BIT 36:2 (1996), 189–206.

ROUNDOFF ERROR ANALYSIS OF ALGORITHMS BASED ON KRYLOV SUBSPACE METHODS *

M. ARIOLI[1] and C. FASSINO[2] †

[1] *Istituto di Analisi Numerica, Consiglio Nazionale delle Ricerche via Abbiategrasso 209, I-27100 Pavia, Italy. email: arioli@dragon.ian.pv.cnr.it*

[2] *II Università degli Studi di Roma, Dipartimento di Matematica via Ricerca Scientifica 5, 00133 Roma, Italy. email: fassino@vax.mat.utovrm.it*

Abstract.

We study the roundoff error propagation in an algorithm which computes the orthonormal basis of a Krylov subspace with Householder orthonormal matrices. Moreover, we analyze special implementations of the classical GMRES algorithm, and of the Full Orthogonalization Method. These techniques approximate the solution of a large sparse linear system of equations on a sequence of Krylov subspaces of small dimension. The roundoff error analyses show upper bounds for the error affecting the computed approximated solutions.

AMS subject classification: 65F05.

Key words: Krylov subspace methods, GMRES, roundoff error analysis.

1 Introduction.

Let A be an $n \times n$ real matrix and b be an n real vector, we define the Krylov subspace $\mathcal{K}^j(A,b)$ as follows [12, 13]:

$$\mathcal{K}^j(A,b) = \text{span}\,\{b, Ab, A^2b, \ldots, A^{j-1}b\} \quad .$$

The Krylov subspaces are one of the most useful and powerful tools for computing approximate solutions in different Linear Algebra problems. It is possible, given the linear system

$$Ax = b,$$

to approximate the problem on a subspace of dimension $j \ll n$ with the use of Galerkin conditions: find

$$\hat{x}_j \in \mathcal{K}^j(A,b)$$

such that

$$A\hat{x}_j - b \quad \perp \mathcal{K}^j(A,b) \quad , \tag{1.1}$$

or

$$A\hat{x}_j - b \quad \perp A\mathcal{K}^j(A,b) \quad . \tag{1.2}$$

*Received March 1995. Revised February 1996.

†This work was carried out with the financial contribution of the Human Capital and Mobility Programme of the European Union grant ERB4050PL921378.

Condition (1.1) generates the class of the so-called Full Orthogonalization Methods (FOM) and condition (1.2) the family of GMRES methods [15]. Similar considerations can be made for eigenvalue problems (see [13]). Many different algorithms, based on Krylov space methods, have been proposed in the literature for the solution of linear systems with non symmetric matrices: GMRES [15], FOM [14], Orthodir [9], Orthomin [18], Axelsson's method [1], Generalized CG [3, 6], Bi-CGSTAB [16], Lanczos based techniques [20]. More references and interesting overviews can be found in [17], [11] and [14].

In this paper, we will restrict ourselves to all the variants which compute an orthonormal basis of Krylov subspaces. In particular, we will analyze the roundoff error propagation for an algorithm proposed in [19], which computes an orthonormal basis for the Krylov subspace with the use of elementary Householder matrices ([21, 10]).

Recently, Drkosova, Greenbaum, Rozloznik and Strakos [4] proved that, if the dimension of the Krylov space is n and its orthonormal basis is computed by a variant of the Arnoldi algorithm (HArnoldi [19]), which is based on the Householder elementary matrices technique, a class of GMRES algorithms is "weakly" backward stable. They prove that the computed solution at step n is the exact solution of a nearby problem, though not necessarily the approximation which would be generated by exact GMRES if applied to that nearby problem.

In our paper we focus on the forward error relative to Krylov spaces of dimension $j << n$. In particular, we compare the local forward roundoff error with the approximation error and we propose an upper bound for the global error at each step of the GMRES and the FOM algorithms.

In Section 2, we refine some of the proofs presented in [4]; in Section 3, we present results on the forward stability of GMRES and FOM algorithms which use the HArnoldi algorithm to compute the Krylov orthonormal basis. Finally, in the last section, we draw some conclusions.

Hereafter, we will consider that the finite precision arithmetic satisfies the following assumption. Let fl$(\cdot)$ denote the result of a floating point computation. We assume that the arithmetic of the computer satisfies the following for real arithmetic:

$$fl(a\Box b) = (a\Box b)(1+\delta(\Box, a, b)) \ ; \quad |\delta(\Box, a, b)| < \varepsilon \tag{1.3}$$

where ε is the machine precision and $\Box$ is one of $+ - */$. To a great extent, modern computers possess arithmetic that satisfies assumption (1.3) with the exception of some CRAY computers (such as CRAY2 and CRAY-YMP).

Furthermore, we assume that the scalar products are accumulated with the use of either mixed precision arithmetic or of the **"Kahan Summation Formula"** (for more details about these techniques we refer to [7]). As a consequence of these assumptions, for x and y real vectors of dimension n, we have:

$$fl(x^Ty) = x^Ty + x^TDy + s, \quad |D| \le 3\varepsilon I, \quad |s| \le \mathcal{O}(n\varepsilon^2|x|^T|y|).$$

This assumption is reasonable, in view of the large values that $n(=10^6, 10^7)$ takes in the practical problems where Krylov based algorithms are used.

Moreover, given a nonsingular $n \times n$ matrix B of entries B_{ij} and an n-vector v of entries v_i, we will denote by $\|B\|_2$ the usual spectral norm for a matrix, by $\|v\|_2$ the 2-norm of a vector, by $\kappa_2(B) = \|B^{-1}\|_2\|B\|_2$ the classical condition number of the matrix B and, finally, by

$$\|B\|_F = (\sum_{i,j=1}^{n} B_{ij}^2)^{1/2}, \qquad \kappa_F(B) = \|B^{-1}\|_F\|B\|_F$$

the Frobenius norm and the Frobenius condition number of B.

2 Roundoff error analysis of HArnoldi.

Let e_k be the kth column of the unit matrix, the Householder unitary transformation $Pc = e_1\rho$ uses

$$P = I - 2ww^T/\|w\|_2^2, \quad w = c - e_1\rho, \quad \rho = sign(c_1)\|c\|_2.$$

The roundoff properties of the Householder matrices are very favourable. In [21] it is shown that the computed version of P is very close to the true one:

$$\|fl(P) - P\|_2 \leq 84\varepsilon + \mathcal{O}(\varepsilon^2),$$

(see also [10] and [8]).

