Numer. Math. (2011) 119:725–764
DOI 10.1007/s00211-011-0398-8

Numerische Mathematik

Acceleration of convergence of general linear sequences by the Shanks transformation

Avram Sidi

Received: 5 January 2010 / Revised: 24 February 2011 / Published online: 14 July 2011

Abstract The Shanks transformation is a powerful nonlinear extrapolation method that is used to accelerate the convergence of slowly converging, and even diverging, sequences $\{A_n\}$. It generates a two-dimensional array of approximations $A_n^{(j)}$ to the limit or anti-limit of $\{A_n\}$ defined as solutions of the linear systems

$$A_l = A_n^{(j)} + \sum_{k=1}^{n} \bar{\beta}_k (\Delta A_{l+k-1}), \quad j \le l \le j+n,$$

where $\bar{\beta}_k$ are additional unknowns. In this work, we study the convergence and stability properties of $A_n^{(j)}$, as $j \to \infty$ with n fixed, derived from general linear sequences $\{A_n\}$, where $A_n \sim A + \sum_{k=1}^{m} \zeta_k^n \sum_{i=0}^{\infty} \beta_{ki} n^{\gamma_k - i}$ as $n \to \infty$, where $\zeta_k \neq 1$ are distinct and $|\zeta_1| = \cdots = |\zeta_m| = \theta$, and $\gamma_k \neq 0, 1, 2, \ldots$. Here A is the limit or the anti-limit of $\{A_n\}$. Such sequences arise, for example, as partial sums of Fourier series of functions that have finite jump discontinuities and/or algebraic branch singularities. We show that definitive results are obtained with those values of n for which the integer programming problems

$$\max_{s_1,\ldots,s_m} \sum_{k=1}^{m} \left[(\Re\gamma_k) s_k - s_k(s_k - 1) \right],$$

$$\text{subject to } \; s_1 \ge 0, \ldots, s_m \ge 0 \quad \text{and} \quad \sum_{k=1}^{m} s_k = n,$$

A. Sidi (✉)
Computer Science Department, Technion-Israel Institute of Technology, Haifa 32000, Israel
e-mail: asidi@cs.technion.ac.il
URL: http://www.cs.technion.ac.il/~asidi

have unique (integer) solutions for $s_1, \ldots, s_m$. A special case of our convergence result concerns the situation in which $\Re\gamma_1 = \cdots = \Re\gamma_m = \alpha$ and $n = m\nu$ with $\nu = 1, 2, \ldots,$ for which the integer programming problems above have unique solutions, and it reads $A_n^{(j)} - A = O(\theta^j j^{\alpha-2\nu})$ as $j \to \infty$. When compared with $A_j - A = O(\theta^j j^{\alpha})$ as $j \to \infty$, this result shows that the Shanks transformation is a true convergence acceleration method for the sequences considered. In addition, we show that it is stable for the case being studied, and we also quantify its stability properties. The results of this work are the first ones pertaining to the Shanks transformation on general linear sequences with $m > 1$.

Mathematics Subject Classification (2000) 40A05 · 40A25 · 41A60 · 65B05 · 65B10

1 Introduction

The *Shanks transformation* [27] is a very powerful convergence acceleration (or extrapolation) method that is used to approximate limits or so-called antilimits of infinite sequences or to sum infinite series, whether convergent or divergent. Given an infinite sequence $\{A_n\}$, this transformation produces a two-dimensional array of quantities $A_n^{(j)}$ that approximate the limit of $\{A_n\}$ in case of convergence and the antilimit of $\{A_n\}$ in case of divergence. These approximations are defined via the linear systems

$$A_l = A_n^{(j)} + \sum_{k=1}^{n} \bar{\beta}_k (\Delta A_{l+k-1}), \quad j \le l \le j+n, \tag{1.1}$$

where $\bar{\beta}_k$ are additional auxiliary unknowns. Here, $\Delta A_k = A_{k+1} - A_k$ for all k.

Solving (1.1) for $A_n^{(j)}$ by Cramer's rule, we obtain the following determinant representation for $A_n^{(j)}$:

$$A_n^{(j)} = \frac{\begin{vmatrix} A_j & \Delta A_j & \Delta A_{j+1} & \cdots & \Delta A_{j+n-1} \\ A_{j+1} & \Delta A_{j+1} & \Delta A_{j+2} & \cdots & \Delta A_{j+n} \\ \vdots & \vdots & \vdots & & \vdots \\ A_{j+n} & \Delta A_{j+n} & \Delta A_{j+n+1} & \cdots & \Delta A_{j+2n-1} \end{vmatrix}}{\begin{vmatrix} 1 & \Delta A_j & \Delta A_{j+1} & \cdots & \Delta A_{j+n-1} \\ 1 & \Delta A_{j+1} & \Delta A_{j+2} & \cdots & \Delta A_{j+n} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & \Delta A_{j+n} & \Delta A_{j+n+1} & \cdots & \Delta A_{j+2n-1} \end{vmatrix}}. \tag{1.2}$$

This representation of $A_n^{(j)}$ turns out to be extremely useful in analyzing the properties of $A_n^{(j)}$. It is not meant to be a computational tool for evaluating $A_n^{(j)}$, however. Instead, the computation of the $A_n^{(j)}$ can be achieved, without resorting to either (1.1) or

(1.2), and in a most efficient way, via the famous *epsilon algorithm* due to Wynn [44]. Another algorithm, called the *FS/qd algorithm*, has been given recently in Sidi [34, Chapter 21], and it turns out to be practically as efficient as the epsilon algorithm. The FS/qd algorithm is a combination of the FS algorithm of Ford and Sidi [11] (see also Sidi [34, Chapter 3]) and the qd algorithm of Rutishauser [25], and it implements the higher-order G transformation of Gray et al. [15] as well.

Two types of sequences of the $A_n^{(j)}$ are normally considered: (i) $\{A_n^{(j)}\}_{j=0}^{\infty}$ with fixed n (called *column sequences*), and (ii) $\{A_n^{(j)}\}_{n=0}^{\infty}$ with fixed j (called *diagonal sequences*). Experience suggests that diagonal sequences have much better convergence properties; they are extremely difficult to study rigorously, however. Column sequences turn out to be relatively easy to study, and, in general, the conclusions drawn from this study seem to be relevant for diagonal sequences as well.

In this work, we will treat some of the convergence and stability properties of column sequences produced by the Shanks transformation as applied to *general linear sequences* $\{A_n\}$, whose terms behave specifically as in

$$A_n \sim A + \sum_{k=1}^{m} \zeta_k^n \sum_{i=0}^{\infty} \beta_{ki} n^{\gamma_k - i} \quad \text{as } n \to \infty; \quad \beta_{k0} \neq 0, \quad k = 1, \dots, m, \tag{1.3}$$

where $m > 1$ is arbitrary and

$$|\zeta_1| = \cdots = |\zeta_m| = \theta; \quad \zeta_k \neq 1 \text{ distinct}, \quad \gamma_k \neq 0, 1, 2, \dots, \text{arbitrary.}^1 \tag{1.4}$$

Note that the assumption that the ζ_k have equal moduli is not a restriction, because the ζ_k that dominate in the asymptotic expansion of A_n, and hence in our results in this paper, are only the largest ones that have the same modulus, while the remaining ones are subdominant and hence do not contribute asymptotically. (This will become obvious as we proceed with our treatment later.) Everything in (1.3) can be real or complex. We also assume that $|\zeta_k| \leq 1$. Hence $\lim_{n\to\infty} A_n$ exists and equals A unconditionally when $|\zeta_k| < 1$. When $|\zeta_k| = 1$, $\lim_{n\to\infty} A_n$ exists and equals A provided $\Re\gamma_k < 0$ for all k. Again, when $|\zeta_k| = 1$, (i) if $\Re\gamma_k \leq 0$, with equality for at least one value of k, $\{A_n\}$ diverges but remains bounded, and (ii) if $\Re\gamma_k > 0$ for at least one value of k, $\{A_n\}$ diverges and is unbounded. In case $|\zeta_k| = 1$, we have that $\arg \zeta_k \in (0, 2\pi)$ because $\zeta_k \neq 1$. Therefore, $\{A_n\}$, whether convergent or divergent, is oscillatory as a function of n since $\zeta_k^n = \exp(\mathrm{i}\, n \arg \zeta_k)$. In case of divergence, A is the antilimit of $\{A_n\}$. As we will see in the next section, such sequences arise naturally as partial sums of Fourier series of functions with various types of singularities, for example.

Before we continue, we would like to discuss briefly the subject of stability. We know that (see [34, p. 307, Eqs. (16.4.23), (16.4.24)]) $A_n^{(j)}$ can be expressed as in

$$A_n^{(j)} = \sum_{i=0}^{n} \gamma_{ni}^{(j)} A_{j+i}, \tag{1.5}$$

[1] When $m = 1$, we call $\{A_n\}$ simply a *linear sequence*. Thus, linear sequences are the simplest special cases of general linear sequences.

for some scalars $\gamma_{ni}^{(j)}$ that satisfy

$$\sum_{i=0}^{n} \gamma_{ni}^{(j)} = 1. \tag{1.6}$$

(The exact nature of the $\gamma_{ni}^{(j)}$ is discussed in Lemma 4.1 of this work.) The stability of the extrapolation process is ultimately connected with the quantity

$$\Gamma_n^{(j)} = \sum_{i=0}^{n} \left|\gamma_{ni}^{(j)}\right|, \tag{1.7}$$

which determines the rate at which errors in the A_l propagate into $A_n^{(j)}$. Note that, in view of (1.6), we have $\Gamma_n^{(j)} \geq 1$ always. In order for the column and diagonal sequences to be stable numerically, it is necessary that $\sup_j \Gamma_n^{(j)}$ and $\sup_n \Gamma_n^{(j)}$, respectively, be finite. For a survey of the issue of stability in extrapolation methods, see the recent paper by Sidi [38].

In this work, we study the convergence and stability behavior of $A_n^{(j)}$ as $j \to \infty$, when $m > 1$ in (1.3). Specifically, we analyze the asymptotic behavior, as $j \to \infty$, of the error $A_n^{(j)} - A$ and of the polynomial $\sum_{i=0}^{n} \gamma_{ni}^{(j)} z^i$. We show that, for infinitely many (but not all) values of n, $\lim_{j\to\infty} \sum_{i=0}^{n} \gamma_{ni}^{(j)} z^i$ exists, from which we conclude that $\lim_{j\to\infty} \gamma_{ni}^{(j)}$ all exist, and hence conclude that $\sup_j \Gamma_n^{(j)}$ is bounded, implying that the column sequences resulting from application of the Shanks transformation to general linear sequences are stable. We also show that convergence acceleration does take place and we quantify it.

The plan of this paper is as follows: Following some examples of sequences $\{A_n\}$ that satisfy (1.3), which we give in the next section. In Sect. 3, we state the main results for $j \to \infty$, while n is being held fixed. In Sect. 4, we provide some technical preliminaries, including determinant representations of $A_n^{(j)} - A$ and of the polynomial $\sum_{i=0}^{n} \gamma_{ni}^{(j)} z^i$. Using the representations of Sect. 4, in Sects. 5 and 6, we give the proofs of the main results. In view of the main results, in Sect. 7, we address the problem of slow convergence of the Shanks transformation that is present when $\zeta_k \approx 1$ in the complex plane for some k, and we justify the use of the so-called *arithmetic progression sampling* (APS) (see [34, Chapter 16, p. 316]) to make $A_n^{(j)}$ converge faster. This amounts to applying the Shanks transformation to a subsequence $\{A_{\kappa n}\}$ for some integer $\kappa \geq 2$. Finally, in Sect. 8, we give a numerical example.

Our main results rely on the existence of unique solutions to an interesting integer programming problem. Unique solutions exist only for certain values of n. Several aspects of this problem that are relevant to our convergence study are treated in detail in the appendix to this work, where we also construct infinitely many values of n, for which the relevant integer programming problem has unique solutions.

The results of the present work are the first ones that pertain to the application of the Shanks transformation to general linear sequences in (1.3)–(1.4) with $m > 1$, as opposed to $m = 1$. As will become clear, the analysis turns out to be quite involved and nontrivial, and the main results are interesting and nonintuitive. Our results here are analogous to, yet different from, those of Sidi [36] that pertain to the application of the author's generalization of the Richardson extrapolation process GREP$^{(m)}$ to general linear sequences with $m > 1$. For GREP$^{(m)}$, see [34, Chapter 4], for example.

The Shanks transformation is treated in various books. See, for example, Brezinski [7] and Brezinski and Redivo Zaglia [9]. For an up-to-date account covering the most recent developments as well, see Sidi [34, Chapter 16].

As shown by Shanks in [27], if $A_n = \sum_{i=0}^{n} c_i z^i$, $n = 0, 1, \ldots,$ then $A_n^{(j)}$ turns out to be the $[j + n/n]$ Padé approximant from the infinite series $\sum_{i=0}^{\infty} c_i z^i$. There is an enormous amount of literature on Padé approximants. See, for example, the extensive treatments in the books by Baker [2], Baker and Graves-Morris [3], and Gilewicz [14]. See also Sidi [34, Chapter 17] for a brief survey.

As already mentioned, the Shanks transformation has presented many difficulties as far as rigorous and interesting convergence studies are concerned, one reason for this being that it is extremely nonlinear. Consequently, there have been very few works dealing with the subjects of convergence and stability relevant to this important sequence transformation. Here is a brief review of these works and their relevant results:

(i) Wynn [45] has shown that convergence acceleration is achieved by column sequences when

$$A_n \sim A + \sum_{k=1}^{\infty} \alpha_k \lambda_k^n \quad \text{as } n \to \infty, \tag{1.8}$$

where the $\lambda_k \neq 1$ are either all positive or all negative and $|\lambda_1| > |\lambda_2| > \cdots$, and $\lim_{k\to\infty} \lambda_k = 0$. The result of Wynn is extended in Sidi [31] to cover the more general case in which the λ_k are real or complex and $|\lambda_1| \geq |\lambda_2| \geq \cdots$, without requiring that the λ_k be either all positive or all negative.

(ii) Sidi [31] has shown that convergence acceleration is achieved by column sequences in the more general case of

$$A_n \sim A + \sum_{k=1}^{\infty} P_k(n) \lambda_k^n \quad \text{as} \quad n \to \infty, \tag{1.9}$$

where $P_k(n)$ are polynomials in n, the $\lambda_k \neq 1$ are arbitrary real or complex scalars, $|\lambda_1| \geq |\lambda_2| \geq \cdots$, and $\lim_{k\to\infty} \lambda_k = 0$. In this case, we may not expect every column sequence to converge; that is, $\lim_{j\to\infty} A_n^{(j)}$ may not exist for every fixed n. A precise classification of those values of n, for which convergence takes place, is given in [29] and [31]. See also Kaminski and Sidi [18] for more on these values of n.

(iii) Garibotti and Grinstein [12] have shown that convergence acceleration is achieved by column sequences when $\{A_n\}$ is a *linear sequence*, that is,

$$A_n \sim A + \zeta^n \sum_{i=0}^{\infty} \alpha_i n^{\gamma - i} \quad \text{as } n \to \infty, \tag{1.10}$$

where $\zeta \neq 1$ and γ are arbitrary. (The special case in which $\zeta = -1$ and $\gamma = -1$ was first treated by Wynn [45]).

We note that all three papers, [12,31,45], provide precise rates of convergence for column sequences. None of the three papers considers the issue of stability, however. This issue is treated in [34, Chapter 16]. We also note that the papers [45] and [31] are actually closely related to the theorem of de Montessus [10] that concerns the convergence of Padé approximants for meromorphic functions. For different treatments of de Montessus' theorem, see Saff [26], Karlsson and Wallin [19], and Sidi [29]. See also the books [2,3,14,34].

(iv) For sequences $\{A_n\}$ whose terms behave as in

$$A_n \sim A + \sum_{i=0}^{\infty} \alpha_i n^{\gamma-i} \quad \text{as } n \to \infty, \tag{1.11}$$

where $\gamma \neq 0, 1, \ldots,$ but is arbitrary otherwise, it is shown in [34, Theorem 16.5.4, p. 316] that column sequences from the Shanks transformation produce no convergence acceleration at all. (The special case in which $\gamma = -1$ was first treated by Wynn [45]).

(v) For sequences $\{A_n\}$, where $A_n = \sum_{k=0}^{n} c_k z^k$, $c_k \neq 0$ for all k, are partial sums of Maclaurin series of entire functions $f(z)$, the convergence of the $A_n^{(j)}$ follows from the results of Lubinsky [22,23]. For example, in [23], Lubinsky has shown that both row and diagonal Padé approximants converge on any compact set of the z-plane provided $\lim_{k\to\infty} c_{k+1}c_{k-1}/c_k^2 = q$, where q is some possibly complex constant such that $|q| < 1$. See the bibliography of [23] for more references on this specific topic.

(vi) Finally, through the study of Padé approximants from Markov functions, we know a lot about the convergence of diagonal sequences of $A_n^{(j)}$ in case $A_n = \sum_{k=0}^{n} c_k z^k$, when $c_k = \int_a^b x^k d\alpha(x)$, where $\alpha(x)$ is a nonnegative function with an infinite number of points of increase on $[a, b]$. This subject is very well developed and is treated in great detail in the books [2,3,14]. Actually, the Gaussian quadrature formulas result from the (diagonal) $[n-1/n]$ Padé approximants from Markov functions. For Gaussian quadrature, see Ralston and Rabinowitz [24], Atkinson [1], Stoer and Bulirsch [43], and Gautschi [13], for example.

