimension m , and $V_m \in \mathbb{R}^{n \times m}$ contain an orthonormal basis of \mathcal{K} . Let $\widehat{X}_m = V_m Y_m V_m^T$ be the approximate solution obtained in \mathcal{K} . With the same notation as in Proposition 2.1, the residual $R_m = A\widehat{X}_m + \widehat{X}_m A^T + bb^T$ satisfies*

$$\|R_m\|_F = \|SJS^T\|_F, \quad J = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

and S is the 3×3 upper triangular matrix in the “skinny” QR factorization of

$$U = [v_{m+1}s_{m+1}, V_m Y H_m^{-T} e_m h_{m+1,m}, -(I - V_m V_m^T) A v_{m+1}].$$

Proof. Starting from $(A - s_{j+1}I)^{-1}v_j = V_{j+1}H_{1:j+1,j}$, $j = 1, \dots, m$ and using the same derivation as in [34] we can write

$$AV_m = V_m T_m + v_{m+1} h_{m+1,m} e_m^\top D_m H_m^{-1} - (I - V_m V_m^\top) A v_{m+1} h e_m^\top H_m^{-1}.$$

Therefore, using the fact that Y_m solves the reduced equation (4.1), and letting $g = (I - V_m V_m^\top) A v_{m+1}$,

$$\begin{aligned} R &= AV_m Y_m V_m^\top + V_m Y_m V_m^\top A^\top + b b^\top \\ &= v_{m+1} h e_m^\top D_m H_m^{-1} Y_m V_m^\top - g h_{m+1,m} e_m^\top H_m^{-1} Y_m V_m^\top \\ &\quad + V_m Y_m H_m^{-\top} D_m^\top e^\top \bar{h}_{m+1} v_{m+1}^\top + V_m Y_m H_m^{-\top} e_m \bar{h}_{m+1,m} g^\top \\ &= U J U^\top, \end{aligned}$$

and using the QR factorization of U the result follows. \square

Proposition 4.1 indicates that the computation of the residual norm requires the computation of the vector $(I - V_m V_m^\top) A v_{m+1} = A v_{m+1} - V_m (V_m^\top A v_{m+1})$. This may be performed by using a matrix-vector product, m inner products and m vector sums. For large problems we can assume that this cost is significantly lower than a system solve with the shifted matrix $(A - s_j I)$. We also observe that if b were not included in the space, the computation of the residual could still be performed without explicitly forming the approximate solution, although the procedure would require the factorization of a slightly larger (5×5) matrix.

The Lyapunov equation is usually stated with the right-hand side being of rank larger than one, that is as BB^\top rather than bb^\top . The Rational Krylov subspace method can be generalized in a straightforward manner to this setting, by generating the block space

$$K_m(A, B, \mathbf{s}) = \text{span}\{B, (A - s_2 I)^{-1}B, \dots, \prod_{j=2}^m (A - s_j I)^{-1}B\}.$$

The changes in the implementation are only technical, and are typical of block Krylov subspace methods; see, e.g., [36].

4.1. Numerical experiments. In the following we report on some computational evidence of the performance of the application of the rational Krylov subspace for solving the Lyapunov equation associated with the reduction of the linear dynamical equation $E\mathbf{x}' = A\mathbf{x} + B\mathbf{u}$, $B = b$. In light of the results in [39] the direct competitor is the Extended Krylov subspace method (called K-PIK in [39]). Since we aim at attacking large scale problems, we propose solving the linear systems by means of either a direct or an iterative method. Timings for both approaches are shown for comparison.

Due to the presence of the symmetric matrix E , the problem actually solved is

$$AXE + EXA^\top + b b^\top = 0.$$

We assume that E is positive definite, therefore letting $E = LL^\top$, we can write the equation as

$$(L^{-1}AL^{-\top})\tilde{X} + L^{-1}\tilde{X}(L^{-1}A^\top L^{-\top}) + L^{-1}b b^\top L^{-\top} = 0.$$

Systems with the shifted matrix can use the equality $(L^{-1}AL^{-\top} - s_j I) = L^{-1}(A - s_j E)L^{-\top}$, so that a system with $(A - s_j E)$ needs to be solved. Note that in many

