



**HAL**  
open science

# Block-Arnoldi and Davidson methods for unsymmetric large eigenvalue problems

Miloud Sadkane

► **To cite this version:**

Miloud Sadkane. Block-Arnoldi and Davidson methods for unsymmetric large eigenvalue problems. [Research Report] RR-1676, INRIA. 1992. inria-00076899

**HAL Id: inria-00076899**

**<https://inria.hal.science/inria-00076899>**

Submitted on 29 May 2006

**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.

# INRIA

UNITÉ DE RECHERCHE  
INRIA-RENNES

Institut National  
de Recherche  
en Informatique  
et en Automatique

Domaine de Voluceau  
Rocquencourt  
BP 105  
78153 Le Chesnay Cedex  
France  
Tél (1) 39.63.55.11

## Rapports de Recherche

1992



25<sup>ème</sup>  
anniversaire  
N° 1676

*Programme 1*

*Architectures parallèles, Bases de données,  
Réseaux et Systèmes distribués*

### **BLOCK-ARNOLDI AND DAVIDSON METHODS FOR UNSYMMETRIC LARGE EIGENVALUE PROBLEMS**

**Miloud SADKANE**

**Mai 1992**



★ R R - 1 6 7 6 ★

## Block-Arnoldi and Davidson methods for unsymmetric large eigenvalue problems.

Miloud SADKANE  
IRISA-INRIA  
Campus de Beaulieu  
35042 Rennes Cedex  
FRANCE

Publication interne n°650 - Avril 1992 - 24 pages  
Programme 6

**Abstract:** We present two methods for computing the leading eigenpairs of large sparse unsymmetric matrices. Namely the block-Arnoldi method and an adaptation of the Davidson method to unsymmetric matrices. We give some theoretical results concerning the convergence of these two methods when restarting is used, and discuss implementation aspects of the two methods on an Alliant FX/80. Finally some results of numerical tests on a variety of matrices including matrices from the Harwell-Boeing test collection, in which we compare these two methods are reported.

**Key words :** eigenclements, sparse matrices, Arnoldi's method, Chebyshev iterations, Davidson's method.

### Méthodes d'Arnoldi par blocs et de Davidson pour le calcul d'éléments propres de matrices creuses non symétriques de grande taille.

#### Résumé:

Nous présentons deux techniques itératives qui permettent le calcul de valeurs propres de plus grandes parties réelles de matrices creuses non symétriques de grande taille : la méthode d'Arnoldi par blocs et une adaptation de la méthode de Davidson au cas non symétrique. Nous étudions la convergence de ces deux méthodes lorsqu'un redémarrage périodique est utilisé et mettons en évidence le rôle joué par le défaut de normalité sur la fiabilité de ces méthodes. Nous discutons leur implantation sur Alliant FX/80 et testons leur efficacité sur un ensemble de matrices provenant de la collection Harwell-Boeing.

**Mots clés :** éléments propres, matrices creuses, méthode d'Arnoldi, itérations de Chebyshev, méthode de Davidson.

# Block-Arnoldi and Davidson methods for unsymmetric large eigenvalue problems.

Miloud Sadkane\*

**Abstract:** We present two methods for computing the leading eigenpairs of large sparse unsymmetric matrices. Namely the block-Arnoldi method and an adaptation of the Davidson method to unsymmetric matrices. We give some theoretical results concerning the convergence of these two methods when restarting is used, and discuss implementation aspects of the two methods on an Alliant FX/80. Finally some results of numerical tests on a variety of matrices including matrices from the Harwell-Boeing test collection, in which we compare these two methods are reported.

**Key words :** eigenelements, sparse matrices, Arnoldi's method, Chebyshev iterations, Davidson's method.

## Méthodes d'Arnoldi par blocs et de Davidson pour le calcul d'éléments propres de matrices creuses non symétriques de grande taille.

### Résumé:

Nous présentons deux techniques itératives qui permettent le calcul de valeurs propres de plus grandes parties réelles de matrices creuses non symétriques de grande taille : la méthode d'Arnoldi par blocs et une adaptation de la méthode de Davidson au cas non symétriques. Nous étudions la convergence de ces deux méthodes lorsqu'un redémarrage périodique est utilisé et mettons en évidence le rôle joué par le défaut de normalité sur la fiabilité de ces méthodes. Nous discutons leur implantation sur Alliant FX/80 et testons leur efficacité sur un ensemble varié de matrices Harwell-Boeing.

**Mots clés :** éléments propres, matrices creuses, méthode d'Arnoldi, itérations de Chebyshev, méthode de Davidson.

---

\*IRISA/INRIA. Campus de Beaulieu, 35042 Rennes Cedex, France.

# 1 Introduction

This paper is concerned with the computation of a few eigenvalues, those with either largest real part or largest magnitude, and the corresponding eigenvectors of large, sparse, unsymmetric matrices. Two methods are discussed both from the theoretical and the practical point of view : a block-Arnoldi method and an adaptation of the Davidson method to the unsymmetric case. The former method and some of its variants have been recently studied [7, 11, 12, 13]. The latter one is very popular in quantum chemistry [4]. The symmetric case has been intensively studied in [2, 9]. We give in Section 2 and 3 some theoretical results concerning the convergence of these two methods. Numerical experiments in which we compare them are described in Section 4.

Throughout this paper, the symbol *MGS* denotes the modified Gram-Schmidt algorithm,  $\| \cdot \|_F$  and  $\| \cdot \|_2$  denote the Frobenius and the Euclidean norm respectively,  $X^H$  and  $Y^T$  denote the transpose of the complex conjugate of a matrix  $X$  and the transpose of a real matrix  $Y$  respectively.

## 2 The block-Arnoldi method

To compute the  $l$  eigenvalues with largest real parts or largest magnitude of a real large sparse unsymmetric matrix  $A \in \mathbf{R}^{n,n}$ , let  $U_1 \in \mathbf{R}^{n,l}$  be a rectangular matrix having at least  $l$  orthonormal columns. If  $m$  is some fixed integer which limits the dimension of the constructed basis, the block-Arnoldi method can be described as follows:

### 2.1 Algorithm 1

```
for  $j = 1, \dots, m - 1$  do
1.  $W_j = AU_j$ ;
2. for  $i = 1, \dots, j$  do
    $H_{i,j} = U_i^T W_j$ 
    $W_j = W_j - U_i H_{i,j}$ ;
end for
3.  $Q_j R_j = W_j$  QR decomposition;
4.  $U_{j+1} = Q_j$  ;  $H_{j+1,j} = R_j$  ;
end for
```

Since the algorithm involves implicitly a Gram-Schmidt process, the resulting block vectors  $V_m = \{U_1, \dots, U_m\}$  have their columns mutually orthogonal provided none of the

upper triangular matrices  $R_j$  are rank-deficient. Furthermore, the columns of  $V_m$  are an orthonormal basis of the generalized Krylov subspace  $K_m = \{U_1, AU_1, \dots, A^{m-1}U_1\}$ . The restriction of the matrix  $A$  to  $K_m$  has a block-Hessenberg form:

$$H_m = V_m^H A V_m = \begin{pmatrix} H_{1,1} & H_{1,2} & \dots & H_{1,m} \\ H_{2,1} & H_{2,2} & & H_{2,m} \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & & \\ 0 & \dots & 0 & H_{m,m-1} & H_{m,m} \end{pmatrix} \quad (1)$$

where the submatrices  $H_{i,k}$  ( $1 \leq i \leq k \leq m$ ) are dense and  $H_{i,i-1}$  ( $2 \leq i \leq m$ ) are dense upper triangular. Each matrix is of order  $l$ . Therefore, the eigenvalues of  $H_m$  constitute the Galerkin approximations of the eigenvalues of  $A$  on the Krylov subspace  $K_m$ .

