Journal of Computational and Applied Mathematics 22 (1988) 35–61
North-Holland

Convergence and stability analyses for some vector extrapolation methods in the presence of defective iteration matrices

Avram SIDI
Computer Science Department, Technion — Israel Institute of Technology, Haifa 32000, Israel

Jacob BRIDGER *
Mathematics Department, Technion — Israel Institute of Technology, Haifa 32000, Israel

Received 25 January 1987
Revised 30 November 1987

Abstract: In two previous papers [10,11] convergence and stability results for the following vector extrapolation methods were presented: Minimal Polynomial Extrapolation, Reduced Rank Extrapolation, Modified Minimal Polynomial Extrapolation, and Topological Epsilon Algorithm. The analyses were carried out for vector sequences that include those arising from iterative methods for linear systems of equations having diagonalizable iteration matrices. In this paper the analyses of [10,11] are extended to vector sequences that include those arising from iterative methods for linear systems having defective iteration matrices. The results are illustrated with numerical examples. The analyses above naturally suggest some old and some new extensions of the well known power method, enabling one to obtain estimates for several dominant eigenvalues of a general matrix.

1. Introduction

Let $\boldsymbol{B}$ be a normed linear space over the field of complex numbers, and denote the norm associated with $\boldsymbol{B}$ by $\|\cdot\|$. In case $\boldsymbol{B}$ is also an inner product space, we adopt the following convention for the homogeneity property of the inner product: For y, $z \in \boldsymbol{B}$ and α, β complex numbers, the inner product $(\cdot, \cdot)$ satisfies $(\alpha y, \beta z) = \bar{\alpha}\beta(y, z)$. The norm in this case is the one induced by the inner product, i.e., if $x \in \boldsymbol{B}$, $\|x\| = \sqrt{(x, x)}$.

Let x_i, $i = 0, 1, \ldots,$ be a sequence in $\boldsymbol{B}$. We shall assume that

$$x_m \sim s + \sum_{i=1}^{\infty} P_i(m)\lambda_i^m \quad \text{as } m \to \infty. \tag{1.1}$$

* Work of this author was performed in partial fulfilment of the requirements for the degree of Master of Science at the Technion-Israel Institute of Technology, Haifa, Israel.

Here s is a vector in $\boldsymbol{B}$ and λ_i are scalars ordered such that

$$|\lambda_1| \geqslant |\lambda_2| \geqslant |\lambda_3| \geqslant \cdots, \tag{1.2}$$

and satisfying $\lambda_i \neq 0$ and $\lambda_i \neq 1$, $i = 1, 2, \ldots$, and $\lambda_i \neq \lambda_j$ if $i \neq j$. In addition, we assume that there can be only a finite number of λ_i having the same modulus. $P_i(m)$ are polynomials in m with vector coefficients (thus $P_i(m)$ are vectors in $\boldsymbol{B}$), which we write in the form

$$P_i(m) = \sum_{l=0}^{p_i} y_{il} \binom{m}{l}, \tag{1.3}$$

where $\binom{m}{l}$ are binomial coefficients and y_{il}, $l = 0, \ldots, p_i$, $i = 1, 2, \ldots$, form a linearly independent set of vectors. We agree to order the λ_i such that if $|\lambda_j| = |\lambda_{j+1}|$, then $p_j \geqslant p_{j+1}$. The meaning of (1.1) is that for any positive integer N there exist a positive constant K and a positive integer m_0 that depend only on N, such that for every $m \geqslant m_0$,

$$\left\| x_m - s - \sum_{i=1}^{N-1} P_i(m) \lambda_i^m \right\| \leqslant K \lambda_N^m m^{p_N}. \tag{1.4}$$

If $|\lambda_1| < 1$, then $\lim_{m \to \infty} x_m$ exists and is simply s. If $|\lambda_1| \geqslant 1$, then $\lim_{m \to \infty} x_m$ does not exist, and s, in this case, is said to be the anti-limit of the sequence x_m, $m = 0, 1, \ldots$.

As will be shown in Section 2, sequences of vectors generated by iterative solution of linear systems of equations having defective iteration matrices are exactly of the form described above. In fact, this has been the source of motivation for the assumptions above.

Our aim is to find a good approximation to s from a small number of terms of the sequence x_m, $m = 0, 1, \ldots$, whether s is the limit or the anti-limit of this sequence. To this effect several vector extrapolation methods have been proposed. In a recent work by Smith, Ford, and Sidi [12] some of these methods have been surveyed and tested numerically. The methods that have been considered in [12] are the Minimal Polynomial Extrapolation (MPE) of Cabay and Jackson [3], the Reduced Rank Extrapolation (RRE) of Eddy [4] and Mešina [8], the Scalar Epsilon Algorithm (SEA) of Wynn [15], the Vector Epsilon Algorithm (VEA) of Wynn [16], and the Topological Epsilon Algorithm (TEA) of Brezinski [1]. In yet another work by Sidi, Ford, and Smith [11] a new method designated the Modified MPE (MMPE) has been proposed. Four of the methods above, namely, MPE, RRE, MMPE, and TEA have been analyzed in Sidi [10] and in [11] for their convergence and stability properties. Their analyses have been carried out for sequences of the form (1.1) with $p_i = \deg P_i(m) = 0$ for all i.

For future reference we will now give a brief description of the above mentioned extrapolation methods based on the developments in [11] for MMPE and TEA and in [10] for MPE and RRE.

Below k denotes a positive integer less than or equal to the dimension of the space B and $u_m = \Delta x_m = x_{m+1} - x_m$, $w_m = \Delta u_m = u_{m+1} - u_m$, $m = 0, 1, \ldots$. Also $s_{n,k}$ denotes the approximation to s obtained by applying any of the methods above to the vector sequence x_m, $m = 0, 1, \ldots$. Clearly, $s_{n,k}$ will be different for each method. For each method $s_{n,k}$ can be shown to be of the form

$$s_{n,k} = \sum_{i=0}^{k} \gamma_i x_{n+i}, \tag{1.5}$$

subject to

$$\sum_{i=0}^{k} \gamma_i = 1. \tag{1.6}$$

It can also be shown that $s_{n,k}$ has the determinant representation

$$s_{n,k}=\frac{D(x_n,\ x_{n+1},\ldots,x_{n+k})}{D(1,\ 1,\ldots,1)}, \tag{1.7}$$

where

$$D(\sigma_0,\ \sigma_1,\ldots,\sigma_k)=\begin{vmatrix} \sigma_0 & \sigma_1 & \cdots & \sigma_k \\ u_{0,0} & u_{0,1} & \cdots & u_{0,k} \\ u_{1,0} & u_{1,1} & \cdots & u_{1,k} \\ \vdots & \vdots & & \vdots \\ u_{k-1,0} & u_{k-1,1} & \cdots & u_{k-1,k} \end{vmatrix}, \tag{1.8}$$

with u_{ij} scalars dependent on the extrapolation method being used. If we let N_i be the cofactor of σ_i in the first row of $D(\sigma_0,\ldots,\sigma_k)$, then

$$D(\sigma_0,\ \sigma_1,\ldots,\sigma_k)=\sum_{i=0}^{k}\sigma_i N_i \tag{1.9}$$

when σ_i are scalars. When σ_i are vectors (1.9) is taken to be the interpretation of $D(\sigma_0,\ldots,\sigma_k)$. The computation of the u_{ij} for the different methods is explained below.

(1) For MMPE

$$u_{ij}=Q_{i+1}(u_{n+j}), \tag{1.10}$$

where $Q_1,\ldots,Q_k$ form a linearly independent set of bounded linear functionals over $\boldsymbol{B}$.

(2) For TEA

$$u_{ij}=Q(u_{n+i+j}), \tag{1.11}$$

where Q is a bounded linear functional over $\boldsymbol{B}$.

(3) For MPE

$$u_{ij}=(u_{n+i},\ u_{n+j}). \tag{1.12}$$

(4) For RRE

$$u_{ij}=(w_{n+i},\ u_{n+j}). \tag{1.13}$$

For all four methods the γ_j satisfy the system of linear equations consisting of (1.6) and

$$\sum_{j=0}^{k}\gamma_j u_{ij}=0,\quad 0\leqslant i\leqslant k-1. \tag{1.14}$$

For more details the reader is referred to [10] and [11].

An extrapolation method almost identical to RRE has been proposed by Kaniel and Stein [7]. In this method

$$s_{n,k}=\sum_{i=0}^{k}\gamma_i x_{n+i+1}, \tag{1.15}$$

where the γ_i are determined exactly as for RRE. Actually, as suggested in [11], one can consider applying all the methods above in the form

$$s_{n,k}=\sum_{i=0}^{k}\gamma_i x_{n+q+i},\quad \text{some fixed } q\geqslant 0, \tag{1.16}$$

where γ_i are determined exactly as before, i.e., $s_{n,k}$ has the determinant representation

$$s_{n,k} = \frac{D(x_{n+q}, x_{n+q+1}, \ldots, x_{n+q+k})}{D(1, 1, \ldots, 1)}, \tag{1.17}$$

c.f. (1.5) and (1.7). Note that the determination of the γ_i for MMPE, MPE, and RRE involves $x_n, x_{n+1}, \ldots, x_{n+k+1}$. This suggests that for computational economy $q \leqslant 1$. For TEA, on the other hand, the γ_i are determined from $x_n, x_{n+1}, \ldots, x_{n+2k}$, which suggests that $q \leqslant k$. Another TEA corresponding to the choice $q = k$ was given in [1, p. 345].

In [10] and [11] it was shown, under the assumption that $p_i = \deg P_i(m) = 0$ and $|\lambda_k| > |\lambda_{k+1}|$ and under additional mild assumptions, that

$$\| s_{n,k} - s \| = O(\lambda_{k+1}^n) \quad \text{as } n \to \infty, \tag{1.18}$$

where $s_{n,k}$ for all four methods is as given in (1.5). In addition, it was shown that the methods are asymptotically stable in the sense that

$$\sup_n \sum_{i=0}^{k} |\gamma_i^{(n,k)}| < \infty \tag{1.19}$$

(we have denoted the γ_i by $\gamma_i^{(n,k)}$ to show their dependence on n and k). In fact, it was shown that

$$\lim_{n \to \infty} \sum_{i=0}^{k} \gamma_i^{(n,k)} \lambda^i = \prod_{i=1}^{k} \left(\frac{\lambda - \lambda_i}{1 - \lambda_i} \right). \tag{1.20}$$

In Section 3 of this work we state the extensions of the results (1.18)–(1.20) to the case of arbitrary p_i. The proofs are carried out in Section 5. In Section 4 we illustrate the results of Section 3 with numerical examples. The analyses of Section 5 naturally suggest some old and some new extensions of the well known power method that enable one to obtain estimates for several dominant eigenvalues of a general matrix. These extensions are considered in Section 6.