The few theoretical studies we have mentioned are of crucial importance, because the conclusions drawn from them suggest that the Shanks transformation will produce convergence acceleration when applied to the sequences described above in (1.8), (1.9), (1.10), and to the partial sums of Taylor series from entire functions and from moment series of Markov functions, but that it will fail on those sequences in (1.11). Our purpose here is to develop this area of research further and to shed some light on the analytic properties of the Shanks transformation as it is applied to the class of general linear sequences. This is a very comprehensive class of sequences that arise commonly in applications, Fourier series, generalized Fourier series, and series of special functions being important members of this class.

The Shanks transformation, along with several other convergence acceleration methods, has also been the subject of an extensive comparative numerical survey by Smith and Ford [41,42]. The important methods compared in this survey are the iterated Aitken Δ^2 process, the u transformation of Levin [20], the θ algorithm of Brezinski [6], and the Shanks transformation (via the epsilon algorithm). The conclusions of this survey pertaining to these four methods are as follows: (i) all four methods accelerate the convergence of sequences satisfying (1.10), (ii) only the u transformation and the θ algorithm accelerate the convergence of sequences satisfying (1.11), and (iii) only the Shanks transformation accelerates the convergence of Fourier series. (The partial sums of the test series used in [42] form general linear sequences exactly of the form we consider in the present work.) Another method, not included in this survey, that is a very effective accelerator for Fourier series and their generalizations is the d transformation of Levin and Sidi [21]. This method is also an effective accelerator for the sequences satisfying (1.10), (1.11), and (1.3) in general; for its application to such series in two different ways, see also Sidi [30]. (For an up-to-date account of all the methods mentioned here, see [34, Chapters 6, 15, 16, 19, 20]).

2 Examples of general linear sequences

General linear sequences $\{A_n\}$ of the form described in (1.3), even with distinct ζ_k that satisfy the more general condition $|\zeta_k| \leq 1$ but do not necessarily satisfy $|\zeta_1| = \cdots = |\zeta_m|$, have been studied rigorously in [34, Chapter 6, subsection 6.8.2]. If we let $a_0 = A_0$ and $a_n = \Delta A_{n-1} = A_n - A_{n-1}, n \geq 1$, then $A_n = \sum_{i=0}^{n} a_i$ for every n. Now, by Theorem 6.8.8 in [34], it follows that if

$$a_n \sim \sum_{k=1}^{m} \zeta_k^n \sum_{i=0}^{\infty} \epsilon_{ki} n^{\gamma_k - i} \quad \text{as } n \to \infty; \quad \epsilon_{k0} \neq 0, \quad k = 1, \ldots, m, \tag{2.1}$$

where

$$|\zeta_k| \leq 1, \quad \zeta_k \neq 1 \text{ distinct}, \quad \gamma_k \text{ arbitrary}, \quad k = 1, \ldots, m, \tag{2.2}$$

then A_n is precisely as in (1.3), A there being $\lim_{n\to\infty} A_n = \sum_{i=0}^{\infty} a_i$ in case of convergence; in case of divergence, A is the *Abel sum* of $\sum_{i=0}^{\infty} a_i$. (For Abel summability of infinite series, see Hardy [17], for example).

2.1 Sums of simple linear sequences

One source of general linear sequences is the set of sequences $\{A_n\}$ that behave as in (1.3) with $m = 1$, that is,

$$A_n \sim A + \zeta^n \sum_{i=0}^{\infty} \beta_i n^{\gamma - i} \quad \text{as } n \to \infty, \quad \zeta \neq 1 \text{ and } \gamma \neq 0, 1, \ldots, \text{arbitrary}. \tag{2.3}$$

Such sequences are quite common and are said to be *linearly convergent* in case $\lim_{n\to\infty} A_n$ exists. $\lim_{n\to\infty} A_n$ exists if either (i) $|\zeta| < 1$ or (ii) $|\zeta| = 1$ and $\Re\gamma < 0$. In case this limit exists, it is equal to A. Such sequences arise as partial sums of infinite series $\sum_{s=0}^{\infty} a_s$, namely, $A_n = \sum_{s=0}^{n} a_s$, $n = 1, 2, \ldots,$ when a_n satisfies

$$a_n \sim \zeta^n \sum_{i=0}^{\infty} \alpha_i n^{\gamma-i} \quad \text{as } n \to \infty, \quad \zeta \neq 1 \text{ and } \gamma \neq 0, 1, \ldots, \text{arbitrary.} \tag{2.4}$$

See [34, Theorem 6.6.6, p. 145] for a proof of this statement.

In view of the above, it is clear that if $A_n = \sum_{i=0}^{n} a_i$, where

$$a_n = \sum_{k=1}^{m} a_n^{(k)}; \quad a_n^{(k)} \sim \zeta_k^n \sum_{i=0}^{\infty} \alpha_{ki} n^{\gamma_k - i} \quad \text{as } n \to \infty,$$
$$\zeta_k \neq 1; \text{ distinct and arbitrary}, \quad \gamma_k \neq 0, 1, \ldots, \text{arbitrary,} \tag{2.5}$$

then A_n has an asymptotic expansion of the form given in (1.3).

2.2 General linear sequences from Fourier series

Another source of interest is the set of Fourier series of functions that have a number of singularities. Recall that if $f(x)$ is a 2π-periodic function, then its Fourier series is $\sum_{s=-\infty}^{\infty} c_s e^{\mathrm{i}sx}$, where $c_s = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{-\mathrm{i}sx}\,dx$. If $f(x)$ is infinitely differentiable on $[-\pi, \pi)$, except at finitely many points in $[-\pi, \pi)$, where it has algebraic singularities of different strengths and/or finite jump discontinuities, then the partial sums $\sum_{s=-n}^{n} c_s e^{\mathrm{i}sx}$ of this series are precisely as described in (1.3) with $|\zeta_k| = 1$ for all k, m being twice the number of the points of singularity in $[0, 2\pi)$.

Before proceeding to examples, we would like to note that the Shanks transformation was applied to Fourier series first by Wynn [46]. Also, the detection and treatment of the Gibbs phenomenon via the Shanks transformation has been the subject of the papers by Brezinski [8], Guilpin et al. [16], and Beckermann et al. [4]. The Gibbs phenomenon can also be treated effectively by using *arithmetic progression sampling* (APS) when applying the d transformation of Levin and Sidi [21], the Shanks transformation, and various other nonlinear acceleration methods. This subject is considered in detail in several chapters of the author's book [34] and in the relevant references therein, as mentioned in the Introduction.

Example 2.1 Let us consider

$$f(x) = \begin{cases} 0 & -\pi \le x < a \\ (x-a)^{\alpha}(b-x)^{\beta} g(x) & a \le x \le b \\ 0 & b < x \le \pi \end{cases},$$

where $g \in C^{\infty}[a, b]$. That is, $f(x)$ has singularities at $x = a$ and $x = b$ only. When α or β is a nonnegative integer, $f(x)$ or some of its derivatives have finite jump discontinuities at the respective points. Otherwise, it has an algebraic singularity at $x = a$ or $x = b$, respectively.

It is known that

$$c_{\pm n} \sim e^{\mp \mathrm{i}na} \sum_{s=0}^{\infty} b_{1s}^{\pm} n^{-\alpha-1-s} + e^{\mp \mathrm{i}nb} \sum_{s=0}^{\infty} b_{2s}^{\pm} n^{-\beta-1-s} \quad \text{as } n \to \infty. \tag{2.6}$$

For a proof of this, see, for example, Bleistein and Handelsman [5, Sect. 3.4]. For a different and simpler proof, see Sidi [37].

For simplicity, let us first assume that $f(x)$ is real so that twice the real part of the series $\frac{1}{2}c_0 + \sum_{s=1}^{\infty} c_s e^{\mathrm{i}sx}$ gives $f(x)$. Thus, the partial sums $A_n = \frac{1}{2}c_0 + \sum_{s=1}^{n} c_s e^{\mathrm{i}sx}$ are of the form $A_n = \sum_{s=0}^{n} a_s$ with $a_s = c_s e^{\mathrm{i}sx}$ for $s \geq 1$. By (2.6), a_n has an asymptotic expansion as $n \to \infty$ that is precisely as in (2.5), with $m = 2$, $\zeta_1 = e^{\mathrm{i}(x-a)}$ and $\zeta_2 = e^{\mathrm{i}(x-b)}$, and $\gamma_1 = -\alpha - 1$ and $\gamma_2 = -\beta - 1$. Clearly, $|\zeta_1| = |\zeta_2| = 1$ and, provided $x \neq a, b$, there holds $\zeta_1, \zeta_2 \neq 1$ in this case, so that (1.4) is satisfied; consequently, A_n satisfies (1.3) with $m = 2$ by [34, p. 156, Theorem 6.8.8]. (It is easy to verify that when $f(x)$ has a finite jump discontinuity at $x = a$, but is infinitely differentiable to the right and to the left of $x = a$, we have $\alpha = -1$ necessarily. Similarly, for $x = b$).

Let us next take a look at the partial sums $A_n = \sum_{s=-n}^{n} c_s e^{\mathrm{i}sx}$ of the whole Fourier series $\sum_{s=-\infty}^{\infty} c_s e^{\mathrm{i}sx}$. Rewriting A_n in the form $A_n = \sum_{s=0}^{n} a_s$, where $a_s = c_s e^{\mathrm{i}sx} + c_{-s} e^{-\mathrm{i}sx}$, and invoking (2.6), it is easy to see that A_n, the nth partial sum of the whole Fourier series, satisfies (1.3) with $m = 4$ and $\zeta_1 = e^{\mathrm{i}(x-a)} = \zeta_3^{-1}$, $\zeta_2 = e^{\mathrm{i}(x-b)} = \zeta_4^{-1}$, $\gamma_1 = -\alpha - 1 = \gamma_3$, and $\gamma_2 = -\beta - 1 = \gamma_4$.

Example 2.2 In general, suppose that $f(x)$ is in $C^{\infty}[-\pi, \pi)$, except at the points $x_1, \ldots, x_p$, where it has discontinuities exactly of the forms described above. Then, in general, the partial sum $A_n = \frac{1}{2}c_0 + \sum_{s=1}^{n} c_s e^{\mathrm{i}sx}$ is precisely as in (1.3) with $m = p$ and $\zeta_k = e^{\mathrm{i}(x-x_k)}$, $k = 1, \ldots, p$, and with appropriate γ_k that are determined exactly as in (2.6). The partial sums $A_n = \sum_{s=-n}^{n} c_s e^{\mathrm{i}sx}$, on the other hand, satisfy (1.3) with $m = 2p$, $\zeta_k = e^{\mathrm{i}(x-x_k)} = \zeta_{p+k}^{-1}$, $k = 1, \ldots, p$, and $\gamma_k = \gamma_{p+k}$. In case $f(x)$ has only finite jump discontinuities at all x_k, we have $\gamma_k = -1$ for all k.

Example 2.3 Now, $\gamma_k = -1$ can also occur when $f(x)$ has a logarithmic singularity. Consider, for example,

$$-\log(1-z) = \sum_{s=1}^{\infty} \frac{z^s}{s}, \quad |z| \leq 1, \quad z \neq 1, \quad -\pi \leq \arg z < \pi.$$

By (2.4) in the preceding subsection, it is clear that $A_n = \sum_{s=1}^{n} z^s/s$ satisfies (2.3) with $\zeta = z$ and $\gamma = -1$.

Letting now $z = e^{\mathrm{i}\theta}$ and taking the real and imaginary parts of this series, we obtain the Fourier series

$$\sum_{s=1}^{\infty} \frac{\cos s\theta}{s} = -\log\left(2\left|\sin\frac{\theta}{2}\right|\right)$$

and

$$\sum_{s=1}^{\infty} \frac{\sin s\theta}{s} = \begin{cases} -\frac{1}{2}(\theta + \pi), & -\pi < \theta < 0, \\ -\frac{1}{2}(\theta - \pi), & 0 < \theta < \pi. \end{cases}$$

For both series, A_n, the nth partial sum, is of the form given in (1.3) with $m = 2$, $\zeta_1 = e^{i\theta} = \zeta_2^{-1}$, and $\gamma_1 = \gamma_2 = -1$. Note that, the sum of the cosine series has a logarithmic singularity at $\theta = 0$, while the sum of the sine series has a finite jump discontinuity at $\theta = 0$. Either sum has no other singularity in $[-\pi, \pi)$.

2.3 General linear sequences from generalized Fourier series

Yet another source of general linear sequences is the family of series of orthogonal polynomials and other generalized Fourier series. This topic is discussed in detail in [30], [34, Chapter 13]. For example, we know that if $A_n = \sum_{s=0}^{n} c_s P_s(x)$, where $P_s(x)$ is the sth Legendre polynomial and

$$c_n \sim \sum_{i=0}^{\infty} \alpha_i n^{\delta-i} \quad \text{as} \quad n \to \infty,$$

then A_n satisfies (1.3) with $m = 2$ and $\zeta_1 = e^{i\theta} = \zeta_2^{-1}$, where $\theta \in (0, \pi)$ is determined via $x = \cos\theta$, and $\gamma_1 = \gamma_2 = \delta - 1/2$. Of course, $-1 \leq x < 1$ must hold.

Cases like this arise, for example, when the infinite series $\sum_{s=0}^{\infty} c_s P_s(x)$ represents a function $f(x)$ on $(-1, 1)$, and $f(x)$ has one endpoint singularity at $x = 1$ and is of the form $f(x) = (1-x)^{\alpha} g(x)$ with $\alpha \neq 0, 1, 2, \ldots$, and $g \in C^{\infty}[-1, 1]$. Then $\gamma_1 = \gamma_2 = -2\alpha - 3/2$. See the recent paper by Sidi [35] for this case and more general cases of endpoint singularities. For yet more general cases involving arbitrarily many algebraic interior singularities as well, see Sidi [39]. In all these cases, the partial sums of the Legendre series form general linear sequences that are precisely of the form described in (1.3) and (1.4).

3 Main results

We assume that the sequence $\{A_n\}_{n=0}^{\infty}$ is precisely as in (1.3) with (1.4), and we order the γ_k in (1.3) as in

$$\Re\gamma_1 \geq \Re\gamma_2 \geq \cdots \geq \Re\gamma_m. \tag{3.1}$$

Throughout, we assume that the positive integer n in $A_n^{(j)}$ is such that the integer programming problem

$$\max_{s_1,\ldots,s_m} g(s_1, \ldots, s_m); \quad g(s_1, \ldots, s_m) = \sum_{k=1}^{m} \left[s_k(\Re\gamma_k) - s_k(s_k - 1) \right]$$
$$\text{subject to} \quad s_1 \geq 0, \ldots, s_m \geq 0 \quad \text{and} \quad \sum_{k=1}^{m} s_k = n; \quad s_k \text{ integers}, \tag{3.2}$$

has a unique solution for integer s_k, which we shall denote by $(s_1', \ldots, s_m')$. This assumption is sufficient to guarantee that the $A_n^{(j)}$ exist for all large j; in general, the convergence and stability study seems to be inconclusive without it. The problem in (3.2) is studied in detail in the appendix to this work (with α_k instead of $\Re\gamma_k$), where it is denoted by $\mathcal{IP}_n$. The value of $g(s_1, \ldots, s_m)$ at the (optimal) solution is denoted by σ_n. The following results are proved in the appendix:

1. For the special case in which $\Re\gamma_k$ are all equal, it is shown in the appendix that a unique solution to $\mathcal{IP}_n$ exists only for $n = m\nu$, with $\nu \in \{1, 2, \ldots\}$, and that this solution is given as $s'_k = \nu$, $k = 1, \ldots, m$.
2. For the general case in which the $\Re\gamma_k$ are not all equal, it is also shown that there exists a smallest integer $\tau \geq 1$ such that the problem $\mathcal{IP}_n$, with $n = \tau + m\nu$, $\nu \in \{0, 1, 2, \ldots\}$, has a unique solution $(s'_1, \ldots, s'_m)$ of the form

$$s'_k = \widehat{s}_k + \nu, \quad k = 1, \ldots, m, \tag{3.3}$$

where $(\widehat{s}_1, \ldots, \widehat{s}_m)$ is the (unique) solution to $\mathcal{IP}_\tau$.

Thus, in any case, there are infinitely many values of n for which $\mathcal{IP}_n$ has a unique solution $(s'_1, \ldots, s'_m)$.[2]

Our first result concerns the stability of the sequence $\{A_n^{(j)}\}_{j=0}^{\infty}$. It says that the sequence $\{A_n^{(j)}\}_{j=0}^{\infty}$ can be computed stably since $\Gamma_n^{(j)}$ is bounded for all large j.

Theorem 3.1 *1. Let n be a fixed integer for which the integer programming problem $\mathcal{IP}_n$ has a unique solution $(s'_1, \ldots, s'_m)$. Then the polynomials $\sum_{i=0}^{n} \gamma_{ni}^{(j)} z^i$ exist for all large j, and satisfy*

$$\lim_{j\to\infty} \sum_{i=0}^{n} \gamma_{ni}^{(j)} z^i = \prod_{k=1}^{m} \left(\frac{z - \zeta_k}{1 - \zeta_k} \right)^{s'_k}. \tag{3.4}$$

Consequently,

$$\lim_{j\to\infty} \Gamma_n^{(j)} = \lim_{j\to\infty} \sum_{i=0}^{n} \left| \gamma_{ni}^{(j)} \right| \leq \prod_{k=1}^{m} \left(\frac{1 + |\zeta_k|}{|1 - \zeta_k|} \right)^{s'_k}. \tag{3.5}$$

2. In case $\Re\gamma_1 = \cdots = \Re\gamma_m$, for $n = m\nu$ with $\nu = 1, 2, \ldots,$ there hold

$$\lim_{j\to\infty} \sum_{i=0}^{n} \gamma_{ni}^{(j)} z^i = \prod_{k=1}^{m} \left(\frac{z - \zeta_k}{1 - \zeta_k} \right)^{\nu} \tag{3.6}$$

and

$$\lim_{j\to\infty} \Gamma_n^{(j)} = \lim_{j\to\infty} \sum_{i=0}^{n} \left| \gamma_{ni}^{(j)} \right| \leq \prod_{k=1}^{m} \left(\frac{1 + |\zeta_k|}{|1 - \zeta_k|} \right)^{\nu}. \tag{3.7}$$

[2] The fact that unique solutions for $\mathcal{IP}_n$ exist for infinitely many n is important. Without it, we would be dealing with the convergence of only a finite number of column sequences, and this diminishes the relevance of the present work.