From the above algorithm, steps 2,3 and 4 give :

$$AU_k = \sum_{i=1}^k U_i H_{i,k} + U_{k+1} H_{k+1,k} \quad k = 1, \dots, m \quad (2)$$

which can easily be condensed in the form

$$AV_m = V_m H_m + [0, \dots, 0, U_{m+1} H_{m+1,m}]. \quad (3)$$

If  $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_{ml})$  denotes the diagonal matrix of eigenvalues of  $H_m$  corresponding to the eigenvectors  $Y_{m,l} = [y_1, \dots, y_{ml}]$ , then relation (3) gives

$$AV_m Y_{m,l} - V_m H_m Y_{m,l} = [0, \dots, 0, U_{m+1} H_{m+1,m}] Y_{m,l}. \quad (4)$$

Let us denote by  $X_m = V_m Y_{m,l}$  the matrix of approximate eigenvectors of  $A$  and by  $\bar{Y}_{m,l}$  the last  $l$  blocks of  $Y_{m,l}$ , that is

$$Y_{m,l} = \begin{pmatrix} * \\ \vdots \\ \bar{Y}_{m,l} \end{pmatrix}. \quad (5)$$

Then equation (4) gives

$$\| AX_m - X_m \Lambda \|_2 = \| H_{m+1,m} \bar{Y}_{m,l} \|_2. \quad (6)$$

Relation (6) is very useful for obtaining the residual norms without having to compute them explicitly.

Convergence of this algorithm can be very slow and the slowness is exacerbated by an unfavourable distribution of the spectrum of  $A$ . To ensure convergence, the dimension of the Krylov subspace must be large, which increases the cost and the storage. One way to overcome this problem is to restart the algorithm periodically. This gives rise to the following algorithm.

## 2.2 Algorithm 2

Choose an orthonormal matrix  $U_1 \in \mathbf{R}^{n,l}$  and an integer  $m$

$Z_1 = [U_1]$

for  $k = 1, 2, \dots$  do

1. Starting with  $Z_k$ , construct  $H_m$  and  $V_m$  from Algorithm 1

2. Compute the  $l$  desired eigenpairs  $\Theta_l = \text{diag}(\lambda_1, \dots, \lambda_l)$

and  $Y_{m,l} = [y_1, \dots, y_l]$  of  $H_m = V_m^H A V_m$ ;

3. if the quantity in (6) is less than some prespecified threshold, then exit

else  $Z_{k+1} = MGS(V_m Y_{m,l})$

endif

end for

Convergence of the above algorithm can be considerably improved by preprocessing  $Z_{k+1}$  by Chebyshev iteration before each restart [11, 12].

From a practical point of view, it is always possible and usually desirable to avoid complex arithmetic by computing separately the real and imaginary parts of eigenvectors of  $H_m$  corresponding to complex eigenvalues. In the following theoretical results, we will not make this assumption and so everything can be treated in complex arithmetic. In the following, the eigenvalues of the matrix  $A$  are denoted by  $\mu_1, \mu_2, \dots, \mu_n$  labelled in decreasing order of their real parts (resp. of their magnitudes) and  $\sigma_1(A) \geq \sigma_2(A) \geq \dots \geq \sigma_n(A)$  denote the singular values of  $A$ .  $\lambda_{\max}(B)$  stands for the largest eigenvalue of a Hermitian (or symmetric) matrix  $B$ .  $\sigma_{\max}(B)$  and  $\sigma_{\min}(B)$  stand for the largest and the smallest singular value of the matrix  $B$  respectively.

Let  $\lambda_k$  be any eigenvalue of the matrix  $H_m$  produced by Algorithm 2, then we have the following

**Proposition 2.1** *The sequences  $\zeta_k = (\text{Re}(\lambda_k))_{k \in \mathbf{N}}$  and  $\xi_k = (|\lambda_k|)_{k \in \mathbf{N}}$  are bounded by  $\lambda_{\max}(\frac{A+A^T}{2})$  and  $\sigma_1(A)$  respectively.*

**Proof**

$$\begin{aligned}
 \zeta_k &\leq \max_{z \neq 0} \text{Re} \frac{z^H H_m z}{z^T z} \\
 &= \max_{z \neq 0} \frac{1}{z^H z} \frac{1}{2} (z^H H_m z + z^H H_m^H z) \\
 &= \max_{z \neq 0} \frac{1}{z^H z} \frac{1}{2} (z^H V_m^H (A + A^T) V_m z) \\
 &= \max_{z \neq 0} \frac{1}{z^H V_m^H V_m z} \frac{1}{2} (z^H V_m^H (A + A^T) V_m z)
 \end{aligned}$$

$$\begin{aligned}
&\leq \max_{x \neq 0} \frac{1}{x^H x} \frac{1}{2} (x^T (A + A^T) x) \\
&= \lambda_{\max} \left( \frac{A + A^T}{2} \right)
\end{aligned}$$

and

$$\begin{aligned}
\xi_k &\leq \rho(V_m^T A V_m) \\
&\leq \|V_m^T A V_m\|_2 \\
&\leq \|A\|_2 = \sigma_1
\end{aligned}$$

□

**Proposition 2.2** *At each restart  $k = 2, 3, \dots$  the matrix  $H_m$  constructed by Algorithm 2 has the following pattern*

$$H_m = \begin{pmatrix} T_{k,l} & H_1 \\ H_2 & H_3 \end{pmatrix}$$

where the matrix  $T_{k,l} = \begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ \mathbf{0} & & \lambda_l \end{pmatrix}$  is upper triangular and  $\lambda_1, \dots, \lambda_l$  are the eigen-

values obtained at the  $k - 1$  st restart.  $H_2 = \begin{pmatrix} \bar{H}_2 \\ \mathbf{0} \end{pmatrix}$  where  $\bar{H}_2$  is an upper triangular matrix of order  $l$ ,  $H_1$  is a rectangular matrix and  $H_3$  is a block-Hessenberg matrix.

Moreover

$$\|H_1\|_F \leq \sqrt{\sum_{i=1}^l (\sigma_i^2(H_m) - |\lambda_i|^2)} \quad (7)$$

and

$$\|\bar{H}_2\|_F = \|AZ_k - Z_k T_{k,l}\|_F \quad (8)$$

$$\leq \sqrt{\sum_{i=1}^l (\sigma_i^2(H_m) - |\lambda_i|^2)} \quad (9)$$

**Proof** It is sufficient to prove this proposition only for the case  $k = 2$ . By construction, the matrix  $H_m$  and hence the matrix  $H_3$  have Hessenberg form. After the first restart  $Z_2 = MGS(V_m Y_{m,l})$  is used as the initial guess. According to the MGS algorithm,  $V_m Y_{m,l} = Z_2 R_2$ , where  $R_2$  is a square and upper triangular matrix and  $Z_2$  is an orthogonal matrix. Thus  $T_{2,l} \equiv Z_2^H A Z_2 = R_2^{-H} Y_{m,l}^H V_m^H A V_m Y_{m,l} R_2^{-1} = R_2^{-H} Y_{m,l}^H Y_{m,l} \Theta_l R_2^{-1}$ . Since  $R_2^H R_2 = Y_{m,l}^H Y_{m,l}$ , we conclude that  $T_{2,l} = R_2 \Theta_l R_2^{-1}$  which is an upper triangular matrix with  $\lambda_1, \dots, \lambda_l$  on its diagonal.



From step 3 and 4 of Algorithm 1 and from (1),  $\overline{H}_2$  is upper triangular and  $\|\overline{H}_2\|_F = \|AZ_2 - Z_2Z_2^HAZ_2\|_F = \|AZ_2 - Z_2T_{2,l}\|_F$ .

