Refined spectral estimates for preconditioned saddle point linear systems in a non-standard inner product

Mattia Tani[1] Valeria Simoncini[2]

November 16, 2012

Abstract

Indefinite Preconditioners can lead to effective strategies for solving algebraic linear systems in saddle point form. Short-term iterative methods such as Conjugate Gradients can be employed if an inner product can be determined that makes the preconditioned coefficient matrix symmetric and positive definite with respect to that inner product. We present new detailed spectral estimates for such preconditioned problem, that may be used to improve our understanding of the expected behavior of indefinite preconditioners when applied to real application problems.

Contents

1 Introduction

We are interested in large saddle point linear systems in the form

$$\mathcal{K}z = b, \qquad \mathcal{K} = \begin{bmatrix} A & B^T \\ B & 0 \end{bmatrix}, \tag{1}$$

with A symmetric and positive semidefinite, and B^T tall and full column rank, with $\ker B \cap \ker A = \emptyset$. This type of linear equations arises in a large variety of applications, and it has recently attracted a lot of attention, as specifically designed solution and preconditioning strategies can be devised to efficiently solve the problem. We refer to Benzi et al. [4] for a recent survey on various theoretical and computational issues associated with the numerical solution of (1). Since $\mathcal{K}$ is in general highly indefinite, symmetric and positive definite block diagonal preconditioning procedures are often employed, which maintain the symmetry of the problem, so that a short-term iterative system solver can be used. On the other hand, it has been observed that *indefinite* preconditioning strategies, that try to mimic the coefficient matrix block structure, may lead to very effective solution methods. Various strategies have been proposed to cope with the resulting nonsymmetry, that aim to exploit the still rich algebraic structure [4, 9, 7, 12]. Here we concentrate on a strategy that allows one to use an iterative solver for positive definite matrices with short-term recurrences, by using a non-standard inner product during the iterative procedure. This strategy was analyzed in detail in Schöberl and Zulehner [13], where the application to linear systems stemming from PDE-constrained optimization problems is also discussed; see also [11] and references therein for alternatives.

In this paper we refine the spectral analysis provided by Schöberl and Zulehner, and we experimentally show that this analysis can provide new insight in the understanding of the actual performance of the linear system solver.

2 CG in a non-standard inner product

Consider the linear system

$$\mathcal{A}x = b \tag{2}$$

with $\mathcal{A} \in \mathbb{R}^{N\times N}$ nonsingular and $b \in \mathbb{R}^N$. Let $\mathcal{B} \in \mathbb{R}^{N\times N}$ be a symmetric and positive definite (SPD) matrix, and define the associated norm on $\mathbb{R}^N$, $\|v\|_{\mathcal{B}} := \sqrt{v^T\mathcal{B}v}$. A Conjugate Gradient (CG) method is an iterative method whose i-th iterate x_i ($i = 1, 2, \ldots$) lies in $x_0 + \mathbb{K}_i(\mathcal{A}, s_0)$, where s_0 is the initial residual, and $\mathbb{K}_i$ is the Krylov subspace $\mathbb{K}_i(\mathcal{A}, s_0) = \operatorname{span}\{s_0, \mathcal{A}s_0, \ldots, \mathcal{A}^{i-1}s_0\}$, such that

$$\|e_i\|_{\mathcal{B}} := \|x^* - x_i\|_{\mathcal{B}} = \min_{x \in x_0 + \mathbb{K}_i} \|x^* - x\|_{\mathcal{B}}\,;$$

see Ashby et al. [1] for a taxonomy. By exploiting some orthogonality properties, the approximate solution at iteration i is obtained from the iterates at the previous iteration. This leads to a short-term recurrence, and only few vectors need to be stored in memory. Necessary and sufficient conditions on $\mathcal{B}$ and $\mathcal{A}$ for a CG method to be computable were discussed by Faber and Manteuffel [6]. In our context, these conditions are met if $\mathcal{B} = \mathcal{D}\mathcal{A}$, with $\mathcal{D}$ SPD and such that $\mathcal{D}s_i$ can be computed at every step of the algorithm.

