

# Convergence Error Estimation and Convergence Acceleration in Iteratively Solved Problems

S. Eyi\*

Corresponding author: seyi@metu.edu.tr

\*Aerospace Engineering, Middle East Technical University Ankara, Turkey

**Abstract:** New methods are developed for convergence error estimation and convergence acceleration in iteratively solved problems. The convergence error estimation method is based on the eigenvalue analysis of linear systems, but it can also be used for nonlinear systems. The convergence of iterative method is accelerated by subtracting convergence error from the iteratively calculated solutions. The performances of these methods are demonstrated for the Laplace, Euler and Navier-Stokes equations.

*Keywords:* Convergence Error, Convergence Acceleration, Iteratively Solved Problems.

## 1 Introduction

In iteratively solved problems, errors in numerical calculations usually come from three different sources. The first is the round-off error due to the finite arithmetic precision of computer and it depends on the type of computer processor and compiler. The second is the convergence error that can be defined as the difference between the exact and iterative solutions of discretized governing equations. Here, the exact solution is defined as the solution that exactly satisfies the discretized governing equations with a zero residual. Although it is not always possible to reduce the residual values to the desired level, the convergence error can be reduced by solving the discretized governing equations with smaller residual tolerance. The third source is the truncation error and it results from neglected terms in the Taylor series expansion. A truncation error depends on the accuracy of the finite-difference stencil used in the discretized equations, and the finite-difference step size. In iteratively solved problems, if the round-off error is neglected, the total error can be defined as the summation of the convergence and truncation errors. This paper concentrates only on the convergence error.

There is a great interest in estimating the convergence error. Knowing when to stop iteration is important in terms of computational efficiency and accuracy. In most of the iteratively solved problems, the reduction in residual is used as a stopping criterion. Unfortunately, the reduction in residual may not be a reliable measure for the convergence error. Different methods were developed to measure the convergence error. Ferziger and Peric [1,2] used eigenvalue analysis and assumed that the system of nonlinear equation has linear behavior as it gets closer to the converged solution. Bergstrom, and Gebart [3] implemented the same method to estimate the convergence error for a flow problem in a draft tube. Roy et al. [4] used an exponential equation to estimate the convergence error of hypersonic flow

problems. Alekseev [5] calculated the convergence error using an adjoint parameter and time derivative. Brezinski [6] developed a method to estimate the error in the solution of linear systems. In the present study, a new method is developed that estimates the convergence error in iteratively solved problems. The method is based on the eigenvalue analysis of linear systems, as presented by Feziger and Peric [1,2]. An equation is developed between the convergence error and correction vectors. The convergence error vector is expressed as the linear combination of the correction vectors and the coefficients of the correction vectors are calculated by the least-squares minimization of the derived equation [7].

Once knowing how to estimate the convergence error, the next step is to develop a convergence acceleration method. In literature, there are only few research on convergence acceleration based on convergence error estimation. Some of the extrapolation based convergence acceleration methods estimate the convergence error. Hafez et al [8] extrapolated the exact solution by using the convergence error which is estimated with the power method. Jespresen and Bunning [9] accelerated the iterative process by annihilating the dominant eigenvalues. Dagan [10] developed a convergence acceleration method based on the power method. Sidi [11, 12] used different extrapolation methods to accelerate the convergence of iterative solutions. The convergence acceleration method presented in this study is based on the estimation of convergence error.

## 2 Convergence Error Estimation

In this section, the new method to estimate the convergence error in iteratively solved problems is presented. Although the developed method is based on the eigenvalue analysis of linear systems, it can also be used for nonlinear systems, especially during near convergence when nonlinear systems behave like linear systems and for which the error estimation is most needed. First, the iterative solution of the system of linear equations is reviewed. Similar reviews can also be found in refs. [1, 2]. The system of linear equations can be defined in the following form:

$$Aw = b \quad . \quad (1)$$

In the equation above,  $w$  is the exact solution of the system. An iterative scheme can be constructed by splitting matrix  $A$  as follows:

$$A = M - N \quad . \quad (2)$$

In this splitting,  $M$  is chosen so that the system can be easily solvable with an iterative scheme,

$$M\tilde{w}^{n+1} = N\tilde{w}^n + b \quad , \quad (3)$$

where  $\tilde{w}^n$  is the iterative solution of the state variable vector after  $n$  iterations. Since  $w$  is the exact solution of the system, it also satisfies the iterative scheme given in Eq. (3)

$$Mw = Nw + b \quad . \quad (4)$$

At iteration  $n$ , the convergence error vector can be defined as the difference between the iterative and exact solution vectors,

$$\mathcal{E}^n = \tilde{w}^n - w \quad . \quad (5)$$

Defining the correction vector  $\delta^n$  as,

$$\delta^n = \tilde{w}^{n+1} - \tilde{w}^n \quad , \quad (6)$$

and using (5), the following relation between the correction and convergence error vectors can be written, thus,

$$\delta^n = \varepsilon^{n+1} - \varepsilon^n . \quad (7)$$

Subtracting Eq. (4) from Eq. (3) gives a relation between the convergence error vectors of two successive iterates,

$$\varepsilon^{n+1} = M^{-1}N\varepsilon^n . \quad (8)$$

The iterative method converges if the spectral radius of the matrix  $M^{-1}N$  is lower than one. The convergence of the iterative scheme can be analyzed with the use of eigenvalues,  $\lambda_k$ , and eigenvectors,  $\varphi_k$ ,

$$M^{-1}N\varphi_k = \lambda_k\varphi_k, \quad k=1, KMAX, \quad (9)$$

where  $KMAX$  is the total number of state variables in the system. For the case of complex eigenvalues, the following equation is also satisfied:

$$M^{-1}N\varphi_k^* = \lambda_k^*\varphi_k^*, \quad k=1, KMAX, \quad (10)$$

where  $\lambda_k^*$ , and  $\varphi_k^*$  are the complex conjugate of the eigenvalues and eigenvectors, respectively. Most of the time, iterative methods have complex eigenvalues. Considering the more general case and assuming real numbers as the special case of complex numbers, in the present derivation, complex eigenvalues and eigenvectors are used. Using the linearly independent eigenvectors, the initial error  $\varepsilon^0$  may be expressed as a linear combination of eigenvectors:

$$\varepsilon^0 = \sum_{k=1}^{KMAX} a_k\varphi_k + a_k^*\varphi_k^*, \quad (11)$$

where  $a_k$  and  $a_k^*$  are generalized Fourier coefficients. The combination of Eqs. (8) and (11) yields:

$$\varepsilon^1 = \sum_{k=1}^{KMAX} M^{-1}N(\varphi_k a_k + \varphi_k^* a_k^*) . \quad (12)$$

Substituting Eqs. (9) and (10) into the equation above gives,

$$\varepsilon^1 = \sum_{k=1}^{KMAX} a_k\lambda_k\varphi_k + a_k^*\lambda_k^*\varphi_k^* . \quad (13)$$

By induction, the error vector at iteration  $n$ , can be written as:

$$\varepsilon^n = \sum_{k=1}^{KMAX} a_k(\lambda_k)^n\varphi_k + a_k^*(\lambda_k^*)^n\varphi_k^* . \quad (14)$$

After a number of iterations, the contribution of the largest eigenvalue  $\lambda_1$  becomes more significant, and the error vector can be approximated as:

$$\varepsilon^n = a_1(\lambda_1)^n\varphi_1 + a_1^*(\lambda_1^*)^n\varphi_1^* . \quad (15)$$

The derivation of above equations can also be found in Refs. [1, 2]. In the present study, the following method is developed to estimate the convergence error, which is one of the original contributions of this paper. Rearranging Eq. (15) yields a relationship between the convergence error vectors at iteration  $n+1$ ,  $n$  and  $n-1$ :

$$\varepsilon^{n+1} = (\lambda_1 + \lambda_1^*)\varepsilon^n - \lambda_1\lambda_1^*\varepsilon^{n-1} . \quad (16)$$

It is not difficult to show that the convergence error vector can be expressed as a function of the correction vectors,

$$\varepsilon^{n+1} = \frac{(C_1 + C_2)\delta^n + C_2\delta^{n-1}}{C_1 + C_2 - 1}. \quad (17)$$

Substituting Eq. (7) into Eq. (16) gives the following relationship between the correction vectors at iteration  $n$ ,  $n-1$  and  $n-2$ :

$$\delta^n = C_1\delta^{n-1} + C_2\delta^{n-2}. \quad (18)$$

In the equation given above,  $C_1$  and  $C_2$  are real numbers and they are functions of eigenvalues,

$$C_1 = \lambda_1 + \lambda_1^*, \quad (19)$$

$$C_2 = -(\lambda_1\lambda_1^*).$$

Assuming that  $l_1$  is the magnitude of the largest eigenvalue and  $\theta_1$  is its phase angle, eigenvalues can be written as:

$$\lambda_1 = l_1 e^{i\theta_1} \quad \text{and} \quad \lambda_1^* = l_1 e^{-i\theta_1}. \quad (20)$$

Then, Eq. (19) becomes

$$C_1 = 2l_1 \cos \theta_1, \quad (21)$$

$$C_2 = -l_1^2.$$

In the calculation of the convergence error vector, first, the coefficients  $C_1$  and  $C_2$  are determined from the least-squares solution of Eq. (18). Then, using Eq. (17), the error vector can be evaluated as a function of the correction vectors of two successive iterations. In order to calculate the coefficients, the correction vectors from the current and previous two iterations must be stored. In this method, the calculation of the values of eigenvalues is not required because the values of coefficients  $C_1$  and  $C_2$  are sufficient to determine the convergence error vector.

