Matrix perturbations: bounding and computing eigenvalues

Reis da Silva, R.J.

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
The Subspace Projected Approximate Matrix method

The Subspace Projected Approximate Matrix method [69] is a fairly recent (2001) and unobserved method for the computation of eigenvalues of large Hermitian matrices. It is known by the acronym of SPAM and is interpreted by its inventors as a modification of Davidson’s method [22]. Its most interesting feature is the way it proposes to reduce the costs of matrix-vector products with the original matrix by using an approximation $A_0$, acting only on a part of the space.

In this chapter, we extend the work initiated in [65] and investigate the Subspace Projected Approximated Matrix method in more detail. We begin with a brief contextualization of SPAM within the class of iterative methods for eigenvalue problems. Next, we show how, for certain special choices of the approximation $A_0$, SPAM turns out to be mathematically equivalent to known eigenvalue methods. More sophisticated approximations $A_0$ turn SPAM into a boosted version of the Lanczos method [49], although, it can also be interpreted as an attempt to enhance a certain instance of the preconditioned Jacobi-Davidson method [72]. The chapter ends with a set of numerical experiments tailored to illustrate certain aspects of SPAM and its variations.
4.1 Introduction

The Subspace Projected Approximate Matrix (SPAM) method belongs to the category of inner-outer iteration methods and is interpreted by its inventors as a modification of the Davidson method \cite{22}. 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 \cite[Section 4.1]{72}. There too, the matrix \( A \) is, in the inner iteration, replaced by a preconditioner. The originality of the approach in \cite{69} 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. The resulting approximation is then different after each outer iteration step, even though only one approximate matrix \( A_0 \) needs to be provided. Consequently, the approximated action of \( A \) in the inner iterations becomes more and more accurate as the number of outer iterations increases. As such, SPAM may be interpreted as a discrete homotopy method. Intuitively, one would expect SPAM to outperform Jacobi-Davidson with one-step approximation.

The expansion vector (see line 15 in Algorithm\cite{1}) is itself a suitable eigenvector of the approximation of \( A \). Thus, the matrix-vector products within the inner iteration are cheaper than for instance in the Jacobi-Davidson \cite{72} and Riccati \cite{9} methods. These methods, which we explain in more detail in Section\cite{4.2} both use \( A \) itself in their inner iteration.
4.1. INTRODUCTION

4.1.1 Preliminaries

Before beginning, we assume that $A$ is a Hermitian $n \times n$ matrix with eigenvalues

$$\lambda_{\text{min}} = \lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_{n-1} \leq \lambda_n = \lambda_{\text{max}}. \quad (4.1)$$

Moreover, we recall from Section 1.2.1 that for the unitary matrix $(V|V_\perp)$, with the columns of $V$ forming an orthonormal basis for the current search space $\mathcal{V}$, we have

$$\hat{A} = (V|V_\perp)^* A (V|V_\perp) = \begin{bmatrix} M & R^* \\ R & S \end{bmatrix} \quad (4.2)$$

where $M = V^*AV$ is $k \times k$ and $S = V_\perp^*AV_\perp$. Recall also that the eigenvalues of $M$ are denoted by $\mu_j$ for $j \in \{1, \ldots, k\}$ and satisfy

$$\mu_1 \leq \mu_2 \leq \ldots \leq \mu_{k-1} \leq \mu_k. \quad (4.3)$$

4.1.2 General description of SPAM.

Let $A_0$ be an approximation of $A$. In Section 4.3 we will comment on how this approximation can be chosen. For now, assume only that $A_0$ is available and define, in view of (4.2)

$$\hat{S} = V_\perp A_0 V_\perp^* \quad (4.4)$$

and

$$A_k = (V|V_\perp) \hat{A}_k (V|V_\perp)^*, \text{ where } \hat{A}_k = \begin{bmatrix} M & R^* \\ R & \hat{S} \end{bmatrix}. \quad (4.5)$$

The subscript $k$ of $A_k$ refers to the number of columns of $V$. In [69], 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 $\mathcal{V}$ (see Section 1.2.1). This will be exploited to derive bounds for the eigenvalues of $A_k$ in Section 4.3. 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

\[
\Pi_k = V_\perp V_\perp^* = I - VV^*.
\]  

(4.6)

This shows, together with equations (4.4) and (4.2), that

\[
A_k = -VV^*AVV^* + AVV^* + VV^*A + \Pi_k A_0 \Pi_k.
\]  

(4.7)

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

\[
A_k v = -VMV^*v + WV^*v + VW^*v + \Pi_k A_0 \Pi_k v
\]  

(4.8)

where we have used (4.6) to avoid the numerically infeasible formation of \( V_\perp \). Note that if \( v \in V_\perp \), the first two terms vanish. We now observe the following.

**Remark 4.1.1.** If SPAM runs for the full \( n \) iterations, it produces a unitary matrix \( U_n \) that depends only on \( A \) and \( A_0 \). The \( k \)th 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 (4.5) contains precisely the basis for the orthogonal complement of \( V \) that SPAM is about to produce in future iterations. With respect to this basis, \( \hat{A}_k \) is an update of \( \hat{A}_{k-1} \) of *arrowhead type*, in the sense that

\[
\hat{A}_k - \hat{A}_{k-1} = \begin{bmatrix}
O_{k-1} & \tau & t^* \\
\tau & e_k^* + e_k \left[ \tilde{o}^* \quad \tau \quad t^* \right] - e_k \tau e_k^* \\
t & O
\end{bmatrix}
\]  

(4.9)

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 the matrix \( \hat{A} \) from (4.2) and the matrix \( \hat{A}_0 = (V|V_\perp)^* A_0 (V|V_\perp) \) from (4.5). Thus, with respect to the basis defined by the columns of \( (V|V_\perp) \), the matrix \( \hat{A}_0 \) simply transforms step by step into \( \hat{A} \) in
the sense that after $k$ steps, the first $k$ columns and rows have changed into those of $\hat{A}$, and the resulting matrix is called $\hat{A}_k$. This is visualized in Figure 1.