For the $\mathcal{B}$-norm of the error of a CG method it holds (see, e.g., [13])

$$\|e_i\|_{\mathcal{B}} \leq \frac{2q^i}{1+q^{2i}}\,\|e_0\|_{\mathcal{B}}\,, \quad q = \frac{\sqrt{\kappa_{\mathcal{B}}(\mathcal{A})}-1}{\sqrt{\kappa_{\mathcal{B}}(\mathcal{A})}+1}, \tag{3}$$

where $\kappa_{\mathcal{B}}(\mathcal{A}) = \|\mathcal{A}\|_{\mathcal{B}} \cdot \|\mathcal{A}^{-1}\|_{\mathcal{B}} = \lambda_{\max}(\mathcal{A})/\lambda_{\min}(\mathcal{A})$ is the real $\mathcal{B}$-condition number of $\mathcal{A}$. Given a matrix $\mathcal{A}$, if there exists $\mathcal{D}$ SPD such that $\mathcal{D}\mathcal{A}$ is SPD, then $\mathcal{A}$ is similar to an SPD matrix [8, Theorem 6.2 and its proof], so that its eigenvalues are real and positive, and $\kappa_{\mathcal{B}}(\mathcal{A})$ is well defined.

The estimate in (3) shows that the error $\mathcal{B}$-norm is bounded by a quantity that only depends on the eigenvalues of the possibly nonsymmetric $\mathcal{A}$, and the use of the $\mathcal{B}$-norm is key for this to happen. In our context, $\mathcal{A}$ is a preconditioned saddle point matrix, that is $\mathcal{A} = \widehat{\mathcal{K}}^{-1}\mathcal{K}$, where $\widehat{\mathcal{K}}$ is the selected preconditioner. Schöberl and Zulehner [13] consider the following symmetric and indefinite matrix:

$$\widehat{\mathcal{K}} = \begin{bmatrix} \hat{A} & B^T \\ B & B\hat{A}^{-1}B^T - \hat{S} \end{bmatrix},$$

where $\hat{A}$ and $\hat{S}$ approximate A and $B\hat{A}^{-1}B^T$, respectively, and satisfy

$$A < \hat{A} \quad \text{and} \quad \alpha x^T\hat{A}x \leq x^TAx \quad \forall\, x \in \ker B, \quad \alpha < 1, \tag{4}$$

$$\hat{S} < B\hat{A}^{-1}B^T \leq \beta\hat{S}, \quad \beta > 1. \tag{5}$$

We recall the following result from Schöberl and Zulehner [13].

Theorem 1 *Let* (4) *and* (5) *hold. Then* $\mathcal{D} := \widehat{\mathcal{K}} - \mathcal{K}$ *is SPD and* $\mathcal{D}\widehat{\mathcal{K}}^{-1}\mathcal{K}$ *is SPD. Moreover,*

$$\lambda_{\max}(\widehat{\mathcal{K}}^{-1}\mathcal{K}) \leq \beta(1 + \sqrt{1 - 1/\beta}) \tag{6}$$

$$\lambda_{\min}(\widehat{\mathcal{K}}^{-1}\mathcal{K}) \;\; \geq \;\; \frac{1}{2}\left(2 + \alpha - 1/\beta - \sqrt{(2+\alpha-1/\beta)^2 - 4\alpha}\right). \tag{7}$$

The result of Theorem 1 allows one to use CG to solve the system $\widehat{\mathcal{K}}^{-1}\mathcal{K}x = \widehat{\mathcal{K}}^{-1}b$, which, at every step, minimizes the error in the norm defined by $\mathcal{B} = \mathcal{D}\widehat{\mathcal{K}}^{-1}\mathcal{K}$. The same result can be employed to give an estimate of the convergence rate, according to (3).

3 Refined spectral estimates

If a matrix $\mathcal{A}$ is SPD in the scalar product defined by $\mathcal{D}$, then it is diagonalizable with real and positive eigenvalues [8, Theorem 6.2]. Moreover, a closer look at the proof of this result reveals that the matrix X of eigenvectors for $\mathcal{A}$ can be chosen to be $\mathcal{D}$-orthogonal, i.e., $X^T\mathcal{D}X = I_N$, where I_N denotes the $N \times N$ identity matrix. We next give a refined result, where we do not restrict ourselves to the saddle point structure.

Proposition 2 *Let $\widehat{\mathcal{K}}, \mathcal{K} \in \mathbb{R}^{\mathcal{N}\times\mathcal{N}}$ be nonsingular symmetric matrices, such that $\mathcal{D} = \widehat{\mathcal{K}} - \mathcal{K} \geq 0$. We suppose that both $\widehat{\mathcal{K}}$ and $\mathcal{K}$ have n positive eigenvalues and $m = N - n$ negative ones. Then $\widehat{\mathcal{K}}^{-1}\mathcal{K}$ has real and positive eigenvalues. Moreover, if $\mathcal{D}$ is positive definite, $\widehat{\mathcal{K}}^{-1}\mathcal{K}$ is diagonalizable and has n eigenvalues strictly smaller than 1 and m eigenvalues strictly greater than 1. If, on the other hand, $\mathcal{D}$ has the eigenvalue 0 with multiplicity ℓ, then $\widehat{\mathcal{K}}^{-1}\mathcal{K}$ has ℓ eigenvectors associated with the eigenvalue 1.*

