This is a postprint version of the following published document:

On the convergence of type I Hermite–Padé apporximants approximants [*]

G. López Lagomasino [*], S. Medina Peralta

Departamento de Matemáticas, Universidad Carlos III de Madrid, Avda. Universidad 30, 28911 Leganés, Madrid, Spain

Abstrac

Padé approximation has two natural extensions to vector rational approximation through the so-called type I and type II Hermite–Padé approximants. The convergence proper-ties of type II Hermite–Padé approximants have been studied. For such approximants Markov and Stieltjes type theorems are available. To the present, such results have not been obtained for type I approximants. In this paper, we provide Markov and Stieltjes type theorems on the convergence of type I Hermite–Padé approximants for Nikishin systems of functions.

Keywords:

Multiple orthogonal polynomials; Nikishin systems; Type I Hermite–Padé approximants

1. Introduction

Let s be a finite Borel measure with constant (not necessarily positive) sign whose support $\operatorname{supp}(s)$ contains infinitely many points and is contained in the real line $\mathbb{R}$. If $\operatorname{supp}(s)$ is an unbounded set we assume additionally that $x^n \in L_1(s)$, $n \in \mathbb{N}$. By

---

[*] Corresponding author.
  E-mail address: lago@math.uc3m.es (G. López Lagomasino).

$\Delta = \mathrm{Co}(\mathrm{supp}(s))$ we denote the smallest interval which contains $\mathrm{supp}(s)$. We denote this class of measures by $\mathcal{M}(\Delta)$. Let

$$\widehat{s}(z) = \int \frac{ds(x)}{z - x}$$

be the Cauchy transform of s.

Given any positive integer $n \in \mathbb{N}$ there exist polynomials Q_n, P_n satisfying:

- $\deg Q_n \leq n$, $\deg P_n \leq n - 1$, $Q_n \not\equiv 0$,
- $(Q_n \widehat{s} - P_n)(z) = \mathcal{O}(1/z^{n+1})$, $z \to \infty$.

The ratio $\pi_n = P_n/Q_n$ of any two such polynomials defines a unique rational function called the nth term of the diagonal sequence of Padé approximants to $\widehat{s}$. Cauchy transforms of measures are important: for example, many elementary functions may be expressed through them, the resolvent function of a bounded selfadjoint operator adopts that form, and they characterize all functions holomorphic in the upper half plane whose image lies in the lower half plane and can be extended continuously to the complement of a finite segment $[a, b]$ of the real line taking negative values for $z < a$ and positive values for $z > b$ (then $\mathrm{supp}(s) \subset [a, b]$), see [13, Theorem A.6]. Providing efficient methods for their approximation is a central question in the theory of rational approximation.

When Δ is bounded, A.A. Markov proved in [15] (in the context of the theory of continued fractions) that

$$\lim_{n \to \infty} \pi_n(z) = \widehat{s}(z) \tag{1.1}$$

uniformly on each compact subset of $\overline{\mathbb{C}} \setminus \Delta$. It is easy to deduce that this limit takes place with geometric rate. When Δ is a half line, T.J. Stieltjes in [22] showed that (1.1) takes place if and only if the moment problem for the sequence $(c_n)_{n \geq 0}$, $c_n = \int x^n ds(x)$, is determinate. It is well known that the moment problem for measures of bounded support is determinate; therefore, Stieltjes' theorem contains Markov's result. In [2], T. Carleman proved when $\Delta \subset \mathbb{R}_+$ that

$$\sum_{n \geq 1} |c_n|^{-1/2n} = \infty \tag{1.2}$$

is sufficient for the moment problem to be determinate. For an arbitrary measure $s \in \mathcal{M}(\Delta)$, where Δ is contained in a half line, we say that it satisfies Carleman's condition if after an affine transformation which takes Δ into $\mathbb{R}_+$ the image measure satisfies Carleman's condition.

Padé approximation has two natural extensions to the case of vector rational approximation. These extensions were introduced by Hermite in order to study the transcendency of e. Other applications in number theory have been obtained. See [23] for

a survey of results in this direction. Recently, these approximants and their associated Hermite–Padé polynomials have appeared in a natural way in certain models coming from probability theory and mathematical physics. A summary of this type of applications can be found in [14].

Given a system of finite Borel measures $S = (s_1, \ldots, s_m)$ with constant sign and a multi-index $\mathbf{n} = (n_1, \ldots, n_m) \in \mathbb{Z}_+^m \setminus \{\mathbf{0}\}$, $|\mathbf{n}| = n_1 + \cdots + n_m$, where $\mathbb{Z}_+$ denotes the set of non-negative integers and $\mathbf{0}$ the m-dimensional zero vector, there exist polynomials $a_{\mathbf{n},j}$, $j = 0, \ldots, m$, not all identically equal to zero, such that:

- $\deg a_{\mathbf{n},j} \leq n_j - 1$, $j = 1, \ldots, m$, $\deg a_{\mathbf{n},0} \leq \max(n_j) - 2$,
- $a_{\mathbf{n},0}(z) + \sum_{j=1}^m a_{\mathbf{n},j}(z) \widehat{s}_j(z) = \mathcal{O}(1/z^{|\mathbf{n}|})$, $z \to \infty$ ($\deg a_{\mathbf{n},j} \leq -1$ means that $a_{\mathbf{n},j} \equiv 0$).

Analogously, there exist polynomials $Q_{\mathbf{n}}$, $P_{\mathbf{n},j}$, $j = 1, \ldots, m$, satisfying:

- $\deg Q_{\mathbf{n}} \leq |\mathbf{n}|$, $Q_{\mathbf{n}} \not\equiv 0$, $\deg P_{\mathbf{n},j} \leq |\mathbf{n}| - 1$, $j = 1, \ldots, m$,
- $Q_{\mathbf{n}}(z) \widehat{s}_j(z) - P_{\mathbf{n},j}(z) = \mathcal{O}(1/z^{n_j+1})$, $z \to \infty$, $j = 1, \ldots, m$.

Traditionally, the systems of polynomials $(a_{\mathbf{n},0}, \ldots, a_{\mathbf{n},m})$ and $(Q_{\mathbf{n}}, P_{\mathbf{n},1}, \ldots, P_{\mathbf{n},m})$ have been called type I and type II Hermite–Padé approximants (polynomials) of $(\widehat{s}_1, \ldots, \widehat{s}_m)$, respectively. When $m = 1$ both definitions reduce to that of classical Padé approximation.

From the definition, type II Hermite–Padé approximation is easy to view as an approximating scheme of the vector function $(\widehat{s}_1, \ldots, \widehat{s}_m)$ by considering a sequence of vector rational functions of the form $(P_{\mathbf{n},1}/Q_{\mathbf{n}}, \ldots, P_{\mathbf{n},m}/Q_{\mathbf{n}})$, $\mathbf{n} \in \Lambda \subset \mathbb{Z}_+^m$, where $Q_{\mathbf{n}}$ is a common denominator for all components. Regarding type I, it is not obvious what is the object to be approximated or even what should be considered as the approximant. Our goal is to clarify these questions providing straightforward analogues of the Markov and Stieltjes theorems.

Before stating our main result, let us introduce what is called a Nikishin system of measures to which we will restrict our study. Let Δ_α, Δ_β be two intervals contained in the real line which have at most one point in common, $\sigma_\alpha \in \mathcal{M}(\Delta_\alpha)$, $\sigma_\beta \in \mathcal{M}(\Delta_\beta)$, and $\widehat{\sigma}_\beta \in L_1(\sigma_\alpha)$. With these two measures we define a third one as follows (using the differential notation)

$$d\langle \sigma_\alpha, \sigma_\beta \rangle(x) := \widehat{\sigma}_\beta(x) d\sigma_\alpha(x).$$

Above, $\widehat{\sigma}_\beta$ denotes the Cauchy transform of the measure σ_β. The more appropriate notation $\widehat{\sigma_\beta}$ causes space consumption and aesthetic inconveniences. We need to take consecutive products of measures; for example,

$$\langle \sigma_\gamma, \sigma_\alpha, \sigma_\beta \rangle := \langle \sigma_\gamma, \langle \sigma_\alpha, \sigma_\beta \rangle \rangle.$$

Here, we assume not only that $\widehat{\sigma}_\beta \in L_1(\sigma_\alpha)$ but also $\widehat{\langle \sigma_\alpha, \sigma_\beta \rangle} \in L_1(\sigma_\gamma)$ where $\widehat{\langle \sigma_\alpha, \sigma_\beta \rangle}$ denotes the Cauchy transform of $\langle \sigma_\alpha, \sigma_\beta \rangle$. Inductively, one defines products of a finite number of measures.

Definition 1.1. Take a collection Δ_j, $j = 1, \ldots, m$, of closed intervals contained in the real line such that, for $j = 1, \ldots, m-1$

$$\Delta_j \cap \Delta_{j+1} = \emptyset, \quad \text{or} \quad \Delta_j \cap \Delta_{j+1} = \{x_{j,j+1}\},$$

where $x_{j,j+1}$ is a single point. Let $(\sigma_1, \ldots, \sigma_m)$ be a system of measures such that $\mathrm{Co}(\mathrm{supp}(\sigma_j)) = \Delta_j$, $\sigma_j \in \mathcal{M}(\Delta_j)$, $j = 1, \ldots, m$, and

$$\langle \sigma_j, \ldots, \sigma_k \rangle := \langle \sigma_j, \langle \sigma_{j+1}, \ldots, \sigma_k \rangle \rangle \in \mathcal{M}(\Delta_j), \quad 1 \leq j < k \leq m. \tag{1.3}$$

When $\Delta_j \cap \Delta_{j+1} = \{x_{j,j+1}\}$ we also assume that $x_{j,j+1}$ is not a mass point of either σ_j or σ_{j+1}. We say that $(s_1, \ldots, s_m) = \mathcal{N}(\sigma_1, \ldots, \sigma_m)$, where

$$s_1 = \sigma_1, \quad s_2 = \langle \sigma_1, \sigma_2 \rangle, \quad \ldots, \quad s_m = \langle \sigma_1, \sigma_2, \ldots, \sigma_m \rangle$$

is the Nikishin system of measures generated by $(\sigma_1, \ldots, \sigma_m)$.

Initially, E.M. Nikishin in [16] restricted himself to measures with bounded support and no intersection points between consecutive Δ_j. Definition 1.1 includes interesting examples which appear in practice (see, [10, Subsection 1.4]). We follow the approach of [10, Definition 1.2] assuming additionally the existence of all the moments of the generating measures. This is done only for the purpose of simplifying the presentation without affecting too much the generality. However, we wish to point out that the results of this paper have appropriate formulations with the definition given in [10] of a Nikishin system.

When $m = 2$, for multi-indices of the form $\mathbf{n} = (n, n)$ E.M. Nikishin proved in [16] that

$$\lim_{n \to \infty} \frac{P_{\mathbf{n},j}(z)}{Q_{\mathbf{n}}(z)} = \widehat{s}_j(z), \quad j = 1, 2,$$

uniformly on each compact subset of $\overline{\mathbb{C}} \setminus \Delta_1$. In [1] this result was extended to any Nikishin system of m measures including generating measures with unbounded support. The convergence for more general sequences of multi-indices was treated in [6,7] and [12].

