KRYLOV SUBSPACE METHODS FOR LARGE SCALE CONSTRAINED SYLVESTER EQUATIONS*

STEPHEN SHANK† AND VALERIA SIMONCINI‡

Abstract. We consider the numerical approximation to the solution of the matrix equation $A_1X + XA_2 - YC = 0$ in the unknown matrices X, Y, under the constraint $XB = 0$, with A_1, A_2 of large dimensions. We propose a new formulation of the problem that entails the numerical solution of an unconstrained Sylvester equation. The spectral properties of the resulting coefficient matrices call for appropriately designed variants of projection-type methods. To this end, we propose new enriched approximation spaces, and provide experimental evidence of their effectiveness on benchmark problems. The application to a control problem is also described.

Key words. Matrix equation, Galerkin method, Krylov subspace, Constrained Sylvester equations.

AMS subject classifications. 93B40, 65F30, 15A06.

1. Introduction.

We consider the following system of matrix equations,

$$A_1X + XA_2 - YC = 0 \tag{1.1}$$

$$XB = 0 \tag{1.2}$$

for given $A_1 \in \mathbb{R}^{n_1\times n_1}$, $A_2 \in \mathbb{R}^{n_2\times n_2}$, $B \in \mathbb{R}^{n_2\times p}$ and $C \in \mathbb{R}^{m\times n_2}$, with unknowns $X \in \mathbb{R}^{n_1\times n_2}$ and $Y \in \mathbb{R}^{n_1\times m}$, where $p < m \ll \min\{n_1, n_2\}$. We assume that A_1 and A_2 are large, sparse, and nonsingular, and that B, C and CB have full rank; in the following we shall refer to this system as a constrained Sylvester equation, with the second equation, $XB = 0$, acting as a constraint. Sylvester equations arise in many areas of control theory and engineering tracking problems, and as workhorse of many numerical methods such as those used in linear and nonlinear eigenvalue problems and certain differential equations; see, e.g., [2],[7],[12],[32],[22]. The above constrained version appears for instance in the design of reduced order observers which achieve precise loop transfer recovery [5].

The two equations in (1.1) may also be viewed as a homogeneous system in the two unknown matrices X and Y, from which it readily follows that the matrix problem has a whole family of solutions. Indeed, by means of the Kronecker product, problem (1.1)-(1.2) can be rewritten as a very large homogeneous linear system, whose unknown vector contains the columns of X and Y, stacked one below the other. The resulting matrix is of size $n_1(p + n_2) \times n_1(m + n_2)$, and therefore overdetermined; thus when a nontrivial solution exists, this is not unique. We note that in general, the Kronecker product formulation is only of interest from a theoretical view point for n_1, n_2 large, since the resulting problem is of exploding size.

In [4], the authors describe a direct method for solving (1.1)-(1.2) when n_2 is small[1]. The method essentially amounts to a change of variables and the formulation of an unconstrained equation for the new variable. However, the method requires a full QR factorization of the matrix B, which is computationally expensive for large n_2. To the best of our knowledge, we are unaware of any method in the literature

*Version of February 5, 2013

†Department of Mathematics, Temple University, Philadelphia, PA 19122 (sshank@temple.edu)

‡Dipartimento di Matematica, Università di Bologna, Piazza di Porta S. Donato, 5,40127 Bologna, Italy (valeria.simoncini@unibo.it).

[1]The name *constrained Sylvester equation* is borrowed from that article.

for solving (1.1)-(1.2) in the large scale setting, at a computational cost and memory requirements that grow only linearly with the problem dimensions n_1 and n_2. We aim to fill this gap. We show that an unconstrained equation can be formulated for the unknown X whose solution automatically satisfies the constraint. The new formulation can be stated as the following standard unconstrained Sylvester equation

$$\mathbf{A}\mathbf{X} + \mathbf{X}\mathbf{B} + \mathbf{E}\mathbf{F}^T = 0, \tag{1.3}$$

where $\mathbf{A} \in \mathbb{R}^{n_1 \times n_1}$, $\mathbf{B} \in \mathbb{R}^{n_2 \times n_2}$, $\mathbf{E} \in \mathbb{R}^{n_1 \times r}$ and $\mathbf{F} \in \mathbb{R}^{n_2 \times r}$ for the unknown $\mathbf{X} \in \mathbb{R}^{n_1 \times n_2}$ with $r \ll n_1, n_2$. Numerical methods for (1.3) with large $\mathbf{A}$, $\mathbf{B}$ have been studied extensively; see, e.g., [2], [8] and other references later in the text. One class of iterative methods is based on projecting the equation into a well chosen approximation subspace and solving a problem of reduced size [30], [27], [28], [36], [37].

We describe how projection-type approaches based on Krylov subspaces can be adapted to effectively handle this new formulation when n_1 and n_2 are large. In particular, since one of the coefficient matrices of the transformed problem is singular, a new strategy for using enriched (extended and augmented) Krylov subspaces is devised. Our numerical experiments on benchmark problems seem to show that the new formulation can be effectively treated with powerful projection spaces, so that small approximation spaces are required to obtain a rather accurate solution.

The outline of the paper is as follows. In section 2, the original method of Barlow et al. ([4]) for solving constrained Sylvester equations is described, along with a proposed modification that is more amenable to iterative solution techniques for the large scale case. Section 3 recalls existing projection methods for Sylvester's equation. Section 3.1 explains how these can be efficiently applied to solve the new unconstrained equation, with appropriate new augmentation strategies to overcome the singularity of the obtained problem. Section 3.2 summarizes the algorithm together with some computational considerations. Section 4 contains some numerical results, while section 5 provides a short description of the control setting where such a constrained matrix equation could arise. Finally, section 6 gives concluding remarks.

Throughout the paper the following notation is used. We denote by range(M) the space spanned by the columns of the matrix M. I_n and 0_n denote the $n \times n$ identity and zero matrices, respectively, and $0_{n\times m}$ denotes the $n \times m$ zero matrix; subscripts will be omitted when clear from the context. $E_{m:r}$ will denote the last r columns of the $mr \times mr$ identity matrix.

2. A new formulation for the constrained Sylvester equations. In [4], the authors describe a direct method for solving (1.1)-(1.2) when the involved matrices have small dimensions. The proposed procedure transforms the original coupled equations into a single unconstrained Sylvester equation in one variable, which can be solved with available methods. We summarize their formulation in the following proposition. We include the proof because it is insightful for later developments.

PROPOSITION 2.1. ([4]) *Let* $B = \begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} R_B \\ 0 \end{bmatrix}$ *denote the QR factorization of* B, *where* $U_1 \in \mathbb{R}^{n_2 \times p}$, $U_2 \in \mathbb{R}^{n_2 \times (n_2 - p)}$, *and* $R_B \in \mathbb{R}^{p \times p}$, *with* $p < m$. *Let also*

$$CU_1 = \begin{bmatrix} Q_1 & Q_2 \end{bmatrix} \begin{bmatrix} R \\ 0 \end{bmatrix} = Q \begin{bmatrix} R \\ 0 \end{bmatrix}$$

denote the QR factorization of CU_1, *with* R *having full rank* p. *Then, a solution pair* (X, Y) *to* (1.1)-(1.2) *can be written as* $Y = [\hat{Y}_1, \hat{Y}_2]Q^T$ *with* $\hat{Y}_2$ *arbitrary,* $X = ZU_2^T$,

where $Z \in \mathbb{R}^{n_1 \times (n_2-p)}$ solves the equation

$$A_1 Z + Z\left((U_2^T A_2 U_2) - (U_2^T A_2 U_1) R^{-1} Q_1^T C U_2\right) = \hat{Y}_2 Q_2^T C U_2, \tag{2.1}$$

and $\hat{Y}_1 = Z(U_2^T A_2 U_1) R^{-1}$.

