Applications of Mathematics

Jan Brandts; Ricardo Reis da Silva
On the subspace projected approximate matrix method

ON THE SUBSPACE PROJECTED APPROXIMATE
MATRIX METHOD

Jan H. Brandts, Ricardo Reis da Silva, Amsterdam

Abstract. We provide a comparative study of the Subspace Projected Approximate Matrix method, abbreviated SPAM, which is a fairly recent iterative method of computing a few eigenvalues of a Hermitian matrix A. It falls in the category of inner-outer iteration methods and aims to reduce the costs of matrix-vector products with A within its inner iteration. This is done by choosing an approximation A_0 of A, and then, based on both A and A_0, to define a sequence $(A_k)_{k=0}^{n}$ of matrices that increasingly better approximate A as the process progresses. Then the matrix A_k is used in the kth inner iteration instead of A.

In spite of its main idea being refreshingly new and interesting, SPAM has not yet been studied in detail by the numerical linear algebra community. We would like to change this by explaining the method, and to show that for certain special choices for A_0, SPAM turns out to be mathematically equivalent to known eigenvalue methods. More sophisticated approximations A_0 turn SPAM into a boosted version of Lanczos, whereas it can also be interpreted as an attempt to enhance a certain instance of the preconditioned Jacobi-Davidson method.

Numerical experiments are performed that are specifically tailored to illustrate certain aspects of SPAM and its variations. For experiments that test the practical performance of SPAM in comparison with other methods, we refer to other sources. The main conclusion is that SPAM provides a natural transition between the Lanczos method and one-step preconditioned Jacobi-Davidson.

Keywords: Hermitian eigenproblem; Ritz-Galerkin approximation; subspace projected approximate matrix; Lanczos method; Jacobi-Davidson method

MSC 2010: 65F10, 65F35

1. Introduction

We provide a comparative study of SPAM [21]. SPAM, which stands for Subspace Projected Approximate Matrix, is a method for the computation of eigenvalues of a large Hermitian matrix A. Like the Davidson method [9], SPAM was originally

developed for matrices that arise from applications in Chemistry. It was only in [8], many years after its conception, that Davidson's method was given proper attention by the numerical linear algebra community. As far as we can tell, also SPAM has been neglected by the numerical linear algebra community. Moreover, even though a number of citations [6], [13], [16], [20], [31] within the Chemistry and Physics communities over the past years demonstrate awareness of its existence, no studies of its mathematical properties seems to exist.

SPAM belongs to the category of inner-outer iteration methods and is interpreted by its inventors as a modification of the above mentioned Davidson method. It is based on the following observation. Even when sparse, the computational effort necessary to carry out a matrix-vector multiplication with A can be significant and often represents the bottleneck of the total computational effort. SPAM reduces the costs of matrix-vector products with A by replacing its action within the inner iteration of the algorithm with a sparser or more structured approximation. By doing so, it attempts to slash the overall computational cost. The idea is not altogether new and is related to a certain type of preconditioning, called *one-step approximation* in the Jacobi-Davidson method. See Section 4.1 of [23]. There too, the matrix A is, in the inner iteration, replaced by a preconditioner. The originality of the approach in [21] lies in the fact that the action of the preconditioner is only applied to the subspace in which the action of A has not yet been computed in the outer iteration of the method. Consequently, the approximated action of A in the inner iterations is likely to become more and more accurate as the number of outer iterations increases. Note that nonetheless, only one approximate matrix A_0 is needed. Intuitively, one would expect that SPAM would outperform Jacobi-Davidson with one-step approximation.

Since the main idea of SPAM is potentially interesting, with links to other eigenvalue methods and selection techniques, we would like to bring it to the attention of the numerical linear algebra community and also to submit it to a more illustrative set of numerical experiments than those in [21]. In the experiments in [21], the overall efficiency of SPAM was the main interest instead of its mathematical position within the class of iterative eigensolvers. In particular, we would like to point out the similarities with and differences from strongly related and well-known iterative methods such as the Lanczos method [15] and the Jacobi-Davidson method of Sleijpen and van der Vorst [23]. In the context of applying the same idea to iteratively solving linear systems, see [5].

1.1. Outline of the results in this paper. We will show that for certain choices of the approximate matrix A_0, the SPAM method is mathematically equivalent to methods such as Lanczos [15] or the Riccati [4] method, another attempt to improve upon the Jacobi-Davidson [23] method. Further, we will see that a Schur

complement-based choice for the action of A outside the Ritz-Galerkin subspace that is being built in the outer iteration naturally leads to a connection with harmonic Rayleigh-Ritz [3] methods. Next, we show that choosing A_0 such that $A - A_0$ is positive semi-definite has, at least in theory, an agreeable effect on the approximations obtained in the inner iteration in comparison to choosing A_0 without such a restriction. Numerical experiments suggest that this also works through into the outer iteration. We comment on how such approximations A_0 can be obtained in the context of discretized elliptic PDEs [7], [30] but also from a purely algebraic point of view. Finally, we present a variety of detailed numerical illustrations of the performance of the method in comparison with the Lanczos and the Jacobi-Davidson method. These illustrations do not merely aim at showing that SPAM is a suitable method to solve eigenvalue problems (which was for a large part already taken care of in [21]), but to emphasize the role of approximations A_0 from below and to show the similarities and discrepancies as compared with Lanczos and Jacobi-Davidson, such that the SPAM method can be put into a proper perspective.

2. SPAM AND SOME OTHER SUBSPACE METHODS

Eigenvalue problems are among the most prolific topics in Numerical Linear Algebra [1], [11]. In particular, the continuous increase in matrix sizes turns the understanding of known iterative methods, as well as the development of more efficient ones, into an ample field of work within the global topic. Successful methods like the Implicitly Restarted Arnoldi [26] and the Krylov-Schur [28] methods and their symmetric counterparts are nowadays among the most competitive. For convenience and in order to set the ground for what is to come, we opt to outline the main concepts. For more detailed considerations on both theoretical and practical aspects of the numerical solution of large eigenvalue problems, we refer to [19], [12], [27], [29].

2.1. Ritz values and vectors. Throughout this paper, A is a Hermitian $n \times n$ matrix with eigenvalues

$$\lambda_n \leqslant \lambda_{n-1} \leqslant \ldots \leqslant \lambda_2 \leqslant \lambda_1. \tag{1}$$

Let V be an $n \times k$ matrix with mutually orthonormal columns and define the k-dimensional subspace $\mathcal{V}$ of $\mathbb{C}^n$ by

$$\mathcal{V} = \{Vy \mid y \in \mathbb{C}^k\}. \tag{2}$$

In the context of iterative methods, $\mathcal{V}$ is called the *search subspace*. Let $V_\perp$ be such that $(V|V_\perp)$ is unitary and write

$$(3) \qquad \widehat{A} = (V|V_\perp)^* A (V|V_\perp) = \begin{bmatrix} M & R^* \\ R & S \end{bmatrix}.$$

The eigenvalues of the $k \times k$ matrix $M = V^* A V$,

$$(4) \qquad \mu_k \leqslant \mu_{k-1} \leqslant \ldots \leqslant \mu_2 \leqslant \mu_1,$$

are called the *Ritz values* of A with respect to $\mathcal{V}$. The vectors $u_i = V z_i$, where $z_1, \ldots, z_k$ is an orthonormal basis for $\mathbb{C}^k$ consisting of eigenvectors of M belonging to the corresponding μ_i, are the *Ritz vectors* of A in $\mathcal{V}$. The *residuals* $\hat{r}_i = A u_i - u_i \mu_i$ for the respective Ritz pairs (μ_i, u_i) satisfy

$$(5) \qquad A u_i - u_i \mu_i = \hat{r}_i \perp \mathcal{V}.$$

Each *Ritz pair* (μ_i, u_i) is also an eigenpair of the $n \times n$ rank-k matrix $V M V^*$ and is interpreted as an approximation of an eigenpair of A. See [14], [19], [22].

2.2. Rayleigh-Ritz and subspace expansion. The *Rayleigh-Ritz* procedure is the first stage in iterative methods for eigenproblems and consists of computing the Ritz pairs from $\mathcal{V}$. The computation of S in (3) is not needed, nor feasible for reasons of efficiency. However, a cheaply available by-product of the computation of AV is the matrix $\widehat{R} = AV - VM$, where $R = V_\perp^* \widehat{R}$. Its columns are the respective residuals (5). In the second stage, the search subspace $\mathcal{V}$ is expanded. Different definitions of the expansion vector distinguish the different iterative methods. Each strategy results in a sequence of nested spaces

$$(6) \qquad \mathcal{V}_1 \subset \mathcal{V}_2 \subset \ldots \subset \mathcal{V}_{n-1} \subset \mathcal{V}_n$$

and has the objective to obtain accurate Ritz pairs while spending only minimal computational effort. One of the strategies will result in SPAM. Other strategies lead to methods with which SPAM will be compared in this paper: the Lanczos [15], Jacobi-Davidson [23], and Riccati [4] methods.

2.3. The Lanczos method. The Lanczos method [15] defines $\mathcal{V}_{j+1}$ as $\mathcal{V}_j \oplus$ span$\{r\}$ where $r \perp \mathcal{V}_j$ is any of the current residuals from (5). This results in a well-defined method: starting with some initial vector v_1 with $\|v_1\| = 1$ that spans the one-dimensional search space $\mathcal{V}_1$, it can be easily verified by induction that regardless

of which residual is used for expansion, the sequence of search spaces that is defined, equals the sequence of Krylov subspaces

$$(7) \qquad\qquad \mathcal{V}_j = \mathcal{K}^j(A, v_1) = \mathrm{span}\{v_1, Av_1, \ldots, A^{j-1}v_1\}.$$

Due to (5), the matrix V_k with the property that its column span equals $\mathcal{K}^k(A, v_1)$ can be chosen to have as columns v_1 and the normalized residuals

$$(8) \qquad\qquad v_{j+1} = \frac{\hat{r}_j}{\|\hat{r}_j\|}, \quad j \in \{1, \ldots, k\},$$

where $\hat{r}_j$ is a residual from the j-dimensional search space. From this the so-called *Lanczos relation* results,

$$(9) \qquad\qquad AV_k = V_{k+1}M_{k,k+1}, \quad \text{with } M_{k,k+1} = \begin{bmatrix} M_k \\ \varrho_k e_k^* \end{bmatrix}.$$

Here, M_k is a $k \times k$ tridiagonal matrix, e_k is the kth canonical basis vector of $\mathbb{C}^k$ and ϱ_k is a scalar. The following trivial observation will have its counterpart in the discussion of the SPAM method in Section 2.4.

Remark 2.1. If the Lanczos method runs for the full number of n iterations, it produces a unitary matrix V_n that depends only on A and the start vector v_1. The kth leading principal submatrix of the $n \times n$ tridiagonal matrix $M = V_n^* AV_n$ is the matrix M_k from (9) whose eigenvalues are the Ritz values after k iteration steps.