Proof: We first assume that $\mathcal{D}$ is positive definite. Then $\mathcal{D}$ defines an inner product on $\mathbb{R}^N$. Since $\mathcal{D}\widehat{\mathcal{K}}^{-1}\mathcal{K} = (\widehat{\mathcal{K}} - \mathcal{K})\widehat{\mathcal{K}}^{-1}\mathcal{K} = \mathcal{K} - \mathcal{K}\widehat{\mathcal{K}}^{-1}\mathcal{K}$ is symmetric, there exists a $\mathcal{D}$-orthogonal matrix X of eigenvectors for $\widehat{\mathcal{K}}^{-1}\mathcal{K}$. It results

$$I_N = X^T\mathcal{D}X = X^T(\widehat{\mathcal{K}} - \mathcal{K})X = X^T\widehat{\mathcal{K}}(I_N - \widehat{\mathcal{K}}^{-1}\mathcal{K})X = X^T\widehat{\mathcal{K}}X(I_N - \Lambda),$$

and hence $X^T\widehat{\mathcal{K}}X = (I_N - \Lambda)^{-1}$ and is thus diagonal. Since $\widehat{\mathcal{K}}$ has m negative and n positive eigenvalues, the Sylvester Law of Inertia ensures that $X^T\widehat{\mathcal{K}}X$ has m negative and n positive diagonal entries. Then Λ must have m eigenvalues greater than 1, and n smaller than 1. Similarly,

$$I_N = X^T(\widehat{\mathcal{K}} - \mathcal{K})X = X^T\mathcal{K}(\mathcal{K}^{-1}\widehat{\mathcal{K}} - I_N)X = X^T\mathcal{K}X(\Lambda^{-1} - I_N),$$

from which we deduce $X^T\mathcal{K}X = (\Lambda^{-1} - I_N)^{-1} = \Lambda(I_N - \Lambda)^{-1}$, and thus $X^T\mathcal{K}X$ is also diagonal. Moreover, this equation shows that Λ must have n eigenvalues lying in the interval $]0, 1[$ and m eigenvalues lying outside $[0, 1]$. Adding these conditions to the previous ones, we can conclude that Λ has n eigenvalues lying in $]0, 1[$ and m eigenvalues lying $]1, +\infty[$.

We now consider the case when $\mathcal{D}$ is positive semidefinite. We define $\widehat{\mathcal{K}}_\epsilon = \widehat{\mathcal{K}} + \epsilon I_N$ and $\mathcal{D}_\epsilon = \widehat{\mathcal{K}}_\epsilon - \mathcal{K}$ for $\epsilon > 0$. Since $\mathcal{D}_\epsilon$ is symmetric and positive definite, from the first part of the proof we deduce that $\widehat{\mathcal{K}}_\epsilon^{-1}\mathcal{K}$ has real and positive eigenvalues. Since $\widehat{\mathcal{K}}_\epsilon^{-1}\mathcal{K} \overset{\epsilon\to 0^+}{\longrightarrow} \widehat{\mathcal{K}}^{-1}\mathcal{K}$, for the continuity of the eigenvalues we can conclude that $\widehat{\mathcal{K}}^{-1}\mathcal{K}$ (which is nonsingular) has real and positive eigenvalues. Finally, from the relation $\mathcal{D} = \widehat{\mathcal{K}}(I - \widehat{\mathcal{K}}^{-1}\mathcal{K})$ one deduces that $\mathcal{D}v = 0$ if and only if $\widehat{\mathcal{K}}^{-1}\mathcal{K}v = v$. ♠

A saddle point matrix of the form (1) has n positive and m negative eigenvalues; the same holds for $\widehat{\mathcal{K}}$. Thus, Theorem 2 ensures that

$$\Lambda(\widehat{\mathcal{K}}^{-1}\mathcal{K}) \subseteq [\lambda_1, \lambda_n] \cup [\lambda_{n+1}, \lambda_{n+m}] \tag{8}$$

with $0 < \lambda_1 \le \lambda_n < 1 < \lambda_{n+1} \le \lambda_{n+m}$.