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

$$A_k = A_{k-1} + uv^* + vu^* = A_{k-1} + (u|v) \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} (u|v)^*, \quad (4.10)$$

where

$$u = \left( \Pi_{k-1} - \frac{1}{2} vv^* \right) (A - A_0)v = \left( \Pi_k + \frac{1}{2} vv^* \right) (A - A_0)v. \quad (4.11)$$

**Proof.** Combining (4.9) with $(V|V_\perp)(\hat{A} - \hat{A}_0)(V|V_\perp)^* = A - A_0$ we find

$$\begin{bmatrix} 0 \\ \tau \\ t \end{bmatrix} = \begin{bmatrix} O & e_k & \cdots & e_n \end{bmatrix}^* (\hat{A} - \hat{A}_0)e_k = (O|v|V_\perp)^*(A - A_0)v \quad (4.12)$$
where \( O \) stands for the \( n \times (k-1) \) zero matrix. Therefore, substituting (4.12) into

\[
A_k - A_{k-1} = (V_{k-1}|v|V_\perp) \begin{pmatrix}
\vec{0} \\
\tau \\
t
\end{pmatrix} e_k^* + e_k \begin{bmatrix}
\vec{0}^* & \tau & t^*
\end{bmatrix} - e_k \tau e_k^*
\begin{pmatrix}
V_{k-1}|v|V_\perp
\end{pmatrix}^*
\]

we arrive, using (4.12) and \( v = (V_{k-1}|v|V_\perp)e_k \), at

\[
A_k - A_{k-1} = \Pi_{k-1}(A - A_0)vv^* + vv^*(A - A_0)\Pi_{k-1} - vv^*(A - A_0)vv^*.
\] (4.13)

Splitting the most right term in two equal parts and rearranging the terms proves the formula for \( u \) in (4.11). Since by (4.6)

\[
\Pi_{k-1} - vv^* = \Pi_k,
\] (4.14)

also the second equality is proved.

\[ \square \]

**Remark 4.1.3.** Note that the result of Proposition 4.1.2 can also be derived in an alternative way, starting from (4.7). The proof presented here also emphasizes the validity of (4.9).

From (4.2) and (4.5) we see that

\[
\text{rank}(A - A_k) = \text{rank}(\hat{A} - \hat{A}_k) \leq n - k.
\]

Thus, even though \( A_k - A_{k-1} \) has rank-2, the update - or maybe better downdate

\[
A - A_k = A - A_{k-1} + (A_{k-1} - A_k)
\]

will generically decrease the rank of \( A - A_{k-1} \) with at most one. This remark goes hand in hand with the observation that even though the approximations \( A_k \) of the matrix \( A \) may seem unusual, the viewpoint of considering the reverse sequence \( O_n = A - A_n, \ A - A_{n-1}, \ldots, \ A - A_1, \ A - A_0 \) 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, in (4.11), both $v^*Av$ and $\Pi_{k-1}Av$ are available 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 (4.10) are cheaply available.

**Remark 4.1.4.** Of course, the update (4.10) itself should not be performed explicitly because it will generally result in fill-in of originally sparse matrices.

### 4.1.3 Choice of the method for the inner iteration of SPAM

In [69], the authors suggest to use Davidson’s method [22] 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 [57] was tested as inner method in [69]. Other possibilities include, for instance, the Jacobi-Davidson [72] 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 study this choice in Section 4.2.

**Remark 4.1.5.** 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 4.4.

In [69] 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 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.

\(^1\)M. Hochstenbach presented this option at the 2006 GAMM/SIAM Applied Linear Algebra conference.
4.1.4 Practical implementation

The implementation of the Subspace Projected Approximate Matrix method is, in theory, given by Algorithm 1 where on line 14 the matrix $A_k$ could be formed explicitly and its target eigenpair determined. Such implementation, however, does not take into account the fact that part of the information needed for the computations with $A_k$ is already available. The subspace projected approximate matrix $A_k$ need not be explicitly assembled. An algorithm closer to the practical implementation is given in Algorithm 3 (see also Algorithm/Figure 2 in [69]).

Remark 4.1.6. In Algorithm 3 we define $V_q := [v_1, \ldots, v_q]$ and $W_q := [w_1, \ldots, w_q]$ and use the indexes $k$ and $p$ to keep track of the dimensions of the inner and outer search subspaces. Moreover, due to the characteristics of the SPAM method, inner and outer iterations are only distinguishable from the type of matrix-vector products. In Algorithm 3 this difference is showed by the variable $\text{it}$ which may assume the value inner or outer. When $\text{it}=$inner the Algorithm computes an approximation for a target eigenpair of $A_k$ and when $\text{it}=$outer for one of $A$.

4.2 SPAM and other subspace methods

The philosophy of SPAM is that only $S$ needs to be approximated. Therefore, all available (computed) information from (4.2), i.e., $M, R$ and $R^*$, is used to determine the expansion vector. The Jacobi-Davidson [72] and Riccati [9] methods partially share this philosophy. Instead of $R$, however, they only use the residual corresponding to the selected eigenvector approximation. On the other hand, in their most simple forms, they do not approximate the matrix $S$ but use its full action. In this section we outline their similarities and differences.