Proof. The constraint (1.2) implies that any solution X can be written as $X = ZU_2^T$ for some matrix $Z \in \mathbb{R}^{n_1 \times (n_2-p)}$. The equation (1.1) is then post-multiplied by U_1 and U_2, yielding

$$Z(U_2^T A_2 U_1) = YCU_1, \qquad A_1 Z + Z(U_2^T A_2 U_2) = YCU_2. \tag{2.2}$$

Define $[\hat{Y}_1, \hat{Y}_2] = YQ$ for $\hat{Y}_1 \in \mathbb{R}^{n_1 \times p}$ and $\hat{Y}_2 \in \mathbb{R}^{n_1 \times (m-p)}$. We can use the first equation in (2.2) to express part of Y in terms of Z; namely, since

$$Z(U_2^T A_2 U_1) = YCU_1 = YQQ^T CU_1 = [\hat{Y}_1, \hat{Y}_2] \begin{bmatrix} R \\ 0 \end{bmatrix} = \hat{Y}_1 R,$$

it follows that $\hat{Y}_1 = Z(U_2^T A_2 U_1) R^{-1}$. With this expression, the second equation in (2.2) is used to formulate an unconstrained equation for Z: since

$$YCU_2 = YQQ^T CU_2 = [\hat{Y}_1, \hat{Y}_2] \begin{bmatrix} Q_1^T \\ Q_2^T \end{bmatrix} CU_2 = \left(Z(U_2^T A_2 U_1) R^{-1} Q_1^T + \hat{Y}_2 Q_2^T\right) CU_2,$$

it follows that Z must satisfy

$$A_1 Z + Z\left((U_2^T A_2 U_2) - (U_2^T A_2 U_1) R^{-1} Q_1^T C U_2\right) = \hat{Y}_2 Q_2^T C U_2. \tag{2.3}$$

The entries of $\hat{Y}_2$ are not forced by the procedure, and can be chosen at random. □

It can be readily shown that the result in Proposition 2.1 is in fact an equivalence; specifically, if $\hat{Y}_2 \in \mathbb{R}^{n_1 \times (m-p)}$ is arbitrary, Z solves (2.3), and $\hat{Y}_1$, X and Y are defined as stated, then (X, Y) satisfies the system (1.1)-(1.2). We also remark that the arbitrariness of $\hat{Y}_2$ reflects the number of degrees of freedom of the system.

A disadvantage of the approach described above is that the use of a full QR factorization of the matrix B is too costly for large n_2, thus making the whole procedure not amenable to attack large problems. Indeed, in our setting $m, p \ll n_2$ with, in fact, often $m, p \leq 10$. Assuming such conditions on m, p and n_2, we propose a modification of the above method, showing that the unknown X satisfies a different unconstrained equation, that can be solved numerically without relying on computing the expensive U_2 term. It is worth mentioning that in our treatment a full QR factorization of the matrix $CU_1 \in \mathbb{R}^{m \times p}$ is still computed, but due to the extremely limited size of m and p, its cost remains very low.

Factoring out some terms in (2.1) enables one to rewrite this as

$$A_1 Z + ZU_2^T A_2 \left(I - U_1 R^{-1} Q_1^T C\right) U_2 = \hat{Y}_2 Q_2^T C U_2. \tag{2.4}$$

Right-multiplying by the full row rank matrix U_2^T and recalling that $X = ZU_2^T$, we obtain the equivalent equation

$$A_1 X + X A_2 \left(I - U_1 R^{-1} Q_1^T C\right) \Pi = \hat{Y}_2 Q_2^T C \Pi, \tag{2.5}$$

where $\Pi = U_2 U_2^T$, and it is an orthogonal projector onto $\text{null}(B^T)$. This simple transformation enables one to rewrite the new equation in terms of the coefficient

matrices of the original equation and projectors based on the problem data, and is summarized in the following proposition.

PROPOSITION 2.2. *With the notation above, suppose (Z, Y) are as in Proposition 2.1. Then the matrix $P = U_1 R^{-1} Q_1^T C$ is a projector onto range(B) and orthogonal to range$(C^T CB)$. Moreover, the matrix $X = ZU_2^T$ solves the unconstrained Sylvester equation*

$$A_1 X + X A_2 (I - P)\Pi = \hat{Y}_2 Q_2^T C \Pi, \tag{2.6}$$

and $\widehat{Y}_1 = X A_2 U_1 R^{-1}$.

Proof. Let $M = U_1 R^{-1}$ and $N = C^T Q_1$, so that $P = MN^T$. Then range(M)= range(U_1)=range(B). Moreover, from $N = C^T CB(RR_B)^{-1}$ it also follows that range(N) = range$(C^T CB)$. Using (2.1), $N^T M = Q_1^T C U_1 R^{-1} = Q_1^T Q_1 R R^{-1} = I$, so that $P = M(N^T M)^{-1} N^T$, i.e., P is a projector. Equation (2.6) coincides with (2.5) with P as defined. Finally, $\hat{Y}_1 = Z(U_2^T A_2 U_1) R^{-1} = X A_2 U_1 R^{-1}$. ∎

The Sylvester equation in (2.6) can be stated in the standard form (1.3). Indeed, let us set $\widehat{Y}_2 = \widehat{Y}_{2,1} \widehat{Y}_{2,2}^T$. Then, we can let

$$\mathbf{A} = A_1, \quad \mathbf{B} = A_2(I - P)\Pi, \quad \mathbf{E} = \hat{Y}_{2,1}, \quad \mathbf{F} = \Pi C^T Q_2 \widehat{Y}_{2,2}. \tag{2.7}$$

Other factorizations of $\hat{Y}_2 Q_2^T C \Pi = \mathbf{E}\mathbf{F}^T$ could be considered, however we found that since $\widehat{Y}_2$ can be chosen arbitrarily, it may be computationally advantageous to choose a rank-one matrix. Indeed, with these choices both $\mathbf{E}$ and $\mathbf{F}$ will be vectors, with potential significant computational and memory savings in the solution of the Sylvester equation in the large scale case. In section 3.1 we describe how this reformulation of the problem is much more accessible to treatment by Krylov subspace methods.

From a computational standpoint, we note that in solving (2.6), the action of Π on a vector may be computed cheaply, as U_1 can be obtained with $\mathcal{O}(n_2)$ complexity; moreover, Π may be replaced by the complementary projector $I - U_1 U_1^T$, whose action only involves $\mathcal{O}(n_2)$ computations. Analogously, the action of $I - P$ to a vector can exploit the low rank of P.

A solution of the new unconstrained equation enjoys the property of automatically satisfying the constraint condition, as described in the following corollary.

COROLLARY 2.3. *Suppose that X solves* (2.6). *Then $XB = 0$.*

Proof. Since $\Pi B = 0$, right-multiplication of (2.6) by B yields $A_1 XB = 0$, so the result follows immediately from the nonsingularity of A_1. ∎

3. Projection methods for Sylvester equations. In this section we recall a general framework for Krylov subspace methods as applied to the Sylvester equation (1.3).

Suppose we have a sequence of subspaces $\mathcal{V}_k$ of $\mathbb{R}^n$ which increase in dimension by a fixed amount[2] r_V, i.e., $\dim(\mathcal{V}_{k+1}) = \dim(\mathcal{V}_k) + r_V$ Let $\mathbb{V}_k \in \mathbb{R}^{n \times k r_V}$ denote a matrix whose columns form an orthonormal basis for $\mathcal{V}_k$, and let V_ℓ denote the ℓth block of $\mathbb{V}_k$, so that $\mathbb{V}_k = \begin{bmatrix} V_1 & V_2 & \cdots & V_k \end{bmatrix}$. We also define the matrix

$$\underline{\mathbf{T}}_k^{(A)} = \mathbb{V}_{k+1}^T \mathbf{A} \mathbb{V}_k = \begin{bmatrix} \mathbf{T}_k^{(A)} \\ V_{k+1}^T \mathbf{A} \mathbb{V}_k \end{bmatrix},$$

[2]In actual practice, the basis may not grow by a fixed amount at each step. For example, a numerical loss of rank in the basis may occur, and technically $\dim(\mathcal{V}_k) = r_V(k)$ is some function of k. We will ignore such technical details for ease of exposition.

whose (k,ℓ) block is denoted by $T^{(A)}_{k,\ell}$. Suppose we also have another such sequence of subspaces $\mathcal{W}_k$ of $\mathbb{R}^{n_2}$ growing by dimension r_W, with the columns of $\mathbb{W}_k \in \mathbb{R}^{s\times n_2 r_W}$ forming an orthonormal basis for $\mathcal{W}_k$. In order to formulate a Galerkin projection method, we need to define a search and constraint space.

