

EXTENDED KRYLOV SUBSPACE FOR PARAMETER DEPENDENT SYSTEMS *

V. SIMONCINI†

Abstract. The Extended Krylov Subspace has recently received considerable attention as a powerful tool for matrix function evaluations and other problems involving large matrices. In this paper we show that this space has a great potential within projection-type methods for effectively solving several other important large-scale algebraic problems: we focus on the solution of shifted systems and of more general parameter-dependent matrix equations, and on the approximation of the transfer function by projection. Numerical experiments stemming from real applications show the effectiveness of the approach.

1. Introduction. The Extended Krylov subspace has recently emerged as a competitive tool for approximating the solution to various problems involving large matrices. Introduced by Druskin and Knizhnerman in [10] for approximating functions of symmetric matrices, it was later explored to approximate functions of general non-symmetric matrices [24], and to approximate the solution to the Lyapunov, Sylvester and Riccati equations by means of projection-type methods [20, 21, 34].

Given a nonsingular matrix A and a vector b , the Extended Krylov subspace provides an enhanced approximation space by using both powers of A and A^{-1} , namely

$$\begin{aligned}\mathbf{EK}_m(A, b) &= \text{span}\{b, A^{-1}b, Ab, A^{-2}b, A^2b, \dots, A^{m-1}b, A^{-m}b\} \\ &= K_m(A, b) + K_m(A^{-1}, A^{-1}b),\end{aligned}$$

where $K_m(A, b) = \text{span}\{b, Ab, \dots, A^{m-1}b\}$ is the *Standard* Krylov subspace. Clearly, $\dim(\mathbf{EK}_m(A, b)) \leq 2m$. It was shown in [34] that the space $\mathbf{EK}_m(A, b)$ and the matrix representing A in that space may be generated iteratively, so that $\mathbf{EK}_m(A, b) \subseteq \mathbf{EK}_{m+1}(A, b)$. This important property makes it possible to implement an iterative method based on the Extended Krylov subspace by means of a simple recurrence.

Projection-type methods determine an approximation to the sought after solution by *projecting* the given problem onto a possibly much smaller approximation space, and in case of equations, by then enforcing some condition to obtain a unique approximation in the given space. Approximation spaces that have been extensively explored in the past in a variety of problems are the standard Krylov space, the shift-invert Krylov space, and the rational Krylov space. The last two spaces have the shortcoming of requiring a possibly expensive parameter selection, and their performance may in some cases be very sensitive to an inaccurate selection of these parameters. On the other hand, neither the Standard nor the recently introduced Extended Krylov subspaces depend on parameters. In addition, the Extended Krylov subspace has the advantage of “approximating both ends of the spectrum”, by generating powers of A and of A^{-1} , and their combination may intuitively explain its effectiveness over using either powers of A (standard Krylov space) or of A^{-1} (invert Krylov space). However, a rigorous understanding of the enhancements obtained by the Extended Krylov space is far less intuitive. A convergence analysis for approximating matrix functions for a large class of functions can be found in [24] (see also [10] for the symmetric case); a convergence analysis for solving the Lyapunov equation can instead be found in [26] for A symmetric, while it is under investigation in [25] for general A .

*Version of 7 August 2009.

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

In this paper we show that the Extended Krylov subspace has a great potential for effectively solving several important large-scale algebraic problems that can be found in many applied fields. Here we focus on the solution of shifted systems and of more general parameter-dependent matrix equations, and on the approximation of the transfer function. Our numerical results show that the new proposed projection methods based on the Extended Krylov subspace significantly outperform state-of-the-art methods.

The solution of parameter dependent linear systems of the form

$$(A + \sigma E)x = b, \quad (1.1)$$

where $A, E \in \mathbb{C}^{n \times n}$, $b \in \mathbb{C}^n$, $\sigma \in \mathbb{C}$ and $x = x(\sigma)$, arises in many applications such as control theory, structural dynamics and time-dependent PDEs; see, e.g., [11, 29, 35] and references therein. Depending on the application problem, the values of the parameter σ may vary either in a neighborhood of a given fixed quantity σ_0 , or in a wide range. In the latter case, strategies based on perturbation methods around $A + \sigma_0 E$ are clearly not applicable. It is important to realize that for shifted systems, the computational challenge is given both by the properties and dimensions of A and E , but also by the number of shifts involved. Even for medium size systems, for which sparse direct methods are usually very efficient, the presence of many parameter values makes it necessary to determine strategies whose cost grows only sublinearly with the number of parameters. We refer to, e.g., [4, 5, 12, 13, 33] for some solvers and preconditioners that take into account these properties; a more complete account may be found in [35].

As a natural extension to the problem (1.1), the right-hand side b could also depend on the parameter, namely $b = b(\sigma)$ (see, e.g., [7, 17, 27]), so that for a set of values $\sigma_1, \dots, \sigma_s$, the following problem becomes of interest

$$AX + EXS = B, \quad S = \text{diag}(\sigma_1, \dots, \sigma_s), \quad B = [b(\sigma_1), \dots, b(\sigma_s)].$$

Another closely related problem with wide application in Control consists of approximating the so-called *transfer* function

$$h(\sigma) = C^*(A - i\sigma E)^{-1}B, \quad \sigma \in \mathbb{R},$$

where A, E, B and C are the matrices representing a dynamical linear system. Here and throughout the paper we use x^\top to denote transposition of a complex vector x , and x^* to denote conjugate transposition. The transfer function may have a quite oscillatory behavior on a range of values for σ that usually spans several orders of magnitude, and such behavior needs to be captured by a reliable approximation.

In the following sections we first characterize the Extended Krylov subspace, describing in some detail its computation. We then focus on the three problems above, providing for each of them relevant numerical experiments with data stemming from real application problems, that emphasize the enhancement obtained by employing the new approximation space. All reported experiments were run using Matlab 7.5 [28].

