HAL
open science

Deflation and augmentation techniques in Krylov linear solvers

Olivier Coulaud, Luc Giraud, Pierre Ramet, Xavier Vasseur

To cite this version:

Olivier Coulaud, Luc Giraud, Pierre Ramet, Xavier Vasseur. Deflation and augmentation techniques in Krylov linear solvers. [Research Report] RR-8265, INRIA. 2013, 25 p. hal-00803225

HAL Id: hal-00803225

https://inria.hal.science/hal-00803225v1

Submitted on 21 Mar 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Deflation and augmentation techniques in Krylov subspace methods for the solution of linear systems

Olivier Coulaud, Luc Giraud, Pierre Ramet, Xavier Vasseur

RESEARCH REPORT
N° 8265
February 2013
Project-Teams HiePACS

ISSN 0249-6399 ISRN INRIA/RR--8265--FR+ENG

Deflation and augmentation techniques in Krylov subspace methods for the solution of linear systems

Olivier Coulaud*, Luc Giraud*, Pierre Ramet†, Xavier Vasseur‡

Project-Teams HiePACS

Research Report n° 8265 — February 2013 — 25 pages

Abstract: In this paper we present deflation and augmentation techniques that have been designed to accelerate the convergence of Krylov subspace methods for the solution of linear systems of equations. We review numerical approaches both for linear systems with a non-Hermitian coefficient matrix, mainly within the Arnoldi framework, and for Hermitian positive definite problems with the conjugate gradient method.

Key-words: Augmentation, Deflation, Krylov subspace methods, Linear systems of equations, Preconditioning.

* Inria Bordeaux-Sud Ouest, France
† Université de Bordeaux 1, France
‡ CERFACS, France

RESEARCH CENTRE
BORDEAUX – SUD-OUEST

200 avenue de la Vielle Tour
33405 Talence Cedex

Techniques de déflation et d'augmentation dans les solveurs linéaires de Krylov

Résumé : Dans ce rapport nous présentons des techniques de déflation et d'augmentation qui ont été développées pour accélérer la convergence des méthodes de Krylov pour la solution de systèmes d'équations linéaires. Nous passons en revue des approches pour des systèmes linéaires dont les matrices sont non-hermitiennes, principalement dans le contexte de la méthode d'Arnoldi, et pour des matrices hermitiennes définies positives avec la méthode du gradient conjugué.

Mots-clés : Augmentation, Déflation, Méthodes de Krylov, Systèmes linéaires d'équations, Préconditionnement.

Contents

1 Introduction **4**

2 Some background on Krylov subspace methods **4**
2.1 Notation 5
2.2 Basic properties of Krylov subspace methods 5
2.3 Minimum residual Krylov subspace method 5
2.4 Conjugate gradient method 6

3 Non-Hermitian matrices **8**
3.1 Augmented Krylov subspace methods 8
3.1.1 Augmentation with an arbitrary subspace 8
3.1.2 Augmentation with approximate invariant subspace 9
3.2 Deflated Krylov subspace methods 11
3.2.1 Deflation based on orthogonal projection 11
3.2.2 Deflation based on oblique projection 12
3.2.3 Deflation by preconditioning 13
3.3 Augmented and deflated Krylov subspace methods 14
3.3.1 Equivalence between deflated and augmented methods 14
3.3.2 Methods based on augmentation and deflation 15

4 Hermitian positive definite matrices **16**
4.1 Augmented conjugate gradient methods 16
4.2 Deflated Krylov subspace methods 18
4.3 Deflation via preconditioning 19

5 Linear systems with multiple right-hand sides given in sequence **19**

6 Conclusions and prospectives **20**

1 Introduction

The solution of linear systems of the form $Ax^\star = b$ plays a central role in many engineering and academic simulation codes. Among the most widely used solution techniques are the iterative schemes based on Krylov subspace methods [3, 23, 69, 80]. Their main advantages are their ability to solve linear systems even if the matrix of the linear system is not explicitly available and their capability to be "easily" parallelizable on large computing platforms. In order to speed up the convergence of these solution techniques, Krylov subspace methods are almost always used in combination with preconditioning. That is, instead of solving directly $Ax^\star = b$, the linear system is transformed into an equivalent one, e.g., $M_1Ax^\star = M_1b$, referred to as left preconditioned system, that is expected to be more amenable to a solution. The definition of an efficient preconditioner M_1, that should be an good approximation of A^{-1} in some sense, is very much problem dependent and is consequently an extremely active research field. We can also consider other equivalent linear systems $AM_2t^\star = b$ with $x^\star = M_2t^\star$ (right preconditioner) or $M_1AM_2t^\star = M_1b$ with $x^\star = M_2t^\star$ (split preconditioner). We refer the reader to [8] for a detailed overview on preconditioning.

There exist two complementary alternatives to speed up the convergence of the Krylov space, namely augmentation and deflation. Roughly speaking, in augmentation techniques, the search space in an enlarged Krylov space that is defined by a direct sum of two subspaces. This search space S_ℓ (of dimension ℓ) has the following form

$$S_\ell = \mathcal{K}_m(A, b) \oplus \mathcal{W} \tag{1}$$

where $\mathcal{K}_m(A, b)$ is a Krylov subspace of dimension m generated by the matrix A and the vector b and $\mathcal{W}$ (of dimension k) is called the augmentation space. A typical goal of augmentation is to add information about the problem into the global search space S_ℓ that is only slowly revealed in the Krylov subspace itself.

Alternatively, deflation is based on the use of a projection operator P to decompose $x^\star$ as $x^\star = Px^\star + (I - P)x^\star$. The general idea is to select P such that the solution of $PAx^\star = Pb$, referred to as the deflated linear system, is more easily amenable to a solution by a Krylov subspace method than the original linear system $Ax^\star = b$. The component $(I - P)x^\star$ can then be computed by solving a linear system of small dimension.

The purpose of this paper is to expose these two latter acceleration techniques that become increasingly popular. We refer the reader to [40] for a recent excellent analysis of these methods together with detailed references and historical comments. Here augmentation and deflation are described in a framework where variable preconditioning can be used as it is nowadays customary when considering large scale linear systems [61, 62, 74]. This paper is organized as follows. In Section 2 we introduce some background on Krylov subspace methods with emphasis on the minimum residual norm approach for systems with a non-Hermitian coefficient matrix and the conjugate gradient method for the solution of Hermitian positive definite problems. In Section 3 we describe the augmentation and deflation techniques and their possible combination in the case of systems with non-Hermitian matrices with references to concrete applications. Similar exposure is performed in Section 4 for Hermitian positive definite linear systems. Finally some concluding remarks and prospectives are drawn in Section 6.

2 Some background on Krylov subspace methods

We briefly describe the basic properties of Krylov subspace methods for the solution of a linear system of equations of the form

$$Ax^\star = b \tag{2}$$

where the nonsingular $n \times n$ coefficient matrix $A \in \mathbb{C}^{n\times n}$ is supposed to be either non-Hermitian or Hermitian positive definite and b a given vector in $\mathbb{C}^n$. First we introduce the notation used throughout this paper.

2.1 Notation

We denote the range of a matrix A by Range(A) and its nullspace as $Ker(A)$. We denote by $\|.\|$ 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. ^{T} denotes the transpose operation, while ^{H} denotes the Hermitian transpose operation. Given a vector $d \in \mathbb{C}^k$ with components d_i, $D = \text{diag}(d_1, \ldots, d_k)$ is the diagonal matrix $D \in \mathbb{C}^{k\times k}$ such that $D_{ii} = d_i$. Given $Z_m = [z_1, \cdots, z_m] \in \mathbb{C}^{n\times m}$ we denote its i-th column as $z_i \in \mathbb{C}^n, (1 \leq i \leq m)$. The vector $e_m \in \mathbb{R}^m$ denotes the m-th canonical basis vector of $\mathbb{R}^m$. We denote by $\kappa(A)$ the Euclidean condition number of A that is defined by $\kappa(A) = \dfrac{\sigma_{\max}}{\sigma_{\min}}$ where $\sigma_{\max}$ ($\sigma_{\min}$) is the largest (respectively smallest) singular value of A. For Hermitian positive definite matrices, the condition number reduces to $\kappa(A) = \dfrac{\lambda_{\max}}{\lambda_{\min}}$ where $\lambda_{\max}$ ($\lambda_{\min}$) is the largest (respectively smallest) eigenvalue of A. Finally, throughout the paper for the sake of readability the integer subscript ℓ denotes the dimension of the search space.

2.2 Basic properties of Krylov subspace methods

The Krylov subspace methods seek for the solution of Equation (2) in a sequence of embedded spaces of increasing dimension $\mathcal{K}_\ell(A, b) = span(b, Ab, \cdots, A^{\ell-1}b)$. This is motivated [44] by the fact that for ℓ large enough these spaces contain the solution of the linear system (2). If we denote by $m_A(x)$ the minimal polynomial associated with A, the Jordan decomposition of this polynomial writes $m_A(t) = \prod_{i=1}^{s}(t - \lambda_i)^{m_i}$ where $(\lambda_1, \cdots, \lambda_s)$ are the distinct eigenvalues of A and $(m_1, \cdots, m_s)$ their indices in the Jordan form. In a canonical form we also have $m_A(t) = \sum_{i=0}^{m} \alpha_i t^i$ with $m = \sum_{i=1}^{s} m_i$ and $\alpha_0 = \prod_{i=1}^{s}(-\lambda_i)^{m_i} \neq 0$ since A is nonsingular. Consequently, $A^{-1} = -\alpha_0^{-1} \sum_{i=0}^{m-1} \alpha_{i+1} A^i$ that portrays $x^\star = A^{-1}b$ as a vector of the Krylov space $\mathcal{K}_{m-1}(A, b)$. This indicates that, in exact arithmetic, Krylov methods must converge in at most $m-1$ steps or less if the right-hand side does not have components in all the eigendirections. This observation also gives some ideas on ways to speed-up the convergence of these methods. As mentioned earlier preconditioning is a widely used approach that consists in transforming (2) in an equivalent nonsingular system where the preconditioned matrix has less [58] or better clustered eigenvalues (see [8] and the references therein).