DEFINITION 1. $\mathcal{S}(\mathcal{V}_k,\mathcal{W}_k) = \left\{ \mathbb{V}_k \widetilde{X} \mathbb{W}_k^T : \widetilde{X} \in \mathbb{R}^{kr_V \times kr_W} \right\}$.

We write $\mathcal{S}_k = \mathcal{S}(\mathcal{V}_k,\mathcal{W}_k)$ when there is no ambiguity regarding the underlying spaces. In fact, we could more generally define $\mathcal{S}(\mathcal{V}_{\ell_V},\mathcal{W}_{\ell_W})$, with ℓ_V and ℓ_W not necessarily equal. Although we shall not dwell on this possibility in this paper, working with different space dimensions may be rewarding when the two operators have different spectral properties, so that a larger approximation space for the "harder" operator may speed up convergence; see the recent analysis in [9]. From a computational stand point there are no extra difficulties in handling different values of ℓ_V and ℓ_W.

We define the standard block Krylov space as

$$\mathcal{K}_k(\mathbf{A},\mathbf{E}) = \text{range}\Big([\mathbf{E}, \mathbf{AE}, \cdots, \mathbf{A}^{k-1}\mathbf{E}]\Big),$$

where it holds that $r_V = r$. A possible choice for the underlying subspaces of $\mathcal{S}_k$ is then $\mathcal{V}_k = \mathcal{K}_k(\mathbf{A},\mathbf{E})$ and $\mathcal{W}_k = \mathcal{K}_k(\mathbf{B}^T,\mathbf{F})$ [29], which may be motivated by truncating a closed form solution of (1.3), found in [19]. When one attempts to either accelerate convergence or lessen memory demands of the standard Krylov subspace approach, standard tools available in the case of linear systems like preconditioning or restarting are nontrivial for linear matrix equations. Much of the recent progress has come in introducing new, richer spaces to project into. In [20], the authors introduced the notion of an "extended" Krylov subspace, and in [37] this notion was used to solve large scale Lyapunov equations. To summarize, one defines the extended block Krylov subspace as

$$\mathcal{K}^{\text{ext}}_k(\mathbf{A},\mathbf{E}) = \mathcal{K}_k(\mathbf{A},\mathbf{E}) + \mathcal{K}_k(\mathbf{A}^{-1},\mathbf{A}^{-1}\mathbf{E}),$$

where we have $r_V = 2r$. One then sets $\mathcal{V}_k = \mathcal{K}^{\text{ext}}_k(\mathbf{A},\mathbf{E})$ and $\mathcal{W}_k = \mathcal{K}^{\text{ext}}_k(\mathbf{B}^T,\mathbf{F})$, though in principle any pair of such spaces (standard or extended) or a combination of them, can be used. Bases for both the standard and extended spaces can be generated iteratively by an orthonormalization scheme such as block Arnoldi [35].

Once the spaces are generated, an approximation $X_k \in \mathcal{S}_k$ is determined by imposing some additional condition. Let $R_k = \mathbf{A}X_k + X_k\mathbf{B} + \mathbf{EF}^T$ be the associated residual matrix. To determine $\widetilde{X}_k$ in $X_k = \mathbb{V}_k\widetilde{X}_k\mathbb{W}_k^T$ we impose the *Galerkin condition*, which requires the residual matrix to be orthogonal to the approximation subspace $\mathcal{S}_k$, where the orthogonality is with respect to the Frobenius inner product. In matrix terms, it can be shown that this condition can be rewritten as

$$\mathbb{V}_k^T R_k \mathbb{W}_k = 0. \tag{3.1}$$

Upon recalling the definition of $\mathbf{T}^{(A)}_k$ and further defining $\mathbf{T}^{(B)}_k = \mathbb{W}_k^T\mathbf{B}^T\mathbb{W}_k$, it can be shown that (3.1) actually provides us with a smaller Sylvester equation for $\widetilde{X}_k$:

$$\mathbf{T}^{(A)}_k\widetilde{X}_k + \widetilde{X}_k\left(\mathbf{T}^{(B)}_k\right)^T + \left(\mathbb{V}_k^T\mathbf{E}\right)\left(\mathbb{W}_k^T\mathbf{F}\right)^T = 0. \tag{3.2}$$

Note that all quantities above, including $\mathbb{V}_k^T\mathbf{E}$ and $\mathbb{W}_k^T\mathbf{F}$, can be either obtained as a byproduct of the basis orthogonalization process, or can be updated as the two

spaces grow. This smaller equation can again be solved by standard Schur form based methods; see for instance [6], [23], [38].

The whole procedure terminates when k is large enough so that, say, the norm of the residual matrix is below some a-priori selected threshold. More precisely, the following stopping strategy involving the backward error can be used

$$\rho_k < \tau, \qquad \text{with} \quad \rho_k := \frac{\|R_k\|}{\|X_k\|(\|\mathbf{A}\| + \|\mathbf{B}\|) + \|\mathbf{E}\|\|\mathbf{F}\|}, \tag{3.3}$$

where $\|\cdot\|$ denotes the Frobenius norm. The use of the backward error accounts for possible unbalance in the magnitudes of the matrices involved. The norms $\|X_k\|$ and $\|R_k\|$ may be calculated without forming the large and dense matrices. First, observe that $\|X_k\| = \|\widetilde{X}_k\|$ by standard properties of the Frobenius norm. Concerning $\|R_k\|$, one always has (see, e.g., [33])

$$R_k = \begin{bmatrix}\mathbb{V}_k & \mathbf{A}\mathbb{V}_k\end{bmatrix} \begin{bmatrix} \mathbb{V}_m^T\mathbf{E}(\mathbb{W}_k^T\mathbf{F})^T & \widetilde{X}_k \\ \widetilde{X}_k & 0 \end{bmatrix} \begin{bmatrix} \mathbb{W}_k^T \\ \mathbb{W}_k^T\mathbf{B}\end{bmatrix}, \tag{3.4}$$

which may be exploited by progressively computing QR decompositions as the iterations of a given method proceeds; specifically, if one computes $\begin{bmatrix}\mathbb{V}_k & \mathbf{A}\mathbb{V}_k\end{bmatrix} = Q_k^{\mathbf{A}} R_k^{\mathbf{A}}$ and $\begin{bmatrix}\mathbb{W}_k & \mathbf{B}\mathbb{W}_k\end{bmatrix} = Q_k^{\mathbf{B}} R_k^{\mathbf{B}}$, it readily follows that

$$\|R_k\| = \left\| R_k^{\mathbf{A}} \begin{bmatrix} \mathbb{V}_k^T\mathbf{E}(\mathbb{W}_k^T\mathbf{F})^T & \widetilde{X}_k \\ \widetilde{X}_k & 0 \end{bmatrix} (R_k^{\mathbf{B}})^T \right\| \tag{3.5}$$

In practice, there is a slight twist; since one wants to compute these decompositions as the iterations proceed, typically one computes the following QR decompositions:

$$\begin{bmatrix}V_1 & \mathbf{A}V_1 & V_2 & \mathbf{A}V_2 & \cdots & V_k & \mathbf{A}V_k\end{bmatrix} = Q_k^{\mathbf{A}} R_k^{\mathbf{A}},$$
$$\begin{bmatrix}W_1 & \mathbf{B}^TW_1 & W_2 & \mathbf{B}^TW_2 & \cdots & W_k & \mathbf{B}^TW_k\end{bmatrix} = Q_k^{\mathbf{B}} R_k^{\mathbf{B}},$$

so that $\begin{bmatrix}\mathbb{V}_k & \mathbf{A}\mathbb{V}_k\end{bmatrix} = Q_k^{\mathbf{A}} R_k^{\mathbf{A}} \Omega_k$ and $\begin{bmatrix}\mathbb{W}_k & \mathbf{B}\mathbb{W}_k\end{bmatrix} = Q_k^{\mathbf{B}} R_k^{\mathbf{B}} \Omega_k$ for an appropriate permutation matrix Ω_k. This can be done progressively as the iterative process proceeds, and the calculation of $\|R_k\|$ is done analogously as in (3.5).