2. The Extended Krylov subspace. An orthogonal basis of the Extended Krylov subspace may be generated iteratively while the space is expanded by means of a Gram-Schmidt procedure, in a way similar to that of the standard Krylov subspace. A less straightforward procedure needs to be devised to iteratively generate the projection and restriction matrix that represents A onto the generated space,

without additional matrix-vector multiplies with A . More precisely, if the (orthonormal) columns of $\mathcal{V}_m \in \mathbb{C}^{n \times 2m}$ span $\mathbf{EK}_m(A, b)$, then an ad-hoc procedure is required to obtain $\mathcal{T}_m = \mathcal{V}_m^* A \mathcal{V}_m$, since the Gram-Schmidt procedure delivers matrix-vector products with A for free only for the odd columns of \mathcal{V}_m . In [34] a recurrence was suggested to obtain \mathcal{T}_m : The odd columns coincide with the corresponding columns of the orthogonalization coefficients matrix in the Gram-Schmidt process, whereas the even columns are recovered by means of an auxiliary recurrence with short-length vectors [34, Prop. 3.2].

The generation of $\mathcal{V}_m = [V_1, V_2, \dots, V_m] \in \mathbb{C}^{n \times 2m}$ and $\mathcal{T}_m = (t_{i,j}) \in \mathbb{C}^{2m \times 2m}$ may be summarized in the following steps:

Algorithm 1.

Set $\mathcal{V}_1 = V_1 = \mathbf{gram}([b, A^{-1}b])$.

For $m = 1, 2, \dots$,

1. Set $V_m^{(1)}$: first column of V_m ; $V_m^{(2)}$: second column of V_m
2. $V'_{m+1} = [AV_m^{(1)}, A^{-1}V_m^{(2)}]$
3. $\widehat{V}_{m+1} \leftarrow$ orthogonalize V'_{m+1} w.r.to \mathcal{V}_m ;
4. $V_{m+1} = \mathbf{gram}(\widehat{V}_{m+1})$
5. Let h (size $(2m+2) \times 2$) contain the orth. coeffs
6. New odd column of \mathcal{T}_m : $t_{:,2m-1} = h_{:,1}$
7. New even column of \mathcal{T}_m :

$$\begin{aligned} (m=1) t_{:,2} &= h_{:,1} (\ell_{11}^{(1)})^{-1} \ell_{12}^{(1)} + e_1 (\ell_{11}^{(1)})^{-1} \ell_{22}^{(1)} & t_{:,4} &= (e_2 - \mathcal{T}_1 h_{1:2,2}) \ell_{22}^{(2)} \\ & \rho^{(2)} = (\ell_{11}^{(2)})^{-1} \ell_{12}^{(2)} \\ (m>1) t_{:,2m} &= t_{:,2m} + t_{:,2m-1} \rho^{(m)} \\ t_{:,2m+2} &= (e_{2m} - [\mathcal{T}_m; O_{2,2m}] h_{1:2m,2}) \ell_{22}^{(m+1)} \\ \rho^{(m+1)} &= (\ell_{11}^{(m+1)})^{-1} \ell_{12}^{(m+1)}. \end{aligned}$$

8. $\mathcal{V}_{m+1} = [\mathcal{V}_m, V_{m+1}]$

Algorithm 1 adds two new vectors to the basis to expand the space, one multiplied by A and the other by A^{-1} . The function \mathbf{gram} performs a Gram-Schmidt orthogonalization within the matrix block.

A closer look at the relations generating \mathcal{T}_m shows that \mathcal{T}_m is block upper Hessenberg, with 2×2 blocks ($2p \times 2p$ if b is a matrix with p columns, see below). In exact arithmetic it is block tridiagonal and Hermitian whenever A is Hermitian; see also [23] for a different derivation for A real symmetric. As soon as the representation matrix \mathcal{T}_m is updated, the underlying problem may be approximated within the enhanced space, and then the approximate solution prolonged back to \mathbb{C}^n ; see, e.g., Algorithm 2 below.

For later convenience, we remark that a block version of the space, namely $\mathbf{EK}_m(A, B)$, with $B \in \mathbb{C}^{n \times p}$, can also be generated, requiring only a ‘‘blocking’’ of the scalar quantities described above, and no other major changes [34].

We also mention that the matrix \mathcal{V}_m formally satisfies the following Arnoldi-type relation (see, e.g., [10, 21]):

$$A \mathcal{V}_m = \mathcal{V}_m \mathcal{T}_m + V_{m+1} \boldsymbol{\tau}_{m+1,m} E_m^\top, \quad (2.1)$$

where E_m contains the last 2 columns of the identity matrix of dimension $2m$, and $\boldsymbol{\tau}_{m+1,m} = V_{m+1}^* A \mathcal{V}_m$; note that the second row of $\boldsymbol{\tau}_{m+1,m}$ is zero [34]. Clearly, the relation in (2.1) cannot be used to explicitly derive the next column vectors in V_{m+1} .

For b a vector, it was shown in [24, 34] that the method may break down, but then it is a happy breakdown, that is, the problem solution is contained in the computed approximation space. For $b = B$ a tall matrix, then unlucky breakdown may occur, and this may be due to the redundancy of some of the columns used to enlarge the space at any iteration. Deflation techniques should then be adopted, as it is the case in block Krylov subspace methods; see [18] for a rather complete list of references on the topic. The special structure of the basis makes the deflation implementation more cumbersome. We refer to [36] for an implementation.

