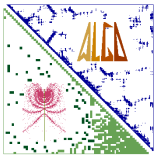
CERFACS

Centre Européen de Recherche et de Formation Avancée en Calcul Scientifique

A Flexible Generalized Conjugate Residual Method with Inner Orthogonalization and Deflated Restarting

L. M. Carvalho, S. Gratton, R. Lago and X. Vasseur

Technical Report TR/PA/10/10



Publications of the Parallel Algorithms Team

http://www.cerfacs.fr/algor/publications/

A FLEXIBLE GENERALIZED CONJUGATE RESIDUAL METHOD WITH INNER ORTHOGONALIZATION AND DEFLATED RESTARTING

L. M. CARVALHO*, S. GRATTON†, R. LAGO‡, AND X. VASSEUR§

Abstract.
This work is concerned with the development and study of a minimum residual norm subspace method based on the Generalized Conjugate Residual method with inner Orthogonalization (GCRO) method that allows flexible preconditioning and deflated restarting for the solution of non-symmetric or non-Hermitian linear systems. First we recall the main features of Flexible Generalized Minimum Residual with deflated restarting (FGMRES-DR), a recently proposed algorithm of the same family but based on the GMRES method. Next we introduce the new inner-outer subspace method named FGCRO-DR. A theoretical comparison of both algorithms is then made in the case of flexible preconditioning. It is proved that FGCRO-DR and FGMRES-DR are algebraically equivalent if a collinearity condition is satisfied. Furthermore we introduce three variants of FGCRO-DR that only differ in the formulation of the generalized eigenvalue problem for the harmonic Ritz pair information and investigate their main properties. Finally we demonstrate the effectiveness of the algorithms on a challenging application in quantum chromodynamics.

Key words. flexible or inner-outer Krylov subspace methods, variable preconditioning, deflation, iterative solver

AMS subject classifications. 65F10, 65N22, 15A06

1. Introduction. In recent years, several authors studied inner-outer Krylov subspace methods that allow variable preconditioning for the iterative solution of large sparse linear systems of equations. One of the first papers describing a subspace method with variable preconditioning is due to Axelsson and Vassilevski who proposed the Generalized Conjugate Gradient method [2]. See also [1, Section 12.3] for additional references. Since then, numerous methods have been proposed to address the symmetric, non-symmetric or non-Hermitian cases; these include Flexible Conjugate Gradient [23], Flexible GMRES (FGMRES) [27], Flexible QMR [34] and GMRESR [39] among others. This class of methods is required when preconditioning with a different (possibly nonlinear) operator at each iteration of a subspace method is considered. This notably occurs when adaptive preconditioners using information obtained from previous iterations [3, 14] are used or when inexact solutions of the preconditioning system using e.g. adaptive cycling strategy in multigrid [24] or approximate interior solvers in domain decomposition methods [35, Section 4.3] are considered. The latter situation is frequent when solving very large systems of linear equations resulting from the discretization of partial differential equations in three dimensions. Thus flexible Krylov subspace methods have gained a considerable interest in the past years and are subject to both theoretical and numerical studies [31]. We refer the reader to [32, Section 10] for additional comments on flexible methods.

When non variable preconditioning is considered, the full GMRES method [30] is often chosen for the solution of non-symmetric or non-Hermitian linear systems because of its robustness and its minimum residual norm property [29]. Nevertheless to control both the memory requirements and the computational cost of the orthogonalization scheme, restarted GMRES is preferred; it corresponds to a scheme where the maximal dimension of the approximation subspace is fixed. It means in practice that the orthonormal basis built is thrown away. Since some information is discarded at the restart, the convergence may stagnate and is expected to be slower compared to full GMRES. Nevertheless to retain the convergence rate a number of techniques have been proposed; they fall in the class of augmented and deflated methods; see e.g. [4, 10, 11, 19, 28]. Deflated methods compute spectral information at a restart and use this information to improve the convergence of the subspace method. One of the most recent procedure based on a deflation approach is GMRES with deflated restarting (GMRES-DR) [21]. This method reduces to restarted GMRES when no

*CNPq fellowship, Brazil. Applied Math. Dep., IME-UERJ, R. S. F. Xavier, 524, 629D, 20559-900, Rio de Janeiro, RJ, Brazil
†INPT-IRIT, University of Toulouse and ENSEEIHT, 2 rue Camichel, BP 7122, F-31071 Toulouse cedex 7, France and CERFACS, 42, Avenue Gaspard Coriolis, F-31057 Toulouse Cedex 1, France
‡FAPERJ Best Student Fellowship, PESC, COPPE-URFJ, 21941-972, Rio de Janeiro, RJ, Brazil
§CERFACS, 42, Avenue Gaspard Coriolis, F-31057 Toulouse Cedex 1, France

deflation is applied, but may provide a much faster convergence than restarted GMRES for well chosen deflation spaces as described in [21].

Quite recently a new minimum residual norm subspace method based on GMRES allowing deflated restarting and variable preconditioning has been proposed in [17]. It mainly attempted to combine the numerical features of GMRES with deflated restarting and the flexibility property of FGMRES. Numerical experiments have shown the efficiency of Flexible GMRES with deflated restarting (FGMRES-DR) on both academic and industrial examples. In this paper we study a new minimum residual norm subspace method based on the Generalized Conjugate Method with inner Orthogonalization (GCRO) [9] allowing deflated restarting and variable preconditioning. It is named Flexible Generalized Conjugate Residual Method with Inner Orthogonalization and Deflated Restarting (FGCRO-DR) and can be viewed as an extension of GCRO-DR [26] to the case of variable preconditioning. A major advantage of FGCRO-DR over FGMRES-DR is its ability to solve sequence of linear systems (where both the left- and right-hand sides could change) through recycling [26]. Although important this latter issue is not addressed in the paper and we concentrate on the case of a single linear system. In [26] Parks et al. mentioned that GCRO-DR and GMRES-DR were algebraically equivalent i.e. both methods produce the same iterates in exact arithmetic when solving the same given linear system starting from the same initial guess. When variable preconditioning is considered, it seems therefore natural to ask whether FGCRO-DR and FGMRES-DR could be also algebraically equivalent. We address this question in this paper and the main theoretical developments that are proposed will help us to answer this question. The main contributions of the paper are then twofold. First we prove that FGCRO-DR and FGMRES-DR can be considered as algebraically equivalent if a collinearity condition between two certain vectors is satisfied at each cycle. When considering non variable preconditioning, these theoretical developments will also allow us to show the unconditional algebraic equivalence between GCRO-DR and GMRES-DR that was stated without proof in [26]. Secondly we extend the initial framework of FGCRO-DR and introduce three variants of FGCRO-DR that only differ in the formulation of the generalized eigenvalue problem for the harmonic Ritz information. While one of them corresponds to the method briefly described in [18], the two others are new to the best of our knowledge. We analyze their corresponding main properties and show their respective interest in an application in quantum chromodynamics, where variable preconditioning is required.

This paper is organized as follows. In Section 2 we introduce the general background of this study. We briefly recall the main properties of FGMRES-DR and then introduce the FGCRO-DR method both from a mathematical and algorithmic points of view. Section 3 is mainly devoted to the analysis of both flexible methods. Therein we show that both methods can be algebraically equivalent in the flexible case if a certain collinearity condition is satisfied at each cycle. In Section 4 we propose three variants of FGCRO-DR, highlight noteworthy differences and finally compare their respective computational costs. Furthermore we demonstrate the effectiveness of the three algorithms on a challenging application in quantum chromodynamics in Section 5. Finally we draw some conclusions and perspectives in Section 6.

2. Flexible Krylov methods with restarting.

2.1. General setting.

Notation. Throughout this paper we denote $\|.\|$ the Euclidean norm, $I_k \in \mathbb{C}^{k\times k}$ the identity matrix of dimension k and $0_{i\times j} \in \mathbb{C}^{i\times j}$ the zero rectangular matrix with i rows and j columns. Given $N \in \mathbb{C}^{n\times m}$ $\Pi_{N^\perp} = I_n - N\,N^\dagger$ will represent the orthogonal projector onto range$(N)^\perp$, where $\dagger$ refers to the Moore-Penrose pseudoinverse operation. Finally given $Z_m = [z_1, \cdots, z_m] \in \mathbb{C}^{n\times m}$, we will usually decompose Z_m into two submatrices defined as $Z_k = [z_1, \cdots, z_k] \in \mathbb{C}^{n\times k}$ and $Z_{m-k} = [z_{k+1}, \cdots, z_m] \in \mathbb{C}^{n\times(m-k)}$.

Setting. We focus on minimum residual norm based subspace methods that allow flexible preconditioning for the iterative solution of

$$Ax = b, \quad A \in \mathbb{C}^{n\times n}, \quad x, b \in \mathbb{C}^n \tag{2.1}$$

given an initial vector $x_0 \in \mathbb{C}^n$. In this paper A is supposed to be nonsingular. Flexible methods refer to a class of methods where the preconditioner is allowed to vary at each iteration. We refer

the reader to e.g. [32] for a general introduction on Krylov subspace methods and to [32, Section 10] and [29, Section 9.4] for a review on flexible methods. The minimum residual norm GMRES method [30] has been extended by Saad [27] to allow variable preconditioning. The resulting algorithm known as FGMRES(m) relies on the Arnoldi relation

$$AZ_m = V_{m+1}\bar{H}_m, \tag{2.2}$$

where $Z_m \in \mathbb{C}^{n\times m}$, $V_{m+1} \in \mathbb{C}^{n\times(m+1)}$ has orthonormal columns and $\bar{H}_m \in \mathbb{C}^{(m+1)\times m}$ is upper Hessenberg. We denote $\mathcal{M}_j$ the preconditioning operator at iteration j and remark that $\mathcal{M}_j$ may be a nonlinear preconditioning function. We will then denote $\mathcal{M}_j(v)$ the action of $\mathcal{M}_j$ on a vector v. In (2.2), the columns of V_{m+1} form an orthonormal basis of the subspace spanned by the following vectors

$$\{r_0, Az_1, \cdots, Az_m\} \quad \text{with} \quad r_0 = b - Ax_0$$

whereas $Z_m = [z_1, \cdots, z_m]$ and $V_m = [v_1, \cdots, v_m]$ are related by

$$Z_m = [\mathcal{M}_1(v_1), \cdots, \mathcal{M}_m(v_m)] \quad \text{with} \quad v_1 = \frac{r_0}{\|r_0\|}.$$

The minimization problem $\min \|b - Ax\|$ is then solved as

$$x_m = x_0 + Z_m y^*,$$

where y^* is the solution of the following least-squares problem of size $(m+1)\times m$

$$y^* = \text{argmin}_{y\in\mathbb{C}^m}\|r_0 - AZ_m y\| = \text{argmin}_{y\in\mathbb{C}^m}\|\|r_0\|\, e_1 - \bar{H}_m y\|,$$

where e_1 denotes the first canonical vector of $\mathbb{C}^{m+1}$. Flexible subspace methods with restarting are based on a procedure where the construction of the subspace is stopped after a certain number of steps (denoted by m in this paper with $m < n$). The method is then restarted mainly to control both the memory requirements and the cost of the orthogonalization scheme. In FGMRES(m) the restarting consists in taking as an initial guess the past iterate x_m associated with the smallest residual norm.

The main focus of this paper is to present minimum residual norm subspace methods with *deflated* restarting that allow *flexible* preconditioning. Deflated restarting aims at determining an approximation subspace of dimension m as a direct sum of two subspaces of smaller dimension, where one of these subspaces will contain relevant spectral information that will be kept for the next cycle. We refer the reader to e.g. [28] and [32, Section 9] for a review on augmented and deflated methods. Flexible methods with deflated restarting will notably satisfy the following flexible Arnoldi relation

$$AZ_m = V_{m+1}\bar{H}_m \quad \text{with} \quad V_{m+1}^H\, V_{m+1} = I_{m+1}, \tag{2.3}$$

where $\bar{H}_m \in \mathbb{C}^{(m+1)\times m}$ is not necessarily of upper Hessenberg form. In this paper we call this relation a flexible Arnoldi-like relation due to its similarity to relation (2.2).

Stagnation and breakdown. We refer the reader to [31, Section 6] for general comments and a detailed discussion on the possibility of both breakdown and stagnation in flexible inner-outer Krylov subspace methods. Although important, these issues are not addressed in this paper and we assume that no breakdown occurs in the inner-outer subspace methods that will be proposed.

2.2. Flexible GMRES with deflated restarting. A number of techniques have been proposed to compute spectral information at a restart and use this information to improve the convergence rate of the Krylov subspace methods; see, e.g., [19, 20, 21, 28]. These techniques have been exclusively developed in the case of a fixed preconditioner. Among others GMRES-DR is one of those methods. It focuses on removing (or deflating) the eigenvalues of smallest magnitude. A full subspace of dimension k, $k < m$ (and not only the approximate solution with minimum residual

norm) is now retained at the restart and the success of this approach has been demonstrated on many academic examples [19]. Approximations of eigenvalues of smallest magnitude are obtained by computing harmonic Ritz pairs of A with respect to a certain subspace [21]. We present here a definition of a harmonic Ritz pair equivalent to the one introduced in [25, 33]; it will be of key importance when defining appropriate deflation strategies.

DEFINITION 2.1. *Harmonic Ritz pair. Consider a subspace $\mathcal{U}$ of $\mathbb{C}^n$. Given $B \in \mathbb{C}^{n\times n}$, $\theta \in \mathbb{C}$ and $y \in \mathcal{U}$, (θ, y) is a harmonic Ritz pair of B with respect to $\mathcal{U}$ if and only if*

$$By - \theta\, y \perp B\mathcal{U}$$

or equivalently, for the canonical scalar product,

$$\forall w \in \mathrm{range}(B\mathcal{U}) \quad w^H\,(By - \theta\, y) = 0.$$

We call y a harmonic Ritz vector associated with the harmonic Ritz value θ.