When a standard or extended Krylov subspace is used for $\mathcal{V}$ and $\mathcal{W}$, one has an Arnoldi relation of the form $\mathbf{A}\mathbb{V}_k = \mathbb{V}_{k+1}\underline{\mathbf{T}}_k^{(A)} = \mathbb{V}_k\mathbf{T}_k^{(A)} + V_{k+1}T_{k+1,k}^{(A)}E_{k:r_V}^T$, and analogously for $\mathbf{B}^T$ and $\mathbb{W}_k$, from which it follows that

$$\begin{bmatrix}\mathbb{V}_k & \mathbf{A}\mathbb{V}_k\end{bmatrix} = \mathbb{V}_{k+1} \begin{bmatrix} I_{kr_V} & \mathbf{T}_k^{(A)} \\ 0_{r_V\times kr_V} & T_{k+1,k}^{(A)} \end{bmatrix}, \tag{3.6}$$

$$\begin{bmatrix}\mathbb{W}_k & \mathbf{B}^T\mathbb{W}_k\end{bmatrix} = \mathbb{W}_{k+1} \begin{bmatrix} I_{kr_W} & \mathbf{T}_k^{(B)} \\ 0_{r_W\times kr_W} & T_{k+1,k}^{(B)} \end{bmatrix}. \tag{3.7}$$

Plugging (3.6)-(3.7) into (3.4), multiplying the resulting matrices, using the fact that $\widetilde{X}_k$ satisfies (3.2), taking norms and finally using standard properties of the Frobenius norm enables one to write

$$\|R_k\| = \left(\|T_{k+1,k}^{(A)}E_{k:r_V}^T\widetilde{X}_k\|^2 + \|\widetilde{X}_kE_{k:r_W}(T_{k+1,k}^{(B)})^T\|^2\right)^{\frac{1}{2}}. \tag{3.8}$$

Note that when only one space, say $\mathcal{W}_k$, has such an Arnoldi relation, a hybrid approach is possible; specifically, plugging (3.7) to (3.4) still yields an efficient method of cheaply calculating the required absolute residual norm.

3.1. Application to the new formulation. With the definitions in (2.7), the Sylvester equation (2.6) may be treated by projecting the matrix equation onto an appropriate Krylov-type subspaces. The quantities $\mathbf{E}$ and $\mathbf{F}$ in (2.7) may be computed a-priori with $\mathcal{O}(n)$ complexity, and the action of $\mathbf{A}$ and $\mathbf{B}^T$ on a vector can be computed with $\mathcal{O}(n)$ complexity at each iteration for sparse A_1, A_2.

As described in the previous section, the solution $\mathbf{X}$ is approximated as $X_k = \mathbb{V}_k \widetilde{X}_k \mathbb{W}_k^T$. For the approximate solution to satisfy the desired constraint equation $X_k B = 0$ it is sufficient that $\mathbb{W}_k^T B = 0$, that is, the right approximation space be orthogonal to B. Such property is naturally maintained when using a standard Krylov subspace as right space, as the following proposition summarizes.

PROPOSITION 3.1. *Suppose that* $X_k \in \mathcal{S}\left(\mathcal{V}_k, \mathcal{K}_k(\mathbf{B}^T, \mathbf{F})\right)$ *for some* k *and* $\mathcal{V}_k$. *Then* $X_k B = 0$.

Proof. We have that range$(\mathbb{W}_k) = \mathcal{K}_k(\mathbf{B}^T, \mathbf{F})$ =range$([\mathbf{F}, \mathbf{B}^T\mathbf{F}, \dots, (\mathbf{B}^T)^{k-1}\mathbf{F}])$. Since $\mathbf{F}^T B = 0$ and $\mathbf{B}B = 0$, we obtain $\mathbb{W}_k^T B = 0$, so that $X_k B = \mathbb{V}\widetilde{X}\mathbb{W}_k^T B = 0$. ∎

The use of more advanced Krylov subspaces as $\mathbb{W}_k$, such as the extended Krylov subspace, is hindered by the singularity of $\mathbf{B} = A_2(I - P)\Pi$, which is due to the presence of the projectors. Nonetheless, we would like to build a richer projection space in analogy with the extended Krylov subspace. To this end, we next define a suitable nonsingular approximation to $\mathbf{B}$.

For a specifically selected $\sigma \in \mathbb{R}$, $\sigma < 0$, our *ideal* choice is

$$\widehat{\mathbf{B}}_{\sigma,\text{ideal}} := (\mathbf{B}^T + \sigma I)^{-1} = (\Pi(I - P^T)A_2^T + \sigma I)^{-1}.$$

The inverse can be explicitly computed by means of the Sherman-Morrison-Woodbury formula. Unfortunately, this procedure still involves computations with large and dense matrices, therefore it requires one further approximation step. In summary, we derive our $\widehat{\mathbf{B}}_\sigma$ as follows, recalling that $\Pi = U_2 U_2^T$. Using the Sherman-Morrison-Woodbury formula we write,

$$(\mathbf{B}^T + \sigma I)^{-1} = \frac{1}{\sigma} I - \frac{1}{\sigma^2} U_2 \left[I + \frac{1}{\sigma} U_2^T (I - P^T) A_2^T U_2 \right]^{-1} U_2^T (I - P^T) A_2^T$$

Let us set $P^T = P_1 P_2^T$ with $P_1 = C^T Q_1 R^{-T}$, $P_2^T = U_1^T$, and let $\widehat{P}_1 = (\sigma I + A_2^T)^{-1} P_1$. One more application of the Sherman-Morrison-Woodbury formula gives

$$\begin{aligned}
&\left[I + \frac{1}{\sigma} U_2^T (I - P^T) A_2^T U_2 \right]^{-1} = \\
&= \sigma \left[\sigma U_2^T U_2 + U_2^T (I - P^T) A_2^T U_2 \right]^{-1} = \sigma \left[U_2^T (\sigma I + (I - P^T) A_2^T) U_2 \right]^{-1} \\
&\approx \sigma U_2^T \left[\sigma I + (I - P^T) A_2^T \right]^{-1} U_2 = \sigma U_2^T \left[\sigma I + A_2^T - P^T A_2^T \right]^{-1} U_2 \\
&= \sigma U_2^T \left[(\sigma I + A_2^T)^{-1} + (\sigma I + A_2^T)^{-1} P_1 (I - P_2^T A_2^T (\sigma I + A_2^T)^{-1} P_1)^{-1} P_2^T A_2^T (\sigma I + A_2^T)^{-1} \right] U_2 \\
&= \sigma U_2^T \left[I + \widehat{P}_1 (I - P_2^T A_2^T \widehat{P}_1)^{-1} P_2^T A_2^T \right] (\sigma I + A_2^T)^{-1} U_2.
\end{aligned}$$

In general, the approximation in the third line of the display formula could be very rough. However, since U_2 usually spans almost the whole space, we expect the approximation to be of good quality. Our numerical results seem to confirm this expectation. Then we obtain

$$\begin{aligned}
(\mathbf{B}^T + \sigma I)^{-1} &\approx \frac{1}{\sigma} I - \frac{1}{\sigma} U_2 U_2^T \left[I + \widehat{P}_1 (I - P_2^T A_2^T \widehat{P}_1)^{-1} P_2^T A_2^T \right] (\sigma I + A_2^T)^{-1} U_2 U_2^T (I - P^T) A_2^T \\
&= \frac{1}{\sigma} I - \frac{1}{\sigma} \Pi \left[I + \widehat{P}_1 (I - P_2^T A_2^T \widehat{P}_1)^{-1} P_2^T A_2^T \right] (\sigma I + A_2^T)^{-1} \mathbf{B}^T =: \widehat{\mathbf{B}}_\sigma.
\end{aligned}$$

We note that the inner matrix $(I - P_2^T A_2^T \widehat{P}_1)$ is in general very small (namely $\mathcal{O}(1)$), so that its inversion is cheap, and could be explicitly formed once and for all at the beginning of the procedure. Moreover, the application of $\widehat{\mathbf{B}}_\sigma$ requires solving a system with the large and sparse matrix $A_2 + \sigma I$ at each iteration, whose computational cost depends on the sparsity pattern of A_2. Such cost is what one would be willing to afford when dealing with a regular extended Krylov subspace on A_2.