Note that (3.4) implies that $\lim_{j\to\infty}\gamma_{ni}^{(j)}$, $i=0,1,\ldots,n$, all exist and are finite. In fact, they are the corresponding coefficients of the polynomial $\prod_{k=1}^{m}\left(\frac{z-\zeta_k}{1-\zeta_k}\right)^{s'_k}$. Consequently, $\lim_{j\to\infty}\Gamma_n^{(j)}$ exists and is finite. This is part of (3.5), which also gives a very simple upper bound on $\lim_{j\to\infty}\Gamma_n^{(j)}$ in terms the ζ_k and the s'_k.

The next theorem says that the Shanks transformation accelerates the convergence of the sequence $\{A_s\}$, in the sense that the sequence $\{A_n^{(j)}\}_{j=0}^{\infty}$, with fixed n, converges faster than $\{A_s\}$, providing at the same time the exact asymptotic behavior of the error in $A_n^{(j)}$ as $j\to\infty$.

Theorem 3.2 *Let θ be as defined in (1.4).*

1. *Let n be a fixed integer for which the integer programming problem $\mathcal{IP}_n$ in (3.2) has a unique solution $(s'_1,\ldots,s'_m)$. Then $A_n^{(j)}$ exist for all large j, and satisfy*

$$A_n^{(j)}-A=O\left(\theta^j\, j^{\sigma_{n+1}-\sigma_n}\right)\quad \text{as}\quad j\to\infty, \tag{3.8}$$

where σ_p is the value of the function $g(s_1,\ldots,s_m)$ in (3.2) at the (optimal) solution to $\mathcal{IP}_p$.

2. *If $n=\tau+m\nu$, where $\tau\geq 1$ and $\nu\in\{0,1,2,\ldots\}$, precisely as described in the first paragraph of this section, so that $\mathcal{IP}_n$ has a unique solution, then $A_n^{(j)}$ satisfies*

$$A_n^{(j)}-A=O\left(\theta^j\, j^{\omega-2\nu}\right)\quad \text{as } j\to\infty,\quad \omega=\Re\gamma_q-2\widehat{s}_q\leq\Re\gamma_1-2\widehat{s}_1+1. \tag{3.9}$$

Here, the index q is as determined in Lemma A.4, and $(\widehat{s}_1,\ldots,\widehat{s}_m)$ is the (unique) solution to $\mathcal{IP}_\tau$.

3. *In case $\Re\gamma_1=\cdots=\Re\gamma_m=\alpha$, for $n=m\nu$ with $\nu\in\{1,2,\ldots\}$, (3.9) simplifies to read*

$$A_n^{(j)}-A=O\left(\theta^j\, j^{\alpha-2\nu}\right)\quad \text{as } j\to\infty. \tag{3.10}$$

Remarks 1. As already mentioned, the uniqueness of the solution to $\mathcal{IP}_n$ is necessary for the results of Theorems 3.1 and 3.2 to be true. In case $\mathcal{IP}_n$ does not have a unique solution, we are not able to make definitive statements about either convergence or stability of the $A_n^{(j)}$.

2. The results of Theorems 3.1 and 3.2 are the best that can be obtained asymptotically under the given conditions. Given the complex structure of the sequence $\{A_n\}$ and the extreme nonlinearity of the Shanks transformation, the simplicity of these results is also quite surprising.
3. The result in (3.9) suggests that, when n is sufficiently large (so that ν is sufficiently large as well), the sequence $\{A_n^{(j)}\}_{j=0}^{\infty}$ converges to A faster than $\{A_j\}_{j=0}^{\infty}$. To make this statement more precise, let us compare (3.9) with

$$A_j-A=O(\theta^j j^{\gamma_1})\quad \text{as } j\to\infty, \tag{3.11}$$

which follows from (1.3) and (3.1). We have convergence acceleration if $\omega - 2\nu < \Re\gamma_1$, which is satisfied when $2\nu > 1 - 2\widehat{s}_1$, and hence when $n > \tau - m\widehat{s}_1 + m/2$. In particular, when $\Re\gamma_1 = \cdots = \Re\gamma_m$, for $n = m\nu$ with $\nu = 1, 2, \ldots,$ convergence acceleration takes place for every such n, by (3.10).

4. For $m = 1$, Theorem 3.2 reduces to Theorem 16.5.1 in [34, Chapter 16], which was proved originally in [12], while Theorem 3.1 reduces to Theorem 16.5.2 in [34, Chapter 16], which is new.

4 Technical preliminaries

The results in the following lemma follow from Sidi [28, Theorem 2.4]. See also [34, Sect. 3.2].

Lemma 4.1 *Define*

$$a_i^{(r)} = a_{i+r} \quad \text{with} \quad a_{i+1} = \Delta A_i = A_{i+1} - A_i. \tag{4.1}$$

Let $\{x_s\}$ *be an arbitrary sequence, and define the* $(n+1) \times (n+1)$ *determinant* $e_n^{(j)}(\{x_s\})$ *via*

$$e_n^{(j)}(\{x_s\}) = \begin{vmatrix} x_j & a_j^{(1)} & a_j^{(2)} & \cdots & a_j^{(n)} \\ x_{j+1} & a_{j+1}^{(1)} & a_{j+1}^{(2)} & \cdots & a_{j+1}^{(n)} \\ x_{j+2} & a_{j+2}^{(1)} & a_{j+2}^{(2)} & \cdots & a_{j+2}^{(n)} \\ \vdots & \vdots & \vdots & & \vdots \\ x_{j+n} & a_{j+n}^{(1)} & a_{j+n}^{(2)} & \cdots & a_{j+n}^{(n)} \end{vmatrix}. \tag{4.2}$$

Then, with $I_s = 1$ *for all* s, *we have*

$$A_n^{(j)} = \frac{e_n^{(j)}(\{A_s\})}{e_n^{(j)}(\{I_s\})} \tag{4.3}$$

and

$$\sum_{i=0}^{n} \gamma_{ni}^{(j)} z^i = \frac{e_n^{(j)}(\{z^s\})}{e_n^{(j)}(\{I_s\})} z^{-j}. \tag{4.4}$$

Note that $e_n^{(j)}(\{A_s\})$ and $e_n^{(j)}(\{I_s\})$ in (4.3) are, respectively, the numerator and denominator determinants in (1.2). Thus, (4.3) is simply (1.2). The validity of the result in (4.4), and also of those in (1.5) and (1.6), can be shown by expanding the determinant $e_n^{(j)}(\{A_s\})$ with respect to its first column and identifying $\gamma_{ni}^{(j)}$ as the cofactor of A_{j+i} divided by $e_n^{(j)}(\{I_s\})$.

The next lemma relates the asymptotic expansions of A_n, a_n, and the $a_n^{(r)}$, $r = 1, 2, \dots,$ defined in (4.1).

Lemma 4.2 *Let $a_n = A_n - A_{n-1}$ and $a_n^{(r)} = a_{n+r}$. Let also $\zeta_1, \dots, \zeta_m$ be distinct scalars different from 1, and let $\gamma_1, \dots, \gamma_k$ be arbitrary scalars. Then the following asymptotic expansions are valid simultaneously:*

$$A_n \sim A + \sum_{k=1}^{m} \zeta_k^n \sum_{i=0}^{\infty} \beta_{ki} n^{\gamma_k - i} \quad as\ n \to \infty; \quad \beta_{k0} \neq 0, \quad k = 1, \dots, m, \tag{4.5}$$

$$a_n \sim \sum_{k=1}^{m} \zeta_k^n \sum_{i=0}^{\infty} \epsilon_{ki} n^{\gamma_k - i} \quad as\ n \to \infty; \quad \epsilon_{k0} \neq 0, \quad k = 1, \dots, m, \tag{4.6}$$

$$a_n^{(r)} \sim \sum_{k=1}^{m} \zeta_k^n \sum_{i=0}^{\infty} \epsilon_{ki}^{(r)} n^{\gamma_k - i} \quad as\ n \to \infty; \quad \epsilon_{k0}^{(r)} \neq 0, \quad k = 1, \dots, m. \tag{4.7}$$

The β_{ki}, ϵ_{ki}, and $\epsilon_{ki}^{(r)}$ are related to each other via

$$\epsilon_{ki} = \sum_{s=0}^{i} \beta_{ks} c_{ks,i-s}, \quad \epsilon_{ki}^{(r)} = \zeta_k^r \sum_{s=0}^{i} \epsilon_{ks} d_{ks,i-s}^{(r)}, \quad i = 0, 1, \dots, \tag{4.8}$$

where

$$c_{ks0} = 1 - \zeta_k^{-1}; \quad c_{ksp} = (-1)^{p-1} \binom{\gamma_k - s}{p} \zeta_k^{-1}, \quad p = 1, 2, \dots, \tag{4.9}$$

and

$$d_{ksp}^{(r)} = \binom{\gamma_k - s}{p} r^p, \quad p = 0, 1, \dots. \tag{4.10}$$

Therefore,

$$\epsilon_{k0} = \beta_{k0}(1 - \zeta_k^{-1}), \quad \epsilon_{k0}^{(r)} = \zeta_k^r \epsilon_{k0}. \tag{4.11}$$

We next order the index pairs $(k, i) \equiv ki$ in the summations of (4.5)–(4.7). Recalling that $1 \leq k \leq m$ and $i \geq 0$, we achieve this as follows:

$$10, 20, \dots, m0;\ 11, 21, \dots, m1;\ 12, 22, \dots, m2;\ \dots. \tag{4.12}$$

With this ordering, we can write

$$ki < k'i' \quad \text{if either } (1) i < i' \text{or } (2) i = i' \text{ and } k < k'. \tag{4.13}$$

In our proofs, we will also be using the next four lemmas:

Lemma 4.3 *Let $i_0, i_1, \ldots, i_k$ be positive integers, and assume that the scalars $v_{i_0,i_1,\ldots,i_k}$ are odd under an interchange of any two of the indices $i_0, i_1, \ldots, i_k$. Let $t_{i,j}$, $i, j \geq 1$, be scalars and let σ_i, $i \geq 1$ be all scalars or vectors. Define*

$$I_{k,N} = \sum_{i_0=1}^{N} \sum_{i_1=1}^{N} \cdots \sum_{i_k=1}^{N} \sigma_{i_0} \left(\prod_{p=1}^{k} t_{i_p,p} \right) v_{i_0,i_1,\ldots,i_k}$$

and

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

Then

$$I_{k,N} = J_{k,N}.$$

This lemma was first stated and proved in Sidi et al. [40, Lemma A.1]; see also [34, p. 303, Lemma 16.4.1]. It has been used by the author in the treatment of different problems in extrapolation methods, including the Shanks transformation and the Padé table.

Lemma 4.4 *Let the $(p+1) \times (p+1)$ matrix H be given as in*

$$H = [\, H_1 \,|\, H_2 \,|\, \cdots \,|\, H_t \,], \tag{4.14}$$

where

$$H_i = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ c_i & c_i 1^1 & c_i 1^2 & \cdots & c_i 1^{s_i-1} \\ c_i^2 & c_i^2 2^1 & c_i^2 2^2 & \cdots & c_i^2 2^{s_i-1} \\ \vdots & \vdots & \vdots & & \vdots \\ c_i^p & c_i^p p^1 & c_i^p p^2 & \cdots & c_i^p p^{s_i-1} \end{bmatrix}, \quad i = 1, \ldots, t, \tag{4.15}$$

and $\sum_{i=1}^{t} s_i = p + 1$. Then

$$\det H = \left[\prod_{i=1}^{t} \left(\prod_{k=0}^{s_i-1} k! \right) c_i^{s_i(s_i-1)/2} \right] \left[\prod_{1 \leq i < k \leq t} (c_k - c_i)^{s_i s_k} \right] \equiv W(c_1, s_1; c_2, s_2; \ldots; c_t, s_t). \tag{4.16}$$

Consequently, with $1+\sum_{i=1}^{t} s_i = p+1$, *for the* $(p+1)\times(p+1)$ *matrix*

$$\widetilde{H} = [\,Z\,|\,H_1\,|H_2\,|\,\cdots\,|H_t\,],\quad Z=[1,z,z^2,\ldots,z^p]^{\mathrm{T}}, \tag{4.17}$$

we have

$$\begin{aligned}\det\widetilde{H} &= W(z,1;c_1,s_1;c_2,s_2;\ldots;c_t,s_t)\\ &= W(c_1,s_1;c_2,s_2;\ldots;c_t,s_t)\left[\prod_{i=1}^{t}(c_i-z)^{s_i}\right].\end{aligned} \tag{4.18}$$

Note that $W(c_1,s_1;c_2,s_2;\ldots;c_t,s_t)$ is actually a multiple of a confluent Vandermonde determinant. For a proof of Lemma 4.4, see Sidi [32, Eq. (3.15) and Appendix].

The following lemma follows from Lemma 4.4 and is new.

Lemma 4.5 *Let the* $p\times p$ *matrix* $\overline{H}$ *be given as in*

$$\overline{H} = \left[\,\overline{H}_1\,|\,\overline{H}_2\,|\,\cdots\,|\,\overline{H}_t\,\right], \tag{4.19}$$

where

$$\overline{H}_i = \begin{bmatrix} c_i & c_i 1^1 & c_i 1^2 & \cdots & c_i 1^{s_i-1}\\ c_i^2 & c_i^2 2^1 & c_i^2 2^2 & \cdots & c_i^2 2^{s_i-1}\\ \vdots & \vdots & \vdots & & \vdots\\ c_i^p & c_i^p p^1 & c_i^p p^2 & \cdots & c_i^p p^{s_i-1}\end{bmatrix},\quad i=1,\ldots,t, \tag{4.20}$$

and $\sum_{i=1}^{t} s_i = p$. *Then*

$$\det\overline{H} = W(c_1,s_1;c_2,s_2;\ldots;c_t,s_t)\left[\prod_{i=1}^{t} c_i^{s_i}\right]. \tag{4.21}$$

Proof Below, we refer to the column $[c_i 1^k, c_i^2 2^k,\ldots,c_i^p p^k]^{\mathrm{T}}$ as "column k" of $\overline{H}_i$. Let us now add $\left[\sum_{\mu=0}^{k-1}(-1)^{k-\mu}\binom{k}{\mu}\times(\text{column }\mu\text{ of }\overline{H}_i)\right]$ to column k and overwrite column k, in the order $k=s_i-1, s_i-2,\ldots,1$, and invoke the binomial theorem $(r-1)^k=\sum_{\mu=0}^{k}(-1)^{k-\mu}\binom{k}{\mu}r^{\mu}$. As a result, $\overline{H}_i$ becomes

$$\check{H}_i = \begin{bmatrix} c_i & 0 & 0 & \cdots & 0\\ c_i^2 & c_i^2 1^1 & c_i^2 1^2 & \cdots & c_i^2 1^{s_i-1}\\ c_i^3 & c_i^3 2^1 & c_i^3 2^2 & \cdots & c_i^3 2^{s_i-1}\\ \vdots & \vdots & \vdots & & \vdots\\ c_i^p & c_i^p(p-1)^1 & c_i^p(p-1)^2 & \cdots & c_i^p(p-1)^{s_i-1}\end{bmatrix},$$

and we have $\det \overline{H} = \det[\check{H}_1|\check{H}_2|\cdots|\check{H}_t]$. Upon factoring out c_i from each of its columns, $\check{H}_i$ becomes H_i in (4.15), with p there replaced by $p-1$. The proof can now be completed by invoking Lemma 4.4. □

Lemma 4.6 *Let $Q_i(x) = \sum_{j=0}^{i} a_{ij}x^j$, with $a_{ii} \neq 0$, $i = 0, 1, \ldots, n$, and let x_i, $i = 0, 1, \ldots, n$, be arbitrary points. Then*

$$\begin{vmatrix} Q_0(x_0) & Q_0(x_1) & \cdots & Q_0(x_n) \\ Q_1(x_0) & Q_1(x_1) & \cdots & Q_1(x_n) \\ \vdots & \vdots & & \vdots \\ Q_n(x_0) & Q_n(x_1) & \cdots & Q_n(x_n) \end{vmatrix} = \left(\prod_{i=0}^{n} a_{ii}\right) V(x_0, x_1, \ldots, x_n), \tag{4.22}$$

where $V(x_0, x_1, \ldots, x_n) = \prod_{0 \le i < j \le n}(x_j - x_i)$ is a Vandermonde determinant.

For a proof of this lemma, see [33, Lemma 1.2]. See also [34, p. 153, Lemma 6.8.1].