Remark 4.2.1. 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 4.3. There we will show that for the choice $A_0 = O_n$, also SPAM needs no inner iteration, and that this choice makes SPAM mathematically equivalent to Lanczos.
Algorithm 3: Subspace Projected Approximate Matrix method

1: Input: $A$, $A_0$, $v_0$, $\epsilon$;
2: Set $k = 0; p = 1$;
3: \textbf{it} = \textbf{inner};
4: \textbf{loop}
5: $v_{k+p} = v_{k+p}/\|v_{k+p}\|_2$;
6: \textbf{if} \textbf{it} = \textbf{inner} \textbf{then}
7: $w_{k+p} = A_0v_{k+p}$;
8: \textbf{else}
9: $w_k = Av_k$;
10: \textbf{end if}
11: for $j = 1, \ldots, k + p - 1$ do
12: $M_{j,k+p} = v_j^*w_{k+p}$; $M_{k+p,j} = M_{j,k+p}^*$;
13: \textbf{end for}
14: $M_{k+p,k+p} = v_{k+p}^*w_{k+p}$;
15: Compute the target eigenpair $(\mu, z)$ of $M_{k+p,k+p}$;
16: Compute $r = W_{k+p}z - \mu V_{k+p}z$;
17: \textbf{if} $\|r\|_2 \leq \epsilon$ and \textbf{it} = \textbf{outer} \textbf{then}
18: \textbf{stop}; \{ $(\mu, Vz)$ is the desired eigenpair of $A$ \}
19: \textbf{else if} $\|r\|_2 \leq \epsilon$ and \textbf{it} = \textbf{inner} \textbf{then}
20: Set \textbf{it} = \textbf{outer} \{ The target eigenvector $Vz$ of $A_k$ was found. \}
21: $v_{k+1} = V_{k+p}z$;
22: $k = k + 1; p = 0$;
23: Orthogonalize $v_{k+1}$ with respect to the columns of $V_k = [v_1, \ldots, v_k]$;
24: \textbf{else}
25: $v_{k+p+1} = V_{k+p}z$
26: Orthogonalize $v_{p+1}$ with respect to the columns of $V_{k+p} = [v_1, \ldots, v_{k+p}]$;
27: $p = p + 1$; \textbf{it} = \textbf{inner};
28: \textbf{end if}
29: \textbf{end loop}

Recall from Section 1.2.4 that Jacobi-Davidson is simpler than the Riccati method in the sense that only a linear system for $\hat{t}$ needs to be solved. Jacobi-Davidson can be interpreted as an accelerated Newton method [23]. However, much more than the Riccati method, Jacobi-Davidson suffers from stagnation in case the term $\hat{t}\hat{r}^*t$ from (1.9) is not small. On the other hand, if $\hat{t}\hat{r}^*t$ 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 (1.10). For instance, in [25], the authors investigated the effect
of solving the following alternative correction equation,
\[ \tilde{t} \perp V \quad \text{and} \quad (I - VV^*)(A - \mu I)(I - VV^*)\tilde{t} = -\hat{r}. \] (4.15)

At first sight this seems to make sense, because it directly looks for a correction orthogonal to \( V \). Also, the conditioning of the linear equation (4.15) may be better than (1.10) in case the search space contains good approximations of eigenvectors belonging to eigenvalues close to \( \mu \). However, orthogonalizing \( \hat{t} \) from (1.10) to \( V \) generally does not result in \( \tilde{t} \) from (4.15) and the price to pay is that \( \|t - \tilde{t}\|_2 \) is not of higher order, as is \( \|t - \hat{t}\|_2 \). Indeed, in [28] 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 [28] confirm the above observations.

Preconditioning in the Jacobi-Davidson method.

In the original paper on the Jacobi-Davidson method [72], reprinted as [74], preconditioning is discussed as follows. Suppose that an approximation \( A_0 \) of \( A \) is available. It is shown in [72] how to apply such a preconditioner to (1.10), which is a linear equation, though not with system matrix \( A \). To be explicit, since \((I - uu^*)\hat{t} = \tilde{t}\), we have that
\[ (A - \mu I)\hat{t} = -\varepsilon u - \hat{r}. \]
where \( \varepsilon \) is such that \( \hat{t} \perp u \). Or equivalently, written as an augmented system,
\[
\begin{bmatrix}
A - \mu I & u \\
0 & 0
\end{bmatrix}
\begin{bmatrix}
\hat{t} \\
\varepsilon
\end{bmatrix}
= \begin{bmatrix}
-\hat{r} \\
0
\end{bmatrix}.
\] (4.16)

Thus, an approximation \( \hat{t}_0 \) of \( \hat{t} \), together with an approximation \( \varepsilon_0 \) of \( \varepsilon \) can be obtained by replacing \( A \) by \( A_0 \) in (4.16). The pair \( \hat{t}_0, \varepsilon_0 \) can be computed as
\[ \hat{t}_0 = -\varepsilon_0 (A_0 - \mu I)^{-1} u - (A_0 - \mu I)^{-1} \hat{r}, \] (4.17)
with
\[ \epsilon_0 = -\frac{u^*(A_0 - \mu I)^{-1}\hat{r}}{u^*(A_0 - \mu I)^{-1}u}. \] (4.18)

This approximation is called a one step approximation in [72]. It was observed that setting \( \epsilon_0 = 0 \) in (4.17), the Davidson method [22] results. With \( A_0 = A \), which corresponds to Jacobi-Davidson with full accuracy solution of the correction equation, (4.17) becomes
\[ \hat{t} = -\epsilon(A - \mu I)^{-1}u - u, \] (4.19)

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 [72], that Jacobi-Davidson combines the best of those two methods. Of course, a natural next stage in preconditioning is to use the matrix
\[ A_0^n = \begin{bmatrix} A_0 - \mu I & u \\ u^* & 0 \end{bmatrix} \]
as preconditioner within the iterative method that aims to solve (4.16). In each step of such a method one would need to solve a system with \( A_0^n \). This can be done by solving two systems as in (4.16)-(4.17) 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.

