J Optim Theory Appl (2015) 164:883–914
DOI 10.1007/s10957-014-0600-0

A Framework of Conjugate Direction Methods for Symmetric Linear Systems in Optimization

Giovanni Fasano

Received: 30 December 2012 / Accepted: 8 June 2014 / Published online: 15 July 2014

Abstract In this paper, we introduce a parameter-dependent class of Krylov-based methods, namely Conjugate Directions (CD), for the solution of symmetric linear systems. We give evidence that, in our proposal, we generate sequences of conjugate directions, extending some properties of the standard conjugate gradient (CG) method, in order to preserve the conjugacy. For specific values of the parameters in our framework, we obtain schemes equivalent to both the CG and the scaled-CG. We also prove the finite convergence of the algorithms in CD, and we provide some error analysis. Finally, preconditioning is introduced for CD, and we show that standard error bounds for the preconditioned CG also hold for the preconditioned CD.

Keywords Krylov-based methods · Conjugate direction methods · Conjugacy loss and error analysis · Preconditioning

Mathematics Subject Classification (2000) 90C30 · 90C06 · 65K05 · 49M15

1 Introduction

The solution of symmetric linear systems arises in a wide range of real applications [1–3], and has been carefully issued in the last 50 years, due to the increasing demand of fast and reliable solvers. *Illconditioning* and *large number of unknowns* are among

Present address:
G. Fasano (✉)
Department of Management, University Ca'Foscari of Venice, Venice, Italy
e-mail: fasano@unive.it

G. Fasano
S.Giobbe, Cannaregio 873, 30121 Venice, Italy

the most challenging issues which may harmfully affect the solution of linear systems, in several frameworks where either structured or unstructured coefficient matrices are considered [1,4,5].

The latter facts have required the introduction of a considerable number of techniques, specifically aimed at tackling classes of linear systems with appointed pathologies [5,6]. We remark that the structure of the coefficient matrix may be essential for the success of the solution methods, both in numerical analysis and optimization contexts. As an example, PDEs and PDE-constrained optimization provide two specific frameworks, where sequences of linear systems often claim for specialized and robust methods, in order to give reliable solutions.

In this paper, we focus on iterative Krylov-based methods for the solution of symmetric linear systems, arising in both numerical analysis and optimization contexts. The theory detailed in the paper is not limited to consider large-scale linear systems; however, since Krylov-based methods have proved their efficiency when the scale is large, without any loss of generality, we will implicitly assume the latter fact.

The accurate study and assessment of methods for the solution of linear systems is naturally expected from the community of people working on numerical analysis, that is due to their expertise and great sensibility to theoretical issues, rather than to practical algorithms implementation or software developments. This has raised a consistent literature, including manuals and textbooks, where the analysis of solution techniques for linear systems has become a keynote subject, and where essential achievements have given strong guidelines to theoreticians and practitioners from optimization [4].

We address here a parameter-dependent class of CG-based methods, which can equivalently reduce to the CG for a suitable choice of the parameters. We firmly claim that our proposal is not primarily intended to provide an efficient alternative to the CG. On the contrary, we mainly detail a general framework of iterative methods, inspired by polarity for quadratic hypersurfaces, and based on the generation of conjugate directions. The algorithms in our class, thanks to the parameters in the scheme, may possibly keep under control the conjugacy loss among directions, which is often caused by finite precision in the computation. The paper is not intended to report also a significant numerical experience. Indeed, we think that there are not yet clear rules on the parameters of our proposal, for assessing efficient algorithms. Similarly, we have not currently evidence that methods in our proposal can outperform the CG. On this guideline, in a separate paper, we will carry on selective numerical tests, considering both symmetric linear systems from numerical analysis and optimization. We further prove that preconditioning can be introduced for the class of methods we propose, as a natural extension of the preconditioned CG (see also [2]).

Section 2 briefly reviews both the CG and the Lanczos process, as Krylov-subspace methods, in order to highlight promising aspects to investigate in our proposal. Section 3 details some relevant applications of conjugate directions in optimization frameworks, motivating our interest for possible extensions of the CG. In Sects. 4 and 5, we describe our class of methods and some related properties. In Sects. 6 and 7, we show that the CG and the scaled-CG may be equivalently obtained as particular members of our class. Then, Sects. 8 and 9 contain further properties of the class of methods we propose. Finally, Sect. 10 analyzes the preconditioned version of our proposal, and a section of Conclusions completes the paper, including some numerical results.

2 The CG Method and the Lanczos Process

In this section, we comment the method in Table 1, and we focus on the relation between the CG and the Lanczos process, as Krylov-subspace methods. In particular, the Lanczos process namely does not generate conjugate directions; however, though our proposal relies on generalizing the CG, it shares some aspects with the Lanczos iteration, too.

As we said, the CG is commonly used to iteratively solving the linear system

$$Ay = b, \tag{1}$$

where $A \in \mathbb{R}^{n\times n}$ is symmetric *positive definite* and $b \in \mathbb{R}^n$. Observe that the CG is quite often applied to a *preconditioned* version of the linear system (1), i.e., $\mathcal{M}Ay = \mathcal{M}b$, where $\mathcal{M} \succ 0$ is the preconditioner [7]. Though the theory for the CG requires A to be positive definite, in several practical applications it is successfully used when A is indefinite, too [8,9]. At Step k the CG generates the pair of vectors r_k (*residual*) and p_k (*search direction*) such that [2]

$$\text{orthogonality property}: \quad r_i^T r_j = 0, \qquad 0 \le i \ne j \le k, \tag{2}$$

$$\text{conjugacy property}: \quad p_i^T A p_j = 0, \qquad 0 \le i \ne j \le k. \tag{3}$$

Moreover, finite convergence holds, i.e., $Ay_h = b$ for some $h \le n$. Relations (2) yield the Ritz-Galerkin condition $r_k \perp \mathcal{K}_{k-1}(r_0, A)$, where $\mathcal{K}_{k-1}(r_0, A)$ is the Krylov-subspace

$$\mathcal{K}_{k-1}(r_0, A) := \text{span}\{b, Ab, A^2b, \ldots, A^{k-1}b\} \equiv \text{span}\{r_0, \ldots, r_{k-1}\}.$$

Furthermore, the direction p_k is computed at Step k imposing the conjugacy condition $p_k^T A p_{k-1} = 0$. It can be easily proved that the latter equality implicitly satisfies relations (3), with $p_0, \ldots, p_k$ linearly independent. We remark that on practical problems, due to *finite precision* and *roundoff* in the computation of the sequences $\{p_k\}$ and $\{r_k\}$, when $|i - j|$ is large, relations (2)–(3) may fail. Thus, in the practical implementation

Table 1 The CG algorithm for solving (1)

The Conjugate Gradient (CG) method

Step 0: Set $k = 0$, $y_0 \in \mathbb{R}$, $r_0 := b - Ay_0$.
If $r_0 = 0$, then STOP. Else, set $p_0 := r_0$; $k = k + 1$.
Set $p_{-1} = 0$ and $\beta_{-1} = 0$.

Step k: Compute $\alpha_{k-1} := r_{k-1}^T p_{k-1} / p_{k-1}^T A p_{k-1}$,
$y_k := y_{k-1} + \alpha_{k-1} p_{k-1}$, $r_k := r_{k-1} - \alpha_{k-1} A p_{k-1}$.
If $r_k = 0$, then STOP. Else, set
– $\beta_{k-1} := \|r_k\|^2 / \|r_{k-1}\|^2$, $p_k := r_k + \beta_{k-1} p_{k-1}$
– (or equivalently set $p_k := -\alpha_{k-1} A p_{k-1} + (1 + \beta_{k-1}) p_{k-1} - \beta_{k-2} p_{k-2}$)
Set $k = k + 1$, go to **Step** k.

of the CG some theoretical properties may not be satisfied, and in particular when $|i - j|$ increases, the conjugacy properties (3) may progressively be lost. As detailed in [10–13], the latter fact may have dramatic consequences also in optimization frameworks (see also Sect. 3 for details). To our purposes we note that in Table 1, at Step k of the CG, the direction p_k is usually computed as

$$p_k := r_k + \beta_{k-1} p_{k-1}, \tag{4}$$

but an equivalent expression is (see also Theorem 5.4 in [14])

$$p_k := -\alpha_{k-1} A p_{k-1} + (1 + \beta_{k-1}) p_{k-1} - \beta_{k-2} p_{k-2}, \tag{5}$$

which we would like to generalize in our proposal. Note also that in exact arithmetics the property (3) is iteratively fulfilled by both (4) and (5).

The Lanczos process (and its preconditioned version) is another Krylov-based method, widely used to tridiagonalize the matrix A in (1). Unlike the CG method, here the matrix A may be possibly indefinite, and the overall method is slightly more expensive than the CG, since further computation is necessary to solve the resulting tridiagonal system. Similarly to the CG, the Lanczos process generates at Step k the sequence $\{u_k\}$ (*Lanczos vectors*) which satisfies

$$\text{orthogonality property}: \qquad u_i^T u_j = 0, \qquad 0 \le i \ne j \le k,$$

and yields finite convergence in at most n steps. However, unlike the CG the Lanczos process is not explicitly inspired by polarity, in order to generate the orthogonal vectors. We recall that the CG and the Lanczos process are 3-term recurrence methods, in other words, for $k \ge 1$

$$\begin{aligned} p_{k+1} &\in \text{span}\{A p_k, p_k, p_{k-1}\}, && \text{for the CG} \\ u_{k+1} &\in \text{span}\{A u_k, u_k, u_{k-1}\}, && \text{for the Lanczos process.} \end{aligned}$$

When A is positive definite, a full theoretical correspondence between the sequence $\{r_k\}$ of the CG and the sequence $\{u_k\}$ of the Lanczos process may be fruitfully used in optimization problems (see also [10,15,16]), being

$$u_k = s_k \frac{r_k}{\|r_k\|}, \qquad s_k \in \{-1, +1\}.$$

The class CD proposed in this paper provides a framework, which encompasses the CG and to some extent resembles the Lanczos iteration, since a 3-term recurrence is exploited. In particular, the CD generates both conjugate directions (as the CG) and orthogonal residuals (as the CG and the Lanczos process). Moreover, similarly to the CG, the CD yields a 3-term recurrence with respect to conjugate directions. As we remarked, our proposal draws its inspiration from the idea of possibly attenuating the conjugacy loss of the CG, which may occur in (3) when $|i - j|$ is large.

3 Conjugate Directions for Optimization Frameworks

Optimization frameworks offer plenty of symmetric linear systems where CG-based methods are often specifically preferable with respect to other solvers. Here we justify this statement by briefly describing the potential use of conjugate directions within truncated Newton schemes. The latter methods strongly prove their efficiency when applied to large scale problems, where they rely on the proper computation of search directions, as well as truncation rules (see [17]).

As regards the computation of search directions, suppose at the outer iteration h of the truncated scheme we perform m steps of the CG, in order to compute the approximate solution d_h^m to the linear system (Newton's equation)

$$\nabla^2 f(z_h)d = -\nabla f(z_h).$$

When z_h is close enough to the solution z^* (minimum point), then possibly $\nabla^2 f(z_h) \succ 0$. Thus, the conjugate directions $p_1, \ldots, p_m$ and the coefficients $\alpha_1, \ldots, \alpha_m$ are generated as in Table 1, so that the following vectors can be formed

$$\begin{aligned}
d_h^m &= \sum_{i=1}^{m} \alpha_i p_i, \\
d_h^P &= \sum_{i \in I_h^P} \alpha_i p_i, \qquad I_h^P = \left\{ i \in \{1, \ldots, m\} : p_i^T \nabla^2 f(z_h) p_i > 0 \right\}, \\
d_h^N &= \sum_{i \in I_h^N} \alpha_i p_i, \qquad I_h^N = \left\{ i \in \{1, \ldots, m\} : p_i^T \nabla^2 f(z_h) p_i < 0 \right\}, \\
s_h &= \frac{p_\ell}{\|r_\ell\|}, \qquad \ell = \arg\min_{i \in \{1,\ldots,m\}} \left\{ \frac{p_i^T \nabla^2 f(z_h) p_i}{\|r_i\|^2} : p_i^T \nabla^2 f(z_h) p_i < 0 \right\}.
\end{aligned} \tag{6}$$

Observe that d_h^m approximates in some sense Newton's direction at the outer iteration h, and as described in [11,12,18,19], the vectors d_h^m, d_h^P and d_h^N can be used/combined to provide fruitful search directions to the optimization framework. Moreover, d_h^N and s_h are suitably used/combined to compute a so called *negative curvature direction* 's_h^m', which can possibly force second order convergence for the overall truncated optimization scheme (see [18] for details). The conjugacy property is essential for computing the vectors (6), i.e., to design efficient truncated Newton methods. Thus, introducing CG-based schemes which deflate conjugacy loss might be of great importance.

