

DIPARTIMENTO DI MATEMATICA
Complesso Universitario, Via Vivaldi 43 - 81100 Caserta



EFFICIENT PRECONDITIONER UPDATES

FOR SHIFTED LINEAR SYSTEMS

Stefania Bellavia¹, Valentina De Simone²,
Daniela di Serafino^{2,3} and Benedetta Morini¹

PREPRINT n. 5, impaginato nel mese di luglio 2010.

2010 Mathematics Subject Classification: primary - 65F08, 65F50; secondary - 65K05, 65M22

- 1 Dipartimento di Energetica "S. Stecco", Università degli Studi di Firenze, via C. Lombroso 6/17, 50134 Firenze, Italy, stefania.bellavia@unifi.it, benedetta.morini@unifi.it.
- 2 Dipartimento di Matematica, Seconda Università degli Studi di Napoli, via Vivaldi 43, 81100 Caserta, Italy, valentina.desimone@unina2.it, daniela.diserafino@unina2.it.
- 3 Istituto di Calcolo e Reti ad Alte Prestazioni, CNR, via P. Castellino 111, 80131 Napoli, Italy.

EFFICIENT PRECONDITIONER UPDATES FOR SHIFTED LINEAR SYSTEMS *

STEFANIA BELLAVIA [†], VALENTINA DE SIMONE [‡], DANIELA DI SERAFINO ^{‡§}, AND
BENEDETTA MORINI [†]

Abstract. We present a new technique for building effective and low cost preconditioners for sequences of shifted linear systems $(A + \alpha I)x_\alpha = b$, where A is symmetric positive definite and $\alpha > 0$. This technique updates a preconditioner for A , available in the form of an LDL^T factorization, by modifying only the nonzero entries of the L factor in such a way that the resulting preconditioner mimics the diagonal of the shifted matrix and reproduces its overall behaviour. The proposed approach is supported by a theoretical analysis as well as by numerical experiments, showing that it works efficiently for a broad range of values of α .

Key words. Shifted linear systems, preconditioner updates, incomplete LDL^T factorization.

AMS subject classifications. Primary: 65F08, 65F50. Secondary: 65K05, 65M22.

1. Introduction. We are concerned with the problem of building efficient preconditioners for the solution, by Krylov methods, of shifted linear systems of the form

$$(1.1) \quad (A + \alpha I)x_\alpha = b,$$

where $A \in \mathfrak{R}^{n \times n}$ is a symmetric positive definite matrix, $I \in \mathfrak{R}^{n \times n}$ is the identity matrix and $\alpha > 0$. Sequences of such linear systems arise in various fields, e.g. in trust-region and regularization techniques for nonlinear least-squares and other optimization problems, as well as in the application of implicit methods for the numerical solution of partial differential equations (PDEs).

The spectrum of the matrix $A + \alpha I$ may considerably change as α varies; thus, reusing the same preconditioner for different matrices is likely to work well for small values of α , but may be inappropriate as α increases. On the other hand, recomputing the preconditioner from scratch for each shifted matrix may result too expensive, either when α is small or when it is large and the spectral properties of $A + \alpha I$ result more favourable.

In this paper we are interested in building efficient and low-cost preconditioners for the shifted matrices, through a suitable update of a seed preconditioner, i.e. of a preconditioner built for the matrix A . Specifically, we assume that the seed preconditioner is available in the form of an LDL^T factorization, with L unit lower triangular and D diagonal positive definite. We note that in case the matrix A is not explicitly available, but can be accessed by evaluating matrix-vector products, it is possible

*Work supported in part by INdAM-GNCS, under grant *Progetti 2010 - Analisi e risoluzione iterativa di sistemi lineari di grandi dimensioni in problemi di ottimizzazione*, and by MIUR, under PRIN grants no. 20079PLLN7 - *Nonlinear Optimization, Variational Inequalities and Equilibrium Problems*, and no. 20074TZWNJ - *Development and analysis of mathematical models and of numerical methods for partial differential equations for applications to environmental and industrial problems*.

[†]Dipartimento di Energetica “S. Stecco”, Università degli Studi di Firenze, via C. Lombroso 6/17, 50134 Firenze, Italy, stefania.bellavia@unifi.it, benedetta.morini@unifi.it.

[‡]Dipartimento di Matematica, Seconda Università degli Studi di Napoli, via Vivaldi 43, 81100 Caserta, Italy, valentina.desimone@unina2.it, daniela.diserafino@unina2.it.

[§]Istituto di Calcolo e Reti ad Alte Prestazioni, CNR, via P. Castellino 111, 80131 Napoli, Italy.

to obtain an incomplete LDL^T factorization of A through matrix-free procedures, without explicitly forming A [8, 20].

Techniques for updating factorized preconditioners were proposed in [3, 6, 9, 15, 16, 23]. More precisely, sequences of systems of the form (1.1) are specifically addressed in [6, 23], while complex matrices which differ for a diagonal matrix are considered in [9]; finally, the algorithms presented in [3, 15, 16] apply to sequences of general nonsymmetric matrices. We outline these preconditioning techniques to better illustrate the contribution given by this paper.

The above techniques are based on the availability of an incomplete factorization of either A or A^{-1} . The algorithm proposed in [23] builds a preconditioner for $A + \alpha I$ by updating an incomplete Cholesky factorization of A . Since the differences between the entries of the Cholesky factor of A and those of $A + \alpha I$ depend on α , asymptotic expansions of order 0 or 1 of the latter entries in terms of α are used to approximately update the incomplete Cholesky factor of A ; furthermore, the recursiveness of the modifications undergone by the matrix entries during the Cholesky factorization is neglected. No theoretical results on the quality of the preconditioner are provided.

The algorithms in [3, 6, 9, 15, 16] share a common ground, since they update a factorized preconditioner $P = LDL^T$, or $P^{-1} = L^{-T}D^{-1}L^{-1}$, by leaving the factor L unchanged and modifying the factor D . The updating technique in [3, 6, 9] exploits the stabilized AINV preconditioner [5], which provides an approximation of A^{-1} :

$$A^{-1} \approx ZD^{-1}Z^T = L^{-T}D^{-1}L^{-1}.$$

It follows that

$$(A + \alpha I)^{-1} \approx Z(D + \alpha Z^T Z)^{-1}Z^T,$$

and a preconditioner for $A + \alpha I$ can be obtained as

$$(\bar{P}_\alpha)^{-1} = Z(D + \alpha E)^{-1}Z^T,$$

where E is a symmetric approximation of $Z^T Z$. This can be formed by extracting a certain number of upper diagonals from Z . If only the main diagonal of Z is taken, the matrix $D + \alpha E$ results to be diagonal and the updated preconditioner can be trivially applied. More generally, the solution of a banded system is required. This updating strategy is supported by a theoretical justification which ensures good spectral properties of the preconditioned matrices if the entries of Z decay fast enough from the main diagonal [6].

The approach presented in [15, 16] has been developed for the general case of preconditioning a sequence of nonsymmetric matrices which differ by a non-diagonal matrix, but we specialize it to the problems of interest in this paper. From

$$A \approx LDL^T,$$

it follows that

$$(1.2) \quad A + \alpha I \approx L(D + \alpha L^{-1}L^{-T})L^T.$$

Therefore, a preconditioner \hat{P}_α for $A + \alpha I$ can be defined as

$$\hat{P}_\alpha = LHL^T,$$

where H is an easily invertible approximation to $D + \alpha L^{-1}L^{-T}$ that preserves the symmetry. Obtaining such an approximation may be quite expensive as it involves the computation of entries of L^{-1} . In the general case of a nonsymmetric matrix for which an incomplete factorization LDU is available, it is tacitly assumed that L and U more or less approximate the identity matrix and one of the two matrices L^{-1} and U^{-1} , or both of them, are neglected [15]. Then, procedures for building a sparse approximation of the resulting mid factor are proposed and theoretically analyzed. However, in our case the previous assumption implies that $A + \alpha I$ is almost diagonal for any $\alpha \geq 0$, thus making the updating strategy little significant.

We propose a different update technique which modifies only the nonzero entries of the L factor of the incomplete LDL^T factorization of A , while leaving D unchanged. Specifically, the modifications applied to L force the preconditioner to mimic the diagonal of the shifted matrix and to reproduce its overall qualitative behaviour. Our procedure preserves the sparsity pattern of L and has a low computational overhead. Furthermore, it does not require tuning any algorithmic parameters, apart those possibly used in the construction of a preconditioner of A . The proposed technique is supported by a theoretical analysis of the updated preconditioners and by numerical experiments, showing that it has the ability of working well for a broad range of values of α .