As in the case of fixed preconditioning, deflated restarting may also improve the convergence rate of flexible subspace methods. In [17] a deflated restarting procedure has been proposed for the FGMRES algorithm. The i-th cycle of the resulting algorithm called FGMRES-DR is now briefly described and we denote $r_0^{(i-1)} = b - Ax_0^{(i-1)}$ the residual obtained at the end of the previous cycle.

Based on the Arnoldi-like relation (2.3), the deflation procedure proposed in [17, Proposition 1] relies on the use of k harmonic Ritz vectors $Y_k = V_m P_k$ of $AZ_m V_m^H$ with respect to range(V_m), where $Y_k \in \mathbb{C}^{n\times k}$ and $P_k \in \mathbb{C}^{m\times k}$. Next, the QR factorization of the following $(m+1)\times(k+1)$ matrix

$$\begin{bmatrix}\begin{bmatrix} P_k \\ 0_{1\times k}\end{bmatrix} & V_{m+1}^H\, r_0^{(i-1)}\end{bmatrix} = \begin{bmatrix}\begin{bmatrix} P_k \\ 0_{1\times k}\end{bmatrix} & c - \bar{H}_m y^*\end{bmatrix} \quad \text{with} \quad r_0^{(i-1)} = V_{m+1}(c - \bar{H}_m y^*)$$

is performed. This allows us to compute new matrices $Z_k \in \mathbb{C}^{n\times k}$, $V_{k+1} \in \mathbb{C}^{n\times(k+1)}$ and $\bar{H}_k \in \mathbb{C}^{(k+1)\times k}$ such that

$$\begin{aligned} A\, Z_k &= V_{k+1}\, \bar{H}_k, \\ V_{k+1}^H\, V_{k+1} &= I_{k+1}, \\ \mathrm{range}([Y_k, r_0^{(i-1)}]) &= \mathrm{range}(V_{k+1}) \end{aligned}$$

where $\bar{H}_k$ is a $(k+1)\times k$ rectangular matrix. FGMRES-DR then carries out $m-k$ Arnoldi steps with flexible preconditioning and starting vector v_{k+1} while maintaining orthogonality to V_k leading to

$$A\,[z_{k+1}, \cdots, z_m] = [v_{k+1}, \cdots, v_{m+1}]\,\bar{H}_{m-k} \quad \text{and} \quad V_{m+1}^H\, V_{m+1} = I_{m+1}.$$

We note that $\bar{H}_{m-k} \in \mathbb{C}^{(m-k+1)\times(m-k)}$ is upper Hessenberg. At the end of the i-th cycle this gives the flexible Arnoldi-like relation

$$A\, Z_m = V_{m+1}\, \bar{H}_m,$$

where $Z_m \in \mathbb{C}^{n\times m}$, $V_{m+1} \in \mathbb{C}^{n\times(m+1)}$ and $\bar{H}_m \in \mathbb{C}^{(m+1)\times m}$. We note that $\bar{H}_m$ is no more upper Hessenberg due to the leading dense $(k+1)\times k$ submatrix $\bar{H}_k$. At the end of the i-th cycle, an approximate solution $x_0^{(i)} \in \mathbb{C}^n$ is then found by minimizing the residual norm $\|b - A(x_0^{(i-1)} + Z_m y)\|$ over the space $x_0^{(i-1)} + \mathrm{range}(Z_m)$, the corresponding residual being $r_0^{(i)} = b - Ax_0^{(i)}$, with $r_0^{(i)} \in \mathrm{range}(V_{m+1})$. We refer the reader to [17] for the complete derivation of the method and numerical experiments showing the efficiency of FGMRES-DR on both academic and industrial examples.

2.3. Flexible GCRO with deflated restarting. GCRO-DR [26] - a combination of GMRES-DR and GCRO - is a Krylov subspace method that allows deflated restarting and subspace recycling simultaneously. This latter feature is particularly interesting when solving sequences of linear systems with possibly different left- or right-hand sides. As pointed out in [26], GCRO-DR is attractive because any subspace may be recycled. In this paper we restrict the presentation to the case of a single linear system as proposed in (2.1).

GCRO and GCRO-DR belong to the family of inner-outer methods [1, Ch. 12] where the outer iteration is based on GCR, a minimum residual norm method proposed by Eisenstat, Elman and Schultz [13]. To this end GCR maintains a correction subspace spanned by range(Z_m) and an approximation subspace spanned by range(V_m), where $Z_m, V_m \in \mathbb{C}^{n\times m}$ satisfy

$$\begin{aligned} A\ Z_m &= V_m, \\ V_m^H\ V_m &= I_m. \end{aligned}$$

The optimal solution of the minimization problem $\min \|b - Ax\|$ over the subspace $x_0 + \text{range}(Z_m)$ is then found as $x_m = x_0 + Z_m\ V_m^H\ r_0$. Consequently $r_m = b - A\ x_m$ satisfies

$$r_m = r_0 - V_m\ V_m^H r_0 = \Pi_{V_m^\perp} r_0, \quad r_m \perp \text{range}(V_m).$$

In [9] de Sturler proposed an improvement to GMRESR [39], an inner-outer method based on GCR in the outer part and GMRES in the inner part respectively. He suggested that the inner iteration takes place in a subspace orthogonal to the outer Krylov subspace. In this inner iteration the projected residual equation

$$\begin{aligned} (I_n - V_m\ V_m^H)Az &= (I_n - V_m\ V_m^H)r_m \\ (I_n - V_m\ V_m^H)Az &= r_m \end{aligned}$$

is solved only approximately. If a minimum residual norm subspace method is used in the inner iteration to solve this projected residual linear system, the residual over both the inner and outer subspaces would be minimized. This leads to the GCRO (Generalized Conjugate Residual method with inner Orthogonalization) Krylov subspace method [9]. Numerical experiments [9] indicate that the resulting method may perform better than inner-outer methods (without orthogonalizations) in some cases.

The GCRO method with deflated restarting (named GCRO-DR) based on harmonic Ritz value information has been proposed in [26]. An approximate invariant subspace is used for deflation following closely the GMRES-DR method. We refer the reader to [26] for a description of this method, algorithms and implementation details. We present now a new variant of GCRO-DR that allows flexible preconditioning by explaining the different steps occurring during the i-th cycle. Again we denote $r_0^{(i-1)} = b - Ax_0^{(i-1)}$ the residual obtained at the end of the previous cycle.

We suppose that a flexible Arnoldi-like relation of type (2.3) holds. As in Section 2.2 an important point is to specify which harmonic Ritz information is selected. Given a certain matrix $W_m \in \mathbb{C}^{n\times m}$, to be specified later on, the deflation procedure relies on the use of k harmonic Ritz vectors $Y_k = W_m P_k$ of $AZ_m W_m^\dagger$ with respect to range(W_m), where $Y_k \in \mathbb{C}^{n\times k}$ and $P_k \in \mathbb{C}^{m\times k}$. W_m will notably satisfy a property detailed in Lemma 3.3 and we point out that the calculation of $W_m^\dagger$ is not needed in the practical implementation of the algorithm (see further discussion in Section 4.2.1). Next, the QR factorization of the $m \times k$ matrix $\bar{H}_m P_k$ is performed. This allows us to obtain new matrices $Z_k, V_k \in \mathbb{C}^{n\times k}$ such that

$$\begin{aligned} A\ Z_k &= V_k, \\ V_k^H\ V_k &= I_k, \end{aligned}$$

by using information related to the QR factorization and the flexible Arnoldi relation (2.3) exclusively. Then the inner iteration is based on the approximate solution of

$$(I_n - V_k\ V_k^H)Az = (I_n - V_k\ V_k^H)r_0^{(i-1)} = r_0^{(i-1)}.$$

For that purpose FGCRO-DR then carries out $m-k$ steps of the Arnoldi method with flexible preconditioning leading to

$$\begin{aligned}(I_n - V_k V_k^H)\ A\ [z_{k+1}, \cdots, z_m] &= [v_{k+1}, \cdots, v_{m+1}]\ \bar{H}_{m-k} \\ (I_n - V_k V_k^H)\ A\ Z_{m-k} &= V_{m-k+1}\ \bar{H}_{m-k}\end{aligned}$$

with $v_{k+1} = r_0^{(i-1)}/||r_0^{(i-1)}||$. At the end of the cycle this gives the flexible Arnoldi-like relation

$$\begin{aligned}A\ [Z_k, Z_{m-k}] &= [V_k, V_{m-k+1}] \begin{bmatrix} I_k & V_k^H\ A\ Z_{m-k} \\ 0_{m-k+1\times k} & \bar{H}_{m-k} \end{bmatrix} \\ A\ Z_m &= V_{m+1}\ \bar{H}_m,\end{aligned}$$

where $Z_m \in \mathbb{C}^{n\times m}$, $V_{m+1} \in \mathbb{C}^{n\times(m+1)}$ and $\bar{H}_m \in \mathbb{C}^{(m+1)\times m}$. At the end of the i-th cycle, an approximate solution $x_0^{(i)} \in \mathbb{C}^n$ is then found by minimizing the residual norm $\|b - A(x_0^{(i-1)} + Z_m y)\|$ over the space $x_0^{(i-1)} + \text{range}(Z_m)$, the corresponding residual being $r_0^{(i)} = b - Ax_0^{(i)}$, with $r_0^{(i)} \in \text{range}(V_{m+1})$.

2.4. Algorithms. Details of the FGCRO-DR method are given in Algorithm 1, where Matlab-like notations are adopted (for instance in step 7b, $Q(1:m, 1:k)$ denotes the submatrix made of the first m rows and first k columns of matrix Q). For the sake of completeness the FGMRES-DR algorithm has been also described with notations chosen as close as possible to FGCRO-DR to make a code comparison as easy as possible. Concerning Algorithm 1 we make the following comments:

- As discussed later the computation of $W_m^\dagger$ in step 5a is not required thanks to the definition of the harmonic Ritz pair (see Definition 2.1).
- As pointed out by Morgan [21] and Parks et al. [26] we might have to adjust k during the algorithm to include both the real and imaginary parts of complex eigenvectors.
- Although notations are similar in steps 6a and 6b, we remark that the respective orthogonal and triangular factors do not have the same dimensions. In FGCRO-DR $Q \in \mathbb{C}^{(m+1)\times k}$ and $R \in \mathbb{C}^{k\times k}$, whereas $Q \in \mathbb{C}^{(m+1)\times(k+1)}$ and $R \in \mathbb{C}^{(k+1)\times(k+1)}$ in the FGMRES-DR algorithm.
- The matrices Z_k and V_k (such that $AZ_k = V_k$ and $V_k^H V_k = I_k$) are obtained in steps 8a and 9a (see also Lemma 3.4).
- In steps 10a and 10b $\mathcal{M}_j^{(i)}$ denotes the possibly nonlinear preconditioning operator at iteration j during the i-th cycle.
- In step 11b $B_{k\times m-k} \in \mathbb{C}^{k\times(m-k)}$ results from the orthogonalization of $[v_{k+2}, \cdots, v_{m+1}]$ against V_{k+1}.
- In FGMRES-DR the computation of the residual at the end of the cycle (step 14) can be performed at a cheaper cost. Indeed it can be shown that $r_0 = V_{k+1}(c - \bar{H}_m y^*)$ [17, Proposition 3].

3. Analysis of FGMRES-DR and FGCRO-DR. We compare now the flexible variants of GMRES-DR and GCRO-DR introduced in Sections 2.2 and 2.3 respectively. In the following we use the superscript $^\#$ to denote quantities related to the FGMRES-DR algorithm e.g. $Y_k^\#$ denote the set of harmonic Ritz vectors computed in the FGMRES-DR algorithm. When analyzing both algorithms we will suppose that identical preconditioning operators are used in steps 10a and 10b respectively i.e.

$$\forall i, \forall j \in \{k+1, \cdots, m\}, \quad \mathcal{M}_j^{(i)}(.) = \mathcal{M}_j^{(i)^\#}(.) \quad . \tag{3.1}$$

3.1. Equivalent preconditioning matrix. LEMMA 3.1. *Equivalent preconditioning matrix. Suppose that $V_p = [v_1, \cdots, v_p] \in \mathbb{C}^{n\times p}$ and $Z_p = [\mathcal{M}_1(v_1), \cdots, \mathcal{M}_p(v_p)] \in \mathbb{C}^{n\times p}$ obtained during a cycle of a flexible method with (standard or deflated) restarting (with $1 \le p \le m < n$) are both of full rank i.e.* $\text{range}(V_p) = \text{range}(Z_p) = p$. *We will then denote $M_{V_p} \in \mathbb{C}^{n\times n}$ a nonsingular equivalent preconditioning matrix defined as*

$$Z_p = M_{V_p}\ V_p. \tag{3.2}$$

Algorithm 1 Flexible GCRO-DR(m,k) and Flexible GMRES-DR(m,k)

1: choose m, k, tol and x_0
2: $r_0 = b - Ax_0$, $\beta = \|r_0\|$, $v_1 = r_0/\beta$, $c = \beta e_1$, $i \leftarrow 0$
3: Apply FGMRES(m) to obtain $\bar{H}_m$, Z_m, V_{m+1} such that $AZ_m = V_{m+1}\bar{H}_m$, $y^* = \arg\min_{y\in\mathbb{C}^m} \|c - \bar{H}_m y\|$, $x_0^{(0)} = x_0 + Z_m y^*$, $r_0^{(0)} = b - Ax_0^{(0)} = V_{m+1}(c - \bar{H}_m y^*)$, $W_m = V_m$
4: **while** $\|r_0^{(i)}\| > \|b\| \times tol$ **do** $i \leftarrow i+1$