**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 \((\mu, u)\) with \( \|u\|_2 = 1 \) 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 \((\mu, u)\) for \( A \) with respect to the current search space \( V_k \) is also a Ritz pair for \( A_k \). Thus we can exploit the fact that \( V_k \) 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 $V_k$ 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:

\[
\begin{bmatrix}
A_k - \mu I & u \\
u^* & 0
\end{bmatrix}
\begin{bmatrix}
t_k \\
\varepsilon_k
\end{bmatrix} = \begin{bmatrix}
-\hat{r} \\
0
\end{bmatrix}.
\] (4.20)

Since quadratic convergence in the outer iteration cannot be expected even if an exact eigenvector of $A_k$ would be 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 $V_k$ for Jacobi-Davidson applied to $A_k$ may be relatively good and may result in quadratic convergence in the inner iteration.

**Remark 4.2.2.** If only one step of Jacobi-Davidson is applied, then after solving $t_k$ from (4.20), the new approximation $v$ for the eigenvector $v_k$ of $A_k$ would lie in the space $V_k \oplus \langle t_k \rangle$. It would not be necessary to actually compute this approximation, because

\[
V_k \oplus \langle v \rangle \subset V_k \oplus \langle t_k \rangle.
\] (4.21)

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.

**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 4.2. 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 investigate the effect of this attempt.
Lemma 4.2.3. Assume that $V_{k+1} = V_k \oplus \langle v \rangle$ with $v \perp V_k$ and $\|v\|_2 = 1$. Then

$$A_{k+1} = -vv^*A + v + vv^*A(I - vv^*)A_k(I - vv^*). \quad (4.22)$$

Proof. By substitution of the defining relation (4.7) for $A_k$.

Corollary 4.2.4. Let $V_1$ be the span of the relevant eigenvector $u$ of $A_0$ and let $\mu = u^*Au$ and $\hat{r} = Au - \mu u$. Then the solution $t_1$ from the system

$$\begin{bmatrix} A_1 - \mu I & u \\ u^* & 0 \end{bmatrix} \begin{bmatrix} t_1 \\ \xi_1 \end{bmatrix} = \begin{bmatrix} -\hat{r} \\ 0 \end{bmatrix},$$

in the first iteration of SPAM(1), to be used to expand $V_1$, coincides with the solution $t_0$ from the system

$$\begin{bmatrix} A_0 - \mu I & u \\ u^* & 0 \end{bmatrix} \begin{bmatrix} t_0 \\ \xi_0 \end{bmatrix} = \begin{bmatrix} -\hat{r} \\ 0 \end{bmatrix},$$

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

Proof. The linear system (4.20) for SPAM(1) with $k = 1$ is equivalent to

$$t_1 \perp u, \quad (I - uu^*)(A_1 - \mu I)(I - uu)t_1 = -\hat{r}, \quad (4.23)$$

where $u$ is a unit vector spanning $V_1$. Substituting the expression (4.22) 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 4.2.5. Given $u$ with $\|u\|_2 = 1$ and $\mu = u^*Au$ and $r = Au - \mu u$. Define for a given approximation $A_0$ of $A$ the matrix

$$A_u = -uu^*Au + uu^*A + Auu^* + (I - uu^*)A_0(I - uu^*)$$

$$= u\mu u^* + ur^* + ru^* + (I - uu^*)A_0(I - uu^*). \tag{4.24}$$

Then $(u, \mu)$ 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

$$(I - uu^*)(A_0 - \mu I)(I - uu^*)t = -r \tag{4.25}$$

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

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

$$(I - uu^*)A_u(I - uu^*) = (I - uu^*)A_0(I - uu^*)$$

its correction equation is precisely (4.25), or equivalently, (4.17). $\square$

Note that $A_u$ from (4.24) is the subspace projected approximate matrix for the one dimensional subspace spanned by $u$. Now, with

$$V_k = \mathcal{U} \oplus \langle u \rangle,$$

where $\mathcal{U}$ is the orthogonal complement of the span of the relevant Ritz vector $u$ in the current search space $V_k$, we have, similar as in (4.22) and (4.24), that

$$A_k = u\mu u^* + ur^* + ru^* + (I - uu^*)A_{\mathcal{U}}(I - uu^*). \tag{4.26}$$

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 (4.24), whereas in SPAM, it is (an
4.3. SELECTING THE MATRIX $A_0$ IN THE SPAM METHOD

arguably good approximation of) the relevant eigenvector of $A_k$ in (4.26). Both matrices have now an appearance that is suitable for studying their differences and similarities.

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 4.2.4, 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 4.4.

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 assume, for simplicity, that the largest eigenvalue of $A$ is the target, although by replacing $A$ by $-A$, we might as well have set the smallest eigenvalue of $A$ as a target.

4.3.1 Some simple choices for $A_0$

It is instructive to study the consequences of the choice $A_0 = O_n$. 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 $A_0 = O_n$, this is a random vector $v \in \mathbb{C}^n$, with $\|v\|_2 = 1$. Thus, we set $V_1 = \text{span}\{v\}$. Write $\mu = v^*Av$ 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
\[
A = (v|V_\perp) \begin{bmatrix} \mu & r^* \\ r & S \end{bmatrix} (v|V_\perp)^*,
\] \hspace{1cm} (4.27)
and consequently, replacing \( S \) by the zero matrix, the next approximating matrix \( A_1 \) from (4.5) is defined by
\[
A_1 = (v|V_\perp) \begin{bmatrix} \mu & r^* \\ r & O_{n-1} \end{bmatrix} (v|V_\perp)^*.
\] \hspace{1cm} (4.28)

As shown already in Proposition 4.1.2, \( A_1 \) 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
\[
w_\pm = \theta_\pm v + \hat{r} \quad \text{and} \quad \theta_\pm = \frac{1}{2} \mu \pm \sqrt{\frac{1}{4} \mu^2 + \|\hat{r}\|_2^2}.
\] \hspace{1cm} (4.29)