Now, Since  $\sigma_i^2(H_m)$  is the  $i$ th largest eigenvalue of the matrix  $H_m^HH_m$ , it follows from the Cauchy interlace theorem [10] that

$$\begin{aligned} \sum_{i=1}^l \sigma_i^2(H_m) &\geq \sum_{i=1}^l \lambda_i (T_{k,l}^HT_{k,l} + H_2^HH_2) \\ &= \text{Trace}(T_{k,l}^HT_{k,l}) + \|H_2\|_F^2 \\ &\geq \sum_{i=1}^l |\lambda_i|^2 + \|H_2\|_F^2. \end{aligned}$$

□

**Corollary 2.1** *If we assume that the matrix  $A$  is nearly normal in the sense that*

$$|\sigma_i - |\mu_i|| = \epsilon\sigma_i \text{ for } i = 1, 2, \dots, n \text{ with } \epsilon \text{ small,}$$

*then the Frobenius norms of  $H_1$  and  $\overline{H}_2$  are bounded by*

$$\left[ \epsilon \sum_{i=1}^l \sigma_i(\sigma_i + |\lambda_i|) + 2\sigma_1 \sum_{i=1}^l |\mu_i - \lambda_i| \right]^{\frac{1}{2}}$$

**Proof** Since for  $i = 1, \dots, l$   $\sigma_i(H_m)$  is the  $i$ th largest singular value of  $H_m$

$$\begin{aligned} \sigma_i(H_m) &= \max_{S \subset \mathbb{R}^{ml}, \dim S=i} \min_{x \in S} \frac{\|H_m x\|_2}{\|x\|_2} \\ &= \max_{S \in \mathbb{R}^{ml}, \dim S=i} \min_{x \in S} \frac{\|V_m^H A V_m x\|_2}{\|V_m x\|_2} \\ &\leq \max_{S \in \mathbb{R}^n, \dim S=i} \min_{z \in S} \frac{\|Az\|_2}{\|z\|_2} \\ &= \sigma_i. \end{aligned}$$

Using the results of proposition 2.2, we see that the square of the Frobenius norms of  $H_1$  and  $\overline{H}_2$  are bounded by

$$\begin{aligned} &\sum_{i=1}^l (\sigma_i^2 - |\lambda_i|^2) \\ &= \sum_{i=1}^l (\sigma_i - |\mu_i|)(\sigma_i + |\lambda_i|) + \sum_{i=1}^l (|\mu_i| - |\lambda_i|)(\sigma_i + |\lambda_i|) \\ &\leq \epsilon \sum_{i=1}^l \sigma_i(\sigma_i + |\lambda_i|) + 2\sigma_1 \sum_{i=1}^l |\mu_i - \lambda_i| \end{aligned}$$

□

As soon as the Ritz values  $\lambda_i$ ,  $i = 1, 2, \dots, l$  start to approximate the eigenvalues  $\mu_i$ ,  $i = 1, 2, \dots, l$ ,  $\|H_1\|_F$  and  $\|\overline{H}_2\|_F$  become very small according to the above inequality and  $AZ_k - Z_kT_{k,l}$  goes to 0, which means that  $Z_k$  tends toward an invariant

subspace of  $A$ . The matrix  $H_m$  tends toward  $H = \begin{pmatrix} \mu_1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \mu_l & \\ & & & H_3 \end{pmatrix}$  where  $H_3$  is

a block Hessenberg matrix. This establishes the convergence of Arnoldi's method in the case of nearly normal matrices.

### Remark

It should be noted that even if  $A$  is nearly normal (or even normal), the matrix  $H_m$  is not.

Experiments have shown that for general matrices, after several restarts the matrix  $H_m$  built by Algorithm 2 "converges" toward a matrix of the form  $\begin{pmatrix} T_l & H_1 \\ \mathbf{0} & H_2 \end{pmatrix}$  where

$T_l = \begin{pmatrix} \mu_1 & & * \\ & \ddots & \\ \mathbf{0} & & \mu_l \end{pmatrix}$  is triangular with  $\mu_1, \dots, \mu_l$ , the sought eigenvalues, on its diagonal.

$H_1$  and  $H_2$  are respectively rectangular and block Hessenberg matrices. This can be viewed as an "incomplete" Schur decomposition of the matrix  $A$ . This last assertion is not obvious, but it can be proved.

## 3 The unsymmetric Davidson method

Davidson's method for symmetric matrices was developed for problems in quantum chemistry [4] as an efficient way of computing the lowest energy levels and the corresponding wave functions of the Schrödinger operator. The matrix dealt with is symmetric and strongly diagonally dominant. It can be viewed as a preconditioned version of the Lanczos method [2]. In the following algorithm, we will generalize the Davidson method to the unsymmetric case.

### 3.1 Algorithm 3

The following algorithm computes the  $l$  eigenpairs of large real parts or large magnitude of the matrix  $A$ ;  $m$  is a given integer which limits the dimension of the constructed basis.  $C_{k,i}$   $i = 1, 2, \dots, l$  are preconditioning matrices.

Choose an initial orthonormal matrix  $V_1 := [v_1, \dots, v_l] \in \mathbf{R}^{n \times l}$ ;

**for**  $k = 1, 2, \dots$  **do**

1.  $W_k := AV_k$ ;
2.  $H_k := V_k^H W_k$ ;
3. Compute the  $l$  desired eigenpairs  $(\lambda_{k,i}, y_{k,i})_{1 \leq i \leq l}$  of  $H_k$ ;
4.  $x_{k,i} := V_k y_{k,i}$ , for  $i = 1, \dots, l$ ;
5.  $r_{k,i} := W_k y_{k,i} - \lambda_{k,i} x_{k,i}$ , for  $i = 1, \dots, l$ ;  
**if** convergence **then** exit;
6.  $t_{k,i} := C_{k,i} r_{k,i}$ , for  $i = 1, \dots, l$ ;
7. **if**  $\dim(V_k) \leq m$   
**then**  $V_{k+1} := MGS(V_k, t_{k,1}, \dots, t_{k,l})$ ;  
**else**  $V_{k+1} := MGS(x_{k,1}, \dots, x_{k,l}, t_{k,1}, \dots, t_{k,l})$ ;  
**end if**

**end for**

All the steps are the same as in the symmetric Davidson method, except that complex arithmetic may occur. Fortunately, this can easily be avoided by dealing with the real and imaginary parts separately in steps 3, 4, 5 and in some cases in step 6. At each iteration  $k$ , the basis  $V_{k+1}$  is obtained from  $V_k$  by incorporating the vectors  $t_{k,i} = C_{k,i} r_{k,i}$ ,  $i = 1, \dots, l$  after orthonormalization. The  $C_{k,i}$  are often of the form  $(M - \lambda_{k,i} I)^{-1}$  where  $M$  is an approximation to  $A$  and  $\lambda_{k,i}$   $i = 1, \dots, l$  are the desired eigenvalues computed at step 3 of the algorithm. The subspace  $\text{span}\{V_k\}$  generated by the columns of  $V_k$  is not a Krylov subspace. If  $M \neq D$  the main diagonal of  $A$ , then a new system must be solved at each iteration. The hope is to reach convergence very quickly with a small value of  $m$ , thus rewarding the extra cost. The choice of a good preconditioner in step 6 is crucial for convergence and has been already studied in the symmetric case [2, 9]. Here, similar conclusions can be derived, namely that the matrix  $M$  must be an approximation to  $A$  in the sense that

$$(M - \mu_k I)^{-1}(A - \mu_k I)\Big|_{\{E_{\mu_k}\}^c} \approx I \quad (10)$$

where  $\{E_{\mu_k}\}^c$  stands for the complement of the set of eigenvectors associated with the eigenvalue  $\mu_k$ . Of course the matrix  $A$  is supposed to be diagonalizable.

