ISSN 2281-4299

Dipartimento di Ingegneria informatica
automatica e gestionale Antonio Ruberti

Sapienza
Università di Roma

An estimation of the condition number for a class of indefinite preconditioned matrices

Giovanni Fasano
Massimo Roma

Technical Report n. 1, 2015

An estimation of the condition number for a class of indefinite preconditioned matrices

Giovanni Fasano
fasano@unive.it
Dipartimento di Management
Università Ca'Foscari Venezia

Massimo Roma
roma@dis.uniroma1.it
Dipartimento di Ingegneria Informatica, Automatica e Gestionale "A. Ruberti"
SAPIENZA, Università di Roma

Abstract. We propose a class of preconditioners for symmetric linear systems arising from numerical analysis and nonconvex optimization frameworks. Our preconditioners are specifically suited for *large* indefinite linear systems and may be obtained as *by-product* of Krylov-subspace solvers, as well as by applying L-BFGS updates. Moreover, our proposal is also suited for the solution of a sequence of linear systems, say $Ax = b_i$ or $A_i x = b_i$, where respectively the right-hand side changes or the system matrix slightly changes, too. Each preconditioner in our class is identified by setting the values of a parameter and two scaling matrices, which are user-dependent, and may be chosen according to the structure of the problem in hand. We specifically focus here on studying the condition number of the preconditioned matrix, where the preconditioner belongs to our class.

Keywords: Preconditioners, large indefinite linear systems, large scale nonconvex optimization, Krylov-subspace methods.

1 Introduction

We study a class of preconditioners for the solution of the symmetric indefinite linear system

$$Ax = b, \qquad A \in \mathbb{R}^{n \times n}, \quad A = A^T,$$

where n is *large* and we do not assume any sparsity pattern for the system matrix A. The solution of large linear systems is sought in a variety of real applications and in different contexts. Moreover, the use of preconditioning is an essential issue to improve the efficiency of iterative solvers. Within different frameworks of complex systems, the solution of sequences of large linear systems often comes up, too. E.g., we often encounter sequences like

$$Ax = b_i \qquad \text{or} \qquad A_i x = b_i, \tag{1.1}$$

where $\{b_i\}$ are possibly arbitrary and the matrices $\{A_i\}$ are *slowly* varying with the index 'i'. Numerical Analysis and Optimization give plenty of frameworks where the solution of a sequence of large linear systems is sought. Truncated Newton methods in unconstrained optimization, KKT systems, interior point methods, and PDE-constrained optimization are just some examples. Similarly, several real applications, ranging from power systems networks to economic models and queuing systems, involve the solution of large linear systems.

Typically, up to one decade ago, the specialized literature was keen on privileging the use of direct methods when n was moderately small, in view to their reasonable cost, since $O(n^3)$ might be unaffordable for large n. However, we have more recently observed an increasing blurred use of techniques, in both sparse direct methods and iterative algorithms, in order to efficiently solve linear systems (see e.g. [5, 7]). Observe that for linear systems where the matrix A is block-diagonal or banded, which typically arise when solving discretized PDEs, specific solvers from the literature can be used [19], which require to include effective preconditioning strategies, too.

In this paper we focus on the use of iterative methods to solve linear systems: the iterative techniques are also used to provide sufficient information on the system matrix, in order to generate the preconditioners.

We propose a general class of preconditioners, which uses information collected by any Krylov-subspace method or possibly using L-BFGS updates, in order to capture the structural properties of the system matrix.

In particular, we iteratively construct our preconditioners by using (but not performing) a factorization of the system matrix (see, e.g. [11, 15, 26]), obtained as by product of Krylov-subspace methods. We show that we can partially control the condition number of the preconditioned matrix, by introducing some care when choosing specific parameters.

The basic idea of our approach draws its inspiration from *Approximate Inverse Preconditioners*, which have proved in general to be remarkably robust and efficient in practice [5, 6]. These methods claim that in principle, an approximate inverse of A should be computed and used as a preconditioner. Though in practice it might be difficult to ensure that the approximate inverse is sparse, suitable factorizations of matrix A can be fruitfully exploited, in order to build the approximate inverse preconditioner. In particular, a generalization of the Gram-Schmidt process can be used to provide a triangular factorization of A^{-1}, where the triangular matrices are in general dense. This is the basic idea of AINV preconditioner (see [5], Section 5.1.2).

In this paper we apply any Krylov-subspace method to implicitly generate a triangular factorization of A^{-1}. The latter is then used to build our preconditioners, namely the AINV$\mathcal{K}$ class, needing to store just a few vectors, without requiring any matrix storage and any product of matrices (see also [2]). As we collect information from Krylov-subspace methods, we assume that the entries of the system matrix are not stored at once and the necessary information is gained by simply using a routine, which computes the product of the system matrix times a vector. Note that, typically, the product of a matrix times a vector allows fast parallel computing, which is another possible advantage of our approach, in large scale settings.

AINV$\mathcal{K}$ can be naturally extended to the solution of a sequence of large linear systems (see also [3, 4]). Indeed, when sequences of systems are tackled, we generate the preconditioner $\mathcal{P}$, for the solution of the first linear system in the sequence, i.e. $Ax = b_1$ or $A_1x = b_1$. Then, we apply $\mathcal{P}$ for solving either $Ax = b_i$ or $A_ix = b_i$, $i = 2, 3, \ldots$ Thus, the cost of computing $\mathcal{P}$, for $i = 1$, is repaid by accelerating the solution for $i = 2, 3, \ldots$; a similar strategy was proposed in [21]. The latter approach might be strongly advantageous in numerical analysis and optimization frameworks, where the cost for computing the preconditioner is relatively small, with respect to solving each linear system in the sequence. Furthermore, when a Krylov-subspace method is adopted to compute the preconditioner, the full storage of system matrix is never required. On the other hand, the same Krylov-subspace method might be used also to compute the solution of the linear system (see also [24, 25]).

Unlike following the idea early developed in [22], where a full-memory quasi-Newton formula is adopted for the preconditioner, we show that a few iterations of any Krylov-subspace method can be used, in order to provide information for building our preconditioners. Even if the resulting matrices are not sparse, they allow to consider preconditioning also for large scale problems, by simply storing k vectors, with $k \ll n$.

AINV$\mathcal{K}$ retains great generality, since it may be applied also when the system matrix is indefinite. Further generality is also provided by AINV$\mathcal{K}$ through the dependency on one parameter and two scaling matrices, which are user-dependent. Finally, we recall that in place of Krylov-subspace methods, in principle also L-BFGS updates can be used to build our preconditioners, so that they may be easily embedded within different numerical frameworks.

The paper is organized as follows: Section 2 reports some preliminaries and Section 3 contains the definition of the AINV$\mathcal{K}$ class. In Section 4 we study more in depth the condition number of the preconditioned matrix $\mathcal{P}A$, where $\mathcal{P}$ belongs to the class AINV$\mathcal{K}$. Finally, Section 5 adds some conclusions.

As regards the notations, for a $n \times n$ real matrix A we denote by $\Lambda[A]$ the spectrum of A. I_k is the identity matrix of order k. With $C \succ 0$ we indicate that the matrix C is positive definite, $tr[C]$, $rk[C]$ and $det[C]$ are the *trace*, the *rank* and the *determinant* of C, respectively. Finally, while $\kappa(B)$ indicates the condition number of B, $\| \cdot \|$ denotes the Euclidean norm, $\bigoplus_{1 \le j \le m} C_j$ represents the direct sum of matrices $\{C_j\}$ (see e.g. [20]) and e_h is the $h-$th unit vector.