FGCRO-DR	
5a:	Compute k harmonic Ritz vectors of $AZ_m W_m^\dagger$ with respect to range(W_m) and store them in Y_k. Define P_k such that $Y_k = W_m P_k$.
6a:	$Q\,R = \bar{H}_m P_k$
7a:	$W_k = W_m P_k R^{-1}$
8a:	$V_k = V_{m+1} Q$
9a:	$Z_k = Z_m P_k R^{-1}$
10a:	Apply $m-k$ flexible preconditioned Arnoldi steps with $(I_n - V_k V_k{}^H)A$ and $v_{k+1} = r_0^{(i-1)}/\|r_0^{(i-1)}\|$ such that $(I_n - V_k V_k{}^H)A\left[z_{k+1},\ldots,z_m\right] = \left[v_{k+1},\ldots,v_{m+1}\right]\bar{H}_{m-k}$ with $z_j = \mathcal{M}_j^{(i)}(v_j)$
11a:	Set $\bar{H}_m = \begin{bmatrix} I_k & V_k^H A Z_{m-k} \\ 0_{m-k+1\times k} & \bar{H}_{m-k} \end{bmatrix}$ yielding $A\left[z_1,\ldots,z_m\right] = \left[v_1,\ldots,v_{m+1}\right]\bar{H}_m$ and define $W_m = \left[W_k \quad V_m(1:n, k+1:m)\right]$

FGMRES-DR	
5b:	Compute k harmonic Ritz vectors of $AZ_m V_m^H$ with respect to range(V_m) and store them in Y_k. Define P_k such that $Y_k = V_m P_k$.
6b:	$QR = \left[\begin{bmatrix} P_k \\ 0_{1\times k}\end{bmatrix} \quad c - \bar{H}_m y^*\right]$
7b:	$\bar{H}_k = Q^H \bar{H}_m Q(1:m, 1:k)$
8b:	$V_{k+1} = V_{m+1} Q$
9b:	$Z_k = Z_m Q(1:m, 1:k)$
10b:	Apply $m-k$ flexible preconditioned Arnoldi steps with A and v_{k+1} while maintaining orthogonality to V_k such that $A\left[z_{k+1},\ldots,z_m\right] = \left[v_{k+1},\ldots,v_{m+1}\right]\bar{H}_{m-k}$ with $z_j = \mathcal{M}_j^{(i)}(v_j)$ and $V_{m+1}^H V_{m+1} = I_{m+1}$
11b:	Set $\bar{H}_m = \left[\begin{bmatrix} \bar{H}_k \\ 0_{m-k\times k}\end{bmatrix} \begin{bmatrix} B_{k\times m-k} \\ \bar{H}_{m-k}\end{bmatrix}\right]$ yielding $A\left[z_1,\ldots,z_m\right] = \left[v_1,\ldots,v_{m+1}\right]\bar{H}_m$

12: $y^* = \arg\min_{y\in\mathbb{C}^m} \|c - \bar{H}_m y\|$ with $c = V_{m+1}^H r_0^{(i-1)}$
13: $x_0^{(i)} = x_0^{(i-1)} + Z_m y^*$
14: $r_0^{(i)} = b - Ax_0^{(i)}$
15: **end while**

Such a matrix represents the action of the nonlinear operators $\mathcal{M}_j$ on the set of vectors v_j (with $j = 1,\cdots,p$). It can be chosen e.g. as $M_{V_p} = [Z_p\ \underline{Z_p}][V_p\ \underline{V_p}]^{-1}$ where $\underline{Z_p}$ (respectively $\underline{V_p}$) denotes an orthogonal complement of Z_p (respectively V_p) in $\mathbb{C}^n$.

3.2. Relations between Z_m and W_m and $Z_m^\#$ and $V_m^\#$. We denote $M_{W_m}^{(0)}$ and $M_{V_m^\#}^{(0)}$ the equivalent preconditioning matrices used in the initialization phase of both algorithms (step 3 in Algorithm 1). With this notation we remark that the following relations hold

$$Z_m = M_{W_m}^{(0)} W_m = Z_m^\# = M_{V_m^\#}^{(0)} V_m^\#. \tag{3.3}$$

We first analyze the relation between $Z_m^\#$ and $V_m^\#$.

LEMMA 3.2. *At the end of the i-th cycle of the FGMRES-DR method $Z_m^\#$ and $V_m^\#$ satisfy*

$$Z_m^\# = M_{V_m^\#}^{(i)\,\#} V_m^\# = [M_{V_m^\#}^{(i-1)\#} V_k^\#,\ M_{V_{m-k}^\#}^{(i)\#} V_{m-k}^\#]. \tag{3.4}$$

Proof.

The initialization phase leads to the relation $Z_m^\# = M_{V_m^\#}^{(0)} V_m^\#$. We suppose that at the end of the $i-1$th cycle the following relation holds: $Z_m^\# = M_{V_m^\#}^{(i-1)\#} V_m^\#$. The orthogonal factor

$Q^{\#} \in \mathbb{C}^{(m+1)\times(k+1)}$ obtained in step 6b can be decomposed as follows

$$Q^{\#} = \begin{bmatrix} Q_k^{\#} & \rho \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} \bar{Q} \\ 0_{1\times k} \end{bmatrix} & \rho \end{bmatrix} \tag{3.5}$$

where $Q_k^{\#} \in \mathbb{C}^{(m+1)\times k}$, $\rho \in \mathbb{C}^{m+1}$ and $\bar{Q} \in \mathbb{C}^{m\times k}$. At step 9b of the i-th cycle $Z_k^{\#}$ is defined as

$$Z_k^{\#} = Z_m^{\#}\bar{Q} = {M_{V_m^{\#}}^{(i-1)}}^{\#} V_m^{\#}\bar{Q} = {M_{V_m^{\#}}^{(i-1)}}^{\#} V_k^{\#}.$$

The proof is then completed since $Z_{m-k}^{\#} = [\mathcal{M}_{k+1}^{(i)\#}(v_{k+1}^{\#}), \cdots, {\mathcal{M}_m^{(i)}}^{\#}(v_m^{\#})] = M_{V_{m-k}^{\#}}^{(i)\#} V_{m-k}^{\#}$ at the end of step 10b. □

The next lemma details a relation between Z_m and W_m that is satisfied in the FGCRO-DR method.

LEMMA 3.3. *At the end of the i-th cycle of the FGCRO-DR method Z_m and W_m satisfy*

$$Z_m = M_{W_m}^{(i)} W_m = [M_{W_m}^{(i-1)} W_k,\ M_{W_{m-k}}^{(i)} W_{m-k}]. \tag{3.6}$$

Proof. The initialization phase leads to the relation $Z_m = M_{W_m}^{(0)} W_m$. We suppose that at the end of the $i-1$th cycle the following relation holds: $Z_m = M_{W_m}^{(i-1)} W_m$. At step 9a of the i-th cycle Z_k is defined as

$$\begin{aligned} Z_k &= Z_m P_k R^{-1} \\ &= M_{W_m}^{(i-1)} W_m P_k R^{-1} \\ &= M_{W_m}^{(i-1)} W_k. \end{aligned}$$

The proof is then completed since $Z_{m-k} = [\mathcal{M}_{k+1}^{(i)}(w_{k+1}), \cdots, \mathcal{M}_m^{(i)}(w_m)] = M_{W_{m-k}}^{(i)} W_{m-k}$ at the end of step 11a. □

Lemma 3.2 and 3.3 show that $Z_m^{\#}$, $V_m^{\#}$, Z_m and W_m satisfy a similar relation that will play a central role in Section 3.3. We investigate next the relation between Z_m and V_m.

LEMMA 3.4. *At the end of the i-th cycle of the FGCRO-DR method Z_m and V_m satisfy*

$$[AZ_k, Z_{m-k}] = [V_k,\ M_{V_{m-k}}^{(i)} V_{m-k}]. \tag{3.7}$$

Proof. Due to the Arnoldi-like relation (2.3), AZ_k can be also written as

$$\begin{aligned} AZ_k &= AZ_m P_k R^{-1} && (3.8)\\ &= V_{m+1}\bar{H}_m P_k R^{-1} && (3.9)\\ &= V_{m+1} Q && (3.10)\\ &= V_k. && (3.11) \end{aligned}$$

The proof is then completed since $Z_{m-k} = [\mathcal{M}_{k+1}^{(i)}(v_{k+1}), \cdots, \mathcal{M}_m^{(i)}(v_m)] = M_{V_{m-k}}^{(i)} V_{m-k}$ at the end of step 11a. □

We conclude this section by presenting a technical lemma related to the FGMRES-DR method.

LEMMA 3.5. *During the i-th cycle of the FGMRES-DR method, $v_{k+1}^{\#}$ satisfies the following relation*

$$v_{k+1}^{\#} = \bar{v}_{k+1}^{\#}/||\bar{v}_{k+1}^{\#}|| \quad \textit{with} \quad \bar{v}_{k+1}^{\#} = \Pi_{[Y_k^{\#}]^{\perp}} r_0^{(i-1)\#} \tag{3.12}$$

where $r_0^{(i-1)\#} = b - Ax_0^{(i-1)\#}$ denotes the residual obtained at the end of the $(i-1)$-th cycle.

Proof. In (3.5) $\rho \in \mathbb{C}^{m+1}$ is defined as

$$\rho = \frac{\bar{\rho}}{||\bar{\rho}||} \quad \text{with} \quad \bar{\rho} = (I_{m+1} - Q_k^{\#} Q_k^{\#H})(c^{\#} - \bar{H}_m^{\#} y^{\#*}). \tag{3.13}$$

A consequence of the representation of $Q_k^\#$ in (3.5) is that the matrix $V_{m+1}^\# Q_k^\#$ can be written as

$$V_{m+1}^\# Q_k^\# = V_m^\# \bar{Q}. \tag{3.14}$$

Using (3.13) and (3.14) we obtain

$$\bar{v}_{k+1}^\# = V_{m+1}^\# \bar{\rho} = r_0^{(i-1)\#} - V_{m+1}^\# Q_k^\# {Q_k^\#}^H V_{m+1}^{\#H} V_{m+1}^\# (c^\# - \bar{H}_m^\# y^{\#*}).$$
$$\bar{v}_{k+1}^\# = r_0^{(i-1)\#} - V_m^\# \bar{Q}\bar{Q}^H V_m^{\#H} V_{m+1}^\# (c^\# - \bar{H}_m^\# y^{\#*}).$$

Since $V_m^\# \bar{Q}$ has orthonormal columns this last expression now becomes

$$\bar{v}_{k+1}^\# = r_0^{(i-1)\#} - V_m^\# \bar{Q}(V_m^\# \bar{Q})^H r_0^{(i-1)\#} = \Pi_{[V_m^\# \bar{Q}]^\perp} r_0^{(i-1)\#}.$$

Since $\bar{Q}$ is the orthogonal factor of the QR decomposition of $P_k^\#$, we have the following relation

$$\text{range}(V_m^\# P_k^\#) = \text{range}(V_m^\# \bar{Q}).$$

Since by definition $Y_k^\# = V_m^\# P_k^\#$ the proof is then completed. □

3.3. Analysis of the FGMRES-DR and FGCRO-DR methods (Algorithm 1). Lemma 3.3 has already described an important property satisfied by W_m in the FGCRO-DR method proposed in Algorithm 1. We will analyze further the relation between the FGMRES-DR and FGCRO-DR methods. The next theorem states that the two flexible methods generate the same iterates in exact arithmetic under some conditions involving notably two vectors.

THEOREM 3.6.
We denote $r_0^{(i)} = b - Ax_0^{(i)}$ the residual obtained at the end of the i-th cycle of the FGCRO-DR method (see step 14 of Algorithm 1). We suppose that Lemma 3.1 holds and that the same equivalent preconditioning matrix is obtained at the end of the i-th cycle of both FGCRO-DR and FGMRES-DR algorithms i.e. $M_{W_m}^{(i)} = M_{V_m^\#}^{(i)\#}$. Under this assumption the harmonic Ritz vectors $Y_k^\#$ and Y_k can be chosen equal during the $i+1$-th cycle. If in addition there exists a real-valued positive coefficient η such that

$$\Pi_{[Y_k, r_0^{(i)}/\|r_0^{(i)}\|]^\perp} A\mathcal{M}_{k+1}^{(i+1)}(\Pi_{Y_k^\perp} r_0^{(i)}/\|\Pi_{Y_k^\perp} r_0^{(i)}\|) = \eta\ \Pi_{[Y_k, r_0^{(i)}/\|r_0^{(i)}\|]^\perp} A\mathcal{M}_{k+1}^{(i+1)}(r_0^{(i)}/\|r_0^{(i)}\|) \tag{3.15}$$

in the FGCRO-DR algorithm, then both algorithms generate the same iterates in exact arithmetic and

$$\text{range}(V_{m+1}) = \text{range}(V_{m+1}^\#), \tag{3.16}$$
$$\text{range}(Z_m) = \text{range}(Z_m^\#), \tag{3.17}$$

with

$$V_{m+1} = [V_{k+1}^\# \widehat{Q}, v_{k+2}, \cdots, v_{m+1}], \quad V_{m+1}^\# = [V_{k+1}^\#, v_{k+2}, \cdots, v_{m+1}], \tag{3.18}$$
$$Z_m = [Z_{k+1}^\# \widehat{X}, z_{k+2}, \cdots, z_m], \quad Z_m^\# = [Z_{k+1}^\#, z_{k+2}, \cdots, z_m], \tag{3.19}$$

where $\widehat{Q} \in \mathbb{C}^{(k+1)\times(k+1)}$ is a unitary matrix and $\widehat{X} \in \mathbb{C}^{(k+1)\times(k+1)}$ is a nonsingular triangular matrix.