With such a matrix, we have from step 6, where the subscript  $i$  is dropped

$$x_k - t_k = x_k - (M - \mu_k I)^{-1}(A - \mu_k I)x_k. \quad (11)$$

If we write  $x_k = \alpha_k x + \beta_k y_k$  where  $Ax = \mu_k x$  and  $y_k \in \{E_{\mu_k}\}^c$ , we have

$$x_k - t_k = \alpha_k \left( I - (\mu_k - \lambda_k)(M - \lambda_k I)^{-1} \right) x + \beta_k \left( I - (M - \mu_k I)^{-1}(A - \mu_k I) \right) y_k \quad (12)$$

Taking into account the condition (10), expression (12) shows that when  $(\lambda_k, x_k)$  starts to approximate  $(\mu_k, x)$ , then, provided that  $(M - \mu_k I)^{-1}$  is bounded, the vector  $x_k - t_k$  has small components except those in the direction of  $x$ , and therefore constitutes an improvement over  $x_k$ .

The main difference between this unsymmetric Davidson's version and the block-Arnoldi method rests in Step 6 of the above algorithm. If no preconditioner is used (i.e.  $C_{k,i} = I$  or  $\gamma I$ ,  $\gamma \neq 0$  for  $i = 1, \dots, l$ ), then the block-Arnoldi method is recovered. Of course this version of the block-Arnoldi method is not recommended because of the unnecessary computational effort at each step. There are some others obvious common points between Arnoldi's method and the unsymmetric Davidson method. First, they both use matrix by block vector multiplication at each step without altering the matrix  $A$  and gradually build an interaction matrix using different projection techniques. Arnoldi's method builds a Hessenberg matrix in a clever way whereas the matrix built by Davidson's method is full, but seems, in some cases, more rich in information about the required eigenvalues. Finally, we mention that both of the methods allow only the computation of right eigenvectors.

In proposition 3.2, we characterize the matrices  $(H_k)_{k \in \mathbf{N}}$  built by Algorithm 3. But first we need the following simple proposition.

**Proposition 3.1** *Let  $B \in \mathbf{C}^{n,n}$  be a Hermitian matrix and  $Q \in \mathbf{C}^{n,k}$  with  $n \geq k$  a rectangular matrix. Then*

$$\lambda_{\max}(Q^H B Q) \leq \lambda_{\max}(B) \sigma_{\max}^2(Q) \quad (13)$$

Moreover if  $\text{rank}(Q) = k$ , then

$$\lambda_{\max} \left( Q(Q^H Q)^{-1} Q^H \right) = 1 \quad (14)$$

**Proof** From the Courant Fisher theorem [10], we have

$$\lambda_{\max}(Q^H B Q) = \max_{x \in \mathbf{C}^k; x \neq 0} \frac{x^H Q^H B Q x}{x^H x}$$

$$\begin{aligned}
&= \max_{x \in \mathbf{C}^k; x \neq 0} \frac{(Qx)^H B(Qx)}{(Qx)^H(Qx)} \frac{x^H Q^H Q x}{x^H x} \\
&\leq \max_{x \in \mathbf{C}^n; x \neq 0} \frac{x^H B x}{x^H x} \max_{x \in \mathbf{C}^k; x \neq 0} \frac{x^H Q^H Q x}{x^H x} \\
&= \lambda_{\max}(B) \lambda_{\max}(Q^H Q).
\end{aligned}$$

To prove (14), let  $Q = U\Sigma V^H$  be the singular value decomposition of  $Q$ , with  $U \in \mathbf{C}^{n,n}$  and  $V \in \mathbf{C}^{k,k}$  orthonormal and  $\Sigma = \begin{pmatrix} I_k \\ \mathbf{0} \end{pmatrix} \Sigma_k \in \mathbf{C}^{n,k}$  where  $I_k$  stands for the identity matrix of order  $k$  and  $\Sigma_k = \text{diag}(\nu_1, \dots, \nu_k)$  are the singular values of  $Q$ . Then  $Q(Q^H Q)^{-1}Q^H = U \begin{pmatrix} I_k & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} U^H$ .  $\square$

For the following proposition, we will make use of the notational conventions  $R_k = [r_{k,1}, \dots, r_{k,l}]$  the residual matrix built in step 5 of Algorithm 3. Step 6 will be denoted by  $T_k = [t_{k,1}, \dots, t_{k,l}]$ .

**Proposition 3.2** *At each restart  $k = 2, 3, \dots$  the matrix  $H_k$  constructed by Algorithm 3 has the following pattern*

$$H_k = \begin{pmatrix} H_{k-1} & H_{2,k} \\ H_{3,k} & H_{4,k} \end{pmatrix}$$

where the matrix  $H_{k-1} = \begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ \mathbf{0} & & \lambda_l \end{pmatrix}$  is upper triangular and  $\lambda_1, \dots, \lambda_l$  are the eigenvalues obtained at the  $k-1$  st restart.  $H_{2,k}$ ,  $H_{3,k}$  and  $H_{4,k}$  are dense matrices with appropriate dimensions. Moreover the Frobenius norms of  $H_{2,k}$ , and  $H_{3,k}$  are bounded by

$$\sqrt{\sum_{i=1}^l (\sigma_i^2(H_k) - |\lambda_i|^2)} \quad (15)$$

and

$$\|H_{3,k}\|_2 \leq \|AV_{k-1} - V_{k-1}H_{k-1}\|_2. \quad (16)$$

If  $Y_{k-1}$  stands for the matrix of eigenvectors of  $H_{k-1}$ , then

$$\|H_{3,k}Y_{k-1}\|_2 \leq \frac{\sigma_{\max}(R_{k-1}^H T_{k-1})}{\sigma_{\min}((I - V_{k-1}V_{k-1}^H)T_{k-1})} \quad (17)$$

and

$$\|H_{3,k}Y_{k-1}\|_2 \leq \sigma_{\max}(R_{k-1}). \quad (18)$$

**Proof** From Algorithm 3, it is clear that the matrices  $H_{2,k}$ ,  $H_{3,k}$  and  $H_{4,k}$  are dense. (15) was shown in proposition 2.2 in the context of Arnoldi's method. The result is still valid here because we did not make use of the fact that the subspace  $V_m$  is a Krylov subspace in the proof of proposition 2.2.

Let us prove the bounds for  $H_{3,k}$ . First, notice that because of the *MGS* algorithm, there is no loss of generality in assuming that the columns of the orthogonal projection of  $T_{k-1}$  on the orthogonal complement of  $V_{k-1}$ , i.e.  $(I - V_{k-1}V_{k-1}^H)T_{k-1}$ , are linearly independent before convergence. If we let  $V_k = [V_{k-1}, U_k]$ , where  $U_k$  is an orthogonal matrix whose columns are those computed at step  $k$  which are mutually orthonormal and orthogonal to  $V_{k-1}$ , then  $H_{3,k} = U_k^H AV_{k-1}$ . Now, if  $\mathcal{Q}_k \mathcal{R}_k$  stands for the *QR* factorization of  $(I - V_{k-1}V_{k-1}^T)T_{k-1}$ , then

$$U_k \equiv \mathcal{Q}_k = (I - V_{k-1}V_{k-1}^H)T_{k-1}\mathcal{R}_k^{-1} \quad (19)$$

and

$$H_{3,k} = \mathcal{R}_k^{-H} T_{k-1}^H (I - V_{k-1}V_{k-1}^H) AV_{k-1}. \quad (20)$$

Hence

$$H_{3,k}^H H_{3,k} = V_{k-1}^H A^T (I - V_{k-1}V_{k-1}^T) T_{k-1} (\mathcal{R}_k^H \mathcal{R}_k)^{-1} T_{k-1}^H (I - V_{k-1}V_{k-1}^T) AV_{k-1}.$$