Given such a $\widehat{\mathbf{B}}_\sigma$, we then propose to build an "augmented" space as follows:

$$\mathcal{K}_k^{\text{aug}}(\mathbf{B}^T, \widehat{\mathbf{B}}_\sigma, \mathbf{F}) = \mathcal{K}_k(\mathbf{B}^T, \mathbf{F}) + \mathcal{K}_k(\widehat{\mathbf{B}}_\sigma, \widehat{\mathbf{B}}_\sigma \mathbf{F}). \tag{3.9}$$

This space is formally not an *extended* Krylov subspace, as $\widehat{\mathbf{B}}_\sigma$ is not the inverse of $\mathbf{B}^{-1}$; however, since $\widehat{\mathbf{B}}_\sigma$ approximates the invariant subspaces of $\mathbf{B}^T$, the term "augmented" seems to be more appropriate.

We explicitly observe that if $x = \Pi x$, then $\widehat{\mathbf{B}}_\sigma x = \Pi \widehat{\mathbf{B}}_\sigma x$, that is, the application of $\widehat{\mathbf{B}}_\sigma$ preserves the space constraint. Therefore, by imposing the right basis $\mathbb{W}_k$ to be such that $\text{range}(\mathbb{W}_k) = \mathcal{K}_k^{\text{aug}}(\mathbf{B}^T, \widehat{\mathbf{B}}_\sigma, \mathbf{F})$, we readily obtain $\mathbb{W}_k^T B = 0$, so that the approximate solution X_k naturally satisfies the constraint, also when using the enriched augmented space.

We are left with the selection of the shift σ. A large body of literature is available on the computation of an appropriate shift for rational Krylov subspaces; see, e.g., the references in [21]. Here the situation is somewhat simplified, since a single shift is used. Our numerical experiments have shown that the geometric mean of the real "spectral interval" is particularly well suited for intervals spanning several orders of magnitude, namely

$$\sigma := -(\alpha_1 \alpha_n)^{\frac{1}{2}},$$

where $|\lambda_1| \geq \cdots \geq |\lambda_n|$ are the eigenvalues of A_2, arranged in decreasing order, and $\alpha_j = \Re(\lambda_j)$, $j = 1, \ldots, n$. An estimate of these quantities usually suffices.

REMARK 3.2. Generalizations of our approach suggest themselves. For instance, the augmented space $\mathcal{K}_k^{\text{aug}}(\mathbf{B}^T, \widehat{\mathbf{B}}_\sigma, \mathbf{F})$ with our choice of $\widehat{\mathbf{B}}_\sigma$ can be generalized to an approximation of a regular *rational* Krylov subspace,

$$\text{range}([(\mathbf{B}^T + \sigma_1 I)^{-1}\mathbf{F}, \ldots, (\mathbf{B}^T + \sigma_k I)^{-1}\mathbf{F}]),$$

where the shift σ_i, $i = 1, \ldots, k$ varies at each iteration, and can be either chosen a-priori, or selected adaptively by means of a greedy algorithm. We refer to [21] and references therein for a more detailed discussion. The associated computational cost may increase significantly, as a different shifted system needs to be solved at each iteration. On the other hand, convergence speed, in terms of number of iterations, may be significantly improved [9], making the trade-off problem dependent. To keep the treatment concise, and because of our satisfactory numerical experiments with $\mathcal{K}_k^{\text{aug}}(\mathbf{B}^T, \widehat{\mathbf{B}}_\sigma, \mathbf{F})$, we decided not to pursue this generalization further in this manuscript. ∎

REMARK 3.3. A similar choice for $\widehat{\mathbf{B}}_\sigma$ could have been made by first shifting the original equation to the following equivalent one:

$$(\mathbf{A} - \sigma I)\mathbf{X} + \mathbf{X}(\mathbf{B} + \sigma I) + \mathbf{E}\mathbf{F}^T = 0$$

so that $(\mathbf{B} + \sigma I)$ is nonsingular. This is a feasible way to address the singularity of $\mathbf{B}$, though a sufficiently large shift to make $\mathbf{B}$ well conditioned could be deleterious for the conditioning of $\mathbf{A} - \sigma I$, thus not solving the problem in practice. ∎

3.2. The algorithm. In this section we give a sketch of the algorithm, together with some algorithmic details.

An important computational aspect concerns the choice of the matrix $\widehat{Y}_2$. Proposition 2.2 shows that $\widehat{Y}_2 = \widehat{Y}_{2,1}\widehat{Y}_{2,2}^T$ can be chosen freely. Moreover, from (2.7) we deduce that the number of columns of $\widehat{Y}_{2,1}$ and $\widehat{Y}_{2,2}$ determines the block size of the Krylov subspaces generated during the process. Therefore, as already mentioned, a computationally advantageous choice consists in selecting $\widehat{Y}_2$ to be of rank one, so that $\mathbf{E} = \widehat{Y}_{2,1}$ and $\mathbf{F} = \Pi C^T Q_2 \widehat{Y}_{2,2}$ both have a single column. We originally experimented with a larger number of columns for $\widehat{Y}_{2,1}$ and $\widehat{Y}_{2,2}$, however the computational cost of dealing with block methods was higher than the increase in convergence rate. Therefore, unless explicitly stated, in all our experiments we defined $\widehat{Y}_{2,1}$ and $\widehat{Y}_{2,2}$ to be rank-one matrices, and we chose them to have all components equal to one. It is also important to mention once again that different space dimensions could be used for the right and left spaces. Without extra information on the spectral properties of the matrices, we did not find any special reason to exploit this extra feature, although the code could be easily adapted to handle this case as well. We also emphasize that in the generic case, the standard Krylov subspace after k iteration has dimension rk, whereas an extended or augmented Krylov subspace has dimension $2rk$, if started with a matrix with r columns.

The Krylov subspace-based method for the Sylvester equation in (1.3) can be summarized as follows; we stress that except for the novelty of the singularity of one of the coefficient matrices, the method follows well known Galerkin-type approaches in the literature, as described in section 3; see, e.g., [11], [26], [24], [36] and references therein.

Algorithm 1 Galerkin projection method for Sylvester equation.

Given $\mathbf{A} = A_1$, $\mathbf{B} = A_2(I - P)\Pi$, $\mathbf{E} = \widehat{Y}_{2,1}$, $\mathbf{F} = \Pi C^T Q_2 \widehat{Y}_{2,2}$

1: Compute V_1 and W_1. Set $\mathbb{V}_1 = V_1$, $\mathbb{W}_1 = W_1$.
2: **For** $k = 2, \ldots,$ until convergence
3: Expand spaces $\mathcal{V}_k, \mathcal{W}_k$ (either $\mathcal{K}$ or $\mathcal{K}^{\text{ext}}$, $\mathcal{K}^{\text{aug}}$):
Compute V_k and expand $\mathbb{V}_k$. Compute W_k and expand $\mathbb{W}_k$,
Compute or update $\mathbf{T}_k^{(A)} = \mathbb{V}_k^T \mathbf{A} \mathbb{V}_k$, $\mathbf{T}_k^{(B)} = \mathbb{W}_k^T \mathbf{B} \mathbb{W}_k$
4: Solve $\mathbf{T}_k^{(A)} \widetilde{X}_k + \widetilde{X}_k \left(\mathbf{T}_k^{(B)}\right)^T + \left(\mathbb{V}_k^T \mathbf{E}\right)\left(\mathbb{W}_k^T \mathbf{F}\right)^T = 0$
5: EndFor

The expanded subspaces are either the standard Krylov subspace or the augmented or extended Krylov subspace; For the last two the choice depends on the singularity of the coefficient matrix, as described in the previous sections. In step 4, the small scale Sylvester equation is solved by a Schur-decomposition based method [6].

Concerning stopping criteria, we observe that the term $\|\mathbf{B}\|$ used in (3.3) is not readily available. Therefore, we propose stopping when

$$\widehat{\rho}_k < \tau \qquad \text{with} \quad \widehat{\rho}_k := \frac{\|R_k\|}{\|X_k\|\|\mathbf{A}\| + \|X_k\mathbf{B}\| + \|\mathbf{E}\|\|\mathbf{F}\|}, \tag{3.10}$$

for a prescribed tolerance τ; in our experiments we used $\tau = 10^{-12}$. Note that the term $\|X_k\mathbf{B}\| = \|\widetilde{X}_k \mathbb{W}_k \mathbf{B}^T\|$ can be computed at a reasonable cost and that $\rho_k \leq \widehat{\rho}_k$.