Proof. The whole proof is performed in three parts assuming that we analyze the $i+1$-th cycle of each algorithm. Suppose that at the beginning of the $i+1$-th cycle (step 4) there exist a unitary matrix $\widehat{Q} \in \mathbb{C}^{(k+1)\times(k+1)}$ and a nonsingular matrix $\widehat{X} \in \mathbb{C}^{(k+1)\times(k+1)}$ such that the following relations hold

$$V_{k+1} = V_{k+1}^\# \widehat{Q}, \tag{3.20}$$
$$Z_{k+1} = Z_{k+1}^\# \widehat{X}, \tag{3.21}$$
$$[v_{k+2}, \cdots, v_{m+1}] = \left[v_{k+2}^\#, \cdots, v_{m+1}^\#\right], \tag{3.22}$$
$$[z_{k+2}, \cdots, z_m] = \left[z_{k+2}^\#, \cdots, z_m^\#\right]. \tag{3.23}$$

We will then prove the existence of a unitary matrix $\widehat{Q}' \in \mathbb{C}^{(k+1)\times(k+1)}$ and of a nonsingular matrix $\widehat{X}' \in \mathbb{C}^{(k+1)\times(k+1)}$ such that at the end of the $i+1$-th cycle

$$V_{k+1} = V_{k+1}^{\#}\widehat{Q}', \tag{3.24}$$

$$Z_{k+1} = Z_{k+1}^{\#}\widehat{X}', \tag{3.25}$$

$$[v_{k+2}, \cdots, v_{m+1}] = \left[v_{k+2}^{\#}, \cdots, v_{m+1}^{\#}\right], \tag{3.26}$$

$$[z_{k+2}, \cdots, z_m] = \left[z_{k+2}^{\#}, \cdots, z_m^{\#}\right]. \tag{3.27}$$

Regarding FGCRO-DR we assume that at the beginning of the $i+1$-th cycle (step 4)

$$\text{range}(W_m) = \text{range}(V_m). \tag{3.28}$$

We will also prove that relation (3.28) holds at the end of the $i+1$-th cycle. Note that relations (3.16), (3.17) and (3.28) are obviously satisfied before the first cycle, because steps 1 to 3 are identical in both algorithms yielding $V_{m+1} = V_{m+1}^{\#}$, $Z_m = Z_m^{\#}$ and $W_m = V_m$. Finally a consequence of (3.20), (3.22), (3.21) and (3.23) is that the residual of the linear system $Ax = b$ in both algorithms are equal at the beginning of the $i+1$-th cycle i.e. $r_0^{(i)} = r_0^{(i)\#}$. We will denote r_0 this residual for ease of notation.

Part I - Steps 5a and 5b. In this part, we prove that we can choose $Y_k^{\#} = Y_k$ with $Y_k = W_m P_k = V_m^{\#} P_k^{\#}$.

FGCRO-DR. Let $y_j = W_m p_j$ be the j-th column of Y_k. Since y_j is a harmonic Ritz vector of $AZ_m W_m^{\dagger}$ with respect to range(W_m), the following relation holds (see Definition (2.1))

$$(AZ_m W_m^{\dagger} W_m)^H \left(AZ_m W_m^{\dagger} y_j - \theta_j y_j\right) = 0$$

which is equivalent to

$$Z_m^H A^H \left(AZ_m p_j - \theta_j W_m p_j\right) = 0. \tag{3.29}$$

Due to (3.21) and (3.23) there exists a nonsingular matrix $X \in \mathbb{C}^{m\times m}$ that relates Z_m and $Z_m^{\#}$

$$Z_m = Z_m^{\#} X. \tag{3.30}$$

Using the last equality (3.30), the harmonic Ritz relation (3.29) now becomes

$$X^H {Z_m^{\#}}^H A^H \left(AZ_m^{\#} X p_j - \theta_j W_m p_j\right) = 0.$$

From Lemma 3.3 and relation (3.30) we deduce

$$X^H {Z_m^{\#}}^H A^H \left(AZ_m^{\#} X p_j - \theta_j {M_{W_m}^{(i)}}^{-1} Z_m p_j\right) = 0,$$

$$X^H {Z_m^{\#}}^H A^H \left(AZ_m^{\#} X p_j - \theta_j {M_{V_m^{\#}}^{(i)\#}}^{-1} Z_m^{\#} X p_j\right) = 0,$$

where we have used explicitly the assumption on the equivalent preconditioning matrix obtained at the end of the i-th cycle i.e. $M_{W_m}^{(i)} = M_{V_m^{\#}}^{(i)\#}$. Next, the application of Lemma 3.2 leads to

$$X^H {Z_m^{\#}}^H A^H \left(AZ_m^{\#} {V_m^{\#}}^H V_m^{\#} X p_j - \theta_j V_m^{\#} X p_j\right) = 0. \tag{3.31}$$

Since X is nonsingular the last equality proves that $V_m^{\#} X p_j$ is a harmonic Ritz vector of $AZ_m^{\#} {V_m^{\#}}^H$ with respect to range$(V_m^{\#})$ associated to the Ritz value θ_j. From relations (3.29) and (3.31) we deduce that the harmonic Ritz vectors can be chosen to be equal and correspond to the same harmonic Ritz values. In this case they notably satisfy the following equality

$$\forall j \in \{1, \cdots, k\}, \quad V_m^{\#} X p_j = W_m p_j \quad \text{i.e.} \quad p_j^{\#} = X p_j. \tag{3.32}$$

We will then denote $Y = Y_k^{\#} = Y_k$ the k harmonic Ritz vectors computed in either FGCRO-DR or FGMRES-DR. We assume that the harmonic Ritz values θ_j $(1 \leq j \leq k)$ are non zero.

Part IIa - Steps 6a to 10a, 6b to 10b. We show that at the end of steps 10a and 10b the following relations hold: $\text{range}(V_{k+1}) = \text{range}(V_{k+1}^{\#}) = \text{range}([Y, r_0^{(i)}/\|r_0^{(i)}\|])$. This result will help us to prove the existence of the matrix $\widehat{Q}'$ introduced in relation (3.24).

FGCRO-DR. $AZ_m p_j - \theta_j W_m p_j \in \text{range}(V_{m+1})$ is orthogonal to AZ_m due to the definition of the harmonic Ritz information (3.29). Since $r_0^{(i)} \in \text{range}(V_{m+1})$ is the non zero optimum residual at the i-th cycle, we have $(AZ_m)^H r_0^{(i)} = 0$. Thus there exists a coefficient $\alpha_j \in \mathbb{C}$ such that

$$AZ_m p_j - \theta_j W_m p_j = \alpha_j r_0^{(i)}. \tag{3.33}$$

Using $AZ_m = V_{m+1}\bar{H}_m$ and $QR = \bar{H}_m P_k$ we deduce

$$\begin{aligned} V_{m+1}\bar{H}_m p_j &= \theta_j W_m p_j + \alpha_j r_0^{(i)} \\ V_{m+1} Q\, R &= Y \text{diag}(\theta_1, \ldots, \theta_k) + r_0^{(i)} \alpha^T \end{aligned}$$

where $\alpha = [\alpha_1, \ldots, \alpha_k] \in \mathbb{C}^{k\times 1}$. This leads to

$$\begin{aligned} V_k R &= Y \text{diag}(\theta_1, \ldots, \theta_k) + r_0^{(i)} \alpha^T \\ \left[V_k, r_0^{(i)}/\|r_0^{(i)}\|\right] &= \left[Y, r_0^{(i)}/\|r_0^{(i)}\|\right] \begin{bmatrix} \text{diag}(\theta_1, \ldots \theta_k) R^{-1} & 0_{k\times 1} \\ \|r_0^{(i)}\| \alpha^T R^{-1} & 1 \end{bmatrix}. \end{aligned} \tag{3.34}$$

This relation leads to the following result

$$\text{range}(V_{k+1}) = \text{range}([Y, r_0^{(i)}/\|r_0^{(i)}\|]). \tag{3.35}$$

Similarly $W_{k+1} = [W_k, \frac{r_0^{(i)}}{\|r_0^{(i)}\|}]$ can be written as, using $Y = W_m P_k$

$$\begin{aligned} [W_k, r_0^{(i)}/\|r_0^{(i)}\|] &= [W_m P_k R^{-1}, \frac{r_0^{(i)}}{\|r_0^{(i)}\|}] \\ &= [Y R^{-1}, r_0^{(i)}/\|r_0^{(i)}\|] \\ &= [Y, r_0^{(i)}/\|r_0^{(i)}\|] \begin{bmatrix} R^{-1} & 0_{k\times 1} \\ 0_{1\times k} & 1 \end{bmatrix}. \end{aligned} \tag{3.36}$$

From relations (3.36) and (3.35) we deduce that

$$\text{range}(W_{k+1}) = \text{range}(V_{k+1}). \tag{3.37}$$

This last result also proves that $\text{range}(W_m) = \text{range}(V_m)$ at the end of the cycle.

FGMRES-DR. Using successively the relations of steps 8b and 6b and $r_0^{(i)} = V_{m+1}^{\#}(c^{\#} - \bar{H}_m^{\#} y^{\#*})$, we deduce

$$\begin{aligned} V_{k+1}^{\#} R^{\#} &= V_{m+1}^{\#} Q^{\#} R^{\#} \\ &= V_{m+1}^{\#} \begin{bmatrix} \begin{bmatrix} P_k^{\#} \\ 0_{1\times k} \end{bmatrix} & c^{\#} - \bar{H}_m^{\#} y^{\#*} \end{bmatrix} \\ &= [V_m^{\#} P_k^{\#}, r_0^{(i)}]. \end{aligned}$$

From the main result of Part I ($V_m^{\#} P_k^{\#} = Y$) we finally obtain

$$V_{k+1}^{\#} R^{\#} = [Y, r_0^{(i)}] = [Y, r_0^{(i)}/\|r_0^{(i)}\|] \begin{bmatrix} I_k & 0_{k\times 1} \\ 0_{1\times k} & \|r_0^{(i)}\| \end{bmatrix}. \tag{3.38}$$

Since $R^{\#}$ is nonsingular we deduce that

$$\text{range}(V_{k+1}^{\#}) = \text{range}([Y, r_0^{(i)}/\|r_0^{(i)}\|]). \tag{3.39}$$

Since both V_{k+1} and $V_{k+1}^{\#}$ have orthonormal columns we deduce from (3.35) and (3.39) that there exists a unitary matrix $\widehat{Q}'$ such that

$$V_{k+1} = V_{k+1}^{\#} \widehat{Q}' \tag{3.40}$$

which proves the relation proposed in equation (3.24).

Part IIb - Steps 6a to 10a, 6b to 10b. We show that at the end of steps 10a and 10b the following relations hold: range(Z_{k+1}) = range$(Z_{k+1}^{\#})$. This result will help us to prove the existence of the matrix $\widehat{X}'$ introduced in relation (3.25).

FGCRO-DR. Concerning $Z_{k+1} = [Z_k, z_{k+1}]$, there exists a nonsingular matrix $M^{(i+1)}_{[W_k, r_0^{(i)}/\|r_0^{(i)}\|]} \in \mathbb{C}^{n\times n}$ (see Lemma 3.1) such that

$$Z_{k+1} = M^{(i+1)}_{[W_k, r_0^{(i)}/\|r_0^{(i)}\|]}[W_k, r_0^{(i)}/\|r_0^{(i)}\|].$$

If $T \in \mathbb{C}^{(k+1)\times(k+1)}$ denotes the following triangular matrix

$$T = \begin{bmatrix} R & 0_{k\times 1} \\ 0_{1\times k} & 1 \end{bmatrix}$$

$Z_{k+1}T$ can be written as

$$\begin{aligned} Z_{k+1}T &= M^{(i+1)}_{[W_k, r_0^{(i)}/\|r_0^{(i)}\|]}[W_k, r_0^{(i)}/\|r_0^{(i)}\|]\ T \\ Z_{k+1}T &= M^{(i+1)}_{[W_k, r_0^{(i)}/\|r_0^{(i)}\|]}[Y, r_0^{(i)}/\|r_0^{(i)}\|]. \end{aligned} \tag{3.41}$$

where the last equality results from (3.36).

FGMRES-DR. Similarly from Lemma 3.2, $Z_{k+1}^{\#}$ can be expressed as

$$Z_{k+1}^{\#} = M^{(i+1)\#}_{V_{k+1}^{\#}} V_{k+1}^{\#}$$

where $M^{(i+1)\#}_{V_{k+1}^{\#}} \in \mathbb{C}^{n\times n}$ is nonsingular (see Lemma 3.1). If $T^{\#} \in \mathbb{C}^{(k+1)\times(k+1)}$ denotes the following triangular matrix

$$T^{\#} = R^{\#} \begin{bmatrix} I_k & 0_{k\times 1} \\ 0_{1\times k} & 1/\|r_0^{(i)}\| \end{bmatrix}$$

$Z_{k+1}^{\#}T^{\#}$ can be expressed as

$$Z_{k+1}^{\#}T^{\#} = M^{(i+1)\#}_{V_{k+1}^{\#}}[Y, r_0^{(i)}/\|r_0^{(i)}\|] \tag{3.42}$$

thanks to the relation (3.38). Relations (3.41) and (3.42) characterize $Z_{k+1}T$ and $Z_{k+1}^{\#}T^{\#}$ with respect to $[Y, r_0^{(i)}/\|r_0^{(i)}\|]$. We can further improve this result by showing the following equality

$$M^{(i+1)}_{[W_k, r_0^{(i)}/\|r_0^{(i)}\|]}[Y, r_0^{(i)}/\|r_0^{(i)}\|] = M^{(i+1)\#}_{V_{k+1}^{\#}}[Y, r_0^{(i)}/\|r_0^{(i)}\|]. \tag{3.43}$$

Lemma 3.3 and Lemma 3.2 respectively give us two useful relations for $M^{(i+1)}_{[W_k, r_0^{(i)}/\|r_0^{(i)}\|]}[Y, r_0^{(i)}/\|r_0^{(i)}\|]$ and $M^{(i+1)\#}_{V_{k+1}^{\#}}[Y, r_0^{(i)}/\|r_0^{(i)}\|]$ i.e.

$$M^{(i+1)}_{[W_k, r_0^{(i)}/\|r_0^{(i)}\|]}[Y, r_0^{(i)}/\|r_0^{(i)}\|] = [M^{(i)}_{W_m} Y,\ \mathcal{M}_{k+1}^{(i+1)}(r_0^{(i)}/\|r_0^{(i)}\|)] \tag{3.44}$$

$$M^{(i+1)\#}_{V_{k+1}^{\#}}[Y, r_0^{(i)}/\|r_0^{(i)}\|] = [M^{(i)\#}_{V_m^{\#}} Y,\ \mathcal{M}_{k+1}^{(i+1)^{\#}}(r_0^{(i)}/\|r_0^{(i)}\|)]. \tag{3.45}$$

Thus we investigate the relation between $M^{(i)}_{W_m} Y$ and $M^{(i)\#}_{V_m^{\#}} Y$. Using successively relation (3.32), $W_k = W_m P_k R^{-1}$, Lemma 3.3, $Z_k = Z_m P_k R^{-1}$, relation (3.30) and finally Lemma 3.2 the following

development can be made

$$\begin{aligned}
M_{W_m}^{(i)} Y &= M_{W_m}^{(i)} W_m P_k \\
M_{W_m}^{(i)} Y &= M_{W_m}^{(i)} W_k R \\
M_{W_m}^{(i)} Y &= Z_k R \\
M_{W_m}^{(i)} Y &= Z_m P_k \\
M_{W_m}^{(i)} Y &= Z_m^{\#} X P_k \\
M_{W_m}^{(i)} Y &= M_{V_m^{\#}}^{(i)\#} V_m^{\#} X P_k \\
M_{W_m}^{(i)} Y &= M_{V_m^{\#}}^{(i)\#} Y.
\end{aligned} \tag{3.46}$$

The fact that identical (possibly nonlinear) preconditioning operators are used in steps 10a and 10b of Algorithm 1 (see relation (3.1)) allows us to write

$$\mathcal{M}_{k+1}^{(i+1)}(r_0^{(i)}/\|r_0^{(i)}\|) = \mathcal{M}_{k+1}^{(i+1)^{\#}}(r_0^{(i)}/\|r_0^{(i)}\|). \tag{3.47}$$

Relations (3.46) and (3.47) finally show the relation (3.43). Consequently from relations (3.41), (3.42) and (3.43) we deduce that there exists a nonsingular matrix $\widehat{X}' \in \mathbb{C}^{(k+1)\times(k+1)}$ such that

$$Z_{k+1} = Z_{k+1}^{\#} \widehat{X}'. \tag{3.48}$$

This proves the relation proposed in equation (3.25). Since T and $T^{\#}$ are both triangular, we note that $\widehat{X}' = T^{\#} T^{-1}$ is also triangular.