Moreover, the computed updates of a vector y with $fl(P)$ are close to the exact updates with P:

$$fl(fl(P)y) = P(y+w)\ , \quad \|w\|_2 \leq 87\varepsilon\|y\|_2 + \mathcal{O}(\varepsilon^2)\ ,$$

$$fl(y^T fl(P)) = (y+w)^T P\ , \quad \|w\|_2 \leq 87\varepsilon\|y\|_2 + \mathcal{O}(\varepsilon^2)\ ,$$

and, in general,

$$fl(fl(P_1)\dots fl(P_j)y) = P_1\dots P_j(y+z),$$
$$\|z\|_2 \leq 87j\varepsilon\|y\|_2 + \mathcal{O}(\varepsilon^2)\ .$$

With the previous notations, the HArnoldi algorithm ([19]) is as follows:

HArnoldi algorithm.

Suppose v_1 is given with $\|v_1\|_2 = 1$.

1) Choose P_1 such that $P_1v_1 = e_1$.

2) For $m = 1, 2, \dots,$ do:

   a) Set $v_m = P_1\dots P_m e_m$.

   b) If $(v_1, Av_1, \dots, Av_m)$ has rank m, then stop;
   otherwise choose P_{m+1} such that
   $P_{m+1}\dots P_1(v_1, Av_1, \dots, Av_m)$ is upper triangular.

We denote by

- $\overline{P}_j$ the approximation, computed at the j-th step, of the true Householder matrix $\hat{P}_j$, which would have produced zeros in positions $j+1$ through n of the vector $fl(Av_{j-1})$;
- $\overline{Q}_j = \overline{P}_1 \ldots \overline{P}_j \begin{bmatrix} I_j \\ 0 \end{bmatrix}$ the matrix which approximates the basis of the Krylov space $\mathcal{K}^j(A,b)$;
- $\hat{Q}^{(j)} = \hat{P}_1 \ldots \hat{P}_j$ and $\hat{Q}_j = \hat{Q}^{(j)} \begin{bmatrix} I_j \\ 0 \end{bmatrix}$;

THEOREM 2.1. *Let $\overline{H}_j$ be the computed upper Hessenberg $(j+1) \times j$ matrix which is obtained after j steps of the HArnoldi algorithm. Thus, there exists a perturbation matrix E and an orthonormal matrix $\hat{Q}$ such that*

$$(A+E)\hat{Q}_j = \hat{Q}_{j+1}\overline{H}_j,$$

where

$\hat{Q}_j$ comprises the first j columns of $\hat{Q}$, and

$\|E\|_2 \leq \sqrt{n}(174n + 3\sqrt{n} + 87)\varepsilon\|A\|_2 + \mathcal{O}(\varepsilon^2)$.

Moreover, there exists a perturbation vector e such that the matrix $\hat{Q}_j$ is an orthonormal basis of $\mathcal{K}^j(A+E, b+e)$, where $\|e\|_2 \leq 87\varepsilon\|b\|_2 + \mathcal{O}(\varepsilon^2)$.

PROOF. We will show by induction that, after m steps of the algorithm, the computed matrix can be written in the form

$$\hat{P}_{m+1} \ldots \hat{P}_1(\overline{v}_1 + g_1, \overline{z}_1 + w_1, \ldots, \overline{z}_m + w_m).$$

Step 1.

Let $v_1 = b/\|b\|_2$, and $\overline{v}_1 = fl(v_1)$.

i) Compute the Householder matrix P_1 such that $P_1 v_1 = e_1$.

If $\overline{P}_1 = fl(P_1)$, from the roundoff analysis of the Householder transformations ([10]) there exists a Householder matrix $\hat{P}_1$ such that

$$fl(\overline{P}_1\overline{v}_1) = \hat{P}_1(\overline{v}_1 + g_1),$$

with

$$\|g_1\|_2 \leq 87\varepsilon\|\overline{v}_1\|_2 + \mathcal{O}(\varepsilon^2) = 87\varepsilon + \mathcal{O}(\varepsilon^2).$$

ii) Compute Av_1.

The computed value is

$$\overline{z}_1 \equiv fl(Av_1) = (A + G_1)\overline{v}_1,$$

where

$$\|G_1\|_2 \leq \varepsilon 3\sqrt{n}\|A\|_2.$$

iii) Choose the Householder matrix P_2 such that $P_2 P_1(v_1, Av_1)$ is upper triangular; note that the matrix P_2 only transforms the vector $P_1 A v_1$, and therefore the matrices $\overline{P}_2$ and $\hat{P}_2$ do not change the first column of

$$fl(\overline{P}_1(fl(v_1), fl(Av_1))).$$

Let $\hat{P}_2$ be the Householder matrix such that $\hat{P}_2 fl(\overline{P}_1(\overline{v}_1, A\overline{v}_1))$ is upper triangular; in practice ([10]), the algorithm computes a matrix $\overline{P}_2$ such that

$$fl(\overline{P}_2 \overline{P}_1 \overline{z}_1) = \hat{P}_2 \hat{P}_1 (\overline{z}_1 + w_1),$$

where

$$\|w_1\|_2 \leq 174\varepsilon \|\overline{z}_1\|_2 \leq 174\varepsilon \left(\|A\|_2 + \|G_1\|_2\right) \|\overline{v}_1\|_2 = 174\varepsilon \|A\|_2 + \mathcal{O}(\varepsilon^2).$$

Finally, we have:

$$fl(\overline{P}_2 \overline{P}_1 (\overline{v}_1, \overline{z}_1)) = \left(fl(\overline{P}_1 \overline{v}_1), fl(\overline{P}_2 fl(\overline{P}_1 z_1))\right) =$$

$$\left(\hat{P}_1(\overline{v}_1 + g_1), \hat{P}_2 \hat{P}_1 (\overline{z}_1 + w_1)\right) = \hat{P}_2 \hat{P}_1 (\overline{v}_1 + g_1, \overline{z}_1 + w_1),$$

where $\|g_1\|_2 \leq 87\varepsilon + \mathcal{O}(\varepsilon^2)$, and $\|w_1\|_2 \leq 174\varepsilon \|A\|_2 + \mathcal{O}(\varepsilon^2)$.

Step m.