The result above shows that the spectral interval used in the convergence rate estimate is actually given by the union of two intervals, which do not include the value 1. We are interested in understanding how far these are from such value, and whether this may influence convergence. In the following we provide new bounds for λ_n and λ_{n+1}, and also a new lower bound for λ_1. We first need to define two new quantities:

$$a = \lambda_{\max}\left(\hat{A}^{-1}A\right), \quad s = \lambda_{\max}\left((B\hat{A}^{-1}B^T)^{-1}\hat{S}\right), \tag{9}$$

with $\alpha \le a < 1$ and $1/\beta < s < 1$ (cf. (4) and (5)). Since $\mathcal{D} = \hat{\mathcal{K}} - \mathcal{K}$,

$$\mathcal{K}z = \lambda\widehat{\mathcal{K}}z \quad \text{is equivalent to} \quad \mathcal{K}z = \mu\mathcal{D}z, \text{ with } \mu = \frac{\lambda}{1-\lambda}. \tag{10}$$

We have that $\lambda < 1$ if and only if $\mu > 0$, and $\lambda > 1$ if and only if $\mu < -1$.

Lemma 3 *Let a, s be as in (9). Let μ be an eigenvalue of $\mathcal{K}z = \mu\mathcal{D}z$. Then either $\mu_- \le \mu < -1$ or $0 < \mu \le \mu_+$, with*

$$\mu_\pm = \frac{1}{2}\left(\frac{a}{1-a} \pm \sqrt{\left(\frac{a}{1-a}\right)^2 + \frac{4}{(1-a)(1-s)}}\right).$$

Proof: Let $z = [x; y]$ be an eigenvector associated with μ. Then

$$Ax + B^T y = \mu(\hat{A} - A)x \quad Bx = \mu E y$$

with $E = B\hat{A}^{-1}B^T - \hat{S}$. Note that $x \neq 0$, otherwise the first equation would give $B^T y = 0$, and since B^T is full column rank, this would imply $y = 0$. From the second equation we can find y and substitute in the first one:

$$Ax + \frac{1}{\mu} B^T E^{-1} Bx = \mu(\hat{A} - A)x.$$

Reordering the terms and premultipling by μx^T we obtain

$$\mu^2 x^T \hat{A} x - (\mu^2 + \mu) x^T A x - x^T B^T E^{-1} B x = 0.$$

Since $\mu \in]-\infty, -1[\cup]0, +\infty[$, $\mu^2 + \mu > 0$. Moreover, $x^T A x \leq a x^T \hat{A} x$. Thus,

$$(1-a)\mu^2 x^T \hat{A} x - a\mu x^T \hat{A} x - x^T B^T E^{-1} B x \leq 0. \tag{11}$$

It holds

$$x^T B^T E^{-1} B x \leq \frac{1}{(1-s)} x^T B^T (B\hat{A}^{-1}B^T)^{-1} B x \leq \frac{1}{(1-s)} x^T \hat{A} x.$$

Using this inequality in (11) and dividing by $(1-a)x^T \hat{A} x$, we find:

$$\mu^2 - \mu \frac{a}{1-a} - \frac{1}{(1-a)(1-s)} \leq 0,$$

from which both extremes μ_- and μ_+ are computed. ♠

We emphasize that the bounds of Lemma 3 are sharp. Indeed, let us consider the case $n = 2$, $m = 1$, with

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 1-\epsilon_A \end{bmatrix}, \quad B^T = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \qquad \hat{A} = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}, \quad \hat{S} = 1 - \epsilon_S,$$

with $\epsilon_A < \frac{1}{2}$ and $\epsilon_S < 1$. Clearly, $a = \lambda_{\max}(\hat{A}^{-1}A) = 1 - \epsilon_A$ and $s = \lambda_{\max}((B\hat{A}^{-1}B^T)^{-1}\hat{S}) = 1 - \epsilon_S$. The eigenvalues of the matrix

$$\mathcal{D}^{-1/2}\mathcal{K}\mathcal{D}^{-1/2} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1-\epsilon_A}{\epsilon_A} & (\epsilon_S\epsilon_A)^{-1/2} \\ 0 & (\epsilon_S\epsilon_A)^{-1/2} & 0 \end{bmatrix}$$

satisfy the characteristic equation

$$(\mu - 1)\left(\mu^2 - \mu\frac{1-\epsilon_A}{\epsilon_A} - \frac{1}{\epsilon_A\epsilon_S}\right) = 0,$$

whose solutions are $\mu = 1$ and both bounds $\mu = \mu_-$, $\mu = \mu_+$.

