# On the theory of the generalized augmented matrix preconditioning method

Maxim Larin* Alexander Padiy †

**Abstract**

The present paper is devoted to an improvement of the theory of the recently proposed generalized augmented matrix preconditioning method [5]. Namely, we compute a sharp lower bound on eigenvalues of the preconditioned matrix based on the properties of the projector involved in its definition.

KEY WORDS linear system, augmented matrix, preconditioning method.

## 1 Introduction

This work concerns the solution process of the linear system of equation

$$A\,\mathbf{x} = \mathbf{b}, \qquad A = A^T \in \mathbb{R}^{n\times n}, \tag{1}$$

where $A$ is a sparse symmetric positive definite matrix. The dimension of $A$ is assumed to be sufficiently large such that the use of direct methods to solve (1) is prohibitive; we assume that the linear system will be solved iteratively.

It is well-known that the smallest eigenvalues and their corresponding eigenvectors of the matrix $A$ plays a crutial role in solving the system (1) by iterative methods. For example, the widely used preconditioned conjugate gradient (PCG) method efficiently suppressed high-frequency components of the error while remaining low-frequency components is performed less efficiently [1]. In certain cases the spectrum of $A$ consists of a relatively regular part and a cluster of small eigenvalues. The latter causes additional difficulties for iterative methods. In order to improve the rate of convergence we need a preconditioner $B$ to $A$, which is capable to deal with this subspace corresponding to the claster of the smallest eigenvalues. Moreover, it should have a reasonable computational cost to keep the whole computational complexity low.

In the present paper we will consider the recently proposed generalized augmented matrix preconditioning method [5] to construct $B$. This method is based on the bordering preconditioning technique, which was proposed in [2] and futher investigated in [5, 6]. The present paper is devoted to an improvement of the theory of the generalized augmented matrix preconditioning method. Namely, we compute the sharp lower bound on the minimal eigenvalue based on the properties of the projector $P_{\mathcal{V}}$ involved in the definition of the preconditioned matrix $BA$.

*Lehrstuhl für Numerische Mathematik und Institut für Geometrie und Praktische Mathematik, RWTH Aachen, Templergraben 55, D-52056 Aachen, Germany. E-mail: larin@igpm.rwth-aachen.de .

†Philips Research Laboratories Eindhoven, WY 32, Prof. Holstlaan 4, 5656 AA Eindhoven, The Netherlands. E-mail: Alexander.Padiy@philips.com .

## 2 The generalized augmented matrix preconditioning method

In this section we give a short introduction into the generalized augmented matrix preconditioning method. The presentation is mainly based on [5]. The augmented matrix preconditioning approach is related to the following theorem:

**Theorem 1** *[2] Let $A$ be of order $n \times n$ and $V_m$ be of order $n \times m$, where $m < n$. Assume that* $\operatorname{rank} V_m = m$. *Consider the augmented matrix*

$$\widetilde{A} = \begin{bmatrix} A & -A\,V_m \\ -V_m^T A & V_m^T A\,V_m \end{bmatrix} = \begin{bmatrix} I_n & 0 \\ -V_m^T & I_m \end{bmatrix} \begin{bmatrix} A & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} I_n & -V_m \\ 0 & I_m \end{bmatrix}.$$

*Then*

*a) $\widetilde{A}$ has at least $m$ zero eigenvalues. The remaining eigenvalues $\widetilde{\lambda}_i$ of $\widetilde{A}$ are equal to those of $(I + V_m V_m^T)A$.*

*b) If $A$ is symmetric positive definite, then for every eigenvalue $\lambda_i$ of $A$ there exists an eigenvalue $\widetilde{\lambda}_i$ of $\widetilde{A}$ such that $\widetilde{\lambda}_i \geq \lambda_i$.*

*c) If $A$ is nonsingular and symmetric and $V_m = [\alpha_1 \mathbf{v}_1, \ldots, \alpha_m \mathbf{v}_m]$, where $\mathbf{v}_i$ are the normalized eigenvectors of $A$, then the nonzero eigenvalues of $\widetilde{A}$ are the following:*

$$\widetilde{\lambda}_i = \begin{cases} (1+\alpha_i^2)\lambda_i, & i = 1, \ldots, m, \\ \lambda_i, & i = m+1, \ldots, n. \end{cases}$$