On the other hand, at the outer iteration h, effective truncation rules typically attempt to assess the parameter m in (6), as described in [17,20,21], i.e., they monitor the decrease of the quadratic local model

$$Q_h(d_h^m) := f(z_h) + \nabla f(z_h)^T (d_h^m) + \frac{1}{2}(d_h^m)^T \nabla^2 f(z_h)(d_h^m)$$

when $\nabla^2 f(z_h) \succ 0$, so that the parameter m is chosen to satisfy some conditions, including

$$\frac{Q_h(d_h^m) - Q_h(d_h^{m-1})}{Q_h(d_h^m)/m} \leq \alpha, \qquad \text{for some } \alpha \in]0, 1[.$$

Thus, again the correctness of conjugacy properties among the directions $p_1, \ldots, p_m$, generated while solving Newton's equation, may be essential both for an accurate solution of Newton's equation (which is a linear system) and to the overall efficiency of the truncated optimization method.

4 Our Proposal: The *CD* Class

Before introducing our proposal for a new general framework of CG-based algorithms, we consider here some additional motivations for using the CG.

The careful use of the latter theory is in our opinion a launching pad for possible extensions of the CG. On this guideline, recalling the contents in Sect. 3, now we summarize some critical aspects of the CG:

1. The CG works iteratively and at any iteration the overall computational effort is only $O(n^2)$ (since the CG is a Krylov-subspace method);
2. The conjugate directions generated by the CG are linearly independent, so that at most n iterations are necessary to address the solution; and
3. The current conjugate direction p_{k+1} is computed by simply imposing the conjugacy with respect to the direction p_k (computed) in the previous iteration. This automatically yields that $p_{k+1}^T A p_i = 0$, for any $i \leq k$, too.

As a matter of fact, for the design of possible general frameworks including CG-based methods, the items 1. and 2. are essential in order to respectively control the *computational effort* and ensure the *finite convergence*.

On the other hand, altering the item 3. might be harmless for the overall iterative process, and might possibly yield some fruitful generalizations, that is indeed the case of our proposal, where the item 3. is modified with respect to the CG. The latter modification depends on a parameter which is user/problem-dependent, and may be set in order to further compensate or correct the conjugacy loss among directions, due to roundoff and finite precision.

We sketch in Table 2 our new CG-based class of algorithms, namely *CD*.

The computation of the direction p_k at Step k reveals the main difference between the CG and *CD*. In particular, in Table 2 the pair of coefficients σ_{k-1} and ω_{k-1} is

Table 2 The parameter-dependent class CD of CG-based algorithms for solving (1)

The CD class

Step 0: Set $k = 0$, $y_0 \in \mathbb{R}^n$, $r_0 := b - Ay_0$, $\gamma_0 \in \mathbb{R} \setminus \{0\}$.
If $r_0 = 0$, then STOP. Else, set $p_0 := r_0$, $k = k + 1$.
Compute $a_0 := r_0^T p_0 / p_0^T A p_0$,
$y_1 := y_0 + a_0 p_0$, $r_1 := r_0 - a_0 A p_0$.
If $r_1 = 0$, then STOP. Else, set $\sigma_0 := \gamma_0 \|Ap_0\|^2 / p_0^T A p_0$,
$p_1 := \gamma_0 A p_0 - \sigma_0 p_0$, $k = k + 1$.

Step k: Compute $a_{k-1} := r_{k-1}^T p_{k-1} / p_{k-1}^T A p_{k-1}$,
$y_k := y_{k-1} + a_{k-1} p_{k-1}$, $r_k := r_{k-1} - a_{k-1} A p_{k-1}$.
If $r_k = 0$, then STOP. Else, set $\sigma_{k-1} := \gamma_{k-1} \frac{\|Ap_{k-1}\|^2}{p_{k-1}^T A p_{k-1}}$,
$\omega_{k-1} := \gamma_{k-1} \frac{(Ap_{k-1})^T A p_{k-2}}{p_{k-2}^T A p_{k-2}} = \frac{\gamma_{k-1}}{\gamma_{k-2}} \frac{p_{k-1}^T A p_{k-1}}{p_{k-2}^T A p_{k-2}}$, $\gamma_{k-1} \in \mathbb{R} \setminus \{0\}$
$p_k := \gamma_{k-1} A p_{k-1} - \sigma_{k-1} p_{k-1} - \omega_{k-1} p_{k-2}$, $k = k + 1$.
Go to **Step** k.

computed so that *explicitly*[1]

$$p_k^T A p_{k-1} = 0$$
$$p_k^T A p_{k-2} = 0, \tag{8}$$

i.e., in Cartesian coordinates the conjugacy between the direction p_k and both the directions p_{k-1} and p_{k-2} is directly imposed, as specified by (3). As detailed in Sect. 2, imposing the double condition (8) allows to possibly recover the conjugacy loss in the sequence $\{p_i\}$.

On the other hand, the residual r_k at Step k of Table 2 is computed by imposing the orthogonality condition $r_k^T p_{k-1} = 0$, as in the standard CG. The resulting method is *evidently a bit more expensive than the CG*, requiring one additional inner product per step, as long as an additional scalar to compute and an additional n-vector to store. From Table 2 it is also evident that CD provides a 3-term recurrence with respect to the conjugate directions.

In addition, observe that the residual r_k is computed at Step k of CD only to check for the stopping condition, and is not directly involved in the computation of p_k. Hereafter, in this section, we briefly summarize the basic properties of the class CD.

Assumption 4.1 The matrix A in (1) is symmetric positive definite. Moreover, the sequence $\{\gamma_k\}$ in Table 2 is such that $\gamma_k \neq 0$, for any $k \geq 0$.

[1] A further generalization might be obtained computing σ_{k-1} and ω_{k-1} so that

$$\begin{cases} p_k^T A(\gamma_{k-1} A p_{k-1} - \sigma_{k-1} p_{k-1}) = 0, \\ p_k^T A p_{k-2} = 0. \end{cases} \tag{7}$$

Note that, as for the CG, Assumption 4.1 is required for theoretical reasons. However, the CD class may in principle be used also in several cases when A is indefinite, provided that $p_k^T A p_k \neq 0$, for any $k \geq 0$.

Lemma 4.1 *Let Assumption 4.1 hold. At Step k of the CD class, with $k \geq 0$, we have*

$$Ap_j \in \text{span}\left\{p_{j+1}, p_j, p_{\max\{0,j-1\}}\right\}, \; j \leq k. \tag{9}$$

Proof From the Step 0 relation (9) holds for $j = 0$. Then, for $j = 1, \ldots, k-1$ the Step $j+1$ of CD directly yields (9). □

Theorem 4.1 (Conjugacy) *Let Assumption 4.1 hold. At Step k of the CD class, with $k \geq 0$, the directions $p_0, p_1, \ldots, p_k$ are mutually conjugate, i.e., $p_i^T A p_j = 0$, with $0 \leq i \neq j \leq k$.*

Proof The statement holds for Step 0, as a consequence of the choice of the coefficient σ_0. Suppose it holds for $k-1$; then, we have for $j \leq k-1$

$$\begin{aligned} p_k^T A p_j &= (\gamma_{k-1} A p_{k-1} - \sigma_{k-1} p_{k-1} - \omega_{k-1} p_{k-2})^T A p_j \\ &= (\gamma_{k-1} A p_{k-1})^T A p_j - \sigma_{k-1} p_{k-1}^T A p_j - \omega_{k-1} p_{k-2}^T A p_j = 0. \end{aligned}$$

In particular, for $j = k-1$ and $j = k-2$ the choice of the coefficients σ_{k-1} and ω_{k-1}, and the inductive hypothesis, yield $p_k^T A p_{k-1} = p_k^T A p_{k-2} = 0$. For $j < k-2$, the inductive hypothesis and Lemma 4.1 again yield the conjugacy property. □

Lemma 4.2 *Let Assumption 4.1 hold. Given the CD class, we have for $k \geq 2$*

$$(Ap_k)^T (Ap_i) = \begin{cases} \|Ap_k\|^2, & \text{if } i = k, \\ \frac{1}{\gamma_{k-1}} p_k^T A p_k, & \text{if } i = k-1, \\ \emptyset, & \text{if } i \leq k-2. \end{cases}$$

Proof The statement is a trivial consequence of Step k of the CD, Lemma 4.1 and Theorem 4.1. □

Observe that from the previous lemma, a simplified expression for the coefficient ω_{k-1}, at Step k of CD is available, inasmuch as

$$\omega_{k-1} = \frac{\gamma_{k-1}}{\gamma_{k-2}} \cdot \frac{p_{k-1}^T A p_{k-1}}{p_{k-2}^T A p_{k-2}}. \tag{10}$$

Relation (10) has a remarkable importance: it avoids the storage of the vector Ap_{k-2} at Step k, requiring only the storage of the quantity $p_{k-2}^T A p_{k-2}$. Furthermore observe that, unlike the CG, the sequence $\{p_k\}$ in CD is computed independently of the sequence $\{r_k\}$. Moreover, as we said, the residual r_k is simply computed at Step k in order to check the stopping condition for the algorithm.

The following result proves that the CD class recovers the main theoretical properties of the standard CG.

Theorem 4.2 (Orthogonality) *Let Assumption 4.1 hold. Let $r_{k+1} \neq 0$ at Step $k+1$ of the CD class, with $k \geq 0$. Then, the directions $p_0, p_1, \ldots, p_k$ and the residuals $r_0, r_1, \ldots, r_{k+1}$ satisfy*

$$r_{k+1}^T p_j = 0, \qquad j \leq k, \tag{11}$$

$$r_{k+1}^T r_j = 0, \qquad j \leq k. \tag{12}$$

Proof From Step $k+1$ of CD we have $r_{k+1} = r_k - a_k A p_k = r_j - \sum_{i=j}^{k} a_i A p_i$, for any $j \leq k$. Then, from Theorem 4.1 and the choice of coefficient α_j we obtain

$$r_{k+1}^T p_j = \left(r_j - \sum_{i=j}^{k} a_i A p_i \right)^T p_j = r_j^T p_j - \sum_{i=j}^{k} a_i p_i^T A p_j = 0, \; j \leq k,$$

which proves (11). As regards relation (12), for $k = 0$ we obtain from the choice of a_0

$$r_1^T r_0 = r_1^T p_0 = 0.$$

Then, assuming by induction that (12) holds for $k-1$, we have

$$\begin{aligned} r_{k+1}^T r_j &= (r_k - a_k A p_k)^T r_j = (r_k - a_k A p_k)^T \left(r_0 - \sum_{i=0}^{j-1} a_i A p_i \right) \\ &= r_k^T r_0 - \sum_{i=0}^{j-1} a_i r_k^T A p_i - a_k p_k^T A r_0 + \sum_{i=0}^{j-1} a_i a_k (A p_k)^T A p_i, \; j \leq k. \end{aligned}$$

The inductive hypothesis and Theorem 4.1 yield for $j \leq k$ (in next relation, when $i = 0$, then $p_{i-1} \equiv 0$)

$$r_{k+1}^T r_j = -\sum_{i=0}^{j-1} \frac{a_i r_k^T}{\gamma_i} (p_{i+1} + \sigma_i p_i + \omega_i p_{i-1}) + \sum_{i=0}^{j-1} a_i a_k (A p_k)^T A p_i. \tag{13}$$

Therefore, if $j = k$ the relation (11) along with Lemma 4.2 and the choice of a_k yield

$$r_{k+1}^T r_k = -\frac{a_{k-1}}{\gamma_{k-1}} r_k^T p_k + \frac{a_{k-1} a_k}{\gamma_{k-1}} p_k^T A p_k = 0.$$

On the other hand, if $j < k$ in (13), the inductive hypothesis, relation (11) and Lemma 4.2 yield (12). □

Finally, we prove that, likewise the CG, in at most n iterations CD determines the solution of the linear system (1), so that finite convergence holds.