Proposition 4 *Let λ_n and λ_{n+1} as in (8). Then*

$$\begin{aligned}\lambda_n &\leq 1-\frac{2(1-a)\sqrt{1-s}}{(2-a)\sqrt{1-s}+\sqrt{a^2(1-s)+4(1-a)}}\\ &\leq 1-\frac{(1-a)\sqrt{1-s}}{\sqrt{1-s}+\sqrt{1-a}}\end{aligned} \tag{12}$$

and

$$\begin{aligned}\lambda_{n+1} &\geq 1+\frac{(2-a)(1-s)+\sqrt{a^2(1-s)^2+4(1-a)(1-s)}}{2s}\\ &\geq 1+\frac{(1-s)\left(2a\sqrt{(1-s)(1-a)}+2-a\right)}{2s}\geq 1+\frac{1-s}{2s}.\end{aligned} \tag{13}$$

Proof: Using Lemma 3 we can show that

$$\lambda_n \leq \frac{\mu_+}{1+\mu_+}=1-\frac{1}{1+\mu_+}, \qquad \lambda_{n+1}\geq \frac{\mu_-}{1+\mu_-}=1-\frac{1}{1+\mu_-}.$$

Bounds (12) and (13) follow from simple, though tedious, calculations. ♠

Proposition 4 shows that the distance of the interior eigenvalues from one depends linearly on s, the eigenvalue of $(B\hat{A}^{-1}B^T)^{-1}\hat{S}$ closest to one, and nonlinearly on s and a. While it can be shown that the upper bound (6) is sharp, the lower bound (7) can still be improved. The approach we follow deviates from that originally proposed by Schöberl and Zulehner.

Proposition 5 *Let λ be an eigenvalue of $\widehat{\mathcal{K}}^{-1}\mathcal{K}$. Then*

$$\lambda \geq \min\left\{\alpha,\bar{\lambda}\right\} \text{ where } \bar{\lambda}=\frac{1}{2}\left(2\beta+\alpha-1-\sqrt{(2\beta+\alpha-1)^2-4\alpha\beta}\right). \tag{14}$$

Proof: We consider the generalized eigenvalue problem

$$\begin{bmatrix} A & B^T\\ B & 0\end{bmatrix}\begin{bmatrix} x\\ y\end{bmatrix}=\lambda\begin{bmatrix}\hat{A} & B^T\\ B & E\end{bmatrix}\begin{bmatrix} x\\ y\end{bmatrix}, \qquad \text{with} \quad E=B\hat{A}^{-1}B^T-\hat{S}>0.$$

We observe that $x\neq 0$, otherwise it would follow that $\lambda=0$. We find y from the second equation and substitute it in the first one, giving

$$\left(\lambda\hat{A}-A\right)x=\frac{(1-\lambda)^2}{\lambda}B^TE^{-1}Bx \tag{15}$$

Let us first consider the case $x \in \ker B$. Premultiplying the last equation by x^T we have

$$0 = x^T \left(\lambda \hat{A} - A\right) x \leq \left(\frac{\lambda}{\alpha} - 1\right) \|x\|^2$$

and then $\lambda \geq \alpha$. In the general case, we write $x = x_1 + x_2$, with $x_1 \in \ker B$ and $0 \neq x_2 \in (\ker B)^{\perp \hat{A}} := \left\{u \in \mathbb{R}^n \mid u^T \hat{A} v = 0 \ \forall v \in \ker B\right\}$. We premultiply equation (15) by x_1^T and by x_2^T, and obtain (note that $x_1^T \hat{A} x_2 = 0$)

$$x_1^T \left(\lambda \hat{A} - A\right) x_1 - x_1^T A x_2 = 0, \tag{16}$$

$$x_2^T \left(\lambda \hat{A} - A\right) x_2 - x_2^T A x_1 = \frac{(1-\lambda)^2}{\lambda} x_2^T B^T E^{-1} B x_2 \tag{17}$$

We first consider the right-hand side of equation (17). Using (5), we write $E \leq \frac{\beta-1}{\beta} B^T \hat{A}^{-1} B$. Hence, $x_2^T B E^{-1} B^T x_2 \geq c_\beta x_2^T B^T (B \hat{A}^{-1} B^T)^{-1} B x_2$, where $c_\beta = \beta/(\beta - 1)$. Since $\hat{A}^{-1} B^T (B \hat{A}^{-1} B^T)^{-1} B$ is a projector on $(\ker B)^{\perp \hat{A}}$, it also holds

$$x_2^T B^T (B \hat{A}^{-1} B^T)^{-1} B x_2 = x_2^T \hat{A} \left(\hat{A}^{-1} B^T (B \hat{A}^{-1} B^T)^{-1} B\right) x_2 = x_2^T \hat{A} x_2. \tag{18}$$