2.4. The Subspace Projected Approximate Matrix (SPAM) method. In the Subspace Projected Approximate Matrix (SPAM) method [21], the expansion vector is a suitable eigenvector of an approximation of A. This approximation has a cheaper action than A itself. Thus, the matrix-vector products within the inner iteration that is needed to compute this eigenvector, will be cheaper than for instance in the Jacobi-Davidson [23] and Riccati [4] methods. These methods, which we will explain in more detail in Section 2.5, both use A itself in their inner iteration.

A central observation in [21] is that the action of A on $\mathcal{V}$ that has already been performed in the outer iteration can be stored in a matrix $W = AV$, and be re-used within the inner iteration at relatively low costs. Thus, the action of A in the inner iteration only needs to be approximated partially. The resulting approximation is then different after each outer iteration step, even though only one approximate matrix A_0 is provided. Its action is merely used on less and less of the total space. As such, SPAM may be interpreted as a discrete homotopy method.

2.4.1. General description of SPAM. Let A_0 be an approximation of A. In Section 3 we will comment on how this approximation can be chosen. For now, assume that A_0 is available and define, in view of (3),

$$(10) \qquad \qquad \widehat{S} = V_\perp A_0 V_\perp^*$$

and

$$(11) \qquad A_k = (V|V_\perp)\widehat{A}_k(V|V_\perp)^*, \quad \text{where} \quad \widehat{A}_k = \begin{bmatrix} M & R^* \\ R & \widehat{S} \end{bmatrix}.$$

The subscript k of A_k refers to the number of columns of V. In [21], the matrix A_k is called a *subspace projected approximate matrix*. This is motivated by the fact that $A_k V = AV$ and $V^* A_k = V^* A$. In particular, since $M = V^* AV = V^* A_k V$, both A_k and A have the same Ritz pairs in $\mathcal{V}$. This will be exploited to derive bounds for the eigenvalues of A_k in Section 3. This is of interest since the search space $\mathcal{V}$ in the outer iteration is expanded with an eigenvector of A_k. Note that the action of A_k on $\mathcal{V}^\perp$ does *not* equal the action of A_0 on $\mathcal{V}^\perp$. Since $A^* = A$, the action of A_k on $\mathcal{V}^\perp$ equals, in fact, the action of A in k linearly independent functionals on $\mathbb{C}^n$.

With the convention that $\Pi_0 = I$, write

$$(12) \qquad \qquad \Pi_k = V_\perp V_\perp^* = I - VV^*.$$

This shows, together with equations (10) and (3), that

$$(13) \qquad A_k = -VV^*AVV^* + AVV^* + VV^*A + \Pi_k A_0 \Pi_k.$$

Thus, the action of A_k can benefit from the stored action of A on $\mathcal{V}$ as follows. With $W = AV$ we have that

$$(14) \qquad A_k v = -VMV^*v + WV^*v + VW^*v + \Pi_k A_0 \Pi_k v$$

where we have used (12) to avoid the numerically infeasible formation of $V_\perp$. Note that if $v \in \mathcal{V}^\perp$, the first two terms vanish.

R e m a r k 2.2. In [21] and [5], the projections Π_k are used to actually *define* A_k. Indeed, one can verify that

$$(15) \qquad \qquad A_k = A + \Pi_k(A_0 - A)\Pi_k.$$

The present set-up, however, is better suited to stress the relations with other methods.

426

In view of Remark 2.1 we now observe the following.

R e m a r k 2.3. If SPAM runs for the full n iterations, it produces a unitary matrix U_n that depends only on A and A_0. The kth leading principal submatrix of the $n \times n$ matrix $M = U_n^* A U_n$ then contains the Rayleigh-Ritz approximations after k steps of the outer iteration.

For theoretical purposes and without loss of generality, we assume that $V_\perp$ in (11) contains precisely the basis for the orthogonal complement of $\mathcal{V}$ that SPAM is about to produce in future iterations. With respect to this basis, $\widehat{A}_k$ is an update of $\widehat{A}_{k-1}$ of *arrowhead type*, in the sense that

$$(16) \qquad \widehat{A}_k - \widehat{A}_{k-1} = \left[\begin{array}{c|c|c} 0 & 0 & 0 \\ \hline 0^* & \tau & t^* \\ \hline 0^* & t & 0 \end{array} \right] = \begin{pmatrix} 0 \\ \tau \\ t \end{pmatrix} e_k^* + e_k (0^* \quad \tau \quad t^*) - e_k \tau e_k^*,$$

where each entry in the arrowhead formed by $t \in \mathbb{C}^{n-k}$, $\tau \in \mathbb{R}$ and t^*, is the difference between the corresponding entries of $\widehat{A}$ from (3) and $\widehat{A}_0 = (V|V_\perp)^* A_0 (V|V_\perp)$ from (11). Thus, with respect to the basis defined by the columns of $(V|V_\perp)$, the matrix $\widehat{A}_0$ simply transforms step by step into $\widehat{A}$ in the sense that after k steps, the first k columns and rows have changed into those of $\widehat{A}$, and the resulting matrix is called $\widehat{A}_k$. This is visualized in Figure 1.
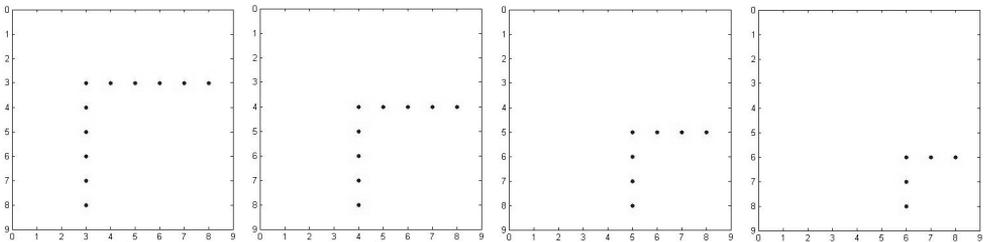

Figure 1. Arrowhead updates from $\widehat{A}_{k-1}$ to $\widehat{A}_k$ for consecutive values of k.

On the original basis, this transformation is described in the next proposition. Note that in this proposition, v is the eigenvector of interest of A_{k-1}, orthonormalized to V_{k-1}.

Proposition 2.4. *Let $k \geqslant 1$. Write $(V_{k-1}|v|V_\perp) = (V|V_\perp)$, thus $v = Ve_k$. Then A_k is the indefinite Hermitian rank-2 update of A_{k-1},*

$$(17) \qquad A_k = A_{k-1} + uv^* + vu^* = A_{k-1} + (u|v) \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} (u|v)^*,$$

where

$$(18) \qquad u = \left(\Pi_{k-1} - \frac{1}{2}vv^*\right)(A - A_0)v = \left(\Pi_k + \frac{1}{2}vv^*\right)(A - A_0)v.$$

P r o o f.   Combining (16) with $(V|V_\perp)(\widehat{A} - \widehat{A}_0)(V|V_\perp)^* = A - A_0$ we find

$$(19) \qquad \begin{pmatrix} 0 \\ \tau \\ t \end{pmatrix} = (0 \quad e_k \quad \dots \quad e_n)^*(\widehat{A} - \widehat{A}_0)e_k = (0|v|V_\perp)^*(A - A_0)v.$$

Therefore, substituting (19) into

$$A_k - A_{k-1} = (V_{k-1}|v|V_\perp)\left(\begin{pmatrix} 0 \\ \tau \\ t \end{pmatrix} e_k^* + e_k(0^* \quad \tau \quad t^*) - e_k \tau e_k^*\right)(V_{k-1}|v|V_\perp)^*$$

we arrive, using that $v = (V_{k-1}|v|V_\perp)$, at

$$(20) \qquad A_k - A_{k-1} = \Pi_{k-1}(A - A_0)vv^* + vv^*(A - A_0)\Pi_{k-1} - vv^*(A - A_0)vv^*.$$

Splitting the most right term in two equal parts and rearranging the terms proves the formula for u in (18). Since by (12)

$$(21) \qquad\qquad\qquad \Pi_{k-1} - vv^* = \Pi_k,$$

also the second equality is proved. □

R e m a r k 2.5. Note that (20) follows more easily from the alternative definition given in Remark 2.2. However, the proof of Proposition 2.4 also shows the validity of (16).

From (3) and (11) we see that

$$(22) \qquad\qquad \text{rank}(A - A_k) = \text{rank}(\widehat{A} - \widehat{A}_k) \leqslant n - k,$$

and thus, even though $A_k - A_{k-1}$ has rank-2, the update—or maybe better downdate

$$(23) \qquad\qquad A - A_k = A - A_{k-1} + (A_{k-1} - A_k)$$

will generically decrease the rank of $A - A_{k-1}$ by at most one. This remark goes hand in hand with the observation that even though the approximations A_k of A may seem unusual, the viewpoint of considering the reverse sequence $0 = A - A_n$, $A - A_{n-1}, \ldots, A - A_1, A - A_0$ as increasingly better approximations of $A - A_0$ is very natural indeed: they form a sequence of Rayleigh-Ritz approximations to $A - A_0$ in the orthogonal complements of the spaces V_k.

In the outer iteration, the products Av and V^*Av were computed. Thus, both v^*Av and $\Pi_{k-1}Av$ are available in (18) without additional computational costs. Furthermore, since v is orthogonal to the $(k-1)$-dimensional search space, we have that $v^*A_{k-1}v = v^*A_0v$. Now, because v is the result of orthogonalization of an eigenvector of A_{k-1} to V_{k-1}, also v^*A_0v can be retrieved from the inner iteration. Thus, the vectors u and v in the updating procedure (17) are cheaply available.

R e m a r k 2.6. Of course, the update (17) itself should not be performed explicitly because it will generally result in fill-in of originally sparse matrices.

2.4.2. Choice of the method for the inner iteration of SPAM. In [21], the authors suggest to use Davidson's method [9] to solve the eigenvalue problem for A_k, but of course any other method can be adopted. Apart from the Lanczos method, also the Generalized Davidson method from [17] was tested as inner method in [21]. Other possibilities include, for instance, the Jacobi-Davidson [23] method[1]. The latter can be a good option because it often needs only a few iterations to converge if a good start vector is available. This start vector may be either the eigenvector approximation of A_{k-1}, or the current eigenvector approximation of A from the outer iteration. We will study this choice in Section 2.5.

R e m a r k 2.7. SPAM should first of all perform well under the assumption that the eigenproblem for A_k is solved exactly. This will be investigated in the numerical illustrations in Section 4.

In [21] it is also noted that SPAM itself can be chosen in the inner iteration. This leads to a recursive multilevel version of the method, and assumes that a whole sequence of approximating matrices of A is available, each having a cheaper action than its predecessor. The eigenvector computed at a given level is then used as expansion vector at the first higher level.

2.5. Comparing SPAM with the Jacobi-Davidson and the Riccati method. The philosophy of SPAM is to use all available (computed) information from (3), i.e., M, R and R^*, to determine the expansion vector, and consequently, that only S needs to be approximated. The Jacobi-Davidson [23] and Riccati [4] methods partially share this philosophy. Instead of R, they only use the residual corresponding to the selected eigenvector approximation. On the other hand, in their simplest forms, they do not approximate the matrix S but use its full action. In this section we will outline their similarities and differences.