Increasing the number of eigenvalues may improve the convergence error estimation. For example, approximating Eq. (14) by using the first and second largest eigenvalues, the convergence error vector can be formulated as:

$$\varepsilon^n = a_1 (\lambda_1)^n \varphi_1 + a_1^* (\lambda_1^*)^n \varphi_1^* + a_2 (\lambda_2)^n \varphi_2 + a_2^* (\lambda_2^*)^n \varphi_2^*. \quad (22)$$

The following relation can be derived to calculate the convergence error vector at iteration  $n+1$ :

$$\varepsilon^{n+1} = \frac{(C_1 + C_2 + C_3 + C_4)\delta^n + (C_2 + C_3 + C_4)\delta^{n-1} + (C_3 + C_4)\delta^{n-2} + (C_4)\delta^{n-3}}{C_1 + C_2 + C_3 + C_4 - 1}. \quad (23)$$

Similarly, the relation between the correction vector at iteration  $n$  and the correction vectors at iterations  $n-1$ ,  $n-2$ ,  $n-3$ , and  $n-4$  can be written as:

$$\delta^n = C_1\delta^{n-1} + C_2\delta^{n-2} + C_3\delta^{n-3} + C_4\delta^{n-4}. \quad (24)$$

In parallel with the previous derivations, the real coefficients  $C_1$ ,  $C_2$ ,  $C_3$ , and  $C_4$  in Eq. (23) are the functions of eigenvalues and these coefficients are calculated from the least-squares solution of Eq. (24),

$$C_1 = \lambda_1 + \lambda_1^* + \lambda_2 + \lambda_2^*, \quad (25)$$

$$C_2 = -(\lambda_1\lambda_1^* + \lambda_1\lambda_2 + \lambda_1\lambda_2^* + \lambda_1^*\lambda_2 + \lambda_1^*\lambda_2^* + \lambda_2\lambda_2^*),$$

$$C_3 = \lambda_1\lambda_1^*\lambda_2 + \lambda_1\lambda_1^*\lambda_2^* + \lambda_1\lambda_2\lambda_2^* + \lambda_1^*\lambda_2\lambda_2^*,$$

$$C_4 = -(\lambda_1 \lambda_1^* \lambda_2 \lambda_2^*)$$

Assuming that  $l_1$  and  $l_2$  are the magnitudes of the largest and the second largest eigenvalues and  $\theta_1$  and  $\theta_2$  are their phase angles, above equation can be written as:

$$C_1 = 2l_1 \cos \theta_1 + 2l_2 \cos \theta_2 , \quad (26)$$

$$C_2 = -[(2l_1 \cos \theta_1)(2l_2 \cos \theta_2) + l_1^2 + l_2^2] ,$$

$$C_3 = l_1^2 (2l_2 \cos \theta_2) + l_2^2 (2l_1 \cos \theta_1) ,$$

$$C_4 = -(l_1^2 l_2^2) .$$

In the calculation of these coefficients, the correction vectors must be stored from the current and last four iterations. As explained above, in the calculation of the convergence error vector, the values of eigenvalues are not required. The convergence error vector can be calculated as a function of coefficients  $C_1$ ,  $C_2$ ,  $C_3$ , and  $C_4$ .

The relations given above can be generalized for an arbitrary number of eigenvalues. Approximating Eq. (14) by using the  $M_{eigen}$  number of eigenvalues, the convergence error becomes:

$$\mathcal{E}^n = a_1 (\lambda_1)^n \phi_1 + a_1^* (\lambda_1^*)^n \phi_1^* + \dots + a_{M_{eigen}} (\lambda_{M_{eigen}})^n \phi_{M_{eigen}} + a_{M_{eigen}}^* (\lambda_{M_{eigen}}^*)^n \phi_{M_{eigen}}^* . \quad (27)$$

By induction, the convergence error vector can be generalized for an arbitrary number of eigenvalues,  $M_{eigen}$ , as:

$$\mathcal{E}^{n+1} = \frac{\left( \sum_{m=1}^{2M_{eigen}} C_m \right) \delta^n + \left( \sum_{m=2}^{2M_{eigen}} C_m \right) \delta^{n-1} + \left( \sum_{m=3}^{2M_{eigen}} C_m \right) \delta^{n-2} + \dots + (C_{2M}) \delta^{n-2M_{eigen}+1}}{\sum_{m=1}^{2M_{eigen}} C_m - 1} . \quad (28)$$

Similarly, the correction vector can be generalized as:

$$\delta^n = \sum_{m=1}^{2M_{eigen}} C_m \delta^{n-m} . \quad (29)$$

As previously explained, the coefficients  $C_m$  in Eq. (28) are real numbers, and they are determined from the least-squares solution of Eq. (29). In the calculation of the coefficients, the correction vectors from the current and last  $2M_{eigen}$  iterations must be stored. Although, increasing the number of eigenvalues may improve the accuracy of convergence error estimation, this improvement may also result in an increase in the amount of memory required to store the correction vectors from the previous iterations.

In many iteratively solved problems, the convergence rate is superlinear, so the convergence accelerates as the error gets smaller. For example, in Newton's method, a quadratic convergence rate can be achieved and the residual can be reduced to the order of machine epsilon in a few iterations. In such cases, the estimation of convergence error with an arbitrary number of eigenvalues may be difficult. There may not be sufficient iterations to determine the coefficients in Eq. (29). Smaller number of eigenvalues may be used in these problems.

### 3 Eigenvalue Estimation

Although the calculation of eigenvalues may not be needed for convergence error estimation, above

relations can also be used in the calculation of eigenvalues. If the number of eigenvalue is one, Eq. (21) can be used in the evaluation of the magnitude and the phase angle of the largest eigenvalue:

$$l_1 = \sqrt{-C_2} \quad (30)$$

$$\cos \theta_1 = \frac{C_1}{2\sqrt{-C_2}}$$

If the number of eigenvalues greater than one, calculations of the magnitude and the phase angles of eigenvalues require the solution of the system of nonlinear equations. The number of equations to be solved is twice the number of eigenvalues. For each eigenvalue, there are two unknowns: magnitude and phase angle. In order to generalize the calculations for an arbitrary number of eigenvalues, the following modifications can be useful. Eq. (21) can be modified as:

$$l_1 \cos \theta_1 = D_1 \quad (31)$$

$$l_1^2 \cos(2\theta_1) = D_2$$

where the new coefficients  $D_1$  and  $D_2$  are the functions of  $C_1$  and  $C_2$  and are given

$$D_1 = \frac{C_1}{2} \quad (32)$$

$$D_2 = C_1 D_1 + C_2$$

Using the Chebyshev polynomials, Eq. (31) can be modified as:

$$l_1 T_1(x_1) = D_1 \quad (33)$$

$$l_1^2 T_2(x_1) = D_2$$

where  $x_1 = \cos \theta_1$ . If the numbers of eigenvalues is two, the magnitude and the phase angles of the largest and the second largest eigenvalues can be calculated by modifying the Eq. (26) in the following form:

$$l_1 \cos \theta_1 + l_2 \cos \theta_2 = D_1 \quad (34)$$

$$l_1^2 \cos(2\theta_1) + l_2^2 \cos(2\theta_2) = D_2$$

$$l_1^3 \cos(3\theta_1) + l_2^3 \cos(3\theta_2) = D_3$$

$$l_1^4 \cos(4\theta_1) + l_2^4 \cos(4\theta_2) = D_4$$

where the coefficients  $D_1$ ,  $D_2$ ,  $D_3$  and  $D_4$  are the functions of  $C_1$ ,  $C_2$ ,  $C_3$  and  $C_4$ :

$$D_1 = \frac{C_1}{2} \quad (35)$$

$$D_2 = C_1 D_1 + C_2$$

$$D_3 = C_1 D_2 + C_2 D_1 + \frac{3}{2} C_3$$

$$D_4 = C_1 D_3 + C_2 D_2 + C_3 D_1 + 2C_4$$

Similarly, using the Chebyshev polynomials Eq. (34) can be modified as:

$$l_1 T_1(x_1) + l_2 T_1(x_2) = D_1 \quad (36)$$

$$l_1^2 T_2(x_1) + l_2^2 T_2(x_2) = D_2$$

$$l_1^3 T_3(x_1) + l_2^3 T_3(x_2) = D_3$$

$$l_1^4 T_4(x_1) + l_2^4 T_4(x_2) = D_4$$

where  $x_2 = \cos \theta_2$ . For an arbitrary number of eigenvalues, above equations can be generalized as:

$$\sum_{m=1}^{M_{eigen}} l_m^n T_n(x_m) = D_n, \quad n = 1, 2M_{eigen} \quad (37)$$

where

$$D_n = \frac{n}{2} C_n + \sum_{k=1}^{n-1} C_k D_{n-k}, \quad n = 1, 2M_{eigen}$$

In the present study, above system of nonlinear equations is solved using Newton's method. Newton's method requires the Jacobian matrix. The entries of the Jacobian matrix are the derivatives of the residual vector with respect to the magnitude and phase angle of the eigenvalues. The size of the Jacobian matrix is  $(2M_{eigen})^*(2M_{eigen})$ .