Part IIIa - Steps 10a and 10b. We first show that $v_{k+2}^{\#} = v_{k+2}$ by expressing these two quantities in function of $r_0^{(i)}$ and Y.

FGCRO-DR. The Arnoldi relation (step 10a) yields $v_{k+2} = \bar{v}_{k+2}/||\bar{v}_{k+2}||$, where $\bar{v}_{k+2} = (I_n - v_{k+1}v_{k+1}^H)(I_n - V_k V_k^H) A\mathcal{M}_{k+1}^{(i+1)}(r_0^{(i)}/\|r_0^{(i)}\|)$. Since $r_0^{(i)}$ is the optimum residual at the i-th cycle, i.e. $(AZ_m)^H r_0^{(i)} = 0$ we have

$$\begin{aligned}
P_k^H (AZ_m)^H r_0^{(i)} &= 0, \\
(V_{m+1}\bar{H}_m P_k)^H r_0^{(i)} &= 0, \\
R^H V_k^H r_0^{(i)} &= 0.
\end{aligned}$$

This shows that $V_k^H v_{k+1} = 0$ since R is nonsingular. Therefore $(I_n - v_{k+1}v_{k+1}^H)$ and $(I_n - V_k V_k^H)$ commute and from Part IIa of the proof, the following expression can be derived

$$\bar{v}_{k+2} = \Pi_{V_{k+1}^{\perp}} A\mathcal{M}_{k+1}^{(i+1)}(r_0^{(i)}/\|r_0^{(i)}\|) = \Pi_{[Y, r_0^{(i)}/\|r_0^{(i)}\|]^{\perp}} A\mathcal{M}_{k+1}^{(i+1)}(r_0^{(i)}/\|r_0^{(i)}\|). \tag{3.49}$$

FGMRES-DR. The following expression for $v_{k+2}^{\#} = \bar{v}_{k+2}^{\#}/||\bar{v}_{k+2}^{\#}||$ is obtained using Lemma 3.5

$$\bar{v}_{k+2}^{\#} = (I_n - V_{k+1}^{\#} V_{k+1}^{\#H}) A\mathcal{M}_{k+1}^{(i+1)}(v_{k+1}^{\#}) = \Pi_{[Y, r_0^{(i)}/\|r_0^{(i)}\|]^{\perp}} A\mathcal{M}_{k+1}^{(i+1)}(\Pi_{Y^{\perp}} r_0^{(i)}/\|\Pi_{Y^{\perp}} r_0^{(i)}\|). \tag{3.50}$$

Due to the assumption (3.15) of Theorem 3.6 we deduce from (3.49) and (3.50) that $\bar{v}_{k+2} = \eta\ \bar{v}_{k+2}^{\#}$ with η positive and therefore $v_{k+2} = v_{k+2}^{\#}$.

Part IIIb - Steps 10a and 10b. In this part we continue the analysis of the Arnoldi procedure with flexible preconditioning and show that $v_{k+2+j} = v_{k+2+j}^{\#}$ for $j = 1, \ldots, m-k-1$.

For the case $j = 1$, we introduce $\bar{v}_{k+3}$ and $\bar{v}_{k+3}^{\#}$ such that $v_{k+3} = \bar{v}_{k+3}/||\bar{v}_{k+3}||$ and $v_{k+3}^{\#} = \bar{v}_{k+3}^{\#}/||\bar{v}_{k+3}^{\#}||$. The application of the Arnoldi procedure in both algorithms leads to

$$\begin{aligned}
\bar{v}_{k+3} &= (I_n - v_{k+2}v_{k+2}^H)(I_n - V_{k+1}V_{k+1}^H)\ A\mathcal{M}_{k+2}^{(i+1)}(\bar{v}_{k+2}) \\
\bar{v}_{k+3}^{\#} &= (I_n - v_{k+2}^{\#}v_{k+2}^{\#H})(I_n - V_{k+1}^{\#}V_{k+1}^{\#H})\ A\mathcal{M}_{k+2}^{(i+1)}(\bar{v}_{k+2}^{\#}).
\end{aligned}$$

Thus from Parts II and IIIa of the proof we obtain that v_{k+3} and $v_{k+3}^{\#}$ are equal. The proof can then be completed by induction.

Results from Parts II and III justify the relation (3.26) i.e. $[v_{k+2}, \cdots, v_{m+1}] = [v_{k+2}^{\#}, \cdots, v_{m+1}^{\#}]$. Consequently from Lemma 3.2, Lemma 3.4 and relation (3.1) we deduce the relation (3.27). This finally shows the main relations (3.16) and (3.17) of Theorem 3.6 that are satisfied at the end of the $i+1$-th cycle.
□

3.3.1. First consequence of Theorem 3.6. COROLLARY 3.7. *If the same flexible preconditioning operators are used in both Arnoldi procedures (steps 10a and 10b of Algorithm 1) and if at each cycle i there exists a real-valued positive coefficient η_i such that*

$$\Pi_{[Y, r_0^{(i-1)}/\|r_0^{(i-1)}\|]^{\perp}} A\mathcal{M}_{k+1}^{(i)}(\Pi_{Y^{\perp}} r_0^{(i-1)}/\|\Pi_{Y^{\perp}} r_0^{(i-1)}\|) = \eta_i \, \Pi_{[Y, r_0^{(i-1)}/\|r_0^{(i-1)}\|]^{\perp}} A\mathcal{M}_{k+1}^{(i)}(r_0^{(i-1)}/\|r_0^{(i-1)}\|),$$

FGCRO-DR and FGMRES-DR are algebraically equivalent.

Proof. We have already emphasized that $M_{W_m}^{(0)} = M_{V_m^{\#}}^{(0)\#}$ in relation (3.3). In Theorem 3.6 we have analyzed the $i+1$-th cycle of both algorithms assuming that $M_{W_m}^{(i)} = M_{V_m^{\#}}^{(i)\#}$. First we have proved in Part IIb the relation (3.43) and secondly in Parts IIIa and IIIb that $[v_{k+2}, \cdots, v_m] = [v_{k+2}^{\#}, \cdots, v_m^{\#}]$ and $[z_{k+2}, \cdots, z_m] = [z_{k+2}^{\#}, \cdots, z_m^{\#}]$ respectively. Consequently the same equivalent preconditioner matrix is obtained at the end of the $i+1$-th cycle i.e. $M_{W_m}^{(i+1)}$ and $M_{V_m^{\#}}^{(i+1)\#}$ can be chosen equal. We deduce that FGCRO-DR and FGMRES-DR are algebraically equivalent. □

3.3.2. About GCRO-DR and GMRES-DR. We propose a second consequence of Theorem 3.6 analyzed now with a fixed preconditioning matrix M. Before, a straightforward reformulation of Lemma 3.3 and Lemma 3.2 is proposed in this context.

LEMMA 3.8. *When a fixed right-preconditioning matrix M is used in FGCRO-DR, Z_m and W_m satisfy*

$$Z_m = MW_m. \tag{3.51}$$

Proof. The application of FGMRES(m) in the initialization phase (step 3 in Algorithm 2) leads to $Z_m = MV_m$ when a fixed preconditioning matrix M is used. Thus $M_{W_m}^{(0)} = M$. Suppose that at the end of the i-th cycle $M_{W_m}^{(i)} = M$. Since

$$\forall j, k+1 \leq j \leq m, \; M_j^{(i+1)} = M,$$

we obtain from Lemma 3.3 that $Z_m = M_{W_m}^{(i+1)} W_m = [MW_k, MW_{m-k}] = MW_m$ i.e. $M_{W_m}^{(i+1)}$ and M can be chosen equal. □

LEMMA 3.9. *When a fixed right-preconditioning matrix M is used in FGMRES-DR, $Z_m^{\#}$ and $V_m^{\#}$ satisfy*

$$Z_m^{\#} = MV_m^{\#}. \tag{3.52}$$

Proof. The proof follows the same steps as in Lemma 3.8 substituting $M_{V_m^{\#}}^{(i)\#}$ for $M_{W_m}^{(i)}$. □

The next corollary details an important result related to the GCRO-DR and GMRES-DR methods.

COROLLARY 3.10. *When a fixed right preconditioner is used, the GCRO-DR and GMRES-DR methods sketched in Algorithm 1 are unconditionally algebraically equivalent.*

Proof. We denote M the fixed right preconditioning operator. Exploiting partial results shown in Part IIa allows us to derive the following relation that holds during the $i+1$-th cycle:

$$AMY = Y \mathrm{diag}(\theta_1, \ldots, \theta_k) + r_0^{(i)} \alpha^T.$$

Thus

$$\Pi_{[Y,r_0^{(i)}]^\perp} AMY = 0. \tag{3.53}$$

From Part IIIa we know that

$$\bar{v}_{k+2} = \Pi_{[Y,r_0^{(i)}]^\perp} AMr_0^{(i)}. \tag{3.54}$$

Due to (3.53) we deduce the following development

$$\begin{aligned}
\bar{v}_{k+2} &= \Pi_{[Y,r_0^{(i)}]^\perp} AM(r_0^{(i)} - YY^\dagger r_0^{(i)}),\\
\bar{v}_{k+2} &= \Pi_{[Y,r_0^{(i)}]^\perp} AM\Pi_{Y^\perp} r_0^{(i)},\\
\bar{v}_{k+2} &= \bar{v}_{k+2}^{\#}.
\end{aligned}$$

By induction it is possible to deduce the rest of the proof regarding $\bar{v}_{k+j}, j > 2$. Using $\text{range}(V_{k+1}^{\#}) = \text{range}(V_{k+1})$ obtained in Part IIa we deduce that

$$\text{range}(V_m^{\#}) = \text{range}(V_m) = \text{range}(W_m). \tag{3.55}$$

From relation (3.55), Lemma 3.8 and Lemma 3.9 we deduce that

$$\text{range}(Z_m^{\#}) = \text{range}(Z_m).$$

Consequently the minimization problem $\min \|r_0^{(i)} - AZ_m y\|$ leads to the same solution for both algorithms at each cycle: GCRO-DR and GMRES-DR sketched in Algorithm 1 are thus unconditionally algebraically equivalent. □

4. Variants of FGCRO-DR. We explore variants of FGCRO-DR that only differ in the formulation of the generalized eigenvalue problem for the harmonic Ritz information. Their computational cost is detailed carefully and their behaviour with respect to fixed preconditioning is finally investigated.

4.1. Derivation and algorithms. In Section 2.3 the deflation procedure relied on the use of k harmonic Ritz vectors of $AZ_m W_m^\dagger$ with respect to $\text{range}(W_m)$, where W_m satisfies the property shown in Lemma 3.3. It is however possible to derive other variants of FGCRO-DR by choosing differently the way the harmonic Ritz information is selected. Indeed at each cycle, Z_m and V_m are also available and it seems natural to exploit this feature. Thus variants of FGCRO-DR can be deduced by computing either k harmonic Ritz vectors of $AZ_m Z_m^\dagger$ with respect to $\text{range}(Z_m)$ or k harmonic Ritz vectors of $AZ_m V_m^H$ with respect to $\text{range}(V_m)$. We summarize the different variants in Algorithm 2 with the same notations as in Algorithm 1. Strategy A corresponds to the algorithm first presented in [18]. As far as we know, Strategies B and C are new. We note that Strategy C has been introduced in Section 2.3 and equivalence with FGMRES-DR has been discussed in Section 3. The harmonic Ritz formulation of Strategy B has been inspired by step 5b of Algorithm 1.

4.2. Computational cost. We first detail the computational cost related to the harmonic Ritz information (step 5 of Algorithm 2) since this is the main difference between the proposed strategies.

4.2.1. Harmonic Ritz information.

Strategy A. The generalized eigenvalue problem of Strategy A presented in Algorithm 2 is

$$(AZ_m)^H\ (AZ_m)y = \theta (AZ_m)^H Z_m y,$$

where $\theta \in \mathbb{C}$ and $y \in \mathbb{C}^m$. Using the Arnoldi-like relation (2.3) it can be written as

$$\bar{H}_m^H\ \bar{H}_m y = \theta \bar{H}_m^H V_{m+1}^H Z_m y.$$

Algorithm 2 Flexible GCRO-DR(m,k) algorithms: strategies A, B and C.