The rest of the paper is dedicated to an overview of proposed techniques for augmentation and deflation both for non-Hermitian and Hermitian positive definite problems.

2.3 Minimum residual Krylov subspace method

In this section we focus on minimum residual norm subspace methods for the solution of linear systems with a non-Hermitian coefficient matrix. We refer the reader to [69, 80] for a general introduction to Krylov subspace methods and to [74] for a recent overview on Krylov subspace methods; see also [20, 21] for an advanced analysis related to minimum residual norm Krylov subspace methods.

Augmented and deflated minimum residual norm Krylov subspace methods are usually characterized by a generalized Arnoldi relation introduced next.

Definition 1 *Generalized Arnoldi relation. The minimum residual norm subspace methods investigated in this paper satisfies the following relation:*

$$AZ_\ell = V_{\ell+1}\bar{H}_\ell \tag{3}$$

where $Z_\ell \in \mathbb{C}^{n\times\ell}$, $V_{\ell+1} \in \mathbb{C}^{n\times(\ell+1)}$ *such that* $V_{\ell+1}^H V_{\ell+1} = I_{\ell+1}$ *and* $\bar{H}_\ell \in \mathbb{C}^{(\ell+1)\times\ell}$. *These methods compute an approximation of the solution of (2) in a* ℓ*-dimensional affine space* $x_0 + Z_\ell\, y_\ell$ *where* $y_\ell \in \mathbb{C}^\ell$. *In certain cases,* $\bar{H}_\ell$ *is an upper Hessenberg matrix.*

We next introduce a minimum residual norm subspace method proposed by Saad [67] since it is the basis for further developments related to augmented and deflated Krylov subspace methods of minimum residual norm type. This method named Flexible GMRES (FGMRES) was primarily introduced to allow variable preconditioning. We denote by M_j the nonsingular matrix that represents the preconditioner at step j of the method. Algorithm 1 depicts the FGMRES(ℓ) method where the dimension of the approximation subspace is not allowed to be larger than a prescribed dimension noted ℓ. Starting from an initial guess $x_0 \in \mathbb{C}^n$, it is based on a generalized Arnoldi relation

$$AZ_\ell = V_{\ell+1}\bar{H}_\ell \quad \text{with} \quad V_{\ell+1}^H V_{\ell+1} = I_{\ell+1}, \tag{4}$$

where $Z_\ell \in \mathbb{C}^{n\times\ell}$, $V_{\ell+1} \in \mathbb{C}^{n\times(\ell+1)}$ and the upper Hessenberg matrix $\bar{H}_\ell \in \mathbb{C}^{(\ell+1)\times\ell}$ are obtained from the Arnoldi procedure described in Algorithm 2. An approximate solution $x_\ell \in \mathbb{C}^n$ is then found by minimizing the residual norm $\|b - A(x_0 + Z_\ell y)\|$ over the space $x_0 + \text{range}(Z_\ell)$, the corresponding residual being $r_\ell = b - Ax_\ell \in \mathbb{C}^n$ with $r_\ell \in \text{range}(V_{\ell+1})$. With notation of Algorithm 1 the current approximation x_ℓ can be written as

$$x_\ell = x_0 + Z_\ell y^\star, \tag{5}$$

whereas the residual $r_\ell = b - Ax_\ell$ satisfies the Petrov-Galerkin orthogonality condition

$$r_\ell \perp A\ \text{range}(Z_\ell).$$

Hence, an optimality property similar to the one that defines GMRES is thus obtained [69]. We note however that no general convergence results are available since the subspace of approximants Range(Z_ℓ) is no longer a standard Krylov subspace. We refer the reader to [67, 69] for the analysis of the breakdown in FGMRES. Furthermore, as it can be seen in Equation (5), the update of the iterate x_ℓ requires to store the complete set of vectors Z_ℓ inducing a large memory footprint for large ℓ. In order to alleviate this memory requirement, a restarting strategy must be implemented as shown in Algorithm 1. The construction of a complete set of Z_ℓ is often name a cycle of the method and corresponds to one iteration of the loop in Algorithm 1.

When the preconditioner is constant, FGMRES(ℓ) reduces to right-preconditioned GMRES(ℓ) whose convergence properties are discussed in [69, Chapter 6].

2.4 Conjugate gradient method

The conjugate gradient [41] is the method of choice for Hermitian Positive Definite (HPD) linear systems. In a shortcut, it relies on an Arnoldi like relation (namely a Lanczos relation [51]) similar to Equation (4) with $Z_\ell = V_\ell$ (case of no preconditioning) and a Ritz-Galerkin condition $r_\ell = b - Ax_\ell \perp \mathcal{K}_\ell(A, r_0)$. At each iteration $x_\ell = x_0 + V_\ell y^\star$ is computed via the solution of the small linear system $H_\ell y^\star = \|r_0\|(1, 0_{1\times(\ell-1)})^T$, where $H_\ell = V_\ell^H A V_\ell$ is the square leading part of $\bar{H}_\ell$. Since A is Hermitian, H_ℓ is also Hermitian. Furthermore, its structure is upper

Algorithm 1 Flexible GMRES(ℓ)

1: *Initialization:* Choose $\ell > 0$, $tol > 0$, $x_0 \in \mathbb{C}^n$. Let $r_0 = b - Ax_0$, $\beta = \|r_0\|$, $c = [\beta, 0_{1\times\ell}]^T$ where $c \in \mathbb{C}^{\ell+1}$, $v_1 = r_0/\beta$.
Loop

2: *Computation of $V_{\ell+1}$, Z_ℓ and $\bar{H}_\ell$ (see Algorithm 2):* Apply ℓ steps of the Arnoldi method with variable preconditioning ($z_j = M_j^{-1}v_j, 1 \leq j \leq \ell$) to obtain $V_{\ell+1} \in \mathbb{C}^{n\times(\ell+1)}$, $Z_\ell \in \mathbb{C}^{n\times\ell}$ and the upper Hessenberg matrix $\bar{H}_\ell \in \mathbb{C}^{(\ell+1)\times\ell}$ such that

$$AZ_\ell = V_{\ell+1}\bar{H}_\ell \quad \text{with} \quad V_{\ell+1}^H V_{\ell+1} = I_{\ell+1}.$$

3: *Minimum norm solution:* Compute the minimum norm solution $x_\ell \in \mathbb{C}^n$ in the affine space $x_0 + \text{range}(Z_\ell)$; that is, $x_\ell = x_0 + Z_\ell y^\star$ where $y^\star = \underset{y\in\mathbb{C}^\ell}{\text{argmin}}\, \|c - \bar{H}_\ell y\|$.

4: *Check the convergence criterion:* If $\|c - \bar{H}_\ell y^\star\|/\|b\| \leq tol$, exit

5: *Restarting:* Set $x_0 = x_\ell$, $r_0 = b - Ax_0$, $\beta = \|r_0\|$, $c = [\beta, 0_{1\times\ell}]^T$, $v_1 = r_0/\beta$.

End of loop

Algorithm 2 Arnoldi procedure: computation of $V_{\ell+1}$, Z_ℓ and $\bar{H}_\ell$

1: **for** $j = 1, \ell$ **do**
2: $\quad z_j = M_j^{-1}v_j$
3: $\quad s = Az_j$
4: $\quad$ **for** $i = 1, j$ **do**
5: $\quad\quad h_{i,j} = v_i^H s$
6: $\quad\quad s = s - h_{i,j}v_i$
7: $\quad$ **end for**
8: $\quad h_{i+1,j} = \|s\|$, $v_{j+1} = s/h_{i+1,j}$
9: **end for**
10: Define $Z_\ell = [z_1, \cdots, z_\ell]$, $V_{\ell+1} = [v_1, \cdots, v_{\ell+1}]$, $\bar{H}_\ell = \{h_{i,j}\}_{1\leq i\leq\ell+1, 1\leq j\leq\ell}$

Hessenberg that, combined with the Hermitian property, implies that H_ℓ is tridiagonal HPD. The first consequence of this structure of H_ℓ is that the orthogonalization of V_ℓ can be performed cheaply with a three term recurrence. The second consequence is that a LU factorization of H_ℓ can be incrementally computed and this factorization without pivoting is known to be stable for positive definite matrices. The conjugate gradient method is a very elegant, sophisticated and powerful algorithm that exploits nicely all the above mentioned properties. It can be implemented through short recurrences that do not require to store the complete set of vector V_ℓ leading to a very low memory consumption. Furthermore, the conjugate gradient enjoys a unique minimum norm property on the forward error that reads $x_\ell = \text{argmin}_{x\in x_0+\mathcal{K}_\ell(A,r_0)} \|x - x^\star\|_A$ where $x^\star$ denotes the exact solution and $\|\cdot\|_A$ is the norm associated with A. In addition, it exists an upper bound on its convergence rate that reads ($\ell \geq 1$)

$$\|x_\ell - x^\star\|_A \leq 2 \cdot \left(\frac{\sqrt{\kappa(A)} - 1}{\sqrt{\kappa(A)} + 1}\right)^\ell \|x_0 - x^\star\|_A. \tag{6}$$

We refer to [52, 69, 80] for an exhaustive and detailed exposure of CG and to [39] for a nice description of its history.