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 \( \hat{r} \), we find that
\[
\mathcal{V}_2 = \text{span}\{v, \hat{r}\} = \mathcal{K}^2(A,v),
\]
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 4.3.1.** *If the goal is to find the largest eigenvalue of a positive definite matrix \( A \), the SPAM method with \( A_0 = O_n \) is mathematically equivalent to the Lanczos method.*

**Proof.** Let \( A_0 = O_n \). Then the eigenvalues of \( A_k \) in (4.5) are those of the \( n \times n \) Hermitian arrowhead
\[
\begin{bmatrix} M & R^* \\ R & O_{n-k} \end{bmatrix}.
\]
The Cauchy Interlace Theorem (Theorem 1.1.10) immediately gives that the largest \( k \) of them are each larger than or equal to the corresponding eigenvalue of \( M \). This
assures that the eigenvector of \( A_k \) that is selected for expansion is not from its null space but from its column span. From (4.7) we see that if \( A_0 = O_n \) this column span is the span of \( V \) and \( AV \). A simple induction argument shows that this span equals \( K_{k+1}(A, v) \).

\[ \text{Remark 4.3.2.} \] If \( A \) is indefinite and we would like to find the eigenvalue closest to zero, the choice \( A_0 = O_n \) 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 \( \alpha \) times the identity \( I \) to make it positive semi-definite. Equivalently, instead of using \( A_0 = O_n \) 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 (4.1) and (4.3), makes this explicit.

\[ \text{Proposition 4.3.3.} \] With \( \theta_+ \) as in (4.29) we have that \( \mu \leq \| Av \|_2 \leq \theta_+ \). Moreover, if \( A \) is positive semi-definite, we find that \( \theta_+ \leq \lambda_+ \), where \( \lambda_+ \) is the largest eigenvalue of \( A \).

\[ \text{Proof.} \] Because \( \hat{r} \perp v \) and \( Av - \mu v = \hat{r} \), by Pythagoras’ Theorem we have

\[ \mu^2 \| v \|_2^2 + \| \hat{r} \|_2^2 = \mu^2 + \| \hat{r} \|_2^2 = \| Av \|_2^2, \quad (4.30) \]

hence \( \mu \leq \| Av \|_2 \). Squaring \( \theta_+ \) from (4.29) gives

\[ \theta_+^2 = \| \hat{r} \|_2^2 + \frac{1}{2} \mu^2 + \mu \sqrt{\frac{1}{4} \mu^2 + \| \hat{r} \|_2^2}, \]

which together with (4.30) shows that \( \| Av \|_2 \leq \theta_+ \). Since \( S \) in (4.27) is positive definite whenever \( A \) is, combining (4.27) and (4.28) with Weyl’s bound (Theorem 1.1.7) yields \( \theta_+ \leq \lambda_+ \).
The key issue here is that the inner eigenvalue approximations are much related to the so-called harmonic Ritz values [6, 71] of $A$. Indeed, assuming that $A$ itself is positive semi-definite, these are the $k$ positive eigenvalues of the at most rank-2 matrix

$$\tilde{A}_k = \begin{bmatrix} M & R^* \\ R & T \end{bmatrix}, \quad \text{where} \quad T = RM^{-1}R^*. \quad (4.31)$$

They can be computed without additional matrix vector multiplications with $A$. Note that we introduced in Section 1.2.1 harmonic Ritz values as the reciprocals of the Rayleigh-Ritz approximations of $A^{-1}$ in the space $A\mathbf{v}$. 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 4.3.5. See also [6].

**Proposition 4.3.4.** The matrix $\tilde{A}_k$ can be decomposed as

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

The blocks

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

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

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

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

**Proof.** The statements are all easy to verify. The positivity of the $k$ nonzero eigenvalues follows from Sylvester’s Theorem of Inertia. \qed

The $k$ eigenpairs $(\tilde{\theta}_j, w_j)$ of $\tilde{A}_k$ in (4.31) with positive eigenvalues we label as

$$0 < \tilde{\theta}_1 \leq \tilde{\theta}_2 \leq \ldots \leq \tilde{\theta}_{k-1} \leq \tilde{\theta}_k.$$
4.3. SELECTING THE MATRIX $A_0$ IN THE SPAM METHOD

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 [6, 63].

**Lemma 4.3.5.** For all $j \in \{1, \ldots, k\}$, we have that

$$\mu_j \leq \tilde{\theta}_j \leq \lambda_j.$$  \hspace{1cm} (4.32)

**Proof.** The left inequalities follow from the Cauchy Interlace Theorem (Theorem 1.1.10) applied to $\tilde{A}_k$. Now, with $\hat{A}$ as in (4.2), recognizing the Schur complement $\hat{A}/M$ shows that,

$$\hat{A} - \tilde{A}_k = \begin{bmatrix} O_k & \hat{A}/M \end{bmatrix}$$  \hspace{1cm} (4.33)

is positive semi-definite, hence the right inequalities follow from Weyl’s bound (Theorem 1.1.7).

Observe that, as was the case with the choice $A_0 = O_n$, assuming that from the start of the SPAM method the eigenvector $w_1$ belonging to $\theta_1$ was selected for expansion, the eigenvectors $w_j$, called the harmonic Ritz vectors, lie in the column span of $A V$ and hence, in $K^{k+1}(A, v)$.

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 $A_0 = O_n$, even after modification of $A_k$ into the positive semi-definite matrix $\tilde{A}_k$. In order to get a method different than Lanczos, we should use less trivial approximations that are based on the structure and properties of $A$ itself, while aiming to retain similar inequalities as in Lemma 4.3.5. For this, we would need the matrices $A - A_k$ to be positive semi-definite. We will investigate this in the following sections.

### 4.3.2 One-sided approximations

Having seen that the trivial choice $A_0 = O_n$, even after 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 4.3.6.** If \( A_0 \) approximates \( A \) from below, then so does each matrix \( A_k \).