Lemma 4.3 (Finite convergence) *Let Assumption 4.1 hold. At Step k of the C D class, with $k \geq 0$, the vectors $p_0, \ldots, p_k$ are linearly independent. Moreover, in at most n iterations the C D class computes the solution of the linear system (1), i.e., $Ay_h = b$, for some $h \leq n$.*

Proof The proof follows very standard guidelines (the reader may also refer to [22]). Thus, by (11) an integer $m \leq n$ exists such that $r_m = b - Ay_m = 0$. Then, if y^* is the solution of (1), we have

$$0 = b - Ay_m = Ay^* - A\left[y_0 + \sum_{i=0}^{m-1} a_i p_i\right] \quad \Longleftrightarrow \quad y^* = y_0 + \sum_{i=0}^{m-1} a_i p_i.$$

□

Remark 4.1 Observe that there is the additional chance to replace the Step 0 in Table 2, with the following CG-like Step 0_b

$$\begin{aligned}
\textbf{Step } 0_b: \ & \text{Set } k = 0,\ y_0 \in \mathbb{R}^n,\ r_0 := b - Ay_0.\\
& \text{If } r_0 = 0, \text{ then STOP. Else, set } p_0 := r_0,\ k = k+1.\\
& \text{Compute } a_0 := r_0^T p_0 / p_0^T A p_0,\\
& y_1 := y_0 + a_0 p_0, \quad r_1 := r_0 - a_0 A p_0.\\
& \text{If } r_1 = 0, \text{ then STOP. Else, set } \sigma_0 := -\|r_1\|^2 / \|r_0\|^2,\\
& p_1 := r_1 + \sigma_0 p_0, \quad k = k+1.
\end{aligned}$$

5 Further Properties for *C D*

In this section, we consider some properties of *C D* which represent a natural extension of analogous properties of the CG. To this purpose, we introduce the *error function*

$$f(y) := \frac{1}{2}(y - y^*)^T A (y - y^*), \qquad \text{with} \quad Ay^* = b, \tag{14}$$

and the quadratic functional

$$g(y) := \frac{1}{2}(y - y_i)^T A (y - y_i), \qquad \text{with} \quad i \in \{1, \ldots, m\}, \tag{15}$$

which satisfy $f(y) \geq 0$, $g(y) \geq 0$, for any $y \in \mathbb{R}^n$, when $A \succeq 0$. Then, we have the following result, where we prove minimization properties of the error function $f(y)$ (see also Theorem 6.1 in [14]) and $g(y)$ (see also [23]), along with the fact that *C D* provides a suitable approximation of the inverse matrix A^{-1}, too.

Theorem 5.1 (Further Properties) *Consider the linear system (1) with $A \succeq 0$, and the functions $f(y)$ and $g(y)$ in (14)–(15). Assume that the C D has performed $m+1$ iterations, with $m+1 \leq n$ and $Ay_{m+1} = b$. Let $\gamma_{i-1} \neq 0$ with $i \geq 1$. Then,*

- σ_0 *minimizes* $g(y)$ *on the manifold* $(y_1 + \gamma_0 A p_0) + \mathrm{span}\{p_0\}$,
- σ_{i-1} *and* ω_{i-1}, $i = 2, \ldots, m$, *minimize* $g(y)$ *on the two dimensional manifold* $(y_i + \gamma_{i-1} A p_{i-1}) + \mathrm{span}\{p_{i-1}, p_{i-2}\}$.

Moreover,

$$f(y_i + a_i p_i) = f(y_i) - \left(\frac{\gamma_{i-1}}{a_{i-1}}\right)^2 \frac{\|r_i\|^4}{p_i^T A p_i}, \qquad i = 1, \ldots, m, \tag{16}$$

and we have

$$\left[A^+ - \sum_{i=0}^{m} \frac{p_i p_i^T}{p_i^T A p_i}\right] r_0 = 0, \qquad \text{for any } y_0 \in \mathbb{R}^n, \tag{17}$$

where A^+ *is the Moore–Penrose pseudoinverse matrix of* A.

Proof Observe that for $i = 1$, indicating in Table 2 $p_1 = \gamma_0 A p_0 + a p_0$, with $a \in \mathbb{R}$, by (15)

$$g(y_2) = g(y_1 + a_1 p_1) = \frac{a_1^2}{2} (\gamma_0 A p_0 + a p_0)^T A (\gamma_0 A p_0 + a p_0),$$

and we have

$$0 = \left.\frac{\partial g(y_2)}{\partial a}\right|_{a=a^*} = a_1^2 p_0^T A (\gamma_0 A p_0 + a^* p_0) \quad \Longleftrightarrow \quad a^* = -\gamma_0 \frac{\|A p_0\|^2}{p_0^T a p_0} = -\sigma_0.$$

For $i \geq 2$, if we indicate in Table 2 $p_i = \gamma_{i-1} A p_{i-1} + b p_{i-1} + c p_{i-2}$, with $b, c \in \mathbb{R}$, then by (15)

$$g(y_i + a_i p_i) = \frac{a_i^2}{2} (\gamma_{i-1} A p_{i-1} + b p_{i-1} + c p_{i-2})^T A (\gamma_{i-1} A p_{i-1} + b p_{i-1} + c p_{i-2}),$$

and by Assumption 4.1, after some computation, the equalities

$$\begin{cases} \left.\frac{\partial g(y_{i+1})}{\partial b}\right|_{b=b^*, c=c^*} = \left.\frac{\partial g(y_i + a_i p_i)}{\partial b}\right|_{b=b^*, c=c^*} = 0 \\ \left.\frac{\partial g(y_{i+1})}{\partial c}\right|_{b=b^*, c=c^*} = \left.\frac{\partial g(y_i + a_i p_i)}{\partial c}\right|_{b=b^*, c=c^*} = 0 \end{cases}$$

imply the unique solution

$$\begin{aligned} b^* &= -\gamma_{i-1} \frac{\|A p_{i-1}\|^2}{p_{i-1}^T A p_{i-1}} = -\sigma_{i-1} \\ c^* &= -\gamma_{i-1} \frac{(A p_{i-1})^T (A p_{i-2})}{p_{i-2}^T A p_{i-2}} = -\frac{\gamma_{i-1}}{\gamma_{i-2}} \frac{p_{i-1}^T A p_{i-1}}{p_{i-2}^T A p_{i-2}} = -\omega_{i-1}. \end{aligned} \tag{18}$$

As regards (16), from Table 2 we have that, for any $i \geq 1$,

$$\begin{aligned} f(y_i + a_i p_i) &= f(y_i) + a_i (y_i - y^*)^T A p_i + \frac{1}{2} a_i^2 p_i^T A p_i \\ &= f(y_i) - a_i r_i^T p_i + \frac{1}{2} a_i^2 p_i^T A p_i \\ &= f(y_i) - \frac{1}{2} \frac{(r_i^T p_i)^2}{p_i^T A p_i}. \end{aligned} \tag{19}$$

Now, since $r_i = r_{i-1} - a_{i-1} A p_{i-1}$ we have

$$\begin{aligned} p_i &= \gamma_{i-1} \left(\frac{r_{i-1} - r_i}{a_{i-1}} \right) - \sigma_{i-1} p_{i-1}, && i = 1, \\ p_i &= \gamma_{i-1} \left(\frac{r_{i-1} - r_i}{a_{i-1}} \right) - \sigma_{i-1} p_{i-1} - \omega_{i-1} p_{i-2}, && i \geq 2, \end{aligned}$$

so that from Theorem 4.2

$$r_i^T p_i = -\frac{\gamma_{i-1}}{a_{i-1}} \|r_i\|^2.$$

The latter relation and (19) yield (16).

As regards (17), since $A y_{m+1} = b$, then $b \in R(A)$, where $R(A)$ is the range of A, and from Table 2 then $r_i \in \mathcal{K}_i(b, A) \subseteq R(A)$, $i = 0, \ldots, m$, where $\mathcal{K}_{i+1}(b, A) \supseteq \mathcal{K}_i(b, A)$. In addition, by the definition of Moore-Penrose pseudoinverse matrix (see [24]), and since y_{m+1} is a solution of (1), we have

$$\begin{aligned} Pr_{R(A)}(y_{m+1}) &= A^+ b = A^+ (r_0 + A y_0) \\ &= A^+ r_0 + Pr_{R(A)}(y_0), \end{aligned} \tag{20}$$

being $Pr_{R(A)}(y_0)$ the projection of y_0 onto $R(A)$. Moreover, we have that $y_{m+1} = y_0 + \sum_{i=0}^{m} a_i p_i$, and by induction $p_i \in \mathcal{K}_i(b, A) \subseteq R(A)$, thus

$$\begin{aligned} Pr_{R(A)}(y_{m+1}) &= Pr_{R(A)}(y_0) + Pr_{R(A)} \left(\sum_{i=0}^{m} a_i p_i \right) \\ &= Pr_{R(A)}(y_0) + \sum_{i=0}^{m} a_i p_i. \end{aligned} \tag{21}$$

By (20), (21) and recalling that for CD we have

$$p_i^T r_i = p_i^T (r_{i-1} - a_{i-1} A p_{i-1}) = p_i^T r_{i-1} = \cdots = p_i^T r_0,$$

we obtain

$$A^+ r_0 = \sum_{i=0}^{m} a_i p_i = \sum_{i=0}^{m} \frac{p_i^T r_i}{p_i^T A p_i} p_i = \sum_{i=0}^{m} \frac{p_i p_i^T}{p_i^T A p_i} r_0,$$

which yields (17). □

Observe that the result in (18) may be seen as a consequence of the Theorem 3.6 in [8], which holds for a general quadratic functional $g(x)$.

Corollary 5.1 (Inverse Approximation) *Let Assumption 4.1 hold and suppose that $Ay_{m+1} = b$, where y_{m+1} is computed by CD and $m = n - 1$. Then, we have*

$$A^{-1} = \sum_{i=0}^{n-1} \frac{p_i p_i^T}{p_i^T A p_i}.$$

Proof The proof follows from (17), recalling that the directions $p_0, \ldots, p_{n-1}$ are linearly independent and when A is nonsingular $A^{-1} \equiv A^+$. □

6 Basic Relation Between the CG and *CD*

Observe that the geometry of vectors $\{p_k\}$ and $\{r_k\}$ in CD might be substantially different with respect to the CG. Indeed, in the latter scheme the relation $p_k = r_k + \beta_{k-1} p_{k-1}$ implies $r_k^T p_k = \|r_k\|^2 > 0$, for any k. On the contrary, for the CD, using relation $r_k = r_{k-1} - a_{k-1} A p_{k-1}$ and Theorem 4.2 we have that possibly $r_k^T p_k \neq \|r_k\|^2$ and

$$\begin{aligned} \frac{p_k^T A p_k}{p_{k-1}^T A p_{k-1}} &= \gamma_{k-1} \frac{(A p_{k-1})^T A p_k}{p_{k-1}^T A p_{k-1}} = -\frac{\gamma_{k-1} \|r_k\|^2}{a_k a_{k-1} p_{k-1}^T A p_{k-1}} \\ &= -\gamma_{k-1} \frac{\|r_k\|^2 p_k^T A p_k}{(r_k^T p_k)(r_{k-1}^T p_{k-1})}, \end{aligned}$$

so that, when $A \succ 0$, we obtain

$$\gamma_{k-1} (r_k^T p_k)(r_{k-1}^T p_{k-1}) < 0. \tag{22}$$

The latter result is a consequence of the fact that in the CD class, the direction p_k is not generated directly using the vector r_k. In addition, a similar conclusion also holds if we compute the quantity $p_k^T p_j > 0$, $k \neq j$, for both the CG and the CD (see also Theorem 5.3 in [14]).