The remainder of this paper is organized as follows. In Section 2 we provide some examples of problems requiring the solution of sequences of shifted linear systems, which provide motivations for our work. In Section 3 we present our procedure, while in Section 4 we analyze the quality of the resulting preconditioner, by providing an estimate of its distance from $A + \alpha I$, as well as results on the clustering of the eigenvalues of the preconditioned matrix for small and large values of α . In Section 5 we discuss numerical experiments showing the effectiveness of the proposed technique. Finally, in Section 6, we give some concluding remarks and outline future work.

In the following, $\|\cdot\|$ denotes the vector or matrix 2-norm. For any square matrix B , $diag(B)$ is the diagonal matrix with the same diagonal entries as B , $off(B)$ is the matrix with null diagonal and the same the off-diagonal entries of B , i.e. $off(B) = B - diag(B)$, $[B]_{ij}$ represents the (i, j) -th entry of B , and $\lambda_{min}(B)$ and $\lambda_{max}(B)$ denote the minimum and the maximum eigenvalue of B .

2. Motivating applications. We briefly describe some problems requiring the solution of sequences of shifted linear systems, which arise in numerical optimization and in the numerical solution of PDEs.

2.1. Trust-region and regularized subproblems in optimization. Let us consider the following problems which arise, e.g., as subproblems in unconstrained nonlinear least squares [13, Chapter 10]:

$$(2.1) \quad \min_p m(p) = \frac{1}{2} \|F + Jp\|_2^2, \quad \|p\|_2 \leq \Delta,$$

and

$$(2.2) \quad \min_p m(p) = \frac{1}{h} \|F + Jp\|_2^h + \frac{\sigma}{k} \|p\|_2^k,$$

where $F \in \mathbb{R}^m$, $J \in \mathbb{R}^{m \times n}$ is full rank, $m \geq n$, $\Delta > 0$, $\sigma > 0$, and h and k are integers such that $h > 0$ and $k > 1$. Specifically, the solution of (2.1) is required to obtain the trial step in trust-region methods [24], while the solution of (2.2) is needed

to compute the trial step in recent regularization approaches where $h = 1$, $k = 2$, or $h = 2$, $k = 3$ [4, 11, 25]. Subproblems of the previous type may also arise in numerical methods for constrained optimization, when computing a “normal step” with the aim of reducing the constraint infeasibility (see, e.g., [26, Chapter 18]).

The solution of (2.1) and (2.2) can be accomplished by solving a so-called secular equation, that is by finding the unique positive root λ_* of a secular function $\phi : \mathbb{R} \rightarrow \mathbb{R}$, whose form depends on the method used. For example, it is well known that any global minimizer p_* of (2.1) satisfies, for a certain $\lambda_* \geq 0$,

$$(2.3) \quad (J^T J + \lambda_* I)p_* = -J^T F, \quad \lambda_*(\|p_*\|_2 - \Delta) = 0$$

(see [24, Lemma 2.1]). Letting $p(\lambda) : \mathbb{R} \rightarrow \mathbb{R}^n$ be such that

$$(2.4) \quad (J^T J + \lambda I)p(\lambda) = -J^T F, \quad \lambda \geq 0,$$

it follows that either $\lambda_* = 0$, $p_* = -(J^T J)^{-1}J^T F$ and $\|p_*\|_2 \leq \Delta$, or $p_* = p(\lambda_*)$, where λ_* is the positive solution of the secular equation

$$(2.5) \quad \phi(\lambda) = \|p(\lambda)\|_2^2 - \Delta^2 = 0.$$

The Newton method or the secant method applied to (2.5) produce a sequence $\{\lambda_l\}$ of scalars and require the evaluation of ϕ at each iterate. This amounts to solve a sequence of shifted linear systems of the form (2.4).

Similarly, the characterization of the minimizer of (2.2) leads to the solution of a secular equation, which in turn requires the solution of a sequence of linear systems of the form (2.4) [4, 11, 25].

When the dimension n of the problem is large, the linear systems (2.4) can be solved by the LSQR method [27] and our preconditioning update technique. It is worth mentioning that, alternatively, LSQR can be used to find constrained minimizers of (2.1) and (2.2) over a sequence of expanding subspaces associated with the Lanczos process for reducing J to bi-diagonal form [11].

We conclude this section noting that shifted matrices of the form $J^T J + \lambda I$ arise also in the very recent method presented in [18]. This method solves the inequality constrained problem (2.1) over a sequence of evolving low-dimensional subspaces and employs preconditioners for a sequence of shifted matrices in the calculation of one of the vectors spanning the subspaces [18, §3.4].

2.2. Solution of PDEs by implicit methods. Let us consider the well-known heat equation:

$$(2.6) \quad \frac{\partial u}{\partial t}(t, x) - \beta \Delta u(t, x) = f(t, x),$$

where $t \geq 0$, $x \in [a, b] \times [c, d]$ and β is a constant, with Dirichlet boundary conditions and an initial condition $u(0, x) = u_0(x)$. By discretizing (2.6) in space by central differences on a Cartesian grid and in time by the backward Euler method, we obtain a sequence of linear systems of the form:

$$(2.7) \quad \left(I + \frac{\Delta t}{(\Delta x)^2} A \right) u^{k+1} = u^k + \Delta t f^k,$$

where A is a symmetric definite positive matrix, Δt and Δx are the grid spacing and the time stepsize, respectively, and k denotes the current time step. A simple

multiplication by $\alpha = (\Delta x)^2/\Delta t$, yields a sequence of systems of the form (1.1). Since the stepsize is usually chosen adaptively at each time step by estimating the accuracy in the computed solution, the value of α changes during the integration process.

More generally, sequences of systems of type (2.7) arise in the solution of initial-boundary value problems for diffusion equations, when implicit methods such as, e.g., BDF and SDIRK are applied [1]. The involved matrices are usually very large and sparse, thus requiring the application of Krylov solvers with suitable preconditioners.

3. A strategy for updating the preconditioner. In this section we introduce a preconditioner P_α for the matrix $A + \alpha I$, $\alpha > 0$, which is an update of an incomplete factorization of A . More precisely, we assume that an incomplete LDL^T factorization of A is available, where L is a unit lower triangular matrix and D is a diagonal matrix with positive entries, i.e. LDL^T is an incomplete square root - free Cholesky factorization of A .

We define P_α as follows:

$$(3.1) \quad P_\alpha = (L + G)D(L + G)^T,$$

where G is the sum of two matrices:

$$(3.2) \quad G = E + F,$$

with E diagonal and F strictly lower triangular. The matrix $E = \text{diag}(e_{11}, \dots, e_{nn})$ is defined by

$$(3.3) \quad e_{ii} = \sqrt{1 + \frac{\alpha}{d_{ii}}} - 1, \quad 1 \leq i \leq n,$$

while the nonzero entries of the matrix $F = (f_{ij})$ are given by

$$(3.4) \quad f_{ij} = \gamma_j l_{ij}, \quad 2 \leq i \leq n, \quad 1 \leq j < i$$

$$(3.5) \quad \gamma_j = \frac{1}{\sqrt{1 + \frac{\alpha}{d_{jj}}}} - 1 = \frac{1}{e_{jj} + 1} - 1.$$

Here and in the following the dependence of e_{ii} , f_{ij} and γ_j on α is neglected for simplicity. Furthermore, we extend the definition of γ_j in (3.5) to $j = n$. It is immediate to see that the entries of E are positive, γ_j is a decreasing function of e_{jj} , for $e_{jj} \in (0, +\infty)$, and $\gamma_j \in (-1, 0)$ for $\alpha > 0$. We note also that, for all i ,

$$(3.6) \quad \lim_{\alpha \rightarrow 0} e_{ii} = 0, \quad \lim_{\alpha \rightarrow +\infty} e_{ii} = +\infty,$$

$$(3.7) \quad \lim_{\alpha \rightarrow 0} \gamma_i = 0, \quad \lim_{\alpha \rightarrow +\infty} \gamma_i = -1.$$

In practice, the update of the LDL^T factorization of A consists in adding the term e_{ii} to the i -th unit diagonal entry of L and scaling the subdiagonal entries of the i -th columns of L with the scalar $\gamma_i + 1$. Remarkably, the sparsity pattern of the factors of P_α in (3.1) is the same of the factors of the incomplete LDL^T factorization of A .