3 Non-Hermitian matrices

In this section we detail augmentation and deflation techniques in Krylov subspace methods when the coefficient matrix A is non-Hermitian. We specifically focus on minimum residual norm subspace methods and assume that a generalized Arnoldi relation (3) holds. We denote by x_0, $r_0 = b - Ax_0$ the initial guess and residual vector respectively, and by $V_{\ell+1}$, $\bar{H}_\ell$ and Z_ℓ the matrices involved in this relation. With notation of Algorithm 1 r_0 can be expressed as $r_0 = V_{\ell+1}(c - \bar{H}_\ell y^\star)$.

3.1 Augmented Krylov subspace methods

We next discuss two possibilities to select the augmentation space and analyze the corresponding Krylov subspace methods.

3.1.1 Augmentation with an arbitrary subspace

Given a basis $W = [w_1, \cdots, w_k]$ of an augmentation subspace $\mathcal{W}$, a slight modification in the Arnoldi procedure (Algorithm 2) is used to deduce an orthogonal basis of S_ℓ defined in (1) (see [14]). It consists of defining z_j (line 2 of Algorithm 2) now as

$$z_j = M_j^{-1} v_j \quad (1 \le j \le m) \text{ and } \; z_j = M_j^{-1} w_{j-m} \quad (m < j \le m+k).$$

With this definition we finally obtain the generalized Arnoldi relation

$$AZ_{m+k} = V_{m+k+1} \bar{H}_{m+k}$$

where

$$Z_{m+k} = [M_1^{-1} v_1, M_2^{-1} v_2, \cdots, M_{m+1}^{-1} w_1, M_{m+2}^{-1} w_2, \cdots, M_{m+k+1}^{-1} w_k], \tag{7}$$

$$V_{m+k+1} = [v_1, v_2, \cdots, v_{m+k+1}], \tag{8}$$

and $\bar{H}_{m+k}$ is a $(m+k+1) \times (m+k)$ upper Hessenberg matrix. Thus the residual minimization property is then deduced similarly as in FGMRES [67]. Hence, the approximate solution from the affine space $x_0 + \text{range}(Z_{m+k})$ can be written as

$$x_{m+k} = x_0 + Z_{m+k} y^\star$$

with $y^\star \in \mathbb{C}^{(m+k)}$ solution of the residual norm minimization problem

$$y^\star = \underset{y \in \mathbb{C}^{(m+k)}}{\text{argmin}} \; \| \|r_0\| e_1 - \bar{H}_{m+k} y \|$$

(with e_1 designing here the first canonical vector of $\mathbb{R}^{(m+k+1)}$). In case of constant right preconditioning the main important property is that if any vector w_j is the solution of $AM^{-1} w_j = v_i, 1 \le i \le m$, then in general the exact solution of the original system (2) can be extracted from S_ℓ; see, e.g., [68, Proposition 2.1]. We refer the reader to [14] for a discussion of possible choices for the augmented subspace $\mathcal{W}$. Vectors obtained with either different iterative methods or with different preconditioners can be incorporated in Z_m quite easily. In block Krylov subspace methods we also mention that $\mathcal{W}$ consists of the sum of a few other Krylov subspaces generated with the same matrix but with different right-hand sides; see [14] for a discussion and numerical experiments on academic problems. A popular idea is to choose $\mathcal{W}$ as an approximate invariant subspace associated with a specific part of the spectrum of A or AM^{-1} in case of fixed preconditioning. This is discussed next.

3.1.2 Augmentation with approximate invariant subspace

A typical goal of augmentation is to add information about the problem into the search space that is only slowly revealed in the Krylov subspace itself. It is often known that eigenvalues of the (preconditioned) operator close to zero tend to slow down the convergence rate of the Krylov subspace methods [14]. Hence, augmentation based on approximate invariant subspaces made of eigenvectors corresponding to small in modulus eigenvalues of the (preconditioned) operator has been proposed; see, e.g.,[54, 55, 56, 68] and references therein.

Harmonic Ritz information In [56] Morgan has suggested to select $\mathcal{W}$ as an approximate invariant subspace and to update this subspace at the end of each cycle. Approximate spectral information is then required to define the augmentation space. This is usually obtained by computing harmonic Ritz pairs of A with respect to a certain subspace [14, 56]. We present here a definition of a harmonic Ritz pair as given in [63, 75].

Definition 2 *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 the vector y a harmonic Ritz vector associated with the harmonic Ritz value θ.

Based on the generalized Arnoldi relation (3), the augmentation procedure proposed in [37, Proposition 1] relies on the use of k harmonic Ritz vectors $Y_k = V_\ell P_k$ of $AZ_\ell V_\ell^H$ with respect to Range(V_ℓ), where $Y_k \in \mathbb{C}^{n\times k}$ and $P_k = [p_1, \cdots, p_k] \in \mathbb{C}^{\ell\times k}$. According to Definition 2, the harmonic Ritz vector $y_j = V_\ell p_j$ then satisfies

$$Z_\ell^H A^H\,(AZ_\ell\; p_j - \theta_j V_\ell\; p_j) \quad = \quad 0. \tag{9}$$

Using the generalized Arnoldi relation (3) we finally obtain the relation

$$\bar{H}_\ell^H\;\bar{H}_\ell\; y_j = \theta \bar{H}_\ell^H V_{\ell+1}^H V_\ell\; y_j. \tag{10}$$

Since

$$\bar{H}_\ell = \begin{bmatrix} H_\ell \\ h_{\ell+1,\ell} e_\ell^T \end{bmatrix}, \quad H_\ell \in \mathbb{C}^{\ell\times\ell}$$

where $H_\ell \in \mathbb{C}^{\ell\times\ell}$ is supposed to be nonsingular, the generalized eigenvalue problem is then equivalent to

$$(H_\ell + h_{\ell+1,\ell}^2 H_\ell^{-H} e_\ell e_\ell^T) y_j = \theta_j y_j. \tag{11}$$

This corresponds to a standard eigenvalue problem of dimension ℓ only, where ℓ is supposed to be much smaller than the problem dimension n. In consequence, the approximate spectral information based on Harmonic Ritz pair is quite inexpensive to compute.

GMRES augmented with approximate invariant subspace The augmentation space $\mathcal{W}$ based on approximate invariant information corresponding to Range(Y_k) is then used. The key point detailed next is to understand how to incorporate this information in a minimum residual norm subspace method such as GMRES. To do so, we recall a useful relation satisfied by the harmonic Ritz vectors $P_k \in \mathbb{C}^{\ell\times k}$ shown in [37, Lemma 3.1]

$$AZ_\ell P_k = V_{\ell+1}\left[\begin{bmatrix} P_k \\ 0_{1\times k}\end{bmatrix}, c-\bar{H}_\ell y^\star\right]\begin{bmatrix} \mathrm{diag}(\theta_1,\ldots,\theta_k) \\ \alpha_{1\times k}\end{bmatrix}, \tag{12}$$

$$AZ_\ell P_k = [V_\ell P_k, r_0]\begin{bmatrix} \mathrm{diag}(\theta_1,\ldots,\theta_k) \\ \alpha_{1\times k}\end{bmatrix}, \tag{13}$$

where $r_0 = V_{\ell+1}(c-\bar{H}_\ell y^\star)$ and $\alpha_{1\times k} = [\alpha_1,\ldots,\alpha_k] \in \mathbb{C}^{1\times k}$. Next, the QR factorization of the $(\ell+1)\times(k+1)$ matrix appearing on the right-hand side of relation (12) is performed as

$$\left[\begin{bmatrix} P_k \\ 0_{1\times k}\end{bmatrix}, c-\bar{H}_\ell y^\star\right] = QR \tag{14}$$

where $Q \in \mathbb{C}^{(\ell+1)\times(k+1)}$ has orthonormal columns and $R \in \mathbb{C}^{(k+1)\times(k+1)}$ is upper triangular, respectively. Then it can be shown that the relations

$$A\ Z_k = V_{k+1}\bar{H}_k, \tag{15}$$

$$V_{k+1}^H V_{k+1} = I_{k+1}, \tag{16}$$

$$\mathrm{Range}([Y_k, r_0]) = \mathrm{Range}(V_{k+1}), \tag{17}$$

hold with *new* matrices $Z_k, V_k \in \mathbb{C}^{n\times k}$ and $\bar{H}_k \in \mathbb{C}^{(k+1)\times k}$ defined as

$$Z_k = Z_\ell\ Q_{\ell\times k}, \tag{18}$$

$$V_{k+1} = V_{\ell+1}\ Q, \tag{19}$$

$$\bar{H}_k = Q^H\ \bar{H}_\ell\ Q_{\ell\times k}, \tag{20}$$

where $V_{\ell+1}$, Z_ℓ and $\bar{H}_\ell$ refer to matrices obtained at the end of the previous cycle; see [37, Proposition 2]. With the augmentation subspace $\mathcal{W} = \mathrm{Range}(Y_k)$, m Arnoldi steps with variable preconditioners and starting vector v_{k+1} are then carried out while maintaining orthogonality to V_k leading to