*Proof.* Combining (4.2) and (4.4) with (4.5) we see that for all \( x \in \mathbb{C}^n \),
\[
x^* (\hat{A} - \hat{A}_k) x = x^* \begin{bmatrix} O_k & \phantom{\sqrt{1}} \\
S - \tilde{S} & \phantom{\sqrt{1}} \end{bmatrix} x = x^* V_\perp^* (A - A_0) V_\perp x \geq 0 \quad (4.34)
\]
where the last inequality holds because \( A - A_0 \) is positive semi-definite. And thus, \( \hat{A} - \hat{A}_k \) is positive semi-definite, and hence, so is \( A - A_k \).

By Proposition 4.1.2, \( A_k - A_{k-1} \) is an indefinite rank-2 matrix, hence it will generally not be true that \( A_{k-1} \) approximates \( A_k \) from below.

**Lemma 4.3.7.** The following inequalities are valid generally,
\[
\theta_{j-1} \leq \mu_j \leq \theta_j \quad \text{for all } j \in \{1, \ldots, k\},
\]
whereas, if \( A_0 \) approximates \( A \) from below, additionally
\[
\theta_j \leq \lambda_j \quad \text{for all } j \in \{1, \ldots, n\}.
\]

*Proof.* The first set of inequalities applies because \( A_k \) has the same Ritz values as \( A \). See also Section 4.1.2. It is well known that the Ritz values interlace the exact eigenvalues \( [63] \) (see also Sections 1.1.4 and 1.2.1). Since \( A - A_k \) is positive semi-definite due to Lemma 4.3.6, the equality \( A_k + (A - A_k) = A \) together with Weyl’s bound (Theorem 1.1.7) proves the second set of inequalities.

Lemma 4.3.7 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.

4.3.3 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 $\mathcal{I} \subset \{1, \ldots, n\}$ of cardinality $m$, let $E_{\mathcal{I}}$ be the matrix with the standard basis vectors $e_j, j \in \mathcal{I}$ as columns. Then set

$$A_0 = A - H, \quad \text{where} \quad H = E_{\mathcal{I}}E_{\mathcal{I}}^*AE_{\mathcal{I}}E_{\mathcal{I}}^*.$$}

The matrix $H$ is the result of a Ritz Galerkin procedure with 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.

Remark 4.3.8. In this particular situation, it is an advantage if $H$ does not approximate $A$ very well, because it is $A_0$ that should be close to $A$, not $H$. Notice also that $A_0$ has zero entries at positions $(j,k)$ for all $j,k \in \mathcal{I}$ and is, thus, always more sparse 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 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 4.2.

Remark 4.3.9. Notice that $\text{rank}(A_0) \leq 2(n - m)$. For large $m$ this may greatly simplify the computation of the eigendata of $A_0$ and of $A_k$ for small values of $k$.

Remark 4.3.10. 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
Figure 4.2: 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 of $A_0$ (with 968 nonzero entries) on the right. See the experiments with this pair of matrices in Section 4.4.4.

the zero block, as was done in (4.31). 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 $A_0 = O_n$.

The rather crude approach that we just described is easy to implement and can be applied in the context of the multilevel version of SPAM, mentioned in Section 4.1.3. 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.

4.3.4 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 [18, 19]. 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

\[
L(u) = \lambda u \quad \text{where} \quad L(u) = -\varepsilon \text{div}(K \nabla u) + cu, \quad \text{and} \quad u = 0 \quad \text{on} \quad \partial \Omega. \quad (4.35)
\]

Here \( \varepsilon > 0 \) is a parameter, \( K : \Omega \to \text{Mat}^{3 \times 3}(\mathbb{R}) \) maps 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 \( \varepsilon \) 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.

Remark 4.3.11. In this context, the algebraic method can be used to approximate the smallest eigenvalue of \( A \) by applying it to \( \alpha I - A \) with \( \alpha \) 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.

4.3.5 Cutting off the bandwidth

Here we describe an obvious choice of approximating matrices that was used in \[69\] 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 4.4 we also intend to apply our algebraic approximations from below to these matrices.
Given $0 \leq \epsilon \leq 1$ and $1 \leq q \leq n - 1$, define a matrix $A = (A_{jk})$ by

$$A_{jk} = \begin{cases} 
  j & \text{if } j = k; \\
  \epsilon^{|j-k|} & \text{if } 1 \leq |j-k| \leq q; \\
  0 & \text{otherwise.}
\end{cases} \quad (4.36)$$

In [69] 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, for

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

For each eigenvalue $\theta$ of $A_0$ there is an eigenvalue $\lambda$ of $A$ with

$$|\lambda - \theta| \leq \frac{\epsilon^{\tilde{q}+1} - \epsilon^{q+1}}{1 - \epsilon} \sqrt{n}. \quad (4.37)$$

Indeed, the difference $A - A_0$ is zero except for the bands $\tilde{q} + 1$ to $q$. Each non-zero row contains at most the numbers $\epsilon^{q+1}$ to $\epsilon^q$, and (4.37) follows from the Bauer-Fike (Theorem 1.1.1) theorem and the finite geometric series sum formula. Thus, for values of $\epsilon$ 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 applications where the boundary element method [84] is used to approximate integral equations. Notice that for such applications also the approach from Section 4.3.3 can be
applied, resulting in both sparse and low rank approximating matrices $A_0$.

4.4 Numerical illustrations

In the previous sections, we have compared SPAM with both the Jacobi-Davidson method and the Lanczos method. We have also studied ways to define an appropriate approximate matrix $A_0$ in the context of the SPAM method. Since the choice $A_0 = O_n$ 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 4.2, SPAM can also be considered as 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.