TABLE 4.1

Example 4.2. Performance of methods with the RAIL (symmetric) matrices. Cholesky preconditioning was used with threshold 10^{-2} for the inner iterative solver, when required.

n		Rational space direct	Extended space direct	Rational space iterative	Extended space iterative
1357	CPU time (s)	0.84	0.36	0.96	1.60
	dim. Approx. Space	21	64	21	68
	Rank of Solution	21	47	21	45
20209	CPU time (s)	11.19	10.97	25.31	201.94
	dim. Approx. Space	25	124	25	126
	Rank of Solution	25	75	25	75
79841	CPU time (s)	51.54	73.03	189.48	2779.95
	dim. Approx. Space	26	168	26	170
	Rank of Solution	26	103	26	98

problems, either E has a very sparse structure (diagonal or block diagonal), or it has a sparsity pattern very similar to that of A .

For all methods, the stopping criterion is based on the backward error as

$$\frac{\|R\|_F}{\|b\|^2 + \|E^{-1}\|_F \|A\|_F \|\widehat{X}_m\|} < \text{tol},$$

where the residual norm $\|R_m\|_F$ is cheaply computed (cf. Proposition 4.1 for the Rational Krylov subspace method), $\|\widehat{X}_m\| = \|Y_m\|$, while $\|E^{-1}\|_F \approx \text{cond}(E)/\|E\|_F$, and the remaining F-norms are computed explicitly at the beginning of the process. All these computations are included in the final timings. In all experiments we used $\text{tol} = 10^{-10}$ and an inner tolerance of $\text{tol}_{\text{inner}} = 10^{-11}$ when the inner systems were solved by an iterative method. For the rational Krylov space, the computation of the starting values $s_0^{(1)}, s_0^{(2)}$ required a low percentage of the whole computation, usually far below 10%, and it is not included.

The rank of the final approximate solution is obtained by evaluating the number of eigenvalues of Y_m that are larger than 10^{-12} .

We should mention that the Extended Krylov subspace method exploits the possible symmetry of the problem, so that only *local* orthogonality needs to be maintained in this case, yielding a short-term recurrence. Nonetheless, the subspace basis needs to be saved to recover the final approximate solution. The factorization of the matrix (or preconditioner) for solving with A at each iteration is computed once for all at the beginning of the process.

All results were obtained with Matlab 7.4 (R2007a) on a Machine running linux with 2Gb memory.

EXAMPLE 4.2. These data correspond to various mesh resolutions of the RAIL Benchmark problem in [12]. The involved matrices A , E are symmetric and of size 1357, 20209 and 79841, for the three available meshes respectively. The two matrices have apparently the same pattern, therefore the same `symamd` reordering of the entries in both matrices was performed prior to computation. The PCG method was used as inner iterative solver, with incomplete Cholesky preconditioning with threshold 10^{-2} .

The high cost of the iterative version of the Extended Krylov method is due to the complexity of the inner preconditioned solver: for the larger data matrices, it takes

TABLE 4.2

Example 4.3. Performance of methods with the FLOW and CHIP matrices. ILU preconditioning was used with threshold 10^{-2} for the inner iterative solver, when required.

n		Rational space direct	Extended space direct	Rational space iterative	Extended space iterative
9669	CPU time (s)	3.16	3.06	3.01	9.95
	dim. Approx. Space	16	36	16	36
	Rank of Solution	16	24	16	24
20082	CPU time (s)	59.99	45.84	13.01	25.28
	dim. Approx. Space	15	26	15	26
	Rank of Solution	15	22	15	22

230 iterations on average to give a sufficiently accurate solution, whereas the inner solver in the Rational Krylov method requires a widely varying number of iterations, depending on the chosen shift, ranging from 7 to 224 iterations, with the large majority of the systems solved in very few iterations. The fact that shifted systems (away from the origin) are solved in the latter method makes the whole procedure significantly cheaper than the Extended Krylov method. However, for this type of problem direct solves outperform iterative ones.

The Rational Krylov space method is the method of choice for the largest test case, fully overcoming the high costs due to the new solves at each iteration, and the orthogonalization of the whole space. It is particularly remarkable that the approximation space does not appreciably increase with the problem dimension, suggesting an extremely welcome “mesh independence” characterization. Moreover, the low space dimension also provides a low rank approximate solution, which in turn ensures further memory savings compared to the other method. This feature is common to all experiments we have tested.