The construction of the matrix G is inspired by the following observations. Suppose that LDL^T is the exact factorization of A . The shift by αI modifies only the diagonal entries of A ; furthermore, the larger is α the stronger is the effect of this

modification on the matrix A . Therefore, we would like to have a preconditioner \tilde{P}_α such that

$$[\tilde{P}_\alpha]_{ii} = a_{ii} + \alpha.$$

This can be achieved by modifying the diagonal of L , i.e. by considering $\tilde{P}_\alpha = (L + E)D(L + E)^T$, with E diagonal, and imposing that

$$(3.8) \quad \sum_{j=1}^{i-1} l_{ij}^2 d_{jj} + (l_{ii} + e_{ii})^2 d_{ii} = \sum_{j=1}^{i-1} l_{ij}^2 d_{jj} + l_{ii}^2 d_{ii} + \alpha, \quad i = 1, \dots, n.$$

From (3.8) and $l_{ii} = 1$ we obtain

$$(3.9) \quad (1 + e_{ii})^2 d_{ii} = d_{ii} + \alpha,$$

which is equivalent to (3.3). On the other hand, the off-diagonal entries of \tilde{P}_α take the following form

$$(3.10) \quad [\tilde{P}_\alpha]_{ij} = [\tilde{P}_\alpha]_{ji} \sum_{k=1}^{j-1} l_{ik} d_{kk} l_{jk} + l_{ij} d_{jj} (1 + e_{jj}), \quad i > j,$$

and, by using the second limit in (3.6), it results that

$$(3.11) \quad \lim_{\alpha \rightarrow +\infty} [\tilde{P}_\alpha]_{ij} = +\infty.$$

In other words, the off-diagonal entries of \tilde{P}_α increase with α , and hence they depart from the corresponding entries of $A + \alpha I$, which are independent of α . This means that \tilde{P}_α cannot be a good approximation of $A + \alpha I$ when α is large.

To overcome the previous drawback, we relax the requirement that the preconditioner has the same diagonal as $A + \alpha I$. By defining the preconditioner as in (3.1), we have that its off-diagonal entries are kept bounded, while the diagonal ones satisfy

$$(3.12) \quad [P_\alpha]_{ii} - (a_{ii} + \alpha) = \mu,$$

where

$$\mu = \sum_{j=1}^{i-1} (1 - (1 + \gamma_j)^2) l_{ij}^2 d_{jj}.$$

By using (3.6) and (3.7), it is immediate to verify that the magnitude of μ is small compared to $a_{ii} + \alpha$ as α tends to zero or infinity. Furthermore, the strictly lower triangle of $L + F$ tends to L when α tends to zero, and vanishes when α tends to infinity, thus preserving the qualitative behaviour of the off-diagonal part of $A + \alpha I$. We note also that the first row and column of P_α are equal to the corresponding row and column of $A + \alpha I$. A rigorous analysis of the behaviour of P_α , showing that it can be effective for a broad range of values of α , is carried out in the next section.

4. Analysis of the preconditioner. For the sake of simplicity, we assume that LDL^T is the exact factorization of A , i.e.

$$(4.1) \quad A = LDL^T;$$

however, the following analysis can be easily extended to the case that the factorization of A is incomplete.

We provide first an estimate of $\|P_\alpha - (A + \alpha I)\|$, i.e. of the accuracy of the preconditioner as an approximation of $A + \alpha I$, for all values of α , showing that the distance between P_α and $(A + \alpha I)$ is bounded independently of α . We study also the behaviour of the preconditioner as α tends to zero or infinity, demonstrating that the preconditioner behaves particularly well in these limiting cases. Finally, we analyse the spectrum of $P_\alpha^{-1}(A + \alpha I)$, showing that a certain number of eigenvalues may be equal to 1, depending on the rank of $P_\alpha - (A + \alpha I)$, and that all the eigenvalues are clustered in a neighbourhood of 1 when α is sufficiently small or large.

Letting

$$(4.2) \quad S = \text{diag}(\gamma_1, \dots, \gamma_n),$$

we can write F as

$$(4.3) \quad F = \text{off}(L)S.$$

Furthermore, it is easy to see that

$$(4.4) \quad S = (I + E)^{-1} - I, \quad \text{and} \quad E + SE + S = 0.$$

The following Lemma provides an expression of the difference between P_α and $A + \alpha I$.

LEMMA 4.1. *The preconditioner P_α given in (3.1) satisfies*

$$(4.5) \quad P_\alpha = A + \alpha I + Z,$$

where

$$(4.6) \quad Z = \text{diag}(B_1) + \text{off}(B_2),$$

and the diagonal entries of B_1 and the off-diagonal entries of B_2 are given by

$$(4.7) \quad [B_1]_{ii} = \sum_{k=1}^{i-1} l_{ik}^2 d_{kk} \gamma_k (\gamma_k + 2), \quad 1 \leq i \leq n,$$

$$(4.8) \quad [B_2]_{ij} = [B_2]_{ji} = \sum_{k=1}^{j-1} l_{ik} l_{jk} d_{kk} \gamma_k (\gamma_k + 2), \quad 2 \leq i \leq n, \quad 1 \leq j < i.$$

Proof. We start noting that

$$\begin{aligned} P_\alpha &= (L + G)D(L + G)^T \\ &= LDL^T + \alpha I + GDL^T + LDG^T + GDG^T - \alpha I \\ &= A + \alpha I + GDL^T + LDG^T + GDG^T - \alpha I. \end{aligned}$$

Then, letting

$$(4.9) \quad Z = GDL^T + LDG^T + GDG^T - \alpha I,$$

we have

$$P_\alpha = A + \alpha I + \text{diag}(Z) + \text{off}(Z).$$

We first prove that $\text{diag}(Z) = \text{diag}(B_1)$ with B_1 given by

$$(4.10) \quad B_1 = FDL^T + LDF^T + FDF^T.$$

From the definition of E in (3.3) it follows that

$$\text{diag}(E^2D + LDE + EDL^T - \alpha I) = 0;$$

then, using (3.2) we have

$$\begin{aligned} \text{diag}(Z) &= \text{diag}(EDL^T + FDL^T + LDE^T + LDF^T + E^2D \\ &\quad + EDF^T + FDE + FDF^T - \alpha I) \\ &= \text{diag}(FDL^T + LDF^T + EDF^T + FDE + FDF^T) \\ &= \text{diag}(FDL^T + LDF^T + FDF^T), \end{aligned}$$

where the last equality depends on the fact that FDE and EDF^T are strictly lower and upper triangular matrices, respectively. Hence, $\text{diag}(Z) = \text{diag}(B_1)$ and from (4.10) it is easy to see that the diagonal entries of B_1 have the form given in (4.7).

Now we show that $\text{off}(Z) = \text{off}(B_2)$ where B_2 is given by

$$(4.11) \quad B_2 = \text{off}(L)(2DS + DS^2)\text{off}(L)^T.$$

Using (3.2) we have

$$GDG^T = (E + F)D(E + F)^T = E^2D + FDE + EDF^T + FDF^T,$$

and, since E^2D is diagonal, (4.9) yields

$$\text{off}(Z) = \text{off}(GDL^T + LDG^T + FDE + EDF^T + FDF^T).$$

From (3.2) and (4.3) it follows that $G = E + \text{off}(L)S$, and, using $L = I + \text{off}(L)$, we obtain

$$\begin{aligned} GDL^T + LDG^T + FDE + EDF^T + FDF^T &= 2DE + (E + S + ES)D\text{off}(L)^T \\ &\quad + \text{off}(L)D(E + S + ES) + 2\text{off}(L)SD\text{off}(L)^T \\ &\quad + \text{off}(L)SDS\text{off}(L)^T. \end{aligned}$$

By taking into account that DE is diagonal and by (4.4), we get

$$\text{off}(Z) = \text{off}(\text{off}(L)(2DS + SDS)\text{off}(L)^T) = \text{off}(B_2).$$

From the previous expression it follows that B_2 is symmetric and its off-diagonal entries have the form given in (4.8), which completes the proof. \square

Note that the equalities (4.7) and (4.8) include also $[B_1]_{11} = 0$ and $[B_2]_{i1} = [B_2]_{1i} = 0$, $2 \leq i \leq n$, which, by (4.5), means that the first row and column of P_α are equal to those of $(A + \alpha I)$.