5 Proof of Theorem 3.1

5.1 Asymptotic expansion for $z^{-j}e_n^{(j)}(\{z^s\})$

We start with the analysis of the determinant

$$z^{-j}e_n^{(j)}(\{z^s\}) = \begin{vmatrix} z^0 & a_j^{(1)} & a_j^{(2)} & \cdots & a_j^{(n)} \\ z^1 & a_{j+1}^{(1)} & a_{j+1}^{(2)} & \cdots & a_{j+1}^{(n)} \\ z^2 & a_{j+2}^{(1)} & a_{j+2}^{(2)} & \cdots & a_{j+2}^{(n)} \\ \vdots & \vdots & \vdots & & \vdots \\ z^n & a_{j+n}^{(1)} & a_{j+n}^{(2)} & \cdots & a_{j+n}^{(n)} \end{vmatrix}. \tag{5.1}$$

At this point, we would like to state that it is important to understand the details of the technique that we develop here, because this technique is used again in the proof of Theorem 3.2. In this technique, we perform only elementary column transformations on $z^{-j}e_n^{(j)}(\{z^s\})$.

Throughout, $\sum_{ki}$ stands for $\sum_{k=1}^{m}\sum_{i=0}^{\infty}$ and "$u_j \sim v_j$" stands for "$u_j \sim v_j$ as $j \to \infty$." In the sequel, we also adopt the short-hand notation

$$\mathbf{ki}(r:s) = k_r i_r, k_{r+1} i_{r+1}, \ldots, k_s i_s. \tag{5.2}$$

Let us replace the $a_{j+s}^{(r)}$ in (5.1) by their asymptotic expansions in Lemma 4.2. This results in

$$z^{-j}e_n^{(j)}(\{z^s\}) \sim \begin{vmatrix} 1 & \sum\limits_{k_1i_1}\epsilon^{(1)}_{k_1i_1}\zeta^{j}_{k_1}j^{\gamma_{k_1}-i_1} & \sum\limits_{k_2i_2}\epsilon^{(2)}_{k_2i_2}\zeta^{j}_{k_2}j^{\gamma_{k_2}-i_2} & \cdots & \sum\limits_{k_ni_n}\epsilon^{(n)}_{k_ni_n}\zeta^{j}_{k_n}j^{\gamma_{k_n}-i_n} \\ z & \sum\limits_{k_1i_1}\epsilon^{(1)}_{k_1i_1}\zeta^{j+1}_{k_1}(j+1)^{\gamma_{k_1}-i_1} & \sum\limits_{k_2i_2}\epsilon^{(2)}_{k_2i_2}\zeta^{j+1}_{k_2}(j+1)^{\gamma_{k_2}-i_2} & \cdots & \sum\limits_{k_ni_n}\epsilon^{(n)}_{k_ni_n}\zeta^{j+1}_{k_n}(j+1)^{\gamma_{k_n}-i_n} \\ z^2 & \sum\limits_{k_1i_1}\epsilon^{(1)}_{k_1i_1}\zeta^{j+2}_{k_1}(j+2)^{\gamma_{k_1}-i_1} & \sum\limits_{k_2i_2}\epsilon^{(2)}_{k_2i_2}\zeta^{j+2}_{k_2}(j+2)^{\gamma_{k_2}-i_2} & \cdots & \sum\limits_{k_ni_n}\epsilon^{(n)}_{k_ni_n}\zeta^{j+2}_{k_n}(j+2)^{\gamma_{k_n}-i_n} \\ \vdots & \vdots & \vdots & & \vdots \\ z^n & \sum\limits_{k_1i_1}\epsilon^{(1)}_{k_1i_1}\zeta^{j+n}_{k_1}(j+n)^{\gamma_{k_1}-i_1} & \sum\limits_{k_2i_2}\epsilon^{(2)}_{k_2i_2}\zeta^{j+n}_{k_2}(j+n)^{\gamma_{k_2}-i_2} & \cdots & \sum\limits_{k_ni_n}\epsilon^{(n)}_{k_ni_n}\zeta^{j+n}_{k_n}(j+n)^{\gamma_{k_n}-i_n} \end{vmatrix}. \tag{5.3}$$

Recalling that determinants are multilinear in their rows and columns, we first take out the summation over $k_1i_1, \dots, k_ni_n$. Following that, we take out the common factors from each column of the remaining determinant. This results in

$$z^{-j}e_n^{(j)}(\{z^s\}) \sim \sum_{k_1i_1}\sum_{k_2i_2}\cdots\sum_{k_ni_n}\left[\prod_{s=1}^{n}\epsilon^{(s)}_{k_si_s}\right]\left[\prod_{s=1}^{n}\zeta_{k_s}\right]^j M^{(j)}_{\mathbf{ki}(1:n)}(z), \tag{5.4}$$

where

$$M^{(j)}_{\mathbf{ki}(1:n)}(z) = \begin{vmatrix} 1 & \zeta^0_{k_1}j^{\gamma_{k_1}-i_1} & \zeta^0_{k_2}j^{\gamma_{k_2}-i_2} & \cdots & \zeta^0_{k_n}j^{\gamma_{k_n}-i_n} \\ z & \zeta^1_{k_1}(j+1)^{\gamma_{k_1}-i_1} & \zeta^1_{k_2}(j+1)^{\gamma_{k_2}-i_2} & \cdots & \zeta^1_{k_n}(j+1)^{\gamma_{k_n}-i_n} \\ z^2 & \zeta^2_{k_1}(j+2)^{\gamma_{k_1}-i_1} & \zeta^2_{k_2}(j+2)^{\gamma_{k_2}-i_2} & \cdots & \zeta^2_{k_n}(j+2)^{\gamma_{k_n}-i_n} \\ \vdots & \vdots & \vdots & & \vdots \\ z^n & \zeta^n_{k_1}(j+n)^{\gamma_{k_1}-i_1} & \zeta^n_{k_2}(j+n)^{\gamma_{k_2}-i_2} & \cdots & \zeta^n_{k_n}(j+n)^{\gamma_{k_n}-i_n} \end{vmatrix}. \tag{5.5}$$

Note that, being a sum of products of a finite number of Poincaré-type asymptotic expansions, the right-hand side of (5.4) is a genuine Poincaré-type asymptotic expansion once its terms are ordered according to their sizes.

Now, the product $M^{(j)}_{\mathbf{ki}(1:n)}(z)\left[\prod_{s=1}^{n}\zeta_{k_s}\right]^j$ is odd under an interchange of any two of the index pairs $k_1i_1, \dots, k_ni_n$, since this amounts to interchanging two columns in the determinant representation of $M^{(j)}_{\mathbf{ki}(1:n)}(z)$, while the term $\left[\prod_{s=1}^{n}\zeta_{k_s}\right]^j$ does not change. Consequently, Lemma 4.3 applies, and we obtain the asymptotic expansion

$$z^{-j}e_n^{(j)}(\{z^s\}) \sim \sum_{k_1i_1<k_2i_2<\cdots<k_ni_n} E_{\mathbf{ki}(1:n)}M^{(j)}_{\mathbf{ki}(1:n)}(z)\left[\prod_{s=1}^{n}\zeta_{k_s}\right]^j, \tag{5.6}$$

where

$$E_{\mathbf{ki}(1:n)} = \begin{vmatrix} \epsilon^{(1)}_{k_1 i_1} & \epsilon^{(1)}_{k_2 i_2} & \cdots & \epsilon^{(1)}_{k_n i_n} \\ \epsilon^{(2)}_{k_1 i_1} & \epsilon^{(2)}_{k_2 i_2} & \cdots & \epsilon^{(2)}_{k_n i_n} \\ \vdots & \vdots & & \vdots \\ \epsilon^{(n)}_{k_1 i_1} & \epsilon^{(n)}_{k_2 i_2} & \cdots & \epsilon^{(n)}_{k_n i_n} \end{vmatrix}. \tag{5.7}$$

Note that the dependence on j in (5.6) comes into play only through the products $M^{(j)}_{\mathbf{ki}(1:n)}(z)\left[\prod_{s=1}^{n}\zeta_{k_s}\right]^j$. Now, by (1.4), the terms $\left[\prod_{s=1}^{n}\zeta_{k_s}\right]^j$ all have the same modulus, namely, θ^{nj}. Consequently, the dominant terms in the asymptotic expansion of (5.6) are those with the largest $|M^{(j)}_{\mathbf{ki}(1:n)}(z)|$ as $j \to \infty$. Therefore, we need to analyze the asymptotic behavior of the determinants $M^{(j)}_{\mathbf{ki}(1:n)}(z)$ carefully.

5.2 Asymptotic expansion of $M^{(j)}_{\mathbf{ki}(1:n)}(z)$

First, we note that the n index pairs $k_1 i_1, \dots, k_n i_n$ consist of s_1 pairs of the form $1 i_{1r}$, s_2 pairs of the form $2i_{2r}$,..., s_m pairs of the form $m i_{mr}$. Here $0 \le i_{t1} < i_{t2} < \cdots < i_{ts_t}$ when $s_t \neq 0$, and $\sum_{t=1}^{m} s_t = n$. Of course, in case $s_t = 0$, there are no pairs of the form $t i_{tr}$ among $k_1 i_1, \dots, k_n i_n$. For simplicity of notation, assume that $s_1, \dots, s_p$ are nonzero ($p \le m$) and that the rest of the s_t are zero. Thus, after a permutation of columns, the determinant $M^{(j)}_{\mathbf{ki}(1:n)}(z)$ becomes

$$\begin{aligned} &M^{(j)}_{\mathbf{ki}(1:n)}(z) \\ &\quad = \pm \det\left[Z_n \mid D(\zeta_1, s_1, \{i_{1r}\}) \mid D(\zeta_2, s_2, \{i_{2r}\}) \mid \cdots \mid D(\zeta_p, s_p, \{i_{pr}\})\right], \end{aligned} \tag{5.8}$$

where

$$Z_n = [z^0, z^1, \dots, z^n]^{\mathrm{T}}, \tag{5.9}$$

and $D(\zeta_t, s_t, \{i_{tr}\})$ is an $(n+1) \times s_t$ matrix given as in

$$D(\zeta_t, s_t, \{i_{tr}\}) = \begin{bmatrix} \zeta_t^0 j^{\gamma_t - i_{t1}} & \zeta_t^0 j^{\gamma_t - i_{t2}} & \cdots & \zeta_t^0 j^{\gamma_t - i_{ts_t}} \\ \zeta_t^1 (j+1)^{\gamma_t - i_{t1}} & \zeta_t^1 (j+1)^{\gamma_t - i_{t2}} & \cdots & \zeta_t^1 (j+1)^{\gamma_t - i_{ts_t}} \\ \zeta_t^2 (j+2)^{\gamma_t - i_{t1}} & \zeta_t^2 (j+2)^{\gamma_t - i_{t2}} & \cdots & \zeta_t^2 (j+2)^{\gamma_t - i_{ts_t}} \\ \vdots & \vdots & & \vdots \\ \zeta_t^n (j+n)^{\gamma_t - i_{t1}} & \zeta_t^n (j+n)^{\gamma_t - i_{t2}} & \cdots & \zeta_t^n (j+n)^{\gamma_t - i_{ts_t}} \end{bmatrix}. \tag{5.10}$$

Now, to determine the asymptotic behavior of $M^{(j)}_{\mathbf{ki}(1:n)}(z)$ in (5.8), it is sufficient to look at the contribution of each of the blocks $D(\zeta_t, s_t, \{i_{tr}\})$, $t = 1, \ldots, p$. We will see that each one of these blocks contributes independently of others.

An intermediate asymptotic result

For this, we start by analyzing the contribution of the $(n+1) \times s$ matrix

$$G = \begin{bmatrix} \zeta^0 j^{\delta_1} & \zeta^0 j^{\delta_2} & \cdots & \zeta^0 j^{\delta_s} \\ \zeta^1 (j+1)^{\delta_1} & \zeta^1 (j+1)^{\delta_2} & \cdots & \zeta^1 (j+1)^{\delta_s} \\ \zeta^2 (j+2)^{\delta_1} & \zeta^2 (j+2)^{\delta_2} & \cdots & \zeta^2 (j+2)^{\delta_s} \\ \vdots & \vdots & & \vdots \\ \zeta^n (j+n)^{\delta_1} & \zeta^n (j+n)^{\delta_2} & \cdots & \zeta^n (j+n)^{\delta_s} \end{bmatrix} \tag{5.11}$$

to the asymptotic behavior, as $j \to \infty$, of a determinant U that has the form

$$U = \det\,[G \,|\, W], \quad W \in \mathbb{C}^{(n+1)\times(n+1-s)}. \tag{5.12}$$

Factoring out j^{δ_q} from the qth column of G, $q = 1, \ldots, s$, we first have

$$U = \left[\prod_{q=1}^{s} j^{\delta_q}\right] \det\left[G' \,|\, W\right], \tag{5.13}$$

where

$$G' = \begin{bmatrix} \zeta^0 (1+0/j)^{\delta_1} & \zeta^0 (1+0/j)^{\delta_2} & \cdots & \zeta^0 (1+0/j)^{\delta_s} \\ \zeta^1 (1+1/j)^{\delta_1} & \zeta^1 (1+1/j)^{\delta_2} & \cdots & \zeta^1 (1+1/j)^{\delta_s} \\ \zeta^2 (1+2/j)^{\delta_1} & \zeta^2 (1+2/j)^{\delta_2} & \cdots & \zeta^2 (1+2/j)^{\delta_s} \\ \vdots & \vdots & & \vdots \\ \zeta^n (1+n/j)^{\delta_1} & \zeta^n (1+n/j)^{\delta_2} & \cdots & \zeta^n (1+n/j)^{\delta_s} \end{bmatrix}. \tag{5.14}$$

Expanding each of the terms $(1+r/j)^{\delta_q}$ in negative powers of j, we obtain

$$G' = \begin{bmatrix} \zeta^0 \sum_{i_1} \binom{\delta_1}{i_1} (0/j)^{i_1} & \zeta^0 \sum_{i_2} \binom{\delta_2}{i_2} (0/j)^{i_2} & \cdots & \zeta^0 \sum_{i_s} \binom{\delta_s}{i_s} (0/j)^{i_s} \\ \zeta^1 \sum_{i_1} \binom{\delta_1}{i_1} (1/j)^{i_1} & \zeta^1 \sum_{i_2} \binom{\delta_2}{i_2} (1/j)^{i_2} & \cdots & \zeta^1 \sum_{i_s} \binom{\delta_s}{i_s} (1/j)^{i_s} \\ \zeta^2 \sum_{i_1} \binom{\delta_1}{i_1} (2/j)^{i_1} & \zeta^2 \sum_{i_2} \binom{\delta_2}{i_2} (2/j)^{i_2} & \cdots & \zeta^2 \sum_{i_s} \binom{\delta_s}{i_s} (2/j)^{i_s} \\ \vdots & \vdots & & \vdots \\ \zeta^n \sum_{i_1} \binom{\delta_1}{i_1} (n/j)^{i_1} & \zeta^n \sum_{i_2} \binom{\delta_2}{i_2} (n/j)^{i_2} & \cdots & \zeta^n \sum_{i_s} \binom{\delta_s}{i_s} (n/j)^{i_s} \end{bmatrix}. \tag{5.15}$$

Here $\sum_i$ stands for $\sum_{i=0}^{\infty}$, and $0^0 = 1$ while $0^i = 0$ for $i \geq 1$. Let us substitute this in (5.13), and take out the summations and the common factors from each of the first s columns. We obtain the convergent expansion

$$U = \left[\prod_{q=1}^{s} j^{\delta_q}\right] \sum_{i_1}\sum_{i_2}\cdots\sum_{i_s}\left[\prod_{q=1}^{s}\binom{\delta_q}{i_q}\right] \times \left[\prod_{q=1}^{s} j^{-i_q}\right] \det\left[H_{i_1,\ldots,i_s}(\zeta)\,|\,W\right], \tag{5.16}$$

where

$$H_{i_1,\ldots,i_s}(\zeta) = \begin{bmatrix} \zeta^0 0^{i_1} & \zeta^0 0^{i_2} & \cdots & \zeta^0 0^{i_s} \\ \zeta^1 1^{i_1} & \zeta^1 1^{i_2} & \cdots & \zeta^1 1^{i_s} \\ \zeta^2 2^{i_1} & \zeta^2 2^{i_2} & \cdots & \zeta^2 2^{i_s} \\ \vdots & \vdots & & \vdots \\ \zeta^n n^{i_1} & \zeta^n n^{i_2} & \cdots & \zeta^n n^{i_s} \end{bmatrix}. \tag{5.17}$$

Since $\left[\prod_{q=1}^{s} j^{-i_q}\right]\det\left[H_{i_1,\ldots,i_s}(\zeta)\,|\,W\right]$ is odd under an interchange of any two of the indices $i_1, \ldots, i_s$, Lemma 4.3 applies to the multiple summation in (5.16), and we have

$$U = \left[\prod_{q=1}^{s} j^{\delta_q}\right] \sum_{i_1<i_2<\cdots<i_s} K_{i_1,\ldots,i_s}\left[\prod_{q=1}^{s} j^{-i_q}\right] \det\left[H_{i_1,\ldots,i_s}(\zeta)\,|\,W\right], \tag{5.18}$$

where

$$K_{i_1,\ldots,i_s} = \begin{vmatrix} \binom{\delta_1}{i_1} & \binom{\delta_1}{i_2} & \cdots & \binom{\delta_1}{i_s} \\ \binom{\delta_2}{i_1} & \binom{\delta_2}{i_2} & \cdots & \binom{\delta_2}{i_s} \\ \vdots & \vdots & & \vdots \\ \binom{\delta_s}{i_1} & \binom{\delta_s}{i_2} & \cdots & \binom{\delta_s}{i_s} \end{vmatrix}. \tag{5.19}$$

It is clear that, *provided it does not vanish*, the dominant term as $j \to \infty$ in the summation of (5.18) is that with the indices $i_1 = 0, i_2 = 1, \ldots, i_s = s-1$, and it is of order $j^{-s(s-1)/2}$. (The rest of the terms are subdominant.) Now, $H_{0,1,\ldots,s-1}(\zeta)$ has full rank on account of Lemma 4.4, which means that $\det\left[H_{0,1,\ldots,s-1}(\zeta)\,|\,W\right]$ cannot vanish on account of $H_{0,1,\ldots,s-1}(\zeta)$. We should then check whether $K_{0,1,\ldots,s-1}$

vanishes or not. By the fact that $\binom{\delta}{i} = \frac{1}{i!}\prod_{k=1}^{i}(\delta - k + 1)$ is a polynomial of degree i in δ with leading coefficient $1/i!$, Lemma 4.6 applies to $K_{0,1,\dots,s-1}$, and we have

$$K_{0,1,\dots,s-1} = \left[\prod_{k=0}^{s-1} k!\right]^{-1} V(\delta_1, \delta_2, \dots, \delta_s), \tag{5.20}$$

which is nonzero since the δ_k are distinct. Thus,

$$U \sim \frac{V(\delta_1, \delta_2, \dots, \delta_s)}{\prod_{k=0}^{s-1} k!} \det\left[\widehat{H}(\zeta, s) \,|\, W\right] j^{\omega}, \quad \omega = \sum_{q=1}^{s} \delta_q - s(s-1)/2, \tag{5.21}$$

where we have defined

$$\widehat{H}(\zeta, s) = H_{0,1,\dots,s-1}(\zeta) = \begin{bmatrix} \zeta^0 0^0 & \zeta^0 0^1 & \cdots & \zeta^0 0^{s-1} \\ \zeta^1 1^0 & \zeta^1 1^1 & \cdots & \zeta^1 1^{s-1} \\ \zeta^2 2^0 & \zeta^2 2^1 & \cdots & \zeta^2 2^{s-1} \\ \vdots & \vdots & & \vdots \\ \zeta^n n^0 & \zeta^n n^1 & \cdots & \zeta^n n^{s-1} \end{bmatrix}, \tag{5.22}$$

provided $\det\left[\widehat{H}(\zeta, s) \,|\, W\right] \neq 0$.