As another difference between the CG and CD, we have that in the first algorithm, the coefficient β_{k-1}, at Step k in Table 1, is always positive. On the other hand, the coefficients γ_{k-1}, σ_{k-1} and ω_{k-1} at Step k of Table 2 might be possibly negative.

We also observe that the CG in Table 1 simply stores at Step k the vectors r_{k-1} and p_{k-1}, in order to compute, respectively, r_k and p_k. On the other hand, at Step k the

CD requires the storage of one additional vector, which contains some information from iteration $k-2$. The idea of storing at Step k some information from iterations preceding Step $k-1$ is not new for Krylov-based methods. Some examples, which differ from our approach, may be found in [7], for unsymmetric linear systems.

In any case, it is not difficult to verify that the CG may be equivalently obtained from CD, setting $\gamma_{k-1} = -\alpha_{k-1}$, for $k = 1, 2, \ldots$, in Table 2. Indeed, though in Table 1 the coefficient β_{k-1} *explicitly* imposes the conjugacy only between p_k and p_{k-1}, the pair $(\alpha_{k-1}, \beta_{k-1})$ implicitly imposes both the conditions (8) for the CG. Now, by (5) and comparing with Step k of Table 2, we want to show that, setting $\gamma_{k-1} = -\alpha_{k-1}$ in Table 2 we obtain

$$\begin{cases} \sigma_{k-1} = -(1+\beta_{k-1}), & k \geq 1, \\ \omega_{k-1} = \beta_{k-2}, & k \geq 2, \end{cases} \tag{23}$$

which implies that CD reduces equivalently to the CG.

For the CG $r_i^T r_j = 0$, for $i \neq j$, and $p_i^T r_i = \|r_i\|^2$, so that

$$\beta_{k-1} := \frac{\|r_k\|^2}{\|r_{k-1}\|^2} = -\frac{r_k^T(\alpha_{k-1} A p_{k-1})}{\|r_{k-1}\|^2} = -\frac{r_k^T A p_{k-1}}{p_{k-1}^T A p_{k-1}}.$$

Thus, recalling that $r_{k-1} = r_{k-2} - \alpha_{k-2} A p_{k-2}$ and $p_{k-1} = r_{k-1} + \beta_{k-2} p_{k-2}$, we obtain for $\gamma_{k-1} = -\alpha_{k-1}$, with $k \geq 2$,

$$\begin{aligned} -(1+\beta_{k-1}) &= -\frac{p_{k-1}^T A p_{k-1} - r_k^T A p_{k-1}}{p_{k-1}^T A p_{k-1}} \\ &= -\frac{(p_{k-1} - r_{k-1} + \alpha_{k-1} A p_{k-1})^T A p_{k-1}}{p_{k-1}^T A p_{k-1}} \\ &= -\alpha_{k-1} \frac{\|A p_{k-1}\|^2}{p_{k-1}^T A p_{k-1}} = \sigma_{k-1} \end{aligned} \tag{24}$$

and

$$\begin{aligned} \beta_{k-2} &= -\frac{r_{k-1}^T A p_{k-2}}{p_{k-2}^T A p_{k-2}} = \frac{\|r_{k-1}\|^2}{\alpha_{k-2}} \frac{1}{p_{k-2}^T A p_{k-2}} \\ &= \frac{\alpha_{k-1}}{\alpha_{k-2}} \frac{p_{k-1}^T A p_{k-1}}{p_{k-2}^T A p_{k-2}} = \omega_{k-1}. \end{aligned} \tag{25}$$

Finally, it is worth noting that for CD, the following two properties hold, for any $k \geq 2$ ((i)–(ii) also hold for $k = 1$, with obvious modifications to (i)):

(i) $r_k^T p_k = r_k^T \left[\gamma_{k-1} \left(\frac{r_{k-1} - r_k}{a_{k-1}} \right) - \sigma_{k-1} p_{k-1} - \omega_{k-1} p_{k-2} \right] = -\frac{\gamma_{k-1}}{a_{k-1}} \|r_k\|^2$

(ii) $r_k^T A p_k = r_k^T \left(\frac{r_k - r_{k+1}}{a_k} \right) = \frac{1}{a_k} \|r_k\|^2 = \frac{\|r_k\|^2}{r_k^T p_k} p_k^T A p_k,$

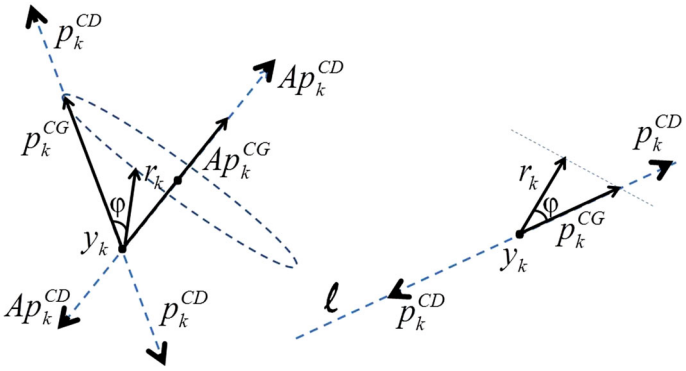

Fig. 1 At the kth iteration of the CG and CD, the directions p_k^{CG} and p_k^{CD} are respectively generated, along the line ℓ. Applying the CG, the vectors p_k^{CG} and r_k have the same orthogonal projection on Ap_k^{CG}, since $(p_k^{CG})^T Ap_k^{CG} = r_k^T Ap_k^{CG}$. Applying CD, the latter equality with p_k^{CD} in place of p_k^{CG} is not necessarily satisfied

which indicate explicitly a difference with respect to the CG. Indeed, for any $\gamma_{k-1} \neq -a_{k-1}$ we have, respectively from (i) and (ii),

$$r_k^T p_k \neq \|r_k\|^2$$
$$r_k^T Ap_k \neq p_k^T Ap_k.$$

Figure 1 clarifies the geometry of items (i) and (ii) for both the CG and CD.

Relations (24)–(25) suggest that the sequence $\{\gamma_k\}$ must satisfy specific conditions in order to reduce CD equivalently to the CG. For a possible generalization of the latter conclusion, consider that equalities (23) are by (5) sufficient conditions in order to reduce CD equivalently to the CG. Thus, now we want to study general conditions on the sequence $\{\gamma_k\}$, such that (23) are satisfied. By (23) we have

$$-(1+\omega_k) = \sigma_{k-1},$$

which is equivalent from Table 2 to

$$-\left(\gamma_{k-1}\|Ap_{k-1}\|^2 + p_{k-1}^T Ap_{k-1}\right) = \frac{\gamma_k}{\gamma_{k-1}} p_k^T Ap_k \tag{26}$$

or

$$-\gamma_{k-1}^2\|Ap_{k-1}\|^2 - \gamma_{k-1} p_{k-1}^T Ap_{k-1} - \gamma_k p_k^T Ap_k = 0. \tag{27}$$

The latter equality, for $k \geq 1$, and the choice of σ_0 in Table 2 yield the following conclusions.

Lemma 6.1 (Reduction of CD) *The scheme CD in Table 2 can be rewritten as in Table 3 (i.e., with the CG-like structure of Table 1), provided that the sequence $\{\gamma_k\}$ satisfies $\gamma_0 := -a_0$ and*

Table 3 The new CD-red class for solving (1), obtained by setting at Step k of CD the parameter γ_k as in relation (28)

The CD-red class

Step 0: Set $k = 0$, $y_0 \in \mathbb{R}^n$, $r_0 := b - Ay_0$.
If $r_0 = 0$, then STOP. Else, set $p_0 := r_0$, $k = k + 1$.
Compute $a_0 := r_0^T p_0 / p_0^T A p_0$, $\gamma_0 := -a_0$,
$y_1 := y_0 + a_0 p_0$, $r_1 := r_0 - a_0 A p_0$.
If $r_1 = 0$, then STOP. Else, set $\sigma_0 := \gamma_0 \|Ap_0\|^2 / p_0^T A p_0$, $\beta_0 = -(1 + \sigma_0)$
$p_1 := r_1 + \beta_0 p_0$, $k = k + 1$.

Step k: Compute $a_{k-1} := r_{k-1}^T p_{k-1} / p_{k-1}^T A p_{k-1}$,
$y_k := y_{k-1} + a_{k-1} p_{k-1}$, $r_k := r_{k-1} - a_{k-1} A p_{k-1}$.
If $r_k = 0$, then STOP. Else, use (28) to compute γ_{k-1}.
Set $\sigma_{k-1} := \gamma_{k-1} \frac{\|Ap_{k-1}\|^2}{p_{k-1}^T A p_{k-1}}$, $\beta_{k-1} := -(1 + \sigma_{k-1})$
$p_k := r_k + \beta_{k-1} p_{k-1}$, $k = k + 1$.
Go to **Step** k.

$$\gamma_k := -\frac{\gamma_{k-1}^2 \|Ap_{k-1}\|^2 + \gamma_{k-1} p_{k-1}^T A p_{k-1}}{p_k^T A p_k}, \qquad k \geq 1. \tag{28}$$

In particular, the positions $\gamma_i = -a_i$, $i \geq 0$, in CD satisfy (28).

Proof By the considerations which led to (26)–(27), relation (28) yields (23), so that the scheme CD-red in Table 3 follows from CD with the position (28), and setting $\gamma_0 = -a_0$.
Furthermore, replacing in (28) the conditions $\gamma_i = -a_i$, $i \geq 1$, and recalling (i)–(ii), we obtain the condition $a_{k-1}^2 \|Ap_{k-1}\|^2 = \|r_{k-1}\|^2 + \|r_k\|^2$, which is immediately fulfilled using condition $r_k = r_{k-1} - a_{k-1} A p_{k-1}$. □

Note that the CD-red scheme substantially is more similar to the CG than to CD. Indeed, the conditions (8), explicitly imposed at Step k of CD, reduce to the unique condition $p_k^T A p_{k-1} = 0$ in CD-red.

The following result is a trivial consequence of Lemma 4.3, where the alternate use of CG and CD steps is analyzed.

Lemma 6.2 (Combined Finite Convergence) *Let Assumption 4.1 hold. Let $y_1, \ldots, y_h$ be the iterates generated by CD, with $h \leq n$ and $Ay_h = b$. Then, finite convergence is preserved (i.e., $Ay_h = b$) if the Step $\hat{k}$ of CD, with $\hat{k} \in \{k_1, \ldots, k_h\} \subseteq \{1, \ldots, h\}$, is replaced by the Step $\hat{k}$ of the CG.*

Proof First observe that both in Tables 1 and 2, for any $k \leq h$, the quantity $\|r_k\| > 0$ is computed. Thus, in Table 1 the coefficient β_{k-1} is well defined for any $n > k \geq 1$. Now, by Table 2, note that if we set at Step $\hat{k} \in \{k_1, \ldots, k_h\} \subseteq \{1, \ldots, h\}$ the following

Table 4 The scaled-CG algorithm for solving (1)

The Scaled-CG method

Step 0: Set $k = 0$, $y_0 \in \mathbb{R}$, $r_0 := b - Ay_0$.
If $r_0 = 0$, then STOP. Else, set $p_0 := \rho_0 r_0$, $\rho_0 > 0$, $k = k + 1$.