We now consider the mth step, and we suppose that the algorithm has computed the matrices $\overline{P}_1, \ldots, \overline{P}_m$, to which the Householder matrices $\hat{P}_1, \ldots, \hat{P}_m$ are associated. Under the inductive hypothesis, the following matrix has been computed after m steps:

$$fl\left(\overline{P}_m \ldots \overline{P}_1 \left(\overline{v}_1, fl(Av_1), \ldots, fl(Av_m)\right)\right) =$$

$$\hat{P}_m \ldots \hat{P}_1 (\overline{v}_1 + g_1, \overline{z}_1 + w_1, \ldots, \overline{z}_m + w_m),$$

where $\overline{z}_i = fl(Av_i)$, and $\|w_i\|_2 \leq 87\varepsilon(i+1)\|A\|_2 + \mathcal{O}(\varepsilon^2)$.

i) Compute $v_m = P_1 \ldots P_m e_m$.

In practice, the algorithm computes

$$\overline{v}_m = fl(\overline{P}_1 \ldots \overline{P}_m e_m) = \hat{P}_1 \ldots \hat{P}_m (e_m + f_m),$$

where

$$\|f_m\|_2 \leq 87m\varepsilon + \mathcal{O}(\varepsilon^2).$$

ii) Compute Av_m.

The computed vector is

$$\overline{z}_m \equiv fl(Av_m) = (A + G_m)\overline{v}_m,$$

where

$$\|G_m\|_2 \leq \varepsilon 3\sqrt{n}\|A\|_2 + \mathcal{O}(\varepsilon^2).$$

iii) Choose a matrix P_{m+1} such that $P_{m+1}P_m \dots P_1(v_1, \dots, Av_m)$ is upper triangular; note that P_{m+1} only transforms the vector $P_m \dots P_1 Av_1$.
In practice, we obtain a matrix $\overline{P}_{m+1}$ such that

$$fl(\overline{P}_{m+1} \dots \overline{P}_1 \overline{z}_m) = \hat{P}_{m+1} \dots \hat{P}_1(\overline{z}_m + w_m),$$

where

$$\begin{aligned} \|w_m\|_2 &\leq 87(m+1)\varepsilon\|\overline{z}_m\|_2 \leq 87(m+1)\varepsilon\left[\|A\|_2 + \|G_m\|_2\right](1 + \|f_m\|_2) \\ &\leq 87(m+1)\varepsilon\|A\|_2 + \mathcal{O}(\varepsilon^2). \end{aligned}$$

After m steps of the algorithm, the computed matrix can be written in the form

$$\hat{P}_{m+1} \dots \hat{P}_1(\overline{v}_1 + g_1, \overline{z}_1 + w_1, \dots, \overline{z}_m + w_m).$$

Let $\hat{Q}^{(j)}$ be the matrix such that $\hat{Q}^{(j)T} = \hat{P}_j \dots \hat{P}_1$; since each $\hat{P}_i$ has the first $(i-1)$ rows equal to $e_1, \dots, e_{i-1}$, there results that, if $q_k^{(i)}$ is the k-th column of the matrix $\hat{Q}^{(i)}$, then

$$q_k^{(i)} = q_k^{(j)}, \qquad \forall j \geq i, \quad k = 1, \dots i.$$

Now, we consider the matrix computed after j steps

$$\hat{Q}^{(j+1)T}(\overline{v}_1 + g_1, \overline{z}_1 + w_1, \dots, \overline{z}_j + w_j);$$

it is a rectangular $n \times (j+1)$ matrix with an upper triangular structure.
We have that, $\forall i \leq j$,

$$\begin{aligned} \overline{z}_i + w_i &= (A + G_i)\overline{v}_i + w_i = (A + G_i)\hat{P}_1 \dots \hat{P}_i(e_i + f_i) + w_i \\ &= A\hat{Q}^{(i)}e_i + G_i\hat{Q}^{(i)}e_i + A\hat{Q}^{(i)}f_i + G_i\hat{Q}^{(i)}f_i + w_i \\ &= Aq_i^{(i)} + G_iq_i^{(i)} + A\hat{Q}^{(i)}f_i + G_i\hat{Q}^{(i)}f_i + w_i. \end{aligned}$$

Since $\forall j \geq i$ $q_i^{(i)}$ is the i-th column q_i of $\hat{Q}^{(j)}$, if we denote by $y_i = G_iq_i^{(i)} + A\hat{Q}^{(i)}f_i + G_i\hat{Q}^{(i)}f_i + w_i$, we have

$$\overline{z}_i + w_i = Aq_i + y_i,$$

and so we obtain

$$\begin{aligned} &\hat{Q}^{(j+1)T}(\overline{v}_1 + g_1, \overline{z}_1 + w_1, \dots, \overline{z}_j + w_j) \\ &\quad = \hat{Q}^{(j+1)T}(\overline{v}_1 + g_1, Aq_1 + y_1, \dots, Aq_j + y_j) \\ &\quad = \hat{Q}^{(j+1)T}\left[(\overline{v}_1, Aq_1, \dots, Aq_j) + (g_1, y_1, \dots, y_j)\right]. \end{aligned}$$

If we neglect the first column of this matrix, we obtain an $n \times j$ matrix $\overline{\overline{H}}_j$ in the form

$$\overline{\overline{H}}_j = \begin{bmatrix} \overline{H}_j \\ 0 \end{bmatrix},$$

where $\overline{H}_j$ is a $(j+1) \times j$ upper Hessenberg matrix.

If we denote $F_p = [y_1, \ldots, y_p]$, $\forall p$, we have

$$\overline{\overline{H}}_j = \hat{Q}^{(j+1)T} \left[A[q_1, \ldots, q_j] + F_j \right];$$

moreover, if $\hat{Q}_p$, $\forall p \leq j$, is the $n \times p$ matrix defined by the first p columns of the matrix $\hat{Q}^{(j)}$ (note that since $p \leq j$ the first p columns of $\hat{Q}^{(j)}$ are equal to the first p columns of $\hat{Q}^{(i)}$, $i > j$), we have

$$\hat{Q}^{(j+1)}\overline{\overline{H}}_j = A\hat{Q}_j + F_j,$$

and, from the structure of $\overline{\overline{H}}_j$,

$$\hat{Q}_{j+1}\overline{H}_j = A\hat{Q}_j + F_j.$$

Since $\hat{Q}_n^T \hat{Q}_j = \begin{bmatrix} I_j \\ 0 \end{bmatrix}$ then $F_j = F_n \hat{Q}_n^T \hat{Q}_j$, and so

$$\hat{Q}_{j+1}\overline{H}_j = (A + F_n \hat{Q}_n^T)\hat{Q}_j.$$

If $E = F_n \hat{Q}_n^T$, we obtain the first part of the thesis, because

$$\begin{aligned}
\|E\|_2 &\leq \|E\|_F = \|F_n\|_F \\
&= \sqrt{\sum_{j=1}^{n} \|y_j\|_2^2} \leq \sqrt{\sum_{j=1}^{n} \|A\hat{Q}^{(j)} f_j + G_j q_j + w_j\|_2^2 + \mathcal{O}(\varepsilon^2)} \\
&\leq \sqrt{n} \max_j \|A\hat{Q}^{(j)} f_j + G_j q_j + w_j\|_2 + \mathcal{O}(\varepsilon^2) \\
&\leq \varepsilon\sqrt{n}(3\sqrt{n} + 174n + 87n)\|A\|_2 + \mathcal{O}(\varepsilon^2).
\end{aligned}$$