An asymptotic equality for $M^{(j)}_{\mathbf{ki}(1:n)}(z)$

Going back to (5.10), we realize that the matrix $D(\zeta_t, s_t, \{i_{tq}\})$ has the same structure as the matrix G in (5.11), with $\zeta = \zeta_t$, $s = s_t$, $\delta_q = \gamma_t - i_{tq}$. Therefore, (5.8) gives

$$M^{(j)}_{\mathbf{ki}(1:n)}(z) \sim \pm C\, Q(z) j^{\phi}, \tag{5.23}$$

where

$$C = \prod_{t=1}^{p} \frac{V(\gamma_t - i_{t1}, \gamma_t - i_{t2}, \dots, \gamma_t - i_{ts_t})}{\prod_{k=0}^{s_t-1} k!} \neq 0, \tag{5.24}$$

$$Q(z) = \det\left[Z_n \,|\, \widehat{H}(\zeta_1, s_1) \,|\, \widehat{H}(\zeta_2, s_2) \,|\, \cdots \,|\, \widehat{H}(\zeta_p, s_p)\right], \tag{5.25}$$

and

$$\phi = \sum_{t=1}^{p}\left[\sum_{r=1}^{s_t}(\gamma_t - i_{tr}) - s_t(s_t - 1)/2\right]. \tag{5.26}$$

We now observe that, by (5.22) and (5.25), Lemma 4.4 applies to $Q(z)$, and gives

$$Q(z) = L\prod_{t=1}^{p}(\zeta_t - z)^{s_t},$$
$$L = \left[\prod_{t=1}^{p}\left(\prod_{k=0}^{s_t-1} k!\right)\zeta_t^{s_t(s_t-1)/2}\right]\left[\prod_{1\le k<t\le p}(\zeta_t-\zeta_k)^{s_k s_t}\right] \neq 0. \tag{5.27}$$

Substituting (5.27) in (5.23), we finally have

$$M^{(j)}_{\mathbf{ki}(1:n)}(z) \sim \pm N_{\mathbf{ki}(1:n)}\left[\prod_{t=1}^{p}(z-\zeta_t)^{s_t}\right] j^{\phi}, \quad N_{\mathbf{ki}(1:n)} = (-1)^n CL \neq 0, \tag{5.28}$$

the constant $N_{\mathbf{ki}(1:n)}$ being independent of z and j, provided, of course, that $z \neq \zeta_t$, $t = 1, \dots, m$.

5.3 The dominant $M^{(j)}_{\mathbf{ki}(1:n)}(z)$

With the precise asymptotic behavior of $M^{(j)}_{\mathbf{ki}(1:n)}(z)$ available, we now return to (5.6), and look for the dominant term in the asymptotic expansion there. As already mentioned, since the products $\left[\prod_{s=1}^{n}\zeta_{k_s}\right]^j$ all have the same modulus by (1.4), the dominant term is that for which $|M^{(j)}_{\mathbf{ki}(1:n)}(z)|$ is largest. Hence, by (5.28), we need to look for that term in the summation of (5.6) for which ϕ in (5.26) and (5.28) has largest real part.

At this point, we note that, by replacing the summation $\sum_{t=1}^{p}$ by the summation $\sum_{t=1}^{m}$, the expression for ϕ given in (5.26) can be rewritten as

$$\phi \equiv \phi(\boldsymbol{s},\boldsymbol{i}) = \sum_{t=1}^{m}\left[\sum_{r=1}^{s_t}(\gamma_t - i_{tr}) - s_t(s_t-1)/2\right], \quad \text{with} \quad \sum_{r=1}^{m} s_t = n, \tag{5.29}$$

to accommodate the zero s_t, as well as the nonzero ones. Clearly, if $s_t = 0$ for some t, then there is no contribution to $\phi(\boldsymbol{s},\boldsymbol{i})$ either from this s_t or from its associated i_{tr}. Here, we have adopted the notation

$$\boldsymbol{s} = (s_1,\dots,s_m) \quad \text{and} \quad \boldsymbol{i} = (\{i_{1r}\}_{r=1}^{s_1},\dots,\{i_{mr}\}_{r=1}^{s_m}).$$

Going back, we want to find the nonnegative integers $s_1,\dots,s_m$, and $i_{t1} < \dots < i_{ts_t}$, $t = 1,\dots,m$, that maximize $\Re\phi(\boldsymbol{s},\boldsymbol{i})$. We first realize that (5.29) can be rewritten also in the form

$$\phi(\boldsymbol{s},\boldsymbol{i}) = \sum_{t=1}^{m} s_t\gamma_t - \sum_{t=1}^{m}\left[\sum_{r=1}^{s_t} i_{tr} + s_t(s_t-1)/2\right]. \tag{5.30}$$

Therefore,

$$\Re\phi(\boldsymbol{s},\boldsymbol{i}) = \sum_{t=1}^{m} s_t(\Re\gamma_t) - \sum_{t=1}^{m}\left[\sum_{r=1}^{s_t} i_{tr} + s_t(s_t-1)/2\right], \tag{5.31}$$

and we need to maximize $\Re\phi(\boldsymbol{s},\boldsymbol{i})$ subject to the constraints that

$$s_t \geq 0, \quad 0 \leq i_{t1} < \cdots < i_{ts_t}, \quad t = 1,\ldots,m, \quad \text{and} \quad \sum_{t=1}^{m} s_t = n. \tag{5.32}$$

It is easy to see that, at the maximum of $\Re\phi(\boldsymbol{s},\boldsymbol{i})$, the i_{tr} must assume their smallest possible values consistent with the constraints in (5.32). Clearly, these are $i_{tr} = r-1, r = 1,\ldots,s_t$, so that $\sum_{r=1}^{s_t} i_{tr} = s_t(s_t-1)/2, t = 1,\ldots,m$. For these optimal values of the i_{tr}, we thus have $\max_{\boldsymbol{i}} \Re\phi(\boldsymbol{s},\boldsymbol{i}) = g(\boldsymbol{s}) = g(s_1,\ldots,s_m)$, where $g(s_1,\ldots,s_m)$ is as defined in (3.2). Consequently, the optimal $\boldsymbol{s}$, which we now denote by $\boldsymbol{s'} = (s'_1,\ldots,s'_m)$ must be the solution to the problem $\mathcal{IP}_n$ in (3.2), and we also have $\max_{\boldsymbol{s},\boldsymbol{i}} \Re\phi(\boldsymbol{s},\boldsymbol{i}) = \max_{\boldsymbol{s}} g(\boldsymbol{s}) = \sigma_n$ at the solution to $\mathcal{IP}_n$.

Summarizing, we have shown that the dominant $M^{(j)}_{\mathbf{ki}(1:n)}(z)$ is that for which $(k_1i_1,\ldots,k_ni_n)$ is a permutation of $(\{1i\}_{i=0}^{s'_1-1},\ldots,\{mi\}_{i=0}^{s'_m-1})$, where $(s'_1,\ldots,s'_m)$ (with $\sum_{t=1}^{m} s'_t = n$) is the (unique) solution of the problem $\mathcal{IP}_n$.

Now, the problem $\mathcal{IP}_n$ does not necessarily have a unique solution for every n. Lemma A.1 and Lemma A.2 discuss the issue of those values of n for which unique solutions to $\mathcal{IP}_n$ exist, and discuss their precise form.

We note here that the issue of uniqueness of the solution to $\mathcal{IP}_n$ is crucial to our treatment. It enables us to show that $A^{(j)}_n$ exist for all large j and that convergence acceleration does take place as $j \to \infty$.

5.4 Study of $E_{\mathbf{ki}(1:n)}$

Let n be such that $\mathcal{IP}_n$ has a unique solution $(s'_1,\ldots,s'_m)$. As already mentioned, $i_{tr} = r-1, r = 1,\ldots,s'_t$ for this solution. Then the corresponding $M^{(j)}_{\mathbf{ki}(1:n)}(z)$ is the dominant term in the summation of (5.6), the rest of the terms being subdominant, provided the corresponding $E_{\mathbf{ki}(1:n)}$ is nonzero.

Let us go back to the first paragraph of Sect. 5.2, where we noted that the n index pairs $k_1i_1,\ldots,k_ni_n$ consist of s_1 pairs of the form $1i_{1r}$, s_2 pairs of the form $2i_{2r}$,…, s_m pairs of the form mi_{mr}. Here $0 \leq i_{t1} < i_{t2} < \cdots < i_{ts_t}$ when $s_t \neq 0$, and $\sum_{t=1}^{m} s_t = n$. Again, for simplicity of notation, let us assume that $s_1,\ldots,s_p$ are nonzero ($p \leq m$) and that the rest of the s_k are zero. Thus, performing on $E_{\mathbf{ki}(1:n)}$ exactly the same column permutations that were performed on $M^{(j)}_{\mathbf{ki}(1:n)}(z)$, we obtain

$$E_{\mathbf{ki}(1:n)} = \pm\det\left[E(1,s_1)\,|\,E(2,s_2)\,|\,\cdots\,|\,E(p,s_p)\right], \tag{5.33}$$

where $E(t,s)$ is an $n \times s$ matrix of the form

$$E(t,s) = \begin{bmatrix} \epsilon_{t0}^{(1)} & \epsilon_{t1}^{(1)} & \cdots & \epsilon_{t,s-1}^{(1)} \\ \epsilon_{t0}^{(2)} & \epsilon_{t1}^{(2)} & \cdots & \epsilon_{t,s-1}^{(2)} \\ \vdots & \vdots & & \vdots \\ \epsilon_{t0}^{(n)} & \epsilon_{t1}^{(n)} & \cdots & \epsilon_{t,s-1}^{(n)} \end{bmatrix}. \tag{5.34}$$

In the sequel, we will refer to the 1st column of $E(t,s)$ as "column 0," to its 2nd column as "column 1," and so on. We now invoke (4.8) and (4.10) in (5.34). By the fact that $\epsilon_{t0} \neq 0$, all the terms in the first column of $E(t,s)$ are nonzero. By adding a suitable multiple of column 0 to column r, we eliminate ϵ_{tr}, from column r, $r = 1, \ldots, s-1$. We then add a multiple of (the new) column 1 to column r to eliminate $\epsilon_{t,r-1}$ from column r, $r = 2, \ldots, s-1$. We next add a multiple of (the new) column 2 to column r to eliminate $\epsilon_{t,r-2}$ from column r, $r = 3, \ldots, s-1$. Continuing this way, we eliminate all of the ϵ_{tr} with $r \neq 0$ from $E(t,s)$, the resulting matrix being

$$E'(t,s) = \begin{bmatrix} \zeta_t^1 \epsilon_{t0} \binom{\gamma_t}{0} 1^0 & \zeta_t^1 \epsilon_{t0} \binom{\gamma_t}{1} 1^1 & \cdots & \zeta_t^1 \epsilon_{t0} \binom{\gamma_t}{s-1} 1^{s-1} \\ \zeta_t^2 \epsilon_{t0} \binom{\gamma_t}{0} 2^0 & \zeta_t^2 \epsilon_{t0} \binom{\gamma_t}{1} 2^1 & \cdots & \zeta_t^2 \epsilon_{t0} \binom{\gamma_t}{s-1} 2^{s-1} \\ \vdots & \vdots & & \vdots \\ \zeta_t^n \epsilon_{t0} \binom{\gamma_t}{0} n^0 & \zeta_t^n \epsilon_{t0} \binom{\gamma_t}{1} n^1 & \cdots & \zeta_t^n \epsilon_{t0} \binom{\gamma_t}{s-1} n^{s-1} \end{bmatrix}. \tag{5.35}$$

Performing these operations on each of the matrices $E(t, s_t)$ separately, we then obtain

$$\begin{aligned} &\det\left[E(1,s_1)\,|\,E(2,s_2)\,|\,\cdots\,|\,E(p,s_p)\right] \\ &\quad = \det\left[E'(1,s_1)\,|\,E'(2,s_2)\,|\,\cdots\,|\,E'(p,s_p)\right]. \end{aligned} \tag{5.36}$$

Factoring out the binomial coefficients from every column of the matrices $E'(t,s_t)$, we obtain

$$\begin{aligned} \det\left[E(1,s_1)\,|\,E(2,s_2)\,|\,\cdots\,|\,E(p,s_p)\right] &= \prod_{t=1}^{p} \left[\epsilon_{t0}^{s_t} \prod_{r=0}^{s_t-1} \binom{\gamma_t}{r} \right] \\ &\quad \times \det\left[E''(1,s_1)\,|\,E''(2,s_2)\,|\,\cdots\,|\,E''(p,s_p)\right], \end{aligned} \tag{5.37}$$

where

$$E''(t,s) = \begin{bmatrix} \zeta_t^1 1^0 & \zeta_t^1 1^1 & \cdots & \zeta_t^1 1^{s-1} \\ \zeta_t^2 2^0 & \zeta_t^2 2^1 & \cdots & \zeta_t^2 2^{s-1} \\ \vdots & \vdots & & \vdots \\ \zeta_t^n n^0 & \zeta_t^n n^1 & \cdots & \zeta_t^n n^{s-1}. \end{bmatrix}. \tag{5.38}$$

We now note that Lemma 4.5 applies, giving

$$\det\left[E''(1,s_1)\,|\,E''(2,s_2)\,|\,\cdots\,|\,E''(p,s_p)\right] = W(\zeta_1,s_1;\ldots;\zeta_p,s_p)\left[\prod_{t=1}^{p}\zeta_t^{s_t}\right] \neq 0.$$

From this and from (5.33) and (5.37), we conclude that $E_{\mathbf{ki}(1:n)} \neq 0$ for the case being considered. This is also the case when $s_t = s'_t$, $t = 1,\ldots,m$.

Remark We would like to draw attention to the fact that if, for some t, $\gamma_t = 0, 1, 2, \ldots,$ then $\binom{\gamma_t}{r} = 0$ for $r > \gamma_t$, and this forces the particular $E_{\mathbf{ki}(1:n)}$ we have just analyzed to vanish, by (5.35)–(5.37). By our assumptions in (1.4), $\gamma_t \neq 0, 1, 2, \ldots,$ hence such a problematic situation cannot occur. Recall also that in the practical examples discussed in Sect. 2, we have $\gamma_t \neq 0, 1, 2, \ldots,$ necessarily.

5.5 Completion of proof of Theorem 3.1

Combining everything, we see that in case the solution $(s'_1,\ldots,s'_m)$ to $\mathcal{IP}_n$ is unique, the determinant $z^{-j}e_n^{(j)}(\{z^s\})$ satisfies the asymptotic equality

$$z^{-j}e_n^{(j)}(\{z^s\}) \sim E_{\mathbf{ki}(1:n)}M^{(j)}_{\mathbf{ki}(1:n)}(z)\left[\prod_{s=1}^{n}\zeta_{k_s}\right]^j$$

where $k_1i_1,\ldots,k_ni_n$ is a permutation of the n index pairs $\{1i\}_{i=0}^{s'_1-1},\ldots,\{mi\}_{i=0}^{s'_m-1}$. We have also seen that $E_{\mathbf{ki}(1:n)} \neq 0$ and that

$$M^{(j)}_{\mathbf{ki}(1:n)}(z) \sim S_1\left[\prod_{t=1}^{m}(z-\zeta_t)^{s'_t}\right] j^{\sigma_n},$$

for some constant $S_1 \neq 0$. As a result, we have

$$z^{-j}e_n^{(j)}(\{z^s\}) \sim S_2\left[\prod_{t=1}^{m}(z-\zeta_t)^{s'_t}\right]\left[\prod_{t=1}^{m}\zeta_t^{s'_t}\right]^j j^{\sigma_n}, \tag{5.39}$$

for some constant $S_2 \neq 0$. Letting $z = 1$ and recalling that $\zeta_k \neq 1$ for all k, we also have

$$e_n^{(j)}(\{I_s\}) \sim S_2\left[\prod_{t=1}^{m}(1-\zeta_t)^{s'_t}\right]\left[\prod_{t=1}^{m}\zeta_t^{s'_t}\right]^j j^{\sigma_n}. \tag{5.40}$$

Because (5.39) and (5.40) are asymptotic equalities, we can divide the former by the latter, invoke (4.4), and obtain the result in (3.4). The result in (3.5) is obtained by invoking [34, p. 31, Lemma 1.4.4].