4.4.1 Objectives

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

- Lanczos (see Section 1.2.3);
- $\text{JD}(\ell)$: Jacobi-Davidson using $\ell$ steps of MinRES to approximate the solution of the exact correction equation (4.16) in augmented form;
- $\text{JD}(1,\ell)$: Jacobi-Davidson with one step approximation as preconditioner (see Section 4.2) using the matrix $A_0$, and $\ell$ steps of MinRES to approximate the solution of the correction equation (4.16), with $A$ replaced by $A_0$;
- Full SPAM: each eigenproblem for $A_k$ solved in full precision;
- $\text{SPAM}(1)$: eigenproblem for $A_k$ approximated with one step of the Jacobi-Davidson method, correction equation (4.20) in augmented form solved to full precision (see Section 4.2);
- $\text{SPAM}(1,\ell)$: using $\ell$ steps of MinRES to approximate the solution of the correction equation for $\text{SPAM}(1)$.
Remark 4.4.1. 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 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 $A_0 = O_n$, 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 [75], 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,\ell) will be compared with JD(1,\ell), both started with the same initial vector, the relevant eigenvector of $A_0$; i.e., both methods will spend the same number $\ell$ of matrix-vector products in their inner iteration, where in SPAM(1,\ell), the matrix will be $A_k$, and in JD(1,\ell), the matrix will be $A_0$. From the viewpoint of this chapter, this is the comparison that is of most interest. Not only is JD(1,\ell) the closest related to SPAM(1,\ell), 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,\ell), whereas in JD(1,\ell), this is not done. See the discussion in, and particularly at the end of Section 4.2.
Finally, we compare SPAM(1,ℓ) with JD(ℓ). This is perhaps the comparison that the authors of SPAM [69] 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.

### 4.4.2 Lanczos versus full SPAM

**Reaction-Diffusion problem, various eigenvalues**

In this section we 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 start Lanczos with the appropriate eigenvector of \( A_0 \). The approximate matrix \( A_0 \) is constructed using both approaches described in Sections 4.3.3 and 4.3.4. For this, we discretized the one-dimensional version of (4.35) on \( \Omega = [0,1] \) using finite differences with grid size \( h = 1/33 \) with the parameters

\[
\varepsilon = \frac{1}{33^2}, \quad K = 1, \quad \text{and} \quad c(x) = x(1-x)e^{3x}.
\]

The resulting 32 \( \times \) 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 4.3.4 we approximate the largest eigenvalue and with the approach from Section 4.3.3 the smallest eigenvalue.

**Natural approximation from below** The left picture of Figure 4.3 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 4.3.1 we could also say that this picture belongs to SPAM with choice $A_0 = O_n$. In the right picture of Figure 4.3 we show in a similar fashion the convergence of SPAM, using $A_0 = R$ as the approximate matrix, as suggested in Section 4.3.4. 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.

**Algebraic approximation from below** We also tested the algebraic approach from Section 4.3.3 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 4.4 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 4.4 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 4.4, we plotted the absolute error in the second and fifth largest eigenvalue 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.

### 4.4.3 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 ap-
4.4. NUMERICAL ILLUSTRATIONS

Figure 4.3: 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.

approximated using one step of the Jacobi-Davidson method, as explained in Section 4.2. The correction equation that results is still solved to full precision. For our experiments, we took the banded matrix from Section 4.3.5 of size $32 \times 32$ and with $q = 5$ and $\epsilon = 0.5$. In [69], 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 $A$, such that here too, $A_k$ is diagonal plus a rank-2k perturbation. The comparison between Lanczos, Full SPAM and SPAM(1) is depicted in the left graph in Figure 4.5. 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 4.3.3 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 4.5 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).
124 CHAPTER 4. THE SPAM METHOD

4.4.4 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 $1024 \times 1024$. As approximating matrix $A_0$ we took approximations from below keeping respectively the largest $12$, $20$ and $180$ diagonal entries. Recall that in Figure 4.2 we displayed the sparsity plots of $A$ and $A_0$ for *bcsstk04*. 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, even if this would not be done to full precision, makes SPAM always more expensive than Lanczos. The experiments in Figures 4.3, 4.4, 4.5 and 4.6 show that
4.4. NUMERICAL ILLUSTRATIONS

4.4.5 Lanczos versus SPAM(1,\(\ell\))

Approximating the correction equation of SPAM(1) using MinRES

In this section we investigate the use of \(\ell\) iterations of the MinRES method [61] for approximating the solution of the Jacobi-Davidson correction equation (4.20) 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
Remark 4.4.2. In SPAM(1,1), MinRES produces an approximation of the Jacobi-Davidson correction equation (4.20) 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,$\ell$) will represent a transition between Lanczos and SPAM(1). Thus, for reference, in the experiments to come, we displayed 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 4.7 and confirm the expectations. Even for $\ell = 2$ and $\ell = 3$, SPAM(1,$\ell$) resembles SPAM(1) much.
more than it resembles Lanczos. 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 4.7), the matrix $A$ itself has 322
nonzero entries and is of full rank, whereas $A_0$ only has 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, if 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 4.2, if SPAM is going to be better, then
probably not by very much.

Figure 4.7: Lanczos and SPAM(1) compared with SPAM(1, ℓ) for small values of ℓ. Left:
reaction-diffusion problem, smallest eigenvalue. Other pictures: largest eigenvalue. Mid-
dle left: banded matrix; middle right: bcsstk07; right: bcsstk10. The graphs for Lanczos
and SPAM(1) can also be found in Figures 4.4, 4.5 and 4.6.