We now turn to the left-hand side of equation (17). We consider

$$\begin{aligned} -x_2^T A x_1 &= x_2^T (\hat{A} - A) x_1 \leq \left(x_1^T (\hat{A} - A) x_1\right)^{1/2} \left(x_2^T (\hat{A} - A) x_2\right)^{1/2} \\ &\leq \sqrt{1-\alpha} \left(x_1^T \hat{A} x_1\right)^{1/2} \left(x_2^T \hat{A} x_2\right)^{1/2}. \end{aligned} \tag{19}$$

From (16) and condition (4) we deduce that $-x_2^T A x_1 \geq (\alpha - \lambda)\, x_1^T \hat{A} x_1$. We suppose $\lambda < \alpha$ (if not, α is the sought after extreme). The last inequality, added to (19), shows that

$$\left(x_1^T \hat{A} x_1\right)^{1/2} \leq \frac{\sqrt{1-\alpha}}{\alpha - \lambda} \left(x_2^T \hat{A} x_2\right)^{1/2}$$

Note that this inequality also holds for $x_1 = 0$. Back to inequality (19) we can now conclude that $-x_2^T A x_1 \leq \frac{1-\alpha}{\alpha-\lambda} x_2^T \hat{A} x_2$, and thus

$$x_2^T \left(\lambda \hat{A} - A\right) x_2 - x_2^T A x_1 \leq \left(\lambda + \frac{1-\alpha}{\alpha - \lambda}\right) x_2^T \hat{A} x_2 = \frac{(1-\lambda)(\lambda - \alpha + 1)}{\alpha - \lambda} x_2^T \hat{A} x_2. \tag{20}$$

Collecting inequalities (18) and (20) we find that λ satisfies

$$\frac{\lambda - \alpha + 1}{\alpha - \lambda} \geq \frac{(1-\lambda)}{\lambda} c_\beta,$$

or, after some algebra, $\lambda^2 - (2\beta + \alpha - 1)\lambda + \alpha\beta \le 0$. We denote this polynomial by $p(\lambda)$. Since $p(0) = \alpha\beta > 0$ then the smallest positive root of $p(\lambda)$, which is precisely $\bar{\lambda}$, is a lower bound for λ, when $\bar{\lambda} < \alpha$. ♠

We next analyze the quality of $\bar{\lambda}$ by comparing it with the lower bound in Theorem 1, which will be denoted by $\bar{\lambda}_{SZ}$ in the following. We note that $\bar{\lambda}_{SZ}$ can too be seen as the smallest positive root of a second degree polynomial, i.e., $p_{SZ}(\lambda) = \lambda^2 - (2 + \alpha - 1/\beta)\lambda + \alpha$. We observe that $p(\lambda) - p_{SZ}(\lambda) = (\beta - 1)\,[(1/\beta - 2)\lambda + \alpha]$. Therefore, $p(\lambda) > p_{SZ}(\lambda)$ if and only if $\lambda < \alpha/(2 - 1/\beta)$. If we show that $\bar{\lambda}_{SZ} < \alpha/(2 - 1/\beta)$, then necessarily $\bar{\lambda}_{SZ} < \bar{\lambda}$, and thus $\bar{\lambda}$ is a sharper lower bound for the eigenvalues of $\widehat{\mathcal{K}}^{-1}\mathcal{K}$. Let $\rho = 2 - 1/\beta$. Our condition reads

$$\frac{1}{2}\left(\rho + \alpha - \sqrt{(\rho + \alpha)^2 - 4\alpha}\right) < \frac{\alpha}{\rho},$$

which is equivalent to

$$\left(\rho + \alpha - \frac{2\alpha}{\rho}\right) - \sqrt{\left(\rho + \alpha - \frac{2\alpha}{\rho}\right)^2 + 4\frac{\alpha^2}{\rho}\left(1 - \frac{1}{\rho}\right)} < 0,$$

which holds since $\rho > 1$, so that $(1 - 1/\rho) > 0$.

4 Numerical experiments

In this section we report on some of our numerical experiments to illustrate our theoretical results. All computations were performed using Matlab [10].