The result in (3.6) is obtained from (3.4) by recalling that $s'_t = \nu$ for all t when $\Re\gamma_1 = \cdots = \Re\gamma_m$ and $n = m\nu$, $\nu = 1, 2, \ldots$. Similarly, the result in (3.7) follows by letting $s'_t = \nu$ in (3.5).

6 Proof of Theorem 3.2

We start with the observation that the error in $A_n^{(j)}$ has a determinant representation given as in

$$A_n^{(j)} - A = \frac{e_n^{(j)}(\{A_s - A\})}{e_n^{(j)}(\{I_s\})}, \tag{6.1}$$

which is easily obtained from (4.3). Note that, by (5.1), the determinant $e_n^{(j)}(\{A_s - A\})$ has the vector $[A_j - A, A_{j+1} - A, \ldots, A_{j+n} - A]^{\mathrm{T}}$ as its first column. Proceeding as in (5.3), we replace the $a_{j+s}^{(r)}$ and $A_{j+s} - A$ by their asymptotic expansions. The asymptotic expansions of the $a_{j+s}^{(r)}$ are exactly as in (5.3). Now, by (1.3) [or by (4.5)], we have

$$A_n - A \sim \sum_{k=1}^{m} \zeta_k^n \sum_{i=0}^{\infty} \beta_{ki} n^{\gamma_k - i} \quad \text{as } n \to \infty. \tag{6.2}$$

Therefore, the asymptotic expansion of $A_{j+s} - A$ in the first column of the determinant $e_n^{(j)}(\{A_s - A\})$ is as in

$$A_{j+s} - A \sim \sum_{k_0=1}^{m} \zeta_{k_0}^{j+s} \sum_{i_0=0}^{\infty} \beta_{k_0 i_0} (j+s)^{\gamma_{k_0} - i_0} \quad \text{as } j \to \infty.$$

Proceeding exactly as before, we obtain

$$e_n^{(j)}(\{A_s - A\}) \sim \sum_{k_0 i_0} \sum_{k_1 i_1} \cdots \sum_{k_n i_n} \beta_{k_0 i_0} \left[\prod_{s=1}^{n} \epsilon_{k_s i_s}^{(s)}\right] \left[\prod_{s=0}^{n} \zeta_{k_s}\right]^j \widehat{M}_{\mathbf{ki}(0:n)}^{(j)}, \tag{6.3}$$

where

$$\widehat{M}_{\mathbf{ki}(0:n)}^{(j)} = \begin{vmatrix} \zeta_{k_0}^0 j^{\gamma_{k_0} - i_0} & \zeta_{k_1}^0 j^{\gamma_{k_1} - i_1} & \cdots & \zeta_{k_n}^0 j^{\gamma_{k_n} - i_n} \\ \zeta_{k_0}^1 (j+1)^{\gamma_{k_0} - i_0} & \zeta_{k_1}^1 (j+1)^{\gamma_{k_1} - i_1} & \cdots & \zeta_{k_n}^1 (j+1)^{\gamma_{k_n} - i_n} \\ \zeta_{k_0}^2 (j+2)^{\gamma_{k_0} - i_0} & \zeta_{k_1}^2 (j+2)^{\gamma_{k_1} - i_1} & \cdots & \zeta_{k_n}^2 (j+2)^{\gamma_{k_n} - i_n} \\ \vdots & \vdots & & \vdots \\ \zeta_{k_0}^n (j+n)^{\gamma_{k_0} - i_0} & \zeta_{k_1}^n (j+n)^{\gamma_{k_1} - i_1} & \cdots & \zeta_{k_n}^n (j+n)^{\gamma_{k_n} - i_n} \end{vmatrix}. \tag{6.4}$$

Note again that, being a sum of products of a finite number of Poincaré-type asymptotic expansions, the right-hand side of (6.3) is a genuine asymptotic expansion once its terms are ordered according to their sizes.

Now, the product $\widehat{M}^{(j)}_{\mathbf{ki}(0:n)}\left[\prod_{s=0}^{n}\zeta_{k_s}\right]^j$ is odd under an interchange of any two of the index pairs $k_0i_0, k_1i_1, \dots, k_ni_n$, since this amounts to interchanging two columns in the determinant representation of $\widehat{M}^{(j)}_{\mathbf{ki}(0:n)}$, while the term $\left[\prod_{s=0}^{n}\zeta_{k_s}\right]^j$ does not change. Consequently, Lemma 4.3 applies, and we obtain the asymptotic expansion

$$e_n^{(j)}(\{A_s - A\}) \sim \sum_{k_0i_0<k_1i_1<\cdots<k_ni_n} \widehat{E}_{\mathbf{ki}(0:n)}\widehat{M}^{(j)}_{\mathbf{ki}(0:n)}\left[\prod_{s=0}^{n}\zeta_{k_s}\right]^j, \tag{6.5}$$

where

$$\widehat{E}_{\mathbf{ki}(0:n)} = \begin{vmatrix} \beta_{k_0i_0} & \beta_{k_1i_1} & \cdots & \beta_{k_ni_n} \\ \epsilon^{(1)}_{k_0i_0} & \epsilon^{(1)}_{k_1i_1} & \cdots & \epsilon^{(1)}_{k_ni_n} \\ \epsilon^{(2)}_{k_0i_0} & \epsilon^{(2)}_{k_1i_1} & \cdots & \epsilon^{(2)}_{k_ni_n} \\ \vdots & \vdots & & \vdots \\ \epsilon^{(n)}_{k_0i_0} & \epsilon^{(n)}_{k_1i_1} & \cdots & \epsilon^{(n)}_{k_ni_n} \end{vmatrix}. \tag{6.6}$$

Note that the dependence on j in (6.5) comes into play only through the products $\widehat{M}^{(j)}_{\mathbf{ki}(0:n)}\left[\prod_{s=0}^{n}\zeta_{k_s}\right]^j$. Now, by (1.4), the terms $\left[\prod_{s=0}^{n}\zeta_{k_s}\right]^j$ all have the same modulus, namely, $\theta^{(n+1)j}$. Consequently, the dominant terms in the asymptotic expansion of (6.5) are those with the largest $|\widehat{M}^{(j)}_{\mathbf{ki}(0:n)}|$ as $j \to \infty$. Comparing $\widehat{M}^{(j)}_{\mathbf{ki}(0:n)}$ with $M^{(j)}_{\mathbf{ki}(1:n)}(z)$, we realize that the analysis we applied to the latter (with n) applies to the former (with $n+1$) without any change, and we have that the dominant $\widehat{M}^{(j)}_{\mathbf{ki}(0:n)}$ are those for which $(k_0i_0, k_1i_1, \dots, k_ni_n)$ are permutations of $(\{1i\}_{i=0}^{s_1''-1}, \dots, \{mi\}_{i=0}^{s_m''-1})$, where $(s_1'', \dots, s_m'')$ (with $\sum_{t=1}^{m} s_t'' = n+1$) are solutions of the problem $\mathcal{IP}_{n+1}$. For these dominant $\widehat{M}^{(j)}_{\mathbf{ki}(0:n)}$, we thus have

$$\widehat{M}^{(j)}_{\mathbf{ki}(0:n)} = O(j^{\sigma_{n+1}}) \quad \text{as } j \to \infty, \tag{6.7}$$

the coefficient of $j^{\sigma_{n+1}}$ being of the form CL, where C and L are as in (5.23) and (5.27), with $\prod_{t=1}^{p}$ replaced by $\prod_{t=1}^{m}$ and with the s_t replaced by the s_t''. Here, σ_{n+1} is the (maximum) value of $\Re\phi(\boldsymbol{s}, \boldsymbol{i})$, with $\phi(\boldsymbol{s}, \boldsymbol{i})$ defined as in

$$\phi(\boldsymbol{s}, \boldsymbol{i}) = \sum_{t=1}^{m}\left[\sum_{r=1}^{s_t}(\gamma_t - i_{tr}) - s_t(s_t-1)/2\right], \quad \text{with} \quad \sum_{r=1}^{m} s_t = n+1, \tag{6.8}$$

cf. (5.29). The result in (6.7) holds even when the problem $\mathcal{IP}_{n+1}$ does not have a unique solution. [In case of nonuniqueness, every solution to $\mathcal{IP}_{n+1}$ contributes a

$\widehat{M}^{(j)}_{\mathbf{ki}(0:n)}$ that is $O(j^{\sigma_{n+1}})$ asymptotically as $j \to \infty$.] As a result of (6.7), we also have that

$$e_n^{(j)}(\{A_s - A\}) = O(\theta^{(n+1)j} j^{\sigma_{n+1}}) \quad \text{as } j \to \infty. \tag{6.9}$$

As for $\widehat{E}_{\mathbf{ki}(0:n)}$ that corresponds to optimal $\widehat{M}^{(j)}_{\mathbf{ki}(0:n)}$, we have

$$\widehat{E}_{\mathbf{ki}(0:n)} = \pm \det\left[\widehat{E}(1, s_1'') \,|\, \widehat{E}(2, s_2'') \,|\, \cdots \,|\, \widehat{E}(m, s_m'')\right], \tag{6.10}$$

where $\widehat{E}(t, s)$ is an $(n+1) \times s$ matrix of the form

$$\widehat{E}(t, s) = \begin{bmatrix} \beta_{t0} & \beta_{t1} & \cdots & \beta_{t,s-1} \\ \epsilon_{t0}^{(1)} & \epsilon_{t1}^{(1)} & \cdots & \epsilon_{t,s-1}^{(1)} \\ \epsilon_{t0}^{(2)} & \epsilon_{t1}^{(2)} & \cdots & \epsilon_{t,s-1}^{(2)} \\ \vdots & \vdots & & \vdots \\ \epsilon_{t0}^{(n)} & \epsilon_{t1}^{(n)} & \cdots & \epsilon_{t,s-1}^{(n)} \end{bmatrix}. \tag{6.11}$$

[When $s_t'' = 0$, the matrix $\widehat{E}(t, s_t'')$ in (6.10) simply does not exist.] Invoking Lemma 4.2, we can show that $\beta_{ki} = \sum_{l=0}^{i} h_{kil}\epsilon_{kl}$ with $h_{ki0} \propto (1-\zeta_k)^{-i-1}$. As a result, by expanding the determinant in (6.11) with respect to its first row, we obtain

$$\widehat{E}_{\mathbf{ki}(0:n)} \propto \prod_{t=1}^{m}(1-\zeta_t)^{-s_t''}. \tag{6.12}$$

To complete the proof of Theorem 3.2, we divide (6.9) by the asymptotic equality in (5.40), and obtain (3.8). It is also clear from (6.12), that the coefficient that multiplies $\theta^j j^{\sigma_{n+1}-\sigma_n}$ in (3.8) is proportional to the term $\left[\prod_{t=1}^{m}(1-\zeta_t)^{s_t'+s_t''}\right]^{-1}$, where $(s_1'', \ldots, s_m'')$ is the unique solution to the problem $\mathcal{IP}_{n+1}$ when the latter has a unique solution, or it is a sum of such terms when $\mathcal{IP}_{n+1}$ has more than one solution.

The rest of the proof can now be achieved by invoking Lemma A.4 from the appendix. We leave the details to the reader.

7 Implications of Theorems 3.1 and 3.2

Comparing Theorems 3.1 and 3.2 of this paper with Theorems 4.1 and 4.2 in [36], we see that they are very similar. Consequently, the implications of both sets of results are the same. These implications are already discussed in [36]. For completeness, we discuss briefly these implications as they are related to the Shanks transformation.

In accelerating the convergence of general linear sequences such as those treated in this work, one is confronted with slow convergence when $\zeta_t \approx 1$ for some t, irrespective of which method of acceleration is used. The presence of this problem when using the Shanks transformation can be deduced from our main results in this paper.

First, by Theorem 3.1, $\lim_{j\to\infty}\gamma_{ni}^{(j)}$ exist and are all proportional to $\prod_{k=1}^{m}(1-\zeta_k)^{-s_k'}$. This suggests that when $\zeta_t \approx 1$ for some t, $\Gamma_n^{(j)}$ will be large, even though it remains bounded as $j\to\infty$. Next, as $j\to\infty$, the (theoretical) error $A_n^{(j)}-A$ is proportional to $\big[\prod_{t=1}^{m}(1-\zeta_t)^{s_t'+s_t''}\big]^{-1}$, or to a sum of such terms, as discussed in the proof of Theorem 3.2 at the end of Sect. 6. This also suggests that when $\zeta_t\approx 1$ for some t, this error will be large, hence the sequence $\{A_n^{(j)}\}_{j=0}^{\infty}$ will converge slowly, mathematically speaking. Numerical experience suggests that these conclusions are valid also for diagonal sequences $\{A_n^{(j)}\}_{n=0}^{\infty}$ with fixed j, which seem to have the best convergence properties, even though Theorems 3.1 and 3.2 do not apply to diagonal sequences.

Both problems can be remedied by applying the Shanks transformation to a subsequence $\{A_{\kappa n}\}_{n=0}^{\infty}$ with some suitable fixed integer $\kappa>1$, as proposed in [34, Chapter 16, pp. 315–316]. This choice of the subsequence has been called *arithmetic progression sampling* (APS for short) in [34]. That this will result in improvement can be deduced from the fact that

$$A_{\kappa n} \sim A + \sum_{k=1}^{m}\zeta_k^{\kappa n}\sum_{i=0}^{\infty}\beta_{ki}(\kappa n)^{\gamma_k - i} \quad \text{as } n\to\infty,$$

which, defining

$$\widetilde{A}_n = A_{\kappa n}, \quad \widetilde{\zeta}_k = \zeta_k^{\kappa}, \quad \widetilde{\beta}_{ki} = \beta_{ki}\kappa^{\gamma_k - i},$$

can be rewritten as

$$\widetilde{A}_n \sim A + \sum_{k=1}^{m}\widetilde{\zeta}_k^{\,n}\sum_{i=0}^{\infty}\widetilde{\beta}_{ki}n^{\gamma_k - i} \quad \text{as } n\to\infty.$$

Note that $\{\widetilde{A}_n\}$ is also a general linear sequence, with $\widetilde{\zeta}_k$ instead of ζ_k. Thus, when the Shanks transformation is applied to this sequence, Theorems 3.1 and 3.2 now hold with $\widetilde{\zeta}_k$ instead of ζ_k. Whether $|\zeta_t|\le 1$ or $|\zeta_t|\ge 1$, when ζ_t is close to 1, we have that $\widetilde{\zeta}_t$ is farther from 1 than ζ_t is, even for $\kappa=2$. Therefore, by choosing κ appropriately, we can cause $\widetilde{\zeta}_t$ to be away from 1 sufficiently to improve the convergence and stability of the column sequences. Numerical examples show that the performance of the diagonal sequences also improves under APS.

APS can be applied to power series $\sum_{n=0}^{\infty}c_n z^n$, where

$$c_n \sim \sum_{k=1}^{m}\sigma_k^n\sum_{i=0}^{\infty}\epsilon_{ki}n^{\gamma_k - i} \quad \text{as } n\to\infty, \quad \sigma_k\neq 1 \quad \text{distinct.}$$

This series converges to a function $A(z)$ that is analytic for $|z| < \min_k(1/|\sigma_k|) = \rho$. In addition, letting $A_n(z) = \sum_{s=0}^{n} c_s z^s$, $n = 0, 1, \ldots,$ we also have

$$A_n(z) \sim A(z) + \sum_{k=1}^{m} (\sigma_k z)^n \sum_{i=0}^{\infty} \beta_{ki}(z) n^{\gamma_k - i} \quad \text{as } n \to \infty,$$

as mentioned in Sect. 2. Clearly, $\zeta_k = \sigma_k z, k = 1, \ldots, m$, in the notation of the preceding sections. In general, the function $A(z)$ can be continued analytically to $|z| \geq \rho$, $z_k = 1/\sigma_k$ being its branch points, with the branch cuts directed appropriately. [As an example, think of $A(z)$ to be the sum of m functions that have branch points at m distinct points $z_1, \ldots, z_m$ in the z-plane.] Then $\zeta_t \approx 1$ means $z \approx z_t = 1/\sigma_t$, which in turn means that z is near a point of singularity. Thus, we conclude that the Shanks transformation will converge slowly when z is close to a point of singularity of $A(z)$. Thus, close to points of singularity, applying the Shanks transformation with APS is likely to be beneficial.

Our numerical experiments with Fourier series do show that the Shanks transformation with APS does improve the convergence rate of diagonal approximations greatly; that is, the convergence of $A_n^{(j)}$ as $n \to \infty$ improves with increasing κ. However, the number of series elements used to obtain a given level of accuracy does not seem to change much with increasing κ in using APS, however. There is a small advantage to using APS for Fourier series: Given that the terms of the series are already available, the overhead of computing the $A_n^{(j)}$ via the epsilon algorithm, for example, decreases by a factor of κ^2. We will observe all this in the example of the next section.

