Filtering methods for finite-time normal modes and atmospheric error growth

Mozheng Wei* Jorgen S. Frederiksen* Steve Kepert*

(Received 7 August 2000)

Abstract

Krylov subspace methods are employed to generate the leading normal modes (NMs) and dominant finite-time normal modes (FTNMs) for atmospheric flows. The filtering techniques that play essential roles in our applications are described. The statistics of 100 evolved error fields are studied and compared with the growth and structures of fast

*Meteorology CRC, CSIRO Atmospheric Research, AUSTRALIA.
mailto:mwei@splash.princeton.edu, mailto:jorgen.frederiksen@dar.csiro.au, mailto:steve.kepert@dar.csiro.au

[0]See http://anziamj.austms.org.au/V42/CTAC99/Wei1 for this article and ancillary services, Published 27 Nov 2000.

growing NMs and FTNMs in a barotropic tangent linear model, with time-dependent basic states taken from observations.

Contents

1 Introduction

We apply Krylov subspace methods for generating the leading normal modes (NMs) and dominant finite-time normal modes (FTNMs) for atmospheric flows. Krylov subspace methods reduce the size of the matrix that needs to be solved in order to calculate the leading eigenvalues and eigenvectors of large systems of differential equations [3, 2, 4]. We describe the filtering techniques that we apply to start the Arnoldi process for generating the dominant NMs and FTNMs. We examine the structural organisation of initially random errors in a barotropic tangent linear model, with time-dependent basic states taken from observations. The statistics of 100 evolved error fields are studied for six day periods and compared with the growth and structures of fast growing NMs and FTNMs.

Section 2 briefly summarises the barotropic tangent linear model used in this study. The Krylov subspace methods and filtering methods used to generate leading NMs and FTNMs are presented in Section 3. In Section 4, we compare the growth and structures of evolved error fields with NMs and FTNMs and our conclusions are presented in Section 5.

2 Model Details

Models of atmospheric flows may be written in the form $\frac{d\mathbf{X}(t)}{dt} = \mathbf{N}[\mathbf{X}(t)]$, where $\mathbf{N}$ denotes a nonlinear operator. Here, $\mathbf{X}(t)$ is the state vector specifying the time-dependent flow in the phase space of either grid point values or spectral components of the dynamical variables. In this paper, the basic state flow will be taken directly from the observations.

We use the tangent linear equations based on the barotropic vorticity equations to describe error growth. This choice is based partly on simplicity and the need to perform large ensembles of simulations. For the flow on a sphere, the nondimensional form of the tangent linear vorticity equation is given by

$$\frac{\partial \zeta}{\partial t} = -J(\psi, \bar{\zeta} + 2\mu) - J(\bar{\psi}, \zeta) - \eta\zeta - \eta'\nabla^4\zeta, \tag{1}$$

where $J(\psi, \zeta) = \frac{\partial \psi}{\partial \lambda}\frac{\partial \zeta}{\partial \mu} - \frac{\partial \psi}{\partial \mu}\frac{\partial \zeta}{\partial \lambda}$, ψ is the streamfunction perturbation, $\zeta = \nabla^2\psi$ is the vorticity perturbation, while $\bar{\zeta}$ and $\bar{\psi}$ are the basic state vorticity and streamfunction respectively. The other parameters in the above equation are as follows: t is time, λ is longitude, μ is sine of latitude, and η and η' are the coefficients of viscosity representing drag and diffusion. All the variables are nondimensional with space coordinates scaled by the earth's radius and time scaled by Ω^{-1}, the inverse of the earth's angular velocity.

The spectral version of the barotropic tangent linear equation is obtained by expanding the streamfunction and vorticity in spherical harmonics, e.g.

$\zeta(\lambda,\mu,t) = \sum_{m=-R}^{R} \sum_{l=|m|}^{|m|+R} \zeta_{ml}(t) P_l^m(\mu) \exp(im\lambda)$, where R is a rhomboidal truncation wave number, which we take to be 15. Here $P_l^m(\mu)$ are orthonormalised Legendre functions, m is the zonal wave number and l is the total wave number. With $\mathbf{x}$ denoting the column vector of real and imaginary parts of spectral coefficients of perturbation fields, the tangent linear spectral equations can be written in the form

$$\frac{d\mathbf{x}(t)}{dt} = \mathbf{M}(t)\mathbf{x}(t), \tag{2}$$

where $\mathbf{M}(t)$ is the tangent linear operator evaluated on the observed trajectory $\mathbf{X}(t)$. The formal solution of equation (2) is

$$\mathbf{x}(t) = \mathbf{G}(t,t_0)\mathbf{x}(t_0), \tag{3}$$

where $\mathbf{G}(t,t_0)$ is called the propagator.