Step k: Compute $\alpha_{k-1} := \rho_{k-1}\|r_{k-1}\|^2 / p_{k-1}^T A p_{k-1}$, $\rho_{k-1} > 0$,
$y_k := y_{k-1} + \alpha_{k-1} p_{k-1}$, $r_k := r_{k-1} - \alpha_{k-1} A p_{k-1}$.
If $r_k = 0$, then STOP. Else, set $\beta_{k-1} := -p_{k-1}^T A r_k / p_{k-1}^T A p_{k-1}$ or
$\beta_{k-1} := \|r_k\|^2 / (\rho_{k-1}\|r_{k-1}\|^2)$
$p_k := \rho_k (r_k + \beta_{k-1} p_{k-1})$, $\rho_k > 0$, $k = k + 1$,
Go to **Step** k.

$$\begin{cases} \gamma_{\hat{k}-1} = -a_{\hat{k}-1}, & \text{if } \hat{k} \geq 1 \\ \sigma_{\hat{k}-1} = -(1 + \beta_{\hat{k}-1}), & \text{if } \hat{k} \geq 1 \\ \omega_{\hat{k}-1} = \beta_{\hat{k}-2}, & \text{if } \hat{k} \geq 2, \end{cases}$$

the Step $\hat{k}$ of CD coincides formally with the Step $\hat{k}$ of CG. Thus, finite convergence with $Ay_h = b$ is proved recalling that Lemma 4.3 holds for any choice of the sequence $\{\gamma_k\}$, with $\gamma_k \neq 0$. □

7 Relation Between the Scaled-CG and *CD*

Similarly to the previous section, here we aim at determining the relation between our proposal in Table 2 and the scheme of the *scaled*-CG in Table 4 (see also [8], page 125). In [8] a motivated choice for the coefficients $\{\rho_k\}$ in the scaled-CG is also given. Here, following the guidelines of the previous section, we first rewrite the relation

$$p_{k+1} := \rho_{k+1}(r_{k+1} + \beta_k p_k),$$

at Step $k + 1$ of the scaled-CG, as follows

$$\begin{aligned} p_{k+1} &= \rho_{k+1}(r_k - \alpha_k A p_k) + \rho_{k+1}\beta_k p_k \\ &= \rho_{k+1}\left[\frac{p_k}{\rho_k} - \beta_{k-1} p_{k-1} - \alpha_k A p_k\right] + \rho_{k+1}\beta_k p_k \\ &= -\rho_{k+1}\alpha_k A p_k + \rho_{k+1}\left(\beta_k + \frac{1}{\rho_k}\right) p_k - \rho_{k+1}\beta_{k-1} p_{k-1}. \end{aligned} \tag{29}$$

We want to show that for a suitable choice of the parameters $\{\gamma_k\}$, the CD yields the recursion (29) of the scaled-CG, i.e., for a proper choice of $\{\gamma_k\}$ we obtain from CD a scheme equivalent to the scaled-CG. On this purpose, let us set in CD

$$\gamma_k = -\rho_{k+1}\alpha_k, \qquad k \geq 0, \tag{30}$$

where α_k is given at Step k of Table 4. Thus, by Table 2

$$\sigma_k = \gamma_k \frac{\|Ap_k\|^2}{p_k^T Ap_k} = -\rho_{k+1}\alpha_k \frac{\|Ap_k\|^2}{p_k^T Ap_k}, \qquad k \geq 0, \tag{31}$$

and for $k \geq 1$

$$\omega_k = \frac{\gamma_k}{\gamma_{k-1}} \frac{p_k^T Ap_k}{p_{k-1}^T Ap_{k-1}} = \frac{\rho_{k+1}\alpha_k}{\rho_k \alpha_{k-1}} \frac{p_k^T Ap_k}{p_{k-1}^T Ap_{k-1}}. \tag{32}$$

Now, comparing the coefficients in (29) with (30), (31) and (32), we want to prove that the choice (30) implies

$$\sigma_k = -\rho_{k+1}\left(\beta_k + \frac{1}{\rho_k}\right), \qquad k \geq 0, \tag{33}$$

$$\omega_k = \rho_{k+1}\beta_{k-1}, \qquad k \geq 1, \tag{34}$$

so that the CD class yields equivalently the scaled-CG.

As regards (33), from Table 4 we have, for $k \geq 0$

$$\begin{aligned}
\beta_k + \frac{1}{\rho_k} &= \frac{\frac{1}{\rho_k} p_k^T Ap_k - r_{k+1}^T Ap_k}{p_k^T Ap_k} = \frac{\left(\frac{1}{\rho_k} p_k - r_{k+1}\right)^T Ap_k}{p_k^T Ap_k} \\
&= \frac{\left(\frac{1}{\rho_k} p_k - r_k + \alpha_k Ap_k\right)^T Ap_k}{p_k^T Ap_k} \\
&= \frac{(r_k + \beta_{k-1}p_{k-1} - r_k + \alpha_k Ap_k)^T Ap_k}{p_k^T Ap_k} = \alpha_k \frac{\|Ap_k\|^2}{p_k^T Ap_k},
\end{aligned}$$

so that from (31) the condition (33) holds, for any $k \geq 0$. As regards (34) from Step k of Table 4 we know that $\beta_{k-1} = \|r_k\|^2/(\rho_{k-1}\|r_{k-1}\|^2)$ and, since $r_k^T p_{k-1} = 0$, we obtain $r_k^T p_k = \rho_k \|r_k\|^2$; thus, relation (30) yields

$$\beta_{k-1} = \frac{\|r_k\|^2}{\rho_{k-1}\|r_{k-1}\|^2} = \frac{\alpha_k}{\rho_k \alpha_{k-1}} \frac{p_k^T Ap_k}{p_{k-1}^T Ap_{k-1}} = \frac{\gamma_k}{\rho_{k+1}\gamma_{k-1}} \frac{p_k^T Ap_k}{p_{k-1}^T Ap_{k-1}}, k \geq 1.$$

Relation (34) is proved using the latter equality and (32).

8 Matrix Factorization Induced by CD

We first recall that considering the CG in Table 1 and setting at Step h

$$P_h := \left(\frac{p_0}{\|r_0\|} \cdots \frac{p_h}{\|r_h\|} \right)$$
$$R_h := \left(\frac{r_0}{\|r_0\|} \cdots \frac{r_h}{\|r_h\|} \right),$$

along with

$$L_h := \begin{pmatrix} 1 & & & & \\ -\sqrt{\beta_0} & 1 & & & \\ & -\sqrt{\beta_1} & 1 & & \\ & & \ddots & 1 & \\ & & & -\sqrt{\beta_{h-1}} & 1 \end{pmatrix} \in \mathbb{R}^{h \times h}$$

and $D_h := \mathrm{diag}_i\{1/\alpha_i\}$, we obtain the three matrix relations

$$P_h L_h^T = R_h \tag{35}$$
$$A P_h = R_h L_h D_h - \frac{\sqrt{\beta_h}}{\alpha_h} \frac{r_{h+1}}{\|r_{h+1}\|} e_h^T \tag{36}$$
$$R_h^T A R_h = T_h = L_h D_h L_h^T. \tag{37}$$

Then, in this section we are going to use the iteration in Table 2 in order to possibly recast relations (35)–(37) for CD.

On this purpose, from Table 2 we can easily draw the following relation between the sequences $\{p_0, p_1, \ldots\}$ and $\{r_0, r_1, \ldots\}$

$$\begin{aligned} p_0 &= r_0 \\ p_1 &= \frac{\gamma_0}{a_0}(r_0 - r_1) - \sigma_0 p_0 \\ p_i &= \frac{\gamma_{i-1}}{a_{i-1}}(r_{i-1} - r_i) - \sigma_{i-1} p_{i-1} - \omega_{i-1} p_{i-2}, \qquad i = 2, 3, \ldots, \end{aligned}$$

and introducing the positions

$$\begin{aligned} P_h &:= (p_0 p_1 \cdots p_h) \\ R_h &:= (r_0 r_1 \cdots r_h) \\ \bar{R}_h &:= \left(\frac{r_0}{\|r_0\|} \cdots \frac{r_h}{\|r_h\|} \right), \end{aligned}$$

along with the matrices

$$U_{h,1} := \begin{pmatrix} 1 & \sigma_0 & \omega_1 & 0 & \cdots & \cdots & 0 \\ & 1 & \sigma_1 & \omega_2 & 0 & \cdots & 0 \\ & & 1 & \sigma_2 & \ddots & 0 & \vdots \\ & & & 1 & \ddots & \ddots & 0 \\ & & & & \ddots & \ddots & \omega_{h-1} \\ & & & & & \ddots & \sigma_{h-1} \\ & & & & & & 1 \end{pmatrix} \in \mathbb{R}^{(h+1)\times(h+1)},$$

$$U_{h,2} := \begin{pmatrix} \|r_0\| & \|r_0\| & 0 & \cdots & \cdots & 0 \\ & -\|r_1\| & \|r_1\| & 0 & \cdots & 0 \\ & & -\|r_2\| & \|r_2\| & 0 & \vdots \\ & & & \ddots & \ddots & 0 \\ & & & & -\|r_{h-1}\| & \|r_{h-1}\| \\ & & & & & -\|r_h\| \end{pmatrix} \in \mathbb{R}^{(h+1)\times(h+1)}$$

and

$$D_h := \operatorname{diag}\left\{ 1, \operatorname*{diag}_{i=0,\ldots,h-1} \{\gamma_i / a_i\} \right\} \in \mathbb{R}^{(h+1)\times(h+1)},$$

we obtain after $h-1$ iterations of CD

$$P_h U_{h,1} = \bar{R}_h U_{h,2} D_h,$$

so that

$$P_h = \bar{R}_h U_{h,2} D_h U_{h,1}^{-1} = \bar{R}_h U_h,$$

where $U_h = U_{h,2} D_h U_{h,1}^{-1}$. Now, observe that U_h is upper triangular since $U_{h,2}$ is upper bidiagonal, D_h is diagonal and $U_{h,1}^{-1}$ may be easily seen to be upper triangular. As a

consequence, recalling that $p_0, \ldots, p_h$ are mutually conjugate we have

$$\bar{R}_h^T A \bar{R}_h = U_h^{-T} \mathrm{diag}_i\{p_i^T A p_i\} U_h^{-1},$$

and in case $h = n-1$, again from the conjugacy of $p_0, \ldots, p_{n-1}$

$$P_{n-1}^T A P_{n-1} = U_{n-1}^T \bar{R}_{n-1}^T A \bar{R}_{n-1} U_{n-1} = \underset{i=0,\ldots,h-1}{\mathrm{diag}} \{p_i^T A p_i\}.$$

From the orthogonality of $\bar{R}_{n-1}$, along with relation

$$\det(U_{n-1}) = \|r_0\| \prod_{j=1}^{n-1} \left(-\frac{\|r_j\| \gamma_{j-1}}{a_{j-1}} \right) = \left(\prod_{i=0}^{n-1} \|r_i\| \right) \left(\prod_{i=0}^{n-2} -\frac{\gamma_i}{a_i} \right),$$

we have

$$\det\left(U_{n-1}^T \bar{R}_{n-1}^T A \bar{R}_{n-1} U_{n-1} \right) = \prod_{i=0}^{n-1} p_i^T A p_i \quad \Longleftrightarrow \quad \det(A) = \frac{\prod_{i=0}^{n-1} p_i^T A p_i}{\det(U_{n-1})^2}.$$

Thus, in the end

$$\det(A) = \left[\prod_{i=0}^{n-1} \frac{p_i^T A p_i}{\|r_i\|^2} \right] \cdot \frac{\left[\prod_{i=0}^{n-2} a_i^2 \right]}{\left[\prod_{i=0}^{n-2} \gamma_i^2 \right]}. \tag{38}$$

Note that the following considerations hold:

– for $\gamma_i = \pm a_i$ (which includes the case $\gamma_i = -a_i$, when by Lemma 6.1 CD reduces equivalently to the CG), by (i) of Sect. 6 $|p_k^T r_k| = \|r_k\|^2$, so that we obtain the standard result (see also [14])

$$\det(A) = \left[\prod_{i=0}^{n-1} \frac{p_i^T A p_i}{\|r_i\|^2} \right] = \prod_{i=0}^{n-1} \frac{1}{a_i};$$

– if $|\gamma_i| \neq |a_i|$ we obtain the general formula (38).