The topic of APS, initiated within the context of the Levin–Sidi $d^{(m)}$ transformation in [21], is discussed and analyzed in Sidi [30] and [34, Chapters 10, 12, and 13]. For applications of APS with the $d^{(m)}$ transformation for arbitrary $m \geq 1$, to power series, Fourier series, and generalized Fourier series, see [21,30], and [34, Chapter 6, Sect. 6.5 and Chapter 12, Sect. 12.9]. Its use in conjunction with the Shanks transformation is proposed in [34, Chapter 16, pp. 315–316]. The results of the present work provide further theoretical justification for the use of APS within the context of the Shanks transformation applied to general linear sequences.

8 A numerical example

Consider the application of the Shanks transformation to the slowly convergent series

$$\frac{h}{\pi} + \sum_{s=1}^{\infty} \frac{2}{\pi} \frac{\sin sh}{s} \cos sx = H(h - |x|), \quad H(x) = \begin{cases} 0 & \text{if } x < 0, \\ 1 & \text{if } x > 0. \end{cases}$$

The function $H(h - |x|)$ has finite jumps at the two points $x = h$ and $x = -h$. We can actually conclude that the sum of this Fourier series is singular at these two points

Table 1 Errors $E_n^{(0)} = |A_n^{(0)} - f(x)|$ obtained from the Shanks transformation with APS for the example of Sect. 8 with $h = 1$ and $x = 0.9$ and $\kappa = 1, 5, 10$

k	$E_{40k}^{(0)}(\kappa = 1)$	$E_{8k}^{(0)}(\kappa = 5)$	$E_{4k}^{(0)}(\kappa = 10)$
0	6.817D − 01	6.817D − 01	6.817D − 01
1	8.941D − 02	8.022D − 02	2.001D − 02
2	8.901D − 03	8.091D − 03	1.177D − 01
3	6.615D − 04	8.429D − 04	3.425D − 05
4	7.316D − 05	4.290D − 05	4.268D − 06
5	9.444D − 05	1.705D − 05	5.536D − 07
6	1.806D − 05	8.717D − 07	3.881D − 08
7	2.429D − 07	3.400D − 07	1.722D − 08
8	5.904D − 08	2.799D − 09	4.498D − 11
9	1.059D − 09	2.860D − 09	1.525D − 11
10	6.171D − 10	9.039D − 11	2.620D − 11
11	1.995D − 11	7.901D − 11	8.193D − 13
12	4.561D − 11	5.531D − 12	3.661D − 15
13	7.601D − 13	6.868D − 13	1.218D − 14
14	2.870D − 12	1.663D − 13	1.438D − 15
15	1.534D − 14	6.744D − 15	6.636D − 17
16	8.339D − 14	9.473D − 14	3.956D − 18
17	5.223D − 16	1.881D − 16	5.180D − 18
18	1.543D − 17	1.272D − 17	1.438D − 20
19	4.072D − 17	3.987D − 18	1.825D − 20
20	2.341D − 17	4.958D − 19	8.844D − 21
21	2.515D − 19	3.746D − 20	6.474D − 21
22	9.175D − 20	2.173D − 21	2.650D − 24
23	4.071D − 21	2.151D − 22	4.046D − 24
24	1.714D − 22	2.724D − 22	9.598D − 25
25	1.473D − 22	1.422D − 23	1.037D − 25

by analyzing its coefficients, namely, $a_n = \frac{2}{\pi}(\sin nh \cos nx)/n$. It is easy to see that

$$a_n = \frac{1}{2\mathrm{i}\pi n}\left[e^{\mathrm{i}n(h+x)} + e^{\mathrm{i}n(h-x)} - e^{-\mathrm{i}n(h+x)} - e^{-\mathrm{i}n(h-x)}\right],$$

which is exactly of the form (2.1) with $m = 4$, $\zeta_1 = e^{\mathrm{i}(h+x)} = \zeta_3^{-1}$, $\zeta_2 = e^{\mathrm{i}(h-x)} = \zeta_4^{-1}$, and $\gamma_k = -1, k = 1, \ldots, 4$. Clearly, $\zeta_1 = 1 = \zeta_3$ when $x = -h$ and $\zeta_2 = 1 = \zeta_4$ when $x = h$.

We have applied the Shanks transformation with $h = 1$ and $x = 0.9$ for which the sum of the series is 1. Clearly, we are very near a point of singularity, and hence will have slow convergence if we work with the whole sequence of the partial sums. In Table 1, we give the errors for the diagonal sequences $\{A_n^{(0)}\}_{n=0}^{\infty}$, obtained with APS taking $\kappa = 1, 5, 10$. Note that the convergence of the diagonal sequences improves with increasing κ. It must be also noted that the number of terms of the Fourier series used in computing the approximations $A_{40k}^{(0)}$ with $\kappa = 1$, $A_{8k}^{(0)}$ with $\kappa = 5$, and $A_{4k}^{(0)}$ with $\kappa = 10$ is $80k$; in other words, we are comparing diagonal approximations obtained from the same number of terms. From this, it is clear that all three approximations have very similar accuracies; the accuracy seems to be improving only marginally with increasing κ. However, the number of the entries in the epsilon table of Wynn that need to be computed decreases as κ^{-2} with increasing κ.

Acknowledgments This research was supported in part by the United States–Israel Binational Science Foundation grant no. 2008399. The author would like to thank Prof. Michael Kaminski for very valuable discussions and help concerning the appendix to this paper.

Appendix

Let $\boldsymbol{s} = (s_1, \ldots, s_m)$, where s_t are nonnegative integers, and let $\{\alpha_t\}_{t=1}^m$ be arbitrary real numbers. Define the function $g(\boldsymbol{s})$ via

$$g(\boldsymbol{s}) = \sum_{t=1}^{m} \alpha_t s_t - \sum_{t=1}^{m} s_t(s_t - 1). \tag{A.1}$$

We are interested in the solution to the integer programming problem

$$\max_{\boldsymbol{s}} g(\boldsymbol{s})$$
$$\text{subject to} \quad s_1 \geq 0, \ldots, s_m \geq 0 \quad \text{and} \quad \sum_{t=1}^{m} s_t = n. \tag{A.2}$$

where n is a given positive integer. We will denote this problem by $\mathcal{IP}_n$, and we will denote by σ_n the value of $g(\boldsymbol{s})$ at the solution to $\mathcal{IP}_n$.

A closed-form solution to $\mathcal{IP}_n$ does not seem to be possible when the α_t are arbitrary. Such a solution is possible, however, when the α_t are all the same. This solution is given in the next lemma.

Lemma A.1 *Let*

$$\alpha_1 = \cdots = \alpha_m \equiv \alpha. \tag{A.3}$$

Let also $\mathbf{s} = (s_1, \ldots, s_m)$ *be a solution to* $\mathcal{IP}_n$. *Then the following are true:*

1. *If* $n = m\nu$, *where* $\nu = 1, 2, \ldots,$ *then* $s_1 = \cdots = s_m = \nu$ *and this solution is unique.*
2. *If* $n = m\nu + \tau$, *where* $\nu = 0, 1, \ldots,$ *and* $1 \leq \tau \leq m - 1$, *then* $m - \tau$ *of the* s_t *are equal to* ν *while the remaining* τ *are equal to* $\nu + 1$, *and there are* $\binom{m}{\tau}$ *such solutions.*

Proof Letting $\mu = n/m$, for arbitrary n, we have

$$\begin{aligned} g(\boldsymbol{s}) &= n(\alpha + 1) - \sum_{t=1}^{m} s_t^2 \\ &= n(\alpha + 1) - \sum_{t=1}^{m} \left[(s_t - \mu)^2 + 2\mu(s_t - \mu) + \mu^2 \right] \\ &= n\,(\alpha + 1 - \mu) - \sum_{t=1}^{m} (s_t - \mu)^2, \end{aligned} \tag{A.4}$$

where we have invoked the constraint $\sum_{t=1}^{m} s_t = n$ several times. Clearly, the largest value of $g(\boldsymbol{s})$, for not necessarily integer s_t, is achieved only when $s_t = \mu$ for all t.

Part 1 of the lemma (with $n = m\nu$, $\nu = 1, 2, \dots$, and hence $\mu = \nu$) follows from (A.4).

For part 2 (with $n = m\nu + \tau$, $1 \leq \tau \leq m - 1$), assume that the optimal solution $\boldsymbol{s} = (s_1, \dots, s_m)$ is not of the form described in the statement of the lemma. Then we can find s_i and s_j such that $s_i - s_j \geq 2$. For simplicity of notation, let $i = 1$ and $j = 2$; therefore, $s_1 - s_2 = k \geq 2$. Let us now decrease s_1 by 1 and increase s_2 by 1, keeping $s_3, \dots, s_m$ unchanged. With $\bar{\boldsymbol{s}} = (s_1 - 1, s_2 + 1, s_3, \dots, s_m)$, we have

$$g(\bar{\boldsymbol{s}}) = g(\boldsymbol{s}) + 2(k - 1) > g(\boldsymbol{s}),$$

since $k \geq 2$. This contradicts the fact that $g(\boldsymbol{s}) \geq g(\bar{\boldsymbol{s}})$. [Note that $g(\bar{\boldsymbol{s}}) = g(\boldsymbol{s})$ with $k = 1$.] Actually, by continuing this process with all the s_t, we increase $g(\boldsymbol{s})$ until we finally reach the situation in which $m - \tau$ of the s_t are equal to ν while the remaining τ are equal to $\nu + 1$. □

As is clear from the preceding lemma, and as can be verified by working out some simple examples with general α_t, a unique solution $\boldsymbol{s}$ to $\mathcal{IP}_n$ may not exist for all values of n. The next lemma shows that there are infinitely many values of n for which a unique solution to $\mathcal{IP}_n$ exists, whether the α_t are all equal or not. In addition, the proof of this lemma also gives a simple procedure by which one can construct the solutions to $\mathcal{IP}_n$ with these special values of n.

Lemma A.2 *There is an infinite sequence of positive integers $n_1 < n_2 < \cdots$, such that $\mathcal{IP}_{n_i}$ has a unique solution $\boldsymbol{s}^{(i)} = (s_1^{(i)}, \dots, s_m^{(i)})$ for each i. If we order the α_t as in*

$$\alpha_1 = \cdots = \alpha_\mu > \alpha_{\mu+1} \geq \cdots \geq \alpha_m, \tag{A.5}$$

then $\mu \leq n_1 \leq m$. In addition, there exists a positive integer κ, for which

$$\mu \leq n_{i+1} - n_i < m \quad \text{when } i < \kappa \quad \text{and} \quad n_{i+1} - n_i = m \quad \text{when } i \geq \kappa. \tag{A.6}$$

The solution to $\mathcal{IP}_{n_i}$ satisfies

$$s_1^{(i)} = \cdots = s_\mu^{(i)} \geq s_{\mu+1}^{(i)} \geq \cdots \geq s_m^{(i)} \quad \text{for every } i. \tag{A.7}$$

In addition, the solutions to $\mathcal{IP}_{n_i}$ and $\mathcal{IP}_{n_{i+1}}$ are related via

$$s_t^{(i+1)} = s_t^{(i)} + 1, \quad t = 1, \dots, m, \quad \text{for } i \geq \kappa. \tag{A.8}$$

Proof Let us consider the problem $\mathcal{IP}_n$ in (A.2), with n undetermined as yet. Let us choose a real number c and set $\alpha_t + c = 2\beta_t$, $t = 1, \dots, m$. Then we can express $g(\boldsymbol{s})$

in (A.1) as in

$$g(\boldsymbol{s}) = \left[n(1-c) + \sum_{t=1}^{m} \beta_t^2 \right] - \sum_{t=1}^{m} (s_t - \beta_t)^2, \tag{A.9}$$

and we have

$$\sigma_n = \left[n(1-c) + \sum_{t=1}^{m} \beta_t^2 \right] - \min_{\boldsymbol{s}} \sum_{t=1}^{m} (s_t - \beta_t)^2,$$

$$\text{subject to} \quad s_1 \geq 0, \ldots, s_m \geq 0 \quad \text{and} \quad \sum_{t=1}^{m} s_t = n. \tag{A.10}$$

We now construct the n_i alluded to above by assigning certain suitable values to c and then by solving the problem $\min_{\boldsymbol{s}} \sum_{t=1}^{m} (s_t - \beta_t)^2$ for nonnegative integers s_t. It is easy to see that the vector $\boldsymbol{s}$, for which s_t is the smallest nonnegative integer closest to β_t, $t = 1, \ldots, m$, is a solution to this problem.

Choosing $c = 2 - \alpha_1$ so that $\beta_1 = 1$, with the α_t ordered as in (A.5), we have

$$1 = \beta_1 = \cdots = \beta_\mu > \beta_{\mu+1} \geq \cdots \geq \beta_m.$$

Let us now construct a solution $\boldsymbol{s}^{(1)} = (s_1^{(1)}, \ldots, s_m^{(1)})$ to the problem $\min_{\boldsymbol{s}} \sum_{t=1}^{m} (s_t - \beta_t)^2$, subject to the constraints $s_t \geq 0$, $t = 1, \ldots, m$: We let $s_1^{(1)} = \cdots = s_\mu^{(1)} = 1$. For $\mu + 1 \leq t \leq m$, we proceed as follows: (i) when $\beta_t \leq 1/2$, we let $s_t^{(1)} = 0$. (ii) when $\beta_t > 1/2$, we let $s_t^{(1)} = 1$. Finally, with all the $s_t^{(1)}$ determined, we set $n = \sum_{t=1}^{m} s_t^{(1)} \equiv n_1$. Clearly, $\mu \leq n_1 \leq m$. It is easy to see that $\sum_{t=1}^{m} (s_t^{(1)} - \beta_t)^2 = \min_{\boldsymbol{s}} \sum_{t=1}^{m} (s_t - \beta_t)^2$ subject to $s_t \geq 0, t = 1, \ldots, m$, and $\sum_{t=1}^{m} s_t = n_1$. Consequently, $\boldsymbol{s}^{(1)}$ is the unique vector that maximizes $g(\boldsymbol{s})$ subject to $\sum_{t=1}^{m} s_t = n_1$, hence is the unique solution to the problem $\mathcal{IP}_{n_1}$.

To construct $\boldsymbol{s}^{(2)}$, we increase c by 2. This causes all the β_t to increase by 1. Thus, we have $\beta_1 = \cdots = \beta_\mu = 2$. We let $s_1^{(2)} = \cdots = s_\mu^{(2)} = 2$. For $\mu + 1 \leq t \leq m$, we proceed as follows: (i) when $\beta_t \neq p + 1/2$, where p is an integer, we let $s_t^{(2)}$ be the nonnegative integer that is closest to β_t. (ii) when $\beta_t = p + 1/2$, where p is an integer, we let $s_t^{(2)} = 0$ if $\beta_t < 0$ and $s_t^{(2)} = \beta_t - 1/2$ otherwise; thus, we have either $s_t^{(2)} = s_t^{(1)}$ or $s_t^{(2)} = s_t^{(1)} + 1$. With all the $s_t^{(2)}$ determined, set $n = \sum_{t=1}^{m} s_t^{(2)} \equiv n_2$. It is clear that $\boldsymbol{s}^{(2)}$ is the unique vector that maximizes $g(\boldsymbol{s})$ subject to $\sum_{t=1}^{m} s_t = n_2$, hence is the unique solution to the problem $\mathcal{IP}_{n_2}$. Clearly, $n_1 + \mu \leq n_2 \leq n_1 + m$.

By increasing c by 2 continually, we generate the vectors $\boldsymbol{s}^{(i)}$ and determine the integers n_i, $\boldsymbol{s}^{(i)}$ being the unique solutions to the corresponding problems $\mathcal{IP}_{n_i}$. As mentioned earlier, these solutions are obtained by choosing the s_t to be the smallest nonnegative integers closest to the respective β_t. First, we choose c such that $\alpha_1 + c = 2i$, and let $\alpha_t + c = 2\beta_t$ for all t, as before. Of course, $\beta_1 = \cdots = \beta_\mu = i$.

Next, we determine $\boldsymbol{s}^{(i)}$:

$$\begin{array}{ll} \text{if } \beta_t < 0, & s_t^{(i)} = 0, \\ \text{if } p \le \beta_t \le p + 1/2, \quad p \ge 0 \text{ integer}, & s_t^{(i)} = p, \\ \text{if } p + 1/2 < \beta_t < p + 1, \quad p \ge 0 \text{ integer}, & s_t^{(i)} = p + 1. \end{array}$$

Clearly, $s_1^{(i)} = \cdots = s_\mu^{(i)} = i$. After a number of such steps, say $\kappa + 1$ steps, we reach a situation in which $\beta_t > 1/2$ for all t so that $s_t^{(\kappa+1)} > 0$ for all t. Thus, $s_t^{(i+1)} = s_t^{(i)} + 1$ for all t, and $n_{i+1} = n_i + m$, for $i \ge \kappa$. This completes the proof of the lemma. □

The next lemma shows how to construct a solution to $\mathcal{IP}_{n+1}$ when $\mathcal{IP}_n$ has a unique solution as constructed in Lemma A.2.