In [10, Lemma 2.9] it was shown that if $(\sigma_1, \ldots, \sigma_m)$ is a generator of a Nikishin system then $(\sigma_m, \ldots, \sigma_1)$ is also a generator (as well as any subsystem of consecutive measures drawn from them). When the supports are bounded and consecutive supports do not intersect this is trivially true. In the following, for $1 \leq j \leq k \leq m$ we denote

$$s_{j,k} := \langle \sigma_j, \sigma_{j+1}, \ldots, \sigma_k \rangle, \qquad s_{k,j} := \langle \sigma_k, \sigma_{k-1}, \ldots, \sigma_j \rangle.$$

To state our main results, the natural framework is that of multi-point type I Hermite–Padé approximation.

Definition 1.2. Let $(s_{1,1}, \ldots, s_{1,m}) = \mathcal{N}(\sigma_1, \ldots, \sigma_m)$, $\mathbf{n} = (n_1, \ldots, n_m) \in \mathbb{Z}_+^m \setminus \{\mathbf{0}\}$, and $w_{\mathbf{n}}$, $\deg w_{\mathbf{n}} \leq |\mathbf{n}| + \max(n_j) - 2$, a polynomial with real coefficients whose zeros lie in $\mathbb{C} \setminus \Delta_1$, be given. We say that $(a_{\mathbf{n},0}, \ldots, a_{\mathbf{n},m})$ is a type I multi-point Hermite–Padé approximation of $(\widehat{s}_{1,1}, \ldots, \widehat{s}_{1,m})$ with respect to $w_{\mathbf{n}}$ if:

i) $\deg a_{\mathbf{n},j} \leq n_j - 1$, $j = 1, \ldots, m$, $\deg a_{\mathbf{n},0} \leq n_0 - 1$, $n_0 := \max_{j=1,\ldots,m}(n_j) - 1$, not all identically equal to 0 ($n_j = 0$ implies that $a_{\mathbf{n},j} \equiv 0$),

ii) $\mathcal{A}_{\mathbf{n},0}/w_{\mathbf{n}} \in \mathcal{H}(\mathbb{C} \setminus \Delta_1)$ and $\mathcal{A}_{\mathbf{n},0}(z)/w_{\mathbf{n}}(z) = \mathcal{O}(1/z^{|\mathbf{n}|})$, $z \to \infty$, where

$$\mathcal{A}_{\mathbf{n},j} := a_{\mathbf{n},j} + \sum_{k=j+1}^{m} a_{\mathbf{n},k}\widehat{s}_{j+1,k}(z), \quad j = 0, \ldots, m-1, \quad \mathcal{A}_{\mathbf{n},m} := a_{\mathbf{n},m}.$$

If $\deg w_{\mathbf{n}} = |\mathbf{n}| + \max(n_j) - 2$ the second part of ii) is automatically fulfilled. Should $\deg w_{\mathbf{n}} = N < |\mathbf{n}| + \max(n_j) - 2$ then $|\mathbf{n}| + \max(n_j) - 2 - N$ (asymptotic) interpolation conditions are imposed at ∞. In general $|\mathbf{n}| + \max(n_j) - 2$ interpolation conditions are imposed at points in $(\mathbb{C} \setminus \Delta_1) \cup \{\infty\}$. The total number of free parameters (the coefficients of the polynomials $a_{\mathbf{n},j}$, $j = 0, \ldots, m$) equals $|\mathbf{n}| + \max(n_j) - 1$; therefore, the homogeneous linear system of equations to be solved in order that i)–ii) take place always has a non-trivial solution. Notice that when $w_{\mathbf{n}} \equiv 1$ we recover the definition given above for classical type I Hermite–Padé approximation.

An analogous definition can be given for type II multi-point Hermite–Padé approximants but we will not dwell into this. Algebraic and analytic properties regarding uniqueness, integral representations, asymptotic behavior, and orthogonality conditions satisfied by type I and type II Hermite–Padé approximants have been studied, for example, in [1,3–7,9,10,12,17], and [18, Chapter 4], which include the case of multi-point approximation.

Let $\overset{\circ}{\Delta}$ denote the interior of Δ with the Euclidean topology of the real line. We have

Theorem 1.3. *Let $(s_{1,1}, \ldots, s_{1,m}) = \mathcal{N}(\sigma_1, \ldots, \sigma_m)$, $\mathbf{n} = (n_1, \ldots, n_m) \in \mathbb{Z}_+^m \setminus \{\mathbf{0}\}$, and $w_{\mathbf{n}}$, $\deg w_{\mathbf{n}} \leq |\mathbf{n}| + \max(n_j) - 2$, a polynomial with real coefficients whose zeros lie in $\mathbb{C} \setminus \Delta_1$, be given. The type I multi-point Hermite–Padé approximation $(a_{\mathbf{n},0}, \ldots, a_{\mathbf{n},m})$ of $(\widehat{s}_{1,1}, \ldots, \widehat{s}_{1,m})$ with respect to $w_{\mathbf{n}}$ is uniquely determined except for a constant factor, and $\deg a_{\mathbf{n},j} = n_j - 1$, $j = 0, \ldots, m$. Moreover*

$$\int x^\nu \mathcal{A}_{\mathbf{n},1}(x)\frac{d\sigma_1(x)}{w_{\mathbf{n}}(x)} = 0, \quad \nu = 0, \ldots, |\mathbf{n}| - 2, \tag{1.4}$$

which implies that $\mathcal{A}_{\mathbf{n},1}$ has exactly $|\mathbf{n}| - 1$ simple zeros in $\overset{\circ}{\Delta}_1$ and no other zeros in $\mathbb{C} \setminus \Delta_2$. Additionally,

$$\frac{\mathcal{A}_{\mathbf{n},0}(z)}{w_{\mathbf{n}}(z)} = \int \frac{\mathcal{A}_{\mathbf{n},1}(x)d\sigma_1(x)}{w_{\mathbf{n}}(x)(z-x)} \tag{1.5}$$

and

$$a_{\mathbf{n},0}(z) = -\int \frac{\sum_{j=1}^{m}(w_{\mathbf{n}}(x)a_{\mathbf{n},j}(z) - w_{\mathbf{n}}(z)a_{\mathbf{n},j}(x))ds_{1,j}(x)}{(z-x)w_{\mathbf{n}}(x)}. \tag{1.6}$$

Notice that nothing has been said about the location of the zeros of the polynomials $a_{\mathbf{n},j}$. For special sequences of multi-indices this information can be deduced from the convergence of type I Hermite–Padé approximants. We have the following result (see also Lemma 3.1).

Theorem 1.4. *Let $S = (s_{1,1}, \ldots, s_{1,m}) = \mathcal{N}(\sigma_1, \ldots, \sigma_m)$, $\Lambda \subset \mathbb{Z}_+^m$ an infinite sequence of distinct multi-indices, and $(w_{\mathbf{n}})_{\mathbf{n}\in\Lambda}, \deg w_{\mathbf{n}} \le |\mathbf{n}|+\max(n_j)-2$, a sequence of polynomials with real coefficients whose zeros lie in $\mathbb{C} \setminus \Delta_1$, be given. Consider the corresponding sequence $(a_{\mathbf{n},0}, \ldots, a_{\mathbf{n},m})$, $\mathbf{n} \in \Lambda$, of type I multi-point Hermite–Padé approximants of S with respect to $(w_{\mathbf{n}})_{\mathbf{n}\in\Lambda}$. Assume that*

$$\sup_{\mathbf{n}\in\Lambda}\left(\max_{j=1,\ldots,m}(n_j) - \min_{k=1,\ldots,m}(n_k)\right) \le C < \infty, \tag{1.7}$$

and that either Δ_{m-1} is bounded away from Δ_m or σ_m satisfies Carleman's condition. Then,

$$\lim_{\mathbf{n}\in\Lambda} \frac{a_{\mathbf{n},j}}{a_{\mathbf{n},m}} = (-1)^{m-j}\widehat{s}_{m,j+1}, \quad j = 0, \ldots, m-1, \tag{1.8}$$

uniformly on each compact subset $\mathcal{K} \subset \mathbb{C} \setminus \Delta_m$. The accumulation points of sequences of zeros of the polynomials $a_{\mathbf{n},j}$, $j = 0, \ldots, m$, $\mathbf{n} \in \Lambda$ are contained in $\Delta_m \cup \{\infty\}$. Additionally,

$$\lim_{\mathbf{n}\in\Lambda} \frac{\mathcal{A}_{\mathbf{n},j}}{a_{\mathbf{n},m}} = 0, \quad j = 0, \ldots, m-1, \tag{1.9}$$

uniformly on each compact subset $\mathcal{K} \subset \mathbb{C} \setminus (\Delta_{j+1} \cup \Delta_m)$.

We wish to underline that Theorem 1.4 requires no special analytic property from the generating measures of the Nikishin system except for Carleman's condition on σ_m.

Notice that the sequences of rational functions $(a_{\mathbf{n},j}/a_{\mathbf{n},m})$, $\mathbf{n} \in \Lambda$, $j = 0, \ldots, m-1$, allow to recover the Cauchy transforms of the measures in the Nikishin system $\mathcal{N}(\sigma_m, \ldots, \sigma_1)$ in contrast with the sequences $(P_{\mathbf{n},j}/Q_{\mathbf{n}})$, $\mathbf{n} \in \Lambda$, $j = 1, \ldots, m$, of type II multi-point Hermite–Padé approximants which recover the Cauchy transforms of the measures in $\mathcal{N}(\sigma_1, \ldots, \sigma_m)$.

In the process of writing this paper, S.P. Suetin sent us [19] and [20]. The first one of these papers announces the results contained in the second one. Those papers deal with

the study of type I Hermite–Padé approximants for an interesting class of systems of two functions ($m = 2$) which form a generalized Nikishin system in the sense that the second generating measure lives on a symmetric (with respect to the real line) compact set made up of finitely many analytic arcs which does not separate the complex plane and may be complex valued. The authors obtain the logarithmic asymptotic of the sequences of Hermite–Padé polynomials $a_{\mathbf{n},j}$, $j = 1, 2$, and an analogue of (1.8) for $j = 1$. Convergence is proved in capacity (see [19, Theorem 1] and [20, Corollary 1]).

For the proof of Theorem 1.4 we need a convenient representation of the reciprocal of the Cauchy transform of a measure. It is known that for each $\sigma \in \mathcal{M}(\Delta)$, where Δ is contained in a half line, there exists a measure $\tau \in \mathcal{M}(\Delta)$ and $\ell(z) = az + b$, $a = 1/|\sigma|$, $b \in \mathbb{R}$, such that

$$1/\widehat{\sigma}(z) = \ell(z) + \widehat{\tau}(z), \tag{1.10}$$

where $|\sigma|$ is the total variation of σ. See [13, Appendix] and [21, Theorem 6.3.5] for measures with compact support, and [10, Lemma 2.3] when the support is contained in a half line.

We call τ the inverse measure of σ. Such measures appear frequently in our reasonings, so we will fix a notation to distinguish them. In relation with measures denoted with s they will carry over to them the corresponding sub-indices. The same goes for the polynomials ℓ. For example,

$$1/\widehat{s}_{j,k}(z) = \ell_{j,k}(z) + \widehat{\tau}_{j,k}(z).$$

We also write

$$1/\widehat{\sigma}_\alpha(z) = \ell_\alpha(z) + \widehat{\tau}_\alpha(z).$$

The following result has independent interest and will be used in combination with Lemma 2.4 below in the proof of Theorem 1.4.