Next, we provide technical results.

LEMMA 4.2. *The matrices L and D in (4.1) satisfy*

$$(4.12) \quad \|\text{off}(L)D\text{off}(L)^T\| \leq 4\|LDL^T\|.$$

Also, if W is a positive semidefinite matrix, then

$$(4.13) \quad \|off(W)\| \leq \|W\|.$$

Proof. The matrix $R = LD^{\frac{1}{2}}$ is the Cholesky factor of A and hence satisfies $\|A\| = \|R\|^2$. Furthermore,

$$\|off(L)Doff(L)^T\| = \|off(L)D^{\frac{1}{2}}D^{\frac{1}{2}}off(L)^T\| = \|off(R)off(R)^T\| = \|off(R)\|^2.$$

Then, (4.12) follows from

$$(4.14) \quad \|off(R)\| = \|R - diag(R)\| \leq \|R\| + \|diag(R)\| \leq \|A\|^{\frac{1}{2}} + \max_i a_{ii}^{\frac{1}{2}} \leq 2\|A\|^{\frac{1}{2}}.$$

Now we prove (4.13). Since W is positive semidefinite, we have $W = off(W) + diag(W)$, where $off(W)$ is indefinite with null trace and $diag(W)$ is positive semidefinite. From [19, Theorem 8.1.5] we have

$$\lambda_{max}(W) \geq \lambda_{max}(off(W)) + \lambda_{min}(diag(W)),$$

which implies

$$\lambda_{max}(W) \geq \lambda_{max}(off(W)).$$

Furthermore, it is easy to see that $(\|W\|I - diag(W))$ is positive semidefinite and hence $\|W\|I + off(W) = \|W\|I - diag(W) + W$ is positive semidefinite too. Thus, $\lambda_{min}(\|W\|I + off(W)) \geq 0$, i.e.

$$|\lambda_{min}(off(W))| = -\lambda_{min}(off(W)) \leq \|W\|,$$

which completes the proof. \square

The quality of P_α as a preconditioner of $A + \alpha I$ is studied in the following theorems. Let us introduce the quantities

$$(4.15) \quad \omega_k = \gamma_k(\gamma_k + 2) = -\frac{\alpha}{\alpha + d_{kk}}, \quad k = 1, \dots, n,$$

where γ_k is defined in (3.5). For $k = 1, \dots, n$, ω_k is a monotonically decreasing function of α such that

$$(4.16) \quad \lim_{\alpha \rightarrow 0} \omega_k = 0, \quad \lim_{\alpha \rightarrow \infty} \omega_k = -1,$$

and

$$(4.17) \quad \max_k |\omega_k| = \frac{\alpha}{\alpha + \min_k d_{kk}}.$$

THEOREM 4.1. *For all $\alpha > 0$, the matrix Z given in (4.6) satisfies*

$$(4.18) \quad \|Z\| = \|P_\alpha - (A + \alpha I)\| \leq \max_k |\omega_k| (\max_i (a_{ii} - d_{ii}) + \|off(L)D^{\frac{1}{2}}\|^2)$$

$$(4.19) \quad \leq \max_k |\omega_k| (5\lambda_{max}(A) - \lambda_{min}(A))$$

and

$$(4.20) \quad \frac{\|P_\alpha - (A + \alpha I)\|}{\|A + \alpha I\|} \leq \max_k |\omega_k| \frac{5\lambda_{max}(A) - \lambda_{min}(A)}{\lambda_{max}(A) + \alpha}$$

Proof. Let us derive the inequality (4.18). By (4.7),

$$[B_1]_{ii} = \sum_{k=1}^{i-1} l_{ik}^2 d_{kk} \omega_k, \quad |l_{ik}^2 d_{kk} \omega_k| = l_{ik}^2 d_{kk} |\omega_k|;$$

then, since $a_{ii} - d_{ii} = \sum_{k=1}^{i-1} l_{ik}^2 d_{kk}$, we have

$$(4.21) \quad \|\text{diag}(B_1)\| \leq \max_k |\omega_k| \max_i (a_{ii} - d_{ii}).$$

Using (4.11) and (4.13) we obtain

$$\begin{aligned} \|\text{off}(B_2)\| &= \|\text{off}(\text{off}(L)D^{\frac{1}{2}}(2S + S^2)D^{\frac{1}{2}}\text{off}(L)^T)\| \\ &\leq \|\text{off}(L)D^{\frac{1}{2}}(2S + S^2)D^{\frac{1}{2}}\text{off}(L)^T\| \\ &\leq \|2S + S^2\| \|\text{off}(L)D^{\frac{1}{2}}\|^2. \end{aligned}$$

Hence, we can conclude that

$$(4.22) \quad \|\text{off}(B_2)\| \leq \max_k |\omega_k| \|\text{off}(L)D^{\frac{1}{2}}\|^2.$$

The definition (4.6) of Z , along with (4.21) and (4.22), yields (4.18). Furthermore, from the equality $d_{ii} = (L^{-1})_i A (L^{-1})_i^T$, where $(L^{-1})_i$ is the i -th row of L , and the properties of the Rayleigh quotient it follows that

$$\max_i (a_{ii} - d_{ii}) \leq \max_i a_{ii} - \min_i d_{ii} \leq \lambda_{max}(A) - \lambda_{min}(A).$$

This inequality and (4.14) provide (4.19) and (4.20). \square

The previous theorem shows that $\|P_\alpha - (A + \alpha I)\|$ remains bounded for all values of α . Moreover, from (4.17) it follows that the absolute distance $\|P_\alpha - (A + \alpha I)\|$ goes to zero with α , while the relative distance $\|P_\alpha - (A + \alpha I)\|/\|A + \alpha I\|$ tends to zero as α increases. A componentwise analysis of the behaviour of P_α for $\alpha \rightarrow 0$ and $\alpha \rightarrow +\infty$ is carried out in the theorem below.

THEOREM 4.2. *The matrix P_α satisfies*

$$(4.23) \quad \lim_{\alpha \rightarrow 0} [P_\alpha]_{ij} = a_{ij}, \quad 1 \leq i, j \leq n,$$

$$(4.24) \quad \lim_{\alpha \rightarrow +\infty} [P_\alpha]_{ii} = \lim_{\alpha \rightarrow +\infty} (\alpha + d_{ii}), \quad 1 \leq i \leq n,$$

$$(4.25) \quad \lim_{\alpha \rightarrow +\infty} [P_\alpha]_{ij} = l_{ij} d_{jj}, \quad 2 \leq i \leq n, \quad 1 \leq j < i.$$

Proof. Let us consider (4.5) along with (4.6). From (4.7) and (3.7) we get

$$(4.26) \quad \lim_{\alpha \rightarrow 0} [B_1]_{ii} = 0,$$

$$(4.27) \quad \lim_{\alpha \rightarrow +\infty} [B_1]_{ii} = \sum_{j=1}^{i-1} -d_{jj} l_{ij}^2 = -a_{ii} + d_{ii}.$$

Thus, equality (4.5) implies (4.23), for $i = j$, and (4.24).

From the expression of the off-diagonal entries of B_2 in (4.8), by using (3.7), for $i > j$ we get

$$(4.28) \quad \lim_{\alpha \rightarrow 0^+} [\text{off}(B_2)]_{ij} = 0,$$

$$(4.29) \quad \lim_{\alpha \rightarrow +\infty} [\text{off}(B_2)]_{ij} = -a_{ij} + l_{ij}d_{jj}.$$

The above equalities, along with (4.5), yield (4.23), for $i \neq j$, and (4.25). \square

The previous theorem shows that for $\alpha \rightarrow 0$ the preconditioner P_α behaves like $A + \alpha I$, since both approach the matrix A ; furthermore, for $\alpha \rightarrow +\infty$ the diagonal of P_α behaves like the diagonal of $A + \alpha I$, which, in turn, dominates over the remaining entries, while the off-diagonal part of P_α is kept bounded.

The analysis of the eigenvalues of $P_\alpha^{-1}(A + \alpha I)$ is performed in the following theorem.