Instead of solving $A\mathbf{x} = \mathbf{b}$ one may solve $\widetilde{A}\widetilde{\mathbf{x}} = \widetilde{\mathbf{b}}$. The convergence rate may be improved if the matrix $V$ is chosen such that $\kappa(\widetilde{A}) < \kappa(A)$. Alternatively, one may use the matrix

$$B = I + V_m V_m^T$$

as *a preconditioner* to $A$ instead of solving the linear system with the augumented matrix $\widetilde{A}$. Here we have to note that the preconditioner $B$ can also be rewritten as

$$B = I + VDV^T, \tag{2}$$

where $D = \operatorname{diag}(\alpha_i^2)$ and $V = [\mathbf{v}_1, \ldots, \mathbf{v}_m]$. By choosing the scaling factors $\alpha_i$ according to

$$\alpha_i = \sqrt{\frac{\lambda_n}{\lambda_i} - 1}, \qquad i = 1, \ldots, n,$$

one can improve the condition number of $\widetilde{A}$ by "*moving*" ("*shifting*") the smallest eigenvalues to the maximal eigenvalue, i.e.,

$$\widetilde{\lambda}_i = (1+\alpha_i^2)\lambda_i = \left(1 + \left[\sqrt{\frac{\lambda_n}{\lambda_i} - 1}\right]^2\right)\lambda_i = \lambda_n, \qquad i = 1, \ldots, m.$$

However, the practical implementation of this method is a very problematic since, in general, neither eigenvalues $\lambda_i$ nor their corresponding eigenvectors $\mathbf{v}_i$ are known, and hence, neither the scaling factors $\alpha_i$ nor the matrix $V$ are easily computable. In the sequel we show how to compute and analyse the preconditioner of the form (2) without knowing the exact eigenivalues and eigenvectors of $A$ . The considered approach was suggested in [2] and futher developed in [3, 5, 6].

In what follows we need the following definitions:

**Definition 1** *Denote by* $\mathrm{Span}\{\mathbf{v}_1,\ldots,\mathbf{v}_m\}$ *the subspace spanned onto the system of vectors* $\mathbf{v}_1,\ldots,\mathbf{v}_m$*:*

$$\mathrm{Span}\{\mathbf{v}_1,\ldots,\mathbf{v}_m\} = \left\{\mathbf{x}\in\mathbb{R}^n : \mathbf{x} = \sum_{i=1}^m x_i\mathbf{v}_i,\ x_i\in\mathbb{R}\right\}.$$

**Definition 2** *Denote by* $\mathrm{Im}\,V$ *the image of the rectangular matrix* $V\in\mathbb{R}^{n\times m}$*:*

$$\mathrm{Im}\,V = \{\mathbf{x}\in\mathbb{R}^n : \mathbf{x} = V\mathbf{y} \quad \textit{for all} \quad \mathbf{y}\in\mathbb{R}^m\}.$$

Here we have to note that if $V = [\mathbf{v}_1,\ldots,\mathbf{v}_m]$, then $\mathrm{Im}\,V = \mathrm{Span}\{\mathbf{v}_1,\ldots,\mathbf{v}_m\}$.

**Definition 3** *Denote by* $P_{\mathcal{V}}$ *an orthogonal projector onto the subspace* $\mathcal{V}\subset\mathbb{R}^{n\times n}$*:*

$$\mathrm{Im}\,P_{\mathcal{V}} = \mathcal{V}, \qquad P_{\mathcal{V}}^2 = P_{\mathcal{V}}, \qquad P_{\mathcal{V}}^T = P_{\mathcal{V}}.$$

A nice property of the orthogonal projector $P_{\mathcal{V}}$ is in that for any vector $\mathbf{x}$ from $\mathbb{R}^n$ we can write

$$\mathbf{x} = P_{\mathcal{V}}\mathbf{x} + P_{\mathcal{V}^\perp}\mathbf{x},$$

where $P_{\mathcal{V}^\perp}$ is the projector onto the orthogonal complement of $\mathcal{V}$ denoted by $\mathcal{V}^\perp$:

$$P_{\mathcal{V}^\perp} = I - P_{\mathcal{V}}.$$

Also

$$\lambda_{\min}(P_{\mathcal{V}}) = 0, \qquad \lambda_{\max}(P_{\mathcal{V}}) = 1.$$

The notation of the orthogonal projector provides a useful tool in analysis of preconditioned methods. For example, in [5] the preconditioner of the form (2) was analysed based on the orthogonal projector

$$P_{\mathcal{V}} = A^{1/2}V(V^TAV)^{-1}V^TA^{1/2} \tag{3}$$

onto the space $\mathcal{V} = \mathrm{Im}\,(A^{1/2}V)$.

**Definition 4** *Let* $A$ *be of order* $n\times n$*, then a subspace* $S\in\mathbb{R}^n$ *with the property that*

$$A\mathbf{x}\in S \quad \textit{for all} \quad \mathbf{x}\in S$$

*is said to be* $A$*-invariant.*

According to the definition of eigenvectors an arbitrary set of eigenvectors of A defines the A-invariant subspace. Indeed, let $P_e$ be the orthogonal projector onto $\mathcal{V}_e = \mathrm{Span}\{\mathbf{v}_1,\ldots,\mathbf{v}_m\}$, then $(P_e\mathbf{x})\in\mathcal{V}_e$ and $A(P_e\mathbf{x})\in\mathcal{V}_e$ for all $\mathbf{x}\in R^n$.