Since  $\mathcal{R}_k^H \mathcal{R}_k = T_{k-1}^H (I - V_{k-1}V_{k-1}^H)^2 T_{k-1}$  and  $(I - V_{k-1}V_{k-1}^H)^2 = I - V_{k-1}V_{k-1}^H$ , we have

$$\begin{aligned} H_{3,k}^H H_{3,k} &= \left[ (I - V_{k-1}V_{k-1}^H) AV_{k-1} \right]^H \left[ (I - V_{k-1}V_{k-1}^H) T_{k-1} \right] \\ &\quad \left[ T_{k-1}^H (I - V_{k-1}V_{k-1}^H)^2 T_{k-1} \right]^{-1} \left[ (I - V_{k-1}V_{k-1}^H) T_{k-1} \right]^H \\ &\quad \left[ (I - V_{k-1}V_{k-1}^H) AV_{k-1} \right]. \end{aligned}$$

If we let  $\mathcal{B}_{k-1} = (I - V_{k-1}V_{k-1}^H)T_{k-1}$  and  $\mathcal{D}_{k-1} = (I - V_{k-1}V_{k-1}^H)AV_{k-1}$ , then

$$H_{3,k}^H H_{3,k} = \mathcal{D}_{k-1}^H \mathcal{B}_{k-1} (\mathcal{B}_{k-1}^H \mathcal{B}_{k-1})^{-1} \mathcal{B}_{k-1}^H \mathcal{D}_{k-1}$$

And by proposition 3.1 we have

$$\begin{aligned} \| H_{3,k} \|_2^2 &= \lambda_{\max}(H_{3,k}^H H_{3,k}) \leq \sigma_{\max}^2(\mathcal{D}_{k-1}) \\ &= \sigma_{\max}^2((I - V_{k-1}V_{k-1}^T)AV_{k-1}) \\ &= \| (I - V_{k-1}V_{k-1}^H)AV_{k-1} \|_2^2 \\ &= \| AV_{k-1} - V_{k-1}H_{k-1} \|_2^2. \end{aligned}$$

On the other hand

$$\begin{aligned} H_{3,k}Y_{k-1} &= \mathcal{R}_k^{-H}T_{k-1}^H(I - V_{k-1}V_{k-1}^H)\Lambda V_{k-1}Y_{k-1} \\ &= \mathcal{R}_k^{-H}T_{k-1}^HR_{k-1}. \end{aligned}$$

Hence

$$\begin{aligned} (H_{3,k}Y_{k-1})^H(H_{3,k}Y_{k-1}) &= R_{k-1}^HT_{k-1}(\mathcal{R}_k^H\mathcal{R}_k)^{-1}T_{k-1}^HR_{k-1} \\ &= R_{k-1}^HT_{k-1}\left[T_{k-1}^H(I - V_{k-1}V_{k-1}^H)^2T_{k-1}\right]^{-1}T_{k-1}^HR_{k-1}. \end{aligned}$$

Since  $V_{k-1}^HR_{k-1} = 0$ , we have

$$\begin{aligned} (H_{3,k}Y_{k-1})^H(H_{3,k}Y_{k-1}) &= R_{k-1}^H\left((I - V_{k-1}V_{k-1}^H)T_{k-1}\right)\left[T_{k-1}^H(I - V_{k-1}V_{k-1}^H)^2T_{k-1}\right]^{-1} \\ &\quad \left((I - V_{k-1}V_{k-1}^H)T_{k-1}\right)^HR_{k-1} \end{aligned}$$

or

$$(H_{3,k}Y_{k-1})^H(H_{3,k}Y_{k-1}) = R_{k-1}^H\mathcal{B}_{k-1}(\mathcal{B}_{k-1}^H\mathcal{B}_{k-1})^{-1}\mathcal{B}_{k-1}^HR_{k-1} \quad (21)$$

(18) is then a simple consequence of proposition 3.1 applied to (21). Now observe that  $V_{k-1}^TR_{k-1} = 0$ , this gives

$$\begin{aligned} (H_{3,k}Y_{k-1})^H(H_{3,k}Y_{k-1}) &= R_{k-1}^HT_{k-1}\left[T_{k-1}^H(I - V_{k-1}V_{k-1}^H)T_{k-1}\right]^{-1} \\ &\quad T_{k-1}^HR_{k-1} \end{aligned}$$

and proposition 3.1 yields

$$\begin{aligned} \|H_{3,k}Y_{k-1}\|_2^2 &\leq \lambda_{\max}\left[T_{k-1}^H(I - V_{k-1}V_{k-1}^H)T_{k-1}\right]^{-1}\sigma_{\max}^2(R_{k-1}^HT_{k-1}) \\ &= \left(\frac{\sigma_{\max}(R_{k-1}^HT_{k-1})}{\sigma_{\min}((I - V_{k-1}V_{k-1}^H)T_{k-1})}\right)^2 \end{aligned}$$

which proves (17).  $\square$

### Remark

1. If we denote by  $X_k = [x_{k,1}, \dots, x_{k,l}]$  the matrix of approximate eigenvectors built at step 4 of Algorithm 3, then relation (17) can be written in the form

$$\begin{aligned} \|H_{3,k}Y_{k-1}\|_2 &\leq \frac{\sigma_{\max}\left(R_{k-1}^H(I - V_{k-1}V_{k-1}^H)(X_{k-1} - T_{k-1})\right)}{\sigma_{\min}\left((I - V_{k-1}V_{k-1}^H)T_{k-1}\right)} \\ &\leq \frac{\sigma_{\max}(R_{k-1})\sigma_{\max}\left((I - V_{k-1}V_{k-1}^H)(X_{k-1} - T_{k-1})\right)}{\sigma_{\min}\left((I - V_{k-1}V_{k-1}^H)T_{k-1}\right)}. \end{aligned}$$

This means that  $\|H_{3,k}\|_2$  behaves like  $\sigma_{max}(R_{k-1})\sigma_{max}((I - V_{k-1}V_{k-1}^H)(X_{k-1} - T_{k-1}))$ . The term  $\sigma_{min}((I - V_{k-1}V_{k-1}^H)T_{k-1})$  in the denominator is just a scaling factor. We have seen in (12) that the columns of  $X_{k-1} - T_{k-1}$  have small components except those in the direction of the wanted eigenvectors, therefore  $\sigma_{max}((I - V_{k-1}V_{k-1}^H)(X_{k-1} - T_{k-1}))$  tends towards zero and the matrix  $H_k$  tends towards the form announced in the remark of paragraph 2.2

2. The same results as given in corollary 2.1 for the case of nearly normal matrices hold for Algorithm 3.

3. Following the proof of proposition 3.2, if the  $C_{k,i}$  are proportional to the identity matrix, we have from (19),  $U_k = R_{k-1}\mathcal{R}_k^{-1}$ , where  $\mathcal{R}_k$  stands for the upper triangular matrix derived from the  $QR$  factorization of the residual matrix  $R_{k-1} = AX_{k-1} - X_{k-1}\Lambda_{k,l}$  where  $X_{k-1} = V_{k-1}Y_{k-1}$  and  $\Lambda_{k,l} = \text{diag}(\lambda_{k,1}, \dots, \lambda_{k,l})$ . Then

$$\begin{aligned} U_k^H AX_{k-1} &= \mathcal{R}_k^{-H} R_{k-1}^H AX_{k-1} \\ &= \mathcal{R}_k^{-H} R_{k-1}^H R_{k-1} \quad \text{since } R_{k-1}^H X_{k-1} = 0 \\ &= \mathcal{R}_k^{-H} \mathcal{R}_k^H \mathcal{R}_k \quad \text{since } R_{k-1}^H R_{k-1} = \mathcal{R}_k^H \mathcal{R}_k \\ &= \mathcal{R}_k \end{aligned}$$

If we assume moreover that the matrix  $Y_{k-1}$  of eigenvectors of the triangular matrix  $H_{k-1}$  is of full rank, then

$$H_{3,k} = U_k^H AV_{k-1} = U_k^H AX_{k-1} Y_{k-1}^{-1} = \mathcal{R}_k Y_{k-1}^{-1}$$

and, under the assumption that the eigenvalues of the upper triangular matrix  $H_{k-1}$  are all distinct, the matrix  $Y_{k-1}$  is upper triangular as is the matrix  $H_{3,k}$ . This shows that the submatrices  $(H_k)_{i,i-1}$   $i = 2, 3, \dots$  are, in this special case, upper triangular.