EXAMPLE 4.3. We consider the pairs (A, E) of matrices with A nonsymmetric in the Oberwolfach Benchmark set stemming from convective thermal flow problems. We label the two matrices FLOW (of size 9669, these are the same data as in Example 3.2) and CHIP (of size 20082). We used IDR [42, 43] as inner iterative solver with an incomplete LU decomposition with threshold 10^{-2} as preconditioner. The performance of the methods is reported in Table 4.2. The iterative version of the rational Krylov method is the best performing approach, in particular on the larger case. The inner system requires only between 7 and 34 iterations to converge (with most systems stopped after much less than 34 its), whereas the iterative version of the Extended Krylov space needs always between 33 and 37 its to converge for the smaller problem, and between 44 and 45 its for the larger problem. These digits justify the good performance of the inner iterative methods, with the Rational Krylov subspace one providing the best performance.

EXAMPLE 4.4. We consider the centered finite difference discretization of the following operator:

$$\mathcal{L}(u) = (\exp(-10xy)u_x)_x + (\exp(10xy)u_y)_y - (10(x+y)u)_x \quad (4.3)$$

on the unit square and homogeneous Dirichlet boundary conditions. We consider two discretization meshes, giving rise to matrices of size 10 000 and 160 000, respectively. The vector b is the vector of all ones, normalized. Once again, we used a stopping

TABLE 4.3

Example 4.4. Performance of methods with the matrix associated with the elliptic operator in (4.3). AMG preconditioning was used for an inner solver, when required.

n		Rational space direct	Extended space direct	Rational space iterative	Extended space iterative
10 000	CPU time (s)	7.78	7.21	6.14	18.83
	dim. Approx. Space	29	162	29	146
	Rank of Solution	27	40	27	39
160 000	CPU time (s)	770.71	-	399.76	-
	dim. Approx. Space	74	>300	74	>300
	Rank of Solution	57	-	57	-

tolerance of 10^{-10} and inner tolerance 10^{-11} , and we preconditioned the inner iterative method (GMRES) with an Algebraic Multigrid preconditioner [9]. This is a difficult test problem, as all methods displayed slow convergence, as reported in Table 4.3. For the larger size problem, the Extended Krylov subspace method stopped with a subspace of dimension 306 with an “out of memory” Matlab breakdown, although the method was (slowly) converging, with a relative residual norm of $3 \cdot 10^{-10}$. The iterative version of the rational Krylov subspace method yields the best performance, in terms of both approximation space dimension and CPU Time.

5. Conclusions. We have proposed a new adaptive procedure for interactively compute the shifts of the rational Krylov space. The motivation for using the rational Krylov subspace is well documented in the literature, although the procedure has been lacking a satisfactory computation of the parameters. In this paper we have shown that the adaptive computation of these parameters makes the rational space very appealing for the dimension reduction of dynamical systems, yielding the best observed accuracy for very low space dimensions. For the solution of the Lyapunov equation, the rational Krylov subspace may be appealing when the overall cost of dealing with the involved shifted systems is kept low, while always delivering an approximate solution of rank (much) lower than that of the Extended space.

The adaptive rational Krylov subspace can be used to solve other problems, such as parameterized and nonlinear systems (see, e.g., [44],[10],[7]) and the approximation of high order transfer functions (see, e.g., [16]). We plan to explore these extensions in future work.

Acknowledgments. This work started while the second author was visiting the Schlumberger Doll Research Center in Cambridge (MA). The warm hospitality received is greatly acknowledged. All reported timings were obtained using the computer facilities of the SINCEM Laboratory at CIRSA, Ravenna.

REFERENCES

- [1] A. C. Antoulas. *Approximation of large-scale Dynamical Systems*. Advances in Design and Control. SIAM, Philadelphia, 2005.
- [2] Z. Bai. Krylov subspace techniques for reduced-order modeling of large-scale dynamical systems. *Applied Numerical Mathematics*, 43:2–44, 2002.
- [3] R. H. Bartels and G. W. Stewart. Algorithm 432: Solution of the Matrix Equation $AX+XB=C$. *Comm. of the ACM*, 15(9):820–826, 1972.