2. Example: linear iterative methods with defective matrices

Let us consider a vector sequence generated by a matrix iterative technique used in solving the linear system of equations

$$x = Ax + b, \tag{2.1}$$

where A is a general and possibly defective $M \times M$ (complex) matrix, and b and x are M dimensional (complex) column vectors. We note that if $\lambda = 1$ is an eigenvalue of A, then the system (2.1) does not have a unique solution. Thus, we assume that all eigenvalues of A are different than 1. Under this assumption, the unique solution vector s satisfies

$$s = As + b. \tag{2.2}$$

For a given vector x_0, we generate the vectors x_m by

$$x_{m+1} = Ax_m + b, \quad m = 0, 1, \ldots . \tag{2.3}$$

From (2.2) and (2.3) we obtain

$$x_m - s = A^m(x_0 - s). \tag{2.4}$$

For any $M \times M$ matrix A we can find a non-singular matrix V such that

$$V^{-1}AV = J = \begin{bmatrix} J_1 & & & \mathbf{0} \\ & J_2 & & \\ & & \ddots & \\ \mathbf{0} & & & J_\nu \end{bmatrix}, \tag{2.5}$$

where the Jordan blocks J_i are of dimension r_i for some integers r_i, and have the form

$$J_i = \begin{bmatrix} \lambda_i & 1 & & \mathbf{0} \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ \mathbf{0} & & & \lambda_i \end{bmatrix}_{r_i \times r_i}, \quad \lambda_i \text{ eigenvalue}. \tag{2.6}$$

It can be shown (see Varga [13]) that

$$J_i^m = \begin{bmatrix} \lambda_i^m & \binom{m}{1}\lambda_i^{m-1} & \cdots & \binom{m}{r_i - 2}\lambda_i^{m-r_i+2} & \binom{m}{r_i - 1}\lambda_i^{m-r_i+1} \\ & \lambda_i^m & \binom{m}{1}\lambda_i^{m-1} & \cdots & \binom{m}{r_i - 2}\lambda_i^{m-r_i+2} \\ & & \lambda_i^m & \ddots & \vdots \\ & & & \ddots & \binom{m}{1}\lambda_i^{m-1} \\ \mathbf{0} & & & & \lambda_i^m \end{bmatrix}, \tag{2.7}$$

where by convention $\binom{m}{j} = 0$ if $j > m$. If we denote the columns of the matrix V by $v_{11}, v_{12}, \ldots, v_{1r_1}, v_{21}, v_{22}, \ldots, v_{2r_2}, \ldots, v_{\nu 1}, v_{\nu 2}, \ldots, v_{\nu r_\nu}$, then v_{j1} is the eigenvector corresponding to the eigenvalue λ_j and v_{ji}, $i = 2, \ldots, r_j$, are the principal vectors corresponding to the same eigenvalue. As is known, the set of the eigenvectors and principal vectors is linearly independent and forms a basis for $\mathbb{C}^M$.

For the initial error vector there exist scalars a_{ji} such that

$$x_0 - s = \sum_{j=1}^{\nu} \sum_{i=1}^{r_j} a_{ji} v_{ji}. \tag{2.8}$$

Here we have assumed that Jordan blocks that do not contribute to (2.8) have been removed, and the remaining blocks renumbered.

Define p_j to be the largest nonnegative integer for which $a_{j,p_j+1} \neq 0$. Consequently, $p_j + 1 \leqslant r_j$. By (2.4), (2.5), (2.7), (2.8) and the fact that $A^m V = VJ^m$, it follows that

$$x_m - s = \sum_{j=1}^{\nu} \sum_{i=1}^{p_j+1} a_{ji} \sum_{l=1}^{i} \binom{m}{i-l} \lambda_j^{m-i+l} v_{jl}. \tag{2.9}$$

We first observe that if 0 is an eigenvalue of A, then for m sufficiently large, this eigenvalue does not contribute to the triple sum in (2.9). Thus it can be assumed that $\lambda_j \neq 0$ for all j.

Changing the index l to $q=i-l$, and interchanging the summations over i and q, (2.9) becomes

$$x_m - s = \sum_{j=1}^{\nu} \left[\sum_{q=0}^{p_j} y_{jq} \binom{m}{q} \right] \lambda_j^m, \tag{2.10}$$

where

$$y_{jq} = \left(\sum_{i=q+1}^{p_j+1} a_{ji} v_{j,i-q} \right) \lambda_j^{-q}, \quad 0 \leqslant q \leqslant p_j. \tag{2.11}$$

By the fact that $a_{j,p_j+1} \neq 0$ and the linear independence of the set of vectors v_{ji}, $1 \leqslant i \leqslant p_j + 1$, it follows that y_{jq}, $0 \leqslant q \leqslant p_j$, form a linearly independent set as well. Thus, the set of vectors y_{jq}, $0 \leqslant q \leqslant p_j$, $1 \leqslant j \leqslant \nu$, is linearly independent.

Assume now that there are several Jordan blocks that have the same eigenvalue. For the sake of argument suppose that $\lambda = \lambda_1 = \lambda_2 = \cdots = \lambda_N \neq \lambda_{N+1}$, and that $p = p_1 \geqslant p_2 \geqslant \cdots \geqslant p_N$. Then the part of the double sum in (2.10) having $j = 1, 2, \ldots, N$, can be expressed as

$$\sum_{j=1}^{N} \sum_{q=0}^{p_j} y_{jq} \binom{m}{q} \lambda_j^m = \left[\sum_{q=0}^{p} \hat{y}_q \binom{m}{q} \right] \lambda^m, \tag{2.12}$$

where

$$\hat{y}_q = \sum_{j=1}^{N} y_{jq}, \quad 0 \leqslant q \leqslant p, \tag{2.13}$$

and $y_{jq} = 0$ when $q > p_j$. It is easy to see that the vectors $\hat{y}_q$, $0 \leqslant q \leqslant p$, form a linearly independent set. We thus have shown that Jordan blocks having the same eigenvalue can be combined into a single block, enabling us to assume that all λ_j in (2.10) are distinct and that the y_{jq} form a linearly independent set.

In summary, we have shown that if the vector sequence x_m, $m = 0, 1, 2, \ldots$, is generated by the matrix iterative process (2.3) and if $(I - A)^{-1}$ exists, then x_m automatically obeys (1.1) in conjunction with all the conditions imposed on the scalars λ_j and the vectors y_{jl}.

3. Statement of convergence and stability results

In accordance with the assumptions of Section 1, let the positive integers t and r be such that

$$|\lambda_t| > |\lambda_{t+1}| = \cdots = |\lambda_{t+r}| > |\lambda_{t+r+1}|. \tag{3.1}$$

Now from (3.1) and the ordering $p_{t+1} \geqslant \cdots \geqslant p_{t+r}$ it follows that there is a greatest integer r' ($r' \leqslant r$), for which

$$p_{t+1} = \cdots = p_{t+r'}. \tag{3.2}$$

Obviously, $r' = r$ when $r = 1$ or $p_{t+r} = p_{t+1}$. Let

$$k = \sum_{j=1}^{t} (p_j + 1). \tag{3.3}$$

Theorem 3.1. *If $s_{n,k}$ is as given in* (1.16) (*equivalently* (1.17)), *then*

$$s_{n,k} - s = \Gamma(n) n^{p_{t+1}} |\lambda_{t+1}|^n, \tag{3.4}$$

where

$$\sup_n \| \Gamma(n) \| < \infty, \tag{3.5}$$

provided

$$Z = \begin{vmatrix} Q_1(y_{10}) & \cdots & Q_1(y_{1p_1}) & \cdots & Q_1(y_{t0}) & \cdots & Q_1(y_{tp_t}) \\ Q_2(y_{10}) & \cdots & Q_2(y_{1p_1}) & \cdots & Q_2(y_{t0}) & \cdots & Q_2(y_{tp_t}) \\ \vdots & & \vdots & & \vdots & & \vdots \\ Q_k(y_{10}) & \cdots & Q_k(y_{1p_1}) & \cdots & Q_k(y_{t0}) & \cdots & Q_k(y_{tp_t}) \end{vmatrix} \neq 0 \tag{3.6}$$

for MMPE, or

$$\prod_{j=1}^{t} Q(y_{jp_j}) \neq 0 \tag{3.7}$$

for TEA. For MPE and RRE there are no additional restrictions.

The reason that MPE and RRE need no additional restrictions arises from the fact that the Gram determinant of the vectors y_{jp}, $0 \leqslant p \leqslant p_j$, $1 \leqslant j \leqslant t$, is nonzero, which follows from their linear independence.

The vector $\Gamma(n)$ for each method has the asymptotic form

$$\Gamma(n) = \sum_{j=1}^{r'} \Gamma_j \, e^{in \arg \lambda_{t+j}} + o(1) \quad \text{as } n \to \infty,$$

where Γ_j are vectors independent of n and dependent on the method used.

For MPE and RRE the vectors Γ_j, $1 \leqslant j \leqslant r'$, turn out to be identical. For complete details see Section 5.

Theorem 3.2. *If $s_{n,k}$ is as given in* (1.16) (*equivalently* (1.17)), *then, under the conditions of Theorem* 3.1, *all four methods are asymptotically stable in the sense that*

$$\sup_n \sum_{i=0}^{k} \left| \gamma_i^{(n,k)} \right| < \infty. \tag{3.8}$$

In fact,

$$\lim_{n \to \infty} \sum_{i=0}^{k} \gamma_i^{(n,k)} \lambda^i = \prod_{i=1}^{t} \left(\frac{\lambda - \lambda_i}{1 - \lambda_i} \right)^{p_i + 1}, \tag{3.9}$$

and (3.8) *is a consequence of* (3.9).

Actually, it is true that

$$\gamma_i^{(n,k)} = \delta_i + O(n^{\alpha} | \lambda_{t+1}/\lambda_t |^n) \quad \text{as } n \to \infty, \quad \text{where } \sum_{i=0}^{k} \delta_i \lambda^i = \prod_{i=1}^{t} \left(\frac{\lambda - \lambda_i}{1 - \lambda_i} \right)^{p_i + 1}.$$

and α is an integer greater than or equal to p_{t+1}. $\alpha = p_{t+1}$ if $p_i = 0$ for all λ_i whose moduli are $|\lambda_t|$. We do not give the details of the proof of this statement, although they can be extracted from the proofs of Section 5.

Remarks. (1) As can be seen from (3.4) and (3.5), all four methods are bona fide acceleration methods in the sense that

$$\frac{\| s_{n,k} - s \|}{\| x_{n+q+k} - s \|} = O\left(n^{p_{t+1} - p_1} \left| \frac{\lambda_{t+1}}{\lambda_1} \right|^n \right) \quad \text{as } n \to \infty. \tag{3.10}$$

Consequently, if λ_{t+1} is well separated from λ_1, and $|\lambda_{t+1}| < 1$, then $s_{n,k}$ converges to s much more quickly than x_n itself as $n \to \infty$, irrespective of whether the sequence $x_0, x_1, x_2, \ldots,$ converges or diverges.

(2) As is suggested by (3.4), if $p_{t+1} > 0$, then the quality of $s_{n,k}$ will deteriorate initially, but will improve for increasing n, provided $|\lambda_{t+1}| < 1$.

(3) Inspection of the leading asymptotic behavior of $\Gamma(n)$ for the different methods shows that if some of $\lambda_1, \ldots, \lambda_t$ are close to 1, then $\| \Gamma(n) \|$ is large, and this has an adverse effect on the accuracy of $s_{n,k}$. When the sequence $x_0, x_1, x_2, \ldots,$ is obtained from iterative solution of a linear system of equations, the closeness of some of $\lambda_1, \ldots, \lambda_t$ to 1 means that the matrix of the system is nearly singular.

(4) When some of $\lambda_1, \ldots, \lambda_t$ are close to 1, Theorem 3.2 implies that the $\gamma_i^{(n,k)}$ will be large in modulus although $\sum_{i=0}^{k} \gamma_i^{(n,k)} = 1$. This causes $\sum_{i=0}^{k} |\gamma_i^{(n,k)}|$ to be very large, which, in turn, causes errors in the x_i to be magnified severely.

The conclusions above are identical to those derived in [10,11] for the case $p_i = 0$, $i = 1, 2, \ldots$.

Note: If we assume that the sequence x_m, $m = 0, 1, \ldots,$ considered above is a scalar sequence, i.e., the polynomials $P_i(m)$ in (1.1) have scalar coefficients, then Theorems 3.1 and 3.2 above are valid for the Shanks [9] transformation. To see this we only need to observe that for this case TEA reduces to the Shanks transformation by taking $u_{ij} = u_{n+i+j}$ in (1.11), i.e., by deleting Q everywhere. Of course, (3.7) is automatically satisfied since $y_{jp_j} \neq 0$, $j = 1, 2, \ldots$. Needless to say, $\| \cdot \|$, the vector norm, is replaced by $| \cdot |$, the modulus. Thus, Theorem 3.1 for the Shanks transformation generalizes the result that was given by Wynn [17] for the case $p_j = 0$, $j = 1, 2, \ldots$.