---

[1] M. Hochstenbach presented this option at the 2006 GAMM/SIAM Applied Linear Algebra conference.

Remark 2.8. Since in the Lanczos method all residuals are linearly dependent, also the Lanczos method uses all information about the residual. However, Lanczos has no inner iteration. This will make sense from the point of view taken in Section 3. There we will show that for the choice $A_0 = 0$, also SPAM needs no inner iteration, and that this choice makes SPAM mathematically equivalent to Lanczos.

2.5.1. The Riccati and the Jacobi-Davidson methods. Given a Ritz pair (μ, u) with residual $\hat{r} \perp \mathcal{V}$, generally each eigenvector of A has a multiple that equals $u + t$, with $t \perp u$ a so-called *orthogonal correction* to u. Indeed, let X be such that $(u|X)$ is unitary. Then with $S = X^* A X$ and $\hat{r} = Xr$,

$$(24) \qquad A(u|X) = (u|X) \begin{bmatrix} \mu & r^* \\ r & S \end{bmatrix},$$

and the orthogonal correction t equals Xp where p can be verified to satisfy the *generalized algebraic Riccati equation*

$$(25) \qquad (S - \mu I)p = -r + pr^* p.$$

Transforming (25) back to the original basis shows that t solves

$$(26) \qquad t \perp u \quad \text{and} \quad (I - uu^*)(A - \mu I)(I - uu^*)t = -\hat{r} + t\hat{r}^* t.$$

In [4], solutions of (26) were approximated by means of Rayleigh-Ritz projection in a l-dimensional subspace $\mathcal{U}$ of $u^\perp$ and a suitable one was selected as expansion vector for $\mathcal{V}$. This idea was intended as an enhancement of the Jacobi-Davidson method [23]. This method neglects the quadratic term $t\hat{r}^* t$ from (26) and uses instead the unique solution $\hat{t}$ of

$$(27) \qquad \hat{t} \perp u \quad \text{and} \quad (I - uu^*)(A - \mu I)(I - uu^*)\hat{t} = -\hat{r}$$

to expand $\mathcal{V}$. Jacobi-Davidson is simpler than the Riccati method in the sense that only a linear system for $\hat{t}$ needs to be solved. It can be interpreted as an *accelerated Newton method* [24]. However, much more than the Riccati method, Jacobi-Davidson suffers from stagnation in case the term $t\hat{r}^* t$ from (26) is not small. On the other hand, if $t\hat{r}^* t$ is small enough, Jacobi-Davidson converges quadratically, as one would expect from a Newton type method. This shows that one should be careful in proposing alternatives to the correction equation (27). For instance, in [10], the authors investigated the effect of solving the alternative correction equation

$$(28) \qquad \tilde{t} \perp V \quad \text{and} \quad (I - VV^*)(A - \mu I)(I - VV^*)\tilde{t} = -\hat{r}.$$

At first sight this seems to make sense, because it directly looks for a correction orthogonal to $\mathcal{V}$. Also, the conditioning of the linear equation (28) may be better than (27) in case the search space contains good approximations of eigenvectors belonging to eigenvalues close to μ. However, orthogonalizing $\hat{t}$ from (27) to $\mathcal{V}$ generally does not result in $\tilde{t}$ from (28) and the price to pay is that $\|t - \tilde{t}\|$ is not of higher order, as is $\|t - \hat{t}\|$. Indeed, in [10] it is shown explicitly that, apart from some exceptional cases, the quadratic convergence of Jacobi-Davidson is lost, whereas in those exceptional cases, both expansions are equivalent. Numerical experiments in [10] confirm the above observations.

R e m a r k 2.9. The method using the correction equation (28) may at first sight also resemble SPAM. However, as we will see in the section to come, the use of V in SPAM is of a different nature than in (28), where the correction is sought orthogonal to $\mathcal{V}$. In SPAM, it is sought in the whole space and only orthogonalized to $\mathcal{V}$ afterwards.

2.5.2. Preconditioning in the Jacobi-Davidson method. In the original paper [23] on the Jacobi-Davidson method, reprinted as [25], preconditioning is discussed as follows. Suppose that an approximation A_0 of A is available. It is shown in [23] how to apply such a preconditioner to (27), which is a linear equation, though not with system matrix A. To be explicit, since $(I - uu^*)\hat{t} = \hat{t}$, we have that

$$(29) \qquad (A - \mu I)\hat{t} = -\varepsilon u - \hat{r},$$

where ε is such that $\hat{t} \perp u$. Or equivalently, written as an augmented system,

$$(30) \qquad \begin{bmatrix} A - \mu I & u \\ u^* & 0 \end{bmatrix} \begin{bmatrix} \hat{t} \\ \varepsilon \end{bmatrix} = \begin{bmatrix} -\hat{r} \\ 0 \end{bmatrix}.$$

Thus, an approximation $\hat{t}_0$ of $\hat{t}$, together with an approximation ε_0 of ε can be obtained simply by replacing A by A_0 in (30). The pair $\hat{t}_0, \varepsilon_0$ can be computed as

$$(31) \qquad \hat{t}_0 = -\varepsilon_0(A_0 - \mu I)^{-1}u - (A_0 - \mu I)^{-1}\hat{r}, \quad \text{with } \varepsilon_0 = -\frac{u^*(A_0 - \mu I)^{-1}\hat{r}}{u^*(A_0 - \mu I)^{-1}u}.$$

This approximation is called a *one step approximation* in [23]. It was observed that setting $\varepsilon_0 = 0$ in (31), the Davidson method [9] results. With $A_0 = A$, which corresponds to Jacobi-Davidson with full accuracy solution of the correction equation, (31) becomes

$$(32) \qquad \hat{t} = -\varepsilon(A - \mu I)^{-1}u - u,$$

and since $\hat{t}$ is then orthogonalized to u, the method is mathematically equivalent to an accelerated shift and invert iteration that works with $(A - \mu I)^{-1}u$. It is argued, and demonstrated by experiments in [23], that Jacobi-Davidson combines the best of those two methods. Of course, a natural next stage in preconditioning is to use the matrix

$$
(33) \qquad A_0^u = \begin{bmatrix} A_0 - \mu I & u \\ u^* & 0 \end{bmatrix}
$$

as preconditioner *within the iterative method* that aims at solving (30). In each step of such a method one would need to solve a system with A_0^u. This can be done by solving two systems as in (30)–(31) in the first step of the inner iteration. In each consecutive step, only one system of the form $(A_0 - \mu I)z = y$ would need to be solved.

2.5.3. One step approximation of the SPAM eigenproblem for A_k. In the SPAM method, the expansion vector for the Ritz Galerkin subspace in the outer iteration is a relevant eigenvector v_k of A_k. In principle, any eigenvalue method can be used to compute an approximation for v_k, but observe that the starting point is as follows. In the outer iteration, we have just solved a $k \times k$ eigenproblem for $M = V^*AV$, and a Ritz pair (μ, u) with $\|u\| = 1$ and with residual $\hat{r} = Au - \mu u$ has been selected. The matrix A_k is now available, either explicitly or implicitly. Since $A_k V = AV$ and $V^*A_k = V^*A$, the Ritz pair (μ, u) for A with respect to the current search space $\mathcal{V}_k$ is also a Ritz pair for A_k. Thus we can exploit the fact that $\mathcal{V}_k$ contains, by definition, good approximations of the relevant eigenvectors of A_j with $j < k$, and use it as initial search space for a Ritz Galerkin method applied to A_k to approximate v_k. Since $\mathcal{V}_k$ is generally not a Krylov subspace, the Lanczos method is not a feasible candidate. The Jacobi-Davidson method is. The correction equation for the first step of the Jacobi-Davidson method in the inner iteration can be set up without any additional computations:

$$
(34) \qquad \begin{bmatrix} A_k - \mu I & u \\ u^* & 0 \end{bmatrix} \begin{bmatrix} t_k \\ \varepsilon_k \end{bmatrix} = \begin{bmatrix} -\hat{r} \\ 0 \end{bmatrix}.
$$

Since quadratic convergence *in the outer iteration* cannot be expected even if an exact eigenvector of A_k were computed, we study the effect of applying only one iteration of Jacobi-Davidson in the inner iteration. This is also motivated by the fact that the initial search space $\mathcal{V}_k$ for Jacobi-Davidson applied to A_k may be relatively good and may result in quadratic convergence *in the inner iteration*.

R e m a r k 2.10. If only one step of Jacobi-Davidson is applied, then after solving t_k from (34), the new approximation v for the eigenvector v_k of A_k would lie in the

space $\mathcal{V}_k \oplus \langle t_k \rangle$. It would not be necessary to actually compute this approximation, because

(35)
$$\mathcal{V}_k \oplus \langle v \rangle \subset \mathcal{V}_k \oplus \langle t_k \rangle.$$

Thus, instead of computing the eigendata of a $(k+1) \times (k+1)$ matrix, the expansion of the Ritz-Galerkin space in the outer iteration can be done immediately with t_k.

SPAM in which the eigenproblem for A_k is approximated with one iteration of Jacobi-Davidson, we will refer to as *one step SPAM*, abbreviated by SPAM(1). We will talk about *Full SPAM* if the eigenproblem for A_k is solved to full precision.

2.5.4. Comparing one step Jacobi-Davidson with SPAM(1). SPAM(1) can best be compared with preconditioned Jacobi-Davidson with *one step approximation*, as described in Section 2.5.2. The only difference between the two methods is that in iteration k of SPAM(1) the preconditioner A_k is used, whereas in one-step Jacobi-Davidson this is A_0. As such, SPAM(1) can be seen as an attempt to enhance this type of preconditioned Jacobi-Davidson. We will now investigate the effect of this attempt.

Lemma 2.11. *Assume that $\mathcal{V}_{k+1} = \mathcal{V}_k \oplus \langle v \rangle$ with $v \perp \mathcal{V}_k$ and $\|v\| = 1$. Then*

(36)
$$A_{k+1} = -vv^* A vv^* + A vv^* + vv^* A + (I - vv^*) A_k (I - vv^*).$$

P r o o f. By substitution of the defining relation for A_k. □

Corollary 2.12. *Let $\mathcal{V}_1$ be the span of the relevant eigenvector u of A_0 and let $\mu = u^* A u$ and $\hat{r} = Au - \mu u$. Then the solution t_1 from the system*

(37)
$$\begin{bmatrix} A_1 - \mu I & u \\ u^* & 0 \end{bmatrix} \begin{bmatrix} t_1 \\ \varepsilon_1 \end{bmatrix} = \begin{bmatrix} -\hat{r} \\ 0 \end{bmatrix}$$

in the first iteration of SPAM(1), to be used to expand $\mathcal{V}_1$, coincides with the solution t_0 from the system

(38)
$$\begin{bmatrix} A_0 - \mu I & u \\ u^* & 0 \end{bmatrix} \begin{bmatrix} t_0 \\ \varepsilon_0 \end{bmatrix} = \begin{bmatrix} -\hat{r} \\ 0 \end{bmatrix},$$

solved in the first iteration of Jacobi-Davidson with one-step approximation using A_0.