2 Preliminaries

In this section we first introduce some preliminaries, then we propose our class of preconditioners. Consider the *indefinite* linear system

$$Ax = b, \tag{2.1}$$

where $A \in \mathbb{R}^{n \times n}$ is *symmetric* and n is *large*. Assume that a Krylov-subspace method is used for the solution of (2.1), e.g. the Lanczos process (SYMMLQ, MINRES [23]) or the Conjugate Gradient (CG) method [15, 18] (but Planar-CG methods [10, 18] may be also an alternative choice). As well known, the Lanczos process and the CG are equivalent as long as $A \succ 0$, whereas the CG, though cheaper, in principle may not cope with the indefinite case.

2.1 The matrix factorization we use

With reference to the definition in [14, 27], we say that a symmetric indefinite matrix c is *factorizable* if the diagonal (or 2×2 block diagonal) matrix B and the unit lower triangular matrix L exist such that $C = LBL^T$.

In the next assumption we consider that a finite number of steps, say $h \ll n$, of the Krylov-subspace method adopted have been performed.

Assumption 2.1 [Factorization] *Let us consider any Krylov-subspace method to solve the symmetric linear system (2.1). Suppose at step h of the Krylov-subspace method, with $h \le n-1$, the matrices $R_h \in \mathbb{R}^{n\times h}$, $T_h \in \mathbb{R}^{h\times h}$ and the vector $u_{h+1} \in \mathbb{R}^n$ are generated, such that*

$$AR_h = R_h T_h + \rho_{h+1} u_{h+1} e_h^T, \qquad \rho_{h+1} \in \mathbb{R}. \tag{2.2}$$

Suppose that the matrix T_h is factorizable, so that there exists the following decomposition:

$$T_h = L_h B_h L_h^T, \tag{2.3}$$

where

$R_h = (u_1 \cdots u_h),\ \ u_i^T u_j = 0,\ \ \|u_i\| = 1,\ \ 1 \le i \ne j \le h+1,$

T_h is tridiagonal, irreducible, nonsingular, with eigenvalues not all coincident,

B_h is 1×1 or 2×2 block diagonal, L_h is unit lower bidiagonal.

To have a better intuition on the reason for which h steps of almost any Krylov-subspace method satisfies Assumption 2.1, we remark that they are essentially all based on the generation of orthogonal vectors (the Lanczos vectors or the residuals for CG-based methods), used to transform the system (2.1) into a tridiagonal one. Then, they substantially differ only in the way the resulting tridiagonal system is solved by factorization.

In particular, also observe that from (2.2) we have $T_h = R_h^T A R_h$, so that whenever $A \succ 0$ then $T_h \succ 0$. The Krylov-subspace method adopted may, in general, perform $m \ge h$ iterations, generating the orthonormal vectors $u_1, \ldots, u_m$. Then, we can set $R_h = (u_{\ell_1}, \ldots, u_{\ell_h})$, where $\{\ell_1, \ldots, \ell_h\} \subseteq \{1, \ldots, m\}$, and change relations (2.2)-(2.3) accordingly; i.e. Assumption 2.1 may hold selecting any h out of the m vectors (among $u_1, \ldots, u_m$) computed by the Krylov-subspace method, up to step m.

Observe also that from Assumption 2.1, if $\rho_{h+1} \ne 0$, the subspace $span\{u_1, \ldots, u_h\}$ is not invariant under the transformation by matrix A. This implies that here we consider a more general case with respect to [1].

3 Our class of preconditioners AINV$\mathcal{K}$

On the basis of Assumption 2.1, we can now define our preconditioners and show their properties. To this aim, suppose $T_h = L_h B_h L_h^T$ in (2.3), where $B_h = \bigoplus_{1\le j\le m}\{E_j^h\}$, and where either $E_j^h \in \mathbb{R}$ or $E_j^h \in \mathbb{R}^{2\times 2}$, for any $j \in \{1, \ldots, m\}$. Moreover, if $E_j^h \in \mathbb{R}^{2\times 2}$ for an index j, assume that we compute the eigen-decomposition

$$E_j^h = U_j^h D_j^h (U_j^h)^T, \tag{3.1}$$

with $D_j^h = diag\{d1_j^h; d2_j^h\}$ and $(U_j^h)^T U_j^h = U_j^h (U_j^h)^T = I_2$. On the other hand, if $E_j^h \in \mathbb{R}$ for an index j, for the sake of notation we again assume that (3.1) holds, setting

$$D_j^h \equiv d1_j^h \equiv E_j^h \qquad \text{and} \qquad U_j^h = 1.$$

Then, we have (for the sake of simplicity and without loss of generality we assume $D_j^h \in \mathbb{R}^{2\times 2}$, for any $j \in \{1, \ldots, m\}$)

$$B_h = \bigoplus_{1\le j\le m} \left\{E_j^h\right\} = \bigoplus_{1\le j\le m} \left\{ U_j^h \begin{pmatrix} d1_j^h & 0 \\ 0 & d2_j^h \end{pmatrix} (U_j^h)^T \right\}$$

and we define (see also [13]) the matrix

$$|B_h| \stackrel{\text{def}}{=} \bigoplus_{1\le j\le m} \left\{ U_j^h \begin{pmatrix} |d1_j^h| & 0 \\ 0 & |d2_j^h| \end{pmatrix} (U_j^h)^T \right\}.$$

Moreover, we have equivalently

$$\begin{aligned} |B_h| &= \left[\bigoplus_{1\le j\le m} \left\{ U_j^h \right\} \right] \cdot \left[\bigoplus_{1\le j\le m} \left\{ \begin{pmatrix} |d1_j^h| & 0 \\ 0 & |d2_j^h| \end{pmatrix} \right\} \right] \cdot \left[\bigoplus_{1\le j\le m} \left\{ (U_j^h)^T \right\} \right] \\ &= U_h \cdot \mathcal{D}_h \cdot (U_h)^T, \end{aligned} \tag{3.2}$$

where

$$U_h = \bigoplus_{1\le j\le m} \left\{ U_j^h \right\},$$

$$\mathcal{D}_h = \bigoplus_{1\le j\le m} \left\{ \begin{pmatrix} |d1_j^h| & 0 \\ 0 & |d2_j^h| \end{pmatrix} \right\},$$

and we also define

$$|T_h| \stackrel{\text{def}}{=} L_h |B_h| L_h^T.$$

Observe that of course, by the definition of $|B_h|$, we have $|T_h| = T_h$ in case T_h is positive definite. Furthermore, it is easily seen that $|T_h|$ is positive definite, for any h, and $|T_h|^{-1} T_h^2 |T_h|^{-1} = I_h$ whenever $T_h \succ 0$. As a consequence, by (2.3) we have $T_h |T_h|^{-1} = (|T_h|^{-1} T_h)^T = L_h \hat{I}_h L_h^{-1}$, where

$$\hat{I}_h = B_h |B_h|^{-1} \tag{3.3}$$

is at most 2×2 block-diagonal with all the eigenvalues in $\{-1, +1\}$.

We are now ready to introduce the following class of preconditioners, which depends on the parameter a and the matrices W_h, D

$$\begin{aligned} M_h^\sharp(a, W_h, D) &\stackrel{\text{def}}{=} D \left[I_n - (R_h \mid u_{h+1}) (R_h \mid u_{h+1})^T \right] D^T \\ &\quad + (R_h \mid Du_{h+1}) \left(\begin{array}{c|c} |T_h(W_h)| & a e_h \\ \hline a e_h^T & 1 \end{array} \right)^{-1} (R_h \mid Du_{h+1})^T, \quad h \le n-1, \end{aligned} \tag{3.4}$$

$$M_n^\sharp(a, W_n, D) \stackrel{\text{def}}{=} R_n |T_n(W_n)|^{-1} R_n^T. \tag{3.5}$$

In (3.4)-(3.5) $a \in \mathbb{R}$, $W_h \in \mathbb{R}^{h\times h}$ is diagonal positive definite and $D \in \mathbb{R}^{n\times n}$ is nonsingular. Finally, using (3.2) we also define the matrix

$$|T_h(W_h)| = L_h U_h (W_h \mathcal{D}_h) U_h^T L_h^T, \tag{3.6}$$

so that the matrix $W_h \mathcal{D}_h$ is diagonal.