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

We note that $\bar{H}_m \in \mathbb{C}^{(m+1)\times m}$ is upper Hessenberg. At the end of the new cycle this gives the generalized Arnoldi relation

$$A\ [Z_k, z_{k+1},\cdots,z_{m+k}] = [V_{m+k+1}]\left[\begin{bmatrix}\bar{H}_k \\ 0_{m\times k}\end{bmatrix}\begin{bmatrix} B_{k\times m} \\ \bar{H}_m\end{bmatrix}\right]$$

i.e.

$$AZ_{m+k} = V_{m+k+1}\bar{H}_{m+k},$$

where $V_{m+k+1} \in \mathbb{C}^{n\times(m+k+1)}$, $\bar{H}_{m+k} \in \mathbb{C}^{(m+k+1)\times(m+k)}$ and $B_{k\times m} \in \mathbb{C}^{k\times m}$ results from the orthogonalization of $[Az_{k+1},\cdots,Az_{m+k+1}]$ against V_{k+1}. We note that $\bar{H}_{m+k}$ is no more upper Hessenberg due to the leading dense $(k+1)\times k$ submatrix $\bar{H}_k$. It is important to notice that the augmentation space varies at each restart since it is built from the search space available at the end of each previous cycle. The resulting algorithm can be viewed as an adaptive augmented Krylov subspace method. We refer the reader to [37, Sections 2 and 3] for the complete derivation of the method and additional comments on its computational cost.

Remarks and applications When the preconditioner is fixed, the previous algorithm proposed by Morgan [56] is known as GMRES with deflated restarting (GMRES-DR). Although the term "deflated" is used, we note that this algorithm does correspond to a GMRES method with an adaptive augmented basis without any explicit deflated matrix. The success of GMRES-DR has been demonstrated on many academic examples [54] and concrete applications such as in lattice QCD [15, 29], reservoir modeling [2, 49] or electromagnetism [37]. We refer the reader to [56, 66] for further comments on the algorithm and computational details. We note that GMRES with deflated restarting is equivalent to other augmented GMRES methods such as GMRES with eigenvectors [54] and implicitly restarted GMRES [55]. Most often the approximate invariant subspace is chosen as the Harmonic Ritz pair corresponding to the smallest in modulus Harmonic Ritz values. Depending on the problem we note however that other specific part of the spectrum of the preconditioned operator can be targeted; see, e.g., [37, Section 4.2] for an application related to a wave propagation problem.

3.2 Deflated Krylov subspace methods

We next briefly describe minimal residual Krylov subspace methods based on deflation. We refer the reader to [30, 31, 40] for a recent excellent overview of deflated Krylov subspace methods in the Hermitian and non-Hermitian cases, where extensive bibliographical references and historical comments can be found. The general idea of deflation is to split the approximation space into two complementary subspaces such that the projected linear system, referred to as the deflated linear system, will be easier to solve iteratively than the original linear system (2). The fact that these subspaces can be chosen in different ways explains the huge literature on deflated Krylov subspace methods. The Krylov subspace method is then confined in one of this subspace, by projecting the initial residual into this space and by replacing A by its restriction to this space. If the projection operator is chosen properly the deflated linear system will be easier to solve iteratively than the original linear system (2). This property will be notably shown for Hermitian positive definite systems in Section 4.2 and can be extended to non-Hermitian situations with additional assumptions on A (see, e.g., [26, Section 2]). We first present a possible strategy based on orthogonal projection and then briefly discuss an extension based on oblique projection proposed in [40].

3.2.1 Deflation based on orthogonal projection

We still denote by $\mathcal{W}$ a subspace of $\mathbb{C}^n$ of dimension k, where k is assumed to be much smaller than the problem dimension n. We later denote by $W \in \mathbb{C}^{n\times k}$ a matrix whose columns form a basis of $\mathcal{W}$ so that $W^H A^H AW$ is HPD (hence invertible). To simplify further developments, we introduce the matrices $Q_1, P_1, P_2 \in \mathbb{C}^{n\times n}$ defined respectively as

$$Q_1 = AW(W^H A^H AW)^{-1} W^H A^H, \tag{21}$$

$$P_1 = I_n - Q_1, \tag{22}$$

$$P_2 = I_n - W(W^H A^H AW)^{-1} W^H A^H A. \tag{23}$$

We can easily show that P_1 and P_2 are orthogonal projectors such that P_1 projects onto $(A\mathcal{W})^\perp$ along $(A\mathcal{W})$, whereas P_2 projects onto $\mathcal{W}^\perp$ along $\mathcal{W}$. Furthermore we note that P_1 is Hermitian and that $AP_2 = P_1 A$. The decomposition based on orthogonal projection reads as

$$\mathbb{C}^n = \mathcal{W} \oplus \mathcal{W}^\perp.$$

Hence, the solution $x^\star$ of the original system (2) can be written as

$$x^\star = (I - P_2)x^\star + P_2 x^\star = W(W^H A^H AW)^{-1} W^H A^H b + P_2 x^\star.$$

With this decomposition the original system (2) then simply becomes

$$P_1 A x = P_1 b. \tag{24}$$

Although the deflated matrix $P_1 A$ is singular, the deflated linear system (24) is consistent so that it can be solved by an appropriate Krylov subspace method. Here we focus on the application of minimum residual Krylov subspace method based on GMRES to solve the deflated linear system (24). Hence, the search space of the Krylov subspace method applied to (24) can be written as

$$\hat{S}_m = \mathcal{K}_m(P_1 A, P_1 r_0),$$

while the current approximation $\hat{x}_m$ and the current residual $\hat{r}_m = P_1(b - A\hat{x}_m)$ at the end of the cycle satisfies the relations

$$\begin{array}{rcl} \hat{x}_m & \in & \hat{x}_0 + \hat{S}_m, \\ P_1(b - A\hat{x}_m) & \perp & P_1 A\, \mathcal{K}_m(P_1 A, P_1 r_0). \end{array}$$

Since $P_1 A W = 0_{n\times k}$, $P_1 A$ is singular. Hence it is of paramount importance to analyze the possibilities of a breakdown when solving the deflated linear system (24). In our context, when GMRES is used to solve the deflated linear system, this feature has been notably analyzed in [40, Section 3] based on theoretical results obtained by Brown and Walker [9]. We refer the reader to [40, Corollary 3] for conditions that characterize the possibility of breakdowns. It is worthwhile to note that a breakdown cannot occur if the condition

$$Ker(P_1 A) \cap \mathrm{Range}(P_1 A) = \{0\}$$

holds; see [30, Theorem 4.1]. This condition is notably satisfied if $\mathcal{W}$ is chosen as an exact A-invariant subspace, i.e., when $A\mathcal{W} = \mathcal{W}$ since $Ker(P_1 A) = \mathcal{W}$ and $Im(P_1 A) = \mathcal{W}^\perp$ due to the nonsingularity of A. Once the solution of the deflated linear system is obtained, we deduce the approximation x_m of the original system as

$$x_m = W(W^H A^H A W)^{-1} W^H A^H b + P_2 \hat{x}_m,$$

and by construction we note that

$$b - A x_m = P_1(b - A\hat{x}_m),$$

i.e.,

$$r_m = \hat{r}_m.$$

We refer to [26] for applications of deflated Krylov subspace methods with orthogonal projection to linear systems with non-Hermitian matrices. As an illustration, a typical choice of subspaces is to choose the columns of W as right eigenvectors of A corresponding to eigenvalues of small absolute value.

3.2.2 Deflation based on oblique projection

We briefly mention a strategy based on oblique projection that is considered as more appropriate for the solution of non-Hermitian linear systems since the eigenspaces of A are in general not mutually orthogonal [40]. As in Section 3.2.1, the search space S_ℓ will be decomposed into a direct sum of two subspaces. More precisely, the following decompositions into nonorthogonal complements are used

$$\mathbb{C}^n = A\mathcal{W} \oplus \tilde{\mathcal{W}}^\perp = A\tilde{\mathcal{W}} \oplus \mathcal{W}^\perp,$$

where $\mathcal{W}$ and $\tilde{\mathcal{W}}$ represent two subspaces of $\mathbb{C}^n$ of dimension k respectively. As before, we denote by $W \in \mathbb{C}^{n\times k}$ ($\tilde{W} \in \mathbb{C}^{n\times k}$) a matrix whose columns form a basis of $\mathcal{W}$ ($\tilde{\mathcal{W}}$ respectively). We assume that both matrices are chosen such that $\tilde{W}^H AW$ is nonsingular. The key idea is then to introduce the matrices $Q_2, P_3 \in \mathbb{C}^{n\times n}$ defined as

$$Q_2 = W(\tilde{W}^H AW)^{-1}\tilde{W}^H, \tag{25}$$

$$P_3 = I_n - W(\tilde{W}^H AW)^{-1}\tilde{W}^H. \tag{26}$$

It is easy to show that Q_2 and $P_3 = I_n - Q_2$ are projection operators; Q_2 is the oblique projection onto $(A\mathcal{W})$ along $\tilde{\mathcal{W}}^\perp$, while P_3 is the oblique projection onto $\tilde{\mathcal{W}}^\perp$ along $(A\mathcal{W})$. Given these oblique projection operators, the deflated linear system is now defined as

$$P_3AP_3x = P_3b$$

with $\hat{r}_0 = P_3r_0$. The use of a Krylov space solver is then now restricted to $\tilde{\mathcal{W}}^\perp$. Hence, it can be shown that the deflated Krylov subspace method based on GMRES yields iterates x_m at the end of the cycle of the form

$$x_m \in x_0 + \mathcal{K}_m(P_3AP_3, P_3r_0) + \mathcal{W}.$$

This also implies the following relation for the residual [40]

$$b - Ax_m \in r_0 + A\mathcal{K}_m(P_3AP_3, P_3r_0) + A\mathcal{W}.$$

We refer the reader to [40, Sections 5 and 6] for the mathematical aspects of deflated Krylov subspace methods based on oblique projection and to [40, Section 11] for an overview of partly related methods that only differ in the choice of the projection operators. A typical choice is to choose the columns of W as right eigenvectors of A and the columns of $\tilde{W}$ as the corresponding left eigenvectors. We refer to [26] for an application of deflated Krylov subspace methods with oblique projection in the general non-Hermitian case.