## 4 Numerical experiments

We now give some experimental results to compare Algorithm 2 and Algorithm 3 discussed above. We have implemented these two algorithms using MATLAB. Algorithm 2 and Algorithm 3 refer to Arnoldi's and Davidson's method respectively. We consider two examples.



**Example 1.** The matrix  $A$  is the Toeplitz matrix of order  $n = 30$  defined by

$$A = \begin{pmatrix} 1 & u_1 & \dots & u_{n-2} & u_{n-1} \\ l_1 & 1 & & & u_{n-2} \\ & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ l_{n-2} & & \ddots & \ddots & u_1 \\ l_{n-1} & l_{n-2} & \dots & l_1 & 1 \end{pmatrix}$$

where  $l_i = i + 1$  and  $u_i = i + \frac{3}{2}$  for  $i = 1, 2, \dots, n - 1$ .

Since the diagonal is constant, we know that Arnoldi and Davidson with diagonal preconditioning are equivalent. Table 1 displays the sequence of residuals corresponding to the ( $l = 1$ ) eigenvalue of largest real part of  $A$  obtained by Arnoldi's method and Davidson's method with tridiagonal preconditioning. In this example Arnoldi's method appears to be superior to Davidson's method.

iter	Arnoldi	Davidson Tridiagonal
1	1.63 E+01	1.67 E+01
2	2.08 E+00	2.71 E+01
3	5.57 E-02	3.65 E+00
4	7.71 E-04	9.60 E-02
5	1.37 E-05	1.40 E-03
6	1.77 E-07	1.76 E-05
7	1.22 E-09	1.98 E-07
8	7.69 E-13	5.50 E-12

Table 1. Sequence of residuals obtained by Arnoldi's method and the unsymmetric Davidson's method with tridiagonal preconditioning.

**Example 2.**  $A = (a_{ij})_{1 \leq i, j \leq 30}$  is a matrix of order  $n = 30$ , nearly band, the main diagonal is such that  $a_{ii} = i$  for  $i = 1, 2, \dots, n$ , the first upper (resp. lower) diagonal is random in  $[0, 1]$ , the second upper (resp. lower) diagonal is random in  $[0, \frac{1}{2}]$ , the third upper (resp. lower) diagonal is random in  $[0, \frac{1}{4}]$ . Moreover  $a(n, 1) = 1$  and  $a(1, n) = 2$ . Table 2 displays the sequence of the maximum of residuals corresponding to the ( $l = 2$ ) two eigenvalues of largest real parts of  $A$  obtained by Arnoldi's method and Davidson's method using diagonal preconditioning, tridiagonal preconditioning, pentadiagonal preconditioning and heptadiagonal preconditioning.

Since we are working in this example with blocks of size 2, at iteration 15, the matrices  $H_k$  constructed by the two methods are of order 30 and are similar to  $A$ . Thus we

obtain the eigensystem after 15 iterations. Arnoldi's method performs very badly on this example. For Davidson's method, the goal of this example is to show the extreme benefit which may be obtained from a good preconditioner. The main difficulty rests in finding preconditioner producing as good results as those with heptadiagonal preconditioning. Unfortunately, in practice, the condition (10) is in general unrealistic. It is impossible to find general purpose preconditioners and each case must be treated in its own way.

iter	Arnoldi	Davidson Diagonal	Davidson Tridiagonal	Davidson Pentadiagonal	Davidson Heptadiagonal
1	8.35 E+00	8.08 E+00	9.37 E+00	7.71 E+00	7.95 E+00
2	6.60 E+00	5.15 E+00	5.65 E+00	4.98 E+00	5.05 E+00
3	4.94 E+00	3.57 E+00	4.87 E+00	4.08 E+00	3.45 E-01
4	4.14 E+00	3.27 E+00	1.74 E+00	1.98 E+00	6.80 E-03
5	3.54 E+00	1.82 E+00	1.72 E+00	1.74 E+00	1.21 E-09
6	2.82 E+00	2.08 E+00	3.32 E+00	9.38 E-01	3.08 E-14
7	1.92 E+00	1.58 E+00	1.20 E+00	1.43 E+00	
8	1.62 E+00	3.21 E+00	1.46 E+00	2.67 E+00	
9	1.33 E+00	3.58 E+00	1.73 E+00	1.49 E+00	
10	1.04 E+00	4.24 E+00	5.00 E-01	4.02 E-01	
11	1.00 E+00	3.60 E+00	9.60 E-03	7.90 E-03	
12	8.53 E-01	1.57 E+00	1.77 E-04	9.33 E-05	
13	7.94 E-01	2.78 E-02	2.00 E-06	5.12 E-07	
14	6.43 E-01	2.59 E-04	2.52 E-08	2.26 E-09	
15	1.14 E-12	3.50 E-14	3.49 E-14	3.90 E-14	

Table 2. Sequence of residuals obtained by Arnoldi's method and the unsymmetric Davidson's method with several kinds of preconditioning.

We have also implemented Algorithm 2 and Algorithm 3 on an Alliant FX/80 using double precision (unit roundoff  $\approx 2.2 \cdot 10^{-16}$ ). Because of the block nature of these algorithms, level 3 BLAS routines were used extensively. The small eigenvalue problems (step 2 in Algorithm 2 and step 3 in Algorithm 3) were solved using the PARA-EISPACK library [8]. The following two examples were used to compare the two algorithms.

**Example 3.** In this example, we compare the results obtained by block-Arnoldi and Davidson's methods with diagonal preconditioning, for computing the  $l = 8$  eigenpairs of largest real parts of the matrices GRE1107, HOR131, LNS3937-S, PORES3-S and PSMIGR1-S from the Harwell-Boeing test collection [5]. These matrices are described in the Table 3 below

Matrix	Order	Number of entries	order of $\ A\ _1$	Description
GRE1107	1107	5664	1	Simulation of computer systems
HOR131	434	4710	0.901	Flow network problem
LNS3937-S	3937	25407	1.42	Fluid flow modelling
PORES3-S	532	3474	2 1.40	Reservoir modelling
PSMIGR1-S	3140	543162	5	Inter-county migration

Table 3. Characteristics of test matrices

The letter -S in the last three matrices means that the original matrices LNS3937, PORES3, and PSMIGR1 which have already been studied in [12], were scaled using techniques based on the Harwell routine MC19 [1, 3], which make the nonzeros near to unity in the sense that the sum of the squares of the logarithms of the nonzeros is minimized. The reason for scaling these matrices is that neither block-Arnoldi nor Davidson methods were capable of computing the wanted eigenvalues of the unscaled matrices. We report hereafter some basic facts that can explain the misconvergence phenomenon of the methods on such matrices. First we notice that the original matrices are badly scaled, and this may lead to numerical problems. For example the magnitudes of the nonzeros vary from  $3.2 \cdot 10^{-7}$  to  $1.9 \cdot 10^{11}$  for LNS3937, from  $4.8 \cdot 10^{-9}$  to 99637 for PORES3 and from 0 to  $2.2 \cdot 10^6$  for PSMIGR1. We also compute, in Table 4, the gap  $gap(\lambda) = \frac{\min_k |\lambda - \lambda_k|}{\max_j |\lambda - \lambda_j|}$  for the eight eigenvalues  $\lambda$  of largest real parts of these matrices. The eigenvalues have been computed by the QR method on CRAY2. These gaps measure the closeness of the wanted eigenvalues and may be considered as indicators of difficulty. Finally an important factor that has to be taken into consideration is the departure from normality. Let  $Q^H A Q = \Lambda + N$  be the Schur decomposition of  $A$ , where  $Q$  is a unitary matrix,  $\Lambda$  is the diagonal matrix whose main diagonal contains the eigenvalues of  $A$ , and  $N$  is a strictly upper triangular matrix. A large factor of  $\|N\|_F$  means that the matrix  $A$  is very far from a normal matrix, which may lead to an ill-conditioned basis of eigenvectors. Since  $\|N\|_F$  is extremely difficult to compute, we can estimate it using Henrici's formula [6]

$$\frac{\|AA^T - A^T A\|_F}{\sqrt{6}\|A\|_F} \leq \|N\|_F \leq \sqrt{\frac{n^3 - n}{12}} \sqrt{\|AA^T - A^T A\|_F} \quad (22)$$

In Table 4, we computed the above bounds for the matrices LNS3937, PORES3, and PSMIGR1. In Figure 1, we plot the spectrum of all test matrices.