When $D = I_n$, $a = 0$ and $W_h = I_h$, $M_h^\sharp(a, W_h, D)$ reduces to the preconditioner in [12].

4 Issues on the condition number of $M_h^\sharp(a, W_h, D)A$

In this section we want to estimate the condition number $\kappa(M_h^\sharp(a, W_h, D)A)$ of the unsymmetric matrix $M_h^\sharp(a, W_h, D)A$ (i.e. the preconditioned matrix A), where $M_h^\sharp(a, W_h, D)$ is computed as in (3.4)-(3.5). We immediately have the general formula

$$\begin{aligned} \kappa(M_h^\sharp(a, W_h, D)A) &\stackrel{\text{def}}{=} \|M_h^\sharp(a, W_h, D)A\| \cdot \|(M_h^\sharp(a, W_h, D)A)^{-1}\| \\ &= \|M_h^\sharp(a, W_h, D)A\| \cdot \|A^{-1}(M_h^\sharp(a, W_h, D))^{-1}\|. \end{aligned} \tag{4.1}$$

Observe that (4.1) provides an upper bound for the spectral condition number of $M_h^\sharp(a, W_h, D)A$ (see also formula (3.10) in [17]), which strongly affects the efficiency of preconditioning. Indeed (see also [20]), if $\lambda_1, \ldots, \lambda_n$ are the (nonzero) eigenvalues of $M_h^\sharp(a, W_h, D)A$, we have in general

$$\frac{\max_i |\lambda_i|}{\min_i |\lambda_i|} \leq \kappa(M_h^\sharp(a, W_h, D)A).$$

Moreover, in this section we consider the general case where A is indefinite. Now we prove the next technical lemma.

Lemma 4.1 *Let $C \in \mathbb{R}^{h \times h}$ be a symmetric and positive definite matrix. Let $0 < \omega_1 \leq \cdots \leq \omega_h$ be the ordered eigenvalues of C, with $\omega_1, \ldots, \omega_{h-1}$ not all coincident, and let $a \in \mathbb{R}$. Then, given the quantities*

$$\alpha = -(h-1)\omega_1 + tr(C) + 1,$$

$$\beta = \frac{det(C)\left[1 - a^2 e_h^T C^{-1} e_h\right]}{(\omega_h)^{h-1}},$$

we have

$$\alpha^2 - 4\beta > 0,$$

and

$$\frac{[tr(C) - (h-1)\omega_1]\,\omega_h^{h-1}}{det(C)} > 1.$$

Proof: By the definition of α and β, and since $C \succ 0$, the condition $\alpha^2 - 4\beta \geq 0$ is satisfied if and only if

$$(e_h^T C^{-1} e_h)^{-1}\left[1 - \frac{\alpha^2(\omega_h)^{h-1}}{4det(C)}\right] \leq a^2. \tag{4.2}$$

Now, observing that $\omega_1, \ldots, \omega_{h-1}$ are not all coincident, $\alpha > \omega_h + 1$ and for any $\omega_1 \geq 0$ we have $(\omega_1 + 1)^2 \geq 4\omega_1$, we obtain

$$\frac{\alpha^2(\omega_h)^{h-1}}{4det(C)} \geq \frac{\alpha^2}{4\omega_1} > \frac{(\omega_h+1)^2}{4\omega_1} \geq \frac{(\omega_1+1)^2}{4\omega_1} \geq 1, \tag{4.3}$$

so that (4.2) holds for any choice of a, which also implies that $\alpha^2 - 4\beta \geq 0$. Also observe that by (4.3) $\alpha^2(\omega_h)^{h-1}/[4det(C)] > 1$, so that (4.2) can never be satisfied as an equality, i.e. $\alpha^2 - 4\beta \neq 0$ for any value of the parameter a.

Finally, note that since $det(C) = \Pi_{i=1}^h \omega_i$ we have

$$\omega_h^{h-1} > \frac{det(C)}{tr(C) - (h-1)\omega_1}, \tag{4.4}$$

inasmuch as $\omega_1, \ldots, \omega_{h-1}$ are not all coincident, and

$$\frac{det(C)}{tr(C) - (h-1)\omega_1} \leq \frac{det(C)}{\omega_h} = \prod_{i=1}^{h-1} \omega_i < \omega_h^{h-1}.$$

As a consequence, we have the condition

$$\frac{[tr(C) - (h-1)\omega_1]\, \omega_h^{h-1}}{det(C)} > 1.$$

$\square$

We remark that from (3.4)-(3.5) and (4.1), in case $A \succ 0$, $a = 0$, $W_h = I_h$ and $D = I_n$, using the Krylov-subspace method in Assumption 2.1 we easily obtain $\kappa(T_h) \leq \kappa(A)$ (see for instance [15]). Thus, in the latter case relation (4.1) yields

$$\kappa\left(M_h^\sharp(a, W_h, D)A\right) = \kappa\left(M_h^\sharp(0, I_h, I_n)A\right) \leq \kappa(T_h)\kappa(A) \leq [\kappa(A)]^2. \tag{4.5}$$

However, the bound (4.5) seems rather poor and does not depend on the parameters in our class of preconditioners. On the contrary, in the following result we provide an estimation of the condition number $\kappa(M_h^\sharp(a, W_h, D)A)$ in (4.1), which depends on the parameter 'a', and the matrices 'W_h' and 'D' in (3.4). Note that for the sake of clarity (but with a little abuse of notation), in the sequel we directly indicate with $\mu_1, \ldots, \mu_h$ the eigenvalues of $|T_h(W_h)|$ and not the eigenvalues of T_h.

Proposition 4.2 [Condition Number] *Let us consider the matrix $M_h^\sharp(a, W_h, D)$ in (3.4)-(3.5), with $h \leq n-1$, where T_h satisfies Assumption 2.1. Let $\mu_1 \leq \cdots \leq \mu_h$ be the (ordered) eigenvalues of $|T_h(W_h)|$, where $\mu_1, \ldots, \mu_{h-1}$ are not all coincident. Then, if*

$$|a| \; < \; (e_h^T |T_h(W_h)|^{-1} e_h)^{-1/2} \tag{4.6}$$

we have

$$\kappa\left(M_h^\sharp(a, W_h, D)A\right) \; \leq \; \xi_h \cdot \kappa(N)^2 \cdot \kappa(A), \tag{4.7}$$

with

$$\xi_h = \frac{\max\left\{1, \frac{\gamma_h + (\gamma_h^2 - 4\sigma_h)^{1/2}}{2}\right\}}{\min\left\{1, \frac{\gamma_h - (\gamma_h^2 - 4\sigma_h)^{1/2}}{2}\right\}}, \tag{4.8}$$

$$N = [R_h \mid Du_{h+1} \mid DR_{n,h+1}],$$

and

$$\gamma_h \;=\; -(h-1)\mu_1 + tr(|T_h(W_h)|) + 1 \tag{4.9}$$

$$\sigma_h \;=\; \frac{det(|T_h(W_h)|)\left[1 - a^2 e_h^T |T_h(W_h)|^{-1} e_h\right]}{(\mu_h)^{h-1}}. \tag{4.10}$$

In particular, when $D = I_n$ in (3.4), then $\kappa(M_h^\sharp(a, W_h, I_n)A) \leq \xi_h \cdot \kappa(A)$.