Whenever the matrix A is *complex* symmetric, namely $A = A^\top \in \mathbb{C}^{n \times n}$, it is still possible to exploit this form of symmetry, to obtain a block tridiagonal matrix, as in the real symmetric case. Following the procedure employed for standard Krylov subspace methods, this may be obtained by using the bilinear form $(x, y) = x^\top y$, with $x, y \in \mathbb{C}^n$ throughout the algorithm, instead of the standard Euclidean inner product x^*y , where x^* denotes the conjugate transpose of x ; we refer to [35] for a description of the use of the bilinear form (x, y) in the complex symmetric context. As a result, the matrix \mathcal{V}_m is complex and satisfies $\mathcal{V}_m^\top \mathcal{V}_m = I_{2pm}$, where p is the number of columns of b . Moreover, as anticipated, \mathcal{T}_m is block tridiagonal with $2p \times 2p$ blocks, and as in the real symmetric case no explicit orthogonalization is required with respect to the older basis vectors.

3. Application problems. We next describe some application problems where the use of the Extended Krylov subspace indeed significantly improves convergence, with respect to the most direct competitors. The fact that each iteration requires the application of A^{-1} readily shows that for the same space dimension, the approximation enhancement obtained by using the new space should be sufficiently high to make up for the extra cost of using A^{-1} , in some forms, rather than merely A . We show that this is indeed the case for many applications we have encountered.

3.1. Shifted linear systems. We consider the solution of the parameterized linear system

$$(A + \sigma I)x = b, \quad \sigma \in \mathbb{C}$$

which need be solved for many values of σ . The solution $x = x(\sigma)$ may be written as $x = (A + \sigma I)^{-1}b \equiv f(A)b$. Therefore, this is a special case of matrix function evaluation, and the Extended Krylov subspace can be employed as done for instance in [10, 24]. However, the linear system setting is particularly convenient to derive several useful properties for shifted systems (see, e.g., [35, sec.14]). In particular, using (2.1), it follows that

$$(A + \sigma I)\mathcal{V}_m = \mathcal{V}_m(\mathcal{T}_m + \sigma I) + \mathbf{V}_{m+1}\boldsymbol{\tau}_{m+1,m}E_m^\top,$$

showing that the same space $\mathbf{EK}_m(A, b)$ can be used to approximate all parameterized systems, as σ varies, the way it is done for the usual space $K_m(A, b)$ [35, sec.14]. Therefore, using the fact that $\mathcal{V}_m^*b = e_1\|b\|$, we have

$$x_m(\sigma) = \mathcal{V}_m y_m(\sigma) := \mathcal{V}_m(\mathcal{T}_m + \sigma I)^{-1}e_1\|b\|. \quad (3.1)$$

Clearly, the parameter σ only enters the computation in the evaluation of $y_m(\sigma)$. Note that, for each σ , this procedure corresponds to the interpolation approximation of f as $y_m(\sigma) = f(\mathcal{T}_m)e_1\|b\|$ [24].

The major difference with respect to the standard Krylov approach is that a system with A must be solved at each iteration to expand the basis, making the

generation of the single space more expensive. The computation of $\mathbf{EK}_m(A, b)$ may be advantageous with respect to solving each system $(A + \sigma I)$ if either of the following conditions applies:

- i) Solving with $(A + \sigma I)$ is more expensive than solving with A (e.g., A is large and real and σ is complex);
- ii) The number of parameters to be considered is large, and the space dimension at convergence is significantly lower.

Both conditions can be satisfied in application problems. In particular, below we consider an example from structural dynamics that does require the solution of a large number of parameterized systems. We also remark that shifted systems are difficult to precondition, while maintaining the shifted structure (cf. the discussion in [35]), therefore the new enhanced space may be viewed as a form of acceleration alternative to spectral transformations.

Whenever A is symmetric, it is known that the solution of a linear system by a Krylov subspace does not require storing and explicitly orthogonalizing the whole basis. As already mentioned, this is the also the case for the Extended Krylov subspace: since \mathcal{T}_m is block tridiagonal, only a three-term block recurrence needs to be performed. More precisely, the orthogonalization in step 4 of Algorithm 1 need be carried out only with respect to the blocks V_m, V_{m-1} . We refer to [23] for additional details of an explicit implementation of this short-term recurrence.

We next give a possible outline of the resulting iterative scheme.

Algorithm 2.

Given $A, r_0(\sigma) = b, x_m(\sigma) = 0, \Sigma$.

Compute $\beta_0(\sigma) = \|b\|, v_1 = b/\beta_0(\sigma)$. Set $\Sigma_c = \emptyset$.

Compute $\mathcal{V}_1 = V_1 = \mathbf{gram}([b, A^{-1}b])$.

for $m = 1, 2, \dots$

1. Generate V_{m+1} , compute new columns of \mathcal{T}_m ;
2. Solve $(\mathcal{T}_m + \sigma I)y(\sigma) = e_1\|b\|, \sigma \in \Sigma \setminus \Sigma_c$;
3. Update set Σ_c of converged systems;
4. if $\Sigma \setminus \Sigma_c = \emptyset$ compute $x_m(\sigma) = \mathcal{V}_m y_m(\sigma), \sigma \in \Sigma_c$; stop;
5. Expand space: $\mathcal{V}_{m+1} = [\mathcal{V}_m, V_{m+1}]$;

The Arnoldi-type relation (2.1) can also be used to generalize other convenient properties. Most importantly, the following result holds for the system residuals.