**Definition 5** *Denote by* $\cos(\mathcal{V},\mathcal{W})$ *the cosine of the angle between two spaces* $\mathcal{V}$ *and* $\mathcal{W}$*:*

$$\cos(\mathcal{V},\mathcal{W}) = \sup_{\mathbf{x}\in\mathcal{V}_e,\mathbf{y}\in\mathcal{W}} \frac{\mathbf{x}^T\mathbf{y}}{\|\mathbf{x}\|\,\|\mathbf{y}\|}. \tag{4}$$

Note that if $\cos(\mathcal{V},\mathcal{W}) = 0$, then $\mathcal{V}$ and $\mathcal{W}$ are orthogonal. On the other hand, $\cos(\mathcal{V},\mathcal{W}) = 1$ shows only that there is a nontrivial intersection of the subspaces $\mathcal{V}$ and $\mathcal{W}$. Thus, $\cos(\mathcal{V},\mathcal{W})$ is a good measure for the orthogonality of two subspaces.

The next theorem introduce a special form of the preconditioner (2) with the scaling matrix $A_V$ and the shift parameter $\sigma\in\mathbb{R}$ instead of the diagonal matrix $D$.

**Theorem 2** *[5] Let* $A$ *be an* $n\times n$ *symmetric positive semidefinite matrix and let a rectangular matrix* $V$ *of order* $n\times m$ *be defined as* $V = [\mathbf{v}_1,\ldots,\mathbf{v}_m]$*. Assume that* $\mathrm{rank}\,V = m$*. Further, define* $\widetilde{A}$ *as* $\widetilde{A} = (I + \sigma VA_V^{-1}V^T)A$*, where* $A_V = V^TAV$*. Then the following statements hold:*

*a)* $\lambda_{\max}(\tilde{A}) \le \lambda_{\max}(A) + \sigma$;

*b)* *if for some* $i$, $1 \le i \le m$, $\mathbf{v}_i$ *is an eigenvector of* $A$ *with eigenvalue* $\lambda_i$, *then it is also an eigenvector of* $\widetilde{A}$ *with eigenvalue* $\lambda_i + \sigma$;

*c)* *let* $(\lambda_i, \mathbf{v}_i)$ *be the eigenpairs of* $A$ *and assume that* $V = [\mathbf{v}_1, \dots, \mathbf{v}_m]$ *contains* $m$ *eigenvectors. Then*

$$\widetilde{A}\mathbf{v}_i = \begin{cases} (\lambda_i + \sigma)\mathbf{v}_i, & i = 1, \dots, m \\ \lambda_i \mathbf{v}_i, & i = m+1, \dots, n \end{cases}$$

In [5, 6] was shown that one can also use the system of linear independent vectors spanned onto the proper subspace instead of exact eigenvectors during the definition of $V$. It was also shown in [5] that the action of $A_V^{-1}$ can be replaced by the action of a preconditioner $B_V^{-1}$ to $A_V^{-1}$, and a choice of $V$ can be relaxed, as summarized in Theorem 3.

**Theorem 3** *[5] Define the preconditioner* $\hat{B}$ *as*

$$\hat{B} = I + \hat{\sigma} V B_V^{-1} V^T, \quad \hat{\sigma} = \lambda_{\max}(A)/\lambda_{\max}(B_V^{-1}A_V), \quad \operatorname{Im} V \supseteq \operatorname{Span}\{\mathbf{v}_1, \dots, \mathbf{v}_m\}, \tag{5}$$

*where* $B_V$ *is an* $m \times m$ *symmetric positive definite approximation of* $A_V$. *The eigenvalues* $\lambda(\hat{B}A)$ *of* $\hat{B}A$ *are bounded as follows:*

$$\min\left\{\lambda_1 + \frac{\lambda_{\max}(A)}{\kappa(B_V^{-1}A_V)}, \lambda_{m+1}\right\} \le \lambda(\hat{B}A) \le 2\lambda_{\max}(A). \tag{6}$$

It should be noted, however, that the preconditioner (5) is still difficult to implement in practice since the condition $\operatorname{Im} V \supseteq \operatorname{Span}\{\mathbf{v}_1, \dots, \mathbf{v}_m\}$ is in general not easy to satisfy. Next, in Theorem 4, it is shown that this condition can be significantly relaxed.