Proof: Consider the matrix

$$\left(\begin{array}{c|c} |T_h(W_h)| & ae_h \\ \hline ae_h^T & 1 \end{array}\right), \tag{4.11}$$

which is positive definite as long as condition (4.6) is fulfilled. Indeed, observe that setting $\Delta_h = 1 - a^2 e_h^T |T_h(W_h)|^{-1} e_h$, by the identity

$$\left(\begin{array}{c|c} |T_h(W_h)| & ae_h \\ \hline ae_h^T & 1 \end{array}\right) = \left(\begin{array}{c|c} I_h & 0 \\ \hline ae_h^T |T_h(W_h)|^{-1} & 1 \end{array}\right) \left(\begin{array}{c|c} |T_h(W_h)| & 0 \\ \hline 0 & \Delta_h \end{array}\right) \left(\begin{array}{c|c} I_h & a|T_h(W_h)|^{-1} e_h \\ \hline 0 & 1 \end{array}\right)$$

we have

$$det \left(\begin{array}{c|c} |T_h(W_h)| & ae_h \\ \hline ae_h^T & 1 \end{array}\right) = det(|T_h(W_h)|)\Delta_h. \tag{4.12}$$

Now, let $\lambda_1 \leq \cdots \leq \lambda_{h+1}$ be the (ordered) eigenvalues of the matrix (4.11). By the Cauchy interlacing properties (Lemma 8.4.4 in [8]) between the sequences $\{\mu_j\}_{j=1,\dots,h}$ and $\{\lambda_i\}_{i=1,\dots,h+1}$ we have the relation

$$\lambda_1 \leq \mu_1 \leq \lambda_2 \leq \mu_2 \leq \cdots \leq \lambda_h \leq \mu_h \leq \lambda_{h+1}. \tag{4.13}$$

By (4.11), (4.12) and (4.13) we can immediately infer the following intermediate results:

i) $\mu_1 \leq \lambda_i \leq \mu_h, \qquad i = 2, \dots, h$

ii) $\sum_{i=1}^{h+1} \lambda_i = tr(|T_h(W_h)|) + 1,$

iii) $\prod_{i=1}^{h+1} \lambda_i = det(|T_h(W_h)|)\Delta_h.$

From *i)* we deduce that

$$(h-1)\mu_1 \leq \sum_{i=2}^{h} \lambda_i \leq (h-1)\mu_h,$$

so that from *ii)*, *iii)*, (4.13) and recalling that the matrix (4.11) is positive definite, we have

$$\max\{0, -(h-1)\mu_h + tr(|T_h(W_h)|) + 1\} \leq \lambda_1 + \lambda_{h+1} \leq -(h-1)\mu_1 + tr(|T_h(W_h)|) + 1 \tag{4.14}$$

and

$$\frac{det(|T_h(W_h)|)\Delta_h}{(\mu_h)^{h-1}} \leq \lambda_1 \cdot \lambda_{h+1} \leq \frac{det(|T_h(W_h)|)\Delta_h}{(\mu_1)^{h-1}}. \tag{4.15}$$

From (4.14) and (4.15) (see Figure 4.1), in order to compute a lower [upper] *bound* $\tilde{\lambda}_1$ [$\tilde{\lambda}_{h+1}$] for the smallest [largest] eigenvalue of matrix (4.11), we have to solve the linear system (σ_h and γ_h are defined in (4.9) and (4.10))

$$\begin{cases} \tilde{\lambda}_1 + \tilde{\lambda}_{h+1} = \gamma_h \\ \tilde{\lambda}_1 \cdot \tilde{\lambda}_{h+1} = \sigma_h, \end{cases}$$

which yields

$$\begin{aligned} \tilde{\lambda}_1 &= \frac{\gamma_h - (\gamma_h^2 - 4\sigma_h)^{1/2}}{2} \\ \tilde{\lambda}_{h+1} &= \frac{\gamma_h + (\gamma_h^2 - 4\sigma_h)^{1/2}}{2}, \end{aligned} \tag{4.16}$$

provided that $\gamma_h^2 - 4\sigma_h \geq 0$. The latter condition directly holds from Lemma 4.1. Now, observe that setting $N = [R_h \mid Du_{h+1} \mid DR_{n,h+1}]$ (where N is nonsingular by hypothesis), for $h \leq n-1$ the preconditioners $M_h^\sharp(a, W_h, D)$ may be rewritten as

$$M_h^\sharp(a, W_h, D) = N \left[\begin{array}{c|c} \left(\begin{array}{c|c} |T_h(W_h)| & ae_h \\ \hline ae_h^T & 1 \end{array}\right)^{-1} & 0 \\ \hline 0 & I_{n-(h+1)} \end{array}\right] N^T, \qquad h \leq n-1. \tag{4.17}$$

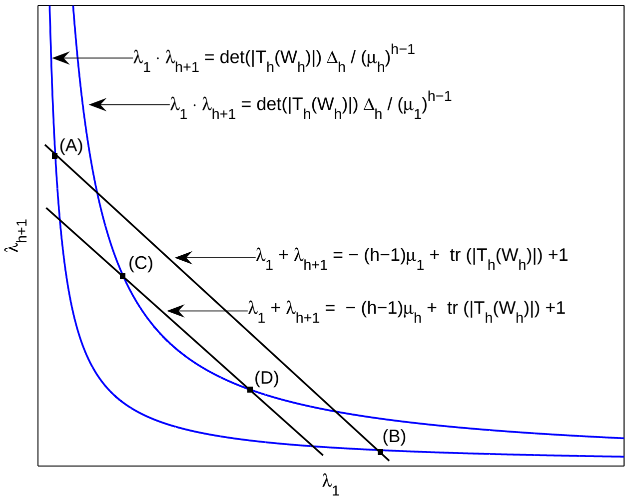

Figure 4.1: Relation between the eigenvalues λ_1 and λ_{h+1} of matrix (4.11).

As a consequence, setting

$$\mathcal{H}_h = \left[\begin{array}{c|c} \left(\begin{array}{c|c} |T_h(W_h)| & ae_h \\ \hline ae_h^T & 1 \end{array} \right) & 0 \\ \hline 0 & I_{n-(h+1)} \end{array} \right],$$

we have for the smallest [largest] eigenvalue of the symmetric matrices $\mathcal{H}_h$ and $\mathcal{H}_h^{-1}$ the expressions

$$\begin{cases} \lambda_m(\mathcal{H}_h) = \min\{1, \lambda_1\} \\ \\ \lambda_M(\mathcal{H}_h) = \max\{1, \lambda_{h+1}\} \end{cases}$$

$$\begin{cases} \lambda_m(\mathcal{H}_h^{-1}) = \frac{1}{\max\{1, \lambda_{h+1}\}} \\ \\ \lambda_M(\mathcal{H}_h^{-1}) = \frac{1}{\min\{1, \lambda_1\}}. \end{cases}$$

Thus, if $\lambda_m(A)$ $[\lambda_m(A^{-1})]$ and $\lambda_M(A)$ $[\lambda_M(A^{-1})]$ are the smallest [largest] eigenvalues of matrix A $[A^{-1}]$, from (4.17) we have

$$\|M_h^\sharp(a, W_h, D)A\| \;\leq\; \lambda_M(A) \cdot \|N\|^2 \cdot \lambda_M(\mathcal{H}_h^{-1}) \;=\; \lambda_M(A) \cdot \|N\|^2 \cdot \frac{1}{\min\{1, \lambda_1\}}$$

and

$$\begin{aligned} \|(M_h^\sharp(a, W_h, D)A)^{-1}\| &= \|A^{-1}(M_h^\sharp(a, W_h, D))^{-1}\| \;\leq\; \lambda_M(A^{-1}) \cdot \|N^{-1}\|^2 \cdot \lambda_M(\mathcal{H}_h) \\ &= \frac{1}{\lambda_m(A)} \cdot \|N^{-1}\|^2 \cdot \max\{1, \lambda_{h+1}\}, \end{aligned}$$

so that from (4.16)

$$\kappa\left(M_h^\sharp(a, W_h, D)A\right) \;=\; \|M_h^\sharp(a, W_h, D)A\| \cdot \|(M_h^\sharp(a, W_h, D)A)^{-1}\| \;\leq\; \frac{\max\left\{1, \tilde{\lambda}_{h+1}\right\}}{\min\left\{1, \tilde{\lambda}_1\right\}} \kappa(N)^2 \kappa(A), \tag{4.18}$$

which is relation (4.7). Finally, when $D = I_n$ in (3.4) then $\kappa(N) = 1$. $\square$

4.1 On the assessment of the bound (4.7)

In order to further address the bound (4.7) we can now prove the next proposition.