Since we are interested in studying the dependence of our results on the choice of viscosity used, we study both the inviscid case and a case with a typical magnitude of the viscosity. Comparison of the respective results gives an indication of the sensitivity to dissipation. In the viscous case, the coefficients of viscosity are chosen as $\eta = 8.4 \times 10^{-7}\mathrm{s}^{-1}$ and $\eta' = 2.5 \times 10^{16}\mathrm{m}^4\mathrm{s}^{-1}$. The tangent linear equations are solved with a half-hour time step. For the basic states we use 300-mb streamfunction fields that are taken from daily observations and linearly interpolated to obtain the time-dependent fields needed every half-hour.

3 Propagator Eigenmodes

3.1 Finite-time normal modes

The natural generalisations of normal modes, to the case of time-dependent instability matrices, $\mathbf{M}(t)$, are the eigenvectors of the propagator, $\mathbf{G}(t, t_0)$ [1]. The eigenvalue equation for $\mathbf{G}(t, t_0)$ is

$$[\lambda^\nu \mathbf{I} - \mathbf{G}(t, t_0)]\phi^\nu = 0, \quad \nu = 1, \ldots, n, \tag{4}$$

where $\lambda^\nu = \lambda^\nu[t, t_0]$ and $\phi^\nu = \phi^\nu[t, t_0]$ are the eigenvalues and eigenvectors. We call these eigenvectors finite-time normal modes (FTNMs) following [1]. In the above equation, $\mathbf{I}$ is the unit matrix. Since the propagator $\mathbf{G}(t, t_0)$ maps any initial perturbation at time t_0 to time t, it must carry some important information about the evolution of the basic state $\mathbf{X}(t)$, and its eigenvectors must play some role in the error growth from t_0 to t. From (4), it is evident that $|\lambda^\nu[t, t_0]|$ represents the amplification of the eigenmode during the period $t - t_0$ and the time-evolving basic state will be unstable if the modulus of $\lambda^\nu[t, t_0]$ is greater than 1.0. If we expand the initial vector $\mathbf{x}(t_0)$ in terms of the eigenvectors of $\mathbf{G}(t, t_0)$

$$\mathbf{x}(t_0) = \sum_{\nu=1}^{n} \kappa^\nu \phi^\nu, \tag{5}$$

then the solution to equation (3) is [1]

$$\mathbf{x}(t) = \sum_{\nu=1}^{n} \kappa^{\nu} \lambda^{\nu}(t, t_0) \phi^{\nu}, \tag{6}$$

where κ^{ν} is given by

$$\kappa^{\nu} = \langle \alpha^{\nu}, \mathbf{x}(t_0) \rangle \, / \, \langle \alpha^{\nu}, \phi^{\nu} \rangle \, . \tag{7}$$

Here $\langle \ , \ \rangle$ is the Euclidean inner product, and α^{ν} are the adjoint modes of $\mathbf{G}(t, t_0)$.

Equation (6) shows that the tangent linear dynamics filters the initial disturbances in favour of the fast growing FTNMs with larger amplification factors $|\lambda^{\nu}|$. This filtering tends to be more effective if the FTNMs are considered for longer time intervals $t - t_0$. However for any finite time interval $t - t_0$, we can increase the effectiveness of the filtering process by repeating it. The purging of subdominant FTNMs is achieved by integrating an initially random perturbation from t_0 to t, then recycling the perturbation back to t_0, integrating to t and continuing the process as many times as is required. The final filtered vector is then used to start the Arnoldi process.

In the Arnoldi method, the propagator $\mathbf{G}(t, t_0)$ is never explicitly calculated; what we need is just the action of $\mathbf{G}(t, t_0)$ on a vector and this is achieved by integrating the tangent linear model (2). In our numerical experiments, we have found that this method is very efficient and accurate. In cases with longer time intervals $t - t_0$, only a few filtering iterations are

needed to obtain a large number of the dominant FTNMs with very high accuracy.