## 4 Convergence Acceleration

Once the convergence error is estimated accurately, the exact solution which satisfies the discretized governing equation can be estimated. A convergence acceleration method can be developed by subtracting the convergence error from the iterative solution.

$$w^* = \tilde{w}^n - \varepsilon^n \quad (38)$$

where  $w^*$  is the estimated exact solution. After estimating the exact solution different approaches may be used for convergence acceleration. In the present study, two convergence acceleration methods are developed. In Method I, at every  $2M_{eigen}$  iterations, the convergence error is calculated and then it is subtracted from the iterative solution. New iterations are started from the estimated exact solution,  $w^*$ . In the next  $2M_{eigen}$  iterations the original iterative procedure is executed without calculating the exact solution. In this method, Eqs. (28) and (29) are implemented at every  $2M_{eigen}$  iterations and hence, the convergence acceleration procedure is not continuous. In Method II, the exact solution is estimated continuously and stored as a new variable. As the original iterative algorithm progresses the exact solution is calculated.

The algorithm in Method I can be given as:

1. Initialize the iterative solution,  $\tilde{w}^0$ .
2. Execute the original iterative procedure  $2^*M_{eigen}$  of times, and store the correction vectors from these iterations.
3. Evaluate the coefficients  $C_n$ 's using Eq. (29).
4. Calculate the convergence error using Eq. (28).
5. Estimate the exact solution,  $w^*$ , using Eq. (38).
6. Initialize the iterative solution with the estimated exact solution.
7. Repeat steps 2 – 6 until the norm value of residual is smaller than a specified residual.

The algorithm in Method II can be given as:

1. Initialize the iterative solution,  $\tilde{w}^0$ .
2. Execute the original iterative procedure  $2 * M_{eigen}$  of times, and store the correction vectors from these iterations.
3. Evaluate the coefficients  $C_n$ 's using Eq. (29).
4. Calculate the convergence error using Eq. (28).
5. Estimate the exact solution,  $w^*$ , using Eq. (38).
6. Define the exact solution as a new variable. (do not initialize the iterative solution with the estimated exact solution)
7. Calculate the norm value of residual based on the new variable.
8. Execute the original iterative procedure to calculate new correction vector (one iteration)
9. Repeat steps 3 – 8 until the norm value of residual based on the new variable is smaller than a specified residual.

Although, in both methods, the same equations are used to calculate the convergence error and exact solutions, there are some differences in the implementation of convergence acceleration algorithms. In Method I, the exact solution is estimated at every  $2 * M_{eigen}$  iterations, and this solution is used as an initial solution for the original iterative procedure. In Method II, new variables are used for the estimated exact solution. At each iterations, while the original iterative solution progresses, the exact solution is calculated.

## 5 Result

The accuracy of the method for estimating the convergence error is tested for both linear and non-linear problems. All calculations in this study are performed with double precision on a 1.5GHz Pentium IV dual core processor. As a linear problem, a two-dimensional Laplace equation in a square domain ( $0 < x < 1$ ;  $0 < y < 1$ ) is solved using the successive over relaxation method with a relaxation parameter of  $\Omega$ . The Laplace equation is one of the few equations whose exact solution can be analytically evaluated with the appropriate boundary conditions. A second-order central-difference scheme is used on a uniform cartesian grid. The boundary condition is chosen to satisfy the solution  $\varphi(x,y)=100xy$ . Since the terms related to the truncation error are eliminated, this analytical solution is also the exact solution of the Laplace equation that is discretized with the second-order and central-difference scheme [1, 2]. The real value of the convergence error is calculated as the difference between the exact solution and the computed solution from the current iteration. The estimated error is defined as the error calculated with the proposed method explained in section II. B. In order to have better understanding of the order of the error in each entry of the convergence error vector, the norm values are redefined according to following relations:

$$\|\mathcal{E}\|_p = \left( \frac{1}{KMAX} \sum_{k=1}^{KMAX} |\mathcal{E}_k|^p \right)^{\frac{1}{p}}, \quad \|\mathcal{E}\|_{\infty} = \max(|\mathcal{E}_1|, \dots, |\mathcal{E}_{KMAX}|) . \quad (39)$$

The first equation given above for  $p=1$  and  $p=2$  corresponds to the absolute-mean and root-mean-square, respectively. Unless otherwise stated, in all figures to predict the convergence error, the number of eigenvalues,  $M_{eigen}$  is set to 4. Solving the Laplace equation with a grid size of 80X80 gives the number of state variables,  $KMAX$ , 6400.

Figure 1a shows the  $\|\mathcal{E}\|_2$  norm values of the real and estimated convergence errors, and their differences. The sole reason for showing the norm value of residual vector in Figure 1a is to see the

effectiveness of using residual values to estimate the convergence error. The norm values of the residual are almost two orders of magnitude smaller than that of the convergence error. The results show that although the residual itself is not a good parameter to predict the convergence error, the proposed method can accurately estimate the convergence error. As the number of iterations increases, the difference between the real and estimated errors becomes smaller. In Figure 1b, the effects of the number of eigenvalues on the estimation of the convergence error are presented. The results show that increasing the number of eigenvalues from 1 to 64 slightly improves the accuracy of the error prediction. However, this improvement requires more memory to store the correction vectors. As shown in Eq. (29), using  $M_{eigen}$  number of eigenvalues entails the storing of the correction vectors from the last  $2M_{eigen}$  iterations. In the estimation of the convergence error with  $M_{eigen}$  number of eigenvalues, the convergence error estimation starts after  $2M_{eigen}$  number of iterations. In the present study, until that iteration is reached, the convergence error is estimated with the maximum number of available eigenvalues. The convergence error estimation starts at the fourth iteration by using only one eigenvalue. Between iterations four and  $2M_{eigen}$ , the convergence error at iteration  $n$  is estimated using the  $n/2$  or  $(n-1)/2$  number of eigenvalues, depending on whether the iteration number is even or odd, respectively.

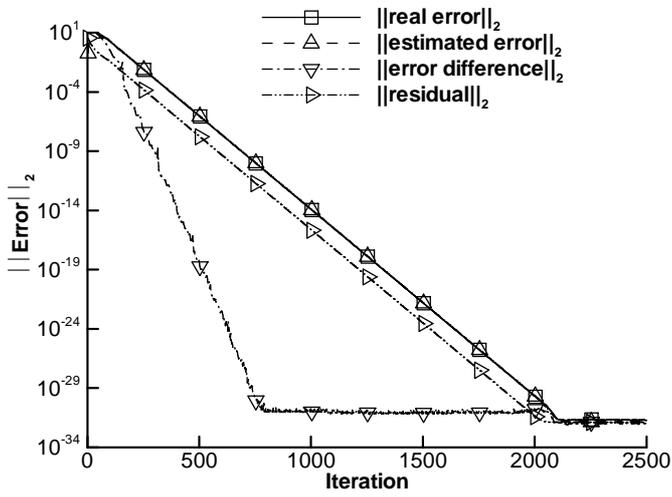
The norm values of the convergence error estimated using the proposed method and the method given in refs. [1, 2] are compared in Figure 1c. In this comparison, the value of the over relaxation parameter,  $\Omega$ , is set to 1.95 which is the maximum attainable value for this grid size. The solution diverges with the higher values of  $\Omega$ . This figure clearly shows the superiority of the proposed method. In the method given in refs. [1, 2], the convergence error is estimated by using the largest eigenvalue hence, the accuracy level of the method is fixed. However, the accuracy level of the proposed method can be adjusted by changing the number of eigenvalues. In Figure 1c, the proposed method uses two different numbers of eigenvalues: 4 and 64. In terms of estimating the convergence error, the proposed method with 4 eigenvalues is more successful than the method given in refs. [1, 2]. The same figure also shows that the accuracy of the proposed method improves as the number of eigenvalues increases. If the proposed method uses 64 eigenvalues, the real and estimated errors cannot be distinguished after 100 iterations. Figure 1d shows the effects of the value of the relaxation parameter,  $\Omega$ , on the estimation of the convergence error. In all cases, the values of estimated and real errors are very close to each other. For the relaxation parameter of 0.7 and 1.3 some oscillations are observed in the norm values of error difference.