Theorem 1.5. *Let $(s_{1,1}, s_{1,2}) = \mathcal{N}(\sigma_1, \sigma_2)$. If σ_1 satisfies Carleman's condition so do $s_{1,2}$ and τ_1.*

This paper is organized as follows. In Section 2 we prove Theorems 1.3 and 1.5. We also present some notions and results necessary for the proof of Theorem 1.4. Section 3 contains the proof of Theorem 1.4 and some extensions of the main result to sequences of multi-indices satisfying conditions weaker than (1.7), estimates of the rate of convergence in (1.8)–(1.9) for the case when Δ_m or Δ_{m-1} is bounded and $\Delta_m \cap \Delta_{m-1} = \emptyset$, and applications to other simultaneous approximation schemes.

2. Proof of Theorem 1.3 and auxiliary results

We begin with a lemma which allows to give an integral representation of the remainder of type I multi-point Hermite–Padé approximants.

Lemma 2.1. *Let $(s_{1,1}, \ldots, s_{1,m}) = \mathcal{N}(\sigma_1, \ldots, \sigma_m)$ be given. Assume that there exist polynomials with real coefficients $a_0, \ldots, a_m$ and a polynomial w with real coefficients whose zeros lie in $\mathbb{C} \setminus \Delta_1$ such that*

$$\frac{\mathcal{A}(z)}{w(z)} \in \mathcal{H}(\mathbb{C} \setminus \Delta_1) \quad and \quad \frac{\mathcal{A}(z)}{w(z)} = \mathcal{O}\left(\frac{1}{z^N}\right), \quad z \to \infty,$$

where $\mathcal{A} := a_0 + \sum_{k=1}^m a_k \widehat{s}_{1,k}$ and $N \geq 1$. Let $\mathcal{A}_1 := a_1 + \sum_{k=2}^m a_k \widehat{s}_{2,k}$. Then

$$\frac{\mathcal{A}(z)}{w(z)} = \int \frac{\mathcal{A}_1(x)}{(z-x)} \frac{d\sigma_1(x)}{w(x)}. \tag{2.1}$$

If $N \geq 2$, we also have

$$\int x^\nu \mathcal{A}_1(x) \frac{d\sigma_1(x)}{w(x)}, \quad \nu = 0, \ldots, N-2. \tag{2.2}$$

In particular, $\mathcal{A}_1$ has at least $N-1$ sign changes in $\overset{\circ}{\Delta}_1$.

Proof. We have

$$\mathcal{A}(z) = a_0(z) + \sum_{k=1}^m a_k(z)\widehat{s}_{1,k}(z) \mp w(z) \int \frac{\mathcal{A}_1(x)}{(z-x)} \frac{d\sigma_1(x)}{w(x)}$$

$$= a_0(z) + \int \frac{\sum_{k=1}^m (w(x)a_k(z) - w(z)a_k(x))ds_{1,k}(x)}{(z-x)w(x)} + w(z) \int \frac{\mathcal{A}_1(x)}{(z-x)} \frac{d\sigma_1(x)}{w(x)}.$$

For each $k = 1, \ldots, m$

$$\bigl(w(x)a_k(z) - w(z)a_k(x)\bigr)/(z-x)$$

is a polynomial in z. Therefore,

$$P(z) := a_0(z) + \int \frac{\sum_{k=1}^m (w(x)a_k(z) - w(z)a_k(x))ds_{1,k}(x)}{(z-x)w(x)}$$

represents a polynomial. Consequently

$$\mathcal{A}(z) = P(z) + w(z) \int \frac{\mathcal{A}_1(x)d\sigma_1(x)}{(z-x)w(x)} = w(z)\mathcal{O}\bigl(1/z^N\bigr), \quad z \to \infty.$$

These equalities imply that

$$P(z) = w(z)\mathcal{O}(1/z), \quad z \to \infty.$$

Therefore, $\deg P < \deg w$ and is equal to zero at all the zeros of w. Hence $P \equiv 0$. (Should w be a constant polynomial likewise we get that $P \equiv 0$.) Thus, we have proved (2.1).

From our assumptions and (2.1), it follows that

$$\frac{\mathcal{A}(z)}{w(z)} = \int \frac{\mathcal{A}_1(x)}{(z-x)} \frac{d\sigma_1(x)}{w(x)} = \mathcal{O}\big(1/z^N\big), \quad z \to \infty.$$

Suppose that $N \geq 2$. We have the asymptotic expansion

$$\int \frac{\mathcal{A}_1(x)}{(z-x)} \frac{d\sigma_1(x)}{w(x)} = \sum_{\nu=0}^{N-2} \frac{d_\nu}{z^{\nu+1}} + \int \frac{x^{N-1}\mathcal{A}_1(x)}{z^{N-1}(z-x)} \frac{d\sigma_1(x)}{w(x)}$$

$$= \sum_{\nu=0}^{N-2} \frac{d_\nu}{z^{N+1}} + \mathcal{O}\big(1/z^N\big), \quad z \to \infty,$$

where

$$d_\nu = \int x^\nu \mathcal{A}_1(x) \frac{d\sigma_1(x)}{w(x)}, \quad \nu = 0, \dots, N-2.$$

Therefore,

$$d_\nu = 0, \quad \nu = 0, \dots, N-2,$$

which is (2.2).

Suppose that $\mathcal{A}_1$ has at most $\widetilde{N} < N-1$ sign changes in $\overset{\circ}{\Delta}_1$ at the points $x_1, \dots, x_N$. Take $q(x) = \prod_{k=1}^{\widetilde{N}}(x - x_k)$. According to (2.2)

$$\int q(x)\mathcal{A}_1(x) \frac{d\sigma_1(x)}{w(x)} = 0$$

which is absurd because $q(a_1 + \sum_{k=2}^m a_k \widehat{s}_{2,k})/w$ has constant sign in $\overset{\circ}{\Delta}_1$ and σ_1 is a measure with constant sign in $\overset{\circ}{\Delta}_1$ whose support contains infinitely many points. Thus, the number of sign changes must be greater or equal to $N-1$ as claimed. $\square$

In [10, Lemma 2.10], several formulas involving ratios of Cauchy transforms were proved. The most useful ones for this paper establish that

$$\frac{\widehat{s}_{1,k}}{\widehat{s}_{1,1}} = \frac{|s_{1,k}|}{|s_{1,1}|} - \big\langle \tau_{1,1}, \langle s_{2,k}, \sigma_1 \rangle \big\rangle^\wedge, \quad 1 = j < k \leq m. \tag{2.3}$$

(Recall that $|\sigma|$ denotes the total variation of a measure σ.)

We are ready for the

Proof of Theorem 1.3. Let $(a_{\mathbf{n},0}, \ldots, a_{\mathbf{n},m})$ be a type I multi-point Hermite–Padé approximation of $(\widehat{s}_{1,1}, \ldots, \widehat{s}_{1,m})$ with respect to $w_{\mathbf{n}}$. From Definition 1.2, formulas (1.4) and (1.5) follow directly from (2.2) and (2.1), respectively. Relation (1.6) is obtained from (1.5) solving for $a_{\mathbf{n},0}$.

In the proof of Lemma 2.1 we saw that (1.4) implies that $\mathcal{A}_{\mathbf{n},1}$ has at least $|\mathbf{n}| - 1$ sign changes in $\overset{\circ}{\Delta}_1$. We have that $(s_{2,2}, \ldots, s_{2,m}) = \mathcal{N}(\sigma_2, \ldots, \sigma_m)$ forms a Nikishin system. According to [10, Theorem 1.1], $\mathcal{A}_{\mathbf{n},1}$ can have at most $|\mathbf{n}| - 1$ zeros in $\mathbb{C} \setminus \Delta_2$. Taking account of what we proved previously, it follows that $\mathcal{A}_{\mathbf{n},1}$ has exactly $|\mathbf{n}| - 1$ simple zeros in $\overset{\circ}{\Delta}_1$ and it has no other zero in $\mathbb{C} \setminus \Delta_2$. This is true for any $\mathbf{n} \in \mathbb{Z}_+^m \setminus \{\mathbf{0}\}$.

Suppose that for some $\mathbf{n} \in \mathbb{Z}_+^m \setminus \{\mathbf{0}\}$ and some $j \in \{1, \ldots, m\}$, we have that $\deg a_{\mathbf{n},j} = \widetilde{n}_j - 1 < n_j - 1$. Then, according to [10, Theorem 1.1] $\mathcal{A}_{\mathbf{n},1}$ could have at most $|\mathbf{n}| - n_j + \widetilde{n}_j - 1 \le |\mathbf{n}| - 2$ zeros in $\mathbb{C} \setminus \Delta_2$. This is absurd because we have proved that it has $|\mathbf{n}| - 1$ zeros in $\overset{\circ}{\Delta}_1$.

Now, suppose that for some $\mathbf{n} \in \mathbb{Z}_+^m \setminus \{\mathbf{0}\}$, there exist two non-collinear type I multi-point Padé approximants $(a_{\mathbf{n},0}, \ldots, a_{\mathbf{n},m})$ and $(\widetilde{a}_{\mathbf{n},0}, \ldots, \widetilde{a}_{\mathbf{n},m})$ of $(\widehat{s}_{1,1}, \ldots, \widehat{s}_{1,m})$ with respect to $w_{\mathbf{n}}$. From (1.6) it follows that $(a_{\mathbf{n},1}, \ldots, a_{\mathbf{n},m})$ and $(\widetilde{a}_{\mathbf{n},1}, \ldots, \widetilde{a}_{\mathbf{n},m})$ are not collinear. We know that $\deg a_{\mathbf{n},j} = \deg \widetilde{a}_{\mathbf{n},j} = n_j - 1$, $j = 1, \ldots, m$. Consequently, there exists some constant C such that $(a_{\mathbf{n},1} - C\widetilde{a}_{\mathbf{n},1}, \ldots, a_{\mathbf{n},m} - C\widetilde{a}_{\mathbf{n},m}) \ne \mathbf{0}$ and $\deg(a_{\mathbf{n},j} - C\widetilde{a}_{\mathbf{n},j}) < n_j - 1$ for some $j \in \{1, \ldots, m\}$. By linearity, $(a_{\mathbf{n},0} - C\widetilde{a}_{\mathbf{n},0}, \ldots, a_{\mathbf{n},m} - C\widetilde{a}_{\mathbf{n},m})$ is a multi-point type I Hermite–Padé approximant of $(\widehat{s}_{1,1}, \ldots, \widehat{s}_{1,m})$ with respect to $w_{\mathbf{n}}$. This is not possible because $\deg(a_{\mathbf{n},j} - C\widetilde{a}_{\mathbf{n},j}) < n_j - 1$. Therefore, non-collinear solutions cannot exist.