4. Numerical examples

In this section we illustrate the convergence results of Section 3 for MPE and MMPE with two examples. In both of the examples the space B is the Euclidean space of dimension $M = 12$, and the sequence $x_0, x_1, \ldots,$ is obtained by the matrix iterative method described in the first paragraph of Section 2 with the notation therein. To simplify matters the solution s to the system (2.1) is taken to be $s = (1, 1, \ldots, 1)^{\mathrm{T}}$ and b is determined by $b = s - As$. The initial vector

x_0 is taken to be zero. The matrix A is defective and is determined from another defective matrix $\tilde{J}$ that has a simple form by the similarity transformation

$$A = W^{-1}\tilde{J}W, \tag{4.1}$$

where the matrix $W = (w_{ij})$ for both examples is given by

$$w_{ij} = \begin{cases} i+j-1 & \text{if } i \neq j, \\ 100 + 10j & \text{if } i = j. \end{cases} \tag{4.2}$$

As such, W is strictly diagonally dominant so that W^{-1} can be computed numerically to very high accuracy.

Example 1. The matrix $\tilde{J}$ is given as the block diagonal matrix

$$\tilde{J} = \begin{pmatrix} C_1 & & \\ & C_2 & \\ & & C_3 \end{pmatrix}, \tag{4.3}$$

where

$$C_1 = \begin{pmatrix} 0 & 0 & 0 & -1.849 \\ 1 & 0 & 0 & -5.44 \\ 0 & 1 & 0 & -6.72 \\ 0 & 0 & 1 & -4 \end{pmatrix}, \tag{4.4}$$

$$C_2 = \begin{pmatrix} 0.6 & 1 & 0 & 0 \\ 0 & 0.6 & 1 & 0 \\ 0 & 0 & 0.6 & 1 \\ 0 & 0 & 0 & 0.6 \end{pmatrix}, \quad C_3 = \begin{pmatrix} 0.1 & 1 & 0 & 0 \\ 0 & 0.1 & 1 & 0 \\ 0 & 0 & 0.1 & 1 \\ 0 & 0 & 0 & 0.1 \end{pmatrix}. \tag{4.5}$$

Each of the matrices C_1, C_2, and C_3 is defective. C_1 is a Frobenius matrix with eigenvalues $\lambda_1 = -1 + 0.6\mathrm{i}$ and $\lambda_2 = \bar{\lambda}_1$, λ_1, and λ_2 having algebraic multiplicity 2 and geometric multiplicity 1. C_2 is a Jordan matrix with eigenvalue $\lambda_3 = 0.6$ having algebraic multiplicity 4 and geometric multiplicity 1. Similarly, C_3 is a Jordan matrix with eigenvalue $\lambda_4 = 0.1$ having algebraic multiplicity 4 and geometric multiplicity 1. Combining this information, we have, for an arbitrary initial vector x_0, the expansion

$$x_m = s + \sum_{i=1}^{4} P_i(m)\lambda_i^m, \tag{4.6}$$

with λ_i as above and $p_1 = p_2 = 1$ and $p_3 = p_4 = 3$.

In Figs. 1 and 2 we give the results of the computations for $\| s_{n,k} - s \|_\infty$ using both MPE and MMPE with $k = 4$ and $k = 8$, respectively. We do not include $\| x_n - s \|_\infty$ as the sequence $x_0, x_1, \ldots,$ diverges by $\rho(A) = |\lambda_1| > 1$. According to the theory of Section 3 we should have

$$-\log_{10} \| s_{n,4} - s \|_\infty = (-\log_{10} 0.6)n - 3 \log_{10} n + O(1) \quad \text{as } n \to \infty, \tag{4.7}$$

and

$$-\log_{10} \| s_{n,8} - s \|_\infty = (-\log_{10} 0.1)n - 3 \log_{10} n + O(1) \text{ as } n \to \infty. \tag{4.8}$$

These results are indeed born out by Figs. 1 and 2.

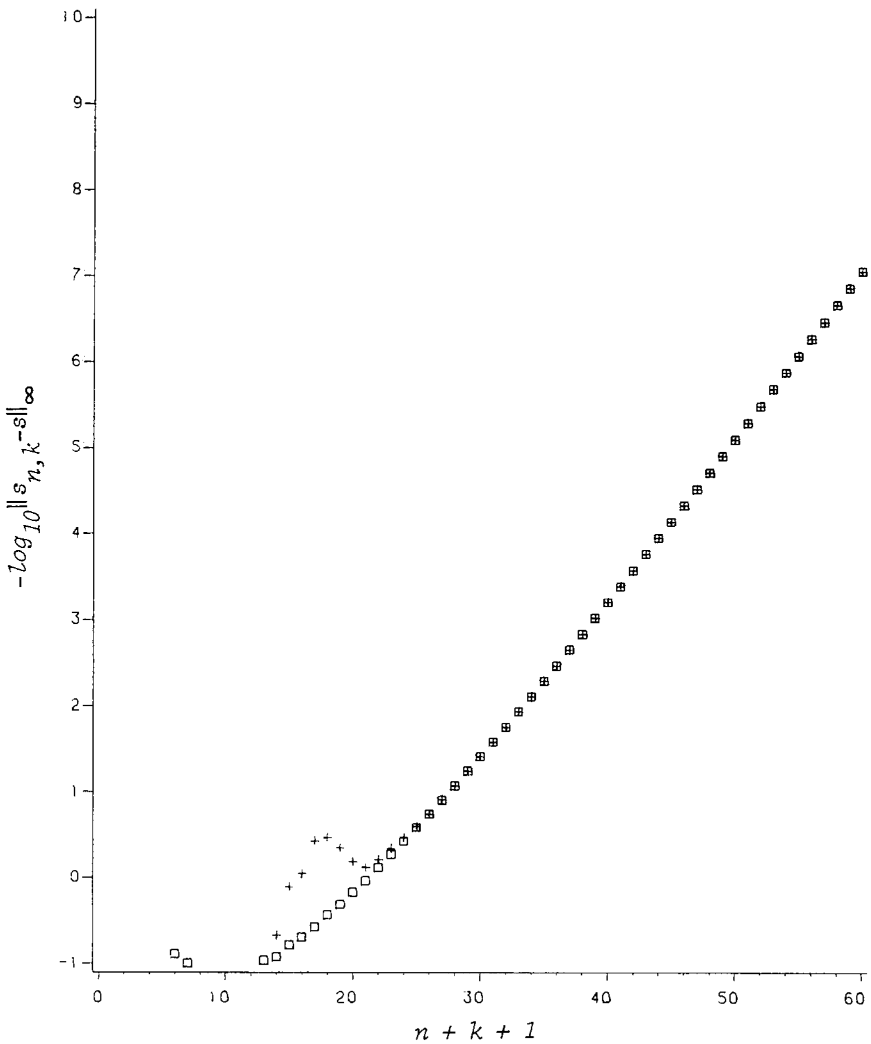

Fig. 1. Results for example 1 taking $k = 4$. $+$: $-\log_{10} \| s_{n,k} - s \|_\infty$ for MMPE. □: $-\log_{10} \| s_{n,k} - s \|_\infty$ for MPE.

Example 2. The matrix $\tilde{J}$ this time is given as the block diagonal matrix

$$\tilde{J} = \begin{pmatrix} C_1 & & & & \\ & C_2' & & & \\ & & C_2' & & \\ & & & C_3' & \\ & & & & C_3' \end{pmatrix}, \tag{4.9}$$

where C_1 is exactly as in (4.4) and

$$C_2' = \begin{pmatrix} 0.6 & 1 \\ 0 & 0.6 \end{pmatrix}, \qquad C_3' = \begin{pmatrix} 0.1 & 1 \\ 0 & 0.1 \end{pmatrix}. \tag{4.10}$$

Again the eigenvalues are $\lambda_1 = -1 + 0.6\text{i}$, $\lambda_2 = \bar{\lambda}_1$, $\lambda_3 = 0.6$, and $\lambda_4 = 0.1$, and, for an arbitrary

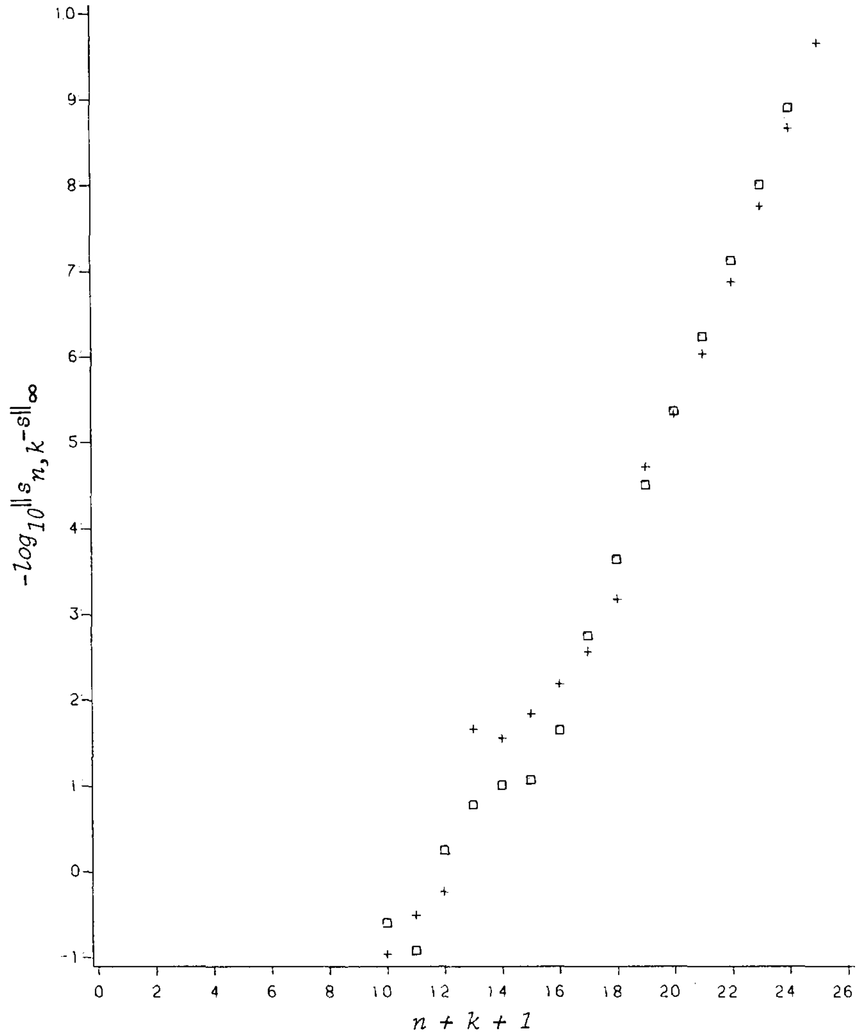

Fig. 2. Results for Example 1 taking $k=8$. +: $-\log_{10}\|s_{n,k}-s\|_\infty$ for MMPE. □: $-\log_{10}\|s_{n,k}-s\|_\infty$ for MPE.

initial vector x_0, x_m has an expansion of the form given in (4.6) with $p_1=p_2=1$, as in Example 1 but with $p_3=p_4=1$ unlike Example 1, as explained at the end of Section 2.

In Figs. 3 and 4 we give the results of the computations for $\|s_{n,k}-s\|_\infty$ using MPE and MMPE with $k=4$ and $k=6$, respectively. We again do not include $\|x_n-s\|_\infty$ as the sequence $x_0, x_1, \ldots,$ diverges by $\rho(A)=|\lambda_1|>1$. This time we should have

$$-\log_{10}\|s_{n,4}-s\|_\infty=(-\log_{10}0.6)n-\log_{10}n+O(1) \quad \text{as } n\to\infty \tag{4.11}$$

and

$$-\log_{10}\|s_{n,6}-s\|_\infty=(-\log_{10}0.1)n-\log_{10}n+O(1) \quad \text{as } n\to\infty. \tag{4.12}$$

These results are born out by Figs. 3 and 4.