THEOREM 4.3. *For all $\alpha > 0$, if the matrix Z has rank $n - k$, then k eigenvalues of $P_\alpha^{-1}(A + \alpha I)$ are equal to 1. Furthermore, for sufficiently small values of α , any eigenvalue λ of $P_\alpha^{-1}(A + \alpha I)$ satisfies*

$$(4.30) \quad |\lambda - 1| = \frac{O(\alpha)}{\lambda_{\min}(A) - O(\alpha)},$$

while, for sufficiently large values of α , it satisfies

$$(4.31) \quad |\lambda - 1| \leq \frac{\|A^{-1}Z\|}{1 + \frac{\alpha}{\lambda_{\max}(A)} - \|A^{-1}Z\|}.$$

Proof. By using (4.5) we get

$$P_\alpha^{-1}(A + \alpha I) = (A + \alpha I + Z)^{-1}(A + \alpha I) = (I + \alpha A^{-1} + A^{-1}Z)^{-1}(I + \alpha A^{-1}).$$

Then, if λ is an eigenvalue of $P_\alpha^{-1}(A + \alpha I)$, we have

$$(4.32) \quad (I + \alpha A^{-1})v = \lambda(I + \alpha A^{-1} + A^{-1}Z)v,$$

where v is an eigenvector corresponding to λ . Without loss of generality, we assume $\|v\| = 1$. From (4.32) it follows that $\lambda = 1$ if and only if $A^{-1}Zv = 0$, i.e. v belongs to the null space of Z . So, if $\text{rank}(Z) = n - k$, it follows that there are at least k unit eigenvalues.

Now suppose that $A^{-1}Zv \neq 0$ and multiply (4.32) by v^T , obtaining

$$(4.33) \quad 1 + \alpha v^T A^{-1}v = \lambda(1 + \alpha v^T A^{-1}v + v^T A^{-1}Zv),$$

and hence

$$(4.34) \quad |\lambda - 1| = \left| \frac{v^T A^{-1}Zv}{1 + \alpha v^T A^{-1}v + v^T A^{-1}Zv} \right| \leq \frac{\|A^{-1}Z\|}{|1 + \alpha v^T A^{-1}v + v^T A^{-1}Zv|},$$

where $1 + \alpha v^T A^{-1}v + v^T A^{-1}Zv \neq 0$ by (4.33). First, we focus on the case where α is small. We observe that

$$\begin{aligned} |\alpha v^T A^{-1}v + v^T A^{-1}Zv| &\leq \|\alpha A^{-1}\| + \|A^{-1}Z\| \\ &\leq \|A^{-1}\|(\alpha + \|Z\|) \\ &= \frac{\alpha + \|Z\|}{\lambda_{\min}(A)}. \end{aligned}$$

From (4.26) and (4.28), it easily follows that $Z = O(\alpha)$, therefore, for sufficiently small values of α we get

$$\begin{aligned} |1 + \alpha v^T A^{-1} v + v^T A^{-1} Z v| &\geq 1 - |\alpha v^T A^{-1} v + v^T A^{-1} Z v| \\ &\geq 1 - \frac{\alpha + \|Z\|}{\lambda_{\min}(A)} > 0. \end{aligned}$$

Then, inequality (4.34) yields

$$(4.35) \quad |\lambda - 1| \leq \frac{\|A^{-1}Z\|}{1 - \frac{\alpha + \|Z\|}{\lambda_{\min}(A)}},$$

which implies (4.30), since $\|Z\| = O(\alpha)$ and $\|A^{-1}\| = 1/\lambda_{\min}(A)$.

Let us consider now the case where α is large. From (4.18)-(4.19) we know that $\|Z\|$ is bounded for all $\alpha > 0$; therefore, for sufficiently large values of α , the denominator in (4.34) is bounded as follows:

$$\begin{aligned} |1 + \alpha v^T A^{-1} v + v^T A^{-1} Z v| &\geq |1 + \alpha v^T A^{-1} v| - |v^T A^{-1} Z v| \\ &\geq \lambda_{\min}(I + \alpha A^{-1}) - |v^T A^{-1} Z v| \\ &\geq 1 + \frac{\alpha}{\lambda_{\max}(A)} - \|A^{-1}Z\| > 0, \end{aligned}$$

where the second inequality follows from the fact that $1 + \alpha v^T A^{-1} v$ is a Rayleigh quotient of $I + \alpha A^{-1}$. Then, (4.31) follows from (4.34). \square

We note that Z has rank at most $n - 1$, since the first row and column of Z are null; then, $P_\alpha^{-1}(A + \alpha I)$ has at least one unit eigenvalue. We observe also that (4.30) means that for small values of α the eigenvalues of $P_\alpha^{-1}(A + \alpha I)$ are clustered in a neighbourhood of 1. Analogously, from (4.31) it follows that for large values of α the eigenvalues are again clustered in a neighbourhood of 1, since $\|Z\|$ is bounded independently of α .

The previous theorems hold under the assumption that we have the exact LDL^T factorization of A , but the results can be easily extended to an incomplete factorization, i.e. $LDL^T = A + B$, with $B \in \mathfrak{R}^{n \times n}$ symmetric. In particular, in this case a term depending on $\|B\|$ must be added to the upper bound on $\|P_\alpha - (A + \alpha I)\|$; furthermore, for large values of α the eigenvalues of $P_\alpha^{-1}(A + \alpha I)$ are still clustered around 1, while for small values of α the distance of the eigenvalues from 1 basically depends on the accuracy of the incomplete factorization, i.e. the more accurate such factorization is, the smaller the distance is expected to be.

5. Numerical experiments. Numerical experiments were carried out on several sequences of shifted linear systems to evaluate the behaviour of our preconditioner update technique, as well as to compare it with other commonly used preconditioning strategies.

The systems of each sequence were solved by using the Conjugate Gradient (CG) method with null initial guess. The CG method was terminated when the 2-norm of the residual was a factor 10^{-6} smaller than the initial residual, or when a maximum of 1000 iterations was reached; in the latter case a failure was declared if the residual did not achieve the required reduction. Three preconditioning strategies were coupled with CG: updating the incomplete LDL^T factorization of A as proposed (UP);

recomputing the incomplete LDL^T factorization for each value of α (RP); “freezing” the incomplete LDL^T factorization of A for all values of α (FP). The CG method was run also without preconditioning (NP).

All the experiments were performed using Matlab 7.10 on an Intel Core 2 Duo E8500 processor, with clock frequency of 3.16 GHz, 4 GB of RAM and 6 MB of cache memory. The preconditioned CG method was applied through the `pcg` function. The incomplete LDL^T factorizations were obtained using the `cholinc` function and recovering the L and D matrices from the Cholesky factor. For each sequence, the drop tolerance $drop$ used in the incomplete factorizations was fixed by trial on the system $Ax = b$. Specifically, a first attempt was made to solve this system using CG preconditioned by an incomplete LDL^T factorization with $drop = 10^{-1}$; if the computation of the incomplete factorization failed or CG did not terminate successfully within 1000 iterations, then $drop$ was reduced by a factor 10. This procedure was repeated until the preconditioned CG achieved the required accuracy in solving $Ax = b$. We remark that in the UP and FP strategies one incomplete LDL^T factorization was performed for each sequence, while in the RP strategy an LDL^T factorization was computed from scratch for each matrix $A + \alpha I$. The elapsed times required by the preconditioned CG were measured, in seconds, using the `tic` and `toc` Matlab commands.

A comparison among the four CG implementations was made using the performance profile proposed by Dolan and Moré [14]. Let $s_{T,A} \geq 0$ be a statistic corresponding to the solution of the test problem T by the algorithm A and suppose that the smaller this value is, the better the algorithm is considered. If s_T denotes the smallest value attained on the test T by one of the algorithms under analysis, the performance profile of the algorithm A , in logarithmic scale, is defined as

$$\pi(\chi) = \frac{\text{number of tests such that } \log_2(s_{T,A}/s_T) \leq \chi}{\text{number of tests}}, \quad \chi \geq 0.$$

5.1. Test sets. A first set of test sequences was built by shifting twenty sparse symmetric positive definite matrices from the University of Florida Sparse Matrix Collection [12]. In Table 5.1 we report the dimension n of each matrix, the density d_A of its lower triangular part, the drop tolerance $drop$ used to compute its incomplete LDL^T factorization, and the density d_L of the corresponding factor L . The density of a triangular matrix is defined as the ratio between the number of nonzero entries and the total number of entries in the triangular part. We remark that d_L is also the density of the preconditioners used by the UP and FP procedures for a whole sequence of shifted systems.