**Theorem 4** *[5] Consider the preconditioner* $\hat{B}$

$$\hat{B} = I + \hat{\sigma} V B_V^{-1} V^T, \qquad \hat{\sigma} = \lambda_{\max}(A)/\lambda_{\max}(B_V^{-1}A_V). \tag{7}$$

*Assume that* $\{(\lambda_i, \mathbf{v}_i)\}_{i=1}^n$ *is an ordered set of eigenpairs of* $A$ *such that* $\lambda_1 \le \dots \le \lambda_n$. *If the subspaces* $\mathcal{W} = (\operatorname{Im}(A^{\frac{1}{2}}V))^\perp$ *and* $\mathcal{V}_e = \operatorname{Span}\{\mathbf{v}_1, \dots, \mathbf{v}_m\}$ *satisfy the condition*

$$\cos(\mathcal{V}_e, \mathcal{W}) \le \hat{\gamma}$$

*for some* $\hat{\gamma} < 1$, *then the minimal eigenvalue of* $\hat{B}A$ *is bounded as*

$$\lambda_{\min}(\hat{B}A) \ge \max\left\{\lambda_1, (1 - \hat{\gamma}) \cdot \min\left\{\frac{\lambda_{\max}(A)}{\kappa(B_V^{-1}A_V)}, \lambda_{m+1}\right\}\right\} \tag{8}$$

*while* $\lambda_{\max}(\hat{B}A)$ *is bounded as* $\lambda_{\max}(\hat{B}A) \le 2\lambda_{\max}(A)$ *for any choice of* $V$ *and* $B_V$.

Unfortunately, the derived estimate for the smallest eigenvalue dose not correspond with the *exact* one from Theorem 3, i.e., if $\hat{\gamma}$ tends to 1, the lower bound from Theorem 4 does not approach $\lambda_1$ directly, and moreover, involves a "strange" subspace $\mathcal{W} = (\operatorname{Im}(A^{\frac{1}{2}}V))^\perp$ rather than a more natural subspace $\mathcal{V} = \operatorname{Im}(A^{\frac{1}{2}}V)$. Thus, Theorem 4 does not clearly show the real nature of changes, which are caused by the inexact approximation of the minimal eigenvalue space. Finally, we have to note that the definition of the cosine between two spaces given in (4) is not a standard metric for the deviation of two spaces. In the next section we derive a sharp and clear estimate for the smallest eigenvalues, which avoids all these problems.

## 3 New results

First of all, in contrast with [5] we will measure the deviation of a subspace $\mathcal{V}$ from a subspace $\mathcal{W}$ through the classical definition of the distance between equidimensional subspaces, see [4], for example. Secondly, we will present our main result, which is a refiement of the Theorem 4.

**Definition 6** *Let $\mathcal{V}$ and $\mathcal{W}$ be subspaces in $R^n$ such that $dim(\mathcal{V}) = dim(\mathcal{W})$. Then we define the distance between these two spaces by*

$$dist(\mathcal{V}, \mathcal{W}) = \|P_{\mathcal{V}} - P_{\mathcal{W}}\|,$$

*where $P_{\mathcal{V}}$ and $P_{\mathcal{W}}$ are orthogonal projectors onto $\mathcal{V}$ and $\mathcal{W}$, respectively.*

Note that if $dist(\mathcal{V}, \mathcal{W}) = 0$, then $\mathcal{V}$ and $\mathcal{W}$ are identical. On the other hand, when $dist(\mathcal{V}, \mathcal{W}) = 1$ it does not mean that $\mathcal{V}$ and $\mathcal{W}$ are orthogonal. Hence, in contrast with the definition of cosine from (4) the definition of distance from (6) is a good measure for the closeness of two subspaces.

**Theorem 5** *Consider the preconditioner $B$*

$$B = I + \hat{\sigma} V B_V^{-1} V^T, \qquad \sigma = \lambda_{\max}(A)/\lambda_{\max}(B_V^{-1} A_V). \tag{9}$$

*Assume that $\{(\lambda_i, \mathbf{v}_i)\}_{i=1}^n$ is an ordered set of eigenpairs of $A$ such that $\lambda_1 \leq \ldots \leq \lambda_n$. If the subspaces $\mathcal{V} = \operatorname{Im}(A^{1/2} V)$ and $\mathcal{V}_e = \operatorname{Span}\{\mathbf{v}_1, \ldots, \mathbf{v}_m\}$ satisfy the condition*

$$dist(\mathcal{V}, \mathcal{V}_e) \leq \gamma \tag{10}$$

*for some $\gamma \leq 1$, then the minimal eigenvalue of $BA$ is bounded as*

$$\lambda_{\min}(BA) \geq \lambda_1 + (1-\gamma) \cdot \min\left\{\lambda_{m+1} - \lambda_1, \frac{\lambda_n}{\kappa(B_V^{-1} A_V)}\right\}, \tag{11}$$

*while $\lambda_{\max}(BA)$ is bounded as $\lambda_{\max}(BA) \leq 2\lambda_n$ for any choice of $V$ and $B_V$.*