3.2 Normal modes

In the special case when $\mathbf{M}(t)$ is independent of time, the propagator has the form

$$\mathbf{G}(t, t_0) = \exp \mathbf{M}(t - t_0). \tag{8}$$

This may be a reasonable approximation if, over a short time interval, $\mathbf{X}(t)$ is slowly varying. The propagator then has eigenvalues $\lambda^\nu = \exp \mu^\nu (t - t_0)$ and eigenvectors ϕ^ν, $\nu = 1, \ldots, n$. Here n is the dimension of $\mathbf{X}$, μ^ν are the complex eigenvalues of $\mathbf{M}$ and ϕ^ν are the eigenvectors of $\mathbf{M}$.

For non-zero and non-degenerate eigenvalues, the eigenmodes form a complete bi-orthogonal system together with the adjoint eigenmodes. The perturbation $\mathbf{x}(t)$ will be dominated by the leading normal modes (NMs), which are those associated with the largest real parts of μ^ν, when $t \gg t_0$. The non-leading modes will gradually become irrelevant with increasing time [1].

The method of purging the sub-dominant eigenvectors in the case when the matrix $\mathbf{M}$ is a constant, is simply to integrate the linear equation (2) forward for a sufficiently long time from a random initial condition. The evolved vector is eventually determined by the leading eigenvectors that grow most rapidly, with more slowly growing eigenvectors becoming irrelevant with increasing time.

If the perturbation becomes too large after a certain time, then the perturbation is normalised and the integration continued. This process is repeated until the integrated perturbation is dominated by the leading eigenvectors of $\mathbf{M}$. The final filtered perturbation vector is then used to start the Arnoldi process.

3.3 Implementation in iterative eigensolvers

For both constant and time-dependent $\mathbf{M}(t)$, the final filtered vector described in the above two subsections is denoted by $\mathbf{x}$. From this vector we then create k vectors, $\mathbf{x}_1 = \mathbf{x}$, $\mathbf{x}_2 = \mathbf{G}\mathbf{x}$, ..., $\mathbf{x}_k = \mathbf{G}^{k-1}\mathbf{x}$ where $\mathbf{G} = \mathbf{G}(t_0, t_0+\Delta t)$ for the case when $\mathbf{M}$ is constant and Δt is a suitable small time step. In our studies we have found that $\Delta t = 450\,\mathrm{s}$ or $900\,\mathrm{s}$ are satisfactory choices. The results with both of these values of Δt are of high accuracy. When $\mathbf{M}(t)$ is time dependent, we have $\mathbf{G} = \mathbf{G}(t_0, t)$ and the time interval $[t_0, t]$ depends on one's interest. The subspace spanned by these k vectors is called the Krylov subspace [3], $K_k(\mathbf{G}, \mathbf{x}) = span\{\mathbf{x}, \mathbf{G}\mathbf{x}, \mathbf{G}^2\mathbf{x}, \ldots, \mathbf{G}^{k-1}\mathbf{x}\}$ where $\mathbf{G}$ is of dimension $n \times n$ and $k \ll n$ in general. One can then extract approximations for $\mathbf{G}$ from this k-dimensional subspace $K_k(\mathbf{G}, \mathbf{x})$.

We start the Arnoldi process ([3]) from an initially normalised vector $\mathbf{w}_1 = \frac{\mathbf{x}(t)}{\|\mathbf{x}(t)\|}$, where $\mathbf{x}(t)$ is the final filtered vector that is dominated by either the leading NMs (for stationary basic state) or dominant FTNMs (for time-dependent basic state). The general terms in the factorisation are determined

through the coupled equations

$$\mathbf{v}_{i+1} = \mathbf{G}\mathbf{w}_i - \sum_{j=1}^{i} (\mathbf{w}_j^T \cdot \mathbf{G}\mathbf{w}_i)\mathbf{w}_j, \quad \mathbf{w}_{i+1} = \frac{\mathbf{v}_{i+1}}{\|\mathbf{v}_{i+1}\|}, \tag{9}$$

where $i = 1, 2, \ldots, k$. Each of the vectors $\mathbf{u}_i = \mathbf{G}\mathbf{w}_i$ is orthonormalised to all those previously generated. It can be shown that $\{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_k\}$ forms an orthonormal basis of the Krylov subspace K_k. In practice, we have used a modified Gram-Schmidt scheme, always with double orthogonalisation. This improves the orthogonality of all the vectors generated during the Arnoldi process, although this re-orthogonalisation increases the computational cost. One of the most reliable orthogonalisation techniques is the Householder algorithm, but it is more expensive than the modified Gram-Schmidt method [3]. Details about the techniques for calculating the eigenvalues and eigenvectors of $\mathbf{G}(t, t_0)$ from Equation (9) can be found in [4, 2].