P r o o f. The linear system (34) for SPAM(1) with $k = 1$ is equivalent to

(39)
$$t_1 \perp u, \quad (I - uu^*)(A_1 - \mu I)(I - uu^*) t_1 = -\hat{r},$$

where u is a unit vector spanning $\mathcal{V}_1$. Substituting the expression (36) for A_1 immediately proves the statement. □

Thus, there is no difference in the first iteration. There is, however, a difference in further iterations. To study this difference we take a different viewpoint. Above, we approximated the SPAM inner eigenvalue problem by a linear correction equation which made it suitable for comparison with one step Jacobi-Davidson. The opposite viewpoint is also possible, which is to interpret the Jacobi-Davidson correction equation with one-step approximation as an *exact* correction equation of a *perturbed* eigenproblem.

Lemma 2.13. *Given u with $\|u\| = 1$ and $\mu = u^*Au$ and $r = Au - \mu u$. Define for a given approximation A_0 of A the matrix*

$$(40) \qquad A_u := -uu^*Auu^* + uu^*A + Auu^* + (I - uu^*)A_0(I - uu^*)$$
$$= u\mu u^* + ur^* + ru^* + (I - uu^*)A_0(I - uu^*).$$

Then (u, μ) is a Ritz pair of A_u in the one-dimensional span of u with residual $r = A_u u - \mu u$ and with the equation

$$(41) \qquad (I - uu^*)(A_0 - \mu I)(I - uu^*)t = -r$$

as its exact (i.e., without preconditioning) Jacobi-Davidson correction equation.

P r o o f.   It is easily verified that $u^*A_u u = \mu$ and $A_u u - \mu u = r$ and thus (μ, u) is a Ritz pair for A_u with residual r. Since, moreover,

$$(42) \qquad (I - uu^*)A_u(I - uu^*) = (I - uu^*)A_0(I - uu^*),$$

its correction equation is precisely (41), or equivalently, (31).   □

Note that A_u from (40) is the subspace projected approximate matrix for the one dimensional subspace spanned by u. Now, with

$$(43) \qquad \mathcal{V}_k = \mathcal{U} \oplus \langle u \rangle,$$

where $\mathcal{U}$ is the orthogonal complement of the span of the relevant Ritz vector u in the current search space $\mathcal{V}_k$, we have, similarly to (36) and (40), that

$$(44) \qquad A_k = u\mu u^* + ur^* + ru^* + (I - uu^*)A_{\mathcal{U}}(I - uu^*).$$

Here, $A_{\mathcal{U}}$ is the subspace projected approximated matrix corresponding to the subspace $\mathcal{U}$. Now, the opposite viewpoint mentioned above is the observation that in one step Jacobi-Davidson, the expansion vector is (an arguably good approximation of) the relevant eigenvector of A_u in (40), whereas in SPAM, it is (an arguably good approximation of) the relevant eigenvector of A_k in (44). Both matrices have now an appearance that is suitable for studying their differences and similarities.

434

The most important observation is that neither correction will lead to the unique correction that results in quadratic convergence in the outer iteration. Second, since both matrices A_u and A_k differ only in their restriction to the orthogonal complement $u^\perp$ of u, the difference in the methods they represent will be marginal if the residual is already small. Since, as already mentioned in Corollary 2.12, not only at the start but also in the first iteration both methods coincide, the difference from the second iteration onwards will probably be very small, especially if A_0 provides a good initial approximation of the relevant eigenvector. Finally, even though A_u uses less information of A than A_k, it does use the *optimal* information in some sense. It may even be a disadvantage to use more information, because this involves approximations of eigenvectors orthogonal to the eigenvector of interest. The above observations will be tested, and confirmed by our numerical illustrations in Section 4.

3. Selecting the matrix A_0 in the SPAM method

Here we study some of the effects of the choice for A_0 on the iterative approximation process in SPAM. We will assume for simplicity that the largest eigenvalue of A is the target, although replacing A by $-A$, we might as well have set the smallest eigenvalue of A as a target.

3.1. Some simple choices for A_0. It is instructive to study the consequences of the choice $A_0 = 0$. This generic parameter-free choice may seem dubious at first sight, but it is not. First note that since the start vector of SPAM is the relevant eigenvector of $A_0 = 0$, this is necessarily a random vector $v \in \mathbb{C}^n$, with $\|v\| = 1$. Thus, we set $\mathcal{V}_1 = \mathrm{span}\{v\}$. Write $\mu = v^* A v$ and $\hat{r} = Av - v\mu$. Then, with $V_\perp$ such that $(v|V_\perp)$ is orthogonal and $r = V_\perp^* \hat{r}$, we find that

$$(45) \qquad A = (v|V_\perp) \begin{bmatrix} \mu & r^* \\ r & S \end{bmatrix} (v|V_\perp)^*,$$

and consequently, replacing S by the zero matrix, the next approximating matrix A_1 from (11) is defined by

$$(46) \qquad A_1 = (v|V_\perp) \begin{bmatrix} \mu & r^* \\ r & 0 \end{bmatrix} (v|V_\perp)^*.$$

As shown already in Proposition 2.4, A_1 is a simple rank-two matrix that on the basis defined by the columns of $(v|V_\perp)$ is of arrow-head type. It has two nontrivial eigenpairs $A_1 w_\pm = \theta_\pm w_\pm$, where

$$(47) \qquad w_\pm = \theta_\pm v + r \quad \text{and} \quad \theta_\pm = \frac{1}{2}\mu \pm \sqrt{\frac{1}{4}\mu^2 + \|r\|^2}.$$

Assuming that A is positive definite will lead to the selection of w_+ for the expansion of the search space. Since w_+ is a linear combination of v and r, we find that

$$(48) \qquad \mathcal{V}_2 = \text{span}\{v, r\} = \mathcal{K}^2(A, v),$$

and the two eigenvalue approximations computed in the outer loop of the SPAM method are the same as in the Lanczos method. This is, of course, not a coincidence.

Theorem 3.1. *If the goal is to find the largest eigenvalue of a positive definite matrix A, the SPAM method with $A_0 = 0$ is mathematically equivalent to the Lanczos method.*

P r o o f. Let $A_0 = 0$. Then the eigenvalues of A_k in (11) are those of the $n \times n$ Hermitian arrowhead

$$(49) \qquad \begin{bmatrix} M & R^* \\ R & 0 \end{bmatrix}.$$

The Cauchy Interlace Theorem immediately gives that each of the largest k of them is larger than or equal to the corresponding eigenvalue of M. This ensures that the eigenvector of A_k that is selected for expansion is not from its null space but from its column span. From (13) we see that if $A_0 = 0$ this column span is the span of V and AV. A simple induction argument shows that this span equals $\mathcal{K}^{k+1}(A, v)$. $\square$

R e m a r k 3.2. If A is indefinite and we would like to find the eigenvalue closest to zero, the choice $A_0 = 0$ would lead to expansion vectors from the null space of A_0, and the method would be worthless. As a solution, A may be shifted by a suitable multiple α times the identity I to make it positive semi-definite. Equivalently, instead of using $A_0 = 0$ we may choose $A_0 = \alpha I$ as approximating matrix. In both cases, it is easy to verify that SPAM will still be equal to Lanczos.

The observant reader may have noticed that a peculiar situation has arisen. In the inner loop of SPAM, a better approximation of the largest eigenvalue of A was computed than the Ritz values from the outer loop. In view of the philosophy of inner-outer iterations, this in itself is not out of the ordinary, but its computation did not require any additional matrix-vector multiplication with A, nor with an elaborate approximation A_0 of A. The following proposition, which uses the notation (1) and (4), makes this explicit.

436

Proposition 3.3. *With θ_+ as in (47) we have that $\mu \leqslant \|Av\| \leqslant \theta_+$. Moreover, if A is positive semi-definite, we find that $\theta_+ \leqslant \lambda_+$, where λ_+ is the largest eigenvalue of A.*

P r o o f.   Because $r \perp v$ and $Av - \mu v = r$, by Pythagoras' Theorem we have

$$(50) \qquad \mu^2 \|v\|^2 + \|r\|^2 = \mu^2 + \|r\|^2 = \|Av\|^2,$$

hence $\mu \leqslant \|Av\|$. Squaring θ_+ from (47) gives

$$(51) \qquad \theta_+^2 = \|r\|^2 + \frac{1}{2}\mu^2 + \mu\sqrt{\frac{1}{4}\mu^2 + \|r\|^2},$$

which together with (50) shows that $\|Av\| \leqslant \theta_+$. Since S in (45) is positive definite whenever A is, combining (45) and (46) with Weyl's bound yields $\theta_+ \leqslant \lambda_+$. □

The key issue here is that the inner eigenvalue approximations are much related to the so-called *harmonic Ritz values* [3], [22] of A. Indeed, assuming that A itself is positive semi-definite, these are the k positive eigenvalues of the at most rank-$2k$ matrix

$$(52) \qquad \widetilde{A}_k = \begin{bmatrix} M & R^* \\ R & T \end{bmatrix}, \quad \text{where } T = RM^{-1}R^*.$$

They can be computed without additional matrix vector multiplications with A. Note that harmonic Ritz values are usually introduced as the reciprocals of the Rayleigh-Ritz approximations of A^{-1} in the space $A\mathcal{V}$. It is well known that for positive semi-definite matrices A, the harmonic Ritz values are better approximations of the larger eigenvalues of A than the standard Ritz values. We provide the short argument in Lemma 3.5. See also [3].

Proposition 3.4. *The matrix $\widetilde{A}_k$ can be decomposed as*

$$(53) \qquad \begin{bmatrix} M & R^* \\ R & T \end{bmatrix} = \begin{bmatrix} M \\ R \end{bmatrix} M^{-1} \begin{bmatrix} M & R^* \end{bmatrix}.$$

The blocks

$$(54) \qquad \begin{bmatrix} M \\ R \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} -M^{-1}R^* \\ I \end{bmatrix}$$

span the range and null space, respectively, and the nonzero eigenvalues are the eigenvalues of the $k \times k$ matrix

$$(55) \qquad UM^{-1}U^* \quad \text{where } \begin{bmatrix} M \\ R \end{bmatrix} = QU$$

is a QR-decomposition. In particular, those eigenvalues are positive.

P r o o f.   The statements are all easy to verify. The positivity of the k nonzero eigenvalues follows from Sylvester's Theorem of Inertia. □

The k eigenpairs (θ_j, w_j) of $\widetilde{A}_k$ in (52) with positive eigenvalues we label as

$$(56) \qquad\qquad 0 < \tilde{\theta}_k \leqslant \tilde{\theta}_{k-1} \leqslant \ldots \leqslant \tilde{\theta}_2 \leqslant \tilde{\theta}_1.$$

The proposition shows that they can be computed by solving a $k \times k$ eigenproblem and that no additional matrix-vector products with A are needed. We can now easily prove the following bounds. See also [3], [19].