1: choose m, k, tol and x_0
2: $r_0 = b - Ax_0$, $\beta = \|r_0\|$, $v_1 = r_0/\beta$, $i \leftarrow 0$
3: Flexible GMRES(m) yields $\bar{H}_m$, Z_m, V_{m+1} such that $AZ_m = V_{m+1}\bar{H}_m$, $y^* = \arg\min_{y\in\mathbb{C}^m} \|c - \bar{H}_m y\|$, $c = \beta e_1$, $x_0^{(0)} = x_0 + Z_m y^*$, $r_0^{(0)} = b - Ax_0^{(0)} = V_{m+1}(c - \bar{H}_m y^*)$, $W_m = V_m$ (only for Strategy C)
4: **while** $\|r_0^{(i)}\| > \|b\| \times tol$ **do** $i \leftarrow i + 1$
5: Compute k eigenvectors of the generalized eigenvalue problem $Dy = \theta Ey$ and store them in P_k.

Strategy A	**Strategy B**	**Strategy C**
• $D = Z_m^H A^H A Z_m$	• $D = Z_m^H A^H A Z_m$	• $D = Z_m^H A^H A Z_m$
• $E = Z_m^H A^H Z_m$	• $E = Z_m^H A^H V_m$	• $E = Z_m^H A^H W_m$
• $Y_k = Z_m P_k$	• $Y_k = V_m P_k$	• $Y_k = W_m P_k$

6: $Q\ R = \bar{H}_m P_k$
7: $W_k = W_m P_k R^{-1}$ (only for Strategy C)
8: $V_k = V_{m+1} Q$
9: $Z_k = Z_m P_k R^{-1}$
10: Apply $m-k$ flexible preconditioned Arnoldi steps with $(I_n - V_k V_k{}^H)A$ and $v_{k+1} = r_0^{(i-1)}/\|r_0^{(i-1)}\|$ such that $(I_n - V_k V_k{}^H)A\,[z_{k+1}, \ldots, z_m] = [v_{k+1}, \ldots, v_{m+1}]\,\bar{H}_{m-k}$ with $z_j = \mathcal{M}_j^{(i)}(v_j)$
11: $d^* = \arg\min_{d\in Z_m} \|r_0^{(i-1)} - Ad\|$, $x_0^{(i)} = x_0^{(i-1)} + d^*$, $r_0^{(i)} = b - Ax_0^{(i)}$
12: $W_m = \left[W_m P_k R^{-1} \quad V_m(1:n, k+1:m)\right]$ (only for Strategy C)
13: **end while**

The computation of $\bar{H}_m^H\ \bar{H}_m$ is cheap since it only involves a matrix of size $(m+1)\times m$, where m is supposed to be small with respect to the problem size n. A block form for $V_{m+1}^H Z_m$ can be found as

$$V_{m+1}^H\ Z_m = \begin{bmatrix} V_k^H Z_k & V_k^H Z_{m-k} \\ V_{m-k+1}^H Z_k & V_{m-k+1}^H Z_{m-k} \end{bmatrix}. \tag{4.1}$$

Thanks to steps 8 and 9 in Algorithm 2, $V_k^H Z_k$ can be also written as

$$(V_k^H Z_k)^{(i)} = Q^H\ (V_{m+1}^H Z_m)^{(i-1)}\ P_k R^{-1}$$

where the superscript is related to the cycle index. Thus storing $V_{m+1}^H Z_m$ at the end of each cycle allows us to compute at a cheap cost a $k\times k$ block of $V_{m+1}^H Z_m$ for the next cycle. Computing the other blocks of $V_{m+1}^H Z_m$ require $2n(m-k+1)m + 2nk(m-k)$ operations.

Strategy B. Similarly the generalized eigenproblem can be written as

$$\bar{H}_m^H\ \bar{H}_m y = \theta \bar{H}_m^H V_{m+1}^H V_m y. \tag{4.2}$$

Exploiting the fact that V_{m+1} has orthonormal columns finally leads to the generalized eigenproblem

$$\bar{H}_m^H\ \bar{H}_m y = \theta \bar{H}_m^H \begin{bmatrix} y \\ 0_{1\times 1} \end{bmatrix}, \tag{4.3}$$

which involves only a matrix of size $(m+1)\times m$.

Strategy C. The corresponding generalized eigenvalue problem can be written as

$$\bar{H}_m^H\ \bar{H}_m y = \theta \bar{H}_m^H V_{m+1}^H W_m y. \tag{4.4}$$

Since $W_m = [W_{k+1}, v_{k+2}, \cdots, v_m]$ (step 12 of Algorithm 2) a new form for $V_{m+1}^H W_m$ can be found as

$$V_{m+1}^H W_m = \begin{bmatrix} V_{k+1}^H\ W_{k+1} & 0_{(k+1)\times(m-k-1)} \\ 0_{(m-k-1)\times(k+1)} & I_{m-k-1} \\ 0_{1\times(k+1)} & 0_{1\times(m-k-1)} \end{bmatrix}. \tag{4.5}$$

The structure of the $(k+1) \times (k+1)$ block $V_{k+1}^H \; W_{k+1}$ is as follows

$$V_{k+1}^H \; W_{k+1} = \begin{bmatrix} V_k^H \; W_k & V_k^H \; w_{k+1} \\ v_{k+1}^H \; W_k & v_{k+1}^H \; w_{k+1} \end{bmatrix} = \begin{bmatrix} V_k^H \; W_k & 0_{k\times 1} \\ v_{k+1}^H \; W_k & 1 \end{bmatrix}.$$

Thanks to steps 7 and 8 in Algorithm 2, $V_k^H \; W_k$ is a $k \times k$ matrix that satisfies the following relation

$$(V_k^H W_k)^{(i)} = Q^H \; (V_{m+1}^H W_m)^{(i-1)} \; P_k R^{-1}$$

where the superscript is related to the cycle index. Thus storing the $(m+1) \times m$ matrix $V_{m+1}^H W_m$ at the end of each cycle can be used to slightly reduce the cost of computing the new matrix $V_{m+1}^H W_m$. It is then sufficient to compute $v_{k+1}^H \; W_k$ at a cost of $2nk$ operations. Comparing (4.1) and (4.5) reveals that Strategy C requires less operations than Strategy A for computing the pair of matrices of the generalized eigenvalue problem. Nevertheless Strategy C requires the additional storage of W_k i.e. k additional vectors of length n (step 7 of Algorithm 2).

4.2.2. Cost of a cycle. We summarize in Table 4.1 the main computational costs associated with each step of the three strategies proposed in Algorithm 2. An Arnoldi method based on the modified Gram-Schmidt procedure has been assumed[1]. We have only included the costs proportional to the size of the original problem n which is supposed to be much greater than m and k. These costs exclude the cost related to both matrix-vector products and preconditioning operations.

Step	Strategy A	Strategy B	Strategy C
5	$2n(m-k+1)m+$ $2nk(m-k)$	-	$2nk$
6	-	-	-
7	-	-	$2nmk$
8	$2n(m+1)k$	$2n(m+1)k$	$2n(m+1)k$
9	$2nmk$	$2nmk$	$2nmk$
10	$(4nk+n)(m-k)+$ $2n(m-k)(m-k+1)+$ $3n(m-k)$	$(4nk+n)(m-k)+$ $2n(m-k)(m-k+1)+$ $3n(m-k)$	$(4nk+n)(m-k)+$ $2n(m-k)(m-k+1)+$ $3n(m-k)$
Total	$C_B + 2n(m^2-k^2+m)$	C_B	$C_B + 2n(k+mk)$

TABLE 4.1

Computational cost of a cycle detailed for each strategy and for each step of a given cycle of Algorithm 2. This excludes the cost of matrix-vector operations and preconditioning operations. The total cost of Strategy B is $C_B = 2n((m+k)^2 - 2k^2 + 3m - 2k)$.

As remarked in Section 4.2.1 Strategy B involves the lowest computational cost among the three variants. Concerning Strategy A and Strategy C it is then interesting to analyze the corresponding additional costs versus m and k. As a first illustration Table 4.2 details three different cases i.e. $k = 1$, $k = m/2$ and $k = m - 1$. When $k = 1$ we remark that Strategy A is the most expensive one. This is mainly due to the construction of $V_{m+1}^H \; Z_m$ in step 5 of Algorithm 2. The additional cost is of order $O(nm^2)$. For the case $k = m - 1$ we note that the additional cost for Strategy C - now the most expensive one - also behaves as $O(nm^2)$.

4.3. Preconditioning. Although the primary focus of this paper is on flexible methods, we propose now two comments on Strategies A, B and C when a nonvariable preconditioner is used. In this setting we note that Strategy A corresponds to the method originally proposed by Parks et al. in [26].

[1] Step 10: during this step the action of $(I_n - V_k V_k^H)$ requires $\sum_{j=k+1}^{m}(4nk+n)$ operations, the Arnoldi method based on modified Gram-Schmidt requires $\sum_{j=k+1}^{m}\sum_{i=k+1}^{j}(4n)$ operations whereas norm computation and normalization cost $\sum_{j=k+1}^{m}(3n)$ operations.

k	Strategy A	Strategy B	Strategy C
1	$C_B^{(1)} + 2n(m^2+m-1)$	$C_B^{(1)} = 2n(m^2+5m-4)$	$C_B^{(1)} + 2n(m+1)$
$m/2$	$C_B^{(m/2)} + 2n(3m^2/4+m)$	$C_B^{(m/2)} = 2n(7m^2/4+2m)$	$C_B^{(m/2)} + 2n(m^2/2+m/2)$
$m-1$	$C_B^{(m-1)} + 2n(3m-1)$	$C_B^{(m-1)} = 2n(2m^2+m+1)$	$C_B^{(m-1)} + 2n(m^2-1)$

TABLE 4.2

Computational cost of a cycle for $k=1$, $k=m/2$ and $k=m-1$. $C_B^{(k)}$ denotes the cost of Strategy B when k harmonic Ritz vectors are used in the deflation procedure.

4.3.1. Behaviour in case of no preconditioning. When no preconditioning occurs, we have the following relation $Z_m = W_m$ thanks to Lemma 3.8. Thus Strategies A and C are equivalent in this case. Since Strategy C is algebraically equivalent to GMRES-DR (Corollary 3.10), we deduce that Strategy A is also algebraically equivalent to GMRES-DR. This shows a remark made by Parks et al. in [26, page 1657]. We refer the reader to Table 5.1 in Section 5 for a numerical illustration. We note that the equivalence between Strategy A and GMRES-DR does *not* hold when preconditioning occurs as will be shown in Section 4.3.2.

4.3.2. Behaviour in case of fixed preconditioning. Suppose that a fixed preconditioner M is used as a right preconditioner for the solution of (2.1). A desirable feature is that applying the Krylov subspace method either on A with right preconditioner M or on $\tilde{A} = AM$ without any preconditioner leads to the same iterates when the same right-hand side is considered. We call this property *right-preconditioning invariance*. We note that GMRES(m) with right-preconditioning satisfies this property. The application of GMRES(m) in the initialization phase (step 3 in Algorithm 2) leads to the relation $Z_m = MV_m$ when a fixed right-preconditioner is used. Table 4.3 collects the different formulations of the *first* generalized eigenvalue problem, where we have used the $Z_m = MV_m$ relation explicitly.

Strategy	Fixed preconditioning matrix M	Equivalent matrix $\tilde{A} = AM$
A	$(AMV_m)^H(AMV_m)y = \theta\ (AMV_m)^H MV_m y$	$(\tilde{A}V_m)^H(\tilde{A}V_m)y = \theta\ (\tilde{A}V_m)^H MV_m y$
B	$(AMV_m)^H(AMV_m)y = \theta\ (AMV_m)^H V_m y$	$(\tilde{A}V_m)^H(\tilde{A}V_m)y = \theta\ (\tilde{A}V_m)^H V_m y$
C	$(AMV_m)^H(AMV_m)y = \theta\ (AMV_m)^H W_m y$	$(\tilde{A}V_m)^H(\tilde{A}V_m)y = \theta\ (\tilde{A}V_m)^H W_m y$

TABLE 4.3

Formulations of the first generalized eigenvalue problem when a fixed right-preconditioning matrix M is used (center) and when an equivalent preconditioned matrix $\tilde{A} = AM$ is used (right) for strategies A, B and C.

From Table 4.3 it can be suspected that Strategy A is not right-preconditioning invariant since this property is not satisfied during the first cycle of the method. A numerical illustration is given in Figure 4.1. In Section 4.2 the generalized eigenvalue problems of Strategies B and C ((4.2) and (4.4) respectively) only involve $\bar{H}_m$, V_{m+1} or W_m i.e. quantities that are preconditioning invariant. This is confirmed in Figures 4.2 and 4.3, where - as expected - Strategies B and C are right-preconditioning invariant.

5. Numerical experiments. We present numerical experiments for a specific class of problems from quantum chromodynamics (QCD). This area is subject to active research to design robust and efficient subspace methods for the efficient approximation of $f(A)\ b$, where f is a function defined on the spectrum of A [16, 38]. Methods based on variable preconditioning [8] have been proven efficient when considering the sign function. Recently adaptive algebraic multigrid methods [5, 6] have been also proposed for the solution of such nearly singular and highly disordered physical systems. We focus here on the solution of a single linear system and investigate the behaviour of various flexible methods with deflated restarting.