**Proof:** Denote by $P_e$, $P_e^{\perp}$ and $P_{\mathcal{V}}$ orthogonal projectors onto $\mathcal{V}_e$, $\mathcal{V}_e^{\perp}$ and $\mathcal{V}$, respectively. Note that $P_e^{\perp} = I - P_e$ and $P_e P_e^{\perp} = 0$. Moreover, due to $\mathcal{V}_e$ is an A-invariant subspace we have

$$\begin{aligned}
\frac{\mathbf{x}^T A \mathbf{x}}{\mathbf{x}^T \mathbf{x}} &= \frac{\mathbf{x}^T (P_e + P_e^{\perp}) A (P_e + P_e^{\perp}) \mathbf{x}}{\mathbf{x}^T \mathbf{x}} \\
&= \frac{\mathbf{x}^T P_e A P_e \mathbf{x}}{\mathbf{x}^T \mathbf{x}} + \frac{\mathbf{x}^T P_e A P_e^{\perp} \mathbf{x}}{\mathbf{x}^T \mathbf{x}} + \frac{\mathbf{x}^T P_e^{\perp} A P_e \mathbf{x}}{\mathbf{x}^T \mathbf{x}} + \frac{\mathbf{x}^T P_e^{\perp} A P_e^{\perp} \mathbf{x}}{\mathbf{x}^T \mathbf{x}} \\
&= \frac{\mathbf{x}^T P_e A P_e \mathbf{x}}{\mathbf{x}^T \mathbf{x}} + \frac{\mathbf{x}^T P_e^{\perp} A P_e^{\perp} \mathbf{x}}{\mathbf{x}^T \mathbf{x}}.
\end{aligned}$$

Now we obtain the upper estimate the maximal eigenvalue $\lambda_{\max}(BA)$ as follows

$$\begin{aligned}
\lambda_{\max}(BA) &= \sup_{\mathbf{x}} \left\{ \frac{\mathbf{x}^T (I + \sigma V B_V^{-1} V^T) \mathbf{x}}{\mathbf{x}^T A^{-1} \mathbf{x}} \right\} \\
&= \sup_{\mathbf{x}} \left\{ \frac{\mathbf{x}^T \mathbf{x}}{\mathbf{x}^T A^{-1} \mathbf{x}} + \sigma \cdot \frac{\mathbf{x}^T V B_V^{-1} V^T \mathbf{x}}{\mathbf{x}^T V A_V^{-1} V^T \mathbf{x}} \cdot \frac{\mathbf{x}^T V A_V^{-1} V^T \mathbf{x}}{\mathbf{x}^T A^{-1} \mathbf{x}} \right\} \\
&\leq \sup_{\mathbf{x}} \left\{ \frac{\mathbf{x}^T \mathbf{x}}{\mathbf{x}^T A^{-1} \mathbf{x}} + \sigma \cdot \sup_{\mathbf{y}} \left\{ \frac{\mathbf{y}^T V B_V^{-1} V^T \mathbf{y}}{\mathbf{y}^T V A_V^{-1} V^T \mathbf{y}} \right\} \cdot \frac{\mathbf{x}^T V A_V^{-1} V^T \mathbf{x}}{\mathbf{x}^T A^{-1} \mathbf{x}} \right\} \\
&\leq \sup_{\mathbf{x}} \left\{ \frac{\mathbf{x}^T A \mathbf{x}}{\mathbf{x}^T \mathbf{x}} + \sigma \cdot \lambda_{\max}(B_V^{-1} A_V) \cdot \frac{\mathbf{x}^T A^{1/2} V A_V^{-1} V^T A^{1/2} \mathbf{x}}{\mathbf{x}^T \mathbf{x}} \right\} \\
&\leq \lambda_n + \lambda_n \cdot \sup_{\mathbf{x}} \frac{\mathbf{x}^T P_{\mathcal{V}} \mathbf{x}}{\mathbf{x}^T \mathbf{x}} \leq 2\lambda_n.
\end{aligned}$$

The lower estimate for the minimal eigenvalue $\lambda_{\min}(BA)$ can be obtained as follows