Lemma 3.5. *For all $j \in \{1, \ldots, k\}$, we have that*

$$(57) \qquad\qquad \mu_j \leqslant \tilde{\theta}_j \leqslant \lambda_j.$$

P r o o f.   The left inequalities follow from the Cauchy Interlace Theorem applied to $\widetilde{A}_k$. Now, with $\widehat{A}$ as in (3), recognizing the Schur complement $\widehat{A}/M$ shows that

$$(58) \qquad\qquad \widehat{A} - \widetilde{A}_k = \begin{bmatrix} 0 & 0^* \\ 0 & \widehat{A}/M \end{bmatrix}$$

is positive semi-definite, hence the right inequalities follow from Weyl's bound. □

Observe that, as was the case with the choice $A_0 = 0$, assuming that from the start of the SPAM method the eigenvector w_1 belonging to θ_1 was selected for expansion, the eigenvectors w_j, called the *harmonic Ritz vectors*, lie in the column span of AV and hence, in $\mathcal{K}^{k+1}(A, v)$.

Thus, even though we may have improved the inner loop eigenvalue approximations, the SPAM method is still equal to the Lanczos method. It does give, however, valuable insight into SPAM: Lanczos results from the choice $A_0 = 0$, even after modification of A_k into the positive semi-definite matrix $\widetilde{A}_k$. In order to get a method different from Lanczos, we should use less trivial approximations that are based on the structure and properties of A itself, while aiming at retaining similar inequalities as in Lemma 3.5. For this, we would need the matrices $A - A_k$ to be positive semi-definite. We will investigate this in the following sections.

3.2. One-sided approximations. Having seen that the trivial choice $A_0 = 0$, even after a correction that turns all approximate matrices positive semi-definite, will generally lead to the Lanczos method, we now turn our attention to *approximations from below*, by which we mean A_0 such that $A - A_0$ is positive semi-definite.

Lemma 3.6. *If A_0 approximates A from below, then so does each matrix A_k.*

P r o o f.   Combining (3) and (10) with (11) we see that for all $x \in \mathbb{C}^n$,

$$(59) \qquad x^*(\widehat{A} - \widehat{A}_k)x = x^* \begin{bmatrix} 0 & 0^* \\ 0 & S - \widetilde{S} \end{bmatrix} x = x^* V_\perp^* (A - A_0) V_\perp x \geqslant 0$$

where the last inequality holds because $A - A_0$ is positive semi-definite. And thus, $\widehat{A} - \widehat{A}_k$ is positive semi-definite, and hence, so is $A - A_k$. $\qquad\square$

By Proposition 2.4, $A_k - A_{k-1}$ is an indefinite rank-2 matrix, hence it will generally not be true that A_{k-1} approximates A_k from below.

Lemma 3.7. *The following inequalities are valid generally:*

$$(60) \qquad\qquad \theta_{j+1} \leqslant \mu_j \leqslant \theta_j \quad \text{for all } j \in \{1, \dots, k\},$$

whereas, if A_0 approximates A from below, additionally

$$(61) \qquad\qquad \theta_j \leqslant \lambda_j \quad \text{for all } j \in \{1, \dots, n\}.$$

P r o o f.   The first set of inequalities applies because A_k has the same Ritz values as A. See also Section 2.4.1. It is well known that the Ritz values interlace the exact eigenvalues [19]. Since $A - A_k$ is positive semi-definite due to Lemma 3.6, the equality $A_k + (A - A_k) = A$ together with Weyl's bound [19] proves the second set of inequalities. $\qquad\square$

Lemma 3.7 shows that if A_0 approximates A from below, the approximations for the larger eigenvalues of A that are produced in the inner iteration will never be worse than the ones obtained in the outer iteration. Moreover, they will never be larger than the corresponding exact eigenvalues. Thus, it indeed makes sense to expand the search space with the eigenvector that is computed in the inner iteration. Question that remains is how to obtain matrices A_0 that approximate A from below.

3.2.1. Algebraic construction of approximations from below. Clearly, for any positive definite matrix H we have that $A_0 = A - H$ approximates A from below, even though A_0 itself may not be positive definite. The problem is of course how to choose H such that A_0 is close to A in an appropriate sense, while its action is considerably less expensive than that of A.

If A itself is a positive definite matrix, a purely algebraic option is at hand. Given an index set $\mathcal{I} \subset \{1, \dots, n\}$ of cardinality m, let $E_{\mathcal{I}}$ be the matrix with the standard basis vectors $e_i, i \in \mathcal{I}$ as columns. Then set

$$(62) \qquad A_0 = A - H, \quad \text{where } H = E_{\mathcal{I}} E_{\mathcal{I}}^* A E_{\mathcal{I}} E_{\mathcal{I}}^*.$$

The matrix H is the result of a Rayleigh-Ritz procedure using as search space the column span of $E_{\mathcal{I}}$. For a randomly chosen index set, this space has no a priori relation with A and thus H is probably a relatively poor approximation of A in comparison with, for example, a Krylov subspace approximation.

R e m a r k 3.8. In this particular situation it is an advantage if H does not approximate A very well, because A_0 should be close to A, not H. Notice also that A_0 has zero entries at positions (i, j) for all $i, j \in \mathcal{I}$, and is thus always sparser than A. A priori knowledge of A may lead to a more sophisticated choice of the index set(s) $\mathcal{I}$. If the goal is to approximate the largest eigenvalues of A, the index set $\mathcal{I}$ could be chosen such that the smallest diagonal entries of A are selected to be put in H. Consequently, A_0 will share with A the largest diagonal entries, and this may increase its approximation quality. This is illustrated in Figure 2.
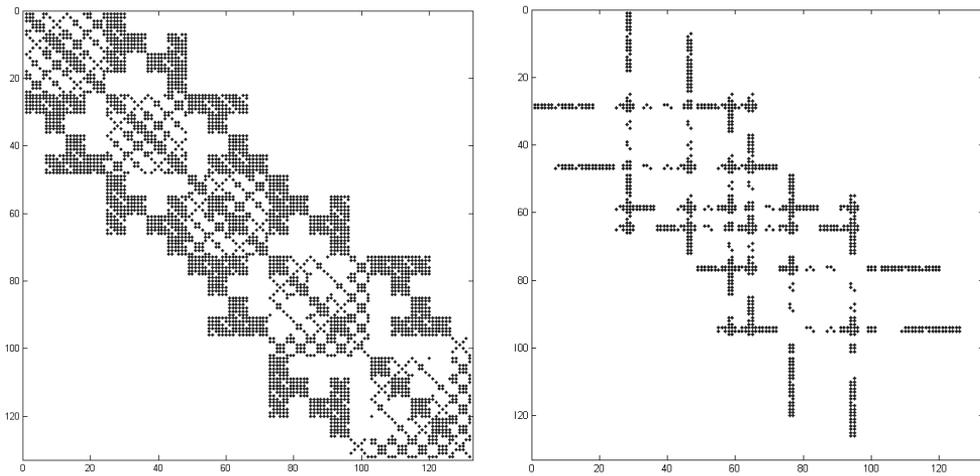

Figure 2. Approximating a matrix A from structural engineering from below by a sparser matrix A_0 by subtracting from A a definite matrix H; the sparsity plot of A (with 3648 nonzero entries) is on the left and that of A_0 (with 968 nonzero entries) on the right. See the experiments with this pair of matrices in Section 4.4.

R e m a r k 3.9. Notice that $\operatorname{rank}(A_0) \leqslant 2(n - m)$. Especially for large m this may greatly simplify the computation of the eigendata of A_0 and of A_k for small values of k.

440

R e m a r k 3.10. Although A itself is positive definite, A_0 generally is not. It can be made positive definite by adding a Schur complement at the position of the zero block, as was done in (52). Since the Schur complement block is the smallest correction to A_0 that makes it positive semi-definite, the result would still approximate A from below. However, its computation involves the evaluation of the inverse of an $(n - m) \times (n - m)$ matrix. It is not clear if the additional computational effort is well spent. Also without the Schur complement, A_0 has as many positive as negative eigenvalues and is likely to perform better than $A_0 = 0$.

The rather crude approach that we have just described is easy to implement, and it can be applied in the context of the *multilevel version* of SPAM mentioned in Section 2.4.2: the solution of the inner iteration is done using SPAM itself, with an approximation of A_0 that is based on a larger index set than the one that was used to construct A_0 itself.

3.2.2. Natural construction of approximations from below. A situation in which approximations from below are naturally available is the setting of discretized partial differential equations including an elliptic term, either by the finite difference method or the finite element method [7]. Then removing the positive definite discrete elliptic operator, either completely or partially, from the total discrete operator, results in an approximation from below. Indeed, let $\Omega \subset \mathbb{R}^3$ be a domain, and consider as an example the problem of finding eigenmodes for the linear operator L, defined by

$$(63) \qquad L(u) = \lambda u \quad \text{where} \quad L(u) = -\varepsilon \operatorname{div}(\mathcal{K}\nabla u) + cu, \quad \text{and} \quad u = 0 \quad \text{on } \partial\Omega.$$

Here $\varepsilon > 0$ is a parameter, $\mathcal{K}\colon \Omega \to \mathbb{M}^{3\times 3}(\mathbb{R})$ maps Ω into the symmetric positive definite matrices, and $c \in C(\overline{\Omega})$ is nonnegative. Discretizing this equation with the finite difference method will lead to an algebraic eigenvalue problem of the form $\varepsilon Kx + Mx = \xi x$. The matrix K that represents the discretized diffusion is positive definite. Although sparse, it will generally have more fill-in than the diagonal matrix M that represents the reaction term, and, if ε is small enough, have smaller entries. Thus, the total discretized operator $A = K + M$ has $A_0 = M$ as a candidate for the approximation matrix: its action is cheaper than the action of A and $A - A_0 = K$ is positive definite. A similar strategy can also be employed in the finite element method, when so-called *mass lumping* is used in the assembly of the mass matrices.

R e m a r k 3.11. In this context, the algebraic method can be used to approximate the smallest eigenvalue of A by applying it to $\alpha I - A$ with α such that $\alpha I - A$ is positive semi-definite. Even though the largest eigenvalue usually has no physical relevance, together they would provide good estimates for the condition number of the A, which is indeed of interest.

3.3. Cutting off the bandwidth. Here we describe an obvious choice of approximating matrices that was used in [21] to illustrate the effectiveness of their method. It concerns symmetric banded matrices. Apart from the approach that we will now describe, in the numerical illustrations of Section 4 we also intend to apply our algebraic approximations from below to these matrices.