We still need to show that $\deg a_{\mathbf{n}_0} = n_0 - 1$. To this end we need to transform $\mathcal{A}_{\mathbf{n},0}$. Let j be the first component of $\mathbf{n}$ such that $n_j = \max_{k=1,\ldots,m} n_k$. Since $n_0 = n_j - 1$, we have that either $j = 1$ or $n_0 \ge n_k$, $k = 1, \ldots, j - 1$. If $j = 1$, using (1.10) and (2.3) it follows that

$$\mathcal{B}_{\mathbf{n},0} := \frac{\mathcal{A}_{\mathbf{n},0}}{\widehat{s}_{1,1}} = \ell_{1,1} a_{\mathbf{n},0} + \sum_{k=1}^{m} \frac{|s_{1,k}|}{|s_{1,1}|} a_{\mathbf{n},k} + a_{\mathbf{n},0} \widehat{\tau}_{1,1} - \sum_{k=2}^{m} a_{\mathbf{n},k} \langle \tau_{1,1}, \langle s_{2,k}, \sigma_1 \rangle \widehat{\rangle},$$

where

$$\mathcal{B}_{\mathbf{n},0}/w_{\mathbf{n}} \in \mathcal{H}(\mathbb{C} \setminus \Delta_1), \qquad \mathcal{B}_{\mathbf{n},0}(z)/w_{\mathbf{n}}(z) = \mathcal{O}\big(1/z^{|\mathbf{n}|-1}\big), \quad z \to \infty.$$

Using Lemma 2.1 it follows that

$$\int x^\nu \mathcal{B}_{\mathbf{n},1}(x) \frac{d\tau_{1,1}(x)}{w_{\mathbf{n}}(x)}, \quad \nu = 0, \ldots, |\mathbf{n}| - 3,$$

where $\mathcal{B}_{\mathbf{n},1} = a_{\mathbf{n},0} - \sum_{k=2}^{m} a_{\mathbf{n},k} \langle \langle \sigma_2, \sigma_1 \rangle, \sigma_3, \ldots, \sigma_k \widehat{\rangle}$. Hence $\mathcal{B}_{\mathbf{n},1}$ has at least $|\mathbf{n}| - 2$ sign changes in $\overset{\circ}{\Delta}_1$. According to [10, Theorem 1.1] the linear form $\mathcal{B}_{\mathbf{n},1}$ has at most

$\deg a_{\mathbf{n},0} + n_2 + \cdots + n_m$ zeros in all of $\mathbb{C} \setminus \Delta_2$. Should $\deg a_{\mathbf{n},0} \leq n_0 - 2$, we would have that $\deg a_{\mathbf{n},0} + n_2 + \cdots + n_m \leq |\mathbf{n}| - 3$ which contradicts that $\mathcal{B}_{\mathbf{n},1}$ has at least $|\mathbf{n}| - 2$ zeros in $\overset{\circ}{\Delta}_1$. Thus, when $j = 1$ it is true that $\deg a_{\mathbf{n},0} = n_0 - 1$. In general, the proof is similar as we will see.

Suppose that j, as defined in the previous paragraph, is ≥ 2. Then, either $n_0 = n_k$, $k = 1, \ldots, j - 1$ or there exists $\check{\jmath} < j$ for which $n_0 = n_k$, $k = 1, \ldots, \check{\jmath} - 1$ and $n_0 > n_{\check{\jmath}}$. In the first case, applying [10, Lemma 2.12], we obtain that there exists a Nikishin system $(s_{1,1}^*, \ldots, s_{1,m}^*) = \mathcal{N}(\sigma_1^*, \ldots, \sigma_m^*)$, a multi-index $\mathbf{n}^* = (n_0^*, \ldots, n_m^*) \in \mathbb{Z}_+^{m+1}$ which is a permutation of $\mathbf{n}$ with $n_0^* = n_j$, and polynomials with real coefficients $a_{\mathbf{n},k}^*$, $\deg a_{\mathbf{n},k}^* \leq n_k^* - 1$, $k = 0, \ldots, m$, such that

$$\frac{\mathcal{A}_{\mathbf{n},0}}{\widehat{s}_{1,j}} = a_{\mathbf{n},0}^* + \sum_{k=1}^{m} a_{\mathbf{n},k}^* \cdot \widehat{s}_{1,k}^*$$

Due to the structure of the values of the components of the multi-index $a_{\mathbf{n},j}^* = (-1)^j a_{\mathbf{n},0}$ and $n_j^* = n_0$ (see formula (31) in [9]). We can proceed as before and find that $\deg a_{\mathbf{n},j}^* = n_j^* - 1$, $j = 1, \ldots, m$. In particular, $\deg a_{\mathbf{n},j} = n_0 - 1$. In the other case, [10, Lemma 2.12] gives that

$$\frac{\mathcal{A}_{\mathbf{n},0}}{\widehat{s}_{1,j}} = a_{\mathbf{n},0}^* + \sum_{k=1}^{m} a_{\mathbf{n},k}^*, \widehat{s}_{1,k}^*$$

where $a_{\mathbf{n}\check{\jmath}}^* = \pm a_{\mathbf{n},0} + C a_{\mathbf{n},\check{\jmath}}$, $C \neq 0$ is some constant, and $n_{\check{\jmath}}^* = n_0$ (see formula (31) in [9]). Repeating the arguments employed above, we obtain that $\deg a_{\mathbf{n},j}^* = n_j^* - 1$, $j = 1, \ldots, m$. In particular, $\deg a_{\mathbf{n},0} = n_0 - 1$, because we already know that $\deg a_{\mathbf{n},\check{\jmath}} = n_{\check{\jmath}} - 1 < n_0 - 1$. $\square$

Remark 2.2. We wish to point out that in the statement of [10, Theorem 1.1] there is a misprint on the last line where $\mathbb{C}$ should replace $\overline{\mathbb{C}}$. That is, it should refer to zeros at finite points. This can be checked looking at the statements of [10, Lemmas 2.1, 2.2] and the proof of [10, Theorem 1.1] itself.

The notion of convergence in Hausdorff content plays a central role in the proof of Theorem 1.4. Let B be a subset of the complex plane $\mathbb{C}$. By $\mathcal{U}(B)$ we denote the class of all coverings of B by at most a numerable set of disks. Set

$$h(B) = \inf \left\{ \sum_{i=1}^{\infty} |U_i| : \{U_i\} \in \mathcal{U}(B) \right\},$$

where $|U_i|$ stands for the radius of the disk U_i. The quantity $h(B)$ is called the 1-dimensional Hausdorff content of the set B.

Definition 2.3. Let $(\varphi_n)_{n\in\mathbb{N}}$ be a sequence of complex functions defined on a domain $D \subset \mathbb{C}$ and φ another function defined on D (the value ∞ is permitted). We say that $(\varphi_n)_{n\in\mathbb{N}}$ converges in Hausdorff content to the function φ inside D if for each compact subset $\mathcal{K}$ of D and for each $\varepsilon > 0$, we have

$$\lim_{n\to\infty} h\big\{z \in K : \big|\varphi_n(z) - \varphi(z)\big| > \varepsilon\big\} = 0$$

(by convention $\infty \pm \infty = \infty$). We denote this writing $h\text{-}\lim_{n\to\infty} \varphi_n = \varphi$ inside D.

To obtain Theorem 1.4 we first prove (1.8) with convergence in Hausdorff content in place of uniform convergence (see Lemma 3.1 below). We need the following notion.

Let $s \in \mathcal{M}(\Delta)$ where Δ is contained in a half line of the real axis. Fix an arbitrary $\kappa \geq -1$. Consider a sequence of polynomials $(w_n)_{n\in\Lambda}$, $\Lambda \subset \mathbb{Z}_+$, such that $\deg w_n = \kappa_n \leq 2n + \kappa + 1$, whose zeros lie in $\mathbb{R} \setminus \Delta$. Let $(R_n)_{n\in\Lambda}$ be a sequence of rational functions $R_n = p_n/q_n$ with real coefficients satisfying the following conditions for each $n \in \Lambda$:

a) $\deg p_n \leq n + \kappa$, $\deg q_n \leq n$, $q_n \not\equiv 0$,
b) $(q_n\widehat{s} - p_n)(z)/w_n = \mathcal{O}(1/z^{n+1-\ell}) \in \mathcal{H}(\mathbb{C} \setminus \Delta)$, $z \to \infty$, where $\ell \in \mathbb{Z}_+$ is fixed.

We say that $(R_n)_{n\in\Lambda}$ is a sequence of incomplete diagonal multi-point Padé approximants of $\widehat{s}$.

Notice that in this construction for each $n \in \Lambda$ the number of free parameters equals $2n + \kappa + 2$ whereas the number of homogeneous linear equations to be solved in order to find q_n and p_n is equal to $2n + \kappa - \ell + 1$. When $\ell = 0$ there is only one more parameter than equations and R_n is defined uniquely coinciding with a (near) diagonal multi-point Padé approximation. When $\ell \geq 1$ uniqueness is not guaranteed. This is the reason why we chose to call them incomplete diagonal multi-point Padé approximants of $\widehat{s}$.

For sequences of incomplete diagonal multi-point Padé approximants, the following Stieltjes type theorem was proved in [1, Lemma 2] in terms of convergence in Hausdorff content.

Lemma 2.4. *Let* $s \in \mathcal{M}(\Delta)$ *be given where* Δ *is contained in a half line. Assume that* $(R_n)_{n\in\Lambda}$ *satisfies* a)–b) *and either the number of zeros of* w_n *lying on a closed bounded segment of* $\mathbb{R} \setminus \Delta$ *tends to infinity as* $n \to \infty$, $n \in \Lambda$, *or* s *satisfies Carleman's condition. Then*

$$h\text{-}\lim_{n\in\Lambda} R_n = \widehat{s}, \quad \text{inside } \mathbb{C} \setminus \Delta.$$

We will need to use Lemma 2.4 for different measures and Theorem 1.5 comes in our aid.

Proof of Theorem 1.5. Without loss of generality, we can assume that $\Delta \subset \mathbb{R}_+$ and that σ_1 is positive. Let $(c_n)_{n\in\mathbb{Z}_+}$ and $(\widetilde{c}_n)_{n\in\mathbb{Z}_+}$ denote the sequences of moments of σ_1 and $s_{1,2}$, respectively. Since $\widehat{\sigma}_2$ has constant sign on $\mathbb{R}_+$, we have that

$$|\widetilde{c}_n| = \int x^n \big|\widehat{\sigma}_2(x)\big| d\sigma_1(x) \leq \int\limits_0^1 x^n \big|\widehat{\sigma}_2(x)\big| d\sigma_1(x) + \int\limits_1^\infty x^n \big|\widehat{\sigma}_2(x)\big| d\sigma_1(x) \leq |s_{1,2}| + Cc_n,$$

where $C = \max\{|\widehat{\sigma}_2(x)| : x \in [1, +\infty)\} < \infty$ because $\lim_{x \to \infty} \widehat{\sigma}_2(x) = 0$. Consequently,

$$\sum_{n \geq 1} |\widetilde{c}_n|^{-1/2n} \geq \sum_{n \geq 1} \big(|s_{1,2}| + Cc_n\big)^{-1/2n}$$

$$\geq \sum_{\{n : Cc_n < |s_{1,2}|\}} \big(2|s_{1,2}|\big)^{-1/2n} + \sum_{\{n : Cc_n \geq |s_{1,2}|\}} \big(2Cc_n\big)^{-1/2n}.$$

If the first sum after the last inequality contains infinitely many terms then that sum is already divergent. If it has finitely many terms then Carleman's condition for σ_1 guarantees that the second sum is divergent. Thus, $s_{1,2}$ satisfies Carleman's condition.

To prove the second part we need to express the moments $(d_n)_{n \in \mathbb{Z}_+}$ of τ_1 in terms of the moments of σ_1. In the proof of [10, Lemma 2.3] we showed that the moments $(d_n)_{n \in \mathbb{Z}_+}$ are finite (since all the moments of σ_1 are finite) and can be obtained solving the system of equations

$$1 = d_{-2}c_0$$
$$0 = d_{-2}c_1 + d_{-1}c_0$$
$$0 = d_{-2}c_2 + d_{-1}c_1 + d_0c_0$$
$$\vdots = \vdots$$
$$0 = d_{-2}c_{n+2} + d_{-1}c_{n+1} + \cdots + d_nc_0.$$

(The values of d_{-2} and d_{-1} turn out to be the coefficients a and b, respectively, of the polynomial ℓ_1 in the decomposition (1.10) of $1/\widehat{\sigma}_1$.) Read the paragraph after formula (9) in [10].