Proposition 4.3 *Let us consider the hypotheses of Proposition 4.2 and let the quantity ξ_h be defined in (4.8). Then, for any choice of the parameter a and the matrix W_h satisfying (4.6) we have*

$$\xi_h = \frac{\gamma_h + (\gamma_h^2 - 4\sigma_h)^{1/2}}{\gamma_h - (\gamma_h^2 - 4\sigma_h)^{1/2}}, \tag{4.19}$$

where γ_h and σ_h are defined in (4.9) and (4.10).

Proof: First observe that ξ_h in (4.8) is independent of the matrix D in (3.4)-(3.5). The proof now consists to analyze the following three cases:

1) $\gamma_h < 2$ (i.e. $1 < 1/[tr(|T_h(W_h)|) - (h-1)\mu_1]$)

2) $\gamma_h = 2$ (i.e. $1 = 1/[tr(|T_h(W_h)|) - (h-1)\mu_1]$)

3) $\gamma_h > 2$ (i.e. $1 > 1/[tr(|T_h(W_h)|) - (h-1)\mu_1]$).

In case 1) is satisfied, observe that the inequality

$$\frac{\gamma_h + (\gamma_h^2 - 4\sigma_h)^{1/2}}{2} < 1$$

cannot hold, since (consider that $\gamma_h - 2 < 0$ and see Lemma 4.1) it requires that

$$\gamma_h < 1 + \sigma_h \qquad \text{if and only if} \qquad a^2 < \left[1 - \frac{(\gamma_h - 1)\mu_h^{h-1}}{det(|T_h(W_h)|)}\right] \frac{1}{e_h^T |T_h(W_h)|^{-1} e_h}$$

which can hold only if

$$\frac{(\gamma_h - 1)\mu_h^{h-1}}{det(|T_h(W_h)|)} \leq 1.$$

However, the last inequality cannot hold because it is equivalent to

$$1 \geq \frac{[tr(|T_h(W_h)|) - (h-1)\mu_1]\mu_h^{h-1}}{det(|T_h(W_h)|)},$$

which cannot be satisfied from Lemma 4.1. Moreover, in case 1), also

$$\frac{\gamma_h - (\gamma_h^2 - 4\sigma_h)^{1/2}}{2} > 1$$

cannot hold, since $\gamma_h - 2 < 0$. Therefore, when $\gamma_h < 2$ relation (4.19) holds.

The case 2) is pretty similar to the case 1), so that again (4.19) follows almost immediately.

In case 3), the inequality

$$\frac{\gamma_h + (\gamma_h^2 - 4\sigma_h)^{1/2}}{2} < 1$$

cannot hold since it is equivalent to $(\gamma_h^2 - 4\sigma_h)^{1/2} < 2 - \gamma_h < 0$. Moreover, from Lemma 4.1 and considering that $\gamma_h - 2 > 0$, the condition

$$\frac{\gamma_h - (\gamma_h^2 - 4\sigma_h)^{1/2}}{2} > 1$$

can be satisfied if

$$\gamma_h < 1 + \sigma_h \qquad \text{if and only if} \qquad a^2 < \left[1 - \frac{(\gamma_h - 1)\mu_h^{h-1}}{det(|T_h(W_h)|)}\right] \frac{1}{e_h^T |T_h(W_h)|^{-1} e_h},$$

which holds only if

$$\frac{(\gamma_h - 1)\mu_h^{h-1}}{det(|T_h(W_h)|)} \leq 1.$$

However, since $\gamma_h - 1 = tr(|T_h(W_h)|) - (h-1)\mu_1$, the last inequality is again equivalent to

$$1 \geq \frac{[tr(|T_h(W_h)|) - (h-1)\mu_1]\mu_h^{h-1}}{det(|T_h(W_h)|)}$$

which cannot hold from Lemma 4.1. Thus relation (4.19) holds.

□

Lemma 4.4 *Consider the matrix $M_h^\sharp(a, W_h, D)$ in (3.4)-(3.5), with $h \leq n-1$. Let $\mu_1 \leq \cdots \leq \mu_h$ be the (ordered) eigenvalues of $|T_h(W_h)|$, with $\mu_1, \cdots, \mu_{h-1}$ not all coincident, and let the parameter a and the matrix W_h satisfy condition (4.6). Then, for any choice of the nonsingular matrix D in (3.4)*

- *the coefficient ξ_h in (4.19) increases when $|a| \to \rho$, with $\rho = (e_h^T |T_h(W_h)|^{-1} e_h)^{-1/2}$, and*

$$\lim_{|a| \uparrow \rho} \xi_h = +\infty$$

- *the coefficient ξ_h in (4.19) attains its minimum when $a = 0$, and in this case we have*

$$\xi_h \;=\; \frac{\gamma_h + \left(\gamma_h^2 - 4\frac{det(|T_h(W_h)|)}{(\mu_h)^{h-1}}\right)^{1/2}}{\gamma_h - \left(\gamma_h^2 - 4\frac{det(|T_h(W_h)|)}{(\mu_h)^{h-1}}\right)^{1/2}}. \tag{4.20}$$

Proof: Observe that $\lim_{|a| \uparrow \rho} \xi_h = +\infty$. Indeed, when $|a| \to \rho$ we have $\sigma_h \to 0$, so that $\gamma_h - (\gamma_h^2 - 4\sigma_h)^{1/2} \to 0$ and $\gamma_h + (\gamma_h^2 - 4\sigma_h)^{1/2} \to 2\gamma_h$, with $\gamma_h > 1$. Thus, since from Lemma 4.1 $\gamma_h^2 - 4\sigma_h \geq 0$, Proposition 4.3 ensures that ξ_h satisfies (4.19), so that ξ_h increases as $|a| \to \rho$, with $\lim_{|a| \uparrow \rho} \xi_h = +\infty$. Moreover, from (4.19) and since ξ_h is a continuous function of the parameter a (see also (4.6)), we have

$$\frac{\partial \xi_h}{\partial a} = \frac{\partial \xi_h}{\partial \sigma_h} \cdot \frac{\partial \sigma_h}{\partial a} = \frac{-4\gamma_h}{[\gamma_h - (\gamma_h^2 - 4\sigma_h)^{1/2}]^2 (\gamma_h^2 - 4\sigma_h)^{1/2}} \cdot \frac{-2a \cdot det(|T_h(W_h)|) e_h^T |T_h(W_h)|^{-1} e_h}{(\mu_h)^{h-1}},$$

so that from (4.6) we have $sgn\{\partial \xi_h / \partial a\} = sgn\{a\}$, which implies that ξ_h attains its minimum for $a = 0$.

Finally, by Lemma 4.1 we obtain $\gamma_h^2 - 4\sigma_h \geq 0$ for any choice of a satisfying (4.6), and when $a = 0$ it results $\sigma_h = det(|T_h(W_h)|)/(\mu_h)^{h-1}$. Thus, from Proposition 4.3 the value of ξ_h when $a = 0$ is given by (4.20). □

Remark 4.1 By (4.20) we observe that as expected, the matrix W_h affects the distribution of eigenvalues of $M_h^\sharp(a, W_h, D)A$, and also its condition number $\kappa(M_h^\sharp(a, W_h, D)A)$, when it is computed according with (4.1).