As functions of n, $-\log_{10}\|s_{n,k}-s\|_\infty$ in the examples above exhibit almost a straight line behavior, which is slightly distorted due to the presence of the terms $-3\log_{10}n$ in (4.7) and (4.8) and of $-\log_{10}n$ in (4.11) and (4.12). The source of these terms is of course in the Jordan

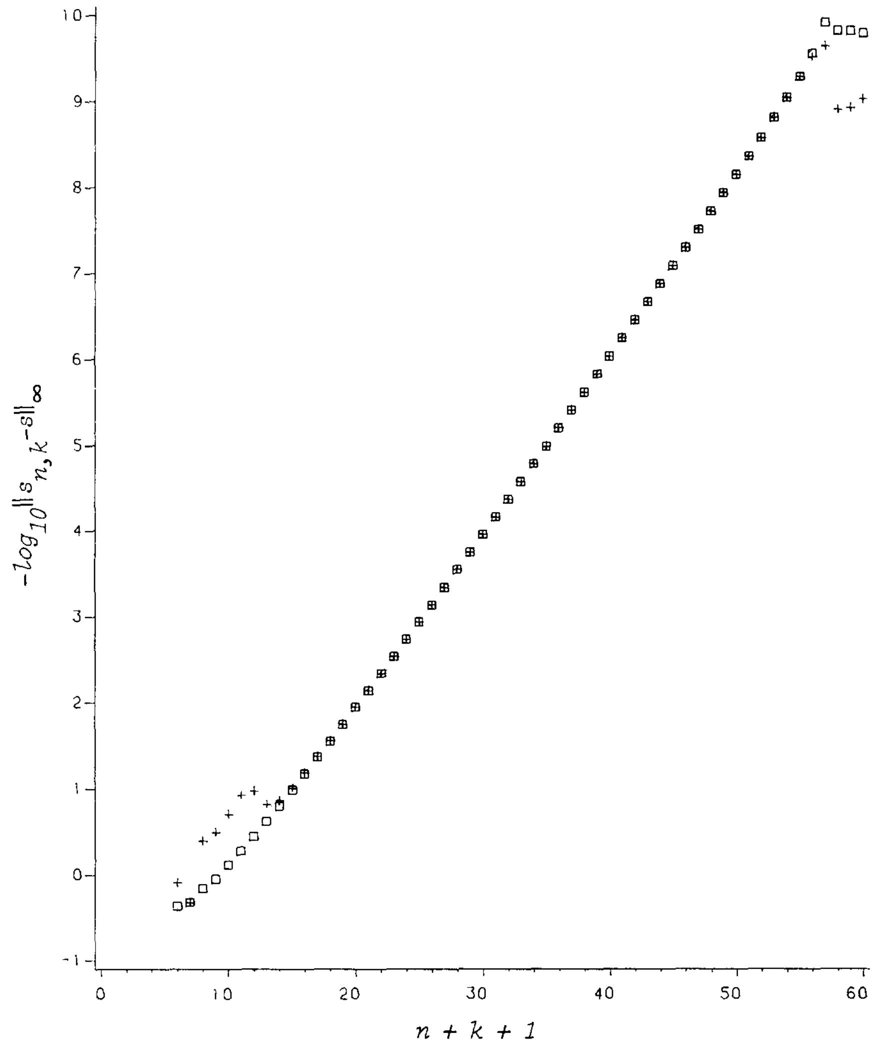

Fig. 3. Results for Example 2 taking $k = 4$. $+$: $-\log_{10} \| s_{n,k} - s \|_\infty$ for MMPE. □: $-\log_{10} \| s_{n,k} - s \|_\infty$ for MPE.

blocks C_2 and C_3 in Example 1 and C_2' and C_3' in Example 2. We see that the behavior of $-\log_{10} \| s_{n,k} - s \|_\infty$ is closer to that of a straight line in Figs. 3 and 4 than in Figs. 1 and 2 since the Jordan blocks C_2' and C_3' have smaller sizes than C_2 and C_3. Also, by the same reason, $s_{n,4}$ and $s_{n,6}$ in Example 2 achieve the same accuracies as $s_{n,4}$ and $s_{n,8}$ respectively in Example 1 with fewer iterations. Also recall that $s_{n,6}$ is obtained with less labor than $s_{n,8}$.

5. Proofs of main results

Definition 5.1. Let λ be a scalar and let m, j, and p be integers. Then the linear operator Δ is defined via

$$\Delta \binom{m}{j} \lambda^p = \binom{m+1}{j} \lambda^{p+1} - \binom{m}{j} \lambda^p. \tag{5.1}$$

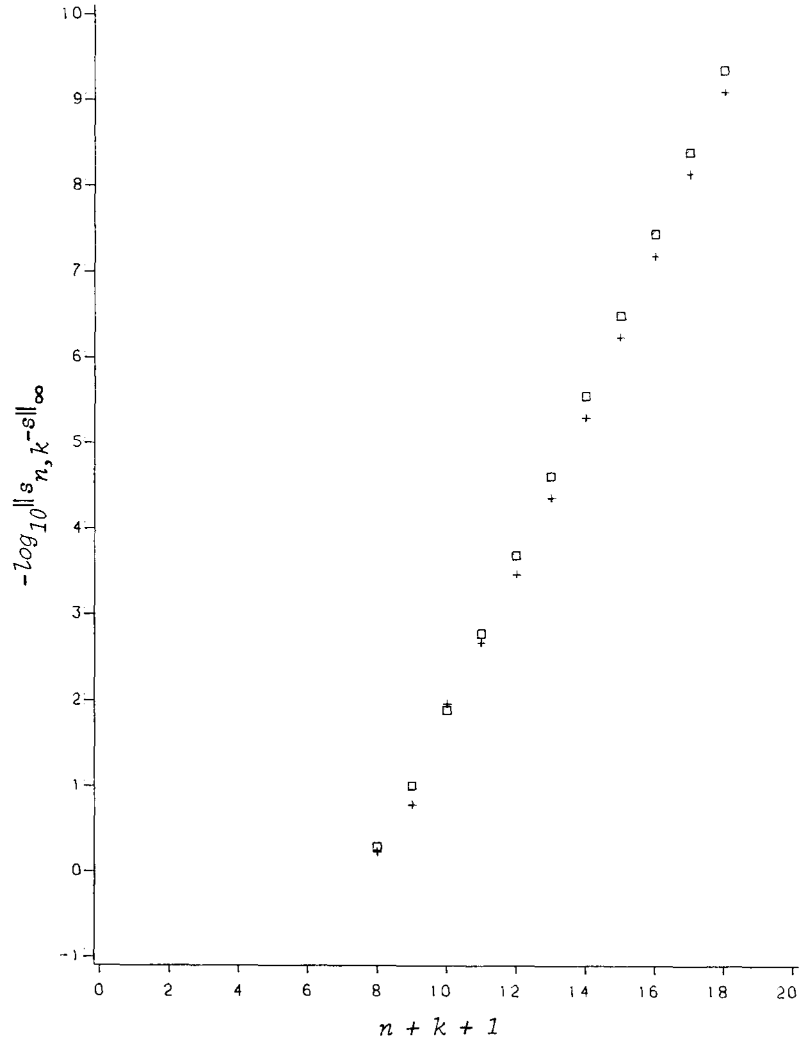

Fig. 4. Results for Example 2 taking $k=6$. $+$: $-\log_{10}\|s_{n,k}-s\|_\infty$ for MMPE. □: $-\log_{10}\|s_{n,k}-s\|_\infty$ for MPE.

Definition 5.2. Let

$$g_i=\sum_{j=1}^{j_i} b_{ij}\binom{n_{ij}}{l_{ij}}\lambda_{ij}^{m_{ij}}, \quad i=1,2,\ldots, \tag{5.2}$$

where b_{ij} and λ_{ij} are scalars and n_{ij}, l_{ij} and m_{ij} are integers. Define the $N\times N$ matrix W by

$$(W)_{ip}=\sum_{j=1}^{j_i} b_{ij}\binom{n_{ij}+p-1}{l_{ij}}\lambda_{ij}^{m_{ij}+p-1}, \quad 1\leqslant i,p\leqslant N, \tag{5.3}$$

and denote

$$Y(g_1, g_2,\ldots, g_N)=\det W. \tag{5.4}$$

When $g_i = 1$ for $i = i_0$, we shall take $(W)_{i_0 p} = 1$, $1 \leqslant p \leqslant N$. Note that the first column of W is composed of the transpose of the vector $(g_1, g_2, \ldots, g_N)$. Note also that if $g_i = g_j$ for some i, j $(i \neq j)$, then $Y(g_1, \ldots, g_N) = 0$ since in this case W has two identical rows.

Example 5.1.

$$Y\left(1, \binom{n}{p}\lambda_1^0, \Delta\binom{n}{p}\lambda_2^0\right)$$
$$= \begin{vmatrix} 1 & 1 & 1 \\ \binom{n}{p} & \binom{n+1}{p}\lambda_1 & \binom{n+2}{p}\lambda_1^2 \\ \binom{n+1}{p}\lambda_2 - \binom{n}{p} & \binom{n+2}{p}\lambda_2^2 - \binom{n+1}{p}\lambda_2 & \binom{n+3}{p}\lambda_2^3 - \binom{n+2}{p}\lambda_2^2 \end{vmatrix}.$$

Example 5.2. Let $a_1, \ldots, a_h$ be arbitrary scalars, and let $q_1, \ldots, q_h$ be non-negative integers. Then,

$$\tilde{Y}(a_1, q_1;\ a_2, q_2;\ \ldots;\ a_h, q_h)$$
$$\equiv Y\left(\binom{0}{0}a_1^0, \binom{0}{1}a_1^{-1}, \ldots, \binom{0}{q_1}a_1^{-q_1}, \ldots, \binom{0}{0}a_h^0, \binom{0}{1}a_h^{-1}, \ldots, \binom{0}{q_h}a_h^{-q_h}\right)$$
$$= \prod_{1 \leqslant i < j \leqslant h} (a_j - a_i)^{(q_j+1)(q_i+1)}. \tag{5.5}$$

Actually, the determinant Y in (5.5) can be shown to be the generalized Vandermonde determinant of $a_1, \ldots, a_h$. For details see, for example, [6].