- [4] B. Beckermann and L. Reichel. Error estimation and evaluation of matrix functions via the Faber transform. *SIAM J. Numer. Anal.*, 47(5):3849–3883, 2009.
- [5] P. Benner. A MATLAB Repository for Model Reduction Based on Spectral Projection. In *Proceedings of the 2006 IEEE Conference on Computer Aided Control Systems Design*, pages 19–24. IEEE Computer Society, 2006.
- [6] P. Benner. *Control Theory*, chapter 57. Chapman & Hall/CRC, 2006. Handbook of Linear Algebra.
- [7] P. Benner. Model Order Reduction for Systems with Non-Rational Transfer Function Arising in Computational Electromagnetics. In *Scientific Computing in Electrical Engineering SCEE 2008, Mathematics in Industry*, volume 14. Springer-Verlag, Berlin/Heidelberg, 2009. To appear.
- [8] P. Benner, V. Mehrmann, and D. Sorensen (eds). *Dimension Reduction of Large-Scale Systems*. Lecture Notes in Computational Science and Engineering. Springer-Verlag, Berlin/Heidelberg, 2005.
- [9] J. Boyle, M. D. Mihajlović, and J.A. Scott. HSL_MI20: an efficient AMG preconditioner. Technical Report RAL-TR-2007-021, RAL, 2007. *Int. J. Numer. Meth. Engrng*, 2009, to appear.
- [10] S. Chaturantabut and D. C. Sorensen. Discrete Empirical Interpolation for Nonlinear Model Reduction. Technical report, Department of Computational and Applied Mathematics, MS-134, Rice University, April 2009.
- [11] Chia-Chi Chua, Ming-Hong Laib, and Wu-Shiung Feng. Model-order reductions for MIMO systems using global Krylov subspace methods. *Mathematics and Computers in Simulation*, 79(4):1153–1164, 2008.
- [12] Benchmark Collection. Oberwolfach model reduction benchmark collection, 2003. <http://www.imtek.de/simulation/benchmark>.
- [13] V. Druskin and L. Knizhnerman. Extended Krylov subspaces: approximation of the matrix square root and related functions. *SIAM J. Matrix Anal. Appl.*, 19(3):755–771, 1998.
- [14] V. Druskin, L. Knizhnerman, and M. Zaslavsky. Solution of large scale evolutionary problems using rational Krylov subspace optimized shifts. *SIAM J. Sci. Comput.*, 31(5):3760–3780, 2009.
- [15] V. Druskin, C. Lieberman, and M. Zaslavsky. On adaptive choice of shifts in rational Krylov subspace reduction of evolutionary problems. Technical report, Schlumberger-Doll Research, 2009.
- [16] V. Druskin and M. Zaslavsky. On convergence of Krylov subspace approximations of time-invariant self-adjoint dynamical systems. Technical report, Schlumberger Doll Research, 2009.
- [17] I. M. Elfadel and D. D. Ling. A block rational Arnoldi algorithm for multipoint passive model-order reduction of multiport RLC networks, San Jose, CA. In *Proceedings of the 1997 IEEE/ACM international conference on Computer-aided design*, pages 66–71. IEEE Computer Society, 1997.
- [18] Nancy S. Ellner and Eugene L. Wachspress. New ADI model problem applications. In *Proceedings of 1986 ACM Fall joint computer conference, Dallas, Texas, United States*, pages 528–534. IEEE Computer Society Press, Los Alamitos, CA, 1986.
- [19] R. W. Freund. Krylov-subspace methods for reduced-order modeling in circuit simulation. *Journal of Computational and Applied Mathematics*, 123:395–421, 2000.
- [20] R. W. Freund. Krylov subspaces associated with higher-order linear dynamical systems. *BIT*, 45:495–516, 2005.
- [21] K. Gallivan, E. Grimme, and P. Van Dooren. Padè approximation of large-scale dynamic systems with Lanczos methods. In *Proc. of the 33rd Conference on Decision and Control*, pages 443–448, 1994.
- [22] K. Gallivan, E. Grimme, and P. Van Dooren. A rational Lanczos algorithm for model reduction. *Numerical Algorithms*, 12(1–2):33–63, 1996.
- [23] E. Grimme. *Krylov projection methods for model reduction*. PhD thesis, The University of Illinois at Urbana-Champaign, 1997.
- [24] S. Gugercin, A. C. Antoulas, and C. Beattie. \mathcal{H}_2 model reduction for large-scale linear dynamical systems. *SIAM J. Matrix Anal. Appl.*, 30:609–638, 2008.
- [25] S. J. Hammarling. Numerical solution of the stable, non-negative definite Lyapunov equation. *IMA J. Numer. Anal.*, 2:303–323, 1982.
- [26] D. Ingerman, V. Druskin, and L. Knizhnerman. Optimal finite difference grids and rational approximations of the square root. I. Elliptic functions. *Communic. on Pure and Appl. Math.*, LIII:1039–1066, 2000.
- [27] L. Knizhnerman, 2010. private communication.