Given $0 \leqslant \varepsilon \leqslant 1$ and $1 \leqslant q \leqslant n - 1$, define a matrix $A = (A_{ij})$ by

(64)
$$A_{ij} = \begin{cases} i & \text{if } i = j; \\ \varepsilon^q & \text{if } 1 \leqslant |i - j| \leqslant q; \\ 0 & \text{otherwise.} \end{cases}$$

In [21] it is proposed to choose as approximating matrix for A a matrix A_0 of the same type with a smaller half-bandwidth $\tilde{q} < q$. For instance,

$$A = \begin{bmatrix} 1 & \varepsilon & \varepsilon^2 & \\ \varepsilon & 2 & \varepsilon & \varepsilon^2 \\ \varepsilon^2 & \varepsilon & 3 & \varepsilon \\ & \varepsilon^2 & \varepsilon & 4 \end{bmatrix}, \quad \text{then } A_0 = \begin{bmatrix} 1 & \varepsilon & & \\ \varepsilon & 2 & \varepsilon & \\ & \varepsilon & 3 & \varepsilon \\ & & \varepsilon & 4 \end{bmatrix}.$$

For each eigenvalue θ of A_0 there is an eigenvalue λ of A with

(65)
$$|\lambda - \theta| \leqslant \frac{\varepsilon^{q_0+1} - \varepsilon^{q+1}}{1 - \varepsilon} \sqrt{n}.$$

Indeed, the difference $A - A_0$ is zero except for the bands $q_0 + 1$ to q. Each non-zero row contains at most the numbers ε^{q_0+1} to ε^q, and (65) follows from the Bauer-Fike [2] theorem and the finite geometric series sum formula. Thus, for values of ε small enough, A_0 may be a good candidate to approximate A even though it is generally not an approximation from below. Nevertheless, its eigenvalues are close to the eigenvalues of A and its action is cheaper than the action of A. The number of floating point operations required to compute Av is approximately twice as much as for an approximation A_0 having half the bandwidth of A. In other words, each two matrix-vector products A_0 are approximately equally costly as a single product with A.

Cutting off the bandwidth in this fashion makes sense especially if the decay of the size of the off-diagonal entries in relation to their distance to the main diagonal is quick enough. Apart from the example above, this is the case in many applications where the boundary element method [30] is used to approximate integral equations. Notice that for such applications also the approach from Section 3.2.1 can be applied, resulting both in sparse and low rank approximating matrices A_0.

In the previous sections, we have compared SPAM with both the Jacobi-Davidson method and the Lanczos method. We have also studied ways how to define an appropriate approximate matrix A_0 in the context of the SPAM method. Since the choice $A_0 = 0$ will effectively lead to the Lanczos method, our starting point in the upcoming numerical illustrations will be to consider SPAM as a boosted version of the Lanczos method. This is, of course, particularly justified if A_0 is an approximation of A from below. As discussed in Section 2.5, SPAM can also be considered an attempt to enhance the Jacobi-Davidson method with one-step approximation as preconditioning. Therefore, we will also present a comparison of SPAM and this version of Jacobi-Davidson. We end this section with a discussion of the numerical results.

4.1. Objectives. First, we list the methods and abbreviations that we use to describe them.

▷ Lanczos (see Section 2.3);

▷ JD(l): Jacobi-Davidson using l steps of MinRES to approximate the solution of the exact correction equation (30) in augmented form;

▷ JD($1, l$): Jacobi-Davidson with one step approximation as preconditioner (see Section 2.5.2) using the matrix A_0, and l steps of MinRES to approximate the solution of the correction equation (30), with A replaced by A_0;

▷ Full SPAM: each eigenproblem for A_k solved in full precision;

▷ SPAM(1): eigenproblem for A_k approximated with one step of the Jacobi-Davidson method, correction equation (34) in augmented form solved to full precision (see Section 2.5.3);

▷ SPAM($1, l$): using l steps of MinRES to approximate the solution of the correction equation for SPAM(1).

R e m a r k 4.1. To minimize the size of the legends in the pictures, we sometimes write LZS for Lanczos, FSP for Full SPAM, SP(1) for SPAM(1), S12 for SPAM(1,2), JD13 for JD(1,3), etcetera.

In the experiments, we will give illustrations of the following aspects of SPAM.

▷ When a nonzero approximation A_0 of A from below is used, less outer iterations of Full SPAM are needed to arrive close to the dominant eigenvalue of A than with the choice $A_0 = 0$, which is equivalent to the Lanczos method with a random start vector.

▷ Even if the Lanczos method is started with the same approximation of the dominant eigenvector of A_0 as Full SPAM, Full SPAM method will still outperform Lanczos in terms of the number of outer iterations.

▷ Also for other eigenvalues, Full SPAM outperforms Lanczos; this may be expected because in Full SPAM the Ritz Galerkin subspace will be forced in the direction of the appropriate eigenvector in every iteration. In Lanczos, this is done only by means of the start vector. Of course, Lanczos allows efficient implicit restart strategies [26], but also Full SPAM may be restarted. We feel that the comparison would become diffuse and thus we refrain from incorporating restarts.

▷ We investigate the effect of approximating the desired eigenvector of A_k with just one step of the Jacobi-Davidson method, i.e., we will be comparing Full SPAM with SPAM(1).

▷ SPAM$(1, l)$ will be compared with JD$(1, l)$, both started with the same initial vector, i.e., the relevant eigenvector of A_0; i.e., both methods will spend the same number l of matrix-vector products in their inner iteration, where in SPAM$(1, l)$, the matrix will be A_k, and in JD$(1, l)$, the matrix will be A_0. From the viewpoint of this paper, this is the comparison that is of most interest. Not only is JD$(1, l)$ the closest related to SPAM$(1, l)$, the difference between the two is solely the fact that the action of A from the outer loop is taken into the inner iteration of SPAM$(1, l)$, whereas in JD$(1, l)$, this is not done. See the discussion in, and particularly at the end of Section 2.5.4.

▷ Finally, we compare SPAM$(1, l)$ with JD(l). This is perhaps the comparison that the authors of SPAM [21] had in mind: in the inner iteration of JD(l) the original matrix A is used, whereas in SPAM$(1, l)$ it will be A_k.

We will comment on the computational costs of an inner iteration in comparison to having no such costs (as in Lanczos) or the full costs (Jacobi-Davidson), although these costs may depend very much on the specific problem and the available approximations and even on hardware parameters like available memory.

4.2. Lanczos versus full SPAM: Reaction-Diffusion problem, various eigenvalues. In this section we will compare Lanczos with Full SPAM. Our comparison is, for the time being, only in terms of the number of outer iterations. A first naive comparison uses a random start vector for Lanczos, but from then onwards, we will start Lanczos with the appropriate eigenvector of A_0. The approximate matrix A_0 will be constructed using both the approaches described in Sections 3.2.1 and 3.2.2. For this, we discretized the one-dimensional version of (63) on $\Omega = [0, 1]$ using finite differences with grid size $h = 1/33$ with the parameters $\varepsilon = 1/33^2$, $\mathcal{K} = 1$, and $c(x) = x(1 - x)\mathrm{e}^{3x}$. The resulting 32×32 algebraic eigenproblem is of the form $Ax = \lambda x$, where $A = D + R$ is the sum of the tridiagonal discretized diffusion D

and the diagonal discretized reaction R. With the approach from Section 3.2.2 we approximate the largest eigenvalue and with the approach from Section 3.2.1 the smallest eigenvalue.

Natural approximation from below. The left picture of Figure 3 illustrates the typical convergence of the Lanczos method with a random start vector. We display the k Ritz values at outer iteration k as circles above the value k of the iteration number on the horizontal axis. Due to the interlacing property, eigenvalue approximations "converge" to either side of the spectrum, and this is emphasized by the connecting lines between Ritz values for different values of k, both upwards and downwards. In view of Theorem 3.1 we could also say that this picture belongs to SPAM with choice $A_0 = 0$. In the right picture of Figure 3, we show in a similar fashion the convergence of SPAM, using $A_0 = R$ as the approximate matrix, as suggested in Section 3.2.2. We see that the convergence towards the largest eigenvalues is stimulated. The costs for this faster convergence is solving a diagonal plus rank $2k$ eigenproblem in iteration step k. There exist efficient methods for such eigenproblems based on the secular equation and Newton's method.
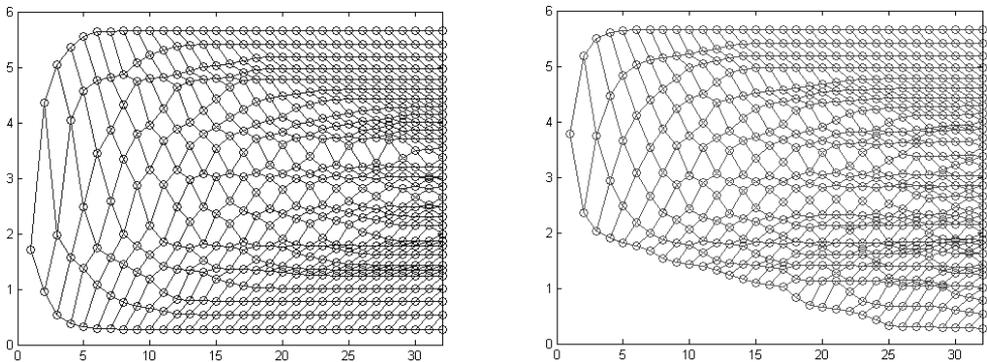

Figure 3. Lanczos method (left) with a random start vector, versus SPAM (right) with the discretized reaction term as approximating matrix A_0, and the largest eigenvalue of A as target.

Algebraic approximation from below. We also tested the algebraic approach from Section 3.2.1 to construct approximations from below. We created a rank-12 approximation A_0 of A based on the largest diagonal elements of A. Full SPAM and Lanczos were started with the dominant eigenvector of A_0 in order to approximate its dominant eigenvalue. The leftmost picture in Figure 4 shows that the incorporation of an approximation A_0 into Full SPAM has an effect that carries beyond that of creating a better start vector for Lanczos. In the rightmost picture of Figure 4, the same was done for the matrix $6I - A$, which is positive definite and has the smallest

eigenvalue of A as dominant eigenvalue. Also here, SPAM outperforms Lanczos in terms of the number of outer iterations. In the middle two pictures of Figure 4, we plotted the absolute error in the second and fifth largest eigenvalues of A. The results are shown from iteration 2 and 5 onwards, respectively, because the Ritz-Galerkin method produces approximations of the second and fifth largest eigenvalues from that iteration onwards. Again, Full SPAM clearly outperforms Lanczos in both cases. Note that, when using Full SPAM to approximate the p-th largest eigenvalue, the Ritz-Galerkin subspace in the outer iteration is in each step expanded with the eigenvector belonging to the p-th largest eigenvalue of A_k. For a fair comparison, we also started Lanczos with the eigenvector of A_0 belonging to the p-th largest eigenvalue.
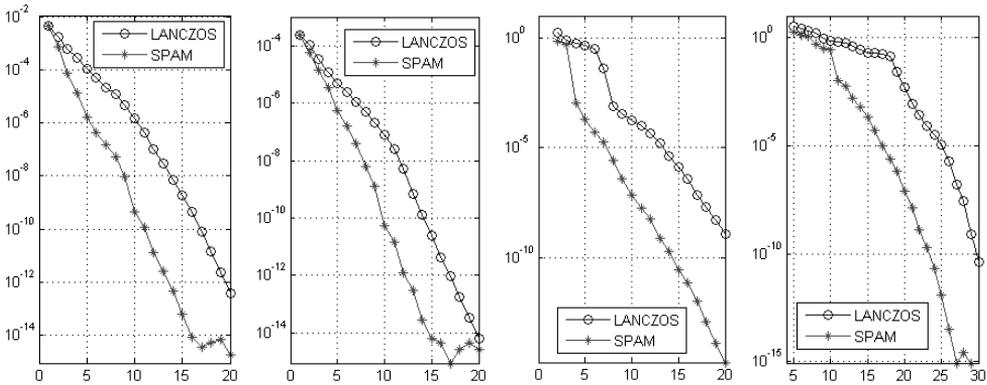

Figure 4. From left to right: Lanczos and Full SPAM approximating the largest, the second, the fifth, and the smallest eigenvalue of A, using algebraic rank-12 approximation from below (for the smallest eigenvalue, we applied both the methods to $6I - A$). Lanczos and Full SPAM used the same start vector in each experiment.