Next, the performance of the convergence error estimation method is analyzed for nonlinear problems. Two dimensional Euler and Navier-Stokes equations are solved around a NACA0012 airfoil at a transonic flow condition ( $M=0.730$ ,  $\alpha=2.78^\circ$ ,  $Re=6.5 \times 10^6$ ). Solving the equations at transonic flow condition increases the nonlinearity in the solution. A finite volume method is used for spatial discretization and the flow variables are defined at the cell centers. Centered differencing is used for the spatial derivatives and the second-order and fourth-order artificial viscosities are added to enforce numerical stability. Time integration is performed using an explicit, four-stage Runge-Kutta scheme. Local time stepping and a multigrid method are implemented to accelerate the convergence to obtain a steady state solution. The multigrid level for the Euler and Navier-Stokes solutions are three and four, respectively. Characteristic boundary conditions are imposed at the far-field boundary based on a one-dimensional eigenvalue analysis. In the solution of the Euler equations, zero normal mass flux is enforced at the airfoil surface and the pressures are extrapolated from the inside cells using the normal momentum equation. In the Navier-Stokes solutions, no-slip and adiabatic wall conditions are used at the airfoil surface, and the Baldwin-Lomax eddy-viscosity model is included for turbulence closure. In the Euler computations, the grid size is  $129 \times 33$ , in the Navier-Stokes computations, the grid size is  $257 \times 65$  and the minimum grid spacing next to the wall is  $10^{-5}$ . Having four variables in each cell, the values of  $KMAX$  are 17028 and 66820 in the solutions of Euler and Navier-Stokes equations, respectively. It is difficult to find analytical relations for the exact solution of the Euler and Navier-Stokes equations. Therefore, in these

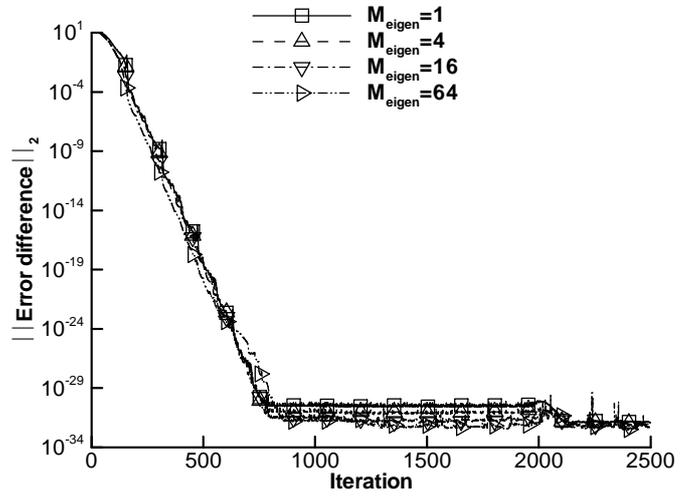
problems, the exact solutions are estimated by iterating solutions until the residual norm becomes the order of a machine epsilon. The real convergence error is calculated as the difference between the solutions for the current iteration and the exact solution. As in the solution of the Laplace equation, the estimated convergence error is calculated by using the method presented in section II.B. The convergence error is calculated on the finest grid of a multigrid cycle and all four conservative flow variables are used in the error estimation.

As a nonlinear problem, the convergence error analysis is first performed on the Euler equations. Figure 2a shows the change of the real and estimated errors and their difference in relation to the number of iterations. In this figure, the convergence error is estimated using 4 eigenvalues. Although the equations are nonlinear and the flow condition enforces the nonlinearity, the real and estimated errors are almost the same. The difference between the estimated and real errors decreases as the number of iterations increases. Figure 2a also shows the variation of the residual with respect to the number of iterations and it can be seen that there is a large difference between the residual and convergence error. Therefore, residual may not be a good parameter to predict the error. The effects of the number of eigenvalues on the estimation of the convergence error are analyzed in Figure 2b. Results show that the accuracy of convergence estimation method can be significantly improved by increasing the number of eigenvalues. Figure 2c shows that in all three norms, the proposed method predicts the convergence error very accurately. Figure 2d shows that as the CFL number increases the norm values of the error difference get smaller.

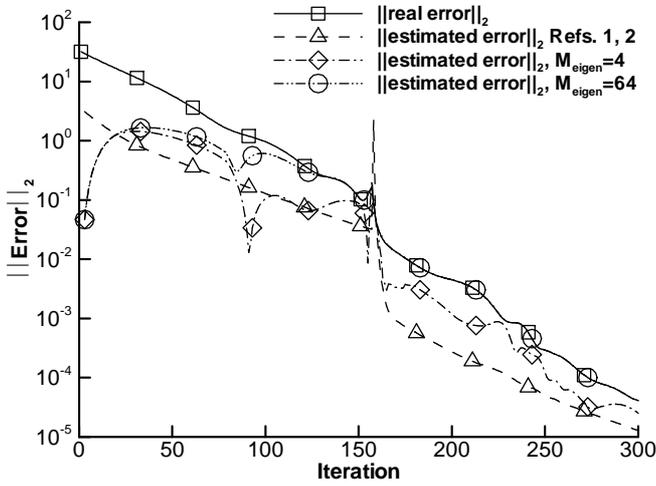
Last, the performance of the proposed method is demonstrated for the Navier-Stokes equations. In the Navier-Stokes equations, there are additional nonlinear terms related to viscous fluxes and turbulence modeling. The real and estimated errors are determined by the similar methods used in Euler equations. In order to show the effects of the precision error on the performance of the convergence error estimation method, computations are performed with single precision. Figure 3a shows that the proposed method can estimate the convergence error very accurately for nonlinear problems. The estimated error almost exactly matches the real error and large reductions are achieved in the difference between the estimated and real errors. In Figure 3b, it can be seen that, increasing the number of eigenvalues decreases the difference between the real and estimated error. However, these improvements in error estimation are achieved at the cost of a large increase in CPU time. For many engineering problems, setting the number of eigenvalues between 1 and 4 may be sufficient. The real and estimated errors comparison using different norms is shown in Figure 3c and again, excellent results are achieved for non-linear problems. In iterative schemes, it is not always possible to reduce the residual values to the order of the round-off error. In the solution of the Navier-Stokes equations, only the CFL number of 1.5 reduces the residual values to that order. Hence, the effects of the CFL numbers on the accuracy of the convergence error estimation method could not be explored for the Navier-Stokes equations. Figure 3d compares real and estimated errors at the initial iterations with the estimated error being calculated with different numbers of eigenvalues. Results show that the convergence error can be accurately estimated in the first twenty iterations.