Each matrix was normalized dividing it by the maximum diagonal entry, to have a largest entry equal to 1. Eleven values of α were used for all the matrices, i.e. $\alpha = 10^{-i}$, $i = 0, 1, \dots, 5$, and $\alpha = 5 \cdot 10^{-i}$, $i = 1, \dots, 5$, for a total of 220 linear systems. We point out that, because of the normalization of the matrices, preconditioning was no longer beneficial as soon as α was of order 10^{-1} (see also [6]). The right-hand sides of the systems were built to obtain the solution vector $(1, \dots, 1)^T$.

A second set of test sequences was obtained by applying the Regularized Euclidean Residual (RER) algorithm described in [4] to three systems of nonlinear equations arising from the discretization of PDEs. Specifically, the RER algorithm was implemented using $h = 1$ and $k = 2$ in (2.2), fixing $\sigma = 1$ as initial parameter and updating σ along the iterations as in [4]. Eight sequences of shifted linear systems were picked up, that arose in the computation of trial steps. All the matrices of the sequences have dimension $n = 10^4$.

Matrix	n	d_A	$drop$	d_L	Matrix	n	d_A	$drop$	d_L
1138_bus	1138	4.0e-3	1.e-1	3.8e-3	cf <code>d1</code>	70656	3.8e-4	1.e-4	5.9e-3
apache1	80800	9.5e-5	1.e-1	6.8e-5	crystm03	24696	1.0e-3	1.e-1	3.1e-4
bcsstk13	2003	2.1e-2	1.e-3	4.1e-2	gyro_m	17361	1.2e-3	1.e-1	6.3e-4
bcsstk14	1806	2.0e-2	1.e-1	4.1e-3	jnlbrng1	40000	1.5e-4	1.e-1	1.0e-4
bcsstk15	3948	7.8e-3	1.e-1	1.5e-3	kuu	7102	6.9e-3	1.e-2	5.6e-3
bcsstk16	4884	1.2e-2	1.e-1	1.3e-3	nd3k	9000	4.1e-2	1.e-4	2.1e-2
bcsstk17	10974	3.7e-3	1.e-3	7.6e-3	s1rmq4m1	5489	8.9e-3	1.e-1	2.2e-3
bcsstk18	11948	1.1e-3	1.e-2	1.0e-3	s2rmq4m1	5489	8.9e-2	1.e-1	1.8e-3
bcsstk25	15439	1.1e-3	1.e-3	3.3e-3	thermal1	82654	9.6e-5	1.e-1	8.2e-5
bcsstk38	8032	5.6e-3	1.e-1	9.7e-4	wathen100	30401	5.4e-4	1.e-1	2.9e-4

TABLE 5.1

Characteristics of test problems from the University of Florida Sparse Matrix Collection.

Sequence	d_A	$drop$	d_L	m
bratu_1	1.3e-3	1.e-3	4.2e-3	4
bratu_2	1.3e-3	1.e-3	4.3e-3	4
fpm_1	1.3e-4	5.e-5	2.8e-2	8
fpm_2	1.3e-4	5.e-5	2.8e-2	8
pdexs_1	1.3e-3	1.e-3	4.1e-3	6
pdexs_2	1.3e-3	1.e-3	4.1e-3	5
pdex1_1	1.3e-3	1.e-4	2.7e-2	6
pdex1_2	1.3e-3	1.e-3	4.3e-3	6

TABLE 5.2

Characteristics of test problems arising in the application of the RER algorithm.

The first of the mentioned nonlinear systems comes from the discretization of the Bratu problem with $\lambda = 6.5$ [28]. We used a null initial guess and extracted the sequences of shifted linear systems at the first and the second nonlinear iteration of RER; we refer to these sequences as `bratu_1` and `bratu_2`. The second nonlinear system was obtained by discretizing a PDE problem modelling a flow in a porous medium as described in [28]. We applied the RER algorithm to the nonlinear system with the initial guess given in [28] and extracted the sequences arising in the first (`fpm_1`) and in the second (`fpm_2`) nonlinear iteration. The discretization of the PDE problem given in [17, §3.2.1] produced our third nonlinear system and gave rise to four sequences. The sequences `pdexs_1` and `pdexs_2` were generated at the first and second nonlinear iteration of RER with the standard initial guess given in [17] (with $k = 100$), while the sequences `pdex1_1` and `pdex1_2` arose at the first and the second nonlinear iteration of RER, starting from the vector of all ones.

In Table 5.2 we specify some details of the sequences of linear systems associated with the RER algorithm, i.e. the density d_A of the matrix $J^T J$, the drop tolerance $drop$ used in the incomplete LDL^T factorization of $J^T J$, the density d_L of the matrix L and the number m of shifted systems forming the sequence. We note that the drop tolerance used with `pdex_1` and `pdex_2` was obtained by dividing by 2, instead of 10, the previous drop tolerance leading to failure, i.e. 10^{-4} , to avoid an incomplete Cholesky factorization too close to the complete one. Concerning the computation of L and D , we are aware of approaches that do not need to form the matrix $J^T J$, such as the incomplete Cholesky factorization proposed in [7], and the Threshold Incomplete Givens Orthogonalization (TIGO) method presented in [2], which provides

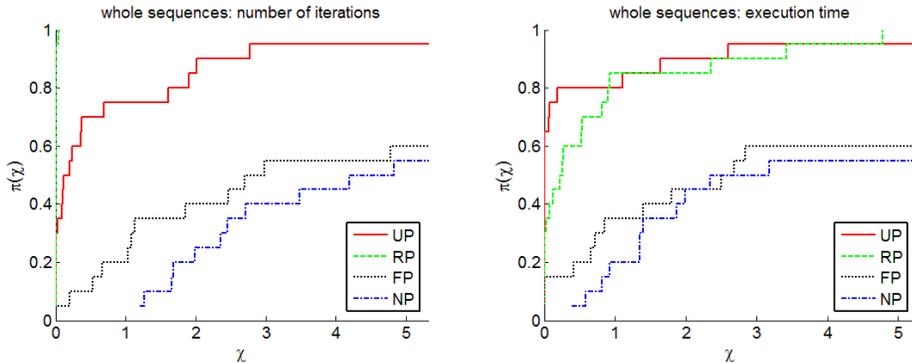


FIG. 5.1. Performance profiles on the first test set for the solution of the whole sequences: number of CG iterations (left) and execution times (right).

an incomplete QR factorization of J . On the other hand, avoiding the computation of $J^T J$ was not crucial for our experiments and we did not address it.

5.2. Analysis of the results. We compare the performance of the UP, RP, FP and NP strategies, focusing on their ability to solve a whole sequence of shifted linear systems. We analyse both the robustness and the efficiency of the four strategies.

The performance profiles in Figure 5.1 concern the total number of CG iterations and the execution times required to solve the whole sequences obtained from the set of matrices in Table 5.1. A failure is declared if the preconditioned CG fails in solving a system of the sequence. As expected, CG with UP requires more iterations than CG with RP, but it is the fastest among all the solvers. Specifically, UP is the most efficient in terms of execution time for 65% of the sequences. On the other hand, it fails in solving the sequence corresponding to the matrix `bcsstk17`, since in this case it fails for $\alpha = 10^{-5}$ and $\alpha = 5 \cdot 10^{-5}$. With the same matrix, FP fails in solving all the systems but the one corresponding to $\alpha = 10^{-5}$; it also fails on all the systems associated with `bcsstk18`, `bcsstk25` and `bcsstk38`. Furthermore, the performance of FP typically deteriorates as α increases, as confirmed by the iterations and timings profiles. More insight into the robustness of the various strategies can be get by looking at the number of successful solutions over the 20 sequences, which is 20 for RP, 19 for UP, 12 for FP and 11 for NP.

To better understand the behaviour of our preconditioning strategy versus the other ones, we analyse also the performance profiles of the CG iterations and the execution times in solving the shifted systems corresponding to three representative values of α , i.e. $\alpha = 10^{-5}$, $5 \cdot 10^{-3}$, 1 (Figures 5.2–5.4).