Finally, we show that the matrix $\hat{Q}_j$ is an orthonormal basis for a Krylov space. We denote $\tilde{A} = A + E$ and $\tilde{b} = \|b\|_2 \hat{Q}_j e_1$ and we consider the Krylov space $\mathcal{K}^j(\tilde{A}, \tilde{b})$. Note that $\tilde{b} = \|b\|_2 \hat{Q}^{(1)} e_1 = \|b\|_2 \hat{P}_1 e_1$, and so, if we denote $e = \tilde{b} - b$, we have, from the error analysis of Householder transformations ([10]),

$$\|e\|_2 = \|b\|_2 \|(\hat{P}_1 - P_1)e_1\|_2 \leq \|b\|_2 \|\hat{P}_1 - P_1\|_2 \leq 87\varepsilon\|b\|_2 + \mathcal{O}(\varepsilon^2).$$

To conclude the proof of the theorem we show, by induction, that

$$\tilde{A}^p \tilde{b} = \hat{Q}_j t_p, \qquad \forall p \leq (j-1),$$

where each vector t_p has only the first $(p+1)$ components which are different from zero.

1) If $p = 1$ we have

$$\tilde{A}\tilde{b} = \|b\|_2 \tilde{A}\hat{Q}_j e_1 = \|b\|_2 \hat{Q}_{j+1}\overline{H}_j e_1 = \hat{Q}_j t_1,$$

since the last component of the vector $\overline{H}_j e_1$ is zero; in particular, t_1 has only the first two components which are different from zero.

2) We suppose that the relation is true for p and we show it for $p+1$:

$$\tilde{A}^{p+1}\tilde{b} = \tilde{A}\hat{Q}_j t_p = \hat{Q}_{j+1}\overline{H}_j t_p.$$

For the inductive hypothesis, t_p has only the first $(p+1)$ components which are different from zero, and so the vector $\overline{H}_j t_p$ has only $(p+2)$ non-zero components, and its last element is zero. Then we can conclude that

$$\tilde{A}^{p+1}\tilde{b} = \hat{Q}_j t_{p+1}.$$

□

REMARK 2.1. $\hat{P}_1$ is the exact Householder matrix which would have produced zeros in positions $2 \dots n$ of v_1, and so we have two different cases.

- If $\|b\|_2 = 1$, since $v_1 = b/\|b\|_2 = b$, then $fl(v_1) = v_1$ and $\hat{P}_1 = P_1$; in this case $\tilde{b} = \hat{P}_1 e_1 = P_1 e_1 = b$ and $e = 0$.
- If $\|b\|_2 \neq 1$, since $v_1 = b/\|b\|_2$, then $fl(v_1) \neq v_1$ and $\hat{P}_1 \neq P_1$; in this case $\tilde{b} \neq b$ and $e \neq 0$.

3 HGMRES and HFOM methods.

In this section we analyze the roundoff error introduced by the HGMRES and the HFOM algorithms, as well as versions of GMRES ([2, 14, 15, 17, 19]) and FOM ([2, 17]) respectively, which use the Householder transformations in order to compute an orthonormal basis of the Krylov space. More precisely, the two algorithms can be described as follows.

HGMRES method.

This algorithm is such that the j-th step computes:

1) by means of the HArnoldi algorithm, an orthonormal matrix Q_j, basis of the Krylov space $\mathcal{K}^j(A,b)$, such that

$$AQ_j = Q_{j+1}H_j,$$

where H_j is a $(j+1)\times j$ upper Hessenberg matrix;

2) a vector $x_j^G \in \mathcal{K}^j(A,b)$ such that

$$\|Ax_j^G - b\|_2 = \min_{z=Q_j u} \|Az - b\|_2.$$

More precisely, since

$$\min_{z=Q_j u} \|Az - b\|_2 = \min_u \|AQ_j u - b\|_2$$
$$= \min_u \|Q_{j+1}H_j u - Q_{j+1}e_1\|_2 = \min_u \|H_j u - e_1\|_2,$$

the algorithm computes a vector u_j^G such that

$$\|H_j u_j^G - e_1\|_2 = \min_u \|H_j u - e_1\|_2,$$

and then the vector x_j^G from the relation $x_j^G = Q_j u_j^G$.

<u>**HFOM** method.</u>

This algorithm is such that the j-th step computes:

1) by means of the HArnoldi algorithm, an orthonormal matrix Q_j, basis of the Krylov space $\mathcal{K}^j(A,b)$, such that

$$AQ_j = Q_{j+1}H_j,$$

where H_j is a $(j+1) \times j$ upper Hessenberg matrix;

2) a vector $x_j^F \in \mathcal{K}^j(A,b)$ such that

$$Q_j^T(Ax_j^F - b) = 0.$$

Since we have

$$Q_j^T(Ax_j^F - b) = Q_j^T(AQ_ju_j^F - b) = Q_j^TQ_{j+1}H_ju_j^F - e_1 = H_{jj}u_j^F - e_1,$$

where H_{jj} is the upper Hessenberg matrix which comprises the first j rows and j columns of H_j, the algorithm solves

$$H_{jj}u_j^F = e_1$$

and then computes $x_j^F = Q_ju_j^F$.

In the following we will assume that, during the HFOM algorithm, no breakdown occurs. In [2], remedies are described which can be used for the solution of these situations. We want to point out that these remedies can be analyzed separately, and that they can be seen as special cases of the roundoff error analysis that follows. In particular, the remedy which suggests the computation of the least squares solution of the system $H_{jj}u_j^F = e_1$ when H_{jj} is singular, is proved to be equivalent to a step of the HGMRES algorithm. Thus, the corresponding roundoff error analysis follows straight from the one done for HGMRES.