3.2.3 Deflation by preconditioning

Finally, we note that deflation based on spectral approximate information can be used to construct nonsingular preconditioners that move small in modulus eigenvalues away from zero. Both Kharchenko and Yeremin [46] and Erhel et al. [24] have proposed GMRES algorithms with augmented basis and a nonsingular right preconditioner that move the small eigenvalues to a (multiple) large eigenvalue. Baglama et al. [5] have proposed a left preconditioned GMRES method with similar effect. In [46] the main idea is to translate a group of small eigenvalues of A via low-rank projections of the form

$$\tilde{A} = A(I_n + u_1w_1^H)\cdots(I_n + u_kw_k^H),$$

where u_j and w_j are the right and left eigenvectors associated with the eigenvalues to be translated respectively. The restarted Krylov subspace method is now applied to the coefficient matrix $\tilde{A}$ leading to an adaptive update of the preconditioner (performed at the end of each cycle). We note that A can correspond to an already preconditioned operator, in such a case this strategy leads to a two-level preconditioning strategy that is found to be effective on real-life applications provided that the spectral information is computed accurately [12]. We also mention the extension of this two-level preconditioning strategy to the case of sequences of linear systems (see, e.g., [36] where additional theoretical results and numerical experiments can be found).

3.3 Augmented and deflated Krylov subspace methods

In the previous sections, we have described how either augmentation or deflation can be incorporated into Krylov subspace methods of minimum residual norm type. We note that it is possible to combine simultaneously deflation and augmentation in a single Krylov subspace method. In such a setting, the search space of the Krylov subspace method is then decomposed as

$$S_\ell = \mathcal{W} + \mathcal{K}_m(\hat{A}, \hat{r}_0)$$

where $\mathcal{W}$ is the augmentation space of dimension k, $\hat{A}$ refers to the deflated operator and $\hat{r}_0$ to the deflated residual. As an illustration, we review the GCRO (Generalized Conjugate Residual with Orthogonalization) method due to de Sturler [16].

3.3.1 Equivalence between deflated and augmented methods

In this section, we describe a general setting that helps us to understand the link between deflated and augmented minimal residual norm Krylov subspace methods. It has been first presented in [40] and we generalize this setting to the case of flexible methods. As discussed in Section 3.1, the search space in augmented methods is of the form

$$S_\ell = \mathcal{W} \oplus \mathcal{K}_m(\hat{A}, \hat{r}_0)$$

where $\mathcal{W}$ is an augmentation subspace of dimension k. The approximation x_m at the end of a given cycle can be written as

$$x_m = x_0 + Z_m y_m + W w_m$$

where $y_m \in \mathbb{C}^m$ and $w_m \in \mathbb{C}^k$. In the augmented Krylov subspace methods that we have considered, the residual r_m satisfies a Petrov-Galerkin condition, i.e., $r_m \perp AS$ which leads to the two orthogonality conditions

$$r_m \perp A\mathcal{W} \quad \text{and} \quad r_m \perp A\mathcal{K}_m(\hat{A}, \hat{r}_0).$$

The first orthogonality condition $r_m \perp A\mathcal{W}$ leads to the relation

$$(W^H A^H A W) w_m = W^H A^H (r_0 - A Z_m y_m).$$

To simplify notation we introduce the matrix $Q_3 \in \mathbb{C}^{n\times n}$ such that

$$Q_3 = W(W^H A^H A W)^{-1} W^H.$$

We then deduce the following relations for the current approximation x_m

$$x_m = (I_n - Q_3 A^H A)(x_0 + Z_m y_m) + Q_3 A^H b, \tag{27}$$

and for the current residual r_m

$$r_m = (I_n - A Q_3 A^H)(r_0 - A Z_m y_m). \tag{28}$$

We then introduce the two matrices

$$\begin{aligned} P_4 &= I_n - Q_3 A^H A, \\ P_5 &= I_n - A Q_3 A^H \end{aligned}$$

where $P_4 \in \mathbb{C}^{n\times n}$ and $P_5 \in \mathbb{C}^{n\times n}$. It is easy to show that both P_4 and P_5 are orthogonal projectors and that $AP_4 = P_5A$. If we define $\tilde{x}_m \in \mathbb{C}^n$ as $\tilde{x}_m = x_0 + Z_m y_m$ then relations (27) and (28) become

$$\begin{aligned} x_m &= P_4\tilde{x}_m + Q_3A^Hb, & (29)\\ r_m &= P_5(b - A\tilde{x}_m). & (30)\end{aligned}$$

Finally the second orthogonality condition $r_m \perp A\mathcal{K}_m(\hat{A}, \hat{r}_0)$ can then be stated as

$$r_m = P_5(b - A\tilde{x}_m) \perp A\mathcal{K}_m(P_5A, P_5r_0).$$

We summarize these developments in the following proposition (see [40, Theorem 2.2]).

Proposition 1 *The following two sets of conditions*

$$\begin{aligned} x_m &\in x_0 + \mathcal{W} + \mathcal{K}_m(\hat{A}, \hat{r}_0),\\ r_m &= b - Ax_m \perp (A\mathcal{W} + A\mathcal{K}_m(\hat{A}, \hat{r}_0)),\end{aligned}$$

and

$$\begin{aligned} \tilde{x}_m &\in x_0 + \mathcal{K}_m(\hat{A}, \hat{r}_0),\\ \tilde{r}_m &= P_5(b - A\tilde{x}_m) \perp A\mathcal{K}_m(\hat{A}, \hat{r}_0)\end{aligned}$$

are equivalent in the sense that

$$x_m = P_4\tilde{x}_m + Q_3A^Hb \quad \textit{and} \quad r_m = \tilde{r}_m. \qquad (31)$$

The first set of conditions corresponds to the standard augmentation approach described in Section 3.1. In this class of methods the augmentation space $\mathcal{W}$ is explicitly included in the search space S of the minimum residual Krylov subspace method and $\hat{A} = A$, $\hat{r}_0 = r_0$. The second set of conditions corresponds to the standard deflation approach described in Section 3.2. Indeed the iteration $\tilde{x}_m$ is first obtained such that the residual $P_5(b - A\tilde{x}_m)$ satisfies the Petrov-Galerkin orthogonality condition. Then a correction is added such that $r_m = \tilde{r}_m$. Both approaches are found to be equivalent. They only differ in the way the augmentation subspace is treated (explicitly or implicitly).

3.3.2 Methods based on augmentation and deflation

Methods based on both augmentation and deflation have been introduced recently; see, e.g., [6, 16, 17, 81]. We focus here on the Generalized Conjugate Residual with inner Orthogonalization (GCRO) [16], which combines augmentation and deflation judiciously as detailed next.

GCRO belongs to the family of inner-outer methods [3, Ch. 12] where the outer iteration is based on the Generalized Conjugate Residual method (GCR), a minimum residual norm Krylov subspace method proposed by Eisenstat, Elman and Schultz [22] while the inner part is based on GMRES respectively. Following the theoretical framework introduced in [21], GCR maintains a correction subspace spanned by Range(Z_k) and an approximation subspace spanned by Range(V_k), where $Z_k, V_k \in \mathbb{C}^{n\times k}$ satisfy the relations

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

The optimal solution of the minimization problem min $\|b-Ax\|$ over the subspace $x_0+\text{Range}(Z_k)$ is then found as $x_k = x_0 + Z_k\ V_k^H\ r_0$. Consequently $r_k = b - Ax_k$ satisfies

$$r_k = r_0 - V_k\ V_k^H r_0 = (I_n - V_k\ V_k^H)r_0, \quad r_k \perp \text{Range}(V_k).$$

In [16] de Sturler suggested that the inner iteration takes place in a subspace orthogonal to the outer Krylov subspace. In this inner iteration the following projected linear system is considered

$$(I_n - V_k\ V_k^H)Az = (I_n - V_k\ V_k^H)r_k = r_k.$$

The inner iteration is then based on a deflated linear system with $(I_n - V_k\ V_k^H)$ as orthogonal projection. If a minimum residual norm subspace method is used in the inner iteration to solve this projected linear system approximately, the residual over both the inner and outer subspaces are minimized. Hence, augmentation is applied in the outer iteration and deflation in the inner part of the method. Numerical experiments (see, e.g., [16] and [27, Chapter 1]) indicate that the resulting method may perform better than other inner-outer methods (without orthogonalization) in some cases.