$$
\begin{aligned}
\lambda_{\min}(BA) &= \inf_{\mathbf{x}} \left\{ \frac{\mathbf{x}^T (I + \sigma V B_V^{-1} V^T) \mathbf{x}}{\mathbf{x}^T A^{-1} \mathbf{x}} \right\} \\
&= \inf_{\mathbf{x}} \left\{ \frac{\mathbf{x}^T \mathbf{x}}{\mathbf{x}^T A^{-1} \mathbf{x}} + \sigma \cdot \frac{\mathbf{x}^T V B_V^{-1} V^T \mathbf{x}}{\mathbf{x}^T V A_V^{-1} V^T \mathbf{x}} \cdot \frac{\mathbf{x}^T V A_V^{-1} V^T \mathbf{x}}{\mathbf{x}^T A^{-1} \mathbf{x}} \right\} \\
&\geq \inf_{\mathbf{x}} \left\{ \frac{\mathbf{x}^T \mathbf{x}}{\mathbf{x}^T A^{-1} \mathbf{x}} + \sigma \cdot \inf_{\mathbf{y}} \left\{ \frac{\mathbf{y}^T V B_V^{-1} V^T \mathbf{y}}{\mathbf{y}^T V A_V^{-1} V^T \mathbf{y}} \right\} \cdot \frac{\mathbf{x}^T V A_V^{-1} V^T \mathbf{x}}{\mathbf{x}^T A^{-1} \mathbf{x}} \right\} \\
&\geq \inf_{\mathbf{x}} \left\{ \frac{\mathbf{x}^T A \mathbf{x}}{\mathbf{x}^T \mathbf{x}} + \sigma \cdot \lambda_{\min}(B_V^{-1} A_V) \cdot \frac{\mathbf{x}^T A^{1/2} V A_V^{-1} V^T A^{1/2} \mathbf{x}}{\mathbf{x}^T \mathbf{x}} \right\} \\
&= \inf_{\mathbf{x}} \left\{ \frac{\mathbf{x}^T A \mathbf{x}}{\mathbf{x}^T \mathbf{x}} + \frac{\lambda_n}{\kappa(B_V^{-1} A_V)} \cdot \frac{\mathbf{x}^T P_{\mathcal{V}} \mathbf{x}}{\mathbf{x}^T \mathbf{x}} \right\} \\
&= \inf_{\mathbf{x}} \left\{ \frac{\mathbf{x}^T P_e A P_e \mathbf{x}}{\mathbf{x}^T \mathbf{x}} + \frac{\mathbf{x}^T P_e^{\perp} A P_e^{\perp} \mathbf{x}}{\mathbf{x}^T \mathbf{x}} + \frac{\lambda_n}{\kappa(B_V^{-1} A_V)} \cdot \frac{\mathbf{x}^T P_{\mathcal{V}} \mathbf{x}}{\mathbf{x}^T \mathbf{x}} \right\} \\
&\geq \inf_{\mathbf{x}} \left\{ \lambda_1 \cdot \frac{\mathbf{x}^T P_e \mathbf{x}}{\mathbf{x}^T \mathbf{x}} + \lambda_{m+1} \cdot \frac{\mathbf{x}^T P_e^{\perp} \mathbf{x}}{\mathbf{x}^T \mathbf{x}} + \frac{\lambda_n}{\kappa(B_V^{-1} A_V)} \cdot \frac{\mathbf{x}^T P_{\mathcal{V}} \mathbf{x}}{\mathbf{x}^T \mathbf{x}} \right\} \\
&= \inf_{\mathbf{x}} \left\{ \lambda_1 \cdot \frac{\mathbf{x}^T (I - P_e^{\perp}) \mathbf{x}}{\mathbf{x}^T \mathbf{x}} + \lambda_{m+1} \cdot \frac{\mathbf{x}^T P_e^{\perp} \mathbf{x}}{\mathbf{x}^T \mathbf{x}} + \frac{\lambda_n}{\kappa(B_V^{-1} A_V)} \cdot \frac{\mathbf{x}^T P_{\mathcal{V}} \mathbf{x}}{\mathbf{x}^T \mathbf{x}} \right\} \\
&= \lambda_1 + \inf_{\mathbf{x}} \left\{ (\lambda_{m+1} - \lambda_1) \cdot \frac{\mathbf{x}^T P_e^{\perp} \mathbf{x}}{\mathbf{x}^T \mathbf{x}} + \frac{\lambda_n}{\kappa(B_V^{-1} A_V)} \cdot \frac{\mathbf{x}^T P_{\mathcal{V}} \mathbf{x}}{\mathbf{x}^T \mathbf{x}} \right\} \\
&\geq \lambda_1 + \min \left\{ \lambda_{m+1} - \lambda_1, \frac{\lambda_n}{\kappa(B_V^{-1} A_V)} \right\} \cdot \inf_{\mathbf{x}} \left\{ \frac{\mathbf{x}^T P_e^{\perp} \mathbf{x} + \mathbf{x}^T P_{\mathcal{V}} \mathbf{x}}{\mathbf{x}^T \mathbf{x}} \right\} \\
&= \lambda_1 + \min \left\{ \lambda_{m+1} - \lambda_1, \frac{\lambda_n}{\kappa(B_V^{-1} A_V)} \right\} \cdot \inf_{\mathbf{x}} \left\{ \frac{\mathbf{x}^T (I - P_e) \mathbf{x} + \mathbf{x}^T P_{\mathcal{V}} \mathbf{x}}{\mathbf{x}^T \mathbf{x}} \right\} \\
&= \lambda_1 + \min \left\{ \lambda_{m+1} - \lambda_1, \frac{\lambda_n}{\kappa(B_V^{-1} A_V)} \right\} \cdot \inf_{\mathbf{x}} \left\{ \frac{\mathbf{x}^T (I + (P_{\mathcal{V}} - P_e)) \mathbf{x}}{\mathbf{x}^T \mathbf{x}} \right\} \\
&\geq \lambda_1 + \min \left\{ \lambda_{m+1} - \lambda_1, \frac{\lambda_n}{\kappa(B_V^{-1} A_V)} \right\} \cdot (1 - \|P_{\mathcal{V}} - P_e\|) \\
&= \lambda_1 + \min \left\{ \lambda_{m+1} - \lambda_1, \frac{\lambda_n}{\kappa(B_V^{-1} A_V)} \right\} \cdot (1 - dist(\mathcal{V}, \mathcal{V}_e)) \\
&\geq \lambda_1 + \min \left\{ \lambda_{m+1} - \lambda_1, \frac{\lambda_n}{\kappa(B_V^{-1} A_V)} \right\} \cdot (1 - \gamma)\,.
\end{aligned}
$$