5.1. Lattice quantum chromodynamics. Quantum chromodynamics [7] is the fundamental theory explaining how neutrons and protons are bound inside nuclei and how their constituents

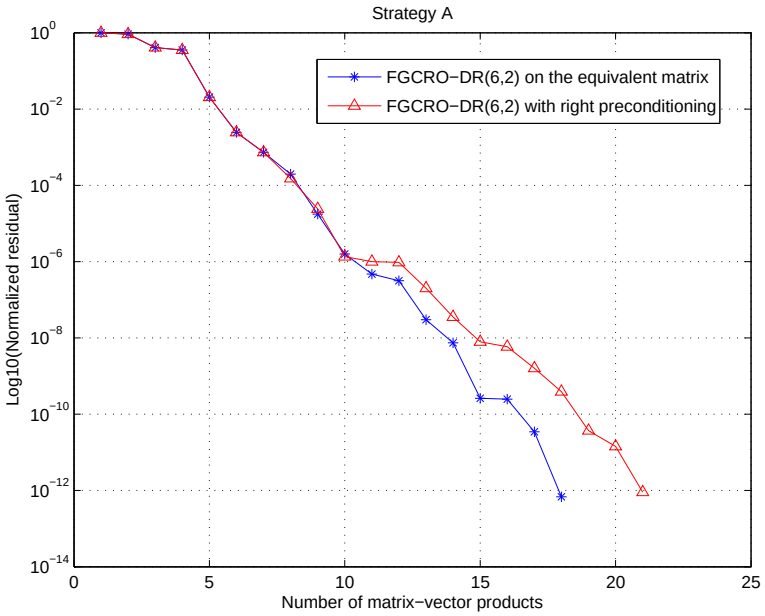


FIG. 4.1. *Strategy A of FGCRO-DR. Behaviour in case of fixed preconditioning. Convergence history of FGCRO-DR(6,2) on the equivalent preconditioned system $AM\phi = b$ and on the original system $Ax = b$ with right preconditioning matrix M. $A \in \mathbb{C}^{400\times 400}$ is here a nonsingular sparse random triangular matrix and Jacobi preconditioning is considered. The right-hand side b is a random vector of unit norm.*

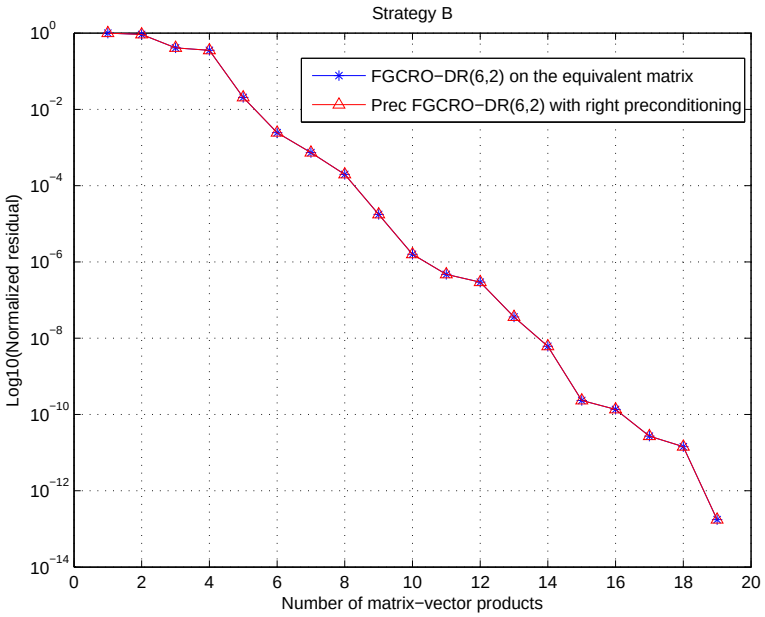


FIG. 4.2. *Strategy B of FGCRO-DR. Behaviour in case of fixed preconditioning. Convergence history of FGCRO-DR(6,2) on the equivalent preconditioned system $AM\phi = b$ and on the original system $Ax = b$ with right preconditioning matrix M. The same linear systems as in Figure 4.1 are considered here.*

- quarks and gluons - interact. Numerical simulations on a four-dimensional hypercube space-time lattice are most often considered as a unique way to solve QCD ab initio [37]. The Wilson fermion matrix - representing periodic nearest neighbour coupling - has the following block structure after

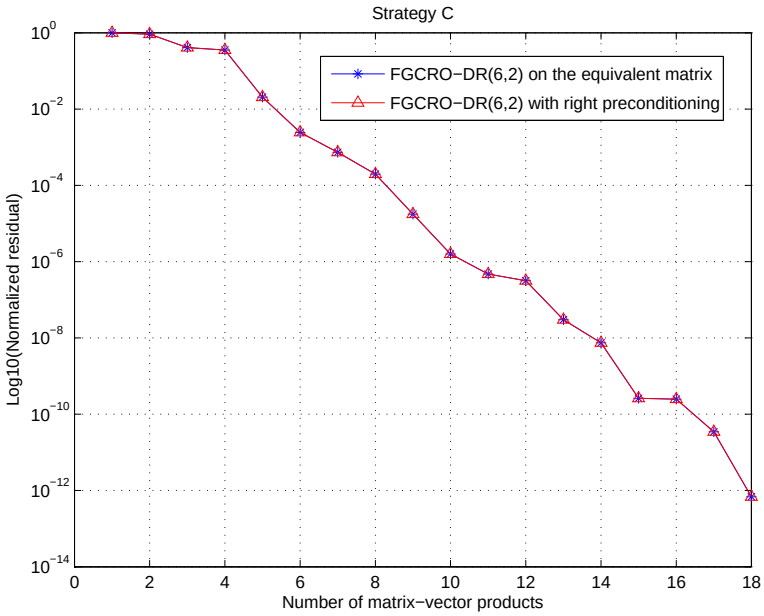


FIG. 4.3. *Strategy C of FGCRO-DR. Behaviour in case of fixed preconditioning. Convergence history of FGCRO-DR(6,2) on the equivalent preconditioned system $AM\phi = b$ and on the original system $Ax = b$ with right preconditioning matrix M. The same linear systems as in Figure 4.1 are considered here.*

a red-black (also named odd-even) ordering of the lattice points [15]

$$A = I_n - \kappa \begin{bmatrix} 0_{n/2\times n/2} & D_b \\ D_r & 0_{n/2\times n/2} \end{bmatrix} = \begin{bmatrix} I_{n/2} & -\kappa D_b \\ -\kappa D_r & I_{n/2} \end{bmatrix} \tag{5.1}$$

where the hopping parameter κ is a real valued positive parameter. The Wilson fermion matrix $A \in \mathbb{C}^{n\times n}$ is a sparse, complex non-Hermitian matrix. It is positive definite as long as $0 \leq \kappa < \kappa_c$. Physically interesting cases are for κ close to the critical parameter κ_c. As a model problem we have used the matrix conf5.0_0014x4.1000.mtx submitted by B. Medeke and publicly available from the Matrix Market collection[2]. This sparse matrix of order 3072 contains 39 nonzero elements per row. The numerical tests were performed on a personal computer running Linux (Intel Dual Core, 2.13 Ghz with 2 GB of memory) using Matlab version 7.1 (release 14).

5.1.1. Solution of the QCD reduced system. QCD computations rely on the use of odd-even preconditioning that aims at exploiting the block structure presented in (5.1). Denoting L and U the strictly lower and triangular parts of A respectively, this odd-even technique is equivalent to apply SSOR preconditioning to the original linear system $Ax = b$ as

$$(I_n - L)^{-1} A (I_n - U)^{-1} y = (I_n - L)^{-1} b \quad \text{with} \quad y = (I_n - U)x. \tag{5.2}$$

This leads to the following linear system

$$\begin{bmatrix} I_{n/2} & 0_{n/2\times n/2} \\ 0_{n/2\times n/2} & I_{n/2} - \kappa^2 D_r D_b \end{bmatrix} \begin{bmatrix} y_r \\ y_b \end{bmatrix} = \begin{bmatrix} b_r \\ b_b + \kappa D_r b_r \end{bmatrix}. \tag{5.3}$$

Thanks to this decoupling, physicists focus on developing efficient methods for the numerical solution of the reduced system

$$(I_{n/2} - \kappa^2 D_r D_b) y_b = b_b + \kappa D_r b_r \tag{5.4}$$

which can be also seen as the Schur complement system of (5.1). Consequently we will next compare different numerical methods for the solution of the reduced system (5.4). The right-hand side b is chosen as the first Cartesian basis vector of $\mathbb{C}^n$. A zero initial iterate is considered as an initial guess and all solvers are required to reduce the true residual to 1.0×10^{-12}.

	$\kappa = 0.200$	$\kappa = 0.202$	$\kappa = 0.204$	$\kappa = 0.206$
GMRES(20)	330	418	550	770
GMRES-DR(20,16)	268	304	334	376
GCRO-DR(20,16) Strategy A	268	304	334	376
GCRO-DR(20,16) Strategy B	274	310	340	394
GCRO-DR(20,16) Strategy C	268	304	334	376
FGMRES(20)	226	328	430	532
FGMRES-DR(20,16)	176	220	242	264
FGCRO-DR(20,16) Strategy A	176	198	220	264
FGCRO-DR(20,16) Strategy B	176	198	242	286
FGCRO-DR(20,16) Strategy C	176	198	242	264

TABLE 5.1
Total number of matrix-vector products required to solve the QCD linear system (5.4) for different values of κ.

In Table 5.1 we collect the total number of matrix-vector products occurring in the different methods for four different linear systems corresponding to increasing values of κ. We give both unpreconditioned and preconditioned convergence results, where all the methods minimize over a subspace of dimension 20 in each cycle. For the variants related to deflated restarting we have fixed the value of k to 16 and selected the eigenvectors related to the k smallest eigenvalues in modulus. In the first five lines of Table 5.1 we have included results related to GMRES, GMRES with deflated restarting and GCRO with deflated restarting. We note that deflated restarting leads to a dramatic improvement with respect to standard restarting. This behaviour has been already observed in [22] for applications in QCD. Strategies A and C lead to the same number of matrix-vector products as for GMRES-DR. This is due to the equivalence discussed in Section 4.3.1. Indeed it has been checked that the three methods produce iterates that are equal up to the machine precision at each restart. Figure 5.1 shows that $\Pi_{[Y_k, r_0^{(i-1)}/\|r_0^{(i-1)}\|]^\perp} A\Pi_{Y_k^\perp} r_0^{(i-1)}/\|\Pi_{Y_k^\perp} r_0^{(i-1)}\|$ and $\Pi_{[Y_k, r_0^{(i-1)}/\|r_0^{(i)}\|]^\perp} A r_0^{(i-1)}/\|r_0^{(i-1)}\|$ are collinear as stated in Corollary 3.10. We also remark that this collinearity property does not hold for the pairs of vectors $(A\Pi_{Y_k^\perp} r_0^{(i-1)}, Ar_0^{(i-1)}/\|r_0^{(i-1)}\|)$ and $(\Pi_{Y_k^\perp} r_0^{(i-1)}, r_0^{(i-1)})$ respectively.

We discuss next the case of flexible methods. We consider the following algorithms: FGMRES, FGMRES with deflated restarting and the three different strategies related to FGCRO-DR presented in Section 4. As a variable preconditioner we consider four iterations of unpreconditioned GMRES. We can notice that flexible variants with deflated restarting lead to additional reductions in terms of matrix-vector products. Variants of FGCRO-DR are most often as efficient as FGMRES-DR, if not better. Thus the interest of the new algorithm FGCRO-DR has been shown already in the case of a single linear system on this application. Furthermore one primary advantage of FGCRO-DR is its ability to handle the solution of linear systems given in sequence. We plan to illustrate this feature in a future research. For this specific choice of m and k parameters, the lowest total number of matrix-vector products always corresponds to Strategy A. As shown in Table 5.2 the vectors $\Pi_{[Y_k, r_0^{(i-1)}]^\perp} A\mathcal{M}_{k+1}^{(i)}(\Pi_{Y_k^\perp} r_0^{(i-1)})$ and $\Pi_{[Y_k, r_0^{(i-1)}]^\perp} A\mathcal{M}_{k+1}^{(i)}(r_0^{(i-1)})$ obtained in Strategy C are not collinear. This might explain why Strategy C and FGMRES-DR(20,16) are not algebraically equivalent in this case.

5.1.2. Computational cost of Strategies A, B and C. We detail now the behaviour of Strategies A, B and C when a variable preconditioning is considered with a fixed value of the restart parameter $m = 20$ and k varying. We select the case of $\kappa = 0.206$ which is the most challenging as shown in Table 5.1. Table 5.3 collects the total number of matrix-vector products and the normalized global computational cost of each strategy for varying k such that $1 \leq k \leq m-1$. To produce a fair comparison between the three strategies this global cost includes both the cost detailed in Section 4.2.2 and the one related to matrix-vector products

[2] http://math.nist.gov/MatrixMarket/data/misc/qcd/qcd.html

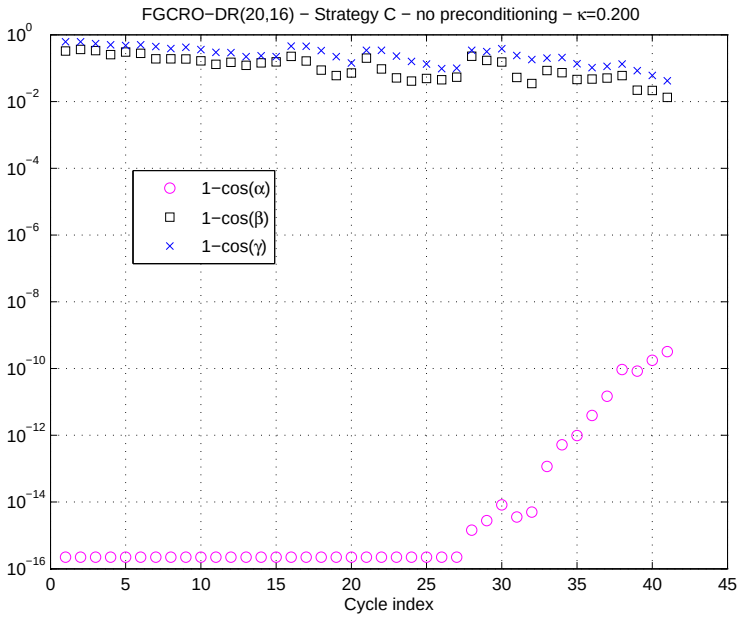


FIG. 5.1. $1-\cos(\alpha)$, $1-\cos(\beta)$, $1-\cos(\gamma)$ *with* $\alpha = \angle\,(\,\Pi_{[Y_k, r_0^{(i-1)}/\|r_0^{(i-1)}\|]^\perp} A\Pi_{Y_k^\perp} r_0^{(i-1)}/\|r_0^{(i-1)}\|,\ \Pi_{[Y_k, r_0^{(i-1)}]^\perp} A r_0^{(i-1)})$ *(circle),* $\beta = \angle\,(\,A\Pi_{Y_k^\perp} r_0^{(i-1)},\ A r_0^{(i-1)}/\|r_0^{(i-1)}\|)$ *(square) and* $\gamma = \angle\,(\,\Pi_{Y_k^\perp} r_0^{(i-1)},\ r_0^{(i-1)})$ *(x-mark) computed at each cycle* i *of FGCRO-DR(20,16) Strategy C obtained during the convergence history. Case of no preconditioning and* $\kappa = 0.200$.