4.2 Guidelines to tighten the bound (4.7) on $\kappa\left(M_h^\sharp(a, W_h, D)A\right)$

This section is devoted to suggest possible values for the matrix W_h in relation (3.6). In particular, we want to show that by suitably setting W_h we can partially bound the spectral condition number of the unsymmetric matrix $M_h^\sharp(a, W_h, D)A$, when $D = I_n$. We recall indeed that the bound (4.5) does not seem so relevant, since it is independent of the parameters of our preconditioners. On this purpose, observe that when $a = 0$ then, by (4.19),

$$\xi_h = \frac{\gamma_h + (\gamma_h^2 - 4\sigma_h)^{1/2}}{\gamma_h - (\gamma_h^2 - 4\sigma_h)^{1/2}}.$$

Now, we want indicate possible values of W_h so that ξ_h is kept as low as possible, and the bound (4.7) is as tight as possible. In particular, let us set

$$\begin{cases} a = 0 \\ \\ W_h = \delta^2 I_h, \end{cases} \tag{4.21}$$

which correspond to a simplification for the choice of W_h. By replacing the latter positions in ξ_h, and indicating with $0 < \varepsilon_1 \le \varepsilon_2 \le \cdots \le \varepsilon_h$ the eigenvalues of $|T_h|$, we obtain

$$\begin{aligned} \xi_h &= \frac{\gamma_h + (\gamma_h^2 - 4\sigma_h)^{1/2}}{\gamma_h - (\gamma_h^2 - 4\sigma_h)^{1/2}} \\ &= \frac{\gamma_h + \left(\gamma_h^2 - 4\frac{\delta^{2h} det(|T_h|)}{(\delta^2)^{h-1}(\varepsilon_h)^{h-1}}\right)^{1/2}}{\gamma_h - \left(\gamma_h^2 - 4\frac{\delta^{2h} det(|T_h|)}{(\delta^2)^{h-1}(\varepsilon_h)^{h-1}}\right)^{1/2}} \\ &= \frac{\gamma_h + \left(\gamma_h^2 - 4\frac{\delta^{2} det(|T_h|)}{(\varepsilon_h)^{h-1}}\right)^{1/2}}{\gamma_h - \left(\gamma_h^2 - 4\frac{\delta^{2} det(|T_h|)}{(\varepsilon_h)^{h-1}}\right)^{1/2}}. \end{aligned} \tag{4.22}$$

Since the positions (4.21) yield

$$\gamma_h = 1 + \delta^2 \left[tr(|T_h|) - (h-1)\varepsilon_1\right]$$

$$\sigma_h = \frac{\delta^{2h} det(|T_h|)}{(\delta^2)^{h-1}\varepsilon_h^{h-1}} = \frac{\delta^2 det(|T_h|)}{\varepsilon_h^{h-1}},$$

setting

$$\begin{cases} \alpha = \dfrac{det(|T_h|)}{\varepsilon_h^{h-1}} > 0 \\ \\ \beta = [tr(|T_h|) - (h-1)\varepsilon_1] > 0 \\ \\ z = \delta^2, \end{cases} \tag{4.23}$$

we immediately obtain $\beta \ge \varepsilon_h \ge \alpha$ and the quantity ξ_h becomes

$$\begin{aligned} \xi_h &= \frac{(1+\beta z) + \left[(1+\beta z)^2 - 4\alpha z\right]^{1/2}}{(1+\beta z) - \left[(1+\beta z)^2 - 4\alpha z\right]^{1/2}} = \frac{\left\{(1+\beta z) + \left[(1+\beta z)^2 - 4\alpha z\right]^{1/2}\right\}^2}{(1+\beta z)^2 - \left[(1+\beta z)^2 - 4\alpha z\right]} \\ &= \frac{(1+\beta z)^2 - 2\alpha z + (1+\beta z)\left[(1+\beta z)^2 - 4\alpha z\right]^{1/2}}{2\alpha z} \\ &= -1 + \frac{(1+\beta z)^2 + (1+\beta z)\left[(1+\beta z)^2 - 4\alpha z\right]^{1/2}}{2\alpha z}. \end{aligned}$$

On this guideline, in order to have a tighter bound (4.7) on the condition number of the preconditioned matrix $M_h^\sharp(a, W_h, D)A$, we want to solve the problem

$$\min_{z\geq 0} \; \xi_h,$$

where possibly *min* stands for *global minimum*. To the latter purpose we observe that

$$\lim_{z\to 0} \xi_h = +\infty,$$

$$\lim_{z\to +\infty} \xi_h = +\infty,$$

which imply that since ξ_h is continuous in $(0, +\infty)$, there will be the values $m > 0$ and $M < +\infty$ such that by Weierstrass theorem ξ_h has at least an unconstrained minimum point for $z \in [m, M]$.
After some computation we obtain

$$\frac{d\xi_h}{dz} = \frac{(\beta^2 z^2 - 1)\left[(1+\beta z)^2 - 4\alpha z\right]^{1/2} + \beta^3 z^3 + \beta(\beta - 2\alpha)z^2 - (\beta - 2\alpha)z - 1}{2\alpha z^2 \left[(1+\beta z)^2 - 4\alpha z\right]^{1/2}}, \tag{4.24}$$

and since by Lemma 4.1 $\left[(1+\beta z)^2 - 4\alpha z\right] > 0$, for z sufficiently small the latter derivative is negative, while for z sufficiently large it is positive, meaning that the derivative has at least one zero in the interval $z \in [m, M]$. This also implies that we can adjust the values of δ, so that we can have a tighter bound (4.7) on the condition number $\kappa(M_h^\sharp(a, W_h, D)A)$. Note that for $z = 1$ (i.e. $\delta = 1$) we have for ξ_h the following bound

$$\xi_h = -1 + \frac{(1+\beta)^2 + (1+\beta)\left[(1+\beta)^2 - 4\alpha\right]^{1/2}}{2\alpha} \leq -1 + \frac{2(1+\beta)^2}{2\alpha} = \frac{(1+\beta)^2}{\alpha} - 1,$$

and $z = 1$ is not necessarily a zero of (4.24), i.e. $\delta = 1$ is not necessarily a stationary value of ξ_h.

Now we investigate more accurately the zeros of (4.24), i.e. the possible stationary values for ξ_h, recalling that since $|T_h| \succ 0$ the denominator of (4.24) is always positive as long as $\delta \neq 0$. Relation (4.24) is zero if and only if the following equality holds

$$(\beta^2 z^2 - 1)^2 \left[(1+\beta z)^2 - 4\alpha z\right] = \left[-\beta^3 z^3 + \beta(2\alpha - \beta)z^2 + (\beta - 2\alpha)z + 1\right]^2. \tag{4.25}$$

For the left and right hand side of the latter relation we have, respectively,

$$\beta^6 z^6 + 2\beta^4(\beta - 2\alpha)z^5 - \beta^4 z^4 - 4\beta^2(\beta - 2\alpha)z^3 - \beta^2 z^2 + 2(\beta - 2\alpha)z + 1$$

and

$$\beta^6 z^6 + 2\beta^4(\beta - 2\alpha)z^5 - (\beta^4 - 4\alpha^2\beta^2)z^4 - \left[4\beta^2(\beta - 2\alpha) + 8\alpha^2\beta\right] z^3 + (4\alpha^2 - \beta^2)z^2 + 2(\beta - 2\alpha)z + 1.$$

Thus, (4.25) holds if and only if

$$4\alpha^2\beta^2 z^4 - 8\alpha^2\beta z^3 + 4\alpha^2 z^2 = 0,$$

and recalling that $z \neq 0$ (i.e. $\delta \neq 0$) and $\alpha \neq 0$ (since $|T_h| \succ 0$), the latter equality implies

$$(\beta z - 1)^2 = 0,$$

i.e. the unique stationary value z^* of ξ_h is

$$z^* = \frac{1}{\beta}.$$

Moreover, even for small values of h (say $h \in [5, 10]$), we often have that $\beta > 2\alpha$. It is not difficult to realize that if the latter inequality holds, for $z > 1/\beta$ we obtain $\xi_h'(z) > 0$, while for $z < 1/\beta$ we obtain $\xi_h'(z) < 0$, so that ξ_h is strictly convex and $z^* = 1/\beta$ is the unique global minimum.