To find d_n we apply Cramer's rule and we get

$$d_n = (-1)^n \Omega_n / c_0^{n+3} \tag{2.4}$$

where c_0^{n+3} gives the value of the determinant of the system and

$$\Omega_n = \begin{vmatrix} c_1 & c_0 & 0 & \cdots \\ c_2 & c_1 & \ddots & \ddots \\ \vdots & \ddots & \ddots & \ddots \\ c_{n+2} & c_{n+1} & \cdots & c_1 \end{vmatrix}$$

is the determinant of a lower Hessenberg matrix of dimension $n+2$ with constant diagonal terms. The expansion of the determinant Ω_n has several characteristics:

- It has exactly 2^{n+1} non-zero terms.
- For each $n \geq 0$, the sum of the subindexes of each non-zero term equals $n + 2$ (if a factor is repeated its subindex is counted as many times as it is repeated).
- The number of factors in each term is equal to $n + 2$.

The last assertion is trivial. To calculate the number of non-zero terms notice that from the first row we can only choose 2 non-zero entries. Once this is done, from the second row we can only choose 2 non-zero entries, and so forth, until we get to the last row where we only have left one non-zero entry to choose.

Regarding the second assertion we use induction. When $n = 0$ it is obvious. Assume that each non-zero term in the expansion of Ω_n has the property that the sum of its subindexes equals $n + 2$ and let us show that each non-zero term in the expansion of Ω_{n+1} has the property that the sum of its subindexes equals $n+3$. Expanding Ω_{n+1} by its first row we have

$$\Omega_{n+1} = c_1 \Omega_n - c_0 \Omega_n^*,$$

where Ω_n^* is obtained substituting the first column of Ω_n by the column vector $(c_2, \ldots, c_{n+3})^t$ (the superscript t means taking transpose). Using the induction hypothesis it easily follows that in each term arising from $c_1 \Omega_n$ and $c_0 \Omega_n^*$ the sum of its subindexes must equal $n + 3$.

Using the properties proved above we obtain that the general expression of Ω_n is

$$\Omega_n = \sum_{j=1}^{n+2} \sum_{\alpha_1 + \cdots + \alpha_j = n+2} \varepsilon_\alpha c_0^{n+2-j} c_{\alpha_1} \cdots c_{\alpha_j},$$

where $\alpha = (\alpha_1, \ldots, \alpha_j)$, $1 \leq \alpha_k \leq n + 2$, $1 \leq k \leq j$ and $\varepsilon_\alpha = \pm 1$. Thus

$$|\Omega_n| \leq \sum_{j=1}^{n+2} \sum_{\alpha_1 + \cdots + \alpha_j = n+2} c_0^{n+2-j} c_{\alpha_1} \cdots c_{\alpha_j}. \tag{2.5}$$

In this sum, there is only one term which contains the factor c_{n+2} and that is when $j = 1$. That term is $c_0^{n+1} c_{n+2}$. In the rest of the terms $1 \leq \alpha_k \leq n+1$. Let us prove that

$$c_0^{n+2-j} c_{\alpha_1} \cdots c_{\alpha_j} \leq c_0^{n+1} c_{n+2} \quad \text{for all } \alpha. \tag{2.6}$$

In fact, using the Hölder inequality on each factor except the first, it follows that

$$c_0^{n+2-j} c_{\alpha_1} \cdots c_{\alpha_j} \leq c_0^{n+2-j} \left(\int x^{n+2} d\sigma_1(x) \right)^{\sum_{k=1}^{j} \alpha_k/(n+2)} \left(\int d\sigma_1(x) \right)^{j - (\sum_{k=1}^{j} \alpha_k)/(n+2)}.$$

It remains to employ that $\sum_{k=1}^{j} \alpha_k = n + 2$ to complete the proof of (2.6).

From (2.4), (2.5), and (2.6), we have that

$$d_n \leq 2^{n+1} c_{n+2}/c_0^2$$

and the Carleman condition for τ_1 readily follows. $\square$

An immediate consequence of Theorem 1.5 is the following

Corollary 2.5. *Let* $(s_{1,1}, \ldots, s_{1,m}) = \mathcal{N}(\sigma_1, \ldots, \sigma_m)$ *be such that* Δ_1 *is contained in a half line and* σ_1 *satisfies Carleman's condition. Then, for all* $j = 1, \ldots, m$ *we have that* $s_{1,j}$ *and* $\tau_{1,j}$ *satisfy Carleman's condition.*

Proof. For $s_{1,1}$ the assertion is the hypothesis. Let $j \in \{2, \ldots, m\}$. Notice that $s_{1,j} = \langle \sigma_1, s_{2,j} \rangle$ and $(s_{1,1}, s_{1,j}) = \mathcal{N}(\sigma_1, s_{2,j})$ so $s_{1,j}$, $j = 2, \ldots, n$ satisfies Carleman's condition due to Theorem 1.5. Since $s_{1,j}$, $j = 1, \ldots, m$ satisfies Carleman's condition then Theorem 1.5 also gives that $\tau_{1,j}$, $j = 1, \ldots, m$ satisfies Carleman's condition. $\square$

Actually we will use this result for $(s_{m,m}, \ldots, s_{m,1}) = \mathcal{N}(\sigma_m, \ldots, \sigma_1)$.

3. Proof of Theorem 1.4

The first step consists in proving a weaker version of (1.8).

Lemma 3.1. *Assume that the conditions of Theorem 1.4 are fulfilled. Then, for each fixed* $j = 0, \ldots, m-1$

$$h\text{-}\lim_{n \in \Lambda} \frac{a_{\mathbf{n},j}}{a_{\mathbf{n},m}} = (-1)^{m-j} \widehat{s}_{m,j+1},$$

$$h\text{-}\lim_{n \in \Lambda} \frac{a_{\mathbf{n},m}}{a_{\mathbf{n},j}} = (-1)^{m-j} \widehat{s}_{m,j+1}^{-1} \quad inside\ \mathbb{C} \setminus \Delta_m. \tag{3.1}$$

There exists a constant C_1*, independent of* Λ*, such that for each* $j = 0, \ldots, m$ *and* $\mathbf{n} \in \Lambda$*, the polynomials* $a_{\mathbf{n},j}$ *have at least* $(|\mathbf{n}|/m) - C_1$ *zeros in* $\overset{\circ}{\Delta}_j$*.*

Proof. If $m = 1$ the statement reduces directly to Lemma 2.4, so without loss of generality we can assume that $m \geq 2$. Fix $\mathbf{n} \in \Lambda$.

In Theorem 1.3 we proved that $\mathcal{A}_{\mathbf{n},1}$ has exactly $|\mathbf{n}| - 1$ simple zeros in $\mathbb{C} \setminus \Delta_2$ and they all lie in $\overset{\circ}{\Delta}_1$. Therefore, there exists a polynomial $w_{\mathbf{n},1}$, $\deg w_{\mathbf{n},1} = |\mathbf{n}| - 1$, whose zeros lie in $\overset{\circ}{\Delta}_1$ such that

$$\frac{\mathcal{A}_{\mathbf{n},1}}{w_{\mathbf{n},1}} \in \mathcal{H}(\mathbb{C} \setminus \Delta_2) \quad \text{and} \quad \frac{\mathcal{A}_{\mathbf{n},1}}{w_{\mathbf{n},1}} = \mathcal{O}\left(\frac{1}{z^{|\mathbf{n}|-\bar{n}_1}}\right), \quad z \to \infty, \tag{3.2}$$

where $\bar{n}_j = \max\{n_k : k = j, \ldots, m\}$.

From (3.2) and Lemma 2.1 it follows that

$$\int x^\nu \mathcal{A}_{\mathbf{n},2}(x) \frac{d\sigma_2(x)}{w_{\mathbf{n},1}(x)} = 0, \quad \nu = 0, \ldots, |\mathbf{n}| - \bar{n}_1 - 2, \tag{3.3}$$

and

$$\frac{\mathcal{A}_{\mathbf{n},1}(z)}{w_{\mathbf{n},1}(z)} = \int \frac{\mathcal{A}_{\mathbf{n},2}(x) d\sigma_2(x)}{w_{\mathbf{n},1}(x)(z-x)}. \tag{3.4}$$

In particular, (3.3) implies that $\mathcal{A}_{\mathbf{n},2}$ has at least $|\mathbf{n}| - \bar{n}_1 - 1$ sign changes in $\mathring{\Delta}_2$. (We cannot claim that $\mathcal{A}_{\mathbf{n},2}$ has exactly $|\mathbf{n}| - \bar{n}_1 - 1$ simple zeros in $\mathbb{C} \setminus \Delta_3$ and that they all lie in $\mathring{\Delta}_2$ except if $\bar{n}_1 = n_1$.) Therefore, there exists a polynomial $w_{\mathbf{n},2}$, $\deg w_{\mathbf{n},2} = |\mathbf{n}| - \bar{n}_1 - 1$, whose zeros lie in $\mathring{\Delta}_2$, such that

$$\frac{\mathcal{A}_{\mathbf{n},2}}{w_{\mathbf{n},2}} \in \mathcal{H}(\mathbb{C} \setminus \Delta_3) \quad \text{and} \quad \frac{\mathcal{A}_{\mathbf{n},2}}{w_{\mathbf{n},2}} = \mathcal{O}\left(\frac{1}{z^{|\mathbf{n}| - \bar{n}_1 - \bar{n}_2}}\right), \quad z \to \infty.$$

Iterating this process, using Lemma 2.1 several times, on step j, $j \in \{1, \ldots, m\}$, we find that there exists a polynomial $w_{\mathbf{n},j}$, $\deg w_{\mathbf{n},j} = |\mathbf{n}| - \bar{n}_1 - \cdots - \bar{n}_{j-1} - 1$, whose zeros are points where $\mathcal{A}_{\mathbf{n},j}$ changes sign in $\mathring{\Delta}_j$ such that

$$\frac{\mathcal{A}_{\mathbf{n},j}}{w_{\mathbf{n},j}} \in \mathcal{H}(\mathbb{C} \setminus \Delta_{j+1}) \quad \text{and} \quad \frac{\mathcal{A}_{\mathbf{n},j}}{w_{\mathbf{n},j}} = \mathcal{O}\left(\frac{1}{z^{|\mathbf{n}| - \bar{n}_1 - \cdots - \bar{n}_j}}\right), \quad z \to \infty. \tag{3.5}$$

This process concludes as soon as $|\mathbf{n}| - \bar{n}_1 - \cdots - \bar{n}_j \le 0$. Since $\lim_{\mathbf{n} \in \Lambda} |\mathbf{n}| = \infty$, because of (1.7) we can always take m steps for all $\mathbf{n} \in \Lambda$ with $|\mathbf{n}|$ sufficiently large. In what follows, we only consider such $\mathbf{n}$'s.

When $n_1 = \bar{n}_1 \ge \cdots \ge n_m = \bar{n}_m$, we obtain that $\mathcal{A}_{\mathbf{n},m} \equiv a_{\mathbf{n},m}$ has $n_m - 1$ sign changes in $\mathring{\Delta}_m$ and since $\deg a_{\mathbf{n},m} \le n_m - 1$ this means that $\deg a_{\mathbf{n},m} = n_m - 1$ and all its zeros lie in $\mathring{\Delta}_m$. (In fact, in this case we can prove that $\mathcal{A}_{\mathbf{n},j}$, $j = 1, \ldots, m$ has exactly $|\mathbf{n}| - n_1 - \cdots - n_{j-1} - 1$ zeros in $\mathbb{C} \setminus \Delta_{j+1}$, that they are all simple and lie in $\mathring{\Delta}_j$, where $\Delta_{m+1} = \emptyset$, compare with [8, Propositions 2.5, 2.7].)