Cycle i	$\kappa = 0.200$ $\cos(\alpha)$	$\kappa = 0.202$ $\cos(\alpha)$	$\kappa = 0.204$ $\cos(\alpha)$	$\kappa = 0.206$ $\cos(\alpha)$
1	0.97351791374476	0.95990773398568	0.95010035729778	0.92807215131352
2	0.95508546906640	0.93164336826626	0.91315357530771	0.86734097857504
3	0.98489849221921	0.97991760584828	0.96982159718869	0.95912459647826
4	0.96879881841919	0.97624255588142	0.96968106888483	0.93832538312347
5	0.93545030140439	0.94039492520401	0.92709902729626	0.88370626674617
6	0.96724690697981	0.97051191983231	0.90407615323902	0.89706063488241
7	0.97170523764768	0.98690054632099	0.98036433795137	0.94843268286296
8		0.94566451575530	0.97935681716533	0.98429649100801
9			0.96995038021057	0.99427074769599
10			0.96271377917353	0.94342942135706
11				0.90360537717426

TABLE 5.2

Cosinus of $\alpha = \angle\,(\,\Pi_{[Y_k, r_0^{(i-1)}]^\perp} A\mathcal{M}_{k+1}^{(i)}(\Pi_{Y_k^\perp} r_0^{(i-1)}),\ \Pi_{[Y_k, r_0^{(i-1)}]^\perp} A\mathcal{M}_{k+1}^{(i)}(r_0^{(i-1)}))$ *computed at each cycle of Strategy C of FGCRO-DR(20,16) obtained during the convergence history. Case of a flexible preconditioner.*

and preconditioning operations. We will denote C_A^g, C_B^g and C_C^g these global costs - related to Strategies A, B and C respectively.

The lowest computational cost obtained corresponds to $k = 13$, $k = 14$ and $k = 16$ for Strategies A, B and C respectively. A relatively large number of harmonic Ritz values is thus required to yield an efficient method. This is in agreement with previous numerical experiments [22]. Bold values in each line of Table 5.3 correspond to the lowest computational cost among the three strategies for a given value of k. It can be noticed that most often a given strategy is more interesting on a certain range of harmonic Ritz values. The interest of the three strategies has been demonstrated on this application in QCD, since the optimal cost for each strategy (20047, 22120, 21728 for Strategies A, B and C respectively) only differ by less than 10%. In addition, when k

k	$\#Mvp$	$C_A^g/(2n)$	$\#Mvp$	$C_B^g/(2n)$	$\#Mvp$	$C_C^g/(2n)$
1	507	40405	410	**32420**	507	40007
2	390	31314	390	**30898**	390	30940
3	370	29785	370	**29374**	370	29437
4	350	28252	432	34244	350	**27932**
5	330	26715	407	32326	330	**26425**
6	310	25174	382	30406	310	**24916**
7	290	23629	357	28484	290	**23405**
8	270	22080	394	31396	270	**21892**
9	307	24973	364	29080	307	**24823**
10	282	23026	334	26762	282	**22916**
11	257	**21075**	304	24442	304	24673
12	274	22396	274	**22120**	274	22372
13	244	**20047**	281	22682	281	22955
14	278	22686	278	**22462**	278	22756
15	265	**21655**	292	23566	265	21775
16	264	**21556**	286	23108	264	21728
17	260	**21219**	294	23740	277	22771
18	262	**21346**	286	23122	274	22564
19	323	26071	358	28742	316	**25865**

TABLE 5.3

Number of matrix-vector products $\#Mvp$ and normalized global computational cost of Strategies A, B and C when solving the QCD linear system (5.4) for $\kappa = 0.206$ and for a variable number of harmonic Ritz values k. Case of a flexible preconditioner.

varies, Strategy A is the best on six cases, Strategy B on five cases and Strategy C on eight cases, indicating again the potential of each of these approaches.

6. Conclusion and perspectives. In this paper we have studied a new minimum residual norm subspace method with deflated restarting that allows flexible preconditioning based on the GCRO subspace method. The resulting method named FGCRO-DR has been presented together with FGMRES-DR, a recently proposed algorithm of the same family but based on the GMRES subspace method. A theoretical comparison analysis of both algorithms has been performed in Section 3. Theorem 3.6 also proves the algebraic equivalence of GMRES-DR and GCRO-DR when a fixed preconditioner is used. Furthermore three variants of the new algorithm - that only differ in the formulation of the generalized eigenvalue problem for the harmonic Ritz information - have been introduced and analyzed in Section 5. Numerical experiments on a challenging application in quantum chromodynamics have shown the interest of these new variants when solving a given linear system.

We have restricted the presentation to the case of a linear system with a single right-hand side. In [26] reusing selected subspaces in GCRO-DR - in the case of fixed preconditioning - has been proved efficient when solving sequence of linear systems where both the left- or right-hand sides could change. A natural perspective could be thus to investigate the numerical properties of FGCRO-DR in this setting. This seems to be especially appealing for applications related to e.g. stochastic finite element methods [12, 36] in three dimensions where variable preconditioning using approximate solvers has to be usually considered. When all right-hand sides are available simultaneously and when the matrix is fixed, block subspace methods may be also suitable. Thus a perspective could be to propose a block variant of FGCRO-DR.

Acknowledgments. The participation of the first author in the initial preparation of this article was guaranteed thanks to grant CNPq-473420/2007-4, coordinated by Professor Nelson Maculan. The first author would like to thank the warm welcome he received at CERFACS, in the Parallel Algorithms Team led by Iain S. Duff, during his sabbatical leave where this work was completed.

REFERENCES

[1] O. Axelsson. *Iterative solution methods.* Cambridge University Press, 1994.

[2] O. Axelsson and P. S. Vassilevski. A black box generalized conjugate gradient solver with inner iterations and variable-step preconditioning. *SIAM J. Matrix Analysis and Applications*, 12(4):625–644, 1991.

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

[4] A. H. Baker, E. R. Jessup, and T. Manteuffel. A technique for accelerating the convergence of restarted GMRES. *SIAM J. Matrix Analysis and Applications*, 26(4):962–984, 2005.

[5] J. Brannick, M. Brezina, D. Keyes, O. Livne, I. Livshits, S. MacLachlan, T. Manteuffel, S. McCormick, J. Ruge, and L. Zikatanov. Adaptive smoothed aggregation in lattice QCD. In Olof B. Widlund and David E. Keyes, editors, *Domain Decomposition Methods in Science and Engineering XVI, Lecture Notes in Computational Science and Engineering, volume 55*, pages 505–512. Springer-Verlag, Berlin, 2007.

[6] J. Brannick, A. Frommer, K. Kahl, S. MacLachlan, and L. Zikatanov. Adaptive reduction-based multigrid for nearly singular and highly disordered physical systems. Technical Report BUW-SC 2009/5, University of Wuppertal, Department of Mathematics, Germany, 2009.

[7] M. Creutz. *Quarks, Gluons and Lattices.* Cambridge University Press, 1983.

[8] N. Cundy, J. van den Eshof, A. Frommer, S. Krieg, Th. Lippert, and K. Schäfer. Numerical methods for the QCD overlap operator: III. nested iterations. *Comput. Phys. Commun.*, 165(3):221–242, 2005.

[9] E. de Sturler. Nested Krylov methods based on GCR. *J. Comput. Appl. Math.*, 67(1):15–41, 1996.

[10] E. de Sturler. Truncation strategies for optimal Krylov subspace methods. *SIAM J. Numerical Analysis*, 36(3):864–889, 1999.

[11] M. Eiermann, O. G. Ernst, and O. Schneider. Analysis of acceleration strategies for restarted minimal residual methods. *J. Comput. Appl. Math.*, 123:261–292, 2000.

[12] M. Eiermann, O. G. Ernst, and E. Ullmann. Computational aspects of the stochastic finite element method. *Comput. Visual. Sci.*, 10(1):3–15, 2007.

[13] S. C. Eisenstat, H. C. Elman, and M. H. Schultz. Variational iterative methods for nonsymmetric systems of linear equations. *SIAM J. Numerical Analysis*, 20(2):345–357, 1983.

[14] J. Erhel, K. Burrage, and B. Pohl. Restarted GMRES preconditioned by deflation. *J. Comput. Appl. Math.*, 69:303–318, 1996.

[15] S. Fischer, A. Frommer, U. Glässner, Th. Lippert, G. Ritzenhöfer, and K. Schilling. A parallel SSOR preconditioner for lattice QCD. *Comput. Phys. Commun.*, 98(1-2):20–34, 1999.

[16] A. Frommer and V. Simoncini. Matrix functions. In H. A. Wil and H. A. van der Vorst, editors, *Model Order Reduction: Theory, Research Aspects and Applications, Mathematics in Industry, volume 13*, pages 275–304. Springer-Verlag, Heidelberg, 2008.

[17] L. Giraud, S. Gratton, X. Pinel, and X. Vasseur. Flexible GMRES with deflated restarting. Technical Report TR/PA/09/111, CERFACS, 2009. In revision in SIAM J. Scientific Computing, extended and revised version of TR/PA/08/128.

[18] C. Jin, X.-C. Cai, and C. Li. Parallel domain decomposition methods for stochastic elliptic equations. *SIAM J. Scientific Computing*, 29(5):2096–2114, 2007.

[19] R. B. Morgan. A restarted GMRES method augmented with eigenvectors. *SIAM J. Matrix Analysis and Applications*, 16:1154–1171, 1995.

[20] R. B. Morgan. Implicitly restarted GMRES and Arnoldi methods for nonsymmetric systems of equations. *SIAM J. Matrix Analysis and Applications*, 21(4):1112–1135, 2000.

[21] R. B. Morgan. GMRES with deflated restarting. *SIAM J. Scientific Computing*, 24(1):20–37, 2002.

[22] R. B. Morgan and W. Wilcox. Deflated Iterative Methods for Linear Equations with Multiple Right-Hand Sides. *ArXiv Mathematical Physics e-prints*, May 2004.

[23] Y. Notay. Flexible conjugate gradients. *SIAM J. Scientific Computing*, 22(4):1444–1460, 2000.

[24] Y. Notay and P. S. Vassilevski. Recursive Krylov-based multigrid cycles. *Numerical Linear Algebra with Applications*, 15:473–487, 2008.

[25] C. C. Paige, B. N. Parlett, and H. A. van der Vorst. Approximate solutions and eigenvalue bounds from Krylov subspaces. *Numerical Linear Algebra with Applications*, 2:115–134, 1995.

[26] M. Parks, E. de Sturler, G. Mackey, D.D. Johnson, and S. Maiti. Recycling Krylov subspaces for sequences of linear systems. *SIAM J. Scientific Computing*, 28(5):1651–1674, 2006.

[27] Y. Saad. A flexible inner-outer preconditioned GMRES algorithm. *SIAM J. Scientific and Statistical Computing*, 14(2):461–469, 1993.

[28] Y. Saad. Analysis of augmented Krylov subspace methods. *SIAM J. Matrix Analysis and Applications*, 18:435–449, 1997.

[29] Y. Saad. *Iterative Methods for Sparse Linear Systems.* SIAM, Philadelphia, 2003. Second edition.

[30] Y. Saad and M. H. Schultz. GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems. *SIAM J. Scientific and Statistical Computing*, 7:856–869, 1986.

[31] V. Simoncini and D. B. Szyld. Flexible inner-outer Krylov subspace methods. *SIAM J. Numerical Analysis*, 40(6):2219–2239, 2003.

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

[33] G. L. G. Sleijpen and H. A. van der Vorst. A Jacobi–Davidson iteration method for linear eigenvalue problems. *SIAM J. Matrix Analysis and Applications*, 17(2):401–425, 1996.

[34] D. B. Szyld and J. A. Vogel. FQMR: A flexible quasi-minimal residual method with inexact preconditioning. *SIAM J. Scientific Computing*, 23(2):363–380, 2001.

[35] A. Toselli and O. Widlund. *Domain Decomposition methods - Algorithms and Theory.* Springer Series on

Computational Mathematics, vol. 34, Springer, 2004.

[36] E. Ullmann. *Solution Strategies for Stochastic Finite Element Discretizations.* PhD thesis, Technische Universität Bergakademie Freiberg, Germany, 2008.

[37] J. van den Eshof, A. Frommer, Th. Lippert, K. Schilling, and H. van der Vorst. Numerical methods for the QCD overlap operator: I. sign-function and error bounds. *Comput. Phys. Commun.*, 146(2):203–224, 2002.

[38] H. A. van der Vorst. Solution of $f(A)$ $x = b$ with projection methods for the matrix A. In A. Frommer, Th. Lippert, B. Medeke, and K. Schilling, editors, *Numerical Challenges in Lattice Quantum Chromodynamics, Lecture Notes in Computational Science and Engineering, volume 15*, pages 18–28. Springer-Verlag, Berlin, 2000.

[39] H. A. van der Vorst and C. Vuik. GMRESR: A family of nested GMRES methods. *Numerical Linear Algebra with Applications*, 1:369–386, 1994.