The Figures 4.2, 4.3, 4.4 and 4.5 show ξ_h vs. z, respectively using the following values of α and β

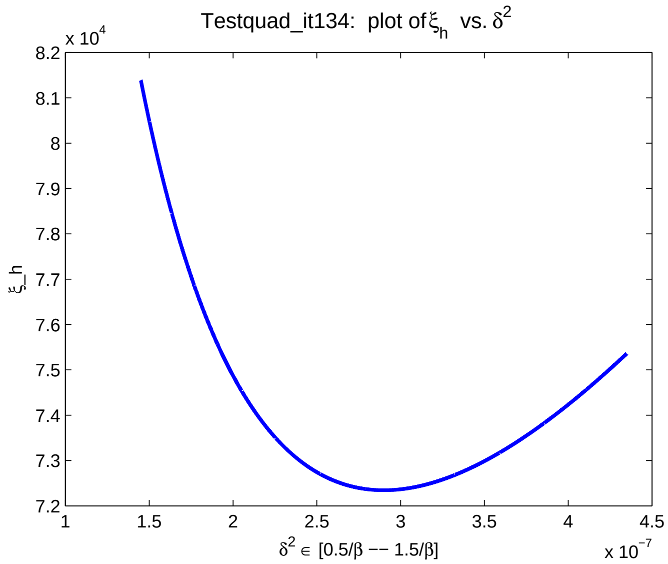

Figure 4.2: Plot of ξ_7 vs. δ^2 for the CUTEst function *Testquad*, where α and β are computed as in (4.23), at outer iteration 134 of the Truncated Newton method described in [12]. The values of δ are such that $\delta^2 \in \left[\frac{0.5}{\beta}, \frac{1.5}{\beta}\right]$.

$$\text{Figure 4.2:} \quad \begin{cases} \alpha = 190.6662 \\ \\ \beta = 3.4485 10^6 \end{cases} \quad \Longrightarrow \quad z^* = \frac{1}{\beta} = 0.2899 10^{-6}$$

$$\text{Figure 4.3:} \quad \begin{cases} \alpha = 2.5857 10^{-4} \\ \\ \beta = 6.7200 \end{cases} \quad \Longrightarrow \quad z^* = \frac{1}{\beta} = 0.1488$$

$$\text{Figure 4.4:} \quad \begin{cases} \alpha = 2.7556 10^{-4} \\ \\ \beta = 6.7510 \end{cases} \quad \Longrightarrow \quad z^* = \frac{1}{\beta} = 0.1481$$

$$\text{Figure 4.5:} \quad \begin{cases} \alpha = 3.8582 \\ \\ \beta = 1.2084 10^5 \end{cases} \quad \Longrightarrow \quad z^* = \frac{1}{\beta} = 0.8275 10^{-5}.$$

The function ξ_h in Figures 4.2, 4.3, 4.4 and 4.5 is obtained by selecting $h = 7$ (see also [12]) in Assumption 2.1, and applying the preconditioner (3.4)-(3.5), with the positions (4.21). The matrix A used in Assumption 2.1 is the Hessian matrix of the sample functions (see CUTEst collection [16]) Testquad, Dixmaani, Dixmaank, Curly20, obtained respectively at the outer iteration 134, 59, 17 and 200 of the preconditiond Truncated Newton method proposed in [12]. The number of unknowns n for the sample functions is respectively 1000, 1500, 1500 and 1000.
We remark that in the outer iterations considered for the samples, a standard preconditioned Conjugate Gradient method was used, to build the matrix T_h in (2.3). The latter choice was motivated since the sample problems above did not show regions of non convexity.

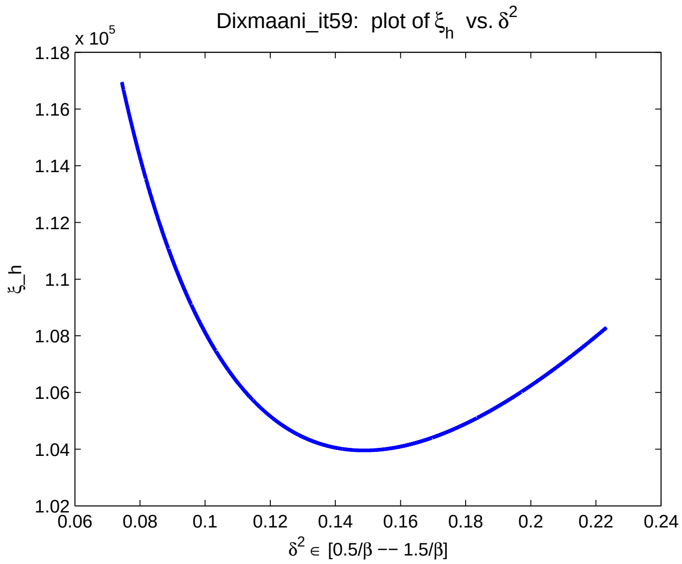

Figure 4.3: Plot of ξ_7 vs. δ^2 for the CUTEst function *Dixmaani*, where α and β are computed as in (4.23), at outer iteration 59 of the Truncated Newton method described in [12]. The values of δ are such that $\delta^2 \in \left[\frac{0.5}{\beta}, \frac{1.5}{\beta}\right]$.

5 Conclusions

We have given theoretical results for a new class of preconditioners, namely the AINV$\mathcal{K}$ class, which is parameter dependent. The preconditioners can be built by using any Krylov-subspace method for the indefinite linear system (2.1), as well as L-BFGS updates, provided that the general conditions (2.2)-(2.3) in Assumption 2.1 are satisfied. In particular, in many problems using a relatively small value of the index h, a significant information on the system matrix A can be captured.

On this guideline our proposal seems tailored also for those cases where a sequence of linear systems of the form

$$A_k x = b_k, \qquad k = 1, 2, \ldots \tag{5.1}$$

requires a solution (e.g., see also [9, 21] for details), where A_k slightly changes with the index k. In the latter case, the preconditioners $M_h^\sharp(a, W_h, D)$ in (3.4)-(3.5) can be computed applying the Krylov-subspace method to the first linear system $A_1 x = b_1$. Then, the resulting preconditioners can be used to efficiently solve (5.1) for $k = 2, 3, \ldots$

A full investigation was also included, where we studied the spectral condition number of the preconditioned matrix. On this guideline, observe that in the proof of Proposition 4.2, the considerations from Figure 4.1 and relations (4.15)-(4.16) are quite conservative; indeed, possibly an alternative estimation of the bounds $\tilde{\lambda}_1$ and $\tilde{\lambda}_{h+1}$ can be provided. In particular, the calculation of the four points A, B, C, D in Figure 4.1 could provide a more practical (say 'on average') bound for $\kappa(M_h^\sharp(a, W_h, D)A)$ in (4.18).

Acknowledgments. G.Fasano thanks the Italian Flagship Project *RITMARE*, coordinated by the Italian National Research Council and funded by the Italian Ministry of Education, within the National Research Program 2011-2013.