Lemma 5.1. *For arbitrary g_i and λ let*

$$R = Y\left(g_1, \ldots, g_h, \binom{n}{0}\lambda^m, \ldots, \binom{n}{j-1}\lambda^m, \Delta\binom{n}{j}\lambda^m, \binom{n}{j+1}\lambda^m, \ldots, \binom{n}{p}\lambda^m\right). \tag{5.6}$$

Then

$$R = (\lambda - 1)Y\left(g_1, \ldots, g_h, \binom{n}{0}\lambda^m, \ldots, \binom{n}{j}\lambda^m, \ldots, \binom{n}{p}\lambda^m\right). \tag{5.7}$$

Proof. That (5.7) holds for $j = 0$ is easily seen since by (5.1)

$$\Delta\binom{n}{0}\lambda^m = (\lambda - 1)\lambda^m = (\lambda - 1)\binom{n}{0}\lambda^m. \tag{5.8}$$

Consider now $j > 0$. Again by virtue of (5.1) we have

$$\Delta\binom{n}{j}\lambda^m = \binom{n}{j-1}\lambda^{m+1} + (\lambda - 1)\binom{n}{j}\lambda^m. \tag{5.9}$$

Substituting (5.9) in (5.6), and using the fact that determinants are multilinear in their rows, we have

$$R = \lambda Y\left(g_1, \ldots, g_h, \binom{n}{0}\lambda^m, \ldots, \binom{n}{j-1}\lambda^m, \binom{n}{j-1}\lambda^m, \binom{n}{j+1}\lambda^m, \ldots, \binom{n}{p}\lambda^m\right)$$
$$+ (\lambda - 1)Y\left(g_1, \ldots, g_h, \binom{n}{0}\lambda^m, \ldots, \binom{n}{j}\lambda^m, \ldots, \binom{n}{p}\lambda^m\right). \tag{5.10}$$

The first of the determinants Y on the right hand side of (5.10) vanishes since it has two identical rows, and what remains is (5.7). □

Corollary. *If in* (5.6) *there are l terms* ($l \leqslant p+1$) *of the form $\Delta\binom{n}{j_i}\lambda^m$, $i=1,\ldots,l$, then any Δ can be removed provided the resulting determinant is multiplied by* $(\lambda-1)$. *In particular,*

$$Y\left(g_1,\ldots,g_h,\Delta\binom{n}{0}\lambda^m,\ldots,\Delta\binom{n}{p}\lambda^m\right)=(\lambda-1)^{p+1}Y\left(g_1,\ldots,g_h,\binom{n}{0}\lambda^m,\ldots,\binom{n}{p}\lambda^m\right). \tag{5.11}$$

Lemma 5.2. *For arbitrary g_i and λ, and non-negative integers q_i, let*

$$\tilde{R}^n_{q_0,\ldots,q_p}=Y\left(g_1,\ldots,g_h,\binom{n+q_0}{0}\lambda^m,\binom{n+q_1}{1}\lambda^m,\ldots,\binom{n+q_p}{p}\lambda^m\right). \tag{5.12}$$

Then $\tilde{R}^n_{q_0,\ldots,q_p}$ is independent of n and the q_i, thus

$$\tilde{R}^n_{q_0,\ldots,q_p}=\tilde{R}^0_{0,\ldots,0}=Y\left(g_1,\ldots,g_h,\binom{0}{0}\lambda^m,\binom{0}{1}\lambda^m,\ldots,\binom{0}{p}\lambda^m\right). \tag{5.13}$$

Proof. We first show that $\tilde{R}^n_{q_0,\ldots,q_p}$ is independent of the q_i. We shall prove this assertion by proving that

$$\tilde{R}^n_{q_0,\ldots,q_p}=Y\left(g_1,\ldots,g_h,\binom{n}{0}\lambda^m,\ldots,\binom{n}{j}\lambda^m,\binom{n+q_{j+1}}{j+1}\lambda^m,\ldots,\binom{n+q_p}{p}\lambda^m\right). \tag{5.14}$$

First, (5.14) holds for $j=0$ since

$$\binom{n+q}{0}=\binom{n}{0}=1 \quad \text{all } q=0,1,2,\ldots\,. \tag{5.15}$$

We now assume that (5.14) holds for j. Substituting the identity

$$\binom{n+q}{l}=\sum_{i=0}^{l}\binom{n}{i}\binom{q}{l-i} \tag{5.16}$$

in (5.14) with $q=q_{j+1}$ and $l=j+1$, and using the fact that Y is multilinear in its arguments, we have

$$\tilde{R}^n_{q_0,\ldots,q_p}=\sum_{i=0}^{j+1}\binom{q_{j+1}}{j+1-i}Y\left(g_1,\ldots,g_h,\binom{n}{0}\lambda^m,\ldots,\binom{n}{j}\lambda^m,\binom{n}{i}\lambda^m,\right.$$
$$\left.\binom{n+q_{j+2}}{j+2}\lambda^m,\ldots,\binom{n+q_p}{p}\lambda^m\right). \tag{5.17}$$

One can see that all terms with $i\leqslant j$ in (5.17) vanish, from which one concludes that $\tilde{R}^n_{q_0,\ldots,q_p}$ is independent of q_{j+1} also.

Now that we have proved $\tilde{R}^n_{q_0,\ldots,q_p}$ to be independent of $q_0,\ldots,q_p$, we can write

$$\tilde{R}^n_{q_0,\ldots,q_p}=\tilde{R}^n_{1,\ldots,1}. \tag{5.18}$$

But by definition of $\tilde{R}^n_{q_0,\ldots,q_p}$

$$\tilde{R}^n_{1,\ldots,1} = \tilde{R}^{n+1}_{0,\ldots,0}, \tag{5.19}$$

thus proving the lemma. □

We now state a lemma whose proof can be found in [11].

Lemma 5.3. *Let $i_0, i_1, \ldots, i_k$ be integers greater than or equal to 1, and assume that the scalars $v_{i_0,\ldots,i_k}$ are odd under an interchange of any two indices $i_0, \ldots, i_k$. Let σ_i, $i \geq 1$, be scalars (or vectors), and let t_{ij}, $i \geq 1$, $1 \leq j \leq k$ be scalars. Define*

$$I_{k,N} = \sum_{i_0=1}^{N} \cdots \sum_{i_k=1}^{N} \sigma_{i_0}\left(\prod_{p=1}^{k} t_{i_p,p}\right) v_{i_0,\ldots,i_k} \tag{5.20}$$

and

$$J_{k,N} = \sum_{1 \leq i_0 < i_1 < \cdots < i_k \leq N} \begin{vmatrix} \sigma_{i_0} & \sigma_{i_1} & \cdots & \sigma_{i_k} \\ t_{i_0,1} & t_{i_1,1} & \cdots & t_{i_k,1} \\ t_{i_0,2} & t_{i_1,2} & \cdots & t_{i_k,2} \\ \vdots & \vdots & & \vdots \\ t_{i_0,k} & t_{i_1,k} & \cdots & t_{i_k,k} \end{vmatrix} v_{i_0,\ldots,i_k}, \tag{5.21}$$

where the determinant in (5.21) *is to be interpreted in the same way as $D(\sigma_0, \ldots, \sigma_k)$ in* (1.9). *Then*

$$I_{k,N} = J_{k,N}. \tag{5.22}$$

Definition 5.3. Let j_h, l_h, j_p, l_p be non-negative integers. We will write

$$j_h l_h < j_p l_p \quad \text{if } j_h < j_p \quad \text{or if } j_h = j_p \quad \text{and} \quad l_h < l_p, \tag{5.23}$$

$$j_h l_h = j_p l_p \quad \text{if } j_h = j_p \quad \text{and} \quad l_h = l_p, \tag{5.24}$$

$$j_h l_h \leq j_p l_p \quad \text{if either (5.23) or (5.24) holds.} \tag{5.25}$$

Note that Definition 5.3 is equivalent to ordering the set of pairs of non-negative integers lexicographically.

For brevity, in the sequel we shall denote

$$\sum_{jl} \equiv \sum_{j=1}^{\infty} \sum_{l=0}^{p_j}, \qquad \sum_{j_1 l_1 < j_2 l_2 < \cdots < j_k l_k} \equiv \sum_{j_1 l_1} \sum_{\substack{j_2 l_2 \\ (j_2 l_2 > j_1 l_1)}} \cdots \sum_{\substack{j_k l_k \\ (j_k l_k > j_{k-1} l_{k-1})}} \tag{5.26}$$

In convergence and stability analyses below, we will make use of the three lemmas above, as well as of the following asymptotic expansion for the vectors u_m, which follows from (1.1):

$$u_m = x_{m+1} - x_m \sim \sum_{j=1}^{\infty} \sum_{l=0}^{p_j} y_{jl} \Delta\binom{m}{l} \lambda_j^m \quad \text{as } m \to \infty. \tag{5.27}$$

In addition, for notational brevity, let us agree that "$\alpha_n \sim \beta_n$" is equivalent to "$\alpha_n \sim \beta_n$ as $n \to \infty$".

Finally, the relation

$$s_{n,k} - s = \frac{D(x_{n+q} - s,\ x_{n+q+1} - s, \ldots, x_{n+q+k} - s)}{D(1, 1, \ldots, 1)} \tag{5.28}$$

and

$$\sum_{j=0}^{k} \gamma_j^{(n,k)} \lambda^j = \frac{D(1,\ \lambda, \ldots, \lambda^k)}{D(1, 1, \ldots, 1)} \tag{5.29}$$

will be of use in the proofs below. Both (5.28) and (5.29) are consequences of (1.6), (1.8), and (1.9).

5.1. Convergence and stability proofs for MMPE

From (1.10) and (5.27) it follows that

$$u_{h,i} \sim \sum_{j=1}^{\infty} \sum_{l=0}^{p_j} z_{jl,h} \Delta \binom{n+i}{l} \lambda_j^{n+i}, \tag{5.30}$$

where

$$z_{jl,h} = Q_{h+1}(y_{jl}). \tag{5.31}$$

Lemma 5.4. *Define*

$$Z_{j_1 l_1, \ldots, j_k l_k} = \begin{vmatrix} z_{j_1 l_1,1} & z_{j_2 l_2,1} & \cdots & z_{j_k l_k,1} \\ z_{j_1 l_1,2} & z_{j_2 l_2,2} & \cdots & z_{j_k l_k,2} \\ \vdots & \vdots & & \vdots \\ z_{j_1 l_1,k} & z_{j_2 l_2,k} & \cdots & z_{j_k l_k,k} \end{vmatrix}. \tag{5.32}$$

Then $H_n(\lambda) \equiv D(1,\ \lambda, \ldots, \lambda^k)$ has the asymptotic behavior

$$H_n(\lambda) \sim \sum_{j_1 l_1 < \cdots < j_k l_k} Z_{j_1 l_1, \ldots, j_k l_k} \left(\prod_{h=1}^{k} \lambda_{j_h}^n \right) Y\left(\binom{n}{0} \lambda^0,\ \Delta \binom{n}{l_1} \lambda_{j_1}^0, \ldots, \Delta \binom{n}{l_k} \lambda_{j_k}^0 \right). \tag{5.33}$$

Proof. Substituting (5.30) into the determinant expression (1.8) for $D(1,\ \lambda, \ldots, \lambda^k)$, we obtain

$$H_n(\lambda) \sim \begin{vmatrix} 1 & \lambda & \cdots & \lambda^k \\ \sum_{j_1 l_1} z_{j_1 l_1,1} \Delta \binom{n}{l_1} \lambda_{j_1}^n & \sum_{j_1 l_1} z_{j_1 l_1,1} \Delta \binom{n+1}{l_1} \lambda_{j_1}^{n+1} & \cdots & \sum_{j_1 l_1} z_{j_1 l_1,1} \Delta \binom{n+k}{l_1} \lambda_{j_1}^{n+k} \\ \sum_{j_2 l_2} z_{j_2 l_2,2} \Delta \binom{n}{l_2} \lambda_{j_2}^n & \sum_{j_2 l_2} z_{j_2 l_2,2} \Delta \binom{n+1}{l_2} \lambda_{j_2}^{n+1} & \cdots & \sum_{j_2 l_2} z_{j_2 l_2,2} \Delta \binom{n+k}{l_2} \lambda_{j_2}^{n+k} \\ \vdots & \vdots & & \vdots \\ \sum_{j_k l_k} z_{j_k l_k,k} \Delta \binom{n}{l_k} \lambda_{j_k}^n & \sum_{j_k l_k} z_{j_k l_k,k} \Delta \binom{n+1}{l_k} \lambda_{j_k}^{n+1} & \cdots & \sum_{j_k l_k} z_{j_k l_k,k} \Delta \binom{n+k}{l_k} \lambda_{j_k}^{n+k} \end{vmatrix}. \tag{5.34}$$

Using the multilinearity property of determinants, and removing common factors from each row, we can express (5.34) in the form

$$H_n(\lambda) \sim \sum_{j_1 l_1} \sum_{j_2 l_2} \cdots \sum_{j_k l_k} \left(\prod_{h=1}^{k} z_{j_h l_h, h} \right) \left(\prod_{h=1}^{k} \lambda_{j_h}^{n} \right) \times Y\left(\binom{n}{0} \lambda^0, \Delta\binom{n}{l_1} \lambda_{j_1}^0, \Delta\binom{n}{l_2} \lambda_{j_2}^0, \ldots, \Delta\binom{n}{l_k} \lambda_{j_k}^0 \right). \tag{5.35}$$

Observing that the product

$$\left(\prod_{h=1}^{k} \lambda_{j_h}^{n} \right) Y\left(\binom{n}{0} \lambda^0, \Delta\binom{n}{l_1} \lambda_{j_1}^0, \ldots, \Delta\binom{n}{l_k} \lambda_{j_k}^0 \right)$$

is odd under an interchange of two pairs of indices $j_h l_h$, we can invoke Lemma 5.3 to obtain (5.33). □

Theorem 5.1. *Provided* $\lambda \neq \lambda_i$, $i = 1, \ldots, t$,

$$D(1, \lambda, \ldots, \lambda^k) = Z\left[\prod_{h=1}^{t} \lambda_h^{(p_h+1)n} \right] \left[\prod_{h=1}^{t} (\lambda_h - 1)^{p_h+1} \lambda_h^{p_h(p_h+1)/2} \right] \times \tilde{Y}(\lambda, 0; \lambda_1, p_1; \ldots; \lambda_t, p_t)[1 + o(1)] \quad \text{as } n \to \infty. \tag{5.36}$$

Note: (5.36) implies that $D(1, 1, \ldots, 1) \neq 0$ as $n \to \infty$ which guarantees the existence of $s_{n,k}$ for large enough n.