PROPOSITION 3.1. *Let $r_m(\sigma) = b - (A + \sigma I)x_m(\sigma)$, with x_m as in (3.1). Then $r_m(\sigma) = -V_{m+1}\tau_{m+1,m}E_m^\top y_m(\sigma)$. In particular, all residuals are collinear to $V_{m+1}e_1$.*

Proof. The expression for the residual is obtained by inserting the Arnoldi-type relation (2.1) into $r_m(\sigma) = b - (A + \sigma I)\mathcal{V}_m y_m(\sigma)$. The collinearity follows from the normality of the basis and from noticing that the second row of $\tau_{m+1,m}$ is zero. \square

This property was first exploited for restarting shifted system solvers in [33]; we also refer to [14] for error estimations of Krylov subspace solvers for shifted systems that make use of such a collinearity property.

Two important features follow from Proposition 3.1. The first one is that the residual norm can be computed cheaply at each iteration for all shifted systems, without the explicit computation of the solution $x_m(\sigma)$ or the residual, namely

$$\|r_m(\sigma)\| = \|e_1^\top \tau_{m+1,m} E_m^\top y_m(\sigma)\|. \quad (3.2)$$

The second consequence is of paramount importance when solving nonsymmetric systems, for which restarting may be required. Since all basis vectors need be explicitly orthogonalized and stored, a maximum subspace dimension is usually allowed, and if accuracy is not satisfactory, the procedure is restarted with the current approximate solution as starting guess, so that the new space is generated with the current residual as starting vector. When dealing with shifted systems, the fact that the residuals are collinear makes it possible to build a single approximation space at restart time, starting with $V_{m+1}e_1$; otherwise, a distinct restart space would be required for each parameterized system. We refer to [34, 35] for a discussion on general Galerkin approximations and their collinearity properties.

The restarted process can be outlined as follows:

Given $A, r_0(\sigma) = b, x_m(\sigma) = 0, \Sigma, m$.

Compute $\beta_0(\sigma) = \|b\|, v_1 = b/\beta_0(\sigma)$. Set $\Sigma_c = \emptyset$.

1. Build $\mathbf{EK}_m(A, v_1)$ as in Algorithm 1;
2. Compute $y_m(\sigma) = (\mathcal{T}_m + \sigma I)^{-1}e_1\beta_0(\sigma), \sigma \in \Sigma \setminus \Sigma_c$;
3. Compute $\|r_m(\sigma)\|$ using (3.2)
4. Compute $x_m(\sigma) = x_m(\sigma) + \mathcal{V}_m y_m(\sigma), \sigma \in \Sigma \setminus \Sigma_c$. Update set Σ_c of converged shifted systems. If $\Sigma \setminus \Sigma_c = \emptyset$ stop;
5. Set $v_1 = V_{m+1}e_1$. Set $\beta_0(\sigma) = v_1^* r_m(\sigma), \sigma \in \Sigma \setminus \Sigma_c$. Goto 1.

3.2. Numerical Experiments. In this section we describe some problems that can benefit of the use of the shifted Extended Krylov subspace. Comparisons are performed with respect to its more direct competitor, shifted FOM [33]. Experiments are also reported in some cases with a sparse direct solver applied to each shifted systems. In all cases, prior to the generation of the space $\mathbf{EK}_m(A, b)$, the matrix A is reordered (`amd` reordering in Matlab). and then factorized once for all.

EXAMPLE 3.2. We consider the nonsymmetric matrix A stemming from the centered finite difference discretization of the operator $L(u) = -\Delta u + 50(x+y)(u_x + u_y)$. The matrix A is scaled by the mesh dimension, and we consider two discretization meshes, resulting into a problem of dimension $n = 2500$ and $n = 10\,000$, respectively. The right-hand side is chosen as the normalized vector of all ones, and the shifts are taken to be 500 values uniformly distributed in the interval $[0, 5]$. The stopping tolerance is set equal to 10^{-8} . The results using the shifted versions of restarted FOM and of restarted EKSM are reported in Table 3.1, and show that the Extended space is able to effectively reduce the number of restarts and the CPU time. The cost of factorizing the coefficient matrix is included.

n	subspace		FOM	EKSM
	dimension	CPU time (#cycles)	CPU time (#cycles)	CPU time (#cycles)
2500	10		1.50 (41)	0.65 (4)
2500	20		1.57 (14)	0.63 (2)
10 000	10		3.23 (49)	1.96 (7)
10 000	20		4.35 (27)	1.75 (3)

TABLE 3.1

Example 3.2. Shifted solvers for nonsymmetric matrices of size n . 500 Shifts uniformly distributed in $[0, 5]$. Stopping tolerance 10^{-8} .

EXAMPLE 3.3. We consider the rational approximation of the action of a matrix function to a vector. More precisely, we compute the partial fraction sum

$$\sum_{k=1}^{\nu} \kappa_k (\omega_k^2 I - A)^{-1} A b \approx f(A) b$$

where $f(\lambda) = \sqrt{\lambda}$. We test the method on the matrix $A \in \mathbb{R}^{2500 \times 2500}$, discretization of the operator $-\Delta u + 10u_x$ on the unit square with homogeneous Dirichlet boundary conditions; b is the vector of all ones. Here we employ the rational approximation described as `method2` in [19, p.2512-2513], which yields complex values of the shifts ω_k^2 's. We use $\nu = 18$ (non-conjugate) terms which in our setting provided a relative error in the function evaluation below 10^{-13} . On the other hand, this requires solving 18 shifted linear systems: we used a stopping tolerance equal to 10^{-10} . A small matrix is considered to be able to use `method2` exactly, without any estimate of the involved quantities. The Extended Krylov subspace method took 0.04 secs to factorize the matrix, and 0.33 secs to solve all systems. We recall here that most computation is carried out in real arithmetic, and only the solution of the small linear system in $\mathcal{T}_k - \omega_k^2 I$ and the solution update are performed in complex arithmetic. The sequential solution of all (complex non-Hermitian) systems by a sparse direct method required about 1.28 secs.