Matrix	$\nu(A)$	Bounds for $\ N\ _F$	$\min_{\lambda} gap(\lambda)$	$\max_{\lambda} gap(\lambda)$
LNS3937	$4.4 \cdot 10^{23}$	$1.2 \cdot 10^{11} \leq \ N\ _F \leq 3.1 \cdot 10^{28}$	$4.4 \cdot 10^{-13}$	$1.2 \cdot 10^{-2}$
PORES3	$3.5 \cdot 10^9$	$2.1 \cdot 10^3 \leq \ N\ _F \leq 3.5 \cdot 10^6$	$4.5 \cdot 10^{-7}$	$7.3 \cdot 10^{-6}$
PSMIGR1	$4.0 \cdot 10^{11}$	$1.4 \cdot 10^4 \leq \ N\ _F \leq 1.4 \cdot 10^8$	$5.5 \cdot 10^{-3}$	$3.7 \cdot 10^{-1}$

Table4. Some spectral characteristics of the unscaled matrix s.

$$gap(\lambda) = \frac{\min_k |\lambda - \lambda_k|}{\max_j |\lambda - \lambda_j|}.$$

Although the scaling changes the eigenvalues, our aim remains the comparison of the block-Arnoldi and Davidson methods. We report on Table 5 the results obtained by these two methods on the test matrices. The gap ratio, as defined above, is also given for these matrices on Table 6. We notice that in all these test matrices, Davidson's method gives better results than block-Arnoldi method.

Matrix	Arnoldi			Davidson		
	Mat-vec	Max. res	Time(sec)	Mat-vec	Max. res	Time(sec)
GRE1107	9200	$9.8 \cdot 10^{-9}$	470	3464	$8.5 \cdot 10^{-9}$	214.53
HOR131	360	$9.0 \cdot 10^{-7}$	11.49	264	$6.0 \cdot 10^{-7}$	8.12
LNS3937-S	1040	$8.3 \cdot 10^{-5}$	243.75	216	$7.1 \cdot 10^{-7}$	37.5
PORES3-S	4000	$1.7 \cdot 10^{-5}$	105.15	2992	$9.9 \cdot 10^{-5}$	82.86
PSMIGR1-S	120	$1.0 \cdot 10^{-11}$	34.19	40	$1.0 \cdot 10^{-14}$	13.15

Table 5. Comparing the block-Arnoldi method and the unsymmetric Davidson method. Size of the basis  $m = 40$  ;  $l = 8$ .

Mat-vec : Matrix-vector multiplications , Max. res :  
Maximum of residual norms.

Matrix	$\min_{\lambda} gap(\lambda)$	$\max_{\lambda} gap(\lambda)$
GRE1107	$2.5 \cdot 10^{-3}$	$5.3 \cdot 10^{-3}$
HOR131	$7.5 \cdot 10^{-3}$	$9.8 \cdot 10^{-2}$
LNS3937-S	$5.3 \cdot 10^{-4}$	$5.3 \cdot 10^{-4}$
PORES3-S	$2.0 \cdot 10^{-7}$	$1.6 \cdot 10^{-6}$
PSMIG1-S	$3.1 \cdot 10^{-7}$	$3.0 \cdot 10^{-5}$

Table6. Gap ratios of test matrices.

$$gap(\lambda) = \frac{\min_k |\lambda - \lambda_k|}{\max_j |\lambda - \lambda_j|}.$$

As show the numerical results, the block-Arnoldi method may not be considered as a powerful technique in itself since the convergence can be very slow. It has, however, a large spectrum of applicability and its efficiency is not preconditioner-dependent as Davidson's method. Chebyshev iteration constitutes a very good way for accelerating the block-Arnoldi method. The algorithm then consists of the following main steps. After construction of the matrix  $H_m$  by Algorithm 2, one computes the eigenvalues of  $H_m$  and splits them into two parts : the wanted ones which constitute an approximation of the sought eigenvalues, and one builds an ellipse which contains all the unwanted eigenvalues

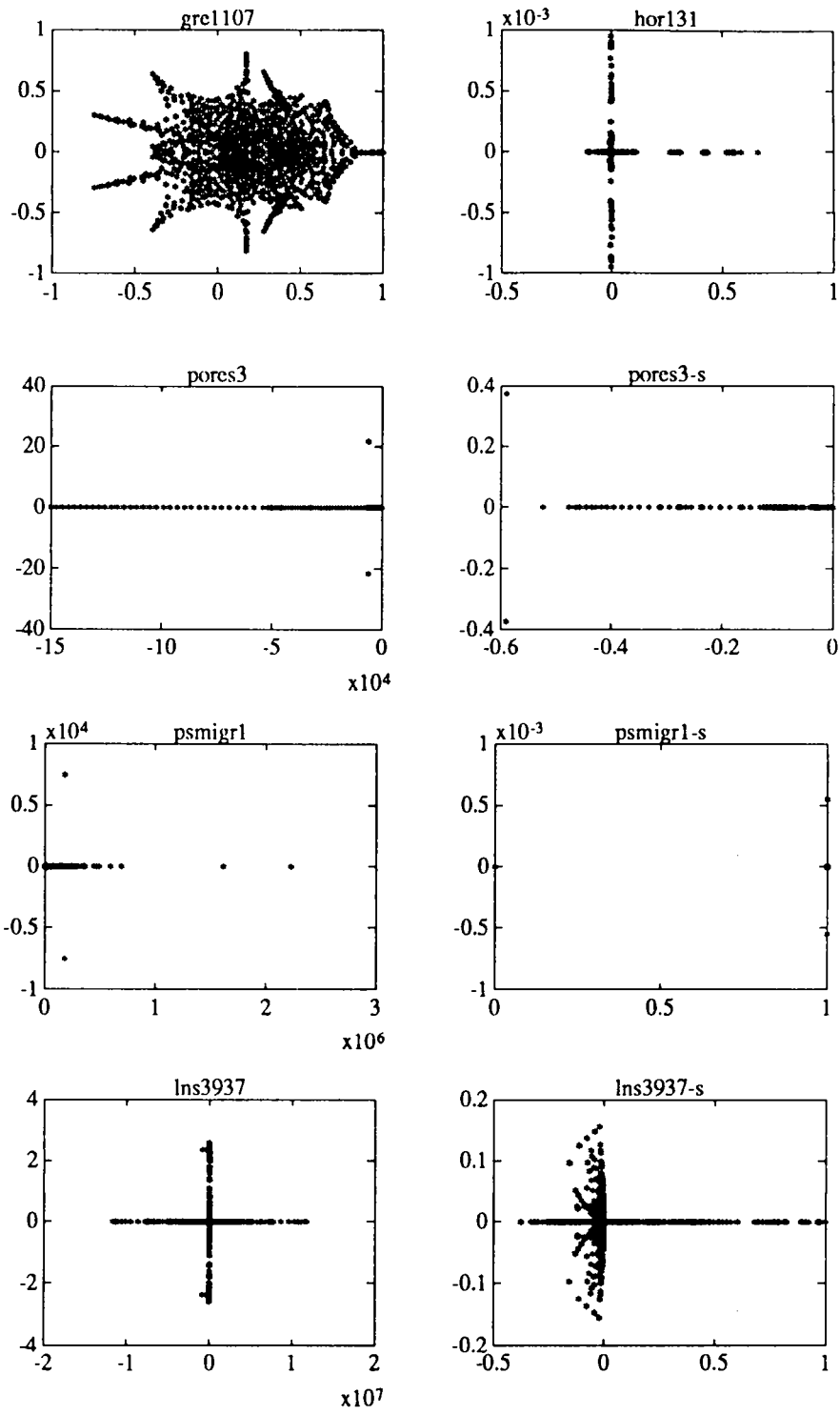


Figure 1: Spectrum of test matrices.