Proof. It can be shown that the dominant term in the expansion (5.33) is the one whose indices $j_1 l_1, \ldots, j_k l_k$ take on the values $10, 11, \ldots, 1p_1, 20, \ldots, 2p_2, \ldots, t0, \ldots, tp_t$, respectively. That is,

$$H_n(\lambda) \sim Z_{10, \ldots, 1p_1, \ldots, t0, \ldots, tp_t} \left[\prod_{h=1}^{t} \lambda_h^{(p_h+1)n} \right] \times Y\left(\binom{n}{0} \lambda^0, \Delta\binom{n}{0} \lambda_1^0, \ldots, \Delta\binom{n}{p_1} \lambda_1^0, \ldots, \Delta\binom{n}{0} \lambda_t^0, \ldots, \Delta\binom{n}{p_t} \lambda_t^0 \right), \tag{5.37}$$

provided that this term is not zero. From (5.32), (5.31) and (3.6), we have $Z_{10, \ldots, 1p_1, \ldots, t0, \ldots, tp_1} = Z$, which is non-zero by assumption. Applying the corollary to Lemma 5.1 and Lemma 5.2 to the determinant Y in (5.37) shows this determinant to be a multiple of $\tilde{Y}(\lambda, 0; \lambda_1, p_1; \ldots; \lambda_t, p_t)$ which is non-zero by virtue of (5.5) and the assumption $\lambda \neq \lambda_i$, $1 \leqslant i \leqslant t$. This proves the theorem. □

Lemma 5.5. $G_n \equiv D(x_{n+q} - s, \ldots, x_{n+q+k} - s)$ *has the asymptotic behavior*

$$G_n \sim \sum_{j_0 l_0} y_{j_0 l_0} \lambda_{j_0}^{q} \sum_{j_1 l_1 < j_2 l_2 < \cdots < j_k l_k} Z_{j_1 l_1, \ldots, j_k l_k} \left(\prod_{h=0}^{k} \lambda_{j_h}^{n} \right) \times Y\left(\binom{n+q}{l_0} \lambda_{j_0}^0, \Delta\binom{n}{l_1} \lambda_{j_1}^0, \ldots, \Delta\binom{n}{l_k} \lambda_{j_k}^0 \right). \tag{5.38}$$

Proof. Substituting (5.30) and (1.1) into the determinant expression (1.8) for $D(x_{n+q}-s,\ldots,x_{n+q+k}-s)$, and proceeding as in the proof of Lemma 5.4, we obtain

$$G_n \sim \sum_{j_0 l_0} y_{j_0 l_0}\lambda_{j_0}^q \sum_{j_1 l_1}\cdots\sum_{j_k l_k}\left(\prod_{h=1}^{k} z_{j_h l_h,h}\right)\left(\prod_{h=0}^{k}\lambda_{j_h}^n\right)Y\left(\binom{n+q}{l_0}\lambda_{j_0}^0,\ \Delta\binom{n}{l_1}\lambda_{j_1}^0,\ldots,\Delta\binom{n}{l_k}\lambda_{j_k}^0\right). \tag{5.39}$$

The product

$$\left(\prod_{h=0}^{k}\lambda_{j_h}^n\right)Y\left(\binom{n+q}{l_0}\lambda_{j_0}^0,\ \Delta\binom{n}{l_1}\lambda_{j_1}^0,\ldots,\Delta\binom{n}{l_k}\lambda_{j_k}^0\right)$$

is odd under an interchange of two pairs of indices $j_h l_h$, $h=1,\ldots,k$. Thus Lemma 5.3 can be invoked, resulting in (5.38). □

Theorem 5.2. *Define the vectors* $\tilde{z}_{i,l}$ *by*

$$\tilde{z}_{i,l}=\begin{vmatrix} d_1 y_{10} & \cdots & d_1 y_{1p_1} & \cdots & d_t y_{t0} & \cdots & d_t y_{tp_t} & d_i y_{il}\\ z_{10,1} & \cdots & z_{1p_1,1} & \cdots & z_{t0,1} & \cdots & z_{tp_t,1} & z_{il,1}\\ \vdots & & \vdots & & \vdots & & \vdots & \vdots\\ z_{10,k} & \cdots & z_{1p_1,k} & \cdots & z_{t0,k} & \cdots & z_{tp_t,k} & z_{il,k}\end{vmatrix}, \tag{5.40}$$

where $d_h=\lambda_h^q/(\lambda_h-1)$. *Then with* r' *as defined in* (3.2),

$$G_n \sim \left[\prod_{h=1}^{t}\lambda_h^{n(p_h+1)}\right]\left[\prod_{h=1}^{t}(\lambda_h-1)^{p_h+1}\lambda_h^{p_h(p_h+1)/2}\right]$$
$$\times\frac{n^{p_{t+1}}}{p_{t+1}!}\sum_{i=1}^{r'}\lambda_{t+i}^n\tilde{z}_{t+i,p_{t+i}}(\lambda_{t+i}-1)\tilde{Y}(\lambda_1,\ p_1;\ \ldots;\ \lambda_t,\ p_t;\ \lambda_{t+i},0). \tag{5.41}$$

Proof. We shall treat the case in which $r=1$ first. In this case the dominant terms in (5.38) are those for which the pairs of indices $j_0l_0,\ldots,j_kl_k$ are permutations of the pairs $10,\ldots,1p_1,\ldots,t0,\ldots,tp_t,(t+1)l$, $(0\leqslant l\leqslant p_{t+1})$, subject to the constraint $j_1l_1<j_2l_2<\cdots<j_kl_k$. Denote the minor of the entry d_jy_{jp} in the first row of the determinant $\tilde{z}_{t+1,l}$ in (5.40) by $\tilde{z}_{t+1,l}^{jp}$. Obviously,

$$\tilde{z}_{t+1,l}^{jp}=Z_{10,\ldots,1p_1,\ldots,j0,\ldots,j(p-1),j(p+1),\ldots,jp_j,\ldots,t0,\ldots,tp_t,(t+1)l}. \tag{5.42}$$

Let $(-1)^{\sigma_{jp}}$ be the phase factor for which $(-1)^{\sigma_{jp}}\tilde{z}_{t+1,l}^{jp}$ is the cofactor of d_jy_{jp}. The dominant term in the asymptotic expansion (5.38) of G_n thus becomes

$$G_n\sim\sum_{l=0}^{p_{t+1}}\sum_{j_0=1}^{t+1}\sum_{l_0}{}^{*}\,y_{j_0l_0}\lambda_{j_0}^q\tilde{z}_{t+1,l}^{j_0l_0}\left[\prod_{i=1}^{t}\lambda_i^{n(p_i+1)}\right]\lambda_{t+1}^n$$
$$\times Y\left(\binom{n+q}{l_0}\lambda_{j_0}^0,\ \Delta\binom{n}{l_1}\lambda_{j_1}^0,\ldots,\Delta\binom{n}{l_k}\lambda_{j_k}^0\right), \tag{5.43}$$

subject to the above mentioned constraints on the indices $j_0 l_0, \ldots, j_k l_k$. Here we have denoted

$$\sum_{l_0}{}^{*} \equiv \begin{cases} \displaystyle\sum_{l_0=0}^{p_{j_0}} & \text{for } 1 \leqslant j_0 \leqslant t, \\ \displaystyle\sum_{l_0=l}^{l} & \text{for } j_0 = t+1. \end{cases}$$

By rearranging rows in the determinant Y in (5.43), and using the corollary to Lemma 5.1 and Lemma 5.2, in this order, we can show that this determinant is actually

$$(-1)^{\sigma_{j_0 l_0}}\left[\prod_{i=1}^{t}(\lambda_i - 1)^{p_i+1}\right](\lambda_{j_0} - 1)^{-1}$$
$$\times\, Y\left(\binom{0}{0}\lambda_1^0, \ldots, \binom{0}{p_1}\lambda_1^0, \ldots, \binom{0}{0}\lambda_t^0, \ldots, \binom{0}{p_t}\lambda_t^0,\ \Delta\binom{n}{l}\lambda_{t+1}^0\right)$$

for $1 \leqslant j_0 \leqslant t$, and

$$(-1)^{\sigma_{j_0 l_0}}\left[\prod_{i=1}^{t}(\lambda_i - 1)^{p_i+1}\right] Y\left(\binom{0}{0}\lambda_1^0, \ldots, \binom{0}{p_1}\lambda_1^0, \ldots, \binom{0}{0}\lambda_t^0, \ldots, \binom{0}{p_t}\lambda_t^0,\ \binom{n+q}{l}\lambda_{t+1}^0\right)$$

for $j_0 = t+1$ and $l_0 = l$. Using the facts that

$$\Delta\binom{n}{l}\lambda_{t+1}^0 = \frac{n^l}{l!}(\lambda_{t+1} - 1) + O(n^{l-1}) \quad \text{as } n \to \infty$$

and

$$\binom{n+q}{l} = \frac{n^l}{l!} + O(n^{l-1}) \quad \text{as } n \to \infty,$$

both cases can be combined in (5.43), to yield

$$G_n \sim \sum_{l=0}^{p_{t+1}} \sum_{j_0=1}^{t+1} \sum_{l_0}{}^{*} d_{j_0} y_{j_0 l_0} (-1)^{\sigma_{j_0 l_0}} \tilde{z}_{t+1,l}^{j_0 l_0} \left[\prod_{i=1}^{t}(\lambda_i - 1)^{p_i+1}\right](\lambda_{t+1} - 1)$$
$$\times \left(\prod_{i=1}^{t} \lambda_i^{n(p_i+1)}\right) \lambda_{t+1}^n \frac{n^l}{l!} Y\left(\binom{0}{0}\lambda_1^0, \ldots, \binom{0}{p_1}\lambda_1^0, \ldots, \binom{0}{0}\lambda_t^0, \ldots, \binom{0}{p_t}\lambda_t^0,\ \binom{0}{0}\lambda_{t+1}^0\right). \tag{5.44}$$

Using the fact that

$$\sum_{j_0=1}^{t+1} \sum_{l_0}{}^{*} (-1)^{\sigma_{j_0 l_0}} d_{j_0} y_{j_0 l_0} \tilde{z}_{t+1,l}^{j_0 l_0} = \tilde{z}_{t+1,l}, \tag{5.45}$$

and noting that $l = p_{t+1}$ yields the dominant term, and exploiting the relation between the determinant Y in (5.44) and the appropriate generalized Vandemonde determinant $\tilde{Y}$, (5.44) reduces to (5.41) with $r' = 1$. It should now be clear that G_n can be expressed via (5.41) for arbitrary r as well. What remains to be shown is that the dominant term in (5.41) is non-zero. For this it is enough to show that the summation in (5.41) does not vanish. By (5.45) we can express this summation as a linear combination of the linearly independent vectors y_{jp},

$10 \leqslant jp \leqslant tp_t$, and $y_{t+i,p_{t+i}}$, $1 \leqslant i \leqslant r'$. Now the coefficient multiplying the vector $y_{t+i,p_{t+i}}$ in this linear combination is the product

$$(-1)^k \lambda_{t+i}^n \tilde{z}_{t+i,p_{t+i}}^{t+i,p_{t+i}} (\lambda_{t+i} - 1) \tilde{Y}(\lambda_1, p_1; \ldots; \lambda_t, p_t; \lambda_{t+i}, 0),$$

and this product is nonzero by $\tilde{z}_{t+i,p_{t+i}}^{t+1,p_{t+i}} = Z$ and the assumption that $Z \neq 0$, and the rest of the assumptions on the λ_i. □

Theorem 3.1 for MMPE can now be proved by dividing the asymptotic behavior of G_n in (5.41) by that of $H_n(1)$ in (5.36), the vector $\Gamma(n)$ in (3.4) being identified as

$$\Gamma(n) = \frac{1}{Zp_{t+1}!} \sum_{j=1}^{r'} \mathrm{e}^{\mathrm{i}n \arg \lambda_{t+j}} \tilde{z}_{t+j,p_{t+j}} (\lambda_{t+j} - 1) \prod_{h=1}^{t} \left(\frac{\lambda_{t+j} - \lambda_h}{\lambda_h - 1} \right)^{p_h+1} + \mathrm{o}(1) \quad \text{as } n \to \infty. \tag{5.46}$$

Clearly this $\Gamma(n)$ satisfies (3.5).