We mention that the augmentation subspace can be based on spectral approximate invariant subspace information. This leads to the GCRO with deflated restarting method (GCRO-DR) [65] that uses Harmonic Ritz information to define the augmentation subspace as in Section 3.1. This method has been further extended to accommodate variable preconditioning leading to the FGCRO-DR method [13]. We also refer the reader to [13] for additional comments on the computational cost of FGCRO-DR and a detailed comparison with the flexible variant of GMRES-DR. When a fixed right preconditioner is used, GMRES-DR and GCRO-DR are equivalent. When variable preconditioning is considered, it is however worthwhile to note that FGMRES-DR and FGCRO-DR are only equivalent if a certain collinearity condition given in [13, Theorem 3.6] is satisfied.

In [17] de Sturler proposed to define an augmentation subspace based on information other than approximate spectral invariant subspace. At the end of each cycle, the strategy (named GCRO with optimal truncation (GCROT)) decides which part of the current global search subspace to keep to define the new augmentation subspace such that the smallest inner residual norm is obtained. This truncation is done by examining angles between subspaces and requires specification of six different parameters that affect the truncation. We refer to [17] for a complete derivation of the method and numerical experiments (see also [21, Section 4.5]). Finally we note that the extension of GCROT to the case of variable preconditioning has been proposed in [42] with application to aerodynamics.

4 Hermitian positive definite matrices

Similarly to unsymmetric problems both augmentation and deflation can be considered to speed-up the convergence of the conjugate gradient method, possibly in combination with preconditioning. However, contrarily to the previous methods based on Arnoldi basis construction, the conjugate gradient method relies on a short term recurrence and restarting mechanisms do not need to be implemented to control the memory consumption. Consequently the space used for augmentation or for deflation should be fully defined before starting the iteration for a given right-hand side.

4.1 Augmented conjugate gradient methods

As discussed in Section 3.1, the search space in augmented methods $S_\ell = W \oplus K_m(A, r_0)$ is a ℓ dimensional space (with $\ell = m + k$) where $\mathcal{W}$ is an augmentation subspace of dimension k

spanned by k linearly independent vectors $W = [w_1, \cdots w_k]$. In order to build a basis of this space a deflated Lanczos algorithm can be used, that consists in applying a standard Lanczos method starting from an unit vector v_1 using the matrix

$$B = A - AW(W^H AW)^{-1} W^H A.$$

Notice that $W^H AW$ is full rank since A is HPD. If v_1 is orthogonal to W, deflated Lanczos builds a sequence of orthonormal vectors $V_m = [v_1, \cdots, v_m]$ ($V_m^H AV_m = I_m$) that spans a space orthogonal to W, i.e., $W^H V_m = 0_{k\times m}$. For the Krylov subpace part of S_ℓ, x_0 is chosen so that $r_0 = b - Ax_0 \perp W$ and $v_1 = r_0/\|r_0\|$. That can be guaranteed by defining x_0 from any x_{-1} as $x_0 = x_{-1} + W(W^H AW)^{-1} W^H r_{-1}$. The augmented CG algorithm seeks for a solution $x_\ell = x_0 + W\mu_\ell + V_m y_\ell \in x_0 + W + K_m(A, r_0)$ with the Ritz-Galerkin condition $r_\ell = b - Ax_\ell \perp (W + K_m(A, r_0))$. Using the above described space and orthogonality condition, it is shown in [71], that the following properties (that are very similar and inherited from the classical CG) still hold.

Proposition 2 *The iterate x_j, the residual r_j and the descent directions p_j satisfy the following relations and properties*

- *r_j is collinear to v_{j+1}, that is, the residual vectors are orthogonal to each other,*
- *the short term recurrences are satisfied:*

$$\begin{aligned} x_j &= x_{j-1} + \alpha_{j-1} p_{j-1} \\ r_j &= r_{j-1} - \alpha_{j-1} A p_{j-1} \\ p_j &= r_j + \beta_{j-1} p_{j-1} - W\mu_j \end{aligned}$$

  where α_{j-1} and β_{j-1} have the same expression as in classical CG and

$$\mu_j = (W^H AW)^{-1} W^H A r_j,$$

- *the vectors p_j are A-orthogonal to each other as well as A-orthogonal to all the w_j's.*

Using theoretical results from [25], the following properties related to convergence rate and optimization property of the iterate are shown in [71].

Proposition 3 *The approximate solution x_ℓ is such that*

- *the convergence history exhibits an upper bound expression on the convergence rate similar to classical CG*

$$\|x_\ell - x^\star\|_A \leq 2 \cdot \left(\frac{\sqrt{\kappa(P_{W^{\perp_A}}^H A P_{W^{\perp_A}})} - 1}{\sqrt{\kappa(P_{W^{\perp_A}}^H A P_{W^{\perp_A}})} + 1} \right)^\ell \|x_0 - x^\star\|_A, \tag{32}$$

  where $\kappa(\cdot)$ denotes the condition number of the matrix and $P_{W^{\perp_A}}$ is the A-orthogonal projection on $W^{\perp_A}$. This projection is defined by $P_{W^{\perp_A}} = I_n - W(W^H AW)^{-1} W^H A$.

- *similarly to classical CG, the iterate complies with a minimum A-norm error on the search space $x_\ell = \underset{x \in x_0 + W + K_m(A, r_0)}{\operatorname{argmin}} \|x - x^\star\|_A$.*

Augmenting using an invariant subspace Let $(\lambda_1 = \lambda_{\min}, \cdots, \lambda_s = \lambda_{\max})$ denote the s distinct eigenvalues of A ordered by increasing magnitude (i.e., values as they are real positive). The invariant subspace spanned by the k extreme (either largest or smallest) eigenvalues can be used in place of $\mathcal{W}$ to build the augmented space. Equation (32) shows that $\kappa(A) = \dfrac{\lambda_{\max}}{\lambda_{\min}}$ (that would appear in this bound for classical CG) is replaced either by $\dfrac{\lambda_{\max}}{\lambda_{k+1}}$ if the left most part of the spectrum is used or by $\dfrac{\lambda_{s-k}}{\lambda_{\min}}$ if the right most part is used. Consequently if $\lambda_{\min} \ll \lambda_{k+1}$ $(\lambda_{s-k} \ll \lambda_{\max})$ the convergence of augmented CG should be significantly faster than the convergence of CG on the original system.

4.2 Deflated Krylov subspace methods

We next briefly describe CG variants based on deflation. As mentioned in Section 3.2, the general idea of deflation is to split the approximation space into two complementary subspaces. Similarly to the notation in the previous sections, we denote by $\mathcal{W}$ a subspace of $\mathbb{C}^n$ of dimension k, where k is assumed to be much smaller than the problem dimension n. We later denote by $W \in \mathbb{C}^{n\times k}$ a matrix whose columns form a basis of $\mathcal{W}$. Because A is Hermitian positive definite, $W^H AW$ is also HPD and hence invertible. We can then define the following projector

$$P_6 = I - W(W^H AW)^{-1} W^H A \tag{33}$$

that is an oblique projector along W (P_6 is equal to $P_{W^{\perp_A}}$). As in the non-Hermitian case, we decompose the solution $x^\star = (I - P_6)x^\star + P_6 x^\star$ and compute each component separately. In particular, $(I - P_6)x^\star = W(W^H AW)^{-1} W^H Ax^\star = W(W^H AW)^{-1} W^H b$ essentially reduces to the solution of a small $k \times k$ system. For the calculation of the second component $P_6 x^\star$, it can be observed that $AP_6 = P_6^H A$ so that $AP_6 x^\star = P_6^H Ax^\star = P_6^H b$. Even though the matrix $P_6^H A$ is Hermitian semi-definite positive of rank $n - k$ (its nullspace is W), CG can still be used because the deflated linear system $P_6^H Ax^\star = P_6^H b$ is consistent [45]. Furthermore, because the null space never enters the iteration, the corresponding zero eigenvalues do not influence the convergence [45] and we can define the effective condition number of the positive semidefinite matrix $P_6^H A$, denoted $\kappa_{eff}(P_6^H A)$, as the ratio of its largest to smallest strictly positive eigenvalues.

Once the linear system $P_6^H A\tilde{x} = P_6^H b$ is solved, one just needs to apply P_6 to this solution to compute the second component of the solution. This technique still requires the solution of a linear system of size n using the CG method, but is expected to be more effective if $\kappa_{eff}(P_6^H A) \ll \kappa(A)$. We refer the reader to [28] for a discussion on the choice of W.

Deflating using an invariant subspace If W defines an invariant subspace of A associated with extreme eigenvalues, the situation becomes much clearer.