We considered the PDE-constraint optimal control problem described by Schöberl and Zulehner [13, Section 4], where the system (2) takes the form

$$\begin{bmatrix} M & 0 & K \\ 0 & \nu M & -M \\ K & -M & 0 \end{bmatrix} \begin{bmatrix} y \\ u \\ q \end{bmatrix} = \begin{bmatrix} f \\ 0 \\ 0 \end{bmatrix},$$

where M is the mass matrix, $K = M + K_0$, where K_0 is the stiffness matrix, and f is the discretized desired state. The actual data we used to construct K, M and f were taken from Thorne [14, Target 1 – 2D]. We first consider the second level of discretization, i.e., the dimension of $\mathcal{K}$ is 675×675.

To construct the preconditioners $\hat{A}$ and $\hat{S}$ we used Algebraic Multigrid [5], 3 Gauss-Seidel iterations, and a scaling as proposed by Schöberl and Zulehner [13, Sections 3 and 4]; These preconditioners depend on two parameters, σ and τ, whose choice is crucial. We set $\sigma = 0.9$, $\tau = 1.1 \cdot \frac{4}{3}$, $\nu = 10^{-4}$. Figure 1 shows the eigenvalues of $\widehat{\mathcal{K}}^{-1}\mathcal{K}$, together with the upper

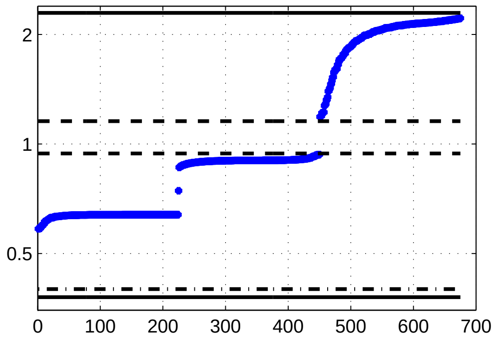

Figure 1: Eigenvalues of $\widehat{\mathcal{K}}^{-1}\mathcal{K}$ and bounds of the spectral interval

.

bound (6) and both interior bounds (12) and (13), which are represented in the plot by the solid and by the dashed lines. The estimates give a very realistic idea of the location of the true eigenvalues. For this example, we also observe that the bound (7) (lower solid line) is not sharp. Bound (14), represented by the dash-dotted line, slightly improves it. To continue with our analysis, we first recall that the two parameters a and s, which can be considered as a quality measure of the preconditioners $\hat{A}$ and $\hat{S}$ (and thus of $\widehat{\mathcal{K}}$), affect the distance of the eigenvalues of $\widehat{\mathcal{K}}^{-1}\mathcal{K}$ from 1, according to Proposition 4. More precisely, if a and s are close to 1, i.e. $\widehat{\mathcal{K}}$ is a good enough preconditioner for $\mathcal{K}$, the two spectral intervals $[\lambda_1, \lambda_n]$ and $[\lambda_{n+1}, \lambda_{n+m}]$ will be close to each other. Otherwise, if a and s are away from 1, the two intervals will be more distant. We also remark that, when $\hat{A}$ and $\hat{S}$ are constructed according to Schöberl and Zulehner [13], a is proportional to σ, and s is proportional to $1/\tau$. The parameter ν also seems to have an influence on the distance between the two intervals.

Figure 2 displays the convergence history of the method, in terms of the relative error $\mathcal{B}$-norm, namely $\|e_k\|_{\mathcal{B}}/\|e_0\|_{\mathcal{B}}$, along with the theoretical upper bound (3). We used the same model but with a finer discretization, yielding $\mathcal{K}$ of size 11907. We used $x^* = \mathtt{randn}(N, 1)$ as exact solution, and $x_0 = 0$ as initial guess. The left plot of Figure 2 considers the previous choice of values for σ, τ and ν. We see that the predicted behavior is in good agreement with the observed one. The right plot refers to parameter values $\sigma = 0.5$, $\tau = 2 \cdot \frac{4}{3}$, $\nu = 1$, and in this case we see that the bound (3) fails to predict the rate of convergence of the method. Note that the convergence curve does not suggest the occurrence of superlinear convergence behavior, for which different bounds would be more suitable [3]. The spectral intervals for the two

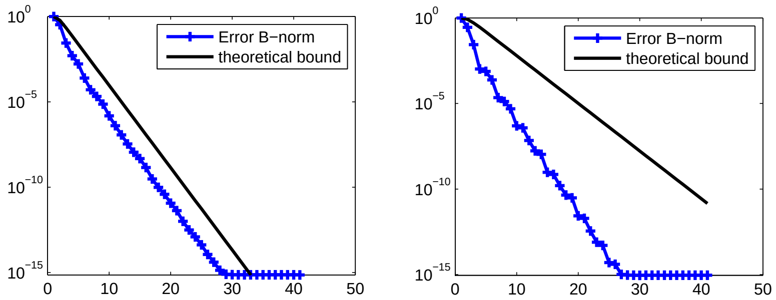

Figure 2: Convergence history and theoretical bound for $\sigma = 0.9$, $\tau = 1.1 \cdot \frac{4}{3}$, $\nu = 10^{-4}$ (left); and $\sigma = 0.5$, $\tau = 2 \cdot \frac{4}{3}$, $\nu = 1$ (right).