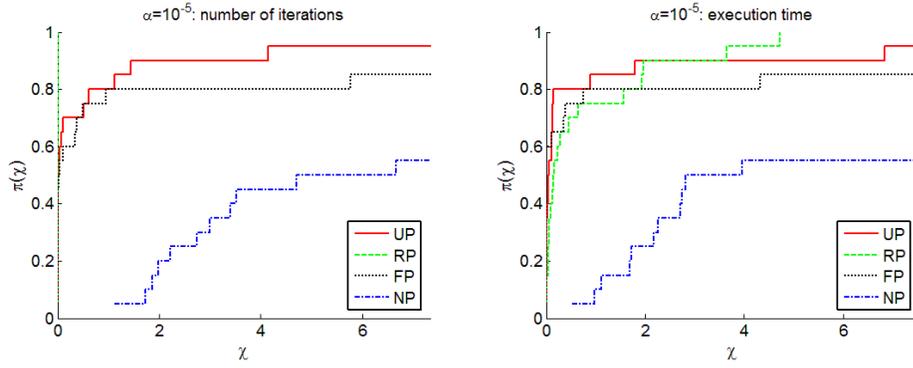


FIG. 5.2. Performance profiles on the first test set for $\alpha = 10^{-5}$: number of CG iterations (left) and execution times (right).

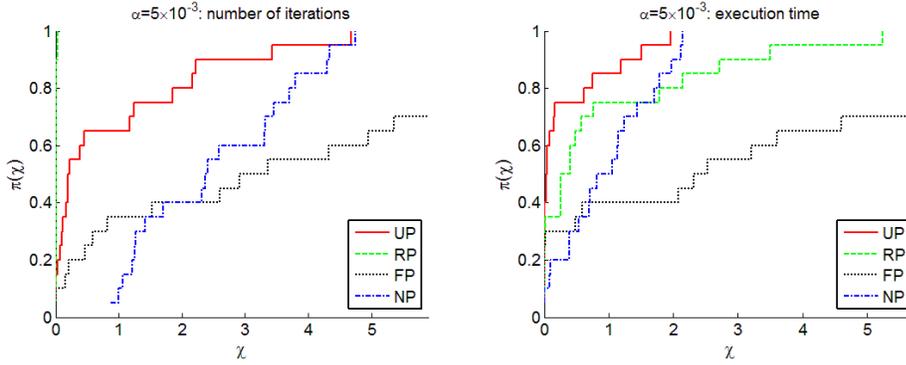


FIG. 5.3. Performance profiles on the first test set for $\alpha = 5 \times 10^{-3}$: number of CG iterations (left) and execution times (right).

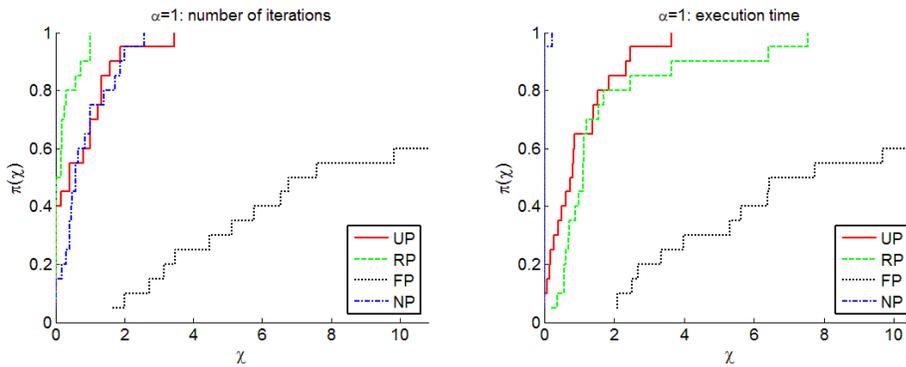


FIG. 5.4. Performance profiles on the first test set for $\alpha = 1$: number of CG iterations (left) and execution times (right).

As expected, for each value of α , RP is the most efficient in terms of CG iterations, but not in terms of time. The performance of FP steadily deteriorates while α increases and, for some tests, it already gives poor results with the small value $\alpha = 10^{-5}$; more precisely, FP fails in solving 3, 6, 8 systems for $\alpha = 10^{-5}$, $5 \cdot 10^{-3}$, 1, respectively. The NP strategy results in the slowest convergence rate and the largest execution time until α becomes sufficiently large, i.e. when preconditioning is no longer needed. Remarkably, the UP strategy is very effective in reducing the number of CG iterations and provides a good tradeoff between the cost of the linear solver and of the construction of the preconditioner. In particular, our strategy is the clear winner in terms of execution time for $\alpha = 10^{-5}$ and $\alpha = 5 \cdot 10^{-3}$, and shows a gain over RP and FP when $\alpha = 1$.

In summary, we see that FP and NP work well for specific values of α , RP is expensive in terms of time to solution, while the UP strategy shows a remarkable reliability and efficiency in solving sequences of shifted systems for a fairly broad range of values of α .

We conclude the analysis of the first set of experiments reporting, in Tables 5.3 and 5.4, the results for the sequences of systems built from the matrices `nd3k` and `s1rmq4m1` (* denotes a CG failure). The test problem `nd3k` is representative of the convergence history of CG with the various preconditioning techniques on several test problems. In this case, NP fails for the first value of α , while it gains in iteration count (and time) as α increases; on the contrary, FP deteriorates for increasing values of α . The UP strategy generally behaves well for all values of α and is considerably faster than RP; in fact, the time for computing the preconditioner from scratch is relatively high. In terms of execution time, the UP strategy outperforms the unpreconditioned CG as long as $\alpha \leq 10^{-3}$. Concerning the sequence associated with `s1rmq4m1`, the time for building the preconditioner from scratch for each matrix $A + \alpha I$ is small, so the RP strategy is fast. However, UP yields a number of CG iterations comparable with that of RP and hence has a smaller execution time. More precisely, the relative time saving of UP vs RP vary between 9% and 46%. Finally, the execution times of UP and RP compare favourably with the times of FP and NP for all values of α .

The results of our experiments on the second set of sequences are summarized in Table 5.5, where the total number of CG iterations and the total execution times for solving all the systems of each sequence are reported. As for the first test set, we declare a failure (denoted by * in the table) when the preconditioned CG fails in solving at least one system of the sequence. We note that these sequences really need a preconditioning strategy; in fact, the unpreconditioned CG fails in solving the first linear system of each sequence, corresponding to the smallest value of α . For five of the eight sequences, the UP strategy results to be the most efficient. Actually, for these sequences RP is the most effective in terms of CG iterations, but it requires a significantly larger execution time than UP, because of the computation of the incomplete LDL^T factorization for each value of α . The sequence `pdex1.1` results hard to be solved, as only the RP strategy succeeds in its solution; our update strategy fails in solving the first linear system of the sequence. We observe also that on the `fpm.1` and `fpm.2` sequences the UP strategy requires a large number of CG iterations compared to the RP one. Moreover, the small drop tolerance $5 \cdot 10^{-5}$ that is needed for computing an effective incomplete Cholesky preconditioner, produces a relatively dense factor L , hence the time required for applying the preconditioner, i.e. for solving the related triangular systems at each CG iteration, is significant. Therefore, for these two sequences, the considerable gain in terms of iterations obtained by the RP

α	CG iterations				execution time (sec.)			
	UP	RP	FP	NP	UP	RP	FP	NP
1.e-5	67	47	66	*	1.7e0	2.1e1	1.6e0	*
5.e-5	46	20	48	806	1.2e0	1.9e1	1.2e0	1.4e1
1.e-4	44	15	48	553	1.1e0	1.8e1	1.2e0	9.8e0
5.e-4	54	8	68	211	1.4e0	1.5e1	1.7e0	3.7e0
1.e-3	57	7	83	142	1.5e0	1.3e1	2.0e0	2.5e0
5.e-3	54	5	155	50	1.4e0	1.0e1	3.8e0	8.9e-1
1.e-2	49	4	197	33	1.3e0	8.9e0	4.8e0	6.0e-1
5.e-2	39	3	375	10	1.0e0	7.1e0	9.1e0	1.9e-1
1.e-1	31	3	481	6	8.2e-1	6.5e0	1.2e0	1.2e-1
5.e-1	15	2	751	2	4.2e-1	5.4e0	1.8e1	4.5e-2
1.e0	11	2	906	1	3.3e-1	5.0e0	2.2e1	2.7e-2

TABLE 5.3

Results on the sequence of shifted systems associated with `nd3k` (* denotes a CG failure).