9 Issues on the Conjugacy Loss for CD

Here we consider a simplified approach to describe the conjugacy loss for both the CG and CD, under Assumption 4.1 (see also [14] for a similar approach). Suppose

that both the CG and CD perform Step $k+1$, and for numerical reasons, a nonzero *conjugacy error* $\varepsilon_{k,j}$ respectively occurs between directions p_k and p_j, i.e.,

$$\varepsilon_{k,j} := p_k^T A p_j \neq 0, \qquad j \leq k-1.$$

Then, we calculate the conjugacy error

$$\varepsilon_{k+1,j} = p_{k+1}^T A p_j, \qquad j \leq k,$$

for both the CG and CD. First observe that at Step $k+1$ of Table 1 we have

$$\begin{aligned} \varepsilon_{k+1,j} &= (r_{k+1} + \beta_k p_k)^T A p_j && (39)\\ &= (p_k - \beta_{k-1} p_{k-1} - \alpha_k A p_k)^T A p_j + \beta_k \varepsilon_{k,j} && (40)\\ &= (1+\beta_k)\varepsilon_{k,j} - \beta_{k-1}\varepsilon_{k-1,j} - \alpha_k (A p_k)^T A p_j. && (41) \end{aligned}$$

Then, from relation $Ap_j = (r_j - r_{j+1})/\alpha_j$ and relations (2)–(3) we have for the CG

$$(Ap_k)^T A p_j = \begin{cases} -\frac{p_k^T A p_k}{\alpha_{k-1}}, & j = k-1, \\ \emptyset, & j \leq k-2. \end{cases}$$

Thus, observing that for the CG we have $\varepsilon_{i,i-1} = 0$ and $\varepsilon_{i,i} = p_i^T A p_i$, for $1 \leq i \leq k+1$, after some computation we obtain from (2), (3) and (41)

$$\varepsilon_{k+1,j} = \begin{cases} \emptyset, & j = k, \\ \emptyset, & j = k-1, \\ (1+\beta_k)\varepsilon_{k,k-2}, & j = k-2, \\ (1+\beta_k)\varepsilon_{k,j} - \beta_{k-1}\varepsilon_{k-1,j} - \Sigma_{kj}, & j \leq k-3, \end{cases} \tag{42}$$

where $\Sigma_{kj} \in \mathbb{R}$ summarizes the contribution of the term $\alpha_k (Ap_k)^T A p_j$, due to a possible conjugacy loss.

Let us consider now for CD a result similar to (42). We obtain the following relations for $j \leq k$

$$\begin{aligned} \varepsilon_{k+1,j} &= p_{k+1}^T A p_j = (\gamma_k A p_k - \sigma_k p_k - \omega_k p_{k-1})^T A p_j \\ &= \gamma_k (Ap_k)^T A p_j - \sigma_k \varepsilon_{k,j} - \omega_k \varepsilon_{k-1,j} \\ &= \frac{\gamma_k}{\gamma_j} (Ap_k)^T \left(p_{j+1} + \sigma_j p_j + \omega_j p_{j-1} \right) - \sigma_k \varepsilon_{k,j} - \omega_k \varepsilon_{k-1,j} \\ &= \frac{\gamma_k}{\gamma_j} \varepsilon_{k,j+1} + \left(\frac{\gamma_k}{\gamma_j} \sigma_j - \sigma_k \right) \varepsilon_{k,j} + \frac{\gamma_k}{\gamma_j} \omega_j \varepsilon_{k,j-1} - \omega_k \varepsilon_{k-1,j}, \end{aligned}$$

and considering now relations (8), the conjugacy among directions $p_0, p_1, \ldots, p_k$ satisfies

$$\varepsilon_{h,l} = p_h^T A p_l = 0, \qquad \text{for any } \mid h - l \mid \in \{1, 2\}. \tag{43}$$

Thus, relation (10) and the expression of the coefficients in CD yields for $\varepsilon_{k+1,j}$ the expression

$$\begin{cases} \emptyset, & j = k, \\ \emptyset, & j = k-1, \\ \frac{\gamma_k}{\gamma_{k-2}}\omega_{k-2}\varepsilon_{k,k-3}, & j = k-2, \\ \left(\frac{\gamma_k}{\gamma_{k-3}}\sigma_{k-3} - \sigma_k\right)\varepsilon_{k,k-3} + \frac{\gamma_k}{\gamma_{k-3}}\omega_{k-3}\varepsilon_{k,k-4}, & j = k-3, \\ \frac{\gamma_k}{\gamma_j}\varepsilon_{k,j+1} + \left(\frac{\gamma_k}{\gamma_j}\sigma_j - \sigma_k\right)\varepsilon_{k,j} + \frac{\gamma_k}{\gamma_j}\omega_j\varepsilon_{k,j-1} - \omega_k\varepsilon_{k-1,j}, & j \le k-4. \end{cases} \tag{44}$$

Finally, comparing relations (42) and (44) we have

– in case $j = k-2$, the conjugacy error $\varepsilon_{k+1,k-2}$ is nonzero for both the CG and CD, as expected. However, for the CG

$$|\varepsilon_{k+1,k-2}| \; > \; |\varepsilon_{k,k-2}|$$

since $(1 + \beta_k) > 1$, which theoretically can lead to an harmful amplification of conjugacy errors. On the contrary, for CD the positive quantity $|\gamma_k\omega_{k-2}/\gamma_{k-2}|$ in the expression of $\varepsilon_{k+1,k-2}$ can be possibly smaller than one.

– choosing the sequence $\{\gamma_k\}$ such that

$$\left|\frac{\gamma_k}{\gamma_{k-i}}\right| \ll 1 \quad \text{and/or} \quad \left|\frac{\gamma_k}{\gamma_{k-i}}\omega_{k-i}\right| \ll 1, \qquad i = 2, 3, \ldots \tag{45}$$

from (44) the effects of conjugacy loss may be attenuated. Thus, a strategy to update the sequence $\{\gamma_k\}$ so that (45) holds might be investigated.

9.1 Bounds for the Coefficients of CD

We want to describe here the sensitivity of the coefficients σ_k and ω_k, at Step $k+1$ of CD, to the condition number $\kappa(A)$. In particular, we want to provide a comparison with the CG, in order to identify possible advantages/disadvantages of our proposal. From Table 2 and Assumption 4.1 we have

$$|\omega_k| = \left|\frac{\gamma_k}{\gamma_{k-1}}\frac{p_k^T A p_k}{p_{k-1}^T A p_{k-1}}\right|, \qquad |\sigma_k| = \left|\gamma_k \frac{\|Ap_k\|^2}{p_k^T A p_k}\right|,$$

so that (we indicate with $\lambda_m(A)$ and $\lambda_M(A)$ the smallest/largest eigenvalue of matrix A)

$$\begin{cases} |\omega_k| \geq \left|\dfrac{\gamma_k}{\gamma_{k-1}}\right| \dfrac{\lambda_m(A)\|p_k\|^2}{\lambda_M(A)\|p_{k-1}\|^2} = \left|\dfrac{\gamma_k}{\gamma_{k-1}}\right| \dfrac{1}{\kappa(A)} \dfrac{\|p_k\|^2}{\|p_{k-1}\|^2} \\ |\omega_k| \leq \left|\dfrac{\gamma_k}{\gamma_{k-1}}\right| \dfrac{\lambda_M(A)\|p_k\|^2}{\lambda_m(A)\|p_{k-1}\|^2} = \left|\dfrac{\gamma_k}{\gamma_{k-1}}\right| \kappa(A) \dfrac{\|p_k\|^2}{\|p_{k-1}\|^2}, \end{cases} \tag{46}$$

and

$$\begin{cases} |\sigma_k| \geq |\gamma_k| \dfrac{\lambda_m^2(A)\|p_k\|^2}{\lambda_M(A)\|p_k\|^2} = |\gamma_k| \dfrac{\lambda_m(A)}{\kappa(A)} \\ |\sigma_k| \leq |\gamma_k| \dfrac{\lambda_M^2(A)\|p_k\|^2}{\lambda_m(A)\|p_k\|^2} = |\gamma_k| \lambda_M(A)\kappa(A). \end{cases} \tag{47}$$

On the other hand, from Table 1 we obtain for the CG

$$\beta_k = -\frac{r_{k+1}^T A p_k}{p_k^T A p_k} = -1 + \alpha_k \frac{\|Ap_k\|^2}{p_k^T A p_k} = -1 + \frac{\|r_k\|^2}{p_k^T A p_k} \frac{\|Ap_k\|^2}{p_k^T A p_k},$$

so that, since $\beta_k > 0$ and using relation $\|r_k\| \leq \|p_k\|$, along with relation $p_k^T A p_k = r_k^T A r_k - \frac{\|r_k\|^4}{\|r_{k-1}\|^4} p_{k-1}^T A p_{k-1} > 0$, we have

$$\begin{cases} \beta_k \geq \max\left\{0, -1 + \frac{\|r_k\|^2}{r_k^T A r_k} \frac{\lambda_m(A)}{\kappa(A)}\right\} \geq \max\left\{0, -1 + \frac{1}{[\kappa(A)]^2}\right\} = 0 \\ \beta_k \leq -1 + \frac{\|p_k\|^2}{p_k^T A p_k} \lambda_M(A)\kappa(A) \leq -1 + [\kappa(A)]^2. \end{cases} \tag{48}$$

In particular, this seems to indicate that, on those problems where the quantity $|\gamma_k|\lambda_M(A)$ is reasonably small, CD might be competitive. However, as expected, high values for $\kappa(A)$ may determine numerical instability for both the CG and CD. In addition, observe that any conclusion on the comparison between the numerical performance of the CG and CD, depends both on the sequence $\{\gamma_k\}$ and on how tight are the bounds (47) and (48) for the problem in hand.

10 The Preconditioned CD Class

In this section, we introduce preconditioning for the class CD, in order to better cope with possible illconditioning of the matrix A in (1).

Let $M \in \mathbb{R}^{n\times n}$ be *nonsingular* and consider the linear system (1). Since we have

$$Ay = b \Longleftrightarrow \left(M^T M\right)^{-1} Ay = \left(M^T M\right)^{-1} b \tag{49}$$

$$\Longleftrightarrow \left(M^{-T} A M^{-1}\right) My = M^{-T} b$$

$$\Longleftrightarrow \bar{A}\bar{y} = \bar{b}, \tag{50}$$

where

$$\bar{A} := M^{-T} A M^{-1}, \qquad \bar{y} := My, \qquad \bar{b} := M^{-T} b, \tag{51}$$

solving (1) is equivalent to solve (49) or (50). Moreover, any eigenvalue λ_i, $i = 1, \ldots, n$, of $M^{-T} A M^{-1}$ is also an eigenvalue of $\left(M^T M\right)^{-1} A$. Indeed, if $(M^T M)^{-1} A z_i = \lambda_i z_i$, $i = 1, \ldots, n$, then

$$\left(M^{-1} M^{-T}\right) A M^{-1} (M z_i) = \lambda_i z_i$$

so that

$$M^{-T} A M^{-1} (M z_i) = \lambda_i (M z_i).$$

Now, let us motivate the importance of selecting a promising matrix M in (50), in order to reduce $\kappa(\bar{A})$ (or equivalently to reduce $\kappa[(M^T M)^{-1} A]$).

Observe that under the Assumption 4.1 and using standard Chebyshev polynomials analysis, we can prove that in exact algebra, for both the CG and CD, the following relation holds (see [2] for details, and a similar analysis holds for CD)

$$\frac{\|y_k - y^*\|_A}{\|y_0 - y^*\|_A} \leq 2 \left(\frac{\sqrt{\kappa(A)} - 1}{\sqrt{\kappa(A)} + 1}\right)^k, \tag{52}$$

where $Ay^* = b$ and $\|v\|_A^2 = v^T A v$, for any $v \in \mathbb{R}^n$. Relation (52) reveals the strong dependency of the iterates generated by the CG and CD, on $\kappa(A)$. In addition, if the CG and CD are used to solve (50) in place of (1), then the bound (52) becomes

$$\frac{\|y_k - y^*\|_A}{\|y_0 - y^*\|_A} \leq 2 \left(\frac{\sqrt{\kappa[(M^T M)^{-1} A]} - 1}{\sqrt{\kappa[(M^T M)^{-1} A]} + 1}\right)^k, \tag{53}$$

which definitely encourages to use the *preconditioner* $(M^T M)^{-1}$ whenever we have $\kappa[(M^T M)^{-1} A] < \kappa(A)$.