4. Numerical experiments. In this section we report on our numerical experience with the proposed formulation and methods. Here we focus on general datasets, whereas in section 5 we will discuss the application of this setting to a control problem.

We experimented with the use of standard and augmented Krylov subspaces. For the sake of clarity we mainly only report the use of the same type of space as left and right spaces (either standard for both $\mathbf{A}$ and $\mathbf{B}^T$ or extended for $\mathbf{A}$ and augmented for $\mathbf{B}^T$). In all examples we defined the matrices A_1 and A_2 so as to have similar large size; unless stated otherwise, the matrices B and C were taken from the dataset associated with A_1, when this was available.

In all plots the backward error in (3.10) is displayed, versus the space dimension as iterations proceed.

EXAMPLE 4.1. We consider the discretization of the Laplace operator, for a variety of dimensions and signs. In all cases, B is the first column of the identity matrix, and C is given by the first five rows of the identity matrix. Let Δ_n be the $n \times n$ matrix stemming from the 5-point stencil finite difference discretization of the Laplace operator on the unit square. The first example (leftmost plot of Figure 4.1) considers $A_1 = n_1\Delta_{n_1}$ and $A_2 = -\Delta_{n_2}$, with $n_1 = 324$ and $n_2 = 400$. Note that the two matrices have eigenvalues on different sides of the complex plane, however the scaling of A_1 avoids any instability problems in the computation. Both Standard and Augmented subspaces converge rapidly, with a comparable CPU time (not reported). Differences are much more pronounced for $n_1 = 2304$ and $n_2 = 2500$ (middle plot of Figure 4.1), for which the use of the enriched spaces significantly improves convergence, in terms of memory used. The computational cost of solving with a shifted version of A_2 is negligible. For larger matrix dimensions, the gap between the two curves substantially increases even more. Finally, the rightmost plot of Figure 4.1 shows the performance for the larger case, where however now $A_1 = -n_1\Delta_{n_1}$, so that both A_1 and A_2 have the same sign. Once again, the augmented space is able to capture good spectral information soon during the iterations, leading to fast convergence. Although we shall not pursue this issue further, we observe that the final subspace dimension of the enriched space method seems to be rather insensitive to the problem dimension, and this is typical of shift-and-invert Krylov subspaces when applied to functionals such as the Laplace operator [40].

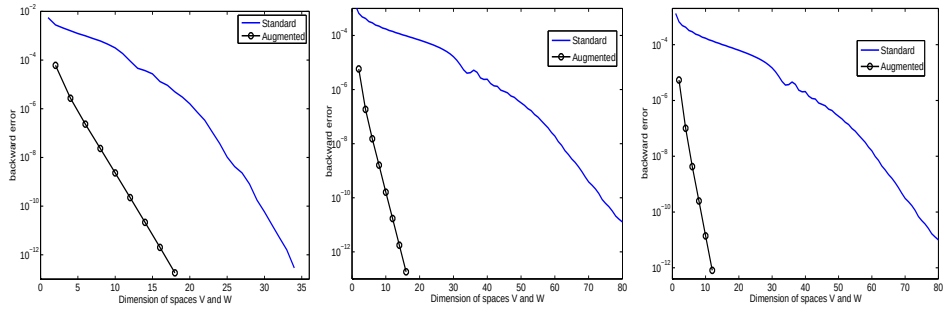

FIG. 4.1. *Example 4.1. Convergence history of Standard and Augmented Krylov subspace solvers.*

EXAMPLE 4.2. We consider the solution of (2.6), where A_2, B and C stem from the nonsymmetric CHIP dataset of the Oberwolfach collection [16]. The matrix A_1 was obtained as a finite difference discretization of the 2D Laplace operator, scaled so as to have the same Frobenius norm as A_2. The two matrices A_1 and A_2 have

size $n_1 = 19881$ and $n_2 = 20082$, respectively. The input matrix B has a single column, whereas C has five rows. Figure 4.2 reports on our experiments comparing two choices of projection spaces: standard Krylov subspace for both $\mathbf{A}, \mathbf{B}^T$, and extended (augmented) Krylov subspaces for $\mathbf{A}$ ($\mathbf{B}^T$). This experiment fully confirms our findings on a more realistic problem than the one of the previous example.

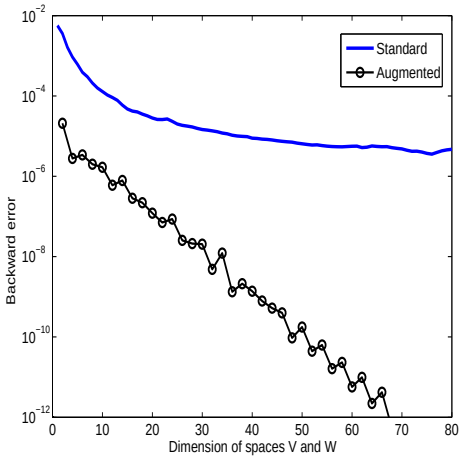

FIG. 4.2. *Example 4.2. Convergence history of Standard/Standard and Extended/Augmented Krylov subspace solvers.*

EXAMPLE 4.3. We consider for A_2 the 9669×9669 nonsymmetric and stable matrix FLOW from the Oberwolfach benchmark collection [16], with B having a single column, and C having five rows. The discretization of the Laplace operator, Δu was used for A_1, giving rise to a 9604×9604 stable matrix. The performance of standard and enriched space methods in terms of space dimension is reported in the left plot of Figure 4.3. The performance of the latter methods confirms our previous experiments. For this problem we further analyze the importance of selecting enriched spaces for *both* matrices $\mathbf{A}$ and $\mathbf{B}^T$. In the right plot of Figure 4.3 we report the convergence history when for the two spaces $\mathcal{V}_k$ and $\mathcal{W}_k$ are taken, respectively: extended/augmented, as before, extended/standard, standard/augmented. The combination of the two richer spaces provides the most effective setting. We also observe that although the use of the extended Krylov subspace for $\mathbf{A}$ seems to be very crucial to speed up convergence, the projection process becomes definitely competitive only when using the enriched (augmented) strategy also on $\mathbf{B}^T$.

5. Application to observer design in Control. The constrained Sylvester equation naturally enters in the design of a deterministic observer in the context of a classical loop transfer recovery [34]. Consider the continuous-time controllable and observable linear dynamical system

$$\begin{aligned} \dot{x} &= Ax + Bu, \quad x \in \mathbb{R}^{n_2}, \quad u \in \mathbb{R}^p, \\ y &= Cx, \quad y \in \mathbb{R}^m, \quad p < m < n, \end{aligned}$$

and its feedback control law $u = -K_c x$. An observer z of order $n - m$ can be written as ([31])

$$\dot{z} = Fz + (TB)u + LCx, \quad \widehat{x} = Nx + Mz, \tag{5.1}$$

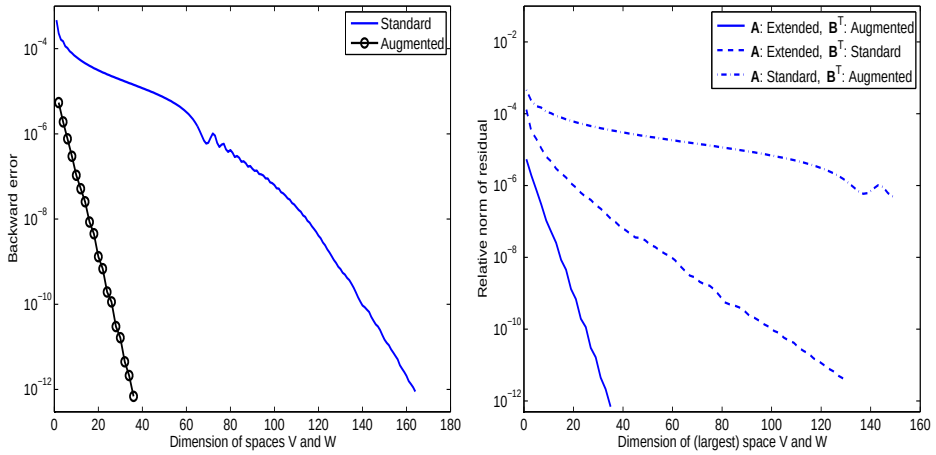

FIG. 4.3. *Example 4.3. Convergence history of Standard and Augmented Krylov subspace solvers.*

where $\widehat{x}$ is the estimated state; (5.1) provides a so-called Luenberger observer of order $n-m$ for linear feedback control lows. In fact, more complex, e.g. nonlinear, Luenberger observers can be derived; see, e.g., [25], [1], [3].