from the rest of the spectrum. Using the parameters of this ellipse, one performs a few Chebyshev iterations aimed at amplifying the components of the restarting matrix  $Z_{k+1}$  in the required directions while removing those in the unwanted ones. For more details on the implementation of this algorithm, see [11, 12]. We call this new modified block-Arnoldi method Arnoldi-Chebyshev. We have considered the matrix *GRE1107* with  $m = 40$  and we have computed the ( $l = 8$ ) eigenpairs of largest real part by the Arnoldi-Chebyshev method using a Chebyshev polynomial of degree 40. We obtained the following results: matrix-vector multiplications = 2920 ; maximum of residual norms = 4.44 E-09 and time=24.90 sec. This constitutes a tremendous improvement over the results given in Table 5. See [12] for other examples.

**Example 4.** In this example, we compare the Davidson method using diagonal preconditioning, the Arnoldi method and the Arnoldi-Chebyshev method with different values of the degree of Chebyshev polynomial. The matrix under consideration is of order 1000 and is generated randomly by setting its density of nonzero elements at 0.01. The nonzero off-diagonal entries are uniformly distributed in the range  $[-1, +1]$ ; the full diagonal entries are in the range  $[0, 20]$ . the matrix has 10763 nonzero elements and is moderately diagonally dominant. Table 7 shows the results obtained when computing the eight eigenpairs of largest real part. Notice the good performance of Davidson's method for this example.

Method	Matrix-vector multiplications	Maximum of residual norms	Time(sec)
Davidson	264	6.79 E-09	9.42
Arnoldi	4467	9.28 E-09	220.53
AC(10)	2352	1.01 E-09	51.78
AC(30)	2064	3.27 E-09	25.19
AC(50)	2288	1.42 E-09	21.47
AC(70)	2480	2.05 E-11	19.54
AC(90)	4656	2.29 E-09	31.49

Table 7. Comparing the unsymmetric Davidson method with Arnoldi and Arnoldi-Chebyshev method. AC(10) means that Arnoldi-Chebyshev is used with a polynomial of degree 10.

## 5 Conclusion

In this paper, we have extended the block Davidson method for solving the symmetric eigenvalue problems to unsymmetric matrices. We have discussed theoretical convergence results and implementation aspects for this adapted version and for the block-Arnoldi method. There are obvious ways to exploit parallelism in the dominant parts of these algorithms. Namely, in the matrix by block-vector multiplications, the matrix-matrix

operations and the eigenvalue problems of full/block Hessenberg matrices. Numerical tests have shown that the principal difficulties with Davidson's method concern the finding of an adequate preconditioner. If a good preconditioner is used, the unsymmetric Davidson method may be quite superior to Arnoldi's method. A typical case is diagonally (resp. block diagonally) dominant matrices where the unsymmetric Davidson method with a diagonal (resp. block diagonal) preconditioner appears to be very suitable and is much superior to the Arnoldi-Chebyshev method.

## References

- [1] P. R. AMESTOY, *Factorization of large sparse matrices based on a multifrontal approach in a multiprocessor environment*, PhD thesis, CERFACS, Toulouse, November 1990.
- [2] M. CROUZEIX, B. PHILIPPE, AND M. SADKANE, *The Davidson method*, Tech. Rep., Report TR/PA/90/45, CERFACS, Toulouse, 1990.
- [3] A. R. CURTIS AND J. K. REID, *On the automatic scaling of matrices for gaussian elimination*, J. Inst. Maths. Applics., 10 (1972), pp. 118–124.
- [4] E. R. DAVIDSON, *The iterative calculation of a few of the lowest eigenvalues and corresponding eigenvectors of large real-symmetric matrices*, Comp. Phys., (1975), pp. 87–94.
- [5] I. S. DUFF, R. G. GRIMES, AND J. G. LEWIS, *Sparse Matrix Test Problems*, ACM Trans. Math. Softw., (1989), pp. 1–14.
- [6] P. HENRICI, *Bounds for iterates, spectral variation and fields of values of non-normal matrices*, Numerische Mathematik 4., (1962), pp. 24–40.
- [7] D. HO, *Tchebychev Acceleration Technique for Large Scale Nonsymmetric Matrices*, Numer. Math., (1990), pp. 721–734.
- [8] KUCK AND A. INC., *Para-Eispack*, Kuck and associates, Urbana, IL, 1988.
- [9] R. B. MORGAN AND D. S. SCOTT, *Generalizations of Davidson's method for computing eigenvalues of sparse symmetric matrices*, SIAM J. Sci. Stat. Comput., vol. 7 (1986), pp. 817–825.
- [10] B. N. PARLETT, *The symmetric eigenvalue problem*, Prentice-Hall, Englewood Cliffs, N.J, 1980.

- [11] Y. SAAD, *Chebyshev acceleration techniques for solving nonsymmetric eigenvalue problems*, Math. Comp., Vol. 42 (1984), pp. 567–588.
- [12] M. SADKANE, *A Block Arnoldi-Chebyshev Method for Computing the Leading Eigenpairs of Large Sparse Unsymmetric Matrices*, Tech. Rep., Report TR/PA/91/46, CERFACS, Toulouse, 1991.
- [13] D. C. SORENSEN, *Implicit Application of Polynomial Filters in a  $k$ -Step Arnoldi Method*, Tech. Rep., Report TR90-27. Rice University, Huston, Texas 77251., October 1990.



## LISTE DES DERNIERES PUBLICATIONS INTERNES IRISA

- PI 642    ARCHE : UN LANGAGE PARALLELE A OBJETS FORTEMENT TYPES  
Marc BENVENISTE, Valérie ISSARNY  
Mars 1992, 132 pages.
- PI 643    CARTESIAN AND SATISTICAL APPROACHES OF THE SATISFIABILITY  
PROBLEM  
Israël-César LERMAN  
Mars 1992, 58 pages.
- PI 648    SET-THEORETIC GRAPH REWRITING  
Jean-Claude RAOULT, Frédéric VOISIN  
Mars 1992, 18 pages.
- PI 649    UNE STRUCTURE D'INFORMATION POUR LES ALGORITHMES  
D'EXCLUSION MUTUELLE FONDES SUR UNE ARBORESCENCE  
Jean-Michel HELARY, Achour MOSTEFAOUI, Michel RAYNAL  
Mars 1992, 18 pages.
- PI 650    BLOCK-ARNOLDI AND DAVIDSON METHODS FOR UNSYMMETRIC LARGE  
EIGENVALUE PROBLEMS  
Miloud SADKANE  
Avril 1992, 24 pages.
- PI 651    COMPILING SEQUENTIAL PROGRAMS FOR DISTRIBUTED MEMORY  
PARALLEL COMPUTERS WITH PANDORE II  
Françoise ANDRE, Olivier CHERON, Jean-Louis PAZAT  
Avril 1992, 18 pages.

ISSN 0249-6399