On this guideline we want to introduce preconditioning in our scheme CD, for solving the linear system (50), where M is non-singular. We do not expect that necessarily, when $M = I$ (i.e., no preconditioning is considered in (50)), CD outperforms the CG. Indeed, as stated in the previous section, $M = I$ along with bounds (46), (47)

and (48) do not suggest a specific preference for CD with respect to the CG. On the contrary, suppose a suitable preconditioner $\mathcal{M} = (M^T M)^{-1}$ is selected when $\kappa(A)$ is large. Then, since the class CD for suitable values of γ_{k-1} at Step k possibly imposes stronger conjugacy conditions with respect to the CG, it may possibly better recover the conjugacy loss.

We will soon see that, if the preconditioner $\mathcal{M}$ is adopted in CD, it is just used throughout the computation of the product $\mathcal{M} \times v$, $v \in \mathbb{R}^n$, i.e., it is not necessary to store the possibly dense matrix $\mathcal{M}$.

The algorithms in CD for (50) are described in Table 5, where each 'bar' quantity has a corresponding quantity in Table 2. Then, after substituting in Table 5 the positions

$$\begin{aligned} \bar{y}_k &:= M y_k \\ \bar{p}_k &:= M p_k \\ \bar{r}_k &:= M^{-T} r_k \\ \mathcal{M} &:= \left(M^T M\right)^{-1}, \end{aligned} \tag{54}$$

the vector $\bar{p}_k$ becomes

$$\bar{p}_k = M p_k = \bar{\gamma}_{k-1} M^{-T} A M^{-1} M p_{k-1} - \bar{\sigma}_{k-1} M p_{k-1} - \bar{\omega}_{k-1} M p_{k-2},$$

hence

$$p_k = \bar{\gamma}_{k-1} \mathcal{M} A p_{k-1} - \bar{\sigma}_{k-1} p_{k-1} - \bar{\omega}_{k-1} p_{k-2}$$

Table 5 The CD class for solving the linear system $\bar{A}\bar{y} = \bar{b}$ in (50)

The CD class for (50)

Step 0: Set $k = 0$, $\bar{y}_0 \in \mathbb{R}^n$, $\bar{r}_0 := \bar{b} - \bar{A}\bar{y}_0$, $\bar{\gamma}_0 \in \mathbb{R} \setminus \{0\}$.
If $\bar{r}_0 = 0$, then STOP. Else, set $\bar{p}_0 := \bar{r}_0$, $k = k + 1$.
Compute $\bar{a}_0 := \bar{r}_0^T \bar{p}_0 / \bar{p}_0^T \bar{A} \bar{p}_0$,
$\bar{y}_1 := \bar{y}_0 + \bar{a}_0 \bar{p}_0$, $\bar{r}_1 := \bar{r}_0 - \bar{a}_0 \bar{A} \bar{p}_0$.
If $\bar{r}_1 = 0$, then STOP. Else, set $\bar{\sigma}_0 := \bar{\gamma}_0 \|\bar{A}\bar{p}_0\|^2 / \bar{p}_0^T \bar{A} \bar{p}_0$,
$\bar{p}_1 := \bar{\gamma}_0 \bar{A} \bar{p}_0 - \bar{\sigma}_0 \bar{p}_0$, $k = k + 1$.

Step k: Compute $\bar{a}_{k-1} := \bar{r}_{k-1}^T \bar{p}_{k-1} / \bar{p}_{k-1}^T \bar{A} \bar{p}_{k-1}$, $\bar{\gamma}_{k-1} \in \mathbb{R} \setminus \{0\}$,
$\bar{y}_k := \bar{y}_{k-1} + \bar{a}_{k-1} \bar{p}_{k-1}$, $\bar{r}_k := \bar{r}_{k-1} - \bar{a}_{k-1} \bar{A} \bar{p}_{k-1}$.
If $\bar{r}_k = 0$, then STOP. Else, set
$\bar{\sigma}_{k-1} := \bar{\gamma}_{k-1} \frac{\|\bar{A}\bar{p}_{k-1}\|^2}{\bar{p}_{k-1}^T \bar{A} \bar{p}_{k-1}}$, $\bar{\omega}_{k-1} := \frac{\bar{\gamma}_{k-1}}{\bar{\gamma}_{k-2}} \frac{\bar{p}_{k-1}^T \bar{A} \bar{p}_{k-1}}{\bar{p}_{k-2}^T \bar{A} \bar{p}_{k-2}}$,
$\bar{p}_k := \bar{\gamma}_{k-1} \bar{A} \bar{p}_{k-1} - \bar{\sigma}_{k-1} \bar{p}_{k-1} - \bar{\omega}_{k-1} \bar{p}_{k-2}$, $k = k + 1$.
Go to **Step** k.

with

$$\bar{\sigma}_{k-1} = \bar{\gamma}_{k-1}\frac{\|M^{-T}Ap_{k-1}\|^2}{p_{k-1}^T Ap_{k-1}} = \bar{\gamma}_{k-1}\frac{(Ap_{k-1})^T\mathcal{M}Ap_{k-1}}{p_{k-1}^T Ap_{k-1}} \tag{55}$$

$$\bar{\omega}_{k-1} = \frac{\bar{\gamma}_{k-1}}{\bar{\gamma}_{k-2}}\frac{p_{k-1}^T M^T M^{-T}AM^{-1}Mp_{k-1}}{p_{k-2}^T M^T M^{-T}AM^{-1}Mp_{k-2}} = \frac{\bar{\gamma}_{k-1}}{\bar{\gamma}_{k-2}}\frac{p_{k-1}^T Ap_{k-1}}{p_{k-2}^T Ap_{k-2}}.$$

Moreover, relation $\bar{r}_0 = \bar{b} - \bar{A}\bar{y}_0$ becomes

$$M^{-T}r_0 = M^{-T}b - M^{-T}AM^{-1}My_0 \qquad \Longleftrightarrow \qquad r_0 = b - Ay_0,$$

and since $\bar{p}_0 = Mp_0 = \bar{r}_0 = M^{-T}r_0$, then $p_0 = \mathcal{M}r_0$, so that the coefficients $\bar{\sigma}_0$ and $\bar{a}_0$ become

$$\begin{aligned}\bar{\sigma}_0 &= \bar{\gamma}_0\frac{p_0^T M^T M^{-T}AM^{-1}M^{-T}AM^{-1}Mp_0}{p_0^T Ap_0} = \bar{\gamma}_0\frac{(Ap_0)^T\mathcal{M}(Ap_0)}{p_0^T Ap_0}\\ &= \bar{\gamma}_0\frac{\|Ap_0\|_{\mathcal{M}}^2}{p_0^T Ap_0}\\ \bar{a}_0 &= \frac{r_0^T M^{-1}Mp_0}{p_0^T M^T M^{-T}AM^{-1}Mp_0} = \frac{r_0^T p_0}{p_0^T Ap_0}.\end{aligned} \tag{56}$$

As regards relation $\bar{p}_1 = \bar{\gamma}_0\bar{A}\bar{p}_0 - \bar{\sigma}_0\bar{p}_0$, we have

$$Mp_1 = \bar{\gamma}_0 M^{-T}AM^{-1}Mp_0 - \bar{\sigma}_0 Mp_0,$$

hence

$$p_1 = \bar{\gamma}_0\mathcal{M}Ap_0 - \bar{\sigma}_0 p_0.$$

Finally, $\bar{r}_k = M^{-T}r_k$ so that

$$\bar{r}_k = M^{-T}r_k = M^{-T}r_{k-1} - \bar{a}_{k-1}M^{-T}AM^{-1}Mp_{k-1}$$

and therefore

$$r_k = r_{k-1} - \bar{a}_{k-1}Ap_{k-1},$$

with

$$\bar{a}_{k-1} = \frac{r_{k-1}^T M^{-1}Mp_{k-1}}{p_{k-1}^T M^T M^{-T}AM^{-1}Mp_{k-1}} = \frac{r_{k-1}^T p_{k-1}}{p_{k-1}^T Ap_{k-1}}.$$

The overall resulting preconditioned algorithm $CD_{\mathcal{M}}$ is detailed in Table 6. Observe that the coefficients a_{k-1} and ω_{k-1} in Tables 2 and 6 are invariant under the introduction of the preconditioner $\mathcal{M}$. Also note that from (55) and (56), now in $CD_{\mathcal{M}}$ the coefficient σ_{k-1} depends on $A\mathcal{M}A$ and not on A^2 (as in Table 2).

Moreover, in Table 6 the introduction of the preconditioner simply requires at Step k the additional cost of the product $\mathcal{M} \times (Ap_{k-1})$ (similarly to the preconditioned CG, where at iteration k the additional cost of preconditioning is given by $\mathcal{M} \times r_{k-1}$).

Furthermore, in Table 6 at Step 0 the products $\mathcal{M}r_0$ and $\mathcal{M}(Ap_0)$ are both required, in order to compute σ_0 and a_0. Considering that Step 0 of CD is equivalent to two iterations of the CG, then the cost of preconditioning either CG or CD is the same. Finally, similar results hold if $CD_{\mathcal{M}}$ is recast in view of Remark 4.1.

11 Numerical Examples

The theory in Sects. 5–9 seems to provide yet premature criteria, for a fruitful choice of the sequence $\{\gamma_k\}$ on applications. Furthermore, we do not have clear ideas about the real importance of the scheme CD-red in Table 3, where the choice (28) is privileged. Anyway, to suggest the reader some numerical clues about our proposal, consider that the apparently simplest choice $\gamma_k = 1, k \geq 0$, proved to be much *inefficient* in practice, while the choices $\gamma_k = \pm a_k$ gave appreciable results on different test problems (but still unclear results on larger test sets).

In particular, we preliminarily tested the CD class on two (small but) illconditioned problems described in Section 4 of [13]. The first problem, whose coefficient matrix is addressed as $A_{1_0} \in \mathbb{R}^{50\times 50}$, is '*obtained from a one-dimensional model, consisting of a line of two-node elements with support conditions at both ends and a linearly varying body force*'. The second problem has the coefficient matrix $A_{2_0} \in \mathbb{R}^{170\times 170}$, which is '*the stiffness matrix from a two-dimensional finite element model of a cantilever beam*'.

Table 6 The preconditioned CD, namely $CD_{\mathcal{M}}$, for solving (1)

The $CD_{\mathcal{M}}$ class

Step 0: Set $k = 0$, $y_0 \in \mathbb{R}^n$, $r_0 := b - Ay_0$, $\bar{\gamma}_0 \in \mathbb{R} \setminus \{0\}$, $\mathcal{M} \succ 0$.
If $r_0 = 0$, then STOP. Else, set $p_0 := \mathcal{M}r_0$, $k = k + 1$.
Compute $a_0 := r_0^T p_0 / p_0^T A p_0$,
$y_1 := y_0 + a_0 p_0$, $r_1 := r_0 - a_0 A p_0$.
If $r_1 = 0$, then STOP. Else, set $\sigma_0 := \bar{\gamma}_0 \|Ap_0\|^2_{\mathcal{M}} / p_0^T A p_0$,
$p_1 := \bar{\gamma}_0 \mathcal{M}(Ap_0) - \sigma_0 p_0$, $k = k + 1$.

Step k: Compute $a_{k-1} := r_{k-1}^T p_{k-1} / p_{k-1}^T A p_{k-1}$, $\bar{\gamma}_{k-1} \in \mathbb{R} \setminus \{0\}$,
$y_k := y_{k-1} + a_{k-1} p_{k-1}$, $r_k := r_{k-1} - a_{k-1} A p_{k-1}$.
If $r_k = 0$, then STOP. Else, set

$$\sigma_{k-1} := \bar{\gamma}_{k-1} \frac{\|Ap_{k-1}\|^2_{\mathcal{M}}}{p_{k-1}^T A p_{k-1}}, \quad \omega_{k-1} := \frac{\bar{\gamma}_{k-1}}{\bar{\gamma}_{k-2}} \frac{p_{k-1}^T A p_{k-1}}{p_{k-2}^T A p_{k-2}},$$

$$p_k := \bar{\gamma}_{k-1} \mathcal{M}(Ap_{k-1}) - \sigma_{k-1} p_{k-1} - \omega_{k-1} p_{k-2}, \quad k = k + 1.$$

Go to **Step** k.