EXAMPLE 3.4. We consider a benchmark test problem stemming from a structural model of a car windscreen [8]. The resulting system to be solved is written as

$$(K - \omega^2 M)x(\omega) = b,$$

and has size $n = 22,692$. The stiffness matrix K is given by a real symmetric matrix multiplied by a complex scalar, while M is a real symmetric matrix. The frequency range for ω is $\omega \in [0.5, 200]$. A typical already challenging situation may be obtained by picking up 11 among the middle interval values, $\omega_k = 0.5k$, $k = 200, \dots, 210$. We solved the systems with a stopping tolerance of 10^{-4} . The *unrestarted* Extended procedure converges for all parameters with a subspace of dimension 36. On the other hand, the shifted FOM approach does not converge in 100 iterations. The restarted version of the latter approach shows similar difficulties. We also observe that this problem was solved in [29] by means of a spectral (shift-and-invert) transformation as preconditioner. A subspace dimension close to that for the Extended Krylov space was reported to achieve similar residual norms for the best selected spectral transformation. However, the effectiveness of the spectral transformation depends on the parameter selection, and it requires a number of solves equal to the space dimension, that is, twice as many as for the Extended Krylov space. Since the back solves is an expensive task in this case, the overall cost is appreciably higher than for the Extended Krylov space.

Finally, we observe that the problem could be formulated as $(\tilde{K} - \tilde{\omega}^2 M)\tilde{x} = b$, where \tilde{K} is real and $\tilde{\omega}^2$ are complex. We have preferred to treat the problem as originally stated.

4. Sylvester-type equations. In some applications the right-hand side may also depend on the parameter. In this case, considering the parameters $\sigma_1, \dots, \sigma_s$, the systems to be solved take the form

$$AX + XS = B, \quad S = \text{diag}(\sigma_1, \dots, \sigma_s), \quad B = [b(\sigma_1), \dots, b(\sigma_s)].$$

This is a particular form of Sylvester equation, where in general S may be nondiagonal. To ensure uniqueness of the solution, A and $-S$ must have disjoint spectra. In fact, we could formulate the problem for any matrix S not necessarily diagonal. Note that the more general equation $AX + EXS = B$ may be restated as the one above, as long as E is nonsingular (see also below).

For $s \ll n$, an approximation to X may be effectively obtained in an approximation space generated by using A and B ; see, e.g., [6, 9, 20, 22, 30, 32]. In the Extended Krylov subspace context, we can generate the matrices $\mathcal{V}_m, \mathcal{T}_m$ by means of Algorithm 1, and then, following the Galerkin requirements, project the problem onto the space and require that the residual be orthogonal to it. More precisely, we can write $X \approx \mathcal{V}_m Y$, for some Y such that

$$\mathcal{V}_m^* R_m \equiv \mathcal{V}_m^* (A\mathcal{V}_m Y + \mathcal{V}_m Y S - \mathcal{V}_m E_1 \beta_0) = 0,$$

where $B = \mathcal{V}_m E_1 \beta_0$, $\beta_0 \in \mathbb{C}^{s \times s}$, which is equivalent to

$$\mathcal{T}_m Y + Y S = E_1 \beta_0. \quad (4.1)$$

Here and in the following, E_1 denotes the first s columns of the identity matrix, whose dimension is made clear from the context. Symmetry (real or complex) can be exploited as described in section 2, lowering the computational cost of the method. This formulation may be viewed as a special case of the generalization proposed in [20], where here S has small dimension and the whole matrix B is used.

As in the previous framework, the residual norm may be computed inexpensively as $\|R_m\| = \|e_1^\top \tau_{m+1,m} E_m^\top Y_m(\sigma)\|$, where Y_m solves the reduced problem (4.1). The only difference with respect to the (scalar) shifted system is that the right-hand side is in general a full rank matrix, whereas in the former case, the right-hand side is the rank-one matrix $e_1 \|b\| \mathbf{1}^\top$.

We explicitly remark that the approach generates the *block* Extended Krylov subspace $\mathbf{EK}_m(A, B)$, whose dimension is less or equal to $2ms$, where s is the number of columns of B . Therefore the method seems to be appealing when s is moderate, to avoid the presence of very large blocks.

EXAMPLE 4.1. This example deals with the solution of parameterized algebraic systems in a direct frequency analysis arising in structural dynamics [11]. The general frequency dependent equation of motion is given by

$$(K^* + i\sigma C_V - \sigma^2 M)x(\sigma) = b(\sigma). \quad (4.2)$$

Here we consider no viscous damping, so that $C_V = 0$. The remaining data matrices are the stiffness and mass matrices of a finite element discretization of a machine foundation resting on ground (Case **B** in [11]), and have size 3627×3627 . The stiffness matrix has in fact the form $K^* = K + iC_D$, where K is the true stiffness matrix and C_D accounts for hysteretic damping, so that K^* is complex symmetric. We consider 9 frequencies f_k uniformly taken in the interval $[10, 20]Hz$, and $\sigma_k = 2\pi f_k$. The right-hand side is a matrix with 9 columns, having nonzero rows $l = 1962, 2046, 2130, 2214, 2298, 2382, 2466, 2550, 2634$, and values $b_{l,k} = \exp(-2\pi f_k(l-1)0.4i)$, $k = 1, \dots, 9$, where f_k is the frequency; we refer to [17] for a detailed description of the problem. The systems $(K^* - \sigma_k^2 M)x(\sigma_k) = b(\sigma_k)$, $k = 1, \dots, 9$ can be written in compact form and with obvious notation as

$$K^* X + M X S = B, \quad S = \text{diag}(\sigma_1, \dots, \sigma_9).$$

For M nonsingular, let $M = LU$ be its LU decomposition. Then we can solve the following equivalent Sylvester equation

$$(L^{-1}K^*U^{-1})\tilde{X} + \tilde{X}S = L^{-1}B, \quad \tilde{X} = UX. \quad (4.3)$$

Note that this formulation allows the iterative solver and the computation of the residual norm to be less sensitive to the conditioning of M , compared to the left-side inversion. It also allows one to exploit the symmetry of M , as discussed next.