4.4.6 SPAM(1, ℓ) versus one-step preconditioned JD

In the $k$-th inner iteration of SPAM(1, ℓ), the Jacobi-Davidson correction equation
(4.20) 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 4.2. This means that in each inner iteration, the initial approximation $A_0$ is used instead of $A_k$. We will still apply $\ell$ steps of MinRES to solve the corresponding correction equation and denote the resulting method by JD($1, \ell$). 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 (4.16) with the matrix $A$ by performing $\ell$ steps of MinRES as well, and denote this method by JD($\ell$). Thus, the differences between these three methods lie in the inner iteration: SPAM($1, \ell$), JD($1, \ell$) and JD($\ell$) all apply $\ell$ steps of MinRES per inner iteration step, to a linear equation with matrix $A_k$, $A_0$, and $A$, respectively. Note that in [69], no explicit comparison of SPAM with Jacobi-Davidson was made, even though the methods are so closely related.

![Figure 4.8](image)

Figure 4.8: Comparing SPAM($1, \ell$) with JD($1, \ell$) and JD($\ell$). The eigenvalue problems are exactly the same as the corresponding ones in Figure 4.7 and the curves for SPAM($1, \ell$) can be found there as well.

The results are shown in Figure 4.8. As expected, JD($\ell$) is the clear winner in all experiments, although the difference with JD($1, \ell$) and SPAM($1, \ell$) 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($\ell$), both methods could be competitive. Having said this, the difference between JD($1, \ell$) and SPAM($1, \ell$) is quite small and not always in favor of SPAM($1, \ell$), even though
SPAM\((1,\ell)\), much more so than JD\((1,\ell)\) uses the information that is available from the outer iteration also in its inner iteration. As argued already in Section 4.2, 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\((1,\ell)\).

### 4.5 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 a 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 more easy to use. There does not seem to be a significant gain in re-using the action of \(A\) on the orthogonal complement \(U\) of the current Ritz vector \(u\) within \(V\) 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 [69] 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 chapter is that the value of SPAM probably lies in providing good initial approximations for the Jacobi-Davidson method.
4.6 Comments on a SPAM linear system solver

The subspace projected approximate matrices $A_k$ could, at least in theory, also be used in the context of approximating the solution $x$ of a linear system of equations $Ax = b$. Indeed, given a sequence of subspaces $V_1 \subset \cdots \subset V_n$ and the corresponding matrices $A_0, A_1, \ldots, A_n$, an obvious idea is to define $x_k$ as the solution of $A_k x_k = b$, assuming that each $A_k$ is invertible. The initial approximation is then, just as in the eigenvalue case, the solution of $A_0 x_0 = b$, and following the approach in the eigenvalue case even further, the expansion vector $v$ should be chosen as the component of the approximation $x_k$ that is orthogonal to the current search space $V_k$.

Remark 4.6.1. Approximation methods for linear systems are outside the scope of this thesis. For this reason, below we only provide an extended abstract of the paper [13], to which we refer the interested reader for further details.

Approximations based on the subspace $V_k$

The main challenge is, that in order for the SPAM linear system solver proposed above to be competitive with methods like the Conjugate Gradient (CG) method [34] or the Minimal Residual (MinRES) method [61], it should be possible to solve the approximate systems $A_k x_k = b$ in a very efficient way. By definition of $A_k$, these systems have the following form,

$$
\begin{bmatrix}
V^t A V & V^t A V_{\perp} \\
V_{\perp}^t A V & V_{\perp}^t A_0 V_{\perp}
\end{bmatrix}
\begin{pmatrix}
y_1 \\
y_2
\end{pmatrix}
= 
\begin{pmatrix}
V^t b \\
V_{\perp}^t b
\end{pmatrix}, \quad (4.38)
$$

after which the approximation $x_k$ equals

$$
x_k = V y_1 + V_{\perp} y_2.
$$

Note that $x_k$ is, in general, not an element from the search space $V_k$, but rather an approximation based on $V_k$. As such, the selection mechanism is different than
Ritz-Galerkin or a Minimal Residual type of selection, which would solve

\[ V^T A \hat{y}_1 = V^T b \quad \text{and} \quad \begin{bmatrix} V^T A \\ V^T AV \end{bmatrix} \begin{bmatrix} \hat{y}_1 \\ V^T AV \end{bmatrix} = \ell.s \begin{bmatrix} V^T b \\ V^T AV \end{bmatrix}, \quad (4.39) \]

where \(=\ell.s\) stands for the least-squares solution, and set \(\hat{x}_k = V \hat{y}_1\) and \(\tilde{x}_k = V \tilde{y}_1\), respectively. Many iterative solvers, such as CG and MinRES, are based on the selection mechanisms in (4.39). To set up the system (4.38), \(k\) matrix-vector products with \(A\) are needed, which is the same as in the other two approaches displayed in [4.39].

**Employing an \(A_0\)-orthonormal basis for \(V_k\)**

Assuming that \(A_0\) is a symmetric positive definite approximation of \(A\), the basis of the search spaces can be chosen \(A_0\)-orthogonal. With the \(\perp\)-sign indicating \(A_0\) orthogonality, the bottom right block \(V^T A_0 V \perp\) in (4.38) becomes the identity matrix. Using block elimination, it can be shown that the computation of \(x_k\) then needs additional \(k\) system solves with \(A_0\). This is not unreasonable, as long as we would use \(A_0\) also as preconditioner within CG method and MinRES. Alternatively, one could precondition the system \(Ax = b\) before setting up the system (4.38), and then substitute \(A_0 = I\) in the above block system, and use orthonormal bases for the search spaces with respect to the standard inner product. It is then possible to prove that expansion of the spaces with the component of \(x_k\) orthogonal to \(V_k\) in fact leads to Krylov subspaces. The approximations, however, are not from those spaces, but as already mentioned, rather based on the spaces. They exploit all available information about the blocks \(V^T A\) and \(V^T AV \perp\) and doing so, go one step further than CG and MinRES. Since these blocks, for Krylov subspaces, have only one single non-zero entry, the gain, although present, is small. For other subspaces, the gain can be more substantial. Again, for more details on this so-called \(LinSPAM\) method, see [13].