4 Numerical Results

In this section, we compare the growth and structures of initially random errors growing on observed Southern Hemisphere basic states with properties of NMs and FTNMs. We focus on dynamical development during the period between 20 and 26 April 1989 when blocking highs developed at the longitudes of eastern Australia and in the central Pacific. We analyse the

statistics of the growth of 100 initial randomly generated error fields. The initial random perturbations are chosen from a Gaussian distribution in which the magnitudes of the streamfunction spectral coefficients are proportional to $(2l+1)^{-1}$; that is, $|\psi_{ml}| \sim \frac{1}{2l+1}$. The corresponding kinetic energy spectra are then consistent with typical estimates of 2-day forecast errors. Each perturbation is evolved in the tangent linear barotropic model from 20 to 26 April.

Figure 1 summarises our results for the increases in the amplification factors with time as the 100 initial random perturbations grow. Here the amplification factor A_f of a perturbation is defined as the ratio of its evolved root mean square (RMS) streamfunction amplitude to its initial RMS streamfunction amplitude. In Figure 1a, we show A_f for 20 of the perturbations; these 20 were the first of the 100 that were chosen by our random number generator. Also shown in Figure 1a is the amplification factor for FTNM 1 for different time periods starting on 20 April and finishing on the day indicated. We note that for all but one of the 20 first perturbations the amplification factors on a given day are smaller than for FTNM 1. In fact, this exception is the only outlier among the 100 error fields that exceeds the FTNM 1 amplification factor at any stage. Figure 1a also shows the mean of the amplification factors for the first 20 perturbations (indicated by □). The variations in the amplification factors of the error fields are reflections of the projections of the initial errors onto FTNM 1.

For the inviscid case, we depict, in Figure 1b, the mean (thin solid) of the 100 amplification factors of the error fields and as well the mean ± the

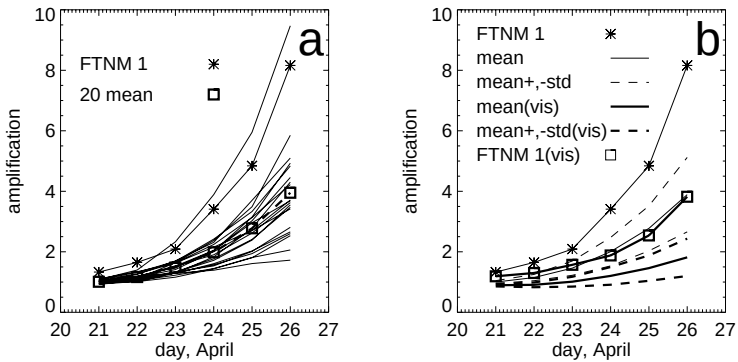

FIGURE 1: The amplification factors of FTNM 1 and evolved errors in both inviscid and viscous cases during periods from 20 to 21, 22, ..., 26 April 1989. (a). The first 20 of the 100 evolved random errors (solid lines), FTNM 1 in inviscid case ($*$) and the mean of the first 20 evolved random errors ($\square$). (b). Inviscid case: FTNM 1 ($*$), the mean of 100 evolved errors (thin solid line) and mean $\pm$ standard deviation of 100 evolved errors (thin dashed lines). Viscous case: FTNM 1 ($\square$), the mean of 100 evolved errors (thick solid line) and mean $\pm$ standard deviation of 100 evolved errors (thick dashed lines).

standard deviation (thin dashed) of the 100 amplification factors. Also shown in Figure 1b are the corresponding results for the viscous case (respective thick lines). Figure 1b again shows the amplification factor for FTNM 1 in the inviscid case and as well in the viscous case. We note that both the mean and mean + standard deviation of the amplification factors of the error fields on a given day are smaller than for FTNM 1 in the respective inviscid and viscous cases. As expected viscosity reduces the amplification factors of both the error fields and FTNMs.