The equation in (4.3) is solved by projection onto the Extended Krylov subspace. In the considered problem, M is real symmetric, so that $L = U^\top$, and K^* is complex symmetric, so that $L^{-1}K^*U^{-1}$ is also complex symmetric. To fully exploit the symmetry of the problem, we have implemented the Extended Krylov method with the bilinear form $(x, y) = x^\top y$ instead of the standard complex inner product x^*y ; see, e.g., [35]. It follows that \top -orthogonalization is only carried out with respect to the previous 2 block vectors (cf. section 2).

We compare the performance of the new solver with the solution of the original equation $Q(\sigma)x(\sigma) = b(\sigma)$, $Q(\sigma) = K^* - \sigma_k^2 M$ by means of the complex symmetric CG method, preconditioned by an LDL $^\top$ -type preconditioner with dropping tolerance 10^{-2} (GMRES on each system $Q(\sigma)x(\sigma) = b(\sigma)$ was significantly more expensive for the same number of iterations). The stopping tolerance was set equal to 10^{-8} .

Complex Sym CG		Extended Krylov	
Total n.its	CPU Time	Space dim	CPU Time
740	21.04	90	5.48

The reported CPU Times show the effectiveness of the new approach. The CPU Time for CG includes the incomplete factorization of $Q(\sigma)$ for each σ , while that of the Extended Krylov subspace includes the time for the factorization of K^* (with a `symamd` pre-reordering).

5. Transfer function. The procedure described in the previous sections can be used to approximate the transfer function

$$h(\sigma) = C^*(A - i\sigma E)^{-1}B, \quad \sigma \in [\alpha, \beta].$$

The approximation is performed by determining smaller representation matrices \tilde{A} , \tilde{E} , \tilde{B} and \tilde{C} , such that $\tilde{h}(\sigma) = \tilde{C}^*(\tilde{A} - i\sigma\tilde{E})^{-1}\tilde{B}$ approximates $h(\sigma)$ for all σ 's in an interval of interest. A crucial feature of any approximation strategy is that, in spite of the size of the original problem, the approximation space should be as small as possible, so as to truly perform a model reduction; see, e.g., [2, 31]. In particular, all relevant oscillatory features of $h(\sigma)$ should be retained in the reduced model.

By building the Extended Krylov subspace associated with AE^{-1} and B , one can construct the matrices $\mathcal{V}_m, \mathcal{T}_m$ and then project the problem, so that

$$h(\sigma) \approx (\mathcal{V}_m^*(E^{-*}C))^*(\mathcal{T}_m - i\sigma I)^{-1}E_1\beta_0, \quad \text{with } B = V_1\beta_0.$$

We stress that we assume here that both A and E are nonsingular. In case of singularity of one of the two matrices, say A , we can proceed by translating the problem so that $A - i\sigma_0 E$ is nonsingular, and then work with the transfer function $h(\sigma) = C^*((A - i\sigma_0 E) - i\delta E)^{-1}B$, $\sigma = \sigma_0 + \delta$.

Similar projection procedures have been classically explored in the control literature for approximating the transfer function, leading to the use of the standard Krylov

space $K_m(A, B)$, the “shift-and-inverted” Krylov space $K_m((A - \omega_0 I)^{-1}, B)$, and the rational Krylov subspace $\text{span}\{B, (A - \omega_1 I)^{-1}B, (A - \omega_2 I)^{-1}B, \dots\}$; we refer to [1] for a thorough analysis of these methods and for an exhaustive list of references. By comparing the Extended and rational Krylov spaces, it appears that the Extended Krylov subspace is nothing but a rational space with repeated poles at zero and infinity, for which a theoretical framework may be found in [1, Cor. 11.9, p.355]. On the other hand, it is interesting that apparently an iterative approach to generate this space was never implemented, although it provides some clear advantages over the general rational space, whose pole selection remains in most cases a trial-and-error procedure [1, 15, 16].

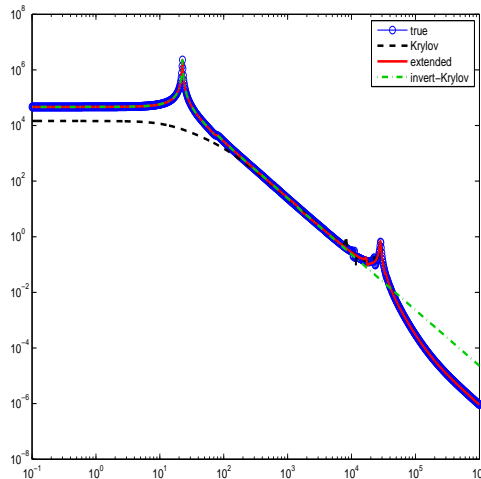


FIG. 5.1. Example 5.1. Bode plot of the first input and output: approximation space dimension equal to 20.

EXAMPLE 5.1. We consider a famous small benchmark problem, the CD Player, available in the SLICOT Library [3], whose matrix A has size $n = 120$ (here $E = I$). The matrices B and C have 2 columns each. For this experiment we have decided to approximate the columns of B, C separately. The block version of the methods could be used to approximate the whole transfer function matrix.