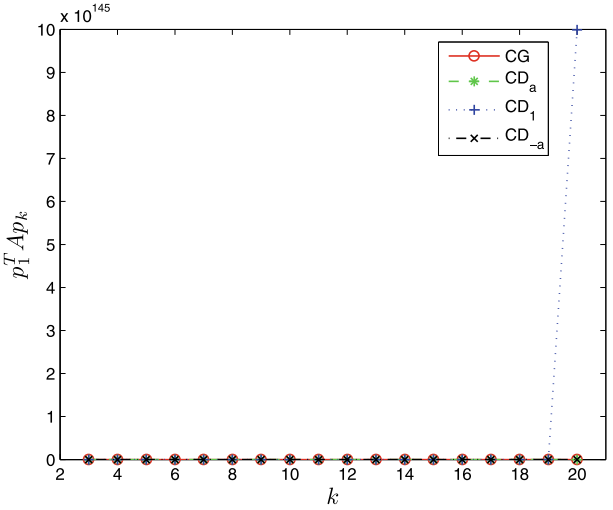

Fig. 2 Conjugacy loss for an illconditioned problem described by the coefficient matrix A_{1_0} in [13], using the CG, CD_a (the CD class setting $\gamma_0 = 1$ and $\gamma_k = a_k$, $k \geq 1$), CD_1 (the CD class setting $\gamma_k = 1$, $k \geq 0$) and CD_{-a} (the CD class setting $\gamma_0 = 1$ and $\gamma_k = -a_k$, $k \geq 1$). The quantity $p_1^T A p_k$ is reported for $k \geq 3$. As evident, the choice $\gamma_k = 1$, $k \geq 0$, can yield very harmful results when the coefficient matrix is illconditioned

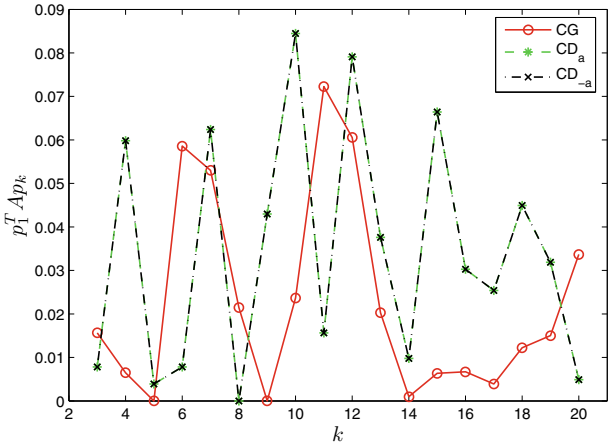

Fig. 3 Conjugacy loss for an illconditioned problem described by the coefficient matrix A_{1_0} in [13], using only the CG, CD_a (the CD class setting $\gamma_0 = 1$ and $\gamma_k = a_k$, $k \geq 1$) and CD_{-a} (the CD class setting $\gamma_0 = 1$ and $\gamma_k = -a_k$, $k \geq 1$). The quantity $p_1^T A p_k$ is reported for $k \geq 3$. The choices $\gamma_k = a_k$ and $\gamma_k = -a_k$ are definitely comparable, and are preferable to the CG for $k \in \{3, 6, 8, 11, 20\}$

In Figs. 2 and 3 we report the resulting experience on just the first of the two problems (similar results hold for the other one), where the CG is compared with algorithms in the class CD, setting $\gamma_k \in \{a_k, 1, -a_k\}$. As a partial justification for the reported numerical experience, we note that in the CD class the coefficient σ_k depends on the quantity $\|Ap_k\|^2$. Thus, $\|Ap_k\|^2$ may be large when A is illconditioned, so that the

choice $\gamma_k = 1$ possibly is inadequate to compensate the effect of illconditioning. On the other hand, setting $\gamma_k = \pm a$ and considering the expression of a_k, the coefficient σ_k is possibly re-scaled, taking into account the condition number of matrix A.

12 Conclusions

We have investigated a novel class of CG-based iterative methods. This allowed us to recast several properties of the CG within a broad framework of iterative methods, based on generating mutually conjugate directions. Both the analytical properties and the geometric insight where fruitfully exploited, showing that general CG-based methods, including the CG and the scaled-CG, may be introduced. Our resulting parameter dependent CG-based framework has the distinguishing feature of including conjugacy in a more general fashion, so that numerical results may strongly rely on the choice of a set of parameters. We urge to recall that in principle, since conjugacy can be generalized to the case of A indefinite (see for instance [8,11,18,25]) potentially further generalizations with respect to CD can be conceived (allowing the matrix A in (1) to be possibly indefinite).

Our study and the present conclusions are not primarily inspired by the aim of possibly beating the performance of the CG on practical cases. On the contrary, we preferred to justify our proposal in the light of a general analysis, which in case (but not necessary) may suggest competitive new iterative algorithms, for solving positive definite linear systems. In a future work, we are committed to consider the following couple of issues:

1. assessing clear rules for the choice of the sequence $\{\gamma_k\}$ in CD;
2. performing an extensive numerical experience, where different choices of the parameters $\{\gamma_k\}$ in our framework are considered, and practical guidelines for new efficient methods might be investigated.

Observe that the algorithms in CD are slightly more expensive than the CG, and they require the storage of one further vector with respect to the CG. However, we proved for CD some theoretical properties, which extend those provided by the CG, in order to possibly prevent from conjugacy loss. In addition, when specific values of the parameters in CD are chosen, then we obtain schemes equivalent to both the CG and the scaled-CG.

Furthermore, we have also introduced preconditioning in our proposal, as a possible extension of the preconditioned CG, so that illconditioned linear systems might be possibly more efficiently tackled. Our methods are also aimed to provide an effective tool in optimization contexts where a sequence of conjugate directions is sought. Truncated Newton methods are just an example of such contexts from unconstrained nonlinear optimization, as detailed in Sect. 3.

We are considering in a further study a numerical experience, over convex optimization problems, where CD and the relative preconditioned scheme are adopted to solve Newton's equation. Indeed, in case the matrix A in (1) is indefinite, the choices $\gamma_k \in \{a_k, |a_k|, -a_k, -|a_k|\}$ are of some interest and might be compared on a significant test set.

In addition, it might be worth also to investigate the choice where the preconditioner $\mathcal{M}$ in Table 6 is computed by a Quasi-Newton approximation of the inverse matrix A^{-1} (see also [13,26]), or by using the conjugate directions generated by CD, for a suitable choice of the parameters (see also [27]).

Furthermore, observe that conditions (8) or (7) cannot be further generalized imposing explicitly relations ($\ell \geq 1$)

$$p_k^T A p_j = 0, \qquad j = k-1, k-2, \ldots, k-\ell,$$

since (8) and (7) automatically imply $p_k^T A p_j = 0$, for any $j \leq k-3$ (see also Lemmas 4.1 and 4.2).

Finally, note that for the minimization of a convex quadratic functional in $\mathbb{R}^n$, the complete relation between the search directions generated by BFGS or L-BFGS updates and the CG was studied (see also [21]). Thus, we think that possible extensions may be considered by replacing the CG with the algorithms in our framework. In this regard, recalling that polarity (see [8]) plays a keynote role for generating conjugate directions, there is the chance that a possible relation between the BFGS update and CD could spot some light on the role of polarity for Quasi-Newton schemes.

Acknowledgments The author is indebted with the anonymous reviewers for their fruitful comments. The author also thanks the Italian national research program 'RITMARE', by *CNR-INSEAN, National Research Council-Maritime Research Centre*, for the support received.

References

1. Axelsson, O.: Iterative Solution Methods. Cambridge University Press, Cambridge (1996)
2. Golub, G.H., Van Loan, C.F.: Matrix Computations, 3rd edn. The John Hopkins University Press, Baltimore (1996)
3. Saad, Y.: Iterative Methods for Sparse Linear Systems, 2nd edn. SIAM, Philadelphia (2003)
4. Higham, N.J.: Accuracy and Stability of Numerical Algorithms. SIAM, Philadelphia (1996)
5. Saad, Y., Van Der Vorst, H.A.: Iterative solution of linear systems in the twentieth century. J. Comput. Appl. Math. **123**, 1–33 (2000)
6. Greenbaum, A., Strakos, Z.: Predicting the behavior of finite precision Lanczos and conjugate gradient computations. SIAM J. Matrix Anal. Appl. **13**, 121–137 (1992)
7. Greenbaum, A.: Iterative Methods for Solving Linear Systems, vol. SIAM. SIAM, Philadelphia (1997)
8. Hestenes, M.R.: Conjugate Direction Methods in Optimization. Springer, New York (1980)
9. Nash, S.G.: A survey of truncated-Newton methods. J. Comput. Appl. Math. **124**, 45–59 (2000)
10. Conn, A.R., Gould, N.I.M., Toint, PhL: Trust Region Methods. MPS-SIAM Series on Optimization. SIAM, Philadelphia (2000)
11. Fasano, G.: Planar-conjugate gradient algorithm for large scale unconstrained optimization. Part 2: Application. J. Optim. Theory Appl. **125**, 523–541 (2005)
12. Grippo, L., Lampariello, F., Lucidi, S.: A truncated Newton method with nonmonotone linesearch for unconstrained optimization. J. Optim. Theory Appl. **60**, 401–419 (1989)
13. Morales, J.L., Nocedal, J.: Automatic preconditioning by limited memory quasi-Newton updating. SIAM J. Optim. **10**, 1079–1096 (2000)
14. Hestenes, M.R., Stiefel, E.: Methods of conjugate gradients for solving linear systems. J. Res. Natl. Bur. Stand. **49**, 409–435 (1952)
15. Fasano, G.: Lanczos conjugate-gradient method and pseudoinverse computation on indefinite and singular systems. J. Optim. Theory Appl. **132**, 267–285 (2007)

16. Stoer, J.: Solution of large linear systems of equations by conjugate gradient type methods. In: Bachem, A., Grötschel, M., Korte, B. (eds.) Mathematical Programming. The State of the Art, pp. 540–565. Springer, Berlin (1983)
17. Nash, S.G., Sofer, A.: Assessing a search direction within a truncated Newton method. Oper. Res. Lett. **9**, 219–221 (1990)
18. Fasano, G., Roma, M.: Iterative computation of negative curvature directions in large scale optimization. Comput. Optim. Appl. **38**, 81–104 (2007)
19. Gould, N.I.M., Lucidi, S., Roma, M., Toint, PhL: Exploiting negative curvature directions in linesearch methods for unconstrained optimization. Optim. Methods Softw. **14**, 75–98 (2000)
20. Fasano, G., Lucidi, S.: A nonmonotone truncated Newton–Krylov method exploiting negative curvature directions, for large scale unconstrained optimization. Optim. Lett. **3**, 521–535 (2009)
21. Nocedal, J., Wright, S.: Numerical Optimization. Springer Series in Operations Research and Financial Engineering, 2nd edn. Springer, New York (2006)
22. Meurant, G.: The Lanczos and Conjugate Gradient Algorithms—From Theory to Finite Precision Computations. SIAM, Philadelphia (2006)
23. Polyak, T.B.: Introduction to Optimization, Translation Series in Mathematics and Engineering. Optimization Software Inc., Publications Division, New York (1987)
24. Campbell, S.L., Meyer, C.D.: Generalized Inverses of Linear Transformations. Dover Publications, New York (1979)
25. Fasano, G.: Planar-conjugate gradient algorithm for large scale unconstrained optimization. Part 1: Theory. J. Optim. Theory Appl. **125**, 543–558 (2005)
26. Gratton, S., Sartenaer, A., Tshimanga, J.: On a class of limited memory preconditioners for large scale linear systems with multiple right-hand sides. SIAM J. Optim. **21**, 912–935 (2011)
27. Fasano, G., Roma, M.: Preconditioning Newton–Krylov methods in non-convex large scale optimization. Comput. Optim. Appl. **56**, 253–290 (2013)