Since the first part of the previous algorithm is the same, the analysis of the roundoff error, which is introduced by the computation of the basis of the Krylov space, is identical. As reported in the previous section (see Theorem 2.1), if $\overline{H}_j$ is the upper Hessenberg matrix computed by the HArnoldi algorithm, then there exists an orthonormal matrix $\hat{Q}_j$ such that

$$(A+E)\hat{Q}_j = \hat{Q}_{j+1}\overline{H}_j,$$

where $\hat{Q}_j$ is an orthonormal basis of the Krylov space $\mathcal{K}^j(A+E, b+e)$. For this reason, the first part of HGMRES and HFOM computes the exact basis of the perturbed Krylov space $\mathcal{K}^j(A+E, b+e)$, and therefore this is equivalent to applying, in exact arithmetic, the HArnoldi algorithm to the matrix $A+E$ and to the vector $b+e$.

If we denote by

- x the solution of the system $Ax = b$,
- $\hat{x}$ the solution of the system $(A+E)\hat{x} = b+e$,

we have that the error introduced by point 1) of the previous two algorithms is the inherent error, due to the perturbations with the matrix E and the vector e of the data of the system $Ax = b$, and so we have

$$\frac{\|x-\hat{x}\|_2}{\|x\|_2} \leq \kappa_2(A)\frac{\frac{\|E\|_2}{\|A\|_2}+\frac{\|e\|_2}{\|b\|_2}}{1-\kappa_2(A)\frac{\|E\|_2}{\|A\|_2}},$$

if $\|A^{-1}\|_2\|E\|_2 < 1$.

Thus, from the analysis of the HArnoldi algorithm, we have

$$\begin{aligned} \|E\|_2 &\leq \varepsilon\sqrt{n}(174n + 3\sqrt{n} + 87n)\|A\|_2 + \mathcal{O}(\varepsilon^2), \\ \|e\|_2 &\leq 87\varepsilon\|b\|_2 + \mathcal{O}(\varepsilon^2), \end{aligned}$$

and we obtain

$$\frac{\|x-\hat{x}\|_2}{\|x\|_2} \leq \varepsilon\left[\sqrt{n}(174n + 3\sqrt{n} + 87n) + 87\right]\kappa_2(A) + \mathcal{O}(\varepsilon^2).$$

3.1 Error analysis of the HGMRES algorithm.

We denote by

- $\hat{x}_j^G = \hat{Q}_j\hat{u}_j^G$ the vector obtained at step j of the HGMRES algorithm which is applied to the system $(A+E)\hat{x} = b+e$, with exact arithmetic, i.e.

  $$\|(A+E)\hat{x}_j^G - (b+e)\|_2 = \min_{\hat{z}=\hat{Q}_j\hat{u}} \|(A+E)\hat{z} - (b+e)\|_2.$$

  Since

  $$\|(A+E)\hat{x}_j^G - (b+e)\|_2 = \min_{\hat{u}} \|\overline{H}_j\hat{u} - e_1\|_2,$$

  then $\hat{u}_j^G$ is the exact solution of the least square problem

  $$\min_{\hat{u}} \|\overline{H}_j\hat{u} - e_1\|_2;$$

- $\overline{x}_j^G = \hat{Q}_j\overline{u}_j^G$, where $\overline{u}_j^G$ is the computed solution of the problem

  $$\min_{\hat{u}} \|\overline{H}_j\hat{u} - e_1\|_2;$$

- $\overline{\overline{x}}_j^G = fl(\overline{Q}_j\overline{u}_j^G)$ the final vector computed at the j-th step by the HGMRES algorithm, which approximates the solution of the system $Ax = b$.

THEOREM 3.1. *Using the previous notations, the global roundoff error introduced by the HGMRES algorithm at step j can be bounded as follows:*

$$\begin{aligned} \|x - \overline{\overline{x}}_j^G\|_2 &\leq \varepsilon\left[\sqrt{n}(174n + 3\sqrt{n} + 87n) + 87\right]\kappa_2(A)\|x\|_2 + \|\hat{x} - \hat{x}_j^G\|_2 \\ &+ \varepsilon 7j\sqrt{j}\kappa_2(\overline{H}_j)\left\{2 + \kappa_2(\overline{H}_j)\frac{\|\overline{H}_j\hat{u}_j^G - e_1\|_2}{\|\overline{H}_j\|_2}\right\}\|\hat{x}_j^G\|_2 \\ &+ 87j\varepsilon\|\overline{x}_j^G\|_2 + \mathcal{O}(\varepsilon^2), \end{aligned}$$

if

$$\varepsilon\left[\sqrt{n}(174n+3\sqrt{n}+87n)\right]\kappa_2(A)<1.$$

PROOF. For simplicity, we have eliminated the superscript G during the proof. The total error can be written in the form

$$x-\overline{\overline{x}}_j=(x-\hat{x})+(\hat{x}-\hat{x}_j)+(\hat{x}_j-\hat{Q}_j\overline{u}_j)+(\hat{Q}_j\overline{u}_j-fl(\overline{Q}_j\overline{u}_j)).$$

We now analyze each part of this error, separately.

1) The first part has been analyzed in the introduction of this section.

2) $\|\hat{x}-\hat{x}_j\|_2$. This value is due to the approximation of the solution of the system $(A+E)\hat{x}=b+e$ by the vector obtained after j steps of the HGMRES algorithm, with the use of exact arithmetic.

3) $\|\hat{x}_j-\hat{Q}_j\overline{u}_j\|_2$. Note that, since $\hat{x}_j=\hat{Q}_j\hat{u}_j$, we have that

$$\|\hat{x}_j-\hat{Q}_j\overline{u}_j\|_2=\|\hat{u}_j-\overline{u}_j\|_2,$$

where the value $\|\hat{u}_j-\overline{u}_j\|_2$ is the roundoff error due to the solution of a least square problem by a Givens based method ([8]). In order to analyze this problem, we can use the results presented in [21] (pp 131-134). We have that the computed solution of a least square problem is the exact solution of a perturbed least square problem; moreover, there exists a study of the sensitivity of such a problem ([10]). Given the problem

$$\min_{\hat{u}}\|\overline{H}_j\hat{u}-e_1\|_2$$

with the exact solution $\hat{u}_j$, the computed solution $\overline{u}_j$ is the exact solution of

$$\min_{\hat{u}}\|(\overline{H}_j+\dot{E}_j)\hat{u}-(e_1+f_j)\|_2,$$

where

$$\|E_j\|_2\le\|E_j\|_F\le\varepsilon 7j\|\overline{H}_j\|_F+\mathcal{O}(\varepsilon^2),$$

and

$$\|f_j\|_2\le\varepsilon 6j+\mathcal{O}(\varepsilon^2).$$

Moreover, if

$$\gamma\equiv max\left\{\frac{\|E_j\|_2}{\|\overline{H}_j\|_2},\|f_j\|_2\right\}=\varepsilon 7j\sqrt{j}+\mathcal{O}(\varepsilon^2),$$

we have ([10])

$$\frac{\|\hat{u}_j-\overline{u}_j\|_2}{\|\hat{u}_j\|_2}\le\gamma\kappa_2(\overline{H}_j)\left\{2+\kappa_2(\overline{H}_j)\frac{\|\overline{H}_j\hat{u}_j-e_1\|_2}{\|\overline{H}_j\|_2}\right\}+O(\gamma^2).$$