In Figure 5.1 we report the Bode plot, namely the log-log graph of the function $|h_{i,j}(\sigma)|$, associated with the problem for the first input and output. The curves represent the true function (circled thin line), the approximation obtained by the standard Krylov subspace of size 20 (dashed line), by the inverted Krylov subspace of size 20 (dash-dotted line) and by the Extended Krylov subspace of size 20 (solid line). Clearly, the Extended method is better capable of capturing the oscillations in both parts of the spectrum by combining the features of both the standard and inverted approaches, thus providing an overall optimal approximation of the transfer function. In Figure 5.2 we report the Bode plot associated with the problem for the first input and second output and a subspace of dimension 20 in both cases (left plot). The standard and inverted spaces are unable to approximate most of the curve for such a small dimension, whereas the Extended space does a good job around its poles. However, the overall approximation is still unsatisfactory. The right plot shows the quality of the approximation for a subspace of dimension 50, which allows the Extended Krylov subspace to satisfactorily capture all transfer function’s

features. The standard procedure is still missing most of the action in the middle of the frequency interval; the inverted method is still not accurate for high frequencies. We stress here that for the same space dimension, the inverted Krylov space requires twice as many matrix solves as the Extended Krylov space.

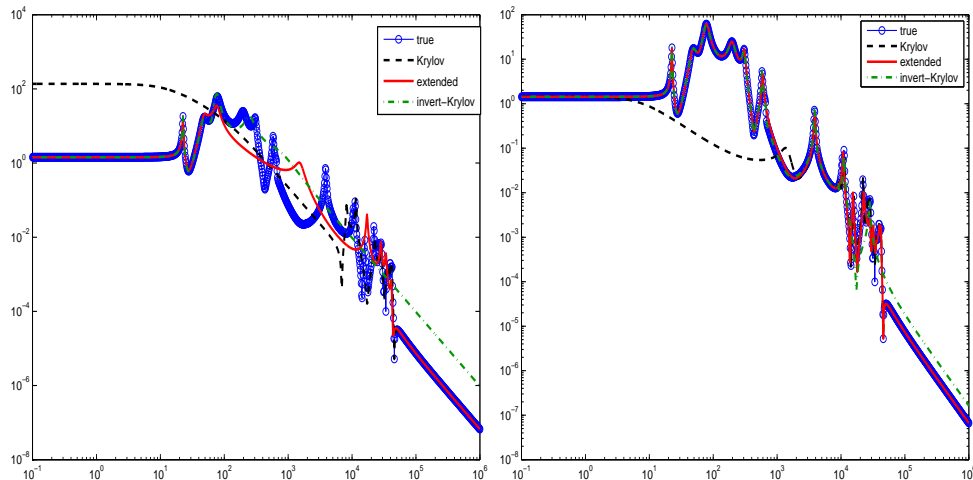


FIG. 5.2. Example 5.1. Bode plot of the first input and second output: approximation space dimension equal to 20 (left), and to 50 (right).

EXAMPLE 5.2. We consider the Benchmark problem Beam [8], which has matrix A of size 348 (also here $E = I$) and single column input-output matrices. Figure 5.3 represents the corresponding transfer function for frequencies in $[10^{-2}, 10^3]$, and its reduced model in a subspace of dimension $m = 20$ (left) and $m = 40$ (right) with both the standard and Extended Krylov spaces, the latter choice delivering a far more accurate approximation of the original function. The inverted Krylov space improves its accuracy on high frequencies when going from $m = 20$ to $m = 40$, but remains unsatisfactory, compared to the combined approach.

6. Conclusions. We have shown that the Extended Krylov subspace may be a competitive alternative to other Krylov-type spaces for solving a variety of algebraic problems of great interest in applications. The effectiveness of the method relies on efficient solves with the involved large-scale nonsingular matrix A . We have shown that many advanced scientific problems are so complex that they can afford such a cost. For larger or denser problems, systems with A may be solved by means of preconditioned iterative techniques, giving rise to so-called inner-outer iterations; we refer to [24, 34] for examples of methods based on the Extended Krylov subspace.

REFERENCES

- [1] A. C. Antoulas. *Approximation of large-scale Dynamical Systems*. Advances in Design and Control. SIAM, Philadelphia, 2005.
- [2] 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.
- [3] P. Benner, V. Mehrmann, V. Sima, S. Van Huffel, and A. Varga. SLICOT - A Subroutine Library in Systems and Control Theory. *Applied and Computational Control, Signals, and Circuits*, 1:499–539, 1999. <http://www.slicot.org>.

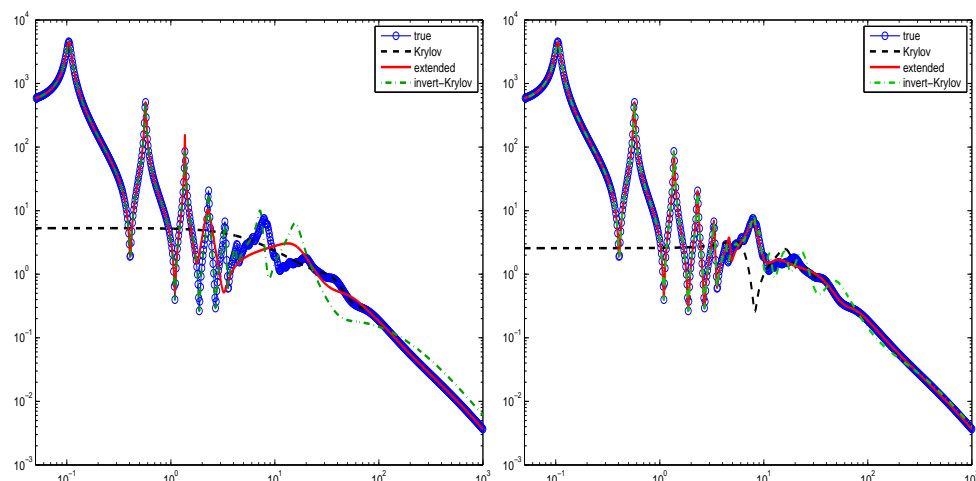


FIG. 5.3. Example 5.2 Bode plot of the second input and first output: approximation space dimension equal to 20 (left), and to 40 (right).