Lemma A.3 *Let $\mathbf{s}' = (s_1', \ldots, s_m')$ be the unique solution to $\mathcal{IP}_n$ when n is one of the integers n_i in Lemma A.2. Let $c = 2s_1' - \alpha_1$ and $\beta_t = (\alpha_t + c)/2$ for all t. Next, let $\mathbf{s}'' = (s_1'', \ldots, s_m'')$ be given as follows: $s_t'' = s_t'$ if $t \ne q$, and $s_q'' = s_q' + 1$, where q is determined via*

$$|s_q' + 1 - \beta_q| = \min_{1 \le t \le m} |s_t' + 1 - \beta_t|. \tag{A.11}$$

Then $\mathbf{s}''$ is a (not necessarily unique) solution to $\mathcal{IP}_{n+1}$.

Remarks 1. Let us define $\mathbb{Z}^+ = \{0, 1, 2, \ldots\}$ and define the sets T_i, $1 \le i \le 4$, as follows:

$$\begin{aligned} T_1 &= \{\beta_t : \ \beta_t < 0\}, \\ T_2 &= \{\beta_t : \ \beta_t = p, \ p \in \mathbb{Z}^+\}, \\ T_3 &= \{\beta_t : \ p < \beta_t \le p + 1/2, \ p \in \mathbb{Z}^+\}, \\ T_4 &= \{\beta_t : \ p + 1/2 < \beta_t < p + 1, \ p \in \mathbb{Z}^+\}. \end{aligned}$$

Let us set $X_t = |s_t' + 1 - \beta_t|$ and observe that

$$\begin{array}{ll} X_t > 1 & \text{if } \beta_t \in T_1 \text{ or } \beta_t \in T_4, \\ X_t = 1 & \text{if } \beta_t \in T_2, \\ 1/2 \le X_t < 1 & \text{if } \beta_t \in T_3. \end{array}$$

From this, it is clear that either (i) $T_3 = \emptyset$ and β_q is any member of T_2, in which case, $s_q' - \beta_q = 0$ and $s_q'' - \beta_q = 1$, or (ii) $T_3 \ne \emptyset$ and $\beta_q \in T_3$, in which case, $-1/2 \le s_q' - \beta_q < 0$ and $1/2 \le s_q'' - \beta_q < 1$, and $\beta_q - \lfloor \beta_q \rfloor = \max_{\beta_t \in T_3}(\beta_t - \lfloor \beta_t \rfloor)$. Note that T_2 is never empty since it contains $\beta_1, \ldots, \beta_\mu$, which are positive integers. Note also that $\beta_q \in T_1$ and $\beta_q \in T_4$ are impossible. These facts are obvious from the construction in the proof of Lemma A.2.

2. By (A.9), the function $g(\boldsymbol{s})$ for the problem $\mathcal{IP}_{n+1}$ is

$$g(\boldsymbol{s}) = \left[(n+1)(1-c) + \sum_{t=1}^{m} \beta_t^2\right] - \sum_{t=1}^{m}(s_t - \beta_t)^2; \quad \sum_{t=1}^{m} s_t = n+1. \quad \text{(A.12)}$$

Proof Define $h(\boldsymbol{s}) = \sum_{t=1}^{m}(s_t - \beta_t)^2$. We need to show that for any $\boldsymbol{s} \neq \boldsymbol{s}''$ satisfying $\sum_{t=1}^{m} s_t = n+1$, we will have $h(\boldsymbol{s}) \geq h(\boldsymbol{s}'')$, and hence that $g(\boldsymbol{s}) \leq g(\boldsymbol{s}'')$. There are two cases to consider: (i) $s_q = s_q''$ and (ii) $s_q \neq s_q''$.

(i) When $s_q = s_q''$, we have $\sum_{t=1, t\neq q}^{m} s_t = n - s_q'$ and $\sum_{t=1, t\neq q}^{m}(s_t - \beta_t)^2 \geq \sum_{t=1, t\neq q}^{m}(s_t'' - \beta_t)^2$, because, being the smallest nonnegative integers closest to the respective β_t, $s_t'' = s_t'$ with $t \neq q$ are optimal for the problem $\min_{s_t, t\neq q} \sum_{t=1, t\neq q}^{m}(s_t - \beta_t)^2$ subject to $\sum_{t=1, t\neq q}^{m} s_t = n - s_q'$. Hence $h(\boldsymbol{s}) \geq h(\boldsymbol{s}'')$.

(ii) When $s_q \neq s_q''$. We distinguish between two cases:

1. The case $s_q \geq s_q'' + 1 = s_q' + 2$ or $s_q \leq s_q'' - 2 = s_q' - 1$: We have $|s_q - \beta_q| \geq |s_q'' - \beta_q| = |s_q' + 1 - \beta_q|$. In addition, $|s_t - \beta_t| \geq |s_t' - \beta_t| = |s_t'' - \beta_t|$ for all $t \neq q$. Therefore, $h(\boldsymbol{s}) \geq h(\boldsymbol{s}'')$.
2. The case $s_q = s_q'' - 1 = s_q'$: First, we note that $s_i \geq s_i'' + 1 = s_i' + 1$ for some $i \neq q$ at least once in this case, since $\sum_{t=1, t\neq q}^{m} s_t = \sum_{t=1, t\neq q}^{m} s_t' + 1$. Then, because $|s_t - \beta_t| \geq |s_t' - \beta_t|$ for all $t \neq i, q$, it is sufficient to show that

$$A \equiv (s_i - \beta_i)^2 + (s_q - \beta_q)^2 \geq (s_i'' - \beta_i)^2 + (s_q'' - \beta_q)^2 \equiv B. \quad \text{(A.13)}$$

Let $s_i = s_i' + k$ for some integer $k \geq 1$. Then, recalling that $s_q = s_q'$ and $s_i'' = s_i'$, we have

$$\begin{aligned} A - B &= 2k(s_i' + 1 - \beta_i) - 2(s_q' + 1 - \beta_q) + (k-1)^2 \\ &= 2k|s_i' + 1 - \beta_i| - 2|s_q' + 1 - \beta_q| + (k-1)^2. \end{aligned}$$

By (A.11) and by the assumption that $k \geq 1$, we realize that $A - B \geq 0$. Therefore, $h(\boldsymbol{s}) \geq h(\boldsymbol{s}'')$.

This completes the proof. □

The next lemma provides an explicit and simple expression for $\sigma_{n+1} - \sigma_n$ for the case in which $\mathcal{IP}_n$ has a unique solution, as in Lemma A.1 or as in Lemma A.2.

Lemma A.4 *The optimal values σ_n and σ_{n+1} of $g(\boldsymbol{s})$ for $\mathcal{IP}_n$ and $\mathcal{IP}_{n+1}$, respectively, satisfy the following:*

1. *If $\alpha_1 = \cdots = \alpha_m \equiv \alpha$, and $n = m\nu$, where ν is a nonnegative integer, then*

$$\sigma_{n+1} - \sigma_n = \alpha - 2\nu. \quad \text{(A.14)}$$

2. *When the* α_t *are not necessarily equal to each other, and are ordered as in* (A.5), *let* n *be one of the integers constructed in Lemma* A.2. *Let* $c = 2\beta_1 - \alpha_1$, *and set* $\beta_t = (\alpha_t + c)/2$, $1 \leq t \leq m$. *Let* $\boldsymbol{s}' = (s'_1, \ldots, s'_m)$ *be the (unique) solution to the problem* $\mathcal{IP}_n$, *and determine the index* q *such that* $|s'_q + 1 - \beta_q| = \min_t |s'_t + 1 - \beta_t|$. *Then*

$$\sigma_{n+1} - \sigma_n = \alpha_q - 2s'_q \quad \textit{and} \quad \sigma_{n+1} - \sigma_n \leq \alpha_1 - 2s'_1 + 1. \tag{A.15}$$

If $n = n_{\kappa+\nu}$, *where* κ *is as in Lemma* A.2 *and* $\nu \geq 0$, *we have*

$$\sigma_{n+1} - \sigma_n = \big(\alpha_q - 2\widehat{s}_q\big) - 2\nu, \quad \textit{and} \quad \sigma_{n+1} - \sigma_n \leq (\alpha_1 - 2\widehat{s}_1 + 1) - 2\nu, \tag{A.16}$$

$(\widehat{s}_1, \ldots, \widehat{s}_m)$ *being the (unique) solution to the problem* $\mathcal{IP}_{n_\kappa}$.

Proof Part 1 follows directly from Lemma A.1. For the proof of part 2, we define $h(\boldsymbol{s}) = \sum_{t=1}^{m} (s_t - \beta_t)^2$, as before. Then, by (A.10) in the proof of Lemma A.3, we have

$$\sigma_n = \left[n(1-c) + \sum_{t=1}^{m} \beta_t^2 \right] - h(\boldsymbol{s}')$$

and, by Lemma A.3, we also have

$$\sigma_{n+1} = \left[(n+1)(1-c) + \sum_{t=1}^{m} \beta_t^2 \right] - h(\boldsymbol{s}''),$$

where $\boldsymbol{s}'' = (s''_1, \ldots, s''_m)$ is a (not necessarily unique) solution to the problem $\mathcal{IP}_{n+1}$ that is described in Lemma A.3. [Note that σ_{n+1} is the (unique) optimal value of the function $g(\boldsymbol{s})$ at any one of the solutions to $\mathcal{IP}_{n+1}$.] Consequently,

$$\sigma_{n+1} - \sigma_n = 1 - c + h(\boldsymbol{s}') - h(\boldsymbol{s}'').$$

Now, by Lemma A.3, we also have

$$h(\boldsymbol{s}'') = h(\boldsymbol{s}') + \big[(s'_q + 1 - \beta_q)^2 - (s'_q - \beta_q)^2\big] = h(\boldsymbol{s}') + 2(s'_q - \beta_q) + 1. \tag{A.17}$$

As a result,

$$\sigma_{n+1} - \sigma_n = 2(\beta_q - s'_q) - c = \alpha_q - 2s'_q,$$

where we have invoked the relation $\alpha_t + c = 2\beta_t$. Next, by (A.17) and by the fact that $-1/2 \leq s'_q - \beta_q \leq 0$,

$$h(\boldsymbol{s}') \leq h(\boldsymbol{s}'') \leq h(\boldsymbol{s}') + 1,$$

from which

$$\sigma_{n+1} - \sigma_n \leq 1 - c = \alpha_1 - 2s_1' + 1.$$

Here we have also invoked the fact that $\alpha_1 + c = 2\beta_1 = 2s_1'$. We have thus shown the validity of (A.15). For (A.16), we make use of (A.6) and (A.8) to conclude that when $n = n_{\kappa+\nu}$, we have $n = n_\kappa + m\nu$ and $s_t' = \widehat{s}_t + \nu$ for all t. We leave the details to the reader. □

References

1. Atkinson, K.E.: An Introduction to Numerical Analysis. Wiley, New York (1978)
2. Baker, G.A. Jr..: Essentials of Padé Approximants. Academic Press, New York (1975)
3. Baker, G.A. Jr.., Graves-Morris, P.R.: Padé Approximants, 2nd edn. Cambridge University Press, Cambridge (1996)
4. Beckermann, B., Matos, A.C., Wielonsky, F.: Reduction of the Gibbs phenomenon for smooth functions with jumps by the ϵ-algorithm. J. Comput. Appl. Math. **219**, 329–349 (2008)
5. Bleistein, N., Handelsman, R.A.: Asymptotic Expansions of Integrals. Holt, Rinehart and Winston, New York (1975)
6. Brezinski, C.: Accélération de suites à convergence logarithmique. C. R. Acad. Sci. Paris **273A**, 727–730 (1971)
7. Brezinski C.: Accélération de la Convergence en Analyse Numérique. In: Lecture Notes in Mathematics, vol. 584. Springer, Berlin (1977)
8. Brezinski, C.: Extrapolation algorithms for filtering series of functions, and treating the Gibbs phenomenon. Numer. Algorithms **36**, 309–329 (2004)
9. Brezinski, C., Redivo Zaglia, M.: Extrapolation Methods: Theory and Practice. North-Holland, Amsterdam (1991)
10. de Montessus de Ballore, R.: Sur les fractions continue algébriques. Bull. Soc. Math. France **30**, 28–36 (1902)
11. Ford, W.F., Sidi, A.: An algorithm for a generalization of the Richardson extrapolation process. SIAM J. Numer. Anal. **24**, 1212–1232 (1987)
12. Garibotti, C.R., Grinstein, F.F.: Recent results relevant to the evaluation of infinite series. J. Comput. Appl. Math. **9**, 193–200 (1983)
13. Gautschi, W.: Orthogonal Poynomials: Computation and Approximation. Numerical Mathematics and Scientific Computation. Oxford University Press, Oxford (2004)
14. Gilewicz, J.: Approximants de Padé. In: Lecture Notes in Mathematics, vol. 667. Springer, New York (1978)
15. Gray, H.L., Atchison, T.A., McWilliams, G.V.: Higher order G-transformations. SIAM J. Numer. Anal. **8**, 365–381 (1971)
16. Guilpin, C., Gacougnolle, J., Simon, Y.: The ϵ-algorithm allows to detect Dirac delta functions. Appl. Numer. Math. **48**, 27–40 (2004)
17. Hardy, G.H.: Divergent Series. Clarendon Press, Oxford (1949)
18. Kaminski, M., Sidi, A.: Solution of an integer programming problem related to convergence of rows of Padé approximants. Appl. Numer. Math. **8**, 217–223 (1991)
19. Karlsson, J., Wallin, H.: Rational approximation by an interpolation procedure in several variables. In: Saff, E.B., Varga, R.S. Padé and Rational Approximation, pp. 83–100. Academic Press, New York (1977)
20. Levin, D.: Development of non-linear transformations for improving convergence of sequences. Int. J. Comput. Math. B **3**, 371–388 (1973)
21. Levin, D., Sidi, A.: Two new classes of nonlinear transformations for accelerating the convergence of infinite integrals and series. Appl. Math. Comp. **9**, 175–215 (1981). Originally appeared as a Tel Aviv University preprint in 1975
22. Lubinsky, D.S.: Padé tables of entire functions of very slow and smooth growth. Constr. Approx. **1**, 349–358 (1985)

23. Lubinsky, D.S.: Padé tables of entire functions of very slow and smooth growth II. Constr. Approx. **4**, 321–339 (1988)
24. Ralston, A., Rabinowitz, P.: A First Course in Numerical Analysis, 2nd edn. McGraw-Hill, New York (1978)
25. Rutishauser, H.: Der Quotienten-Differenzen-Algorithmus. Z. Angew. Math. Phys. **5**, 233–251 (1954)
26. Saff, E.B.: An extension of Montessus de Ballore theorem on the convergence of interpolating rational functions. J. Approx. Theory, **6**, 63–67 (1972)
27. Shanks, D.: Nonlinear transformations of divergent and slowly convergent sequences. J. Math. Phys. **34**, 1–42 (1955)
28. Sidi, A.: On a generalization of the Richardson extrapolation process. Numer. Math. **57**, 365–377 (1990)
29. Sidi, A.: Quantitative and constructive aspects of the generalized Koenig's and de Montessus's theorems for Padé approximants. J. Comput. Appl. Math. **29**, 257–291 (1990)
30. Sidi, A.: Acceleration of convergence of (generalized) Fourier series by the d-transformation. Ann. Numer. Math. **2**, 381–406 (1995)
31. Sidi, A.: Extension and completion of Wynn's theory on convergence of columns of the epsilon table. J. Approx. Theory **86**, 21–40 (1996)
32. Sidi, A.: Further results on convergence and stability of a generalization of the Richardson extrapolation process. BIT Numer. Math. **36**, 143–157 (1996)
33. Sidi, A.: The Richardson extrapolation process with a harmonic sequence of collocation points. SIAM J. Numer. Anal. **37**, 1729–1746 (2000)
34. Sidi, A.: Practical Extrapolation Methods: Theory and Applications. In: Cambridge Monographs on Applied and Computational Mathematics, vol. 10. Cambridge University Press, Cambridge (2003)
35. Sidi, A.: Asymptotic expansions of Legendre series coefficients for functions with endpoint singularities. Asymptot. Anal. **65**, 175–190 (2009)
36. Sidi, A.: Asymptotic analysis of a generalized Richardson extrapolation process on linear sequences. Math. Comput. **79**, 1681–1695 (2010)
37. Sidi, A.: A simple approach to asymptotic expansions for Fourier integrals of singular functions. Appl. Math. Comput. **216**, 3378–3385 (2010)
38. Sidi, A.: Survey of numerical stability issues in convergence acceleration. Appl. Numer. Math. **60**, 1395–1410 (2010)
39. Sidi, A.: Asymptotic expansions of Legendre series coefficients for functions with interior and endpoint singularities. Math. Comput. **80**, 1663–1684 (2011)
40. Sidi, A., Ford, W.F., Smith D.A.: Acceleration of convergence of vector sequences. SIAM J. Numer. Anal. **23**, 178–196 (1986). Originally appeared as NASA TP-2193 (1983)
41. Smith, D.A., Ford, W.F.: Acceleration of linear and logarithmic convergence. SIAM J. Numer. Anal. **16**, 223–240 (1979)
42. Smith, D.A., Ford, W.F.: Numerical comparisons of nonlinear convergence accelerators. Math. Comput. **38**, 481–499 (1982)
43. Stoer, J., Bulirsch, R.: Introduction to Numerical Analysis, 3rd edn. Springer, New York (2002)
44. Wynn, P.: On a device for computing the $e_m(S_n)$ transformation. Math. Tables Aids Comput. **10**, 91–96 (1956)
45. Wynn, P.: On the convergence and stability of the epsilon algorithm. SIAM J. Numer. Anal. **3**, 91–122 (1966)
46. Wynn, P.: Transformations to accelerate the convergence of Fourier series. In: Gertrude Blanche Anniversary Volume, pp. 339–379. Wright Patterson Air Force Base (1967)