Our interest here is in determining the statistics of pattern correlations (calculated over the Southern Hemisphere) between the random perturbations and the FTNMs. We note that the structure of FTNMs may change with time or phase (for generalised travelling FTNMs). We therefore calculate the pattern correlation (A_c) at the phase of the FTNM that gives the largest A_c . The mean (solid) and mean $\pm$ the standard deviation (dashed) of the pattern correlations between each of the 100 error fields and FTNMs are displayed in Figure 2. Figure 2a shows these quantities for the largest correlations taken over the five fastest growing FTNMs and for both the inviscid (thin lines) and viscous (thick lines) cases. The largest correlation is calculated as follows. For each error field at time t we calculate the pattern correlation with each of the five fastest growing FTNMs for the period t_0 to t and take the largest of these five pattern correlations. Then, we calculate the mean and mean $\pm$ standard deviations over 100 error fields. Figure 2b depicts the corresponding results for the correlations between the error fields and FTNM 1 in the inviscid (thin lines) and viscous (thick lines) cases.

In Figure 2a, the mean and the mean ± the standard deviations increase monotonically with time and there is little difference between the inviscid and viscous cases. However, correlations with individual FTNMs may be more variable as shown for FTNM 1 in Figure 2b. We note in particular that FTNM 1 generally, but not always, gives the largest average pattern correlations with the error fields. It appears that FTNM 1 and FTNM 3 swap roles on 24 April in the viscous case. From Figures 2a and b, we might also expect that the average correlation taken over the five fastest growing FTNMs and the mean correlation with FTNM 1 would increase further if the time interval were increased beyond six days. We have confirmed that this is the case (not shown).

Next, we consider probability distributions of correlations between the 100 evolved error fields on 26 April and FTNMs for the period 20 to 26 April. We compare these with corresponding probability distributions where the correlations involve NMs of the instantaneous basic state at 0000 UTC on 25 April instead of the FTNMs. Our purpose is to examine whether FTNMs are more likely predictors of the structure of evolved errors than NMs. That this should be the case has been suggested by the work of [1]. Figures 3a and b shows the probability distribution for the largest correlations taken over the five fastest growing FTNMs and taken over the five fastest growing NMs respectively. The probability distribution for the correlations with FTNM 1 and with NM 1 are displayed in Figures 3c and d. All diagrams are for the viscous case; very similar results are found for the inviscid case (not shown).

It is clear that there is considerable spread in the correlations for the 6

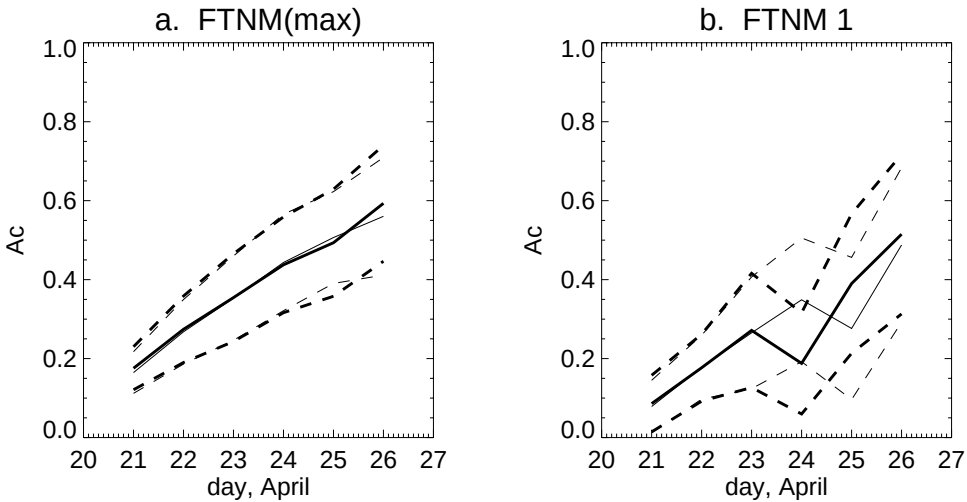

FIGURE 2: Pattern correlations (A_c) between the dominant FTNMs and the 100 evolved random errors in both inviscid and viscous cases during the periods from 20 to 21, 22, ..., 26 April 1989. (a). The mean (solid) and mean $\pm$ the standard deviation (dashed) of the largest correlations taken over the five fastest growing FTNMs (viscous—thick, inviscid—thin). (b). As in (a) but for the correlations between the 100 evolved random errors and FTNM 1.

day period ending on 26 April. However, the correlations taken over the five fastest growing FTNMs and the correlations with FTNM 1 are generally high. The average correlation taken over the five fastest growing FTNMs is about 0.6, and the mean correlation with FTNM 1 is slightly larger than 0.5 (Figs. 2a and b). The correlations taken over the five fastest growing NMs and the correlations with NM 1 tend to be lower than with the respective FTNMs. They are nevertheless significant and indicate that it may be possible to obtain a reasonable representation of evolved error fields through expansions in terms of a subset of the faster growing NMs [1].