- [4] M. Benzi and D. Bertaccini. Approximate inverse preconditioning for shifted linear systems. *BIT, Numerical Mathematics*, 43(2):231–244, 2003.
- [5] D. Bertaccini. Efficient solvers for sequences of complex symmetric linear systems. *ETNA*, 18:49–64, 2004.
- [6] A. Bouhamidi and K. Jbilou. A note on the numerical approximate solutions for generalized Sylvester matrix equations with applications. Technical report, L.M.P.A., 2008.
- [7] C. Cabos and F. Ihlenburg. Vibrational analysis of ships with coupled finite and boundary elements. *J. Computational Acoustics*, 11(1):91–114, 2003.
- [8] Benchmark Collection. Oberwolfach model reduction benchmark collection, 2003. <http://www.imtek.de/simulation/benchmark>.
- [9] B. Datta and Y. Saad. Arnoldi methods for large Sylvester-like observer problems, and an associated algorithm for partial spectrum assignment. *Lin. Alg. Appl.*, 154-156:225–244, 1991.
- [10] 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.
- [11] A. Feriani, F. Perotti, and V. Simoncini. Iterative system solvers for the frequency analysis of linear mechanical systems. *Computer Methods in Applied Mechanics and Engineering*, 190(13–14):1719–1739, 2000.
- [12] R. F. Freund. Solution of shifted linear systems by quasi-minimal residual iterations. In L Reichel, A. Ruttan, and R. S. Varga, editors, *Numerical Linear Algebra*, pages 101–121, Berlin, 1993. W. de Gruyter.
- [13] A. Frommer and U. Glässner. Restarted GMRES for shifted linear systems. *SIAM J. Sci. Comput.*, 19(1):15–26, 1998.
- [14] A. Frommer and V. Simoncini. Stopping criteria for rational matrix functions of Hermitian and symmetric matrices. *SIAM J. Sci. Comput.*, 30(3):1387–1412, 2008.
- [15] K. Gallivan, E. Grimme, and P. Van Dooren. Padé approximation of large-scale dynamic systems with Lanczos methods. In *Proc. of the 33rd Conference on Decision and Control*, pages 443–448, 1994.
- [16] K. Gallivan, E. Grimme, and P. Van Dooren. A rational Lanczos algorithm for model reduction. *Numerical Algorithms*, 12(1–2):33–63, 1996.
- [17] G. Gu and V. Simoncini. Numerical solution of parameter-dependent linear systems. *J. Numer. Linear Algebra w/Appl.*, 12(9):923–940, 2005.
- [18] M. H. Gutknecht. Krylov subspace algorithms for systems with multiple right hand sides: a collection of references, 2004. available at <http://www.sam.math.ethz.ch/~mhg/talks/abstr/abstr.html>.
- [19] N. Hale, N. Higham, and L. N. Trefethen. Computing A^α , $\log(A)$ and Related Matrix Functions by Contour Integrals. *SIAM J. Numer. Anal.*, 46(5):2505–2523, 2008.
- [20] M. Heyouni. Extended Arnoldi Methods for Large Sylvester Matrix Equations. Technical

- report, L.M.P.A., 2008.
- [21] M. Heyouni and K. Jbilou. An extended block method for large-scale continuous-time algebraic Riccati equations. Technical report, L.M.P.A., Université du Littoral, 2007.
 - [22] D. Y. Hu and L. Reichel. Krylov Subspace Methods for the Sylvester Equation. *Linear Algebra and Appl.*, 174:283–314, 1992.
 - [23] C. Jagels and L. Reichel. The extended Krylov subspace method and orthogonal Laurent polynomials. *Lin. Alg. Appl.*, 431:441–458, 2009.
 - [24] L. Knizhnerman and V. Simoncini. A new investigation of the extended Krylov subspace method for matrix function evaluations. *Numerical Linear Algebra with Applications*, pages xx–xx, 2009. To appear.
 - [25] L. Knizhnerman and V. Simoncini. Convergence analysis of EKSM for the Lyapunov equation, 2009. In preparation.
 - [26] D. Kressner and C. Tobler. Krylov subspace methods for linear systems with tensor product structure. Technical Report SAM-2009-16, ETH Zurich, 2009.
 - [27] M. Kuzuoglu and R. Mittra. Finite element solution of electromagnetic problems over a wide frequency range via the Padé approximation. *Computer Methods in Applied Mech. and Eng.*, 169(3–4):263–277, 1999.
 - [28] The MathWorks, Inc. *MATLAB 7*, September 2004.
 - [29] K. Meerbergen. The solution of parametrized symmetric linear systems. *SIAM J. Matrix Anal. Appl.*, 24(4):1038–1059, 2003.
 - [30] 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.
 - [31] 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.
 - [32] V. Simoncini. On the numerical solution of $AX - XB = C$. *BIT*, 36(4):814–830, 1996.
 - [33] V. Simoncini. Restarted full orthogonalization method for shifted linear systems. *BIT, Numerical Mathematics*, 43:459–466, 2003.
 - [34] V. Simoncini. A new iterative method for solving large-scale Lyapunov matrix equations. *SIAM J. Sci. Comput.*, 29(3):1268–1288, 2007.
 - [35] V. Simoncini and D. B. Szyld. Recent computational developments in Krylov subspace methods for linear systems. *J. Numerical Linear Algebra with Appl*, 14(1):1–59, 2007.
 - [36] T. Stykel and V. Simoncini. Krylov subspace methods for projected Lyapunov equations, 2009. In preparation.