choices of parameter sets are $\Lambda(\widehat{\mathcal{K}}^{-1}\mathcal{K}) \subset [0.5821,\, 0.9468] \cup [1.1282,\, 2.2891]$ (left), and $\Lambda(\widehat{\mathcal{K}}^{-1}\mathcal{K}) \subset [0.4591,\, 0.7116] \cup [3.1391,\, 4.7747]$ (right), resp. In the latter case, a much bigger gap can be seen between the two intervals. We used the results of Proposition 4 to obtain the interior extremes of the intervals just showed, so that the true gap might be even larger. The theoretical bound is not representative of the actual convergence rate when the distance between the two intervals is relevant. We expect that bounds such as those described by Axelsson [2], tailored to the presence of more than one spectral interval, might be more descriptive. These considerations, and their applicability to saddle point linear systems will be more closely analyzed in future work.

References

[1] S. F. Ashby, T. A. Manteuffel, and P. E. Saylor, *A Taxonomy for Conjugate Gradients Methods*, SIAM J. Numer. Anal., 27:1542-1568, 1990.

[2] Owe Axelsson. Solution of linear systems of equations: iterative methods. In V.A. Barker, editor, *Sparse matrix techniques*, volume 572 of *Lecture notes in Mathematics*, pp 1–51. Springer, 1997.

[3] B. Beckermann and A. B. J. Kuijlaars. Superlinear Convergence of Conjugate Gradients. *SIAM J. Numer. Anal.*, v.39, pp.300–329, 2001.

[4] M. Benzi, G. H. Golub, and J. Liesen, *Numerical Solution of Saddle Point Problems*, Acta Numerica (2005), pp. 1-137.

[5] J. Boyle, M. D. Mihajlović, and J.A. Scott. HSL_MI20: an efficient AMG preconditioner for finite element problems in 3D. *Int. J. Numer. Meth. Engnrg,*, v.82, pp.64-98, 2010.

[6] V. Faber and T. A. Manteffel *Necessary and Sufficient Conditions for the Existence of a Conjugate Gradient Method*, SIAM J. Numer. Anal., v. 21, pp. 352-362, 1984.

[7] C. Keller, N. I. M. Gould, and A. J. Wathen. *Constraint preconditioning for indefinite linear systems*, SIAM J. Matrix Anal. Appl., v.21, pp. 1300–1317, 2000.

[8] W. D. Joubert and D. M. Young, *Necessary and Sufficient Conditions for the Simplification of Generalized Conjugate-Gradients Algorithms*, Linear Algebra Appl., v.88/89, pp. 449-485, 1987.

[9] L. Luksan and J. Vlcek. *Indefinitely preconditioned inexact Newton method for large sparse equality constrained non-linear programming problems*, Numer. Linear Algebra Appl., v.5, pp.219–247, 1998.

[10] The MathWorks, Inc. *MATLAB 7*, September 2004.

[11] J. Pestana and A. J. Wathen. Combination preconditioning of saddle point systems for positive definiteness. *Numer.Lin.Algebra Appl.*, to appear.

[12] M. Rozložník and V. Simoncini. *Krylov subspace methods for saddle point problem with indefinite preconditioning*, SIAM J. Matrix Anal. Appl., v.24, pp.368–391, 2002.

[13] J. Schöberl and W. Zulehner, *Symmetric Indefinite Preconditioners for Saddle Point Problems with Applications to PDE-Constrained Optimization Problems*, SIAM J. Matrix Anal. Appl., 29:752-773, 2007.

[14] H. S. Thorne, *Distributed Control and Constraint Preconditioners*, Rutherford Appleton Lab. Tech. Rep.RAL-TR-2010-016, 2010.

Author addresses

1. **Mattia Tani**, Dipartimento di Matematica, Università di Bologna, Piazza di Porta San Donato 5, Bologna, Italia
   `mailto:mattia.tani2@unibo.it`

2. **Valeria Simoncini**, Dipartimento di Matematica, Università di Bologna, Piazza di Porta San Donato 5, Bologna, Italia
   `mailto:valeria.simoncini@unibo.it`