4) $\|\hat{Q}_j\overline{u}_j - fl(\overline{Q}_j\overline{u}_j)\|_2$. From the analysis of the roundoff error of the Householder transformation of a vector ([10]) we obtain

$$\|\hat{Q}_j\overline{u}_j - fl(\overline{Q}_j\overline{u}_j)\|_2 \leq \|\hat{Q}_j\overline{u}_j - fl(\overline{Q}_j\overline{u}_j)\|_F \leq 87j\varepsilon\|\overline{u}_j\|_2 + \mathcal{O}(\varepsilon^2).$$

□

REMARK 3.1. In Theorem 3.1, we have shown that the computed vector $\overline{u}_j$ achieves the minimum of

$$\min_{\hat{u}} \|(\overline{H}_j + E_j)\hat{u} - (e_1 + f_j)\|_2,$$

and so, denoting $\overline{x}_j = \hat{Q}_j\overline{u}_j$ and $\tilde{E}_j = \hat{Q}_{j+1}E_j\hat{Q}_j^T$, we have

$$\|(A + E + \tilde{E}_j)\overline{x}_j - (\tilde{b} + \hat{Q}_{j+1}f_j)\| = \min_{\overline{x}=\hat{Q}_j\hat{u}} \|(A + E + \tilde{E}_j)\overline{x} - (\tilde{b} + \hat{Q}_{j+1}f_j)\|.$$

Unfortunately, we cannot conclude that $\overline{x}_j$ is the exact solution of a perturbed problem obtained by the HGMRES algorithm because $\overline{x}_j$ belongs to $\mathcal{K}^j(A+E, \tilde{b})$, but it does not belong to $\mathcal{K}^j(A + E + \tilde{E}_j, \tilde{b} + \hat{Q}_{j+1}f_j)$.

3.2 Error analysis of the HFOM algorithm.

We denote by

- $\hat{x}_j^F = \hat{Q}_j\hat{u}_j^F$ the vector obtained at step j of the HFOM algorithm which is applied to the system $(A + E)\hat{x} = b + e$, with exact arithmetic, i.e. $\hat{x}_j^F$ is the exact solution of

$$\hat{Q}_j^T\left((A + E)\hat{x}_j^F - (b + e)\right) = 0.$$

Since

$$\hat{Q}_j^T\left((A + E)\hat{x}_j^F - (b + e)\right) = \hat{Q}_j^T\hat{Q}_{j+1}\overline{H}_j\hat{u}_j^F - e_1 = \overline{H}_{jj}\hat{u}_j^F - e_1,$$

where $\overline{H}_{jj}$ is the $j \times j$ upper Hessenberg matrix which comprises the first j rows and columns of $\overline{H}_j$, then $\hat{u}_j^F$ is the exact solution of

$$\overline{H}_{jj}\hat{u}_j^F = e_1;$$

- $\overline{x}_j^F = \hat{Q}_j\overline{u}_j^F$, where $\overline{u}_j^F$ is the computed solution of the system

$$\overline{H}_{jj}\hat{u}_j^F = e_1;$$

- $\overline{\overline{x}}_j^F = fl(\overline{Q}_j\overline{u}_j^F)$ the final vector computed at the j-th step by the HFOM algorithm, which approximates the solution of the system $Ax = b$.

THEOREM 3.2. *Using the previous notations, the global roundoff error introduced by the HFOM algorithm at step j can be bounded as follows:*

$$\begin{aligned}\|x - \overline{\overline{x}}_j^F\|_2 &\le \varepsilon\left[\sqrt{n}(3\sqrt{n} + 174n + 87n) + 87\right]\kappa_2(A)\|x\|_2 + \|\hat{x} - \hat{x}_j^F\|_2 + \\ &\qquad \varepsilon j(3\sqrt{j} + 10j^2)\kappa_2\left(\overline{H}_{jj}\right)\|\hat{x}_j^F\|_2 + \\ &\qquad 87j\varepsilon\|\overline{x}_j^F\|_2 + \mathcal{O}(\varepsilon^2),\end{aligned}$$

if $\varepsilon\left[\sqrt{n}(3\sqrt{n} + 174n + 87n\right]\kappa_2(A) < 1$.

PROOF.

For simplicity, we have eliminated the superscript F during the proof. With these notations, the total error can be written in the form

$$x - \overline{\overline{x}}_j = (x - \hat{x}) + (\hat{x} - \hat{x}_j) + (\hat{x}_j - \hat{Q}_j\overline{u}_j) + (\hat{Q}_j\overline{u}_j - fl(\overline{Q}_j\overline{u}_j)).$$

We now analyze each part of this error, separately.

1) The first part has been analyzed in the introduction of this section.

2) $\|\hat{x} - \hat{x}_j\|_2$. This value is due to the approximation of the solution of the system $(A+E)\hat{x} = b+e$ by the vector obtained after j steps of the HFOM algorithm, using exact arithmetic.

3) $\|\hat{x}_j - \hat{Q}_j\overline{u}_j\|_2$. Note that, since $\hat{x}_j = \hat{Q}_j\hat{u}_j$, we obtain

$$\|\hat{x}_j - \hat{Q}_j\overline{u}_j\|_2 = \|\hat{u}_j - \overline{u}_j\|_2,$$

where the value $\|\hat{u}_j - \overline{u}_j\|_2$ is the roundoff error due to the solution of a linear system.