Let assume that W defines an invariant subspace associated with the smallest eigenvalues $(\lambda_1, ..., \lambda_k)$ of A. We have $P_6^H AW = 0_{m\times k}$ so that $P_6^H A$ has k zero eigenvalues. Because A is HPD, $Z = W^\perp$, the orthogonal complement of W (i.e., $W^H Z = 0$ so that $P_6^H Z = Z$) defines an invariant subspace associated with the eigenvalues $\lambda_{k+1}, ..., \lambda_n = \lambda_{\max}$. Therefore, we have $AZ = ZB$ for some nonsingular B. Consequently we have $P_6^H AZ = P_6^H ZB = ZB$ so that Z is an invariant subspace of $P_6^H A$ associated with the same eigenvalues $\lambda_{k+1}, ..., \lambda_{\max}$. This shows that

$$\kappa_{eff}(P_6^H A) = \frac{\lambda_{\max}}{\lambda_{k+1}},$$

that indicates that deflating using an invariant subspace cancels the corresponding eigenvalues, leaving the rest of the spectrum unchanged. If $\lambda_{k+1} \gg \lambda_1 = \lambda_{\min}$ the convergence of CG is significantly speeded-up.

4.3 Deflation via preconditioning

Using spectral information, it is possible to design preconditioners that enable to exhibit a condition number for the preconditioned matrix similar to $\kappa_{eff}(P_6^H A)$.

Let $W = [w_1, \dots w_k] \in \mathbb{C}^{n\times k}$ be the normalized eigenvectors of A associated with $\{\lambda_i\}_{i=1,\dots,k}$ the set of smallest eigenvalues. Let ν be a real positive value. We can then define the preconditioner

$$M^{def} = I_n + W(\nu(W^H AW)^{-1} - I_k)W^H.$$

This preconditioner is such that $M^{def}AW = \nu W$ and $M^{def}Aw = Aw$ if $W^H w = 0$ (in particular any eigenvectors of A not in W), which shows that M^{def} moves the eigenvalues $\{\lambda_i\}_{i=1,\dots,k}$ to ν and leaves the rest of the spectrum unchanged. If $\nu = \lambda_{k+1}$, the condition number of the preconditioned matrix is the same as the one of the deflated matrix in the previous section.

Furthermore we can define additive coarse space correction preconditioners inspired from domain decomposition techniques. They lead to preconditioned matrices with similar condition number as well. We then define

$$M^{coarse} = I_n + \nu W(W^H AW)^{-1}W^H.$$

This preconditioner is such that $M^{coarse}Aw_i = (\nu + \lambda_i)w_i$ and $M^{coarse}Aw = Aw$ if $W^H w = 0$. That is, the eigenvalues $\{\lambda_i\}_{i=1,\dots,k}$ are shifted to $\nu + \lambda_i$, while the rest of the spectrum is unchanged. If it exists ν so that $\lambda_{k+1} \leq \lambda_{\min} + \nu \leq \lambda_k + \nu \leq \lambda_{\max}$, the preconditioned matrix would have again the same condition number as the one of the deflated system κ_{eff}.

We refer the reader to [35] for an analysis of the condition number of this class of preconditioners when approximated spectral information is used. We also refer to [78] and the references therein for the exposure of various preconditioning techniques that can be defined using various combinations of these building box components.

5 Linear systems with multiple right-hand sides given in sequence

Although our primary focus is the solution of a single linear system with preconditioned Krylov subspace methods, it is however possible to include deflation and augmentation in a broader setting. Indeed in many applications in computational science and engineering, linear systems with multiple right-hand sides have to be solved. More precisely we are interested in solving a sequence of linear systems defined as $A^l x^l = b^l$ where both the non-Hermitian matrix $A^l \in \mathbb{C}^{n\times n}$ and the right-hand side $b^l \in \mathbb{C}^n$ may change from one system to the next, and the linear systems may typically not be available simultaneously. If we consider a sequence of identical or *slowly* changing matrices, Krylov subspace methods based on augmentation and deflation are appropriate since subspace recycling is then possible. The key idea is to extract relevant information (e.g. approximate invariant subspace but not only) while solving a given system, and then to use this information to further accelerate the convergence of the Krylov subspace method for the subsequent linear systems. At this point, augmented or deflated Krylov subspace methods of Sections 3.1, 3.2 and 3.3 can then be used. We refer the reader to [64, Chapter 3] for a detailed analysis of subspace recycling in the non-Hermitian case and to [65] where

the GCRO method augmented with approximate spectral information is shown to be efficient on applications related to fatigue and fracture of engineering components, electronic structure calculation and quantum chromodynamics; see also [47] for an application in optical tomography. Recent applications are related to model reduction [7] (see also [1] for recycling methods based on BiCG).

For HPD matrices, if a sequence of linear systems with the same matrix but different right-hand sides has to be solved different alternatives can be considered to define the space to augment the search space from one solve to the next. In [25], an approach based on harmonic Ritz values is described that might be implemented using only the first $m \geq k$ steps of augmented CG iteration. Still to reduce the memory footprint of the eigenvector calculation, in [76] a thick-restart Lanczos is embedded in the CG iterations to extract accurate spectral information.

6 Conclusions and prospectives

We have briefly reviewed the main features and mathematical properties of augmented and deflated Krylov subspace methods for the solution of certain linear systems of equations where the coefficient matrix was either non-Hermitian or Hermitian positive definite. These increasingly popular procedures combined with preconditioning have been shown effective on a wide range of applications in computational science and engineering as mentioned in this paper. We are certainly aware that this brief overview is far from being complete. Results related to two-sided Krylov subspace methods in the non-Hermitian case or the treatment of the Hermitian indefinite case are indeed missing; see, e.g., [1, 30, 31, 83] for additional comments and references. Similarly, the solution of linear systems with multiple right-hand sides given at once has not been covered. For such a class of problems, augmented and deflated block Krylov subspace methods have been studied (see, e.g., [57, 84]) and their efficiency has been proved on realistic applications. Finally we would like to mention that algebraic connections between deflation, multigrid and domain decomposition have been made in recent papers [48, 59, 78].

Concerning implementation aspects, some augmentation and deflation procedures are already present in the main software projects such as either PETSc[1] or Trilinos[2] for the solution of large-scale, complex multi-physics engineering and scientific problems. More precisely, in its scalable linear equation solvers (KSP) component, PETSc includes an algorithm described in [24], while the Belos package in Trilinos notably proposes an augmented and deflated approach based on GCRO-DR [65]. Most likely there will be a growing effort to incorporate augmented and deflated Krylov subspace methods in such libraries in a near future. Finally designing variants or new Krylov subspace methods for the next generation of massively parallel computing platforms is currently a topic of active research in the numerical linear algebra community; see [33, 34, 43] for algorithms, comments and references. Thus in a near future it is highly probable that variants of augmented and deflated Krylov subspace methods will be proposed as well.

References

[1] K. Ahuja, E. de Sturler, S. Gugercin, and E. Chang. Recycling BiCG with an application to model reduction. *SIAM J. Scientific Computing*, 34(4):A1925–A1949, 2012.

[2] B. Aksoylu and H. Klie. A family of physics-based preconditioners for solving elliptic equations on highly heterogeneous media. *Applied Numerical Mathematics*, 59:1159–1186, 2009.

[1] http://www.mcs.anl.gov/petsc/
[2] http://trilinos.sandia.gov/

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

[4] 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.

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

[6] 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.

[7] P. Benner and L. Feng. Recycling Krylov subspaces for solving linear systems with successively changing right-hand sides arising in model reduction. In Peter Benner, Michael Hinze, and E. Jan W. ter Maten, editors, *Model Reduction for Circuit Simulation*, volume 74 of *Lecture Notes in Electrical Engineering*, pages 125–140. Springer Netherlands, 2011.

[8] M. Benzi. Preconditioning techniques for large linear systems: A survey. *J. Comp. Phys.*, 182:418–477, 2002.

[9] P. N. Brown and H. F. Walker. GMRES on (nearly) singular systems. *SIAM J. Matrix Analysis and Applications*, 18:37–51, 1997.

[10] C. Le Calvez and B. Molina. Implicitly restarted and deflated GMRES. *Numerical Algorithms*, 21:261–285, 1999.

[11] M. H. Carpenter, C. Vuik, P. Lucas, M. B. van Gijzen, and H. Bijl. A general algorithm for reusing Krylov subspace information. I. Unsteady Navier-Stokes. NASA/TM 2010216190, NASA, Langley Research Center, 2010.

[12] B. Carpentieri, I.S. Duff, and L. Giraud. A class of spectral two-level preconditioners. *SIAM J. Scientific Computing*, 25(2):749–765, 2003.

[13] L. M. Carvalho, S. Gratton, R. Lago, and X. Vasseur. A flexible Generalized Conjugate Residual method with inner orthogonalization and deflated restarting. *SIAM J. Matrix Analysis and Applications*, 32(4):1212–1235, 2011.

[14] A. Chapman and Y. Saad. Deflated and augmented Krylov subspace techniques. *Numerical Linear Algebra with Applications*, 4(1):43–66, 1997.

[15] D. Darnell, R. B. Morgan, and W. Wilcox. Deflation of eigenvalues for iterative methods in lattice QCD. *Nuclear Physics B - Proceedings Supplements*, 129-130:856–858, 2004.

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

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

[18] Z. Dostál. Conjugate gradient method with preconditioning by projector. *Int. J. Comput. Math.*, 23:315–323, 1988.

[19] Z. Dostál. Projector preconditioning and domain decomposition methods. *Appl. Math. Comput.*, 37:75–81, 1990.

[20] M. Eiermann and O. G. Ernst. Geometric aspects of the theory of Krylov subspace methods. *Acta Numerica*, 10:251–312, 2001.

[21] 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.

[22] 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.

[23] J. Erhel. Some properties of Krylov projection methods for large linear systems. *Computational Technology Reviews*, 3:41–70, 2011.