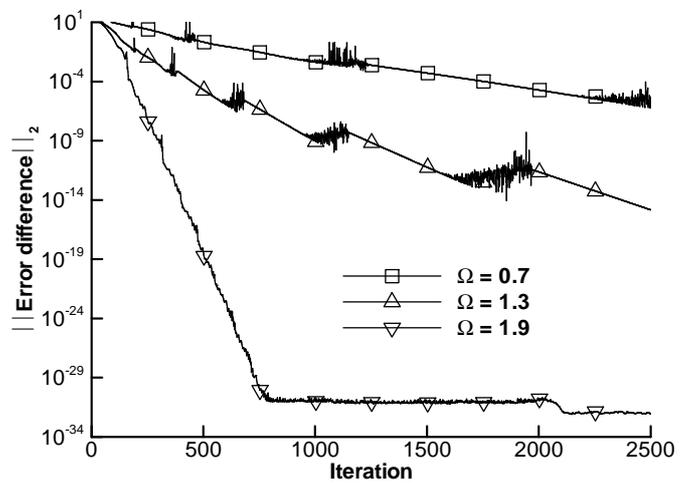
a) Error, error difference and residual



b) Error difference

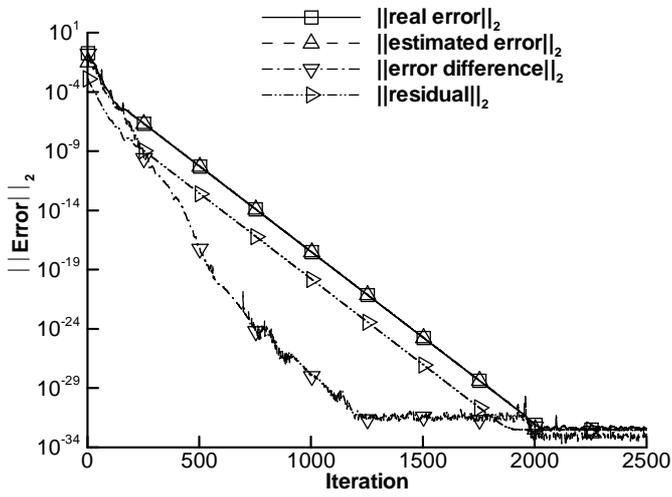


c) Comparison of error estimation methods

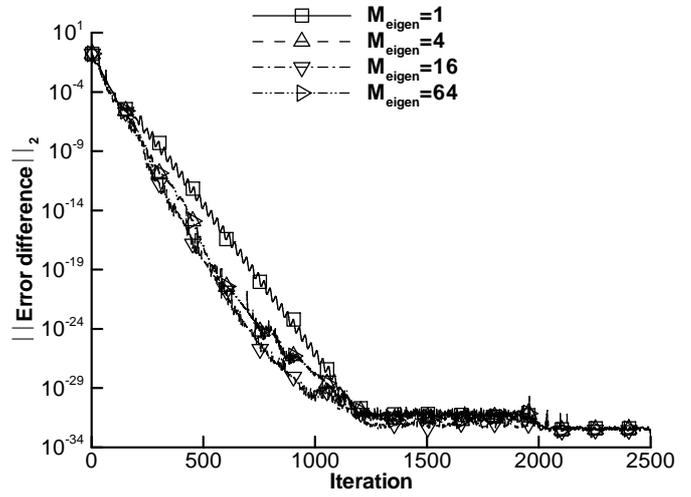


d) Change of error with relaxation factor

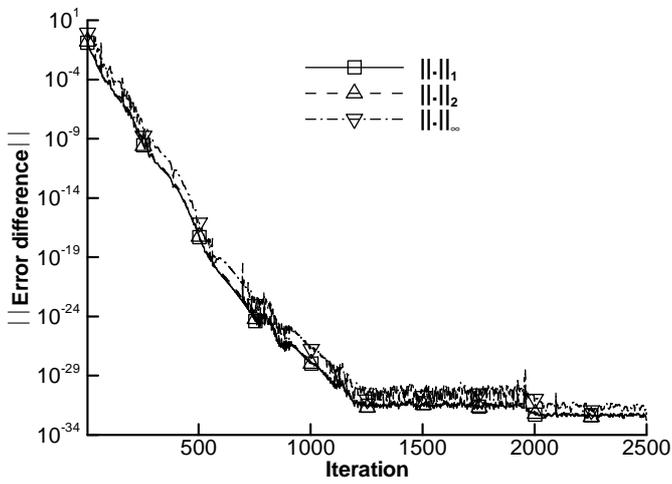
Fig. 1 Convergence error estimation for the Laplace equation.



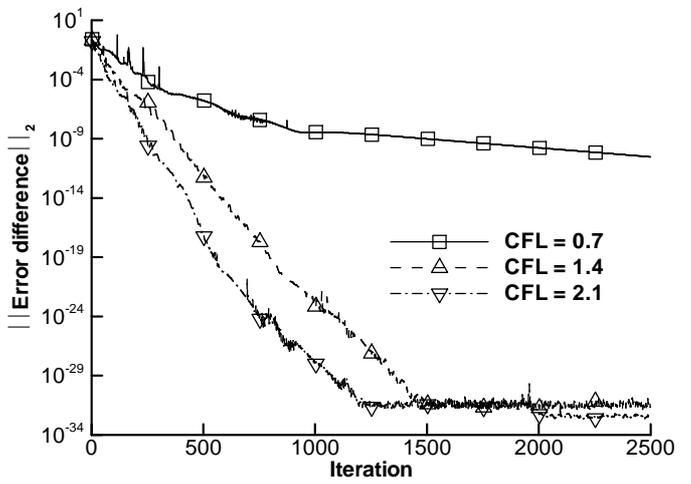
a) Error, error difference and residual



b) Error difference

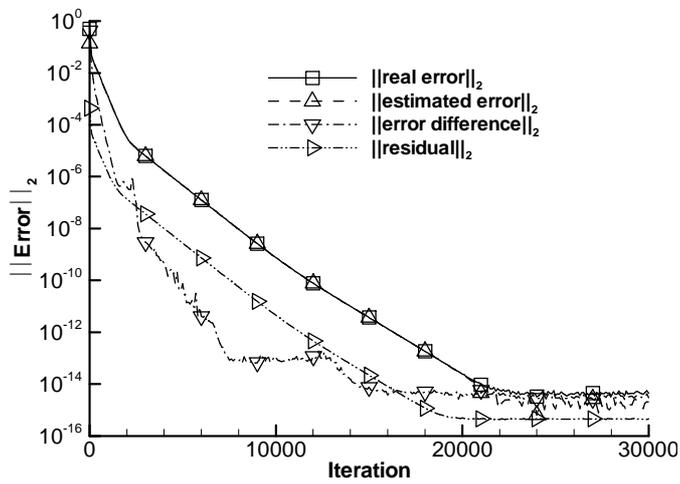


c) Different norms of error

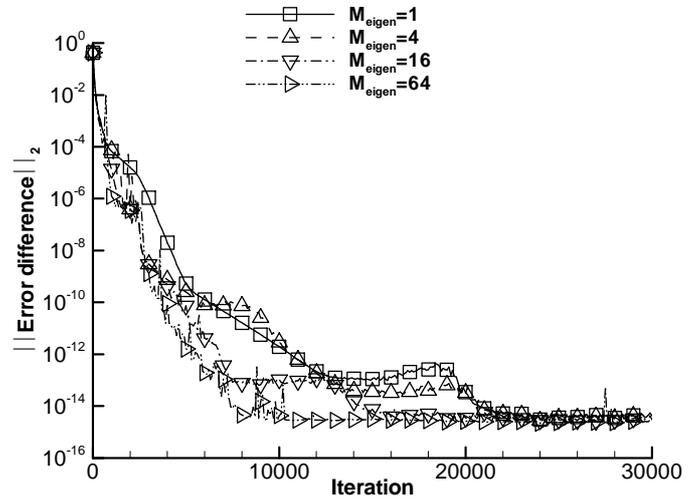


d) Change of error with CFL number

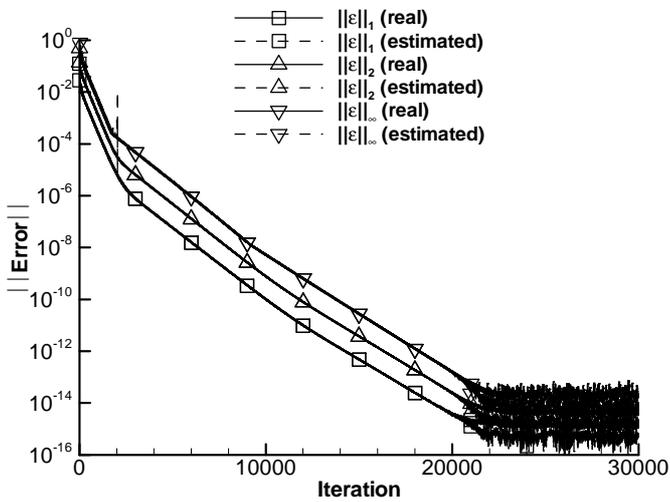
Fig. 2 Convergence error estimation for the Euler equations.



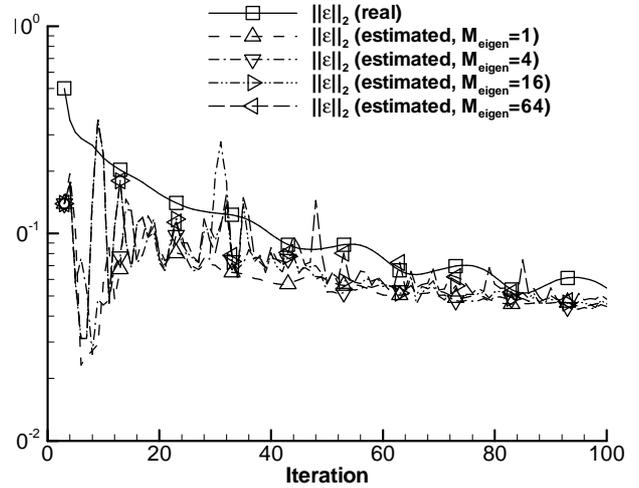
a) Error, error difference and residual



b) Error difference



c) Different norms of error



d) Initial error

Fig. 3 Convergence error estimation for the Navier-Stokes equations.