If we solve the system $\overline{H}_{jj}\hat{u}_j = e_1$ by means of Gaussian elimination with partial pivoting, we obtain a computed solution $\overline{u}_j$ which is the exact solution of a perturbed system ([8])

$$\left(\overline{H}_{jj} + \Delta_{jj}\right)\overline{u}_j = e_1,$$

where

$$|\Delta_{jj}| \le \varepsilon j\left[3\left|\overline{H}_{jj}\right| + 5\hat{\Pi}^T\left|\overline{L}_j\right|\left|\overline{U}_j\right|\right] + \mathcal{O}(\varepsilon^2),$$

where $\left|\overline{L}_j\right|\left|\overline{U}_j\right|$ is the computed LU factorization of $\overline{H}_{jj}$, and $\hat{\Pi}^T$ is a permutation matrix.

Since the matrix $\overline{H}_{jj}$ is upper Hessenberg, it is possible to show that $\hat{\Pi}^T\left|\overline{L}_j\right|\left|\overline{U}_j\right|$ and, consequently, $|\Delta_{jj}|$ are also upper Hessenberg matrices. Moreover, since we use Gaussian elimination with partial pivoting, we obtain

i) $\|\,|\overline{L}_j|\,\|_2 \le 2$;

ii) $\|\,|\overline{U}_j|\,\|_2 \le j^2 \max_{ik}|(\overline{H}_{jj})_{ik}| \le j^2\|\overline{H}_{jj}\|_2$.

From previous remarks it follows that

$$\frac{\|\Delta_{jj}\|_2}{\|\overline{H}_{jj}\|_2} \le \varepsilon j(3\sqrt{j} + 10j^2) + \mathcal{O}(\varepsilon^2),$$

from this relation, we can bound the norm of the difference between the exact solution $\hat{u}_j$ and the computed solution $\overline{u}_j$ ([8])

$$\frac{\|\hat{u}_j - \overline{u}_j\|_2}{\|\hat{u}_j\|_2} \le \varepsilon j(3\sqrt{j} + 10j^2)\kappa_2\left(\overline{H}_{jj}\right) + \mathcal{O}(\varepsilon^2).$$

4) $\|\hat{Q}_j\overline{u}_j - fl(\overline{Q}_j\overline{u}_j)\|_2$. From the analysis of the roundoff error of the Householder transformation of a vector ([10]) we obtain

$$\|\hat{Q}_j\overline{u}_j - fl(\overline{Q}_j\overline{u}_j)\|_2 \le \|\hat{Q}_j\overline{u}_j - fl(\overline{Q}_j\overline{u}_j)\|_F \le 87j\varepsilon\|\overline{u}_j\|_2 + \mathcal{O}(\varepsilon^2).$$

□

REMARK 3.2. We can show that the vector $\overline{u}_j$ computed at the j-th step can be obtained by applying j steps of the HFOM algorithm to a perturbed problem with the use of exact arithmetic. As shown before, we have that

$$\left(\overline{H}_{jj} + \Delta_{jj}\right)\overline{u}_j^F = e_1, \qquad (A+E)\hat{Q}_j = \hat{Q}_{j+1}\overline{H}_j.$$

Since $\overline{H}_{jj} = (I_j, 0)\overline{H}_j$, then we obtain

$$(I_j, 0)\overline{H}_j\overline{u}_j^F + \Delta_{jj}\overline{u}_j^F = e_1$$

$$\hat{Q}_j^T\hat{Q}_{j+1}\overline{H}_j\overline{u}_j^F + \Delta_{jj}\overline{u}_j^F = e_1,$$

and so

$$\hat{Q}_j^T\left((A+E)\hat{Q}_j\overline{u}_j^F + \hat{Q}_j\Delta_{jj}\hat{Q}_j^T\hat{Q}_j\overline{u}_j^F\right) = \hat{Q}_j^T\tilde{b}.$$

Denoting $\overline{x}_j^F = \hat{Q}_j\overline{u}_j^F$ and $E_j = \hat{Q}_j\Delta_{jj}\hat{Q}_j^T$, we obtain

$$\hat{Q}_j^T\left[(A+E+E_j)\overline{x}_j^F - \tilde{b}\right] = 0;$$

moreover, note that, denoting $\Delta_j = \begin{bmatrix} \Delta_{jj} \\ 0 \end{bmatrix}$, since the relations

$$(A+E+E_j)\hat{Q}_j = \hat{Q}_{j+1}(\overline{H}_j + \Delta_j),$$

$$\hat{Q}_je_1 = \tilde{b},$$

hold and Δ_j is a Hessenberg matrix, then $\hat{Q}_j$ is a basis of the Krylov space $\mathcal{K}^j(A+E+E_j, \tilde{b})$.

Thus we can conclude that, by applying j steps of the HFOM algorithm, with exact arithmetic, to the system

$$(A+E+E_j)x = \tilde{b}$$

we obtain the vectors $\overline{u}_j$ and $\overline{x}_j = \hat{Q}_j\overline{u}_j$; nevertheless, two remarks are necessary.

1) Although the computed vector $\overline{u}_j$ coincides with the vector obtained by the perturbed problem, the computed vector $\overline{\overline{x}}_j$ does not coincide with $\overline{x}_j$. Nevertheless, the difference between $\overline{x}_j$ and $\overline{\overline{x}}_j$ is upper bounded, as presented in Theorem 3.2, by

$$\|\overline{x}_j - \overline{\overline{x}}_j\|_2 \leq 87j\varepsilon\|\overline{u}_j\|_2 + \mathcal{O}(\varepsilon^2).$$

2) Since the perturbation matrix E_j depends on j, we cannot conclude that the sequence $\{\overline{u}_1, \ldots, \overline{u}_j\}$ is obtained by applying the HFOM algorithm to a unique perturbed problem with exact arithmetic. However, for any fixed index j, it is possible to find a particular perturbed problem which enables us to compute the vector $\overline{u}_j$ with exact arithmetic.

As an alternative to the Gaussian elimination, it is possible to solve the system $\overline{H}_{jj}\hat{u}_j^F = e_1$, by a QR factorization technique (see [21] and [8]). In this case the computed solution $\overline{u}_j^F$ is the exact solution of the perturbed system

$$\left(\overline{H}_{jj} + \Delta_{jj}\right)\overline{u}_j^F = e_1 + f_j,$$

where $\|\Delta_{jj}\|_2 \leq \varepsilon 7j\sqrt{j}\|\overline{H}_{jj}\|_2 + \mathcal{O}(\varepsilon^2)$ and $\|f_j\|_2 \leq \varepsilon 6j + \mathcal{O}(\varepsilon^2)$.

This approach is fairly attractive because it offers a natural way for computing the least squares solution of the system when a breakdown occurs. Nevertheless, what has been said in the first part of this Remark is no longer true, for the same reasons outlined in Remark 2.