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

[25] J. Erhel and F. Guyomarc'h. An augmented conjugate gradient method for solving consecutive symmetric positive definite linear systems. *SIAM J. Matrix Analysis and Applications*, 21(4):1279–1299, 2000.

[26] Y. Erlangga and R. Nabben. Deflation and balancing preconditioners for Krylov subspace methods applied to nonsymmetric matrices. *SIAM J. Matrix Analysis and Applications*, 30(2):684–699, 2008.

[27] D. Fokkema. *Subspace methods for linear, nonlinear and eigen problems.* PhD thesis, University of Utrecht, The Netherlands, 1996.

[28] J. Frank and C. Vuik. On the construction of deflation-based preconditioners. *SIAM J. Scientific Computing*, 23:442–462, 2011.

[29] A. Frommer, A. Nobile, and P. Zingler. Deflation and flexible SAP-preconditioning of GMRES in lattice QCD simulation. Technical Report BUW-IMACM 12/11, University of Wuppertal, Department of Mathematics, 2012.

[30] A. Gaul, M. Gutknecht, J. Liesen, and R. Nabben. Deflated and augmented Krylov subspace methods: Basic facts and a breakdown-free deflated MINRES. Preprint Preprint 759, DFG Research Center MATHEON, TU Berlin, 2011.

[31] A. Gaul, M. Gutknecht, J. Liesen, and R. Nabben. A framework for deflated and augmented Krylov subspace methods. arXiv preprint 1206.1506, 2012. submitted to SIAM J. Matrix Analysis and Applications.

[32] H. De Gersem and K. Hameyer. A deflated iterative solver for magnetostatic finite element models with large differences in permeability. *Eur. Phys. J. Appl. Phys.*, 67:45–49, 2000.

[33] P. Ghysels, T. Ashby, K. Meerbergen, and W. Vanroose. Hiding global communication latency in the GMRES algorithm on massively parallel machines. *SIAM J. Scientific Computing*, 35(1):C48–C71, 2013.

[34] P. Ghysels and W. Vanroose. Hiding global synchronization latency in the preconditioned Conjugate Gradient algorithm. Technical Report 12.2012.1, Intel ExaScience Lab Flanders, Leuven, December 2012.

[35] L. Giraud and S. Gratton. On the sensitivity of some spectral preconditioners. *SIAM J. Matrix Analysis and Applications*, 27(4):1089–1105, 2006.

[36] L. Giraud, S. Gratton, and E. Martin. Incremental spectral preconditioners for sequences of linear systems. *Applied Numerical Mathematics*, 57:1164–1180, 2007.

[37] L. Giraud, S. Gratton, X. Pinel, and X. Vasseur. Flexible GMRES with deflated restarting. *SIAM J. Scientific Computing*, 32(4):1858–1878, 2010.

[38] L. Giraud, D. Ruiz, and A. Touhami. A comparative study of iterative solvers exploiting spectral information for SPD systems. *SIAM J. Scientific Computing*, 27(5):1760–1786, 2006.

[39] G.H. Golub and D.P. O'Leary. Some history of the conjugate gradient and Lanczos algorithms: 1948-1976. *SIAM Review*, 31(1):50–102, 1989.

[40] M. Gutknecht. Spectral deflation in Krylov solvers: A theory of coordinate space based methods. *Electron.Trans. Numer. Anal.*, 39:156–185, 2012.

[41] M.R. Hestenes and E. Stiefel. Methods of conjugate gradients for solving linear systems. *J. Res. Nat. Bur. Standards*, 49:409–435, 1952.

[42] J. Hicken and D. Zingg. A simplified and flexible variant of GCROT for solving nonsymmetric linear systems. *SIAM J. Scientific Computing*, 32(3):1672–1694, 2010.

[43] M. Hoemmen. *Communication-avoiding Krylov subspace methods*. PhD thesis, University of California, USA, 2010.

[44] I. Ipsen and C. D. Meyer. The idea behind Krylov methods. *The American mathematical monthly*, 105(10):889–899, 1998.

[45] E. F. Kaasschieter. Preconditioned conjugate gradients for solving singular systems. *J. Comput. Appl. Math.*, 24:265–275, 1988.

[46] S. A. Kharchenko and A. Yu. Yeremin. Eigenvalue translation based preconditioners for the GMRES(k) method. *Numerical Linear Algebra with Applications*, 2:51–77, 1995.

[47] M. Kilmer and E. de Sturler. Recycling subspace information for diffuse optical tomography. *SIAM J. Scientific Computing*, 27(6):2140–2166, 2006.

[48] A. Klawonn and O. Rheinbach. Deflation, projector preconditioning, and balancing in iterative substructuring methods: connections and new results. *SIAM J. Scientific Computing*, 34(1):A459–A484, 2012.

[49] H. Klie, M.F. Wheeler, T. Clees, and K. Stueben. Deflation AMG solvers for highly ill-conditioned reservoir simulation problems. Paper SPE 105820 presented at the 2007 SPE Reservoir Simulation Symposium, Houston, TX, Feb. 28–30 2007, 2007.

[50] L. Y. Kolotilina. Twofold deflation preconditioning of linear algebraic systems. I. Theory. *Journal of Mathematical Sciences*, 89:1652–1689, 1998.

[51] C. Lanczos. Solution of systems of linear equations by minimized iterations. *J. Res. Nat. Bur. Standards*, 49(1):33–53, 1952.

[52] G. Meurant. *The Lanczos and Conjugate Gradient Algorithms: From Theory to Finite Precision Computations*. Software, Environments and Tools. Society for Industrial and Applied Mathematics, 2006.

[53] R. B. Morgan. Computing interior eigenvalues of large matrices. *Linear Algebra and its Applications*, 154-156:289–309, 1991.

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

[55] 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.

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

[57] R. B. Morgan. Restarted block GMRES with deflation of eigenvalues. *Applied Numerical Mathematics*, 54(2):222–236, 2005.

[58] M.F. Murphy, G.H. Golub, and A.J. Wathen. A note on preconditioning for indefinite linear systems. *SIAM J. Scientific Computing*, 21(6):1969–1972, 2000.

[59] R. Nabben and C. Vuik. A comparison of abstract versions of deflation, balancing and additive coarse grid correction preconditioners. *Numerical Linear Algebra with Applications*, 15:355–372, 2008.

[60] R. Nicolaides. Deflation of conjugate gradients with applications to boundary value problems. *SIAM J. Numerical Analysis*, 24:355–365, 2000.

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

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

[63] 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.

[64] M. Parks. *The iterative solution of a sequence of linear systems arising from nonlinear finite elements*. PhD thesis, University of Illinois at Urbana-Champaign, USA, 2005.

[65] 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.

[66] S. Röllin and W. Fichtner. Improving the accuracy of GMRes with deflated restarting. *SIAM J. Scientific Computing*, 30(1):232–245, 2007.

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

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

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

[70] 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.

[71] Y. Saad, M. Yeung, J. Erhel, and F. Guyomarc'h. A deflated version of the conjugate gradient algorithm. *SIAM J. Scientific Computing*, 21(5):1909–1926, 2000.

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

[73] V. Simoncini and D. B. Szyld. Theory of inexact Krylov subspace methods and applications to scientific computing. *SIAM J. Scientific Computing*, 25(2):454–477, 2003.

[74] 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.

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

[76] A. Stathopoulos and K. Orginos. Computing and deflating eigenvalues while solving multiple right hand side linear systems with an application to quantum chromodynamics. *SIAM J. Scientific Computing*, 32(1):439–462, 2010.

[77] J.M. Tang, S.P. Mac Lachlan, R. Nabben, and C. Vuik. A comparison of two-level preconditioners based on multigrid and deflation. *SIAM J. Matrix Analysis and Applications*, 31(4):1715–1739, 2010.

[78] J.M. Tang, R. Nabben, C. Vuik, and Y.A. Erlangga. Comparison of two-level preconditioners derived from deflation, domain decomposition and multigrid methods. *Journal of Scientific Computing*, 39:340–370, 2009.

[79] A. Toselli and O. Widlund. *Domain Decomposition methods - Algorithms and Theory.* Springer Series on Computational Mathematics, vol. 34, Springer, 2004.

[80] H. A. van der Vorst. *Iterative Krylov Methods for Large Linear Systems.* Cambridge University Press, 2003.

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

[82] C. Vuik, A. Segal, and J. A. Meijerink. An efficient preconditioned CG method for the solution of a class of layered problems with extreme contrasts in the coefficients. *J. Comp. Phys.*, 152:385–403, 1999.

[83] S. Wang, E. de Sturler, and G. H. Paulino. Large-scale topology optimization using preconditioned Krylov subspace methods with recycling. *Int J. Numerical Methods in Engineering*, 69(12):2441–2468, 2007.

[84] R. Yu, E. de Sturler, and D. D. Johnson. A block iterative solver for complex non-hermitian systems applied to large-scale electronic-structure calculations. Technical Report UIUCDCS-R-2002-2299, University of Illinois at Urbana-Champaign, Department of Computer Science, 2002.

RESEARCH CENTRE
BORDEAUX – SUD-OUEST

200 avenue de la Vielle Tour
33405 Talence Cedex

Publisher
Inria
Domaine de Voluceau - Rocquencourt
BP 105 - 78153 Le Chesnay Cedex
inria.fr

ISSN 0249-6399