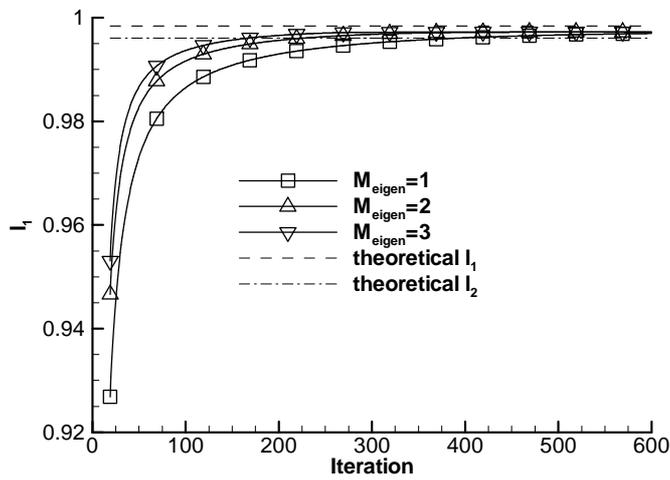
The theory used for convergence error estimation is tested for eigenvalue analysis. Both the absolute value and phase angle of eigenvalues are evaluated. Different numbers of eigenvalues are used. First the eigenvalue analysis is performed in the iterative solution of the Laplace equation. The equation is solved using the Gauss-Seidel method. Theoretical studies show that the magnitude of the largest and the second largest eigenvalues in this method can be estimated with the following equations:

$$l_1 = 1 - \pi^2 h^2, \quad l_2 = 1 - \frac{5}{2} \pi^2 h^2 \quad (40)$$

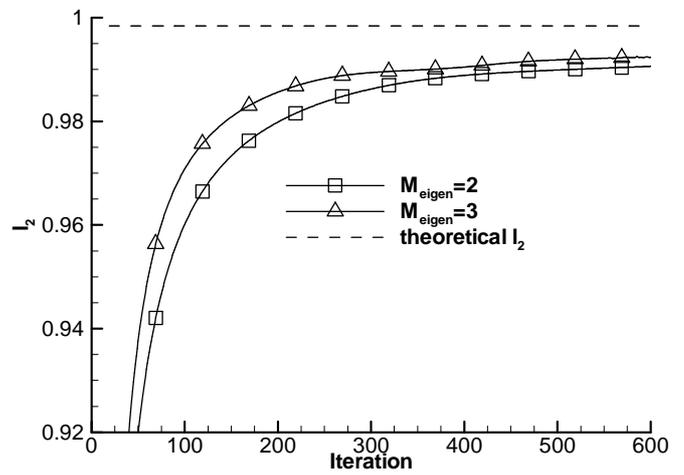
In the estimation of the largest eigenvalue, the performance of the method is given in Figure 4a. The number of eigenvalues is chosen as 1, 2 and 3. As shown in the figure, in all cases, the magnitude of the largest eigenvalue converges to the same value. The converged values are in between the theoretical values of the largest and the second largest eigenvalues. In Figure 4b, the second largest eigenvalue estimated with the proposed and theoretical approaches are compared. In the calculation of the second largest eigenvalue, the number of eigenvalues is fixed as 2 and 3. The figure shows that the estimated eigenvalues approach to the theoretical value. In Figure 4c, the magnitude of the largest eigenvalue is estimated for different values of the over relaxation parameter. These calculations are performed with only one eigenvalue. The result show that as the value of the over relaxation parameter increase, the magnitude of the largest eigenvalue decrease. Since the cosine of the phase angles is one in all calculations, the phase angles of the eigenvalues are not shown.

Eigenvalues are also calculated in the iterative solution of nonlinear equations. Figure 5a shows the estimation of the largest eigenvalue with different number of eigenvalues. The result show that, in the solution of the Euler equations, the estimated largest eigenvalues converge to the same value. Compared to calculations with Laplace equations, the convergence of the largest eigenvalue with three eigenvalues shows more oscillations. This may be related to the nonlinearity of the problem. The effects of the CFL number on the magnitude of the largest eigenvalue are shown in Figure 5b. The result shows that the increase in the CFL number reduces the magnitude of the largest eigenvalue and produces some oscillations. The magnitude of the largest eigenvalue is also estimated in the solution of the Navier-Stokes equations. Figure 5c shows the result calculated with one and two eigenvalues. As in previous calculations, the magnitude of the largest eigenvalues converges to the same value in both cases.

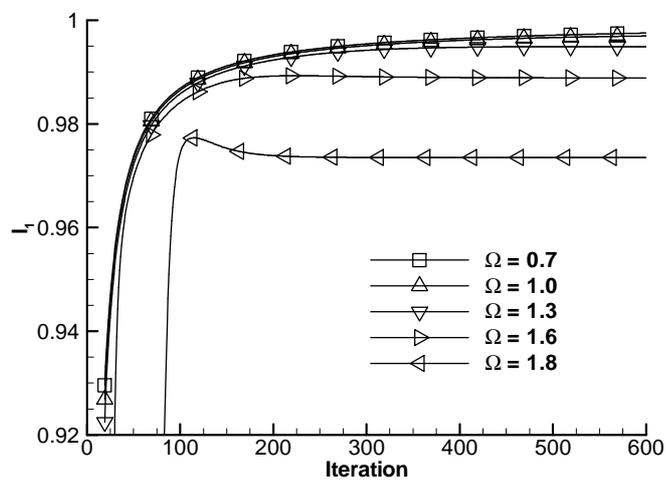
In the present study, the magnitude and the phase angles of eigenvalues are calculated by solving a system of nonlinear equations. This system is solved using, Newton's method. In Newton's method, Jacobian matrix is solved at each iteration. Result show that Newton's method do not converge at some iterations. This problem is frequently encountered when the number of eigenvalues is large. For the large number of eigenvalues, the Jacobian matrix becomes singular. In the solution of the Jacobian matrix singular value decomposition is used.



a) Largest eigenvalue

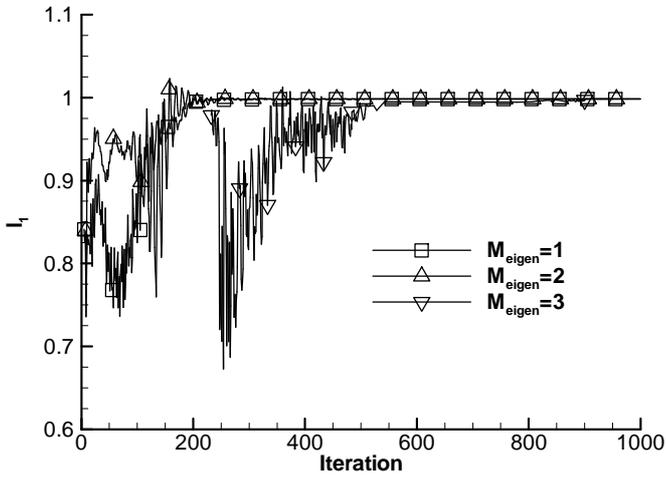


b) Second largest eigenvalue

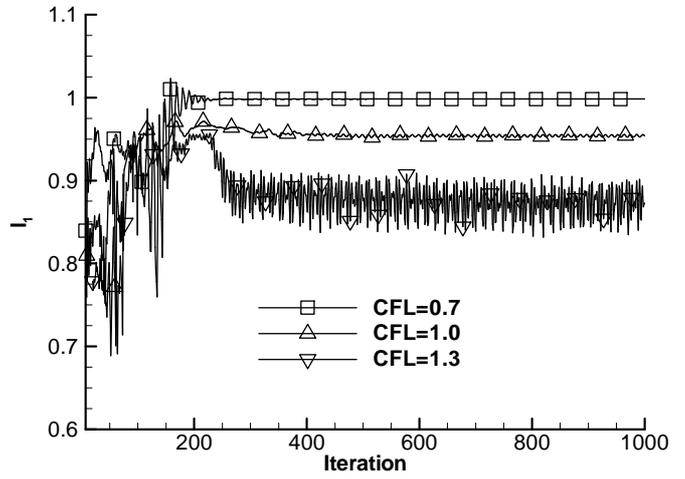


c) Largest eigenvalue

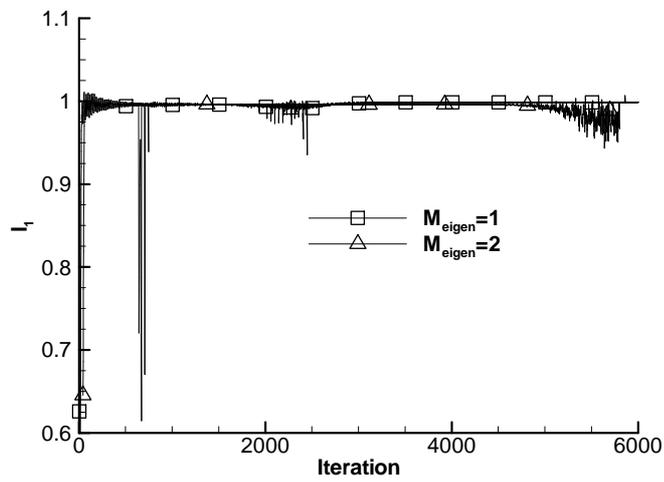
Fig. 4 Eigenvalue estimation for the Laplace equation.