4 Conclusions.

In this paper we presented a local forward error analysis for two variants of the GMRES and FOM algorithms. We analyzed two independent sources of roundoff errors. The first, which is common to both the algorithms, comes from the Krylov process. The second source of roundoff errors is the solution of a least squares problem for HGMRES and the solution of a linear system for HFOM.

We cannot consider the presence of the square of the condition number of $\overline{H}_j$ in the error bound of Theorem 3.1 too negatively because this is multiplied by the residual $\|\overline{H}_j\hat{u}_j^G - e_1\|_2$, which normally becomes small when $\overline{H}_j$ becomes ill conditioned. On the other hand, in the error bound of Theorem 3.2, the condition number of $\overline{H}_{jj}$ becomes very large in the presence of a breakdown in the HFOM algorithm. However, it is interesting to notice that the presence of a breakdown is locally dangerous at step j when the phenomenon happens, and $\overline{H}_{j+1,j+1}$ may be better conditioned than $\overline{H}_{j,j}$, thereby resulting in a better error estimate at step $j+1$.

We are aware that the presence of complicated expressions of the step value j and of the dimension of the problem n in the error bounds is rather cumbersome. However, in practical examples, they are normally linear in j and n. Unfortunately, even if A is quite sparse, the roundoff analysis presented here does not take advantage of the sparsity. The main reason for this is that the Householder algorithm operates not on A but on the full matrix $[b, Ab, \ldots, A^j b]$ with obvious consequences.

The results of this paper can only be partially extended to the analysis of the classical Arnoldi process and to the classical GMRES algorithm. In particular, the HGMRES method costs the double of the classical GMRES method in terms of floating point operations. Nevertheless, the orthonormal properties of the Krylov basis, which is computed by the HArnoldi process, strongly support the use of these Householder matrix based algorithms. This is particularly true when we need to be sure that the perturbed problem we are solving has to conserve some spectral similarity properties. This will be especially relevant when we need to compute approximations of the matrix eigenvalues.

Finally, the bounds we introduced in Theorems 3.1 and 3.2 show that approximation errors $\|\hat{x} - \hat{x}_j^G\|_2$ and $\|\hat{x} - \hat{x}_j^F\|_2$ are predominant until they reach the level of the roundoff error. This suggests that in many practical problems, the stopping criteria can be safely chosen on the basis of theoretical properties of the original problem, thereby diminishing the need for an extremely precise approximation of the algebraic problem solution. If $\frac{1}{2}(A^T + A)$ is symmetric and positive definite, it is possible to predict, a-priori, how many iterations we will need in order to reach the level of the roundoff error, on the basis of the analysis described in [5]. Unfortunately, as reported in [11], the same kind of analysis is not as accurate in more general cases.

REFERENCES

1. O. Axelsson, *Conjugate gradient type methods for unsymmetric and inconsistent systems of equations*, Linear Algebra Appl., 29 (1980), pp. 1–16.

2. P. N. Brown, *A theoretical comparison of the Arnoldi and GMRES algorithms*, SIAM J. Sci. Stat. Comput., 12 (1991), pp. 58–78.

3. P. Concus and G. H. Golub, *A generalized conjugate gradient method for non-symmetric systems of linear equations*, Tech. Rep. STAN–CS–76–535, Stanford University, Stanford CA, 1976.

4. J. Drkošová, A. Greenbaum, M. Rozložník, and Z. Strakoš, *Numerical stability of GMRES*, BIT 35 (1995), pp. 309–330.

5. S. C. Eisenstat, H. C. Elman, and M. H. Schultz, *Variational iterative methods for nonsymmetric systems of linear equations*, SIAM J. Numer. Anal., 20 (1983), pp.245–357.

6. H. C. Elman, *Iterative Methods for Large Sparse Nonsymmetric Systems of Linear Equations*, PhD thesis, Yale University, New Haven CT, 1982.

7. D. Goldberg, *What every computer scientist should know about floating-point arithmetic*, ACM Computing Surveys, 23, (1991), pp. 5–48.

8. G. H. Golub and C. VanLoan, *Matrix Computations. 2nd ed.*, The Johns Hopkins University Press, Baltimore, MD, 1989.

9. K. C. Jea and D. M. Young, *Generalized conjugate-gradient acceleration of non-symmetrizable iterative methods*, Linear Algebra Appl., 34 (1980), pp. 159–194.

10. C. Lawson and R. Hanson, *Solving Least Squares Problems*, Prentice Hall, Englewood Cliffs, NJ, 1974.

11. N. M. Nachtigal, S. C. Reddy, and L. N. Trefethen, *How fast are nonsymmetric matrix iterations?*, SIAM J. Matrix. Anal. Appl. 13 (1992), pp. 778–795.

12. B. N. Parlett, *The Symmetric Eigenvalue Problem*, Prentice Hall, Englewood Cliffs, NJ, 1980.

13. Y. Saad, *Numerical Methods for Large Eigenvalue problems*, Manchester University Press, Manchester, UK, 1992.

14. Y. Saad and M. H. Schultz, *Conjugate gradient-like algorithms for solving nonsymmetric linear systems*, Math. Comp., 44 (1985), pp. 417–424.

15. ———, *GMRES: a generalized minimal residual algorithm for solving nonsymmetric linear systems*, SIAM J. Sci. Statist. Comput., 7 (1986), pp. 856–869.

16. H. A. Van der Vorst, *Bi-CGSTAB: A fast and smoothly converging variant of Bi-CG for the solution of nonsymmetric linear systems*, SIAM J. Sci. Stat. Comput., 13 (1992), pp. 631–644.

17. H. A. Van der Vorst, *Lecture notes on iterative methods*, Tech. Rep., CERFACS, Toulouse, France, 1992.

18. P. K. W. Vinsome, *ORTHOMIN: an iterative method for solving sparse sets of simultaneous linear equations*, in Proc. Fourth Symposium on Reservoir Simulation, Society of Petroleum Engineers of AIME, 1976, pp. 149–159.

19. H. F. Walker, *Implementation of the GMRES method using Householder transformations*, SIAM J. Sci. Stat. Comput., 9 (1988), pp. 152–163.

20. O. Widlund, *A Lanczos method for a class of nonsymmetric systems of linear equations*, SIAM J. Numer. Anal., 15 (1978), pp. 801–812.

21. J. H. Wilkinson, *The Algebraic Eigenvalue Problem*, Clarendon Press, Oxford, UK, 1965.