In general, we have that $a_{\mathbf{n},m}$ has at least $|\mathbf{n}| - \bar{n}_1 - \cdots - \bar{n}_{m-1} - 1$ sign changes in $\mathring{\Delta}_m$; therefore, the number of zeros of $a_{\mathbf{n},m}$ which may lie outside of Δ_m is bounded by

$$\deg a_{\mathbf{n},m} - \left(|\mathbf{n}| - \bar{n}_1 - \cdots - \bar{n}_{m-1} - 1\right) \le \sum_{k=1}^{m-1} \bar{n}_k - n_k \le (m-1)C,$$

where C is the constant given in (1.7), which does not depend on $\mathbf{n} \in \Lambda$.

For $j = m-1$ we have that there exists $w_{\mathbf{n},m-1}$, $\deg w_{\mathbf{n},m-1} = |\mathbf{n}| - \bar{n}_1 - \cdots - \bar{n}_{m-2} - 1$, whose zeros lie in $\mathring{\Delta}_{m-1}$ such that

$$\frac{\mathcal{A}_{\mathbf{n},m-1}}{w_{\mathbf{n},m-1}} = \frac{a_{\mathbf{n},m-1} + a_{\mathbf{n},m}\widehat{\sigma}_m}{w_{\mathbf{n},m-1}} \in \mathcal{H}(\mathbb{C} \setminus \Delta_m) \quad \text{and}$$

$$\frac{\mathcal{A}_{\mathbf{n},m-1}}{w_{\mathbf{n},m-1}} = \mathcal{O}\left(\frac{1}{z^{|\mathbf{n}|-\bar{n}_1-\cdots-\bar{n}_{m-1}}}\right), \quad z \to \infty,$$

where $\deg a_{\mathbf{n},m-1} \le n_{m-1} - 1$, $\deg a_{\mathbf{n},m} \le n_m - 1$. Thus, using (1.7) it is easy to check that $(a_{\mathbf{n},m-1}/a_{\mathbf{n},m})_{n \in \Lambda}$ forms a sequence of incomplete diagonal multi-point Padé approximants of $-\widehat{\sigma}_m$ satisfying a)–b) with appropriate values of n, κ and ℓ. Due to Lemma 2.4 it follows that

$$h\text{-}\lim_{n \in \Lambda} \frac{a_{\mathbf{n},m-1}}{a_{\mathbf{n},m}} = -\widehat{\sigma}_m, \quad \text{inside } \mathbb{C} \setminus \Delta_m.$$

Dividing by $\widehat{\sigma}_m$ and using (1.10), we also have

$$\frac{\mathcal{A}_{\mathbf{n},m-1}}{\widehat{\sigma}_m w_{\mathbf{n},m-1}} = \frac{a_{\mathbf{n},m-1}\widehat{\tau}_m + b_{\mathbf{n},m-1}}{w_{\mathbf{n},m-1}} \in \mathcal{H}(\mathbb{C} \setminus \Delta_m),$$

where $b_{\mathbf{n},m-1} = a_{\mathbf{n},m} + \ell_m a_{\mathbf{n},m-1}$ and

$$\frac{\mathcal{A}_{\mathbf{n},m-1}}{\widehat{\sigma}_m w_{\mathbf{n},m-1}} = \mathcal{O}\left(\frac{1}{z^{|\mathbf{n}|-\bar{n}_1-\cdots-\bar{n}_{m-1}-1}}\right), \quad z \to \infty.$$

Consequently, $(b_{\mathbf{n},m-1}/a_{\mathbf{n},m-1})_{n \in \Lambda}$ forms a sequence of incomplete diagonal multi-point Padé approximants of $-\widehat{\tau}_m$ satisfying a)–b) with appropriate values of n, κ and ℓ. Then Lemma 2.4 and Corollary 2.5 imply that

$$h\text{-}\lim_{n \in \Lambda} \frac{b_{\mathbf{n},m-1}}{a_{\mathbf{n},m-1}} = -\widehat{\tau}_m, \quad \text{inside } \mathbb{C} \setminus \Delta_m,$$

or, equivalently,

$$h\text{-}\lim_{n \in \Lambda} \frac{a_{\mathbf{n},m}}{a_{\mathbf{n},m-1}} = -\widehat{\sigma}_m^{-1}, \quad \text{inside } \mathbb{C} \setminus \Delta_m.$$

We have proved (3.1) for $j = m - 1$.

For $j = m - 2$, we have shown that there exists a polynomial $w_{\mathbf{n},m-2}$, $\deg w_{\mathbf{n},m-2} = |\mathbf{n}| - \bar{n}_1 - \cdots - \bar{n}_{m-3} - 1$, whose zeros lie in $\overset{\circ}{\Delta}_{m-2}$ such that

$$\frac{\mathcal{A}_{\mathbf{n},m-2}}{w_{\mathbf{n},m-2}} = \frac{a_{\mathbf{n},m-2} + a_{\mathbf{n},m-1}\widehat{\sigma}_{m-1} + a_{\mathbf{n},m}\widehat{\langle \sigma_{m-1}, \sigma_m \rangle}}{w_{\mathbf{n},m-2}} \in \mathcal{H}(\mathbb{C} \setminus \Delta_{m-1})$$

and

$$\frac{\mathcal{A}_{\mathbf{n},m-2}}{w_{\mathbf{n},m-2}} = \mathcal{O}\left(\frac{1}{z^{|\mathbf{n}|-\bar{n}_1-\cdots-\bar{n}_{m-2}}}\right), \quad z \to \infty.$$

However, using (1.10) and (2.3), we obtain

$$
\frac{a_{\mathbf{n},m-2} + a_{\mathbf{n},m-1}\widehat{\sigma}_{m-1} + a_{\mathbf{n},m}\widehat{\langle\sigma_{m-1},\sigma_m\rangle}}{\widehat{\sigma}_{m-1}}
$$

$$
= (\ell_{m-1}a_{\mathbf{n},m-2} + a_{\mathbf{n},m-1} + C_1 a_{\mathbf{n},m}) + a_{\mathbf{n},m-2}\widehat{\tau}_{m-1} - a_{\mathbf{n},m}\widehat{\langle\tau_{m-1},\langle\sigma_m,\sigma_{m-1}\rangle\rangle},
$$

where $\deg \ell_{m-1} = 1$ and C_1 is a constant. Consequently, $\mathcal{A}_{\mathbf{n},m-2}/\widehat{\sigma}_{m-1}$ adopts the form of $\mathcal{A}$ in Lemma 2.1, $\mathcal{A}_{\mathbf{n},m-2}/(\widehat{\sigma}_{m-1}w_{\mathbf{n},m-2}) \in \mathcal{H}(\mathbb{C} \setminus \Delta_{m-1})$, and

$$
\frac{\mathcal{A}_{\mathbf{n},m-2}}{\widehat{\sigma}_{m-1}w_{\mathbf{n},m-2}} = \mathcal{O}\left(\frac{1}{z^{|\mathbf{n}|-\bar{n}_1-\cdots-\bar{n}_{m-2}-1}}\right), \quad z \to \infty. \tag{3.6}
$$

From (2.2) in Lemma 2.1 it follows that for $\nu = 0, \ldots, |\mathbf{n}| - \bar{n}_1 - \cdots - \bar{n}_{m-2} - 3$

$$
\int x^\nu \left(a_{\mathbf{n},m-2}(x) - a_{\mathbf{n},m}(x)\widehat{\langle\sigma_m,\sigma_{m-1}\rangle}(x)\right)\frac{d\tau_{m-1}(x)}{w_{\mathbf{n},m-2}(x)} = 0.
$$

Therefore, $a_{\mathbf{n},m-2} - a_{\mathbf{n},m}\widehat{\langle\sigma_m,\sigma_{m-1}\rangle} \in \mathcal{H}(\mathbb{C} \setminus \Delta_m)$ must have at least $|\mathbf{n}| - \bar{n}_1 - \cdots - \bar{n}_{m-2} - 2$ sign changes in $\mathring{\Delta}_{m-1}$. This means that there exists a polynomial $w^*_{\mathbf{n},m-2}$, $\deg w^*_{\mathbf{n},m-2} = |\mathbf{n}| - \bar{n}_1 - \cdots - \bar{n}_{m-2} - 2$, whose zeros are simple and lie in $\mathring{\Delta}_{m-1}$ such that

$$
\frac{a_{\mathbf{n},m-2} - a_{\mathbf{n},m}\widehat{\langle\sigma_m,\sigma_{m-1}\rangle}}{w^*_{\mathbf{n},m-2}} \in \mathcal{H}(\mathbb{C} \setminus \Delta_m)
$$

and

$$
\frac{a_{\mathbf{n},m-2} - a_{\mathbf{n},m}\widehat{\langle\sigma_m,\sigma_{m-1}\rangle}}{w^*_{\mathbf{n},m-2}} = \mathcal{O}\left(\frac{1}{z^{|\mathbf{n}|-\bar{n}_1-\cdots-\bar{n}_{m-3}-2\bar{n}_{m-2}-1}}\right), \quad z \to \infty.
$$

Due to (1.7), this implies that $(a_{\mathbf{n},m-2}/a_{\mathbf{n},m})$, $n \in \Lambda$, is a sequence of incomplete diagonal Padé approximants of $\langle\sigma_m,\sigma_{m-1}\rangle$. By Lemma 2.4 and Corollary 2.5 we obtain its convergence in Hausdorff content to $\widehat{\langle\sigma_m,\sigma_{m-1}\rangle}$. To prove the other part in (3.1), we divide by $\widehat{\langle\sigma_m,\sigma_{m-1}\rangle}(z)$ use (1.10) and proceed as we did in the case $j = m-1$.

Let us prove (3.1) in general. Fix $j \in \{0, \ldots, m-3\}$ (for $j = m-2, m-1$ it's been proved). Having in mind (3.5) we need to reduce $\mathcal{A}_{\mathbf{n},j}$ so as to eliminate all $a_{\mathbf{n},k}$, $k = j+1, \ldots, m-1$. We start out eliminating $a_{\mathbf{n},j+1}$. Consider the ratio $\mathcal{A}_{\mathbf{n},j}/\widehat{\sigma}_{j+1}$. Using (1.10) and (2.3) we obtain that

$$
\frac{\mathcal{A}_{\mathbf{n},j}}{\widehat{\sigma}_{j+1}} = \left(\ell_{j+1}a_{\mathbf{n},j} + \sum_{k=j+1}^{m}\frac{|s_{j+1,k}|}{|\sigma_{j+1}|}a_{\mathbf{n},j+1}\right) + a_{\mathbf{n},j}\widehat{\tau}_{j+1}
$$

$$
- \sum_{k=j+2}^{m}a_{\mathbf{n},k}\widehat{\langle\tau_{j+1},\langle s_{j+2,k},\sigma_{j+1}\rangle\rangle}
$$

has the form of $\mathcal{A}$ in Lemma 2.1, where $\mathcal{A}_{\mathbf{n},j}/(\widehat{\sigma}_{j+1}w_{\mathbf{n},j}) \in \mathcal{H}(\mathbb{C} \setminus \Delta_{j+1})$, and

$$\frac{\mathcal{A}_{\mathbf{n},j}}{\widehat{\sigma}_{j+1}w_{\mathbf{n},j}} \in \mathcal{O}\left(\frac{1}{z^{|\mathbf{n}|-\bar{n}_1-\cdots-\bar{n}_j-1}}\right), \quad z \to \infty.$$

From (2.2) of Lemma 2.1, we obtain that for $\nu = 0, \dots, |\mathbf{n}| - \bar{n}_1 - \cdots - \bar{n}_j - 3$

$$0 = \int x^\nu \left(a_{\mathbf{n},j}(x) - \sum_{k=j+2}^m a_{\mathbf{n},k} \widehat{\langle s_{j+2,k}, \sigma_{j+1}\rangle}(x) \right) \frac{d\tau_{j+1}(x)}{w_{\mathbf{n},j}(x)}$$

which implies that the function in parenthesis under the integral sign has at least $|\mathbf{n}| - \bar{n}_1 - \cdots - \bar{n}_j - 2$ sign changes in $\mathring{\Delta}_{j+1}$. In turn, it follows that there exists a polynomial $\widetilde{w}_{\mathbf{n},j}$, $\deg \widetilde{w}_{\mathbf{n},j} = |\mathbf{n}| - \bar{n}_1 - \cdots - \bar{n}_j - 2$, whose zeros are simple and lie in $\mathring{\Delta}_{j+1}$ such that

$$\frac{a_{\mathbf{n},j} - \sum_{k=j+2}^m a_{\mathbf{n},k} \widehat{\langle s_{j+2,k}, \sigma_{j+1}\rangle}}{\widetilde{w}_{\mathbf{n},j}} \in \mathcal{H}(\mathbb{C} \setminus \Delta_{j+2})$$

and

$$\frac{a_{\mathbf{n},j} - \sum_{k=j+2}^m a_{\mathbf{n},k} \widehat{\langle s_{j+2,k}, \sigma_{j+1}\rangle}}{\widetilde{w}_{\mathbf{n},j}} = \mathcal{O}\left(\frac{1}{z^{|\mathbf{n}|-\bar{n}_1-\cdots-\bar{n}_{j-1}-2\bar{n}_j-1}}\right), \quad z \to \infty.$$

Notice that $a_{\mathbf{n},j+1}$ has been eliminated and that

$$\langle s_{j+2,k}, \sigma_{j+1}\rangle = \langle\langle \sigma_{j+2}, \sigma_{j+1}\rangle, \sigma_{j+3}, \dots, \sigma_k\rangle, \quad k = j+3, \dots, m.$$

Now we must do away with $a_{\mathbf{n},j+2}$ in $a_{\mathbf{n},j} - \sum_{k=j+2}^m a_{\mathbf{n},k} \widehat{\langle s_{j+2,k}, \sigma_{j+1}\rangle}$ (in case that $j+2 < m$). To this end, we consider the ratio

$$\frac{a_{\mathbf{n},j} - \sum_{k=j+2}^m a_{\mathbf{n},k} \widehat{\langle s_{j+2,k}, \sigma_{j+1}\rangle}}{\widehat{\langle \sigma_{j+2}, \sigma_{j+1}\rangle}}$$

and repeat the arguments employed above with $\mathcal{A}_{\mathbf{n},j}$. After $m-j-2$ reductions obtained applying consecutively Lemma 2.1, we find that there exists a polynomial which we denote $w^*_{\mathbf{n},j}$, $\deg w^*_{\mathbf{n},j} = |\mathbf{n}| - \bar{n}_1 - \cdots - \bar{n}_{j-1} - (m-j-1)\bar{n}_j - 2$, whose zeros are simple and lie in $\mathring{\Delta}_{m-1}$ such that

$$\frac{a_{\mathbf{n},j} - (-1)^{m-j} a_{\mathbf{n},m} \widehat{\langle \sigma_m, \dots, \sigma_{j+1}\rangle}}{w^*_{\mathbf{n},j}} \in \mathcal{H}(\mathbb{C} \setminus \Delta_m)$$

and

$$\frac{a_{\mathbf{n},j} - (-1)^{m-j} a_{\mathbf{n},m} \widehat{\langle \sigma_m, \dots, \sigma_{j+1}\rangle}}{w^*_{\mathbf{n},j}} = \mathcal{O}\left(\frac{1}{z^{|\mathbf{n}|-\bar{n}_1-\cdots-\bar{n}_{j-1}-(m-j)\bar{n}_j-1}}\right), \quad z \to \infty.$$

Dividing by $(-1)^{m-j}\langle\sigma_m,\ldots,\sigma_{j+1}\widehat{\rangle}$, from here we also get that

$$\frac{a_{\mathbf{n},j}(-1)^{m-j}\langle\sigma_m,\ldots,\sigma_{j+1}\widehat{\rangle}^{-1} - a_{\mathbf{n},m}}{w^*_{\mathbf{n},j}} \in \mathcal{H}(\mathbb{C}\setminus\Delta_m)$$

and

$$\frac{a_{\mathbf{n},j}(-1)^{m-j}\langle\sigma_m,\ldots,\sigma_{j+1}\widehat{\rangle}^{-1} - a_{\mathbf{n},m}}{w^*_{\mathbf{n},j}} = \mathcal{O}\left(\frac{1}{z^{|\mathbf{n}|-\bar{n}_1-\cdots-\bar{n}_{j-1}-(m-j)\bar{n}_j-2}}\right), \quad z\to\infty.$$

On account of (1.7), these relations imply that $(a_{\mathbf{n},j}/a_{\mathbf{n},m})$, $\mathbf{n}\in\Lambda$, is a sequence of incomplete diagonal multi-point Padé approximants of $(-1)^{m-j}\langle\sigma_m,\ldots,\sigma_{j+1}\widehat{\rangle}$ and $(a_{\mathbf{n},m}/a_{\mathbf{n},j})$, $\mathbf{n}\in\Lambda$, is a sequence of incomplete diagonal multi-point Padé approximants of $(-1)^{m-j}\langle\sigma_m,\ldots,\sigma_{j+1}\widehat{\rangle}^{-1}$. Since $\langle\sigma_m,\ldots,\sigma_{j+1}\widehat{\rangle}^{-1} = \widehat{\tau}_{m,j+1}+\ell_{m,j+1}$, from Lemma 2.4 and Corollary 2.5 we obtain (3.1).

Going one step further using Lemma 2.1, we also obtain that

$$0 = \int x^\nu a_{\mathbf{n},j}(x)\frac{d\tau_{m,j+1}}{w^*_{\mathbf{n},j}(x)}, \quad \nu = 0,\ldots,|\mathbf{n}| - \bar{n}_1 - \cdots - \bar{n}_{j-1} - (m-j)\bar{n}_j - 4$$

which implies that $a_{\mathbf{n},j}$ has at least $|\mathbf{n}| - \bar{n}_1 - \cdots - \bar{n}_{j-1} - (m-j)\bar{n}_j - 3$ sign changes in $\overset{\circ}{\Delta}_m$. From (1.7) we obtain that there exists a non-negative constant C_1, independent of $n\in\Lambda$, such that the number of zeros of $a_{\mathbf{n},j}$, $j = 0,\ldots,m$, in Δ_m is bounded from below by $(|\mathbf{n}|/m) - C_1$. This settles the last statement. $\square$

In the case of decreasing components in $\mathbf{n}$, we saw that all the zeros of $a_{\mathbf{n},m}$ lie in Δ_m and [11, Lemma 1] would allow us to derive immediately uniform convergence on each compact subset of $\mathbb{C}\setminus\Delta_m$ from the convergence in Hausdorff content. For other configurations of the components we have to work a little harder.

Proof of Theorem 1.4. Let $\bar{\jmath}$ be the last component of $(n_0,\ldots,n_m)$ such that $n_{\bar{\jmath}} = \min_{j=0,\ldots,m}(n_j)$. Let us prove that $\deg a_{\mathbf{n},\bar{\jmath}} = n_{\bar{\jmath}} - 1$, that all its zeros are simple and lie in $\overset{\circ}{\Delta}_m$.

From [10, Theorem 3.2] (see also [9, Theorem 1.3]) we know that there exists a permutation λ of $(0,\ldots,m)$ which reorders the components of $(n_0,n_1,\ldots,n_m)$ decreasingly, $n_{\lambda(0)} \geq \cdots \geq n_{\lambda(m)}$, and an associated Nikishin system $(r_{1,1},\ldots,r_{1,m}) = \mathcal{N}(\rho_1,\ldots,\rho_m)$ such that

$$\mathcal{A}_{\mathbf{n},0} = \left(q_{\mathbf{n},0} + \sum_{k=1}^m q_{\mathbf{n},k}\widehat{r}_{1,k}\right)\widehat{s}_{1,\lambda(0)}, \quad \deg q_{\mathbf{n},k} \leq n_{\lambda(k)} - 1, \ k = 0,\ldots,m.$$

The permutation may be taken so that for all $0 \leq j < k \leq n$ with $n_j = n_k$ then also $\lambda(j) < \lambda(k)$. In this case, see formulas (31) in the proof of [9, Lemma 2.3], it follows that

$q_{\mathbf{n},m} = \pm a_{\mathbf{n},\jmath}$. Reasoning with $q_{\mathbf{n},0} + \sum_{k=1}^{m} q_{\mathbf{n},k}\widehat{r}_{1,k}$ as we did with $\mathcal{A}_{\mathbf{n},0}$ we obtain that $\deg q_{\mathbf{n},m} = n_{\lambda(m)} - 1$ and that its zeros are all simple and lie in $\overset{\circ}{\Delta}_m$. However, $n_{\lambda(m)} = n_{\jmath}$ and $q_{\mathbf{n},m} = \pm a_{\mathbf{n},\jmath}$ so the statement holds.

The index $\jmath$ as defined above may depend on the multi-index $\mathbf{n} \in \Lambda$. Given $\jmath \in \{0, \ldots, m\}$, let us denote by $\Lambda(\jmath)$ the set of all $\mathbf{n} \in \Lambda$ such that $\jmath$ is the last component of $(n_0, \ldots, n_m)$ such that $n_{\jmath} = \min_{j=0,\ldots,m}(n_j)$. Fix $\jmath$ and suppose that $\Lambda(\jmath)$ contains infinitely many multi-indices. If $\jmath = m$, then [11, Lemma 1] and the first limit in (3.1) imply that

$$\lim_{n \in \Lambda(m)} \frac{a_{\mathbf{n},j}}{a_{\mathbf{n},m}} = (-1)^{m-j}\widehat{s}_{m,j+1}, \quad j = 0, \ldots, m-1,$$

uniformly on each compact subset of $\mathbb{C} \setminus \Delta_m$, as needed.

Assume that $\jmath \in \{0, \ldots, m-1\}$. Since all the zeros of $a_{\mathbf{n},\jmath}$ lie in $\overset{\circ}{\Delta}_m$, using [11, Lemma 1] and the second limit in (3.1) for $j = \jmath$, we obtain that

$$\lim_{\mathbf{n} \in \Lambda(\jmath)} \frac{a_{\mathbf{n},m}}{a_{\mathbf{n},\jmath}} = \frac{1}{(-1)^{m-\jmath}\widehat{s}_{m,j+1}}, \tag{3.7}$$

uniformly on each compact subset of $\mathbb{C} \setminus \Delta_m$. The function on the right hand side of (3.7) is holomorphic and never zero on $\mathbb{C} \setminus \Delta_m$ and the approximating functions are holomorphic on $\mathbb{C} \setminus \Delta_m$. Using Rouche's theorem it readily follows that on any compact subset $\mathcal{K} \subset \mathbb{C} \setminus \Delta_m$ for all sufficiently large $|\mathbf{n}|$, $n \in \Lambda(\jmath)$, the polynomials $a_{\mathbf{n},m}$ have no zero on $\mathcal{K}$. This is true for any $\jmath \in \{0, \ldots, m\}$ such that $\Lambda(\jmath)$ contains infinitely many multi-indices. Therefore, the only accumulation points of the zeros of the polynomials $a_{\mathbf{n},m}$ are in $\Delta_m \cup \{\infty\}$.

Hence, on any bounded region D such that $\overline{D} \subset \mathbb{C} \setminus \Delta_m$ for each fixed $j = 0, \ldots, m-1$, and all sufficiently large $|\mathbf{n}|$, $\mathbf{n} \in \Lambda$, we have that $a_{\mathbf{n},j}/a_{\mathbf{n},m} \in \mathcal{H}(D)$. From [11, Lemma 1] and the first part of (3.1) it follows that

$$\lim_{\mathbf{n} \in \Lambda} \frac{a_{\mathbf{n},j}}{a_{\mathbf{n},m}} = (-1)^{m-j}\widehat{s}_{m,j+1}, \quad j = 0, \ldots, m-1, \tag{3.8}$$

uniformly on each compact subset of D. Since D was chosen arbitrarily, as long as $\overline{D} \subset \mathbb{C} \setminus \Delta_m$, it follows that the convergence is uniform on each compact subset of $\mathbb{C} \setminus \Delta_m$ and we have (1.8). Since the right hand side of (1.8) is a function which does not vanish in $\overline{D} \subset \mathbb{C} \setminus \Delta_m$, Rouche's theorem implies that for each $j = 0, \ldots, m-1$ the accumulation points of the zeros of the polynomials $a_{\mathbf{n},j}$ must be in $\Delta_m \cup \{\infty\}$ as claimed. (For $j = m$ this was proved above.)

Now,

$$\frac{\mathcal{A}_{\mathbf{n},j}}{a_{\mathbf{n},m}} = \frac{a_{\mathbf{n},j}}{a_{\mathbf{n},m}} + \sum_{k=j+1}^{m-1} \frac{a_{\mathbf{n},k}}{a_{\mathbf{n},m}}\widehat{s}_{j+1,k} + \widehat{s}_{j+1,m}.$$

According to formula (17) in [10, Lemma 2.9]

$$0 \equiv (-1)^{m-j}\widehat{s}_{m,j+1} + \sum_{k=j+1}^{m-1} (-1)^{m-k}\widehat{s}_{m,k+1}\widehat{s}_{j+1,k} + \widehat{s}_{j+1,m}, \quad z \in \mathbb{C} \setminus (\Delta_{j+1} \cup \Delta_m).$$

Deleting one expression from the other we have that

$$\frac{\mathcal{A}_{\mathbf{n},j}}{a_{\mathbf{n},m}} = \left(\frac{a_{\mathbf{n},j}}{a_{\mathbf{n},m}} - (-1)^{m-j}\widehat{s}_{m,j+1} \right) + \sum_{k=j+1}^{m-1} \left(\frac{a_{\mathbf{n},k}}{a_{\mathbf{n},m}} - (-1)^{m-k}\widehat{s}_{m,k+1} \right) \widehat{s}_{j+1,k} \quad (3.9)$$

Consequently, for each $j = 0, \ldots, m-1$, from (1.8) we obtain (1.9). $\square$

Remark 3.2. Observe that in the process of the proof of Theorem 1.4 we showed that if $\check{\jmath}$ is the last component of an arbitrary multi-index $(n_0, \ldots, n_m)$ for which $n_{\check{\jmath}} = \min_{j=0,\ldots,m}(n_j)$, then $\deg a_{\mathbf{n},\check{\jmath}} = n_{\check{\jmath}} - 1$, all the zeros of $a_{\mathbf{n},\check{\jmath}}$ are simple and lie in $\overset{\circ}{\Delta}_m$. For sequences of multi-indices verifying (1.7) one can show that for all $a_{\mathbf{n},j}$, $j = 0, \ldots, m$ the number of zeros that are not simple and/or lie in the complement of $\overset{\circ}{\Delta}_m$ is bounded by a constant (related with the constant C in (1.7)) which is independent of $\mathbf{n}$. We will not dwell into this, instead we will prove some relations concerning the asymptotic configuration of the zeros of these polynomials.

Suppose that Δ_m is bounded. Let Γ be a positively oriented closed simple Jordan curve that surrounds Δ_m. Define $\kappa_{\mathbf{n},j}(\Gamma)$, $j = 0, \ldots, m$ to be the number of zeros of $a_{\mathbf{n},j}$ outside Γ. As above, given $\check{\jmath} \in \{0, \ldots, m\}$, let us denote by $\Lambda(\check{\jmath})$ the set of all $\mathbf{n} \in \Lambda$ such that $\check{\jmath}$ is the last component of $(n_0, \ldots, n_m)$ which satisfies $n_{\check{\jmath}} = \min_{j=0,\ldots,m}(n_j)$.

Corollary 3.3. *Suppose that the assumptions of Theorem 1.4 hold and Δ_m is bounded. Then for all sufficiently large $|\mathbf{n}|$, $\mathbf{n} \in \Lambda(\check{\jmath})$,*

$$\kappa_{\mathbf{n},j}(\Gamma) = \begin{cases} n_j - n_{\check{\jmath}}, & j = 0, \ldots, m-1, \\ n_m - n_{\check{\jmath}} - 1, & j = m. \end{cases} \quad (3.10)$$

The rest of the zeros of the polynomials $a_{\mathbf{n},j}$ accumulate (or lie) on Δ_m.

Proof. Fix $\check{\jmath} \in \{0, \ldots, m-1\}$. Assume that $\Lambda(\check{\jmath})$ contains infinitely many multi-indices. Using the argument principle and (3.7) it follows that

$$\lim_{\mathbf{n} \in \Lambda(\check{\jmath})} \frac{1}{2\pi i} \int_{\Gamma} \frac{(a_{\mathbf{n},m}/a_{\mathbf{n},\check{\jmath}})'(z)}{(a_{\mathbf{n},m}/a_{\mathbf{n},\check{\jmath}})(z)} dz = \frac{1}{2\pi i} \int_{\Gamma} \frac{(1/\widehat{s}_{m,\check{\jmath}+1})'(z)}{(1/\widehat{s}_{m,\check{\jmath}+1})(z)} dz = 1,$$

because $1/\widehat{s}_{m,j+1}$ has one pole and no zeros outside Γ (counting the point ∞). Recall that $\deg a_{\mathbf{n},j} = n_j - 1$, $j = 0, \ldots, m$ and that all the zeros of $a_{\mathbf{n},j}$ lie on Δ_m. Then, for all sufficiently large $|\mathbf{n}|$, $\mathbf{n} \in \Lambda(\check{j})$,

$$(n_m - 1) - (n_{\check{j}} - 1) - \kappa_{\mathbf{n},m}(\Gamma) = 1.$$

Consequently,

$$\kappa_{\mathbf{n},m}(\Gamma) = n_m - n_{\check{j}} - 1, \quad \mathbf{n} \in \Lambda(\check{j}). \tag{3.11}$$

Analogously, from (3.8), for $j = 0, \ldots, m-1$, we obtain

$$\lim_{\mathbf{n}\in\Lambda} \frac{1}{2\pi i} \int_\Gamma \frac{(a_{\mathbf{n},j}/a_{\mathbf{n},m})'(z)}{(a_{\mathbf{n},j}/a_{\mathbf{n},m})(z)} dz = \frac{1}{2\pi i} \int_\Gamma \frac{\widehat{s}'_{m,j+1}(z)}{\widehat{s}_{m,j+1}(z)} dz = -1.$$

Therefore, for all sufficiently large $|\mathbf{n}|$, $\mathbf{n} \in \Lambda$,

$$n_j - n_m + \kappa_{\mathbf{n},m}(\Gamma) - \kappa_{\mathbf{n},j}(\Gamma) = -1, \quad j = 0, \ldots, m-1,$$

which together with (3.11) gives (3.10). The last statement follows from the fact that the only accumulation points of the zeros of the $a_{\mathbf{n},j}$ are in $\Delta_m \cup \{\infty\}$. $\quad\square$

Remark 3.4. The thesis of Theorem 1.4 remains valid if in place of (1.7) we require that

$$n_j = \frac{|\mathbf{n}|}{m} + o(|\mathbf{n}|), \quad |\mathbf{n}| \to \infty, \; j = 1, \ldots, m. \tag{3.12}$$

To prove this we need an improved version of Lemma 2.4 in which the parameter ℓ in b) depends on n but $\ell(n) = o(n)$, $n \to \infty$. The proof of Lemma 2 in [1] admits this variation with some additional technical difficulties in part resolved in the proof of [7, Corollary 1].

Remark 3.5. If either Δ_m or Δ_{m-1} is a compact set and $\Delta_{m-1} \cap \Delta_m = \emptyset$, it not difficult to show that convergence takes place in (1.8) and (1.9) with geometric rate. More precisely, for $j = 0, \ldots, m-1$, and $\mathcal{K} \subset \mathbb{C} \setminus \Delta_m$, we have

$$\limsup_{\mathbf{n}\in\Lambda} \left\| \frac{a_{\mathbf{n},j}}{a_{\mathbf{n},m}} - (-1)^{m-j}\widehat{s}_{m,j+1} \right\|_{\mathcal{K}}^{1/|\mathbf{n}|} = \delta_j < 1. \tag{3.13}$$

For $j = 0, \ldots, m-1$, and $\mathcal{K} \subset \mathbb{C} \setminus (\Delta_{j+1} \cup \Delta_m)$

$$\limsup_{\mathbf{n}\in\Lambda} \left\| \frac{\mathcal{A}_{\mathbf{n},j}}{a_{\mathbf{n},m}} \right\|_{\mathcal{K}}^{1/|\mathbf{n}|} \leq \max\{\delta_k : j \leq k \leq m-1\} < 1. \tag{3.14}$$

The second relation trivially follows from the first and (3.9). The proof of the first is similar to that of [7, Corollary 1]. It is based on the fact that the number of interpolation

points on Δ_{m-1} is $\mathcal{O}(|\mathbf{n}|)$, $|\mathbf{n}| \to \infty$, and that the distance from Δ_m to Δ_{m-1} is positive. Relations (3.13) and (3.14) are also valid if (1.7) is replaced with (3.12).

Asymptotically, (3.12) still means that the components of $\mathbf{n}$ are equally valued. One can relax (3.12) requiring, for example, that the generating measures are regular in the sense of [21, Chapter 3] in which case the exact asymptotics of (3.13) and (3.14) can be given (see [17], [18, Chapter 5, Section 7], [8, Theorem 5.1, Corollary 5.3], and [19, Theorem 1]).

Remark 3.6. The previous results can be applied to other approximation schemes. Let $S^1 = \mathcal{N}(\sigma_0^1, \ldots, \sigma_{m_1}^1)$, $S^2 = \mathcal{N}(\sigma_0^2, \ldots, \sigma_{m_2}^2)$, $\sigma_0^1 = \sigma_0^2$ be given. Fix $\mathbf{n}_1 = (n_{1,0}, n_{1,1}, \ldots, n_{1,m_1}) \in \mathbb{Z}_+^{m_1+1}$ and