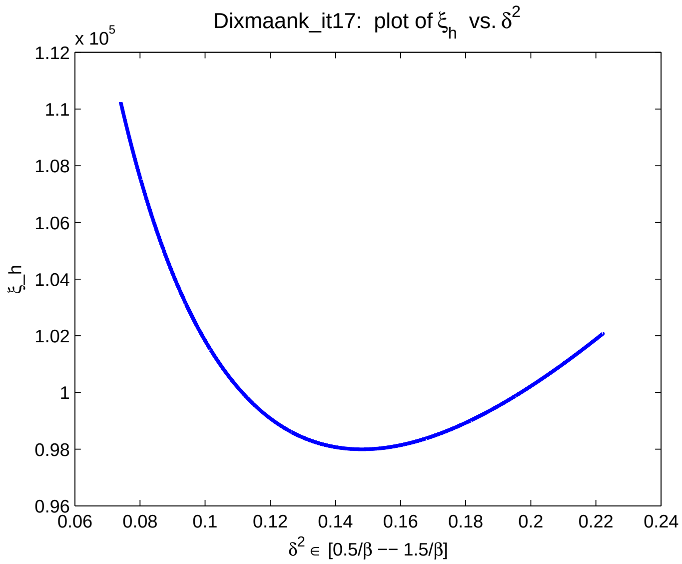

Figure 4.4: Plot of ξ_7 vs. δ^2 for the CUTEst function *Dixmaank*, where α and β are computed as in (4.23), at outer iteration 17 of the Truncated Newton method described in [12]. The values of δ are such that $\delta^2 \in \left[\frac{0.5}{\beta}, \frac{1.5}{\beta}\right]$.

References

[1] J. Baglama, D. Calvetti, G. Golub, and L. Reichel, *Adaptively preconditioned GMRES algorithms*, SIAM Journal on Scientific Computing, 20 (1998), pp. 243–269.

[2] S. Bellavia, J. Gondzio, and B. Morini, *A matrix–free preconditioner for sparse symmetric positive definite systems and least–squares problems*, SIAM Journal on Scientific Computing, 35 (2013), pp. A192–A211.

[3] S. Bellavia, V. D. Simone, D. D. Serafino, and B. Morini, *A preconditioning framework for sequences of diagonally modified linear systems arising in optimization*, SIAM Journal on Numerical Analysis, 50 (2013), pp. 3280–3302.

[4] S. Bellavia, V. D. Simone, D. D. Serafino, and B. Morini, *Updating constraint preconditioners for KKT systems in quadratic programming via low-rank corrections*, Optimization Online, (2013).

[5] M. Benzi, *Preconditioning techniques for large linear systems: a survey*, Journal of Computational Physics, 182 (2002), pp. 418–477.

[6] M. Benzi, J. Cullum, and M. Tuma, *Robust approximate inverse preconditioner for the conjugate gradient method*, SIAM Journal on Scientific Computing, 22 (2000), pp. 1318–1332.

[7] M. Benzi and M. Tuma, *A comparative study of sparse approximate inverse preconditioners*, Applied Numerical Mathematics, 30 (1999), pp. 305–340.

[8] D. Bernstein, *Matrix Mathematics: Theory, Facts, and Formulas – Second Edition*, Princeton University Press, Princeton, 2009.

[9] A. R. Conn, N. I. M. Gould, and P. L. Toint, *Trust–region methods*, MPS–SIAM Series on Optimization, Philadelphia, PA, 2000.

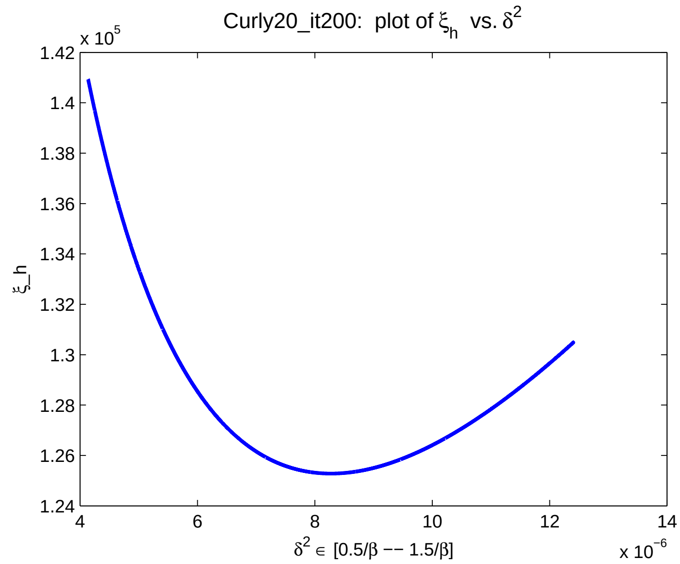

Figure 4.5: Plot of ξ_7 vs. δ^2 for the CUTEst function *Curly20*, where α and β are computed as in (4.23), at outer iteration 200 of the Truncated Newton method described in [12]. The values of δ are such that $\delta^2 \in \left[\frac{0.5}{\beta}, \frac{1.5}{\beta}\right]$.

[10] G. Fasano, *Planar–conjugate gradient algorithm for large–scale unconstrained optimization, Part 1: Theory*, Journal of Optimization Theory and Applications, 125 (2005), pp. 523–541.

[11] G. Fasano and M. Roma, *Iterative computation of negative curvature directions in large scale optimization*, Computational Optimization and Applications, 38 (2007), pp. 81–104.

[12] G. Fasano and M. Roma, *Preconditioning Newton–Krylov methods in non-convex large scale optimization*, Computational Optimization and Applications, 56 (2013), pp. 253–290.

[13] P. E. Gill, W. Murray, D. B. Ponceleon, and M. A. Saunders, *Preconditioners for indefinite systems arising in optimization*, SIAM Journal on Matrix Analysis and Applications, 13 (1992), pp. 292–311.

[14] P. E. Gill, M. A. Saunders, and J. R. Shinnerl, *On the stability of Cholesky factorization for symmetric quasidefinite systems*, SIAM Journal on Matrix Analysis and Applications, 17 (1996), pp. 35–46.

[15] G. Golub and C. Van Loan, *Matrix Computations*, The John Hopkins Press, Baltimore, 1996. Third Edition.

[16] N. I. M. Gould, D. Orban, and P. L. Toint, CUTEr *(and* sifdec*), a constrained and unconstrained testing environment, revised*, ACM Transaction on Mathematical Software, 29 (2003), pp. 373–394.

[17] S. Gratton, A. Sartenaer, and J. Tshimanga, *On a class of limited memory preconditioners for large scale linear systems with multiple right-hand sides*, SIAM Journal on Optimization, 21 (2011), pp. 912–935.

[18] M. Hestenes, *Conjugate Direction Methods in Optimization*, Springer Verlag, New York, 1980.

[19] N. Higham, *Accuracy and stability of numerical algorithms*, SIAM, Philadelphia, PA, 2002. Second Edition.

[20] R. Horn and C. Johnson, *Matrix Analysis*, Cambridge University Press, Cambridge, UK, 1999.

[21] J. Morales and J. Nocedal, *Automatic preconditioning by limited memory quasi–Newton updating*, SIAM Journal on Optimization, 10 (2000), pp. 1079–1096.

[22] D. O'Leary and A. Yeremin, *The linear algebra of block quasi-Newton algorithms*, Linear Algebra and its Applications, 212/213 (1994), pp. 153–168.

[23] C. Paige and M. Saunders, *Solution of sparse indefinite systems of linear equations*, SIAM Journal on Numerical Analysis, 12 (1975), pp. 617–629.

[24] Y. Saad, *Iterative Methods for Sparse Linear Systems – Second Edition*, SIAM, Philadelphia, PA, 2003.

[25] V. Simoncini and D. Szyld, *Recent computational developments in Krylov subspace methods for linear systems*, Numerical Linear Algebra with Applications, 14 (2007), pp. 1–59.

[26] J. Stoer, *Solution of large linear systems of equations by conjugate gradient type methods*, in Mathematical Programming. The State of the Art, A. Bachem, M.Grötschel, and B. Korte, eds., Berlin Heidelberg, 1983, Springer-Verlag, pp. 540–565.

[27] R. J. Vanderbei, *Symmetric quasi-definite matrices*, SIAM Journal of Optimization, 5 (1995), pp. 100–113.