□

Finally, as it was mentioned in [5] the method presented can be combined with a smoother, which essentially improves the conditioning by making the largest eigenvalues smaller.

The preconditioner is constructed as

$$\tilde{B} = M^{-1} + \tilde{\sigma} V B_V^{-1} V^T, \qquad \tilde{\sigma} = \lambda_{\max}(M^{-1}A)/\lambda_{\max}(B_V^{-1} A_V), \tag{12}$$

where $M$ and $B_V$ are symmetric positive definite preconditioners for $A$ and $A_V$, respectively.

**Theorem 6** *Consider the preconditioner (12). Assume that* $\{(\lambda_i, \mathbf{v}_i)\}_{i=1}^n$ *is an ordered set of eigenpairs of* $M^{-1}A$ *such that* $\lambda_1 \leq \ldots \leq \lambda_n$*. Define the matrix* $V_e$ *as* $V_e = [\mathbf{v}_1, \ldots, \mathbf{v}_m]$*. If the subspaces* $\mathcal{V} = \operatorname{Im}(A^{1/2}V)$ *and* $\mathcal{V}_e = \operatorname{Im} V_e$ *satisfy the condition (10) for some* $\gamma < 1$ *then the minimal eigenvalue of* $\tilde{B}A$ *is bounded as*

$$\lambda_{\min}(\tilde{B}A) \geq \lambda_1 + (1-\gamma) \cdot \min\left\{\lambda_{m+1} - \lambda_1, \frac{\lambda_n}{\kappa(B_V^{-1}A_V)}\right\}. \tag{13}$$

*The maximal eigenvalue of* $\tilde{B}A$ *is bounded as*

$$\lambda_{\max}(\tilde{B}A) \leq 2\lambda_{\max}(M^{-1}A) \tag{14}$$

*for any choice of* $V$ *and* $B_V$*.*

**Proof:** Taking into account that now $P_e$ is an orthogonal projector onto the eigenspace of $M^{-1}A$ the maximal and minimal eigenvalues of $BA$ can be estimated as in the proof of Theorem 5. □

Unfortunately, the above estimates need a good approximation $\mathcal{V}$ to the exact low-frequency eigensubspace $\mathcal{V}_e$. The following theorem relaxes this problem.

**Theorem 7** *Consider the preconditioner (12). Assume that there exist two matrices* $\hat{M}$ *and* $\hat{A}$ *such that* $\hat{M} = \hat{M}^T > 0$, $\hat{A} = \hat{A}^T \geq 0$, $\hat{M} \geq M$ *and* $\hat{A} \leq A$ *(all inequalities here are meant in positive definite sense). Assume also that* $\{(\hat{\lambda}_i, \hat{\mathbf{v}}_i)\}_{i=1}^n$ *is an ordered set of eigenpairs of* $\hat{M}^{-1}\hat{A}$ *such that* $\hat{\lambda}_1 \leq \cdots \leq \hat{\lambda}_n$*. Define the matrix* $\hat{V}_e$ *as* $\hat{V}_e = [\hat{\mathbf{v}}_1, \ldots, \hat{\mathbf{v}}_m]$*. If the subspaces* $\mathcal{V} = \operatorname{Im}(A^{1/2}V)$ *and* $\hat{\mathcal{V}}_e = \operatorname{Im} \hat{V}_e$ *satisfy the condition (10) for some* $\hat{\gamma} < 1$ *then the minimal eigenvalue of* $\tilde{B}A$ *is bounded as*

$$\lambda_{\min}(\tilde{B}A) \geq \hat{\lambda}_1 + (1-\hat{\gamma}) \cdot \min\left\{\hat{\lambda}_{m+1} - \hat{\lambda}_1, \frac{\lambda_n}{\kappa(B_V^{-1}A_V)}\right\}. \tag{15}$$