- [28] D. Kressner. Block variants of Hammarling’s method for solving Lyapunov equations. *ACM Trans. Math. Software*, 34(1):1–15, 2008.
- [29] An Lu and Eugene L. Wachspress. Solution of Lyapunov equations by Alternating Direction Implicit iteration. *Computers Math. Applic.*, 21(9):43–58, 1991.
- [30] V. Mehrmann and T. Penzl. Benchmark collections in slicot. Technical Report SLWN1998-5, SLICOT Working Note, ESAT, KU Leuven, K. Mercierlaan 94, Leuven-Heverlee 3100, Belgium, 1998. Available under <http://www.win.tue.nl/niconet/NIC2/reports.html>.
- [31] K. H. A. Olsson and A. Ruhe. Rational Krylov for eigenvalue computation and model order reduction. *BIT Numerical Mathematics*, 46(1):99–111, 2006.
- [32] T. Penzl. A cyclic low-rank Smith method for large sparse Lyapunov equations. *SIAM J. Sci. Comput.*, 21(4):1401–1418, 2000.
- [33] A. Ruhe. Rational Krylov sequence methods for eigenvalue computation. *Lin. Alg. Appl.*, 58:391–405, 1984.
- [34] A. Ruhe. The rational Krylov algorithm for nonsymmetric eigenvalue problems. III: complex shifts for real matrices. *BIT*, 34:165–176, 1994.
- [35] Y. Saad. Numerical solution of large Lyapunov equations. In M. A. Kaashoek, J. H. van Schuppen, and A. C. Ran, editors, *Signal Processing, Scattering, Operator Theory, and Numerical Methods. Proceedings of the international symposium MTNS-89, vol III*, pages 503–511, Boston, 1990. Birkhauser.
- [36] Y. Saad. *Iterative methods for sparse linear systems*. The PWS Publishing Company, 1996.
- [37] W. H. A. Schilders, H. A. van der Vorst, and J. Rommes. *Model Order Reduction: Theory, Research Aspects and Applications*. Springer-Verlag, Berlin/Heidelberg, 2008.
- [38] G. Shi and C.-J. R. Shi. Model-order reduction by dominant subspace projection: error bound, subspace computation, and circuit applications. *IEEE Transactions on circuits and systems*, 52(5):975–993, 2005.
- [39] Valeria Simoncini. A new iterative method for solving large-scale Lyapunov matrix equations. *SIAM J. Sci. Comput.*, 29(3):1268–1288, 2007.
- [40] Valeria Simoncini. The Extended Krylov subspace for parameter dependent systems. Technical report, Università di Bologna, 2009.
- [41] D. Skoogh. A rational Krylov method for model order reduction. Technical Report 1998-47, Department of Mathematics, Chalmers University, Göteborg, Sweden, 1998.
- [42] Martin B. van Gijzen. IDR website, 2009. Available at <http://ta.twi.tudelft.nl/nw/users/gijzen/IDR.html>.
- [43] Martin B. van Gijzen and Peter Sonneveld. An elegant IDR(s) variant that efficiently exploits bi-orthogonality properties. Technical Report 08-21, Department of Applied Mathematical Analysis, Delft University of Technology, 2008.
- [44] Z. Bai Y.-T. Li and Y. Su. A two-directional Arnoldi process and its application to parametric model order reduction. *J. Comput. and Appl. Math.*, 226:10–21, 2009.