Necessary and sufficient conditions for this observer to be well defined are that (F, T, L, N, M) can be found such that they satisfy (see, e.g., [31])

$$TA - FT = LC, \quad N + MT = I; \tag{5.2}$$

Other choices can also be effective, see, e.g., [39]. Some of these matrices may be selected a-priori, such as $F \in \mathbb{R}^{(n-m)\times(n-m)}$. The remaining ones are computed accordingly. Natural conditions are that F be stable, and (F, L), (A^T, C^T) be observable; see, e.g., [2] for these definitions. For the Sylvester equation in (5.2) to have a unique solution it is only necessary that A and F have no common eigenvalues, therefore F may have any eigenvalues, as long as these are different from those of A. Such freedom has allowed the selection of particularly suitable structured matrices with pre-assigned spectrum; see, e.g., [10], [13]. In our experiments we did not dwell with the determination of F with a-priori chosen spectrum; instead, the matrix F was taken as the discretization of an elliptic operator, with known spectral interval for F. From a computational point of view, an important fact concerning the solvability of (5.2) is that for A and F stable, the operator $X \mapsto FX - XA$ may be very ill-conditioned, in case portions of the two spectra happen to be very close. Such risk should be taken into account when choosing F.

For achieving accurate loop transfer recovery, so that the observer-based system has the same robustness as the original system, a necessary (and also sufficient) condition is that $TB = 0$ [39], which gives a constraint for the Sylvester equation appearing in (5.2). Let $\bar{z} = z - Tx$ and $e = \widehat{x} - x$. Under these conditions, one obtains ([31])

$$\dot{\bar{z}} = F\bar{z}, \quad e = M\bar{z}.$$

As discussed in [31], if initially $\widehat{x}(0) = x(0)$, so that $\bar{z}(0) = 0$, then the equality will be maintained for all subsequent times. In general, we see that $e(t) \to 0$ as $t \to \infty$ if F is stable. We also recall from [31] that *the solution T must have rank great enough to guarantee the recovery of the unmeasurable state variables.* Here, the reference is

to the relation $u = -K_c x$, between the input and state variables. Such consideration appears to be crucial in the small scale case, where a full rank assumption on T is natural for an exact recovery; in particular, such assumption allows for a more general form of the equation defining the approximate state $\widehat{x}$ in (5.1); see, for instance, [4], [5], [15] and their references, and the analysis in [39].

In the large scale case, a full rank matrix T is very unlikely to be determined, unless the large matrix F is chosen ad-hoc. Memory limitations are also an issue. We stress that, as opposed to, e.g., [13], [17], [14], [18], the size of the matrix F is fixed and usually large, so that an a-priori pole assignment is infeasible. In this setting, the aim is thus to determine reduced systems for x and z, which should be able to preserve the main properties of the original state and observer systems, so as to conveniently approximate the, e.g., target, loop transfer function. With an appropriate determination of T, the estimated state $\widehat{x}$ will still be able to well represent the leading properties of the original system; the interpretation of the role of T in the large scale setting deserves further analysis, however its examination is beyond the scope of this paper.

With no pretense of being exhaustive, in the following we provide a demonstrative example of the applicability of our algorithm in the context of state estimation via these classical observers, with the substitutions $A_2 = A$ and $A_1 = -F$.

EXAMPLE 5.1. Here A_2 is the finite difference discretization of the operator $L(u) = (e^{-4xy}u_x)_x + (e^{4xy}u_y)_y$, $(x, y) \in (0, 1)^2$. The (scaled) symmetric matrix A_2 of size $n_2 = 6400$ has eigenvalues in $[-1.7061 \cdot 10^2, -4.7543 \cdot 10^{-3}]$. We considered C with $m = 10$ rows and B with $p = 5$ columns, corresponding to the first m and p columns of the identity matrix, respectively. The matrix $A_1 = -F$ is the finite difference discretization of the negative Laplace operator in $(0, 1)^2$, with eigenvalues[3] in $[1.9779 \cdot 10^1, 5.1100 \cdot 10^4]$; here we used $n_1 = n_2 - m = 6390$, and the matrix was obtained by using 90 and 71 interior nodes in the two directions, respectively. The performance of the methods with the use of standard and augmented spaces is reported in Figure 5.1. For the left plot an a-priori computed rank-one matrix $\widehat{Y}_2 = \widehat{Y}_{2,1}\widehat{Y}_{2,2}^T$ with $\widehat{Y}_{2,1}, \widehat{Y}_{2,2}$ of all ones was used, whereas in the right plot a rank-p matrix with random entries for $\widehat{Y}_{2,1}, \widehat{Y}_{2,2}$ was employed. Defining the relative rank of a matrix X as the number of singular values greater than 10^{-12} times the largest singular value, we notice that the augmented method delivered an approximate solution matrix $\widetilde{X}_k$ with relative rank 6 and 22, respectively, for the two choices of $\widehat{Y}_2$. Therefore, the rank-one choice should be preferred, at least in terms of memory requirements for the approximation space (cf. Figure 5.1) as well as for the factors of the solution matrix.

6. Conclusions. We have devised a new formulation of a constrained Sylvester matrix equation, which allows one to solve large scale problems by means of advanced projection-type methods. To be able to efficiently solve the resulting Sylvester equation, we have introduced a new augmented space, which overcomes the singularity of the coefficient matrix $\mathbf{B}$. Though the implementation is completely analogous to that of more familiar extended Krylov subspaces, the use of such an approximation to a shift-and-invert Krylov subspace as extension to the standard space, appears to be new. From a theoretical point of view, known results for rational Krylov subspace methods on the Sylvester equation can be applied to the "exact" augmented space $\mathcal{K}_k(\mathbf{B}^T, \mathbf{F}) + \mathcal{K}_k((\mathbf{B}^T + \sigma I)^{-1}, (\mathbf{B}^T + \sigma I)^{-1}\mathbf{F})$; see [9]. The space we actually

[3] Only the 10 smallest in magnitude eigenvalues of $-A_1$ interlace those of A_2, with apparent no convergence deterioration.

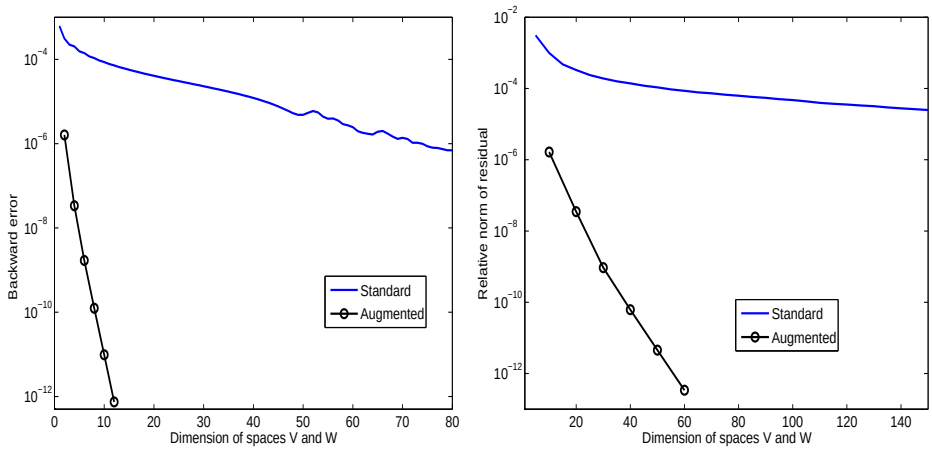

FIG. 5.1. *Example 5.1. Convergence history of Standard and Augmented Krylov subspace solvers. Left: rank-one* $\widehat{Y}_2$*. Right: rank-p* $\widehat{Y}_2 = \widehat{Y}_{2,1}\widehat{Y}_{2,2}^T$ *(random entries in* $\widehat{Y}_{2,1}$*,* $\widehat{Y}_{2,2}$*).*

adopt, $\mathcal{K}_k^{\text{aug}} = \mathcal{K}_k(\mathbf{B}^T, \mathbf{F}) + \mathcal{K}_k(\widehat{\mathbf{B}}_\sigma, \widehat{\mathbf{B}}_\sigma \mathbf{F})$, may somewhat differ from the exact case, depending on the properties of the matrix of inputs, B. Nonetheless, our promising numerical experiments seem to reinforce the intuition that the space generated with $\widehat{\mathbf{B}}_\sigma$ appropriately captures spectral information that is otherwise missed in the standard Krylov subspace $\mathcal{K}_k(\mathbf{B}^T, \mathbf{F})$ at an early stage.