4.3. Lanczos versus Full SPAM and SPAM(1): Banded matrices. In this section we not only compare Lanczos with Full SPAM, but also with SPAM(1), by which we mean that the eigenproblem for A_k in Full SPAM is approximated using one step of the Jacobi-Davidson method, as explained in Section 2.5.3. The correction equation that results is still solved to full precision. For our experiments, we took the banded matrix from Section 3.3 of size 32×32 and with $q = 5$ and $\varepsilon = 0.5$. In [21], it was suggested to take for A_0 the matrix A with a number of its outer diagonals put to zero. We repeated that experiment with A_0 equal to the diagonal of A, so that here too, A_k is diagonal plus a rank-$2k$ perturbation. The comparison of Lanczos, Full SPAM and SPAM(1) is depicted in the left graph in Figure 5. This comparison is initially in favor of Lanczos. This may be due to the fact that the difference $A - A_0$ is indefinite: it has only eleven nonnegative eigenvalues. In the middle left picture we

took the tridiagonal part as approximation. This approximation is indeed positive definite and gives better results. Taking the simple algebraic approximation A_0 from below from Section 3.2.1 based on the three largest diagonal entries of A, which is of rank 6, gives comparable results, but its low rank and higher sparsity make this choice more interesting. In the right graph in Figure 5, the largest eigenvalue of $\alpha I - A$ was approximated with the approximation from below that kept the largest three diagonal entries of $\alpha I - A$, and thus the smallest three diagonal entries of A. In all cases, the positive effect of incorporating an approximation A_0 goes beyond delivering a good start vector for Lanczos. Also in all cases, there is virtually no difference between Full SPAM and SPAM(1).
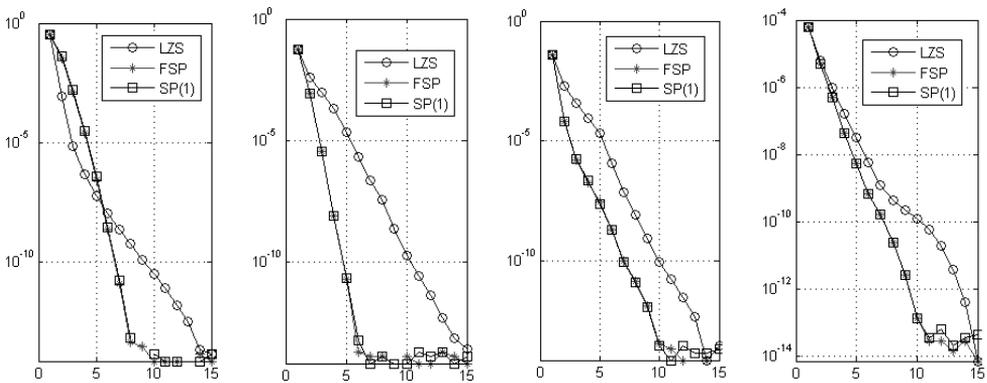

Figure 5. Lanczos versus Full SPAM and SPAM(1) with diagonal (left), tridiagonal (middle left), and with algebraic rank-6 approximation from below (both middle right and right). In the three leftmost pictures, the target was the largest eigenvalue, on the right it was the smallest eigenvalue (i.e., the largest of $\alpha I - A$).

4.4. Lanczos versus Full SPAM and SPAM(1): matrices from structural engineering. As a next set of experiments, we took some matrices from the Harwell-Boeing collection of test matrices. They have their origin in the area of structural engineering and are called *bcsstk04, bcsstk07* and *bcsstk10* and have respective sizes $132 \times 132, 420 \times 420$ and 1024×1024. As approximating matrix A_0 we took approximations from below keeping respectively the largest $12, 20$ and 180 diagonal entries. Recall that in Figure 2 we displayed the sparsity plots of A and A_0 for bcsstk04.mtx. As was the case for the banded matrix in the previous section, Full SPAM and SPAM(1) behave virtually the same, and need less outer iterations than the Lanczos method to arrive at a given accuracy.

Conclusions. The main goal of the experiments so far was, first of all, to investigate if *Full* SPAM is competitive with the Lanczos method in terms of the number of iterations in the outer loop. Solving the eigenproblems for A_k in the inner loop,

447

even if this were not be done to full precision, makes SPAM always more expensive than Lanczos. The experiments in Figures 3, 4, 5 and 6 show that this is the case for different matrices A and different types of approximations A_0. Not only for the largest, but also for other eigenvalues of A. Secondly, we have illustrated that the use of one step of the Jacobi-Davidson method to approximate the eigenvalue problem for A_k in the inner iteration for SPAM hardly influences its behavior. We will now investigate what happens if the linear system in SPAM(1) is approximated with a small number of steps of the Minimal Residual method, MinRES [18].
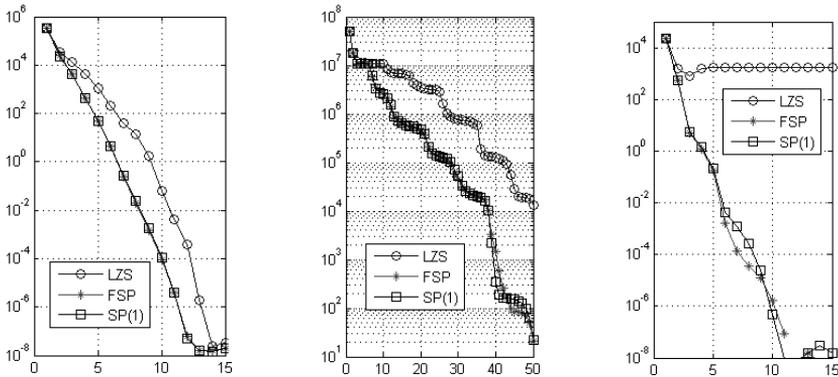

Figure 6. Lanczos versus Full SPAM and SPAM(1) for bcsstk04, bcsstk07 and bcsstk10, with approximating matrices from below. All three methods used the same start vector.

4.5. Lanczos versus SPAM$(1, l)$: approximating the correction equation of SPAM(1) using MinRES. In this section we investigate the use of l iterations of the MinRES method [18] for approximating the solution of the Jacobi-Davidson correction equation (34) for the eigenproblem for A_k in the inner iteration of SPAM(1). Each iteration of MinRES requires one matrix vector product with A_k, which, for small values of k, will be approximately as costly as a matrix vector product with A_0. The initial vector of all methods to which the comparison is applied, i.e., the eigenvector of interest of A_0, will still be computed in full precision. The resulting method is abbreviated by SPAM$(1, l)$.

R e m a r k 4.2. In SPAM$(1, 1)$, MinRES produces an approximation of the Jacobi-Davidson correction equation (34) for A_k in the one-dimensional Krylov subspace with the right-hand side of that correction equation as start vector. Since this is the current residual from the outer iteration, the expansion is the same as for the Lanczos method. We will therefore not display the convergence curves of SPAM$(1, 1)$.

In the light of the previous remark, it is reasonable to expect that SPAM$(1, l)$ will represent a transition between Lanczos and SPAM(1). Thus, for reference, in the

448

experiments to come, we display the convergence graphs of Lanczos and SPAM(1) in solid black lines without any additional symbols. The experiments concern four of the situations that we have already studied. First, we approximated the smallest eigenvalue of the reaction-diffusion problem. For the other three experiments we approximated the largest eigenvalue: in the second, we took the banded matrix with low rank approximation from below, and in the third and fourth the matrices bcsstk07 and bcsstk10 with the respective approximations from the previous section. The results are displayed in Figure 7 and confirm the expectations. Even for $l = 2$ and $l = 3$, SPAM$(1, l)$ resembles SPAM(1) much more than it resembles Lanczos.
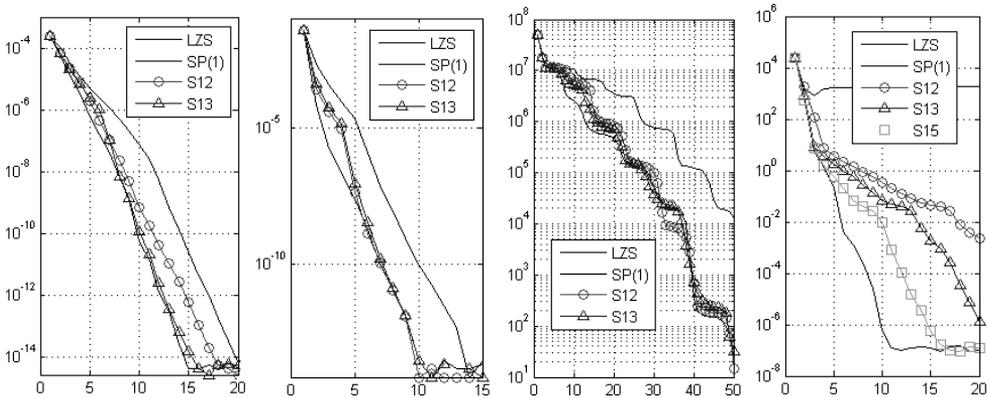

Figure 7. Lanczos and SPAM(1) compared with SPAM$(1, l)$ for small values of l. Left: reaction-diffusion problem, smallest eigenvalue. Other pictures: largest eigenvalue. Middle left: banded matrix; middle right: bcsstk07; right: bcsstk10. The graphs for Lanczos and SPAM(1) can also be found in Figures 4, 5, and 6.

It depends, however, very much on the actual application if the gain in the number of iterations is not undone by the costs of l steps of MinRES per outer iteration with the matrix A_k. For instance, in the banded matrix example (second picture in Figure 7), the matrix A itself has 322 nonzero entries and is of full rank, whereas A_0 has only 33 nonzero elements and is of rank 6, and especially when k is small, the action of A_k will not be very much more expensive than the action of A_0. This, however, brings us to the next question that we would like to investigate, which is, whether using A_k in the k-th inner iteration instead of A_0 all along, is going to make any difference, because this is what distinguishes SPAM from Jacobi-Davidson with one-step approximation as preconditioner. As argued in Section 2.5.4, if SPAM is going to to better, then probably not by very much.