a) Largest eigenvalue in Euler equations



b) Largest eigenvalue Euler equations

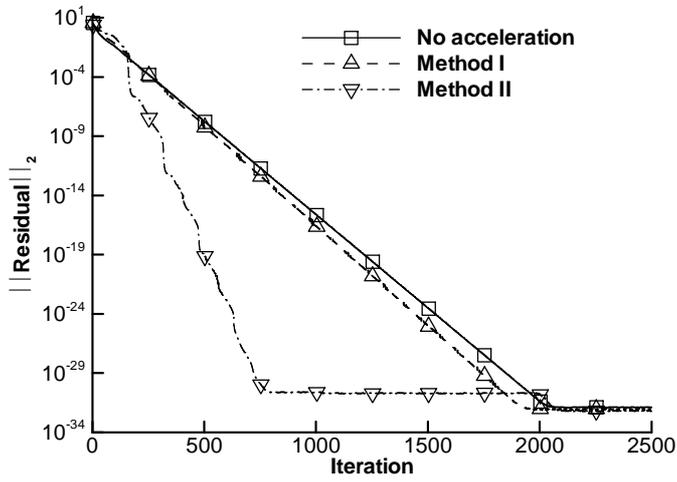


c) Largest eigenvalue in Navier-Stokes equations

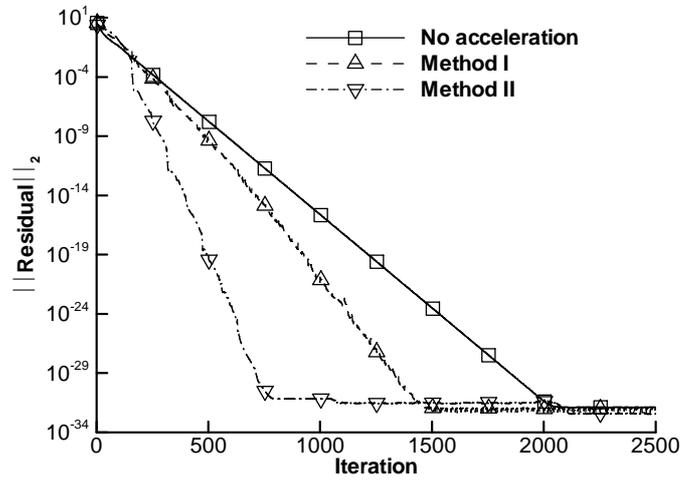
Fig. 5 Eigenvalue estimation for the Euler and Navier-Stokes equation.

The performances of the convergence acceleration methods are evaluated for both linear and nonlinear problems. Figure 6 shows the result with the linear Laplace equation. In the solution of the Laplace equation the Gauss-Seidel iterative method is used. The results are calculated with the maximum value of the over relaxation parameter. If no acceleration method is used, the norm value of residual converges to order of round of error level around 2000 iterations. If only one eigenvalue is used, Method I improves the convergence slightly. A large reduction in the number of iterations can be achieved in Method II. The norm value of residual can be reduced to the same level around 750 iterations. This is a big reduction in the number of iterations. If the number of eigenvalues is increased to four, the improvement in the convergence of Method I can be clearly seen. However, the same improvement is not observed in Method II. If the number of eigenvalues is chosen as 16, more improvements can be achieved in Method I. Further improvements can be gained if the number of eigenvalues is increased to 64. For this case, the convergences of Method I and II are almost same. Results show that with only one eigenvalue, large improvements can be achieved in the convergence of Method II. The performance of Method I is poor if only one eigenvalue is used. However, better performances can be achieved by increasing the number of eigenvalues. Since increasing the number of eigenvalues also increases the CPU time, in the solution of linear Laplace equation, the performance of Method II is better.

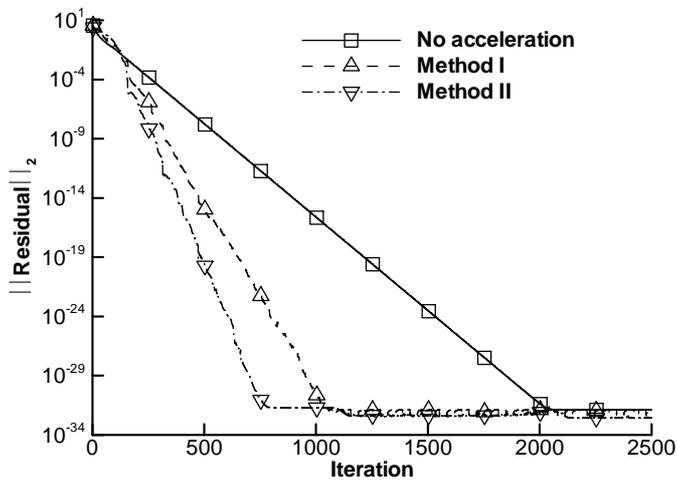
The performances of the acceleration methods for nonlinear problems are studied in the iterative solution of the Euler and Navier-Stokes equations. In these solutions, four-stage Runge-Kutta scheme, local time stepping and a multigrid method are used. The proposed convergence acceleration method is implemented at the finest grid level. Results are presented for the maximum value of the CFL number. Figure 7 shows the performances of the methods for Euler equations. As in the solution of the Laplace equation, Method II achieves a large reduction in the number of iterations with only one eigenvalue. With the original iterative algorithm, the norm value of residual can be reduced to order of round of error in 2000 iterations. With the implementation of Method II, the same residual reduction can be achieved in 1000 iterations. For the number of eigenvalues of 1, 4 and 16, the convergence of Method I is slower. The performances of both convergence acceleration methods are getting closer to each other if the number of eigenvalues is set to 64. In Method II, better convergences can be achieved with a smaller number of eigenvalues. This may be important as far as the CPU time is concerned. The results with Navier-Stokes equations are shown in Figure 8. With the original iterative algorithm, 60,000 iterations are needed to reduce the norm value of residual to order of round off error. If Method I is activated, a significant reduction in the number of iteration is achieved. The same convergence level is achieved in 6000 and 16000 iterations with a number of eigenvalues of 16 and 4, respectively. An attempt to increase the number of eigenvalues further is failed. The performance of Method II is deteriorated in the solution of Navier-Stokes equations. Compared to the convergence improvements achieved in the solution of the Laplace and Euler equations, the improvement is small in the solution of the Navier-Stokes equations. In Figure 8c, the pressure distributions of the experimental and computational results are compared. As seen from the figure, the pressure distributions are accurately matched.



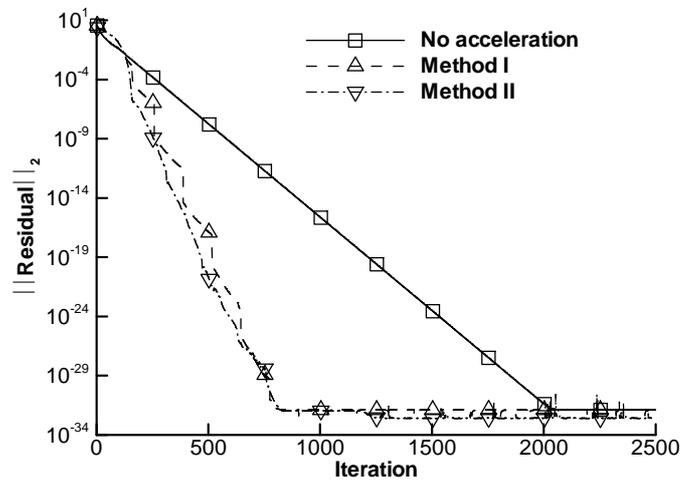
a)  $M_{\text{eigen}}=1$



b)  $M_{\text{eigen}}=4$

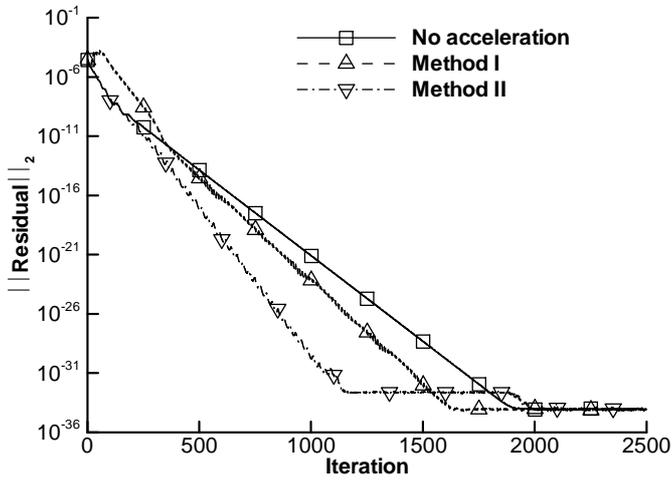


c)  $M_{\text{eigen}}=16$

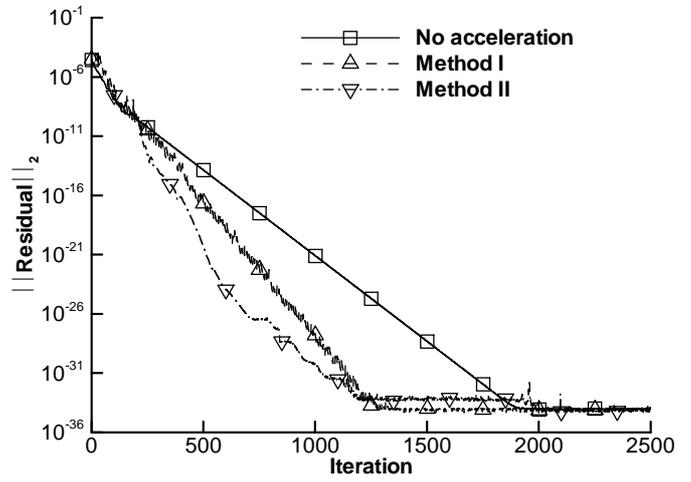


d)  $M_{\text{eigen}}=64$

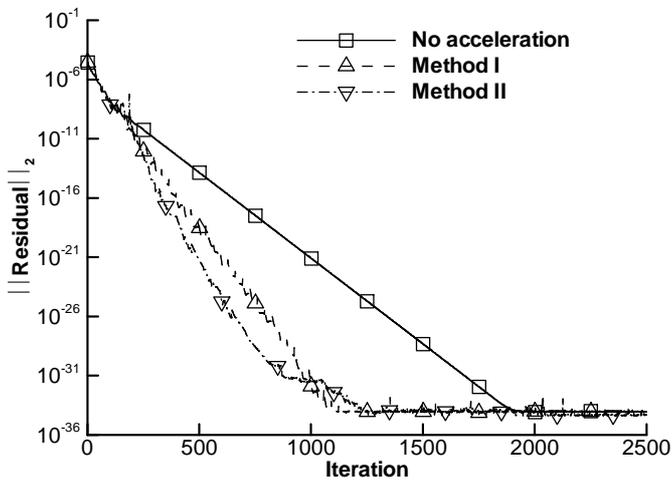
Fig. 6 Convergence acceleration for the Laplace equation.