Finally, we expect that our approach may be of value also when attacking the non-homogeneous form of the equation (1.1)-(1.2), which is currently a rather open problem in the large scale case.

Acknowledgments. This work started while the first author was visiting the Dipartimento di Matematica, Università di Bologna during the Spring semester 2012, with the support of the CAP/Atlantis grant. The first author also fully acknowledges the support from Daniel Szyld's National Science Foundation grant (number DMS-1115520).

REFERENCES

[1] V. Andrieu and L. Praly. Remarks on the existence of a Kazantzis-Kravaris/Luenberger observer. In *43rd IEEE Conference on Decision and Control*, pages 3874–3879, 2004.

[2] A. C. Antoulas. *Approximation of large-scale Dynamical Systems*. Advances in Design and Control. SIAM, Philadelphia, 2005.

[3] M. Ayati and H. Khaloozadeh. Designing a novel adaptive impulsive observer for nonlinear continuous systems using LMIs. *IEEE Trans. on Circuits and SystemsI*, 59(1), 2012.

[4] J. B. Barlow, M. M. Monahemi, and D. P. O'Leary. Constrained matrix Sylvester equations. *SIAM J. Matrix Anal. Appl.*, 13(1):1–9, 1992.

[5] J. B. Barlow, M. M. Monahemi, and D. P. O'Leary. Design of reduced-order observers with precise loop transfer recovery. *Journal of guidance, control, and dynamics*, 15(6):1320–1326, 1992.

[6] 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.

[7] 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.

[8] P. Benner, E. S. Quintana-Ortí, and G. Quintana-Ortí. Solving stable Sylvester equations via

rational iterative schemes. Technical Report 04-08, Technische Universität, Chemnitz, D, 2004. To appear in J. Scientific Computing.

[9] B. Beckermann. An Error Analysis for Rational Galerkin Projection Applied to the Sylvester Equation. *SIAM J. Numerical Analysis*, 49(6):2430–2450, 2011.

[10] C. H. Bischof, B. N. Datta, and A. Purkayastha. A Parallel Algorithm for the Multi-input Sylvester-Observer Equation. In *American Control Conference*, pages 152 – 156, 1992.

[11] A. Bouhamidi and K. Jbilou. A note on the numerical approximate solutions for generalized Sylvester matrix equations with applications. *Applied Math. Comp.*, 206:687–694, 2008.

[12] S. Brahma and B. Datta. An optimization approach for minimum norm and robust partial quadratic eigenvalue assignment problems for vibrating structures. *Journal of Sound and Vibration*, 324(3-5):471–489, 2009.

[13] D. Calvetti, B. Lewis, and L. Reichel. On the solution of large Sylvester-observer equations. *Numer. Linear Algebra Appl.*, 8:435–451, 2001.

[14] J. Carvalho, K. Datta, and Y. Hong. A new block algorithm for full-rank solution of the Sylvester-observer equation. *Automatic Control, IEEE Transactions on*, 48(12):2223–2228, Dec. 2003.

[15] B. M. Chen and Y.-L. Chen. Loop transfer recovery design via new observer-based and CSS architecture-based controllers. *Internat. J. Robust and Nonlinear Control*, 5:649–669, 1995.

[16] Collection. Oberwolfach model reduction benchmark collection, 2003. `http://www.imtek.de/simulation/benchmark`.

[17] B. N. Datta and C. Hetti. Generalized Arnoldi methods for the Sylvester-observer equation and the multi-input pole placement problem. In *Proceedings of the 36th Conference on Decision & Control*, San Diego, California USA, December 1997.

[18] B. N. Datta and Y. Saad. Arnoldi methods for large Sylvester-like observer matrix equations, and an associated algorithm for partial spectrum assignment. *Linear Algebra and Appl.*, 154–156:225–244, 1991.

[19] E. de Souza and S. P. Bhattacharyya. Controllability, observability and the solution of $AX - XB = C$. *Linear Algebra Appl.*, 39:167–188, 1981.

[20] 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.

[21] V. Druskin and V. Simoncini. Adaptive rational Krylov subspaces for large-scale dynamical systems. *Systems and Control Letters*, 60:546–560, 2011.

[22] M. Epton. Methods for the solution of ADXD-BXC=E and its application in the numerical solution of implicit ordinary differential equations. *BIT*, 20:341–345, 1980.

[23] G. H. Golub, S. Nash, and C. Van Loan. A Hessenberg-Schur method for the problem $AX + XB = C$. *IEEE Trans. Automat. Control*, 24(6):909–913, 1979.

[24] A. E. Guennouni, K. Jbilou, and A. Riquet. Block Krylov subspace methods for solving large Sylvester equations. *Numerical Algorithms*, 29:75–96, 2002.

[25] S. M. N. Hasan and I. Husain. A Luenberger-sliding mode observer for online parameter estimation and adaptation in high-performance induction motor drives. *IEEE Transactions on Industry Appl.*, 45(2), 2009.

[26] M. Heyouni. Extended Arnoldi Methods for Large Sylvester Matrix Equations. Technical report, L.M.P.A., 2008.

[27] M. Heyouni. Extended Arnoldi methods for large low-rank Sylvester matrix equations. *Appl. Numer. Math.*, 60(11):1171–1182, 2010.

[28] D. Y. Hu and L. Reichel. Krylov Subspace Methods for the Sylvester Equation. *Linear Algebra and Appl.*, 172:283–313, July 1992.

[29] D. Y. Hu and L. Reichel. Krylov-subspace methods for the Sylvester equation. *Linear Algebra Appl.*, 172:283–313, 1992. Second NIU Conference on Linear Algebra, Numerical Linear Algebra and Applications (DeKalb, IL, 1991).

[30] K. Jbilou. Low rank approximate solutions to large Sylvester matrix equations. *Applied Mathematics and Computation*, 177:365–376, 2006.

[31] D. Luenberger. Observers for multivariable systems. *IEEE Trans. Automat. Control*, 11:190–197, 1966.

[32] K. Meerbergen and A. Spence. Inverse iteration for purely imaginary eigenvalues with application to the detection of Hopf bifurcations in large scale problems. *SIAM Journal on Matrix Analysis and Applications*, 31(4):1982–1999, May 2010.

[33] T. Penzl. A cyclic low-rank Smith method for large sparse Lyapunov equations. *SIAM J. Scientific Computing*, 21(4):1401–1418, 2000.

[34] A. Radke and Z. Gao. A survey of state and disturbance observers for practitioners. In *Proceedings of the 2006 American Control Conference Minneapolis, Minnesota, USA, June 14-16, 2006*, 2006.

[35] Y. Saad. *Iterative methods for sparse linear systems*. Society for Industrial and Applied Mathematics, Philadelphia, PA, second edition, 2003.
[36] V. Simoncini. On the numerical solution of $AX - XB = C$. *BIT*, 36(4):814–830, 1996.
[37] V. Simoncini. A new iterative method for solving large-scale Lyapunov matrix equations. *SIAM J. Sci. Comput.*, 29(3):1268–1288, 2007.
[38] D. C. Sorensen and Y. Zhou. Direct methods for matrix Sylvester and Lyapunov equations. *J. Appl. Math.*, 6:277–303, 2003.
[39] C.-C. Tsui. A new approach to robust observer design. *Internat. J. Control*, 47:745–751, 1988.
[40] J. van den Eshof and M. Hochbruck. Preconditioning Lanczos approximations to the matrix exponential. *SIAM J. Scientific Computing*, 27(4):1438–1457, 2006.