The proof of Theorem 3.2 can be achieved by combining the asymptotic behaviour of $H_n(\lambda)$ and $H_n(1)$ from (5.36) and (5.29).

We conclude this subsection by exploring the meaning of the constraint $Z \neq 0$ for the example described in Section 2. Substituting (2.11) in (3.6) we find after some elementary column transformations that $Z \neq 0$ is equivalent to

$$\begin{vmatrix} Q_1(v_{11}) & \cdots & Q_1(v_{1,p_1+1}) & \cdots & Q_1(v_{t1}) & \cdots & Q_1(v_{t,p_t+1}) \\ \vdots & & \vdots & & \vdots & & \vdots \\ Q_k(v_{11}) & \cdots & Q_k(v_{1,p_1+1}) & \cdots & Q_k(v_{t1}) & \cdots & Q_k(v_{t,p_t+1}) \end{vmatrix} \neq 0, \tag{5.47}$$

where v_{ji} are eigenvectors or principal vectors.

5.2. Convergence and stability proofs for TEA

From (1.11) and (5.27) it follows that

$$u_{h,i} = Q(u_{n+h+i}) \sim \sum_{j=1}^{\infty} \sum_{p=0}^{p_j} Q(y_{jp}) \Delta \binom{n+h+i}{p} \lambda_j^{n+h+i}. \tag{5.48}$$

Using the binomial identity in (5.16) with n and q there replaced by $n+i$ and h respectively, (5.48) becomes

$$u_{h,i} \sim \sum_{j=1}^{\infty} \sum_{p=0}^{p_j} Q(y_{jp}) \sum_{l=0}^{p} \binom{h}{p-l} \Delta \binom{n+i}{l} \lambda_j^{n+h+i}. \tag{5.49}$$

Interchanging the summations over p and l, and letting

$$z_{jl,h} = \sum_{p=l}^{p_j} \binom{h}{p-l} Q(y_{jp}) \lambda_j^h, \tag{5.50}$$

we see that $u_{h,i}$ for TEA has the same form as that for MMPE given in (5.30). Therefore, the proofs for TEA are identical to those for MMPE, provided we replace $Q_{h+1}(y_{hl})$ in the definition of Z in (3.6) with $z_{jl,h}$ of (5.50). The only thing that remains to be shown is that (3.7) implies that $Z \neq 0$. By performing elementary column transformations on this new Z we find that

$$Z = (-1)^{\sum_{j=1}^{t}\sigma_j}\left(\prod_{j=1}^{t}\lambda_j^{\sigma_j}\left[Q(y_{jp_j})\right]^{p_j+1}\right)\tilde{Y}(\lambda_1, p_1; \ldots; \lambda_t, p_t), \tag{5.51}$$

where $\sigma_j = p_j(p_j+1)/2$. The desired result now follows.

In the context of the example of Section 2, the condition (3.7) is equivalent to

$$\prod_{j=1}^{t} Q(v_{j1}) \neq 0, \tag{5.52}$$

which can be seen by observing that $y_{jp_j} = \lambda_j^{-p_j} a_{j,p_j+1} v_{j1}$. It is interesting to note that (5.52) imposes no conditions on the the operator Q with respect to the principal vectors.

5.3. Convergence and stability proofs for MPE

From (1.12) and (5.27) we have

$$u_{h,i} = (u_{n+h}, u_{n+i}) \sim \sum_{mp}\sum_{jl} z_{jl}^{mp}\left[\Delta\binom{n+h}{p}\bar{\lambda}_m^{n+h}\right]\left[\Delta\binom{n+i}{l}\lambda_j^{n+i}\right], \tag{5.53}$$

where

$$z_{jl}^{mp} = (y_{mp}, y_{jl}). \tag{5.54}$$

Theorem 5.3. *Let*

$$Z_{j_1l_1,\ldots,j_kl_k}^{h_1i_1,\ldots,h_ki_k} = \begin{vmatrix} z_{j_1l_1}^{h_1i_1} & z_{j_2l_2}^{h_1i_1} & \cdots & z_{j_kl_k}^{h_1i_1} \\ \vdots & \vdots & & \vdots \\ z_{j_1l_1}^{h_ki_k} & & \cdots & z_{j_kl_k}^{h_ki_k} \end{vmatrix}. \tag{5.55}$$

Provided that $\lambda \neq \lambda_i$, $i = 1,\ldots,t$,

$$\begin{aligned} D(1, \lambda,\ldots,\lambda^k) = {} & Z_{10,\ldots,1p_1,\ldots,t0,\ldots,tp_t}^{10,\ldots,1p_1,\ldots,t0,\ldots,tp_t}\left|\prod_{m=1}^{t}\lambda_m^{2n(p_m+1)}\right| \\ & \times\left|\prod_{m=1}^{t}(\lambda_m-1)^{2(p_m+1)}\lambda_m^{p_m(p_m+1)}\right|\tilde{Y}(\bar{\lambda}_1, p_1; \ldots; \bar{\lambda}_t, p_t) \\ & \times\tilde{Y}(\lambda, 0; \lambda_1, p_1; \ldots; \lambda_t, p_t)[1+o(1)] \quad \text{as } n\to\infty. \end{aligned} \tag{5.56}$$

Proof. The proof of this theorem is similar to that of Theorem 5.1. Therefore, we shall provide an outline for the reader. Substituting (5.53) into the determinant expression (1.8) for $D(1, \lambda,\ldots,\lambda^k)$, and using the multilinearity property of determinants, we obtain an expansion similar to (5.35) involving the summation indices $j_1l_1,\ldots, j_kl_k$ and $h_1i_1,\ldots,h_ki_k$. It turns our that Lemma 5.3 can be applied twice—first to the summation over $j_1l_1, \ldots, j_kl_k$ and then to that over $h_1i_1, \ldots, h_ki_k$—resulting in

$$H_n(\lambda) \sim \sum_{h_1i_1 < \cdots < h_ki_k} \left(\prod_{p=1}^{k} \bar{\lambda}_{h_p}^n \right) Y\left(\Delta \binom{n}{i_1} \bar{\lambda}_{h_1}^0, \ldots, \Delta \binom{n}{i_k} \bar{\lambda}_{h_k}^0 \right)$$
$$\times \sum_{j_1l_1 < \cdots < j_kl_k} \left(\prod_{p=1}^{k} \lambda_{j_p}^n \right) Y\left(\binom{n}{0} \lambda^0, \Delta \binom{n}{l_1} \lambda_{j_1}^0, \ldots, \Delta \binom{n}{l_k} \lambda_{j_k}^0 \right) Z_{j_1l_1, \ldots, j_kl_k}^{h_1i_1, \ldots, h_k, i_k}, \tag{5.57}$$

which is the analogue of (5.33) in Lemma 5.4. (5.56) now follows from (5.57) in the same way (5.36) follows from (5.33). To complete the argument, we note that $Z_{10, \ldots, tp_t}^{10, \ldots, tp_t}$ in (5.56) is the Gram determinant of the linearly independent vectors y_{jl}, $10 \leqslant jl \leqslant tp_t$, and is thus non-zero. □

Theorem 5.4. *Define the vectors*

$$\tilde{z}_{i,l} = \begin{vmatrix} d_1 y_{10} & \cdots & d_1 y_{1p_1} & \cdots & d_t y_{t0} & \cdots & d_t y_{tp_t} & d_i y_{il} \\ z_{10}^{10} & \cdots & z_{1p_1}^{10} & \cdots & z_{t0}^{10} & \cdots & z_{tp_t}^{10} & z_{il}^{10} \\ \vdots & & \vdots & & \vdots & & \vdots & \vdots \\ z_{10}^{tp_t} & \cdots & z_{1p_1}^{tp_t} & \cdots & z_{t0}^{tp_t} & \cdots & z_{tp_t}^{tp_t} & z_{il}^{tp_t} \end{vmatrix}, \tag{5.58}$$

where $d_h = \lambda_h^q/(\lambda_h - 1)$. *Then with* r' *as defined in* (3.2), $G_n \equiv D(x_{n+q} - s, \ldots, x_{n+q+k} - s)$ *has the asymptotic behavior*

$$G_n \sim \left| \prod_{m=1}^{t} \lambda_m^{2n(p_m+1)} \right| \left| \prod_{m=1}^{t} (\lambda_m - 1)^{2(p_m+1)} \lambda_m^{p_m(p_m+1)} \right| \tilde{Y}(\bar{\lambda}_1, p_1; \ldots; \bar{\lambda}_t, p_t) \frac{n^{p_{t+1}}}{p_{t+1}!}$$
$$\times \sum_{i=1}^{r'} \lambda_{t+i}^n \tilde{z}_{t+i, p_{t+i}} (\lambda_{t+i} - 1) \tilde{Y}(\lambda_1, p_1; \ldots; \lambda_t, p_t; \lambda_{t+i}, 0). \tag{5.59}$$

Proof. The proof of (5.59) can be achieved using arguments similar to those found in the proofs of Theorems 5.2 and 5.3. The details are left to the reader. □

Theorem 3.1 for MPE can now be proved by dividing the asymptotic behavior of G_n in (5.59) by that for $D(1, 1, \ldots, 1)$ in (5.56), where the vector $\Gamma(n)$ of (3.4) is exactly of the form (5.46) with Z and $\tilde{z}_{i,l}$ of (5.46) being replaced by $Z_{10, \ldots, tp_t}^{10, \ldots, tp_t}$ of (5.55) and $\tilde{z}_{i,l}$ of (5.58) respectively.

Theorem 3.2 for MPE can be proved by considering the asymptotic behavior of $H_n(\lambda)$ divided by the asymptotic behavior of $H_n(1)$.

5.4. Convergence and stability proofs for RRE

From (1.13) and (5.27) we have

$$u_{h,i} = (w_{n+h}, u_{n+i}) \sim \sum_{mp} \sum_{jl} z_{jl}^{mp} \left[\Delta^2 \binom{n+h}{p} \bar{\lambda}_m^{n+h} \right] \left[\Delta \binom{n+i}{l} \lambda_j^{n+i} \right], \tag{5.60}$$

where z_{jl}^{mp} is as defined in (5.54) and $\Delta^2 \equiv \Delta\Delta$.