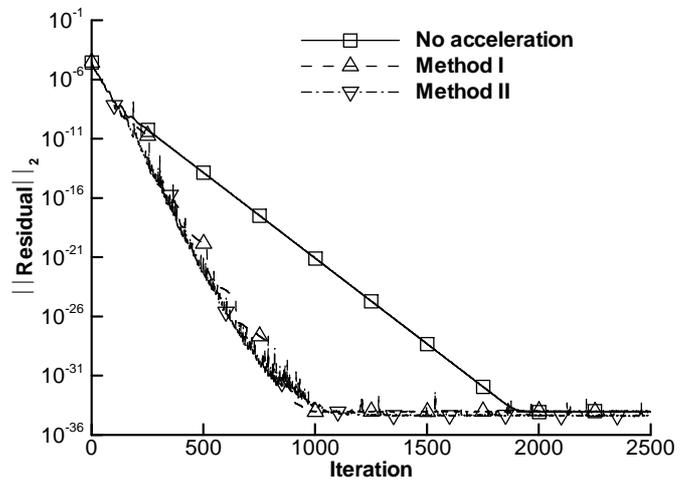
a)  $M_{\text{eigen}}=1$



b)  $M_{\text{eigen}}=4$

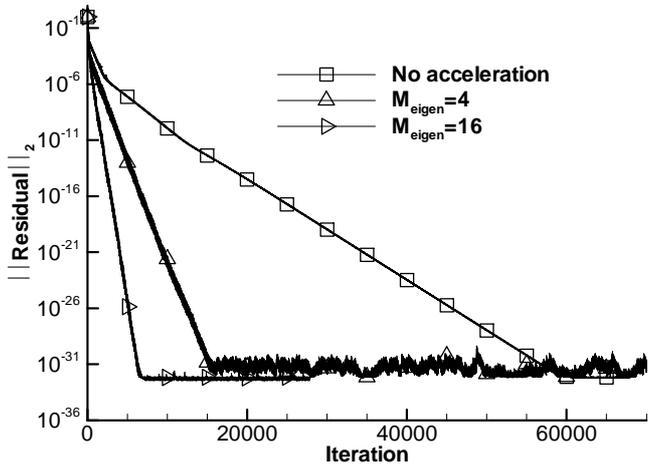


c)  $M_{\text{eigen}}=16$

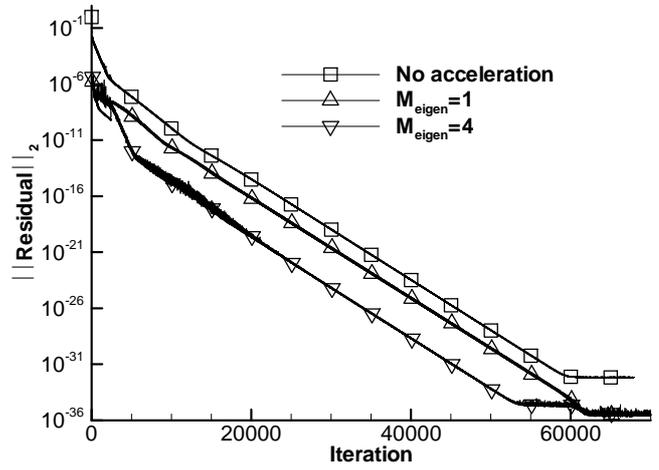


d)  $M_{\text{eigen}}=64$

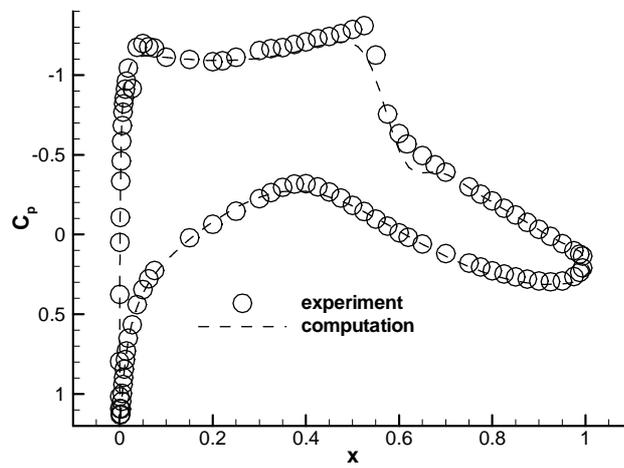
Fig. 7 Convergence acceleration for the Euler equations.



c) Method I



d) Method II



c) Validation of Navier-Stokes solution

Fig. 8 Convergence acceleration for the Navier-Stokes equations.

## 6 Conclusion and Future Work

In iteratively solved problems, new methods are developed to estimate the convergence error and to accelerate the convergence. The convergence error estimation method is based on the eigenvalue analysis of linear systems, the accuracy of which is verified for both linear and nonlinear problems. The results show that the convergence error can be accurately estimated with the developed method. The residual itself, on the other hand, is not considered to be a reliable parameter to predict the convergence error. The performance of the proposed method on the estimation of the magnitude of eigenvalues is also studied. Results show that the developed method can be used to estimate the magnitude of the eigenvalues. The use of Newton's method in eigenvalue analysis may cause some problems. These problems may be related to having singular Jacobian matrices. Two methods are developed for convergence acceleration. The convergence acceleration is achieved by subtracting the convergence error from iterative solution. The acceleration methods show good performances in both linear and nonlinear problems.

## Acknowledgments

This research is supported by The Scientific and Technological Research Council of Turkey (TUBITAK).

## References

- [1] Ferziger, J. H. and Peric, M., "Further Discussion of Numerical Errors in CFD," *International Journal for Numerical Methods in Fluids*, Vol. 23, 1996, pp. 1263-1274.
- [2] Ferziger, J. H. and Peric, M., *Computational Methods for Fluid Dynamics*, Springer, Berlin, 2002.
- [3] Bergstrom, J. and Gebart, R., "Estimation of Numerical Accuracy for the Flow Field in a Draft Tube," *International Journal of Numerical Methods for Heat and Fluid Flow*, Vol. 9, 1999, pp. 472-486.
- [4] Roy, C. J., McWherter-Payne, M. A., and Oberkampf, W. L., "Verification and Validation for Laminar Hypersonic Flowfields," Part1: Verification, *AIAA Journal*, Vol. 21, 2003, pp. 1934-1943.
- [5] Alekseev, A., "An Adjoint-Based A Posteriori Estimation of Iterative Convergence Error," *Computers & Mathematics with Applications*, Vol. 52, 2006, pp. 1205-1212.
- [6] Brezinski, C., "Error Estimates for the Solution of Linear Systems," *SIAM Journal on Scientific Computing*, Vol. 21, 1999, pp. 764-781.
- [7] Eyi, S., "Finite-Difference Sensitivity Calculation in Iteratively Solved Problems," 20th AIAA Computational Fluid Dynamics Conference, AIAA Paper 2011-3071, 27 - 30 June 2011, Honolulu, Hawaii.
- [8] Hafez, M., Parlette E. and Salas, M., "Convergence acceleration of iterative solutions of Euler equations for transonic flow computations," *Computational Mechanics*, Vol. 1, 1986, pp. 165-176.
- [9] Jespersen, D. C., and P. G. Buning, "Accelerating an Iterative Process by Explicit Annihilation," *SIAM Journal of Sci. Stat. Comput.*, Vol. 6, 1985, pp. 639-651.
- [10] Dagan, A., "A Convergence Accelerator of a Linear System of Equations Based upon the Power Method," *International Journal for Numerical Methods in Fluids*, Vol. 35, 2001, pp. 721-741.
- [11] Sidi, A., "Review of two vector extrapolation methods of polynomial type with applications to large-scale problems," *Journal of Computational Science* Vol. 3, 2012, pp. 92-101.
- [12] Sidi, A., "Vector extrapolation methods with applications to solution of large systems of equations and to PageRank computations," *Computers and Mathematics with Applications*, Vol. 56, 2008, pp. 1-24.