5 Conclusions

We have applied Arnoldi methods for generating leading NMs and FTNMs for observed Southern Hemisphere atmospheric basic states during April 1989. We have described the filtering techniques needed to prepare the initial perturbations for starting the Arnoldi process for both NMs and FTNMs. The iterative methods are capable of yielding the leading eigenvalues and eigenvectors to high accuracy.

The structural organisation of initially random errors evolving in a barotropic tangent linear model, with time-dependent basic states taken from observations, has been examined for cases of block development. The statistics of 100 evolved errors have been studied for six day periods and compared with the growth and structures of fast growing NMs and FTNMs. The ampli-

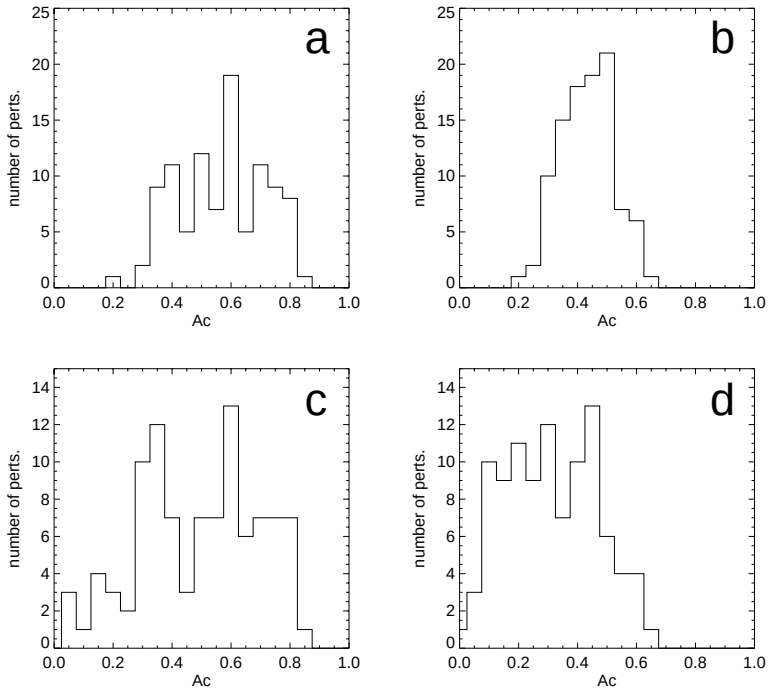

FIGURE 3: The probability distributions of correlations between the 100 evolved random errors on 26 April and FTNMs for the period 20 to 26 April for the viscous case. (a). Shown are the probability distributions for the largest correlations taken over the five fastest growing FTNMs. (b) As in (a) but taken over the five fastest growing NMs. (c). The probability distributions for the correlations with FTNM 1. (d). As in (c) but for NM 1

fication factors of most random errors are comparable to those of the fastest growing FTNM for the same time interval. The FTNMs are shown to provide a good representation of the statistics of error growth with mean pattern correlation between the 100 error fields and the dominant FTNMs increasing to a value close to 0.6 or larger after six days.

References

[1] J.S. Frederiksen. Adjoint sensitivity and finite-time normal mode disturbances during blocking. *J. Atmos. Sci.*, 54:1144–1165, 1997. C1487, C1487, C1488, C1489, C1495, C1497

[2] D.L. Harrar II and M.R. Osborne. Composite Arnoldi-Newton methods for large nonsymmetric eigenvalue problems. In B.J. Noye, M.D. Teubner and A.W. Gill, editors, *Computational Techniques and Applications: CTAC97*, pages 273–280, Singapore, 1998. World Sci. Pub. C1484, C1491

[3] Y. Saad. *Numerical Methods for Large Eigenvalue Problems*. Halsted Press-John Wiley & Sons Inc., New York, 1992. C1484, C1490, C1490, C1491

[4] M. Wei and J.S. Frederiksen. Applications of iterative eigensolvers to error growth during blocking in the southern hemisphere. In B.J. Noye, M.D. Teubner and A.W. Gill, editors, *Computational Techniques and Applications: CTAC97*, pages 727–734, Singapore, 1998. World Sci. Pub. C1484, C1491