*The maximal eigenvalue of* $\tilde{B}A$ *is bounded as*

$$\lambda_{\max}(\tilde{B}A) \leq 2\lambda_{\max}(M^{-1}A) \tag{16}$$

*for any choice of* $V$ *and* $B_V$*.*

**Proof:** The maximal eigenvalue $\lambda_{\max}(\tilde{B}A)$ can be estimated in the standard way as in the proof of Theorem 5, whereas the minimal eigenvalue $\lambda_{\min}(\tilde{B}A)$ need a special treatment. Indeed, we have

$$\begin{aligned}
\lambda_{\min}(\tilde{B}A) &= \inf_{\mathbf{x}} \left\{ \frac{\mathbf{x}^T(M^{-1} + \tilde{\sigma} V B_V^{-1} V^T)\mathbf{x}}{\mathbf{x}^T A^{-1} \mathbf{x}} \right\} \\
&\geq \inf_{\mathbf{x}} \left\{ \frac{\mathbf{x}^T M^{-1} A \mathbf{x}}{\mathbf{x}^T \mathbf{x}} + \tilde{\sigma} \cdot \lambda_{\min}(B_V^{-1} A_V) \cdot \frac{\mathbf{x}^T A^{1/2} V A_V^{-1} V^T A^{1/2} \mathbf{x}}{\mathbf{x}^T \mathbf{x}} \right\} \\
&\geq \inf_{\mathbf{x}} \left\{ \frac{\mathbf{x}^T \hat{M}^{-1} \hat{A} \mathbf{x}}{\mathbf{x}^T \mathbf{x}} + \frac{\lambda_n}{\kappa(B_V^{-1} A_V)} \cdot \frac{\mathbf{x}^T A^{1/2} V A_V^{-1} V^T A^{1/2} \mathbf{x}}{\mathbf{x}^T \mathbf{x}} \right\} \\
&\geq \inf_{\mathbf{x}} \left\{ \hat{\lambda}_1 \cdot \frac{\mathbf{x}^T \hat{P}_e \mathbf{x}}{\mathbf{x}^T \mathbf{x}} + \hat{\lambda}_{m+1} \cdot \frac{\mathbf{x}^T \hat{P}_e^{\perp} \mathbf{x}}{\mathbf{x}^T \mathbf{x}} + \frac{\lambda_n}{\kappa(B_V^{-1} A_V)} \cdot \frac{\mathbf{x}^T P_{\mathcal{V}} \mathbf{x}}{\mathbf{x}^T \mathbf{x}} \right\} \\
&\geq \hat{\lambda}_1 + \min\left\{\hat{\lambda}_{m+1} - \hat{\lambda}_1, \frac{\lambda_n}{\kappa(B_V^{-1} A_V)}\right\} \cdot \inf_{\mathbf{x}} \left\{ \frac{\mathbf{x}^T (I + (P_{\mathcal{V}} - \hat{P}_e)) \mathbf{x}}{\mathbf{x}^T \mathbf{x}} \right\} \\
&\leq \hat{\lambda}_1 + \min\left\{\hat{\lambda}_{m+1} - \hat{\lambda}_1, \frac{\lambda_n}{\kappa(B_V^{-1} A_V)}\right\} \cdot (1 - \hat{\gamma}),
\end{aligned}$$

where $\hat{P}_e$, $\hat{P}_e^{\perp}$ and $P_{\mathcal{V}}$ are the orthogonal projectors onto $\hat{\mathcal{V}}_e$, $\hat{\mathcal{V}}_e^{\perp}$ and $\mathcal{V}$, respectively. □

## 4 Concluding remarks

In the present paper we presented a number of improved estimates for the class of preconditioners developed in [5]. In particular, we replaced the cumbersome cosine condition involving the subspace $\mathcal{W}$ by a much more natural distance condition between $\mathcal{V}$ and $\mathcal{V}_e$. This greatly simplifies conditions and analysis of the present class of preconditioners when applied in practice.

## References


[1] O. Axelsson, *Iterative Solution Methods*, Cambridge University Press, New York, 1994.

[2] O. Axelsson, M. Neytcheva and B. Polman, *The bordering method as a preconditioning method*, internal report #9348, University of Nijmegen (1993).

[3] O. Axelsson and A. Padiy, *On the additive version of the algebraic multilevel iteration method for anisotropic elliptic problems*, SIAM J. Sci. Comput., **20** (1999), pp. 1807–1830.

[4] G. Golub and C. Van Loan, *Matrix Computations*, The Johns Hopkins University Press, 1996.

[5] A. Padiy, O. Axelsson and B. Polman, *Generalized augmented matrix preconditioning approach and its application to iterative solution of ill-conditioned algebraic systems*, SIAM J. Matrix Anal. Appl., **22** (2000), pp. 793–818.

[6] A. Padiy and M. Larin, *Model analysis of a subspace correction technique for anisotropic diffusion problems*, internal report #9818, University of Nijmegen, 1998.