α	CG iterations				execution time (sec.)			
	UP	RP	FP	NP	UP	RP	FP	NP
1.e-5	153	154	155	*	4.1e-1	4.5e-1	4.2e-1	*
5.e-5	103	103	109	937	2.8e-1	3.1e-1	2.9e-1	1.9e0
1.e-4	86	86	96	730	2.3e-1	2.7e-1	2.5e-1	1.5e0
5.e-4	58	54	110	389	1.6e-1	1.8e-1	2.9e-1	7.8e-1
1.e-3	49	45	131	295	1.3e-1	1.6e-1	3.5e-1	5.9e-1
5.e-3	32	28	170	145	9.0e-2	1.1e-1	4.5e-1	2.9e-1
1.e-2	27	24	180	109	7.7e-2	9.5e-2	4.8e-1	2.2e-1
5.e-2	18	21	183	51	5.3e-2	8.2e-2	4.9e-1	1.1e-1
1.e-1	14	20	187	36	4.3e-2	8.0e-2	5.0e-1	7.5e-2
5.e-1	9	12	312	16	3.0e-2	4.9e-2	8.2e-1	3.5e-2
1.e0	7	8	378	11	2.5e-2	3.7e-2	9.9e-1	2.5e-2

TABLE 5.4

Results on the sequence of shifted systems associated with `s1rmq4m1` (* denotes a CG failure).

strategy offsets the cost of the computation of the preconditioner from scratch and makes RP the most efficient in the solution of these two sequences. The FP strategy is generally worse than UP and RP.

6. Conclusions and future work. We proposed a new preconditioning technique for the iterative solution of sequences of symmetric positive definite shifted linear systems. For each shifted matrix $A + \alpha I$, a preconditioner is obtained by a suitable update of a preconditioner for A , available in the LDL^T form. This update is low cost, preserves the sparsity pattern of the initial factors and is very easy to be implemented. A theoretical analysis shows that our technique provides preconditioners that are good approximations of the corresponding shifted matrices and it is able to mimic the behaviour of the matrix sequence not only for small values of α , but also when α increases. Numerical experiments on several test problems confirm the theoretical results and show that our approach is more efficient than freezing the preconditioner or recomputing it from scratch.

Future work will be devoted to extending our procedure to sequences of general symmetric matrices which differ for a diagonal matrix, such as the KKT matrices arising in interior point methods for linear and quadratic programming.

Sequence	total CG iterations				total execution time (sec.)			
	UP	RP	FP	NP	UP	RP	FP	NP
bratu_1	117	66	300	*	5.1e-1	1.2e0	1.1e1	*
bratu_2	126	110	187	*	3.6e-1	1.6e0	5.5e-1	*
fpm_1	443	103	*	*	5.9e0	5.6e0	*	*
fpm_2	671	100	5250	*	8.6e0	5.9e0	6.2e1	*
pdexs_1	114	59	586	*	5.5e-1	1.5e0	2.3e0	*
pdexs_2	127	60	314	*	5.9e-1	1.4e0	1.2e0	*
pdex1_1	*	209	*	*	*	5.3e0	*	*
pdex1_2	96	57	850	*	4.8e-1	1.5e0	3.4e0	*

TABLE 5.5

Results on the second test set (* denotes a CG failure).

REFERENCES

- [1] U.M. Ascher and L.R. Petzold, *Computer Methods for Ordinary Differential Equations and Differential-Algebraic Equations*, SIAM, 1998.
- [2] Z.-Z. Bai, I.S. Duff and A.J. Wathen, *A Class of Incomplete Orthogonal Factorization Methods. I: Methods and Theories*, BIT Numerical Mathematics, 41 (2001), pp. 53–70.
- [3] S. Bellavia, D. Bertaccini and B. Morini, *Nonsymmetric preconditioner updates in Newton-Krylov methods for nonlinear systems*, Report 10/01, Dipartimento di Energetica “S.Stecco”, Università di Firenze, 2010.
- [4] S. Bellavia, C. Cartis, N.I.M. Gould, B. Morini and Ph.L. Toint, *Convergence of a Regularized Euclidean Residual Algorithm for Nonlinear Least-Squares*, SIAM J. Num. Anal., 48 (2010), pp. 1–29.
- [5] M. Benzi, J.K. Cullum and M. Tüma, *Robust approximate inverse preconditioning for the conjugate gradient method*, SIAM J. Sci. Comput., 22 (2000), pp. 1318–1332.
- [6] M. Benzi and D. Bertaccini, *Approximate inverse preconditioning for shifted linear systems*, BIT, 43 (2003), pp. 231–244.
- [7] M. Benzi and M. Tüma, *A Robust Preconditioner with Low Memory Requirements for Large Sparse Least Squares Problems*, SIAM J. Sci. Comput., 25 (2003), pp. 499–512.
- [8] M. Benzi and M. Tüma, *A Robust Incomplete Factorization Preconditioner for Positive Definite Matrices*, Numer. Linear Algebra with Appl., 10 (2003), pp. 385–400.
- [9] D. Bertaccini, *Efficient preconditioning for sequences of parametric complex symmetric linear systems*, ETNA, 18 (2004), pp. 49–64.
- [10] C. Cartis, N.I.M. Gould and Ph.L. Toint, *Adaptive cubic overestimation methods for unconstrained optimization. Part I: motivation, convergence and numerical results*, Math. Prog. A, to appear.
- [11] C. Cartis, N.I.M. Gould and Ph.L. Toint, *Trust-region and other regularisations of linear least-squares problems*, BIT, 49 (2009), pp. 21–53.
- [12] T.A. Davis and Y.F. Hu, *The University of Florida sparse matrix collection*, submitted to ACM Trans. on Math. Soft., 2010, <http://www.cise.ufl.edu/research/sparse/matrices>.
- [13] J.E. Dennis, Jr., and R.B. Schnabel *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*, SIAM, 1996.
- [14] E.D. Dolan and J.J. Moré, *Benchmarking optimization software with performance profiles*, Math. Prog., 91 (2002), pp. 201–213.
- [15] J. Duintjer Tebbens and M. Tüma, *Efficient Preconditioning of Sequences of Nonsymmetric Linear Systems*, SIAM J. Sci. Comput., 29 (2007), pp. 1918–1941.
- [16] J. Duintjer Tebbens and M. Tüma, *Preconditioner Updates for Solving Sequences of Linear Systems in Matrix-free Environment*, Numer. Linear Algebra with Appl., to appear.
- [17] S.C. Eisenstat, H.F. Walker, *Choosing the forcing term in an inexact Newton method*, SIAM J. Sci. Comput., 17 (1996), pp. 16–32.
- [18] J.B. Erway and P.E. Gill, *A Subspace Minimization Method for the Trust-Region Step*, SIAM J. Optim., 20 (2009), pp. 1439–1461.
- [19] G.H. Golub and C.F. Van Loan, *Matrix Computations*, third edition, The Johns Hopkins University Press, 1996.
- [20] J. Gondzio, *Matrix-Free Interior Point Method*, Technical Report ERGO-09-012, School of Mathematics, The University of Edinburgh, revised 29 April 2010.

- [21] N.I.M. Gould, D.P. Robinson and H.S. Thorne, *On solving trust-region and other regularised subproblems in optimization*, Math. Program. Comput., 2 (2010), pp. 21–57.
- [22] R.A. Horn and C.R. Johnson, *Topics in Matrix Analysis*, Cambridge University Press, 1991.
- [23] G. Meurant, *On the incomplete Cholesky decomposition of a class of perturbed matrices*, SIAM J. Sci. Comput., 23 (2001), pp. 419–429.
- [24] J. J. Moré and D.C. Sorensen, *Computing a trust-region step*, SIAM J. Sci. Stat. Comput., 4 (1983), pp. 553–572.
- [25] Y. Nesterov, *Modified Gauss-Newton scheme with worst-case guarantees for global performance*, Optim. Methods Softw., 22 (2007), pp.469–483.
- [26] J. Nocedal and S.J. Wright, *Numerical Optimization*, Springer Series in Operation Research, Springer-Verlag, 1999.
- [27] C.C. Paige and M.A. Saunders, *LSQR: An algorithm for sparse linear equations and sparse least squares*, ACM Trans. Math. Soft. 8 (1982), pp. 43–71.
- [28] M. PERNICE AND H.F. WALKER, *NITSOL: a new iterative solver for nonlinear systems*, SIAM J. Sci. Comput., 19 (1998), pp. 302–318.