4.6. Comparing SPAM$(1, l)$ **with one-step preconditioned Jacobi-Davidson.** In the k-th inner iteration of SPAM$(1, l)$, the Jacobi-Davidson correction equation (34) for A_k is solved using l steps of MinRES. We will now compare this with

the Jacobi-Davidson method with one-step approximation as preconditioner, as was described in Section 2.5.2. This means that in each inner iteration, the initial approximation A_0 is used instead of A_k. We will still apply l steps of MinRES to solve the corresponding correction equation and denote the resulting method by $JD(1, l)$. Since one of the aims of SPAM was to save on the costs of the matrix-vector products in the inner iteration, we will also apply Jacobi-Davidson *without* preconditioning, and approximate the exact correction equation (30) with matrix A by performing l steps of MinRES as well, and denote this method by $JD(l)$. Thus, the differences between these three methods lie in the inner iteration: $SPAM(1, l)$, $JD(1, l)$ and $JD(l)$ all apply l steps of MinRES per inner iteration step, to a linear equation with matrices A_k, A_0, and A, respectively. Note that in [21], no explicit comparison of SPAM with Jacobi-Davidson was made, even though the methods are so closely related.
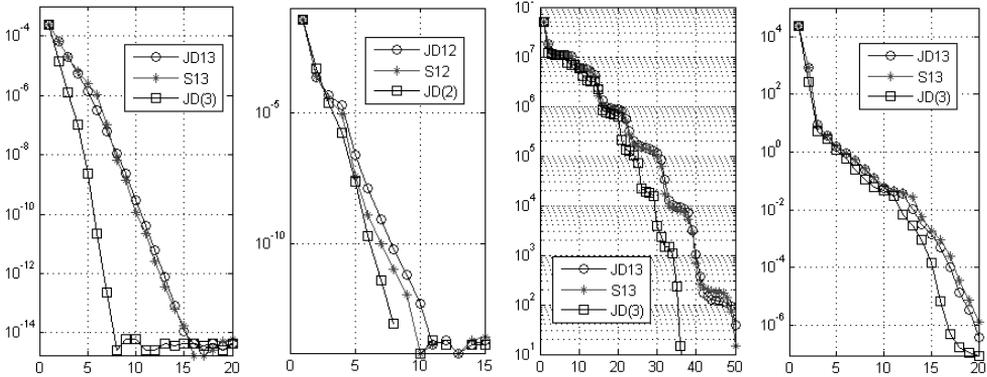

Figure 8. Comparing $SPAM(1, l)$ with $JD(1, l)$ and $JD(l)$. The eigenvalue problems are exactly the same as the corresponding ones in Figure 7, and the curves for $SPAM(1, l)$ can be found there as well.

As expected, $JD(l)$ is the clear winner in all experiments, although the difference from $JD(1, l)$ and $SPAM(1, l)$ is not enough to automatically disqualify the latter two. Since the matrix-vector products in their inner iterations are, in general, considerably cheaper than in $JD(l)$, both methods could be competitive. Having said this, the difference between $JD(1, l)$ and $SPAM(1, l)$ is quite small and not always in favor of $SPAM(1, l)$, even though $SPAM(1, l)$, much more so than $JD(1, l)$, uses the information that is available from the outer iteration also in its inner iteration. As argued already in Section 2.5.4, this may actually be less effective than using only the best information that is available from the outer iteration, as the Jacobi-Davidson method does. So far, the numerical experiments are in favor of using Jacobi-Davidson with one-step preconditioning instead of $SPAM(1, l)$.

5. Conclusions

The experiments above illustrate mathematical aspects of SPAM as a method for approximating eigenvalues of a Hermitian matrix. Using approximations from below, SPAM can be seen as a boosted version of the Lanczos method in the sense that convergence towards the largest eigenvalues is stimulated. Since Lanczos itself is often used to provide a good start vector for Jacobi-Davidson, SPAM is therefore a good candidate for this task, too. Since the difference between SPAM and Jacobi-Davidson with one step approximation is small, it may be preferred to use the latter, especially since the latter is even easier to use. There does not seem to be a significant gain in re-using the action of A on the orthogonal complement $\mathcal{U}$ of the current Ritz vector u within $\mathcal{V}$ also in the inner iterations in comparison with only re-using the action of A on u, as Jacobi-Davidson with one step approximation does. This does not mean that the original idea of the authors [21] of SPAM, to save on the costs of the inner iterations of for instance Jacobi-Davidson, was incorrect. It may well pay off to do so, but this may be done with Jacobi-Davidson with one step approximation just as well. Thus, the main conclusion of this paper is that the value of SPAM probably lies in providing good initial approximations for the Jacobi-Davidson method.

References

[1] *Z. Bai, J. Demmel, J. Dongarra, A. Ruhe, H. van der Vorst (eds.)*: Templates for the Solution of Algebraic Eigenvalue Problems: A Practical Guide. Software-Environments-Tools. 11, SIAM, Philadelphia, 2000.

[2] *F. L. Bauer, C. T. Fike*: Norms and exclusion theorems. Numer. Math. *2* (1960), 137–141.

[3] *C. Beattie*: Harmonic Ritz and Lehmann bounds. ETNA, Electron. Trans. Numer. Anal. *7* (1998), 18–39.

[4] *J. Brandts*: The Riccati algorithm for eigenvalues and invariant subspaces of matrices with inexpensive action. Linear Algebra Appl. *358* (2003), 335–365.

[5] *J. Brandts, R. R. da Silva*: A subspace-projected approximate matrix method for systems of linear equations. East Asian J. Appl. Math. *3* (2013), 120–137.

[6] *W. Chen, B. Poirier*: Parallel implementation of efficient preconditioned linear solver for grid-based applications in chemical physics. II: QMR linear solver. J. Comput. Phys. *219* (2006), 198–209.

[7] *P. G. Ciarlet, J. L. Lions (eds.)*: Handbook of Numerical Analysis. Volume II: Finite Element Methods (Part 1). North-Holland, Amsterdam, 1991.

[8] *M. Crouzeix, B. Philippe, M. Sadkane*: The Davidson method. SIAM J. Sci. Comput. *15* (1994), 62–76.

[9] *E. R. Davidson*: The iterative calculation of a few of the lowest eigenvalues and corresponding eigenvectors of large real-symmetric matrices. J. Comput. Phys. *17* (1975), 87–94.

[10] *M. Genseberger, G. L. G. Sleijpen*: Alternative correction equations in the Jacobi-Davidson method. Numer. Linear Algebra Appl. *6* (1999), 235–253.

[11] *G. H. Golub, H. A. van der Vorst*: Eigenvalue computation in the 20th century. J. Comput. Appl. Math. *123* (2000), 35–65.

[12] *G. H. Golub, C. F. van Loan*: Matrix Computations. The Johns Hopkins Univ. Press, Baltimore, 1996.

[13] *W. Győrffy, P. Seidler, O. Christiansen*: Solving the eigenvalue equations of correlated vibrational structure methods: preconditioning and targeting strategies. J. Chem. Phys. *131* (2009), 024108.

[14] *Z. Jia, G. W. Stewart*: An analysis of the Rayleigh-Ritz method for approximating eigenspaces. Math. Comput. *70* (2001), 637–647.

[15] *C. Lanczos*: An iteration method for the solution of the eigenvalue problem of linear differential and integral operators. J. Research Nat. Bur. Standards *45* (1950), 255–282.

[16] *D. M. Medvedev, S. K. Gray, A. F. Wagner, M. Minkoff, R. Shepard*: Advanced software for the calculation of thermochemistry, kinetics, and dynamics. J. Phys.: Conf. Ser. *16* (2005), 247–251.

[17] *R. B. Morgan, D. S. Scott*: Generalizations of Davidson's method for computing eigenvalues of sparse symmetric matrices. SIAM J. Sci. Stat. Comput. *7* (1986), 817–825.

[18] *C. C. Paige, M. S. Saunders*: Solution of sparse indefinite systems of linear equations. SIAM J. Numer. Anal. *12* (1975), 617–629.

[19] *B. N. Parlett*: The Symmetric Eigenvalue Problem. Classics in Applied Mathematics 20, SIAM, Philadelphia, 1998.

[20] *F. Ribeiro, C. Lung, C. Leforestier*: A Jacobi-Wilson description coupled to a block-Davidson algorithm: an efficient scheme to calculate highly excited vibrational levels. J. Chem. Phys. *123* (2005), 054106.

[21] *R. Shepard, A. F. Wagner, J. L. Tilson, M. Minkoff*: The subspace projected approximate matrix (SPAM) modification of the Davidson method. J. Comput. Phys. *172* (2001), 472–514.

[22] *G. L. G. Sleijpen, J. van den Eshof*: On the use of harmonic Ritz pairs in approximating internal eigenpairs. Linear Algebra Appl. *358* (2003), 115–137.

[23] *G. L. G. Sleijpen, H. A. van der Vorst*: A Jacobi-Davidson iteration method for linear eigenvalue problems. SIAM J. Matrix Anal. Appl. *17* (1996), 401–425.

[24] *G. L. G. Sleijpen, H. A. van der Vorst*: The Jacobi-Davidson method for eigenvalue problems and its relation with accelerated inexact Newton schemes. Iterative Method in Linear Algebra II (S. D. Margenov, P. S. Vassilevski, eds.). IMACS Ann. Comput. Appl. Math. 3, 1996, pp. 377–389.

[25] *G. L. G. Sleijpen, H. A. van der Vorst*: A Jacobi-Davidson iteration method for linear eigenvalue problems. SIAM Rev. *42* (2000), 267–293.

[26] *D. C. Sorensen*: Implicit application of polynomial filters in a k-step Arnoldi method. SIAM J. Matrix Anal. Appl. *13* (1992), 357–385.

[27] *G. W. Stewart*: Matrix Algorithms. Vol. 2: Eigensystems. SIAM, Philadelphia, 2001.

[28] *G. W. Stewart*: A Krylov-Schur algorithm for large eigenproblems. SIAM J. Matrix Anal. Appl. *23* (2002), 601–614; addendum ibid. *24* (2002), 599–601.

[29] *G. W. Stewart, J. G. Sun*: Matrix Perturbation Theory. Academic Press, Boston, 1990.

[30] *L. C. Wrobel, M. H. Aliabadi*: The Boundary Element Method. Vol. 1: Applications in Thermo-Fluids and Acoustics. Wiley, Chichester, 2002.

[31] *Y. Zhou, R. Shepard, M. Minkoff*: Computing eigenvalue bounds for iterative subspace matrix methods. Comput. Phys. Commun. *167* (2005), 90–102.

Authors' address: *Jan H. Brandts* (corresponding author), *Ricardo Reis da Silva*, Korteweg-de Vries Institute for Mathematics, University of Amsterdam, P.O. Box 94248, 1090 GE Amsterdam, The Netherlands, e-mail: `janbrandts@gmail.com`.