The proofs of this section are nearly identical to those for MPE, the only difference being that in the determinants Y having $\bar{\lambda}_j$ as their arguments, the operator Δ is replaced by Δ^2, c.f. (5.57). Ultimately, however, these determinants have no effect on the final results since they disappear from the dominant terms of $G_n/H_n(1)$ and $H_n(\lambda)/H_n(1)$. Consequently, the vector $\Gamma(n)$ for RRE is asymptotically equivalent to $\Gamma(n)$ for MPE.

6. Extensions of power method

From Theorem 3.2 it is apparent that, as $n \to \infty$, the zeros of the polynomial $\Sigma_{i=0}^{k} \gamma_i^{(n,k)} \lambda^i$ approach the λ_i, $i = 1, \ldots, t$, with corresponding multiplicities $p_i + 1$. Based on this observation , in this section we propose some old and some new extensions to the well known power method that is used to estimate the largest eigenvalue (in modulus) of a matrix A. These extensions enable us to estimate the first few dominant eigenvalues of the matrix A.

Let x_0, x_1, $x_2, \ldots,$ be a sequence of vectors in $\boldsymbol{B}$ satisfying

$$x_m \sim \sum_{i=1}^{\infty} P_i(m) \lambda_i^m \quad \text{as } m \to \infty, \tag{6.1}$$

where $P_i(m)$ and λ_i are exactly as described in Section 1 with the notation therein, with the exception that $\lambda_i \neq 1$, $i = 1, 2, \ldots,$ is not required.

A natural example for a sequence of this kind is one generated by the iterative procedure

$$x_{j+1} = Ax_j, \quad j = 0, 1, \ldots, \quad x_0 \text{ given}, \tag{6.2}$$

where A is the matrix of Section 2, with no restrictions being imposed on its spectrum. In fact, (6.1) can be obtained for this example, beginning with (2.4) and deleting s everywhere in Section 2.

Now a close look at the power method for the matrix A above reveals that this method actually approximates λ_1 in (6.1) provided $p_1 = 0$ and $|\lambda_1| > |\lambda_2|$, by utilizing only the vector sequence x_0, $x_1, \ldots,$ with any reference to the the matrix A being indirectly through the vectors x_0, x_1, ... With this in mind we now propose the following extensions to the power method for estimating the first few dominant λ_i in (6.1) counting multiplicities:

Let the vector sequence x_0, $x_1, \ldots,$ be as above. Construct the polynomial

$$P^{(n,k)}(\lambda) = \sum_{i=0}^{k} c_i^{(n,k)} \lambda^i, \quad c_k^{(n,k)} = 1, \tag{6.3}$$

where the coefficients $c_i = c_i^{(n,k)}$, $0 \leqslant i \leqslant k-1$, are determined in one of the following ways:

(1) MMPE extension

$$\sum_{j=0}^{k-1} c_j Q_i(x_{n+j}) = -Q_i(x_{n+k}), \quad 1 \leqslant i \leqslant k; \tag{6.4}$$

(2) TEA extension

$$\sum_{j=0}^{k-1} c_j Q(x_{n+i+j}) = -Q(x_{n+i+k}), \quad 0 \leqslant i \leqslant k-1; \tag{6.5}$$

(3) MPE extension

$$\sum_{j=0}^{k-1} c_j(x_{n+i}, x_{n+j}) = -(x_{n+i}, x_{n+k}), \quad 0 \leqslant i \leqslant k-1; \tag{6.6}$$

(4) RRE extension

$$\sum_{j=0}^{k-1} c_j(u_{n+i}, x_{n+j}) = -(u_{n+i}, x_{n+k}), \quad 0 \leqslant i \leqslant k-1. \tag{6.7}$$

Here Q_i, Q are as described in Section 1. Finally, the zeros of the polynomial $P^{(n,k)}(\lambda)$ are taken to be estimates of the first k most dominant λ_i including their multiplicities.

Note that all the extensions make use of the given vector sequence only.

If $\boldsymbol{B}$ is a complete inner product space, then for every bounded linear functional F on $\boldsymbol{B}$ there exists a unique vector $f \in \boldsymbol{B}$, such that $F(x) = (f, x)$ for every $x \in \boldsymbol{B}$. f is called the representer of F. A finite dimensional Euclidean space is a complete inner product space.

Taking $k = 1$ and letting $Q_1 = Q$ be represented by the vector q, we see that the MMPE and TEA extensions above are equivalent to the standard power method. The MPE extension reduces to the Rayleigh quotient for $k = 1$. The RRE extension, however, has no analogue that we know of.

The following theorem provides the justification for the extensions above.

Theorem 6.1. *Let the vector sequence $x_0, x_1, \ldots,$ be as described in the beginning of this section. Assume, in addition, that the λ_i satisfy* (3.1) *and let k be as in* (3.3). *If we write*

$$\prod_{i=1}^{t} (\lambda - \lambda_i)^{p_i+1} = \sum_{i=0}^{k} \delta_i \lambda^i,$$

then

$$c_j^{(n,k)} = \delta_j + O(n^{\alpha} |\lambda_{t+1}/\lambda_t|^n) \quad \text{as } n \to \infty, \tag{6.8}$$

where α is as described following the statement of Theorem 3.2. *Hence*

$$\lim_{n \to \infty} P^{(n,k)}(\lambda) = \prod_{i=1}^{t} (\lambda - \lambda_i)^{p_i+1}, \tag{6.9}$$

provided (3.6) *holds for MMPE,* (3.7) *for TEA, and $\lambda_i \neq 1$, $1 \leqslant i \leqslant t$, for RRE. No additional conditions are needed for MPE.*

Proof. In analogy to (5.29) it can be shown that $P^{(n,k)}(\lambda)$ can be expressed as

$$P^{(n,k)}(\lambda) = \frac{D(1, \lambda, \ldots, \lambda^k)}{D_k}, \tag{6.10}$$

where $D(1, \lambda, \ldots, \lambda^k)$ is as defined in (1.8) with u_{ij} there redefined as

$$u_{ij} = Q_{i+1}(x_{n+j}) \quad \text{for MMPE,} \tag{6.11}$$

$$u_{ij} = Q(x_{n+i+j}) \quad \text{for TEA,} \tag{6.12}$$

$$u_{ij} = (x_{n+i}, x_{n+j}) \quad \text{for MPE,} \tag{6.13}$$

$$u_{ij} = (u_{n+i}, x_{n+j}) \quad \text{for RRE,} \tag{6.14}$$

and D_k is the cofactor of λ^k in the expansion of $D(1, \lambda, \ldots, \lambda^k)$ with respect to its first row.

The proof now proceeds along the same lines as the proofs of Theorems 3.1 and 3.2. A cursery look at the asymptotic expansions for $D(1, \lambda, \dots, \lambda^k)$ reveals that it is not necessary to require $\lambda_i \neq 1$ for the cases of MMPE, TEA, and MPE. This condition is necessary for RRE, however. □

A method akin to the MMPE extension is suggested in Householder [5, p. 186, equation (15)]. The TEA extension is equivalent to that suggested in [5, p. 186, equation (16)]. These assertions can be verified by (6.10), (1.8), and (6.11) for the MMPE extension and (6.10), (1.8), and (6.12) for the TEA extension.

The MPE extension is a general version of that suggested by Wilkinson [14, pp. 583–584]. This assertion can be verified by observing that the coefficients c_i in the MPE extension are the solution to the problem

$$\underset{c_0,\dots,c_{k-1}}{\text{minimize}} \left\| \sum_{i=0}^{k-1} c_i x_{n+i} + x_{n+k} \right\|. \tag{6.15}$$

A different method is obtained if we normalize $P^{(n,k)}(\lambda)$ such that $P^{(n,k)}(0) = c_0^{(n,k)} = 1$. In this case (6.6) is replaced by

$$\sum_{j=1}^{k} c_j (x_{n+i}, x_{n+j}) = -(x_{n+i}, x_n), \quad 1 \leqslant i \leqslant k, \tag{6.16}$$

and (6.15) is replaced by

$$\underset{c_1,\dots,c_k}{\text{minimize}} \left\| x_n + \sum_{i=1}^{k} c_i x_{n+i} \right\|. \tag{6.17}$$

Theorem 6.1 holds true for this extension too.

Needless to say, in (6.15) and (6.17) the norm $\|\cdot\|$ induced by the inner product $(\cdot, \cdot)$ in $\boldsymbol{B}$ can be replaced by any other whenever this is possible. This results in further extensions of the power method. For example, when $\boldsymbol{B}$ is a finite dimensional space all l_p norms can be employed. In particular, the minimization problems associated with the l_1 and l_∞ norms can be solved using linear programming techniques.

It seems that the convergence results stated in Theorem 6.1, under conditions as general as those assumed there, have not been given before.

Different extensions of the power method based on all three epsilon algorithms, SEA, VEA, and TEA have been given by Brezinski [2]. Brezinski shows convergence under the restrictions that $\lambda_i \neq 1$, $p_i = 0$, and $|\lambda_1| > |\lambda_2| > \dots$.

Note: If we now assume that the sequence x_m, $m = 0, 1, \dots$, is a scalar sequence, i.e., the polynomials $P_i(m)$ in (6.1) have scalar coefficients, then Theorem 6.1 is valid for the method obtained from the TEA extension by deleting Q from (6.5). In this case the expansion given in (6.1) is a generalized Dirichlet series, and the method described above provides estimates for $\lambda_1, \dots, \lambda_t$ in this expansion, taking their respective multiplicities into account at the same time.

References

[1] C. Brezinski, Généralisations de la transformation de Shanks, de la table de Padé et de l'ϵ-algorithm, *Calcolo* **12** (1975) 317–360.

[2] C. Brezinski, Computation of the eigenelements of a matrix by the ϵ-algorithm, *Lin. Algebra Appl.* **11** (1975) 7–20.

[3] S. Cabay and L.W. Jackson, A polynomial extrapolation method for finding limits and antilimits of vector sequences, *SIAM J. Numer. Anal.* **13** (1976) 734–752.

[4] R.P. Eddy, Extrapolating to the limit of a vector sequence, in: P.C.C. Wang, Ed., *Information Linkage Between Applied Mathematics and Industry* (Academic Press, New York, 1979) 387–396.

[5] A.S. Householder, *The Theory of Matrices in Numerical Analysis* (Blaisdell, New York, 1964).

[6] D. Kalman, The generalized Vandermonde matrix, *Mathematics Magazine* **57** (1984) 15–21.

[7] S. Kaniel and J. Stein, Least-square acceleration of iterative methods for linear equations, *J. Optim. Theory Appl.* **14** (1974) 431–437.

[8] M. Mešina, Convergence acceleration for the iterative solution of the equations $X = AX + f$, *Comp. Meth. Appl. Mech. Engn.* **10** (1977) 165–173.

[9] D. Shanks, Non-linear transformations of divergent and slowly convergent sequences, *J. Math. Phys.* **34** (1955) 1–42.

[10] A. Sidi, Convergence and stability properties of minimal polynomial and reduced rank extrapolation algorithms, *SIAM J. Numer. Anal.* **23** (1986) 197–209.

[11] A. Sidi, W.F. Ford and D.A. Smith, Acceleration of convergence of vector sequences, *SIAM J. Numer. Anal.* **23** (1986) 178–196.

[12] D.A. Smith, W.F. Ford and A. Sidi, Extrapolation methods for vector sequences, *SIAM Rev.* **29** (1987) 199–233; see also Correction to "Extrapolation methods for vector sequences", submitted.

[13] R.S. Varga, *Matrix Iterative Analysis* (Prentice-Hall, Englewood Cliffs, NJ, 1962).

[14] J.H. Wilkinson, *The Algebraic Eigenvalue Problem* (Clarendon Press, Oxford, 1965).

[15] P. Wynn, On a device for computing the $e_m(S_n)$ transformation, *MTAC* **10** (1956) 91–96.

[16] P. Wynn, Acceleration techniques for iterated vector and matrix problems, *Math. Comp.* **16** (1962) 301–322.

[17] P. Wynn, On the convergence and stability of the epsilon-algorithm, *SIAM J. Numer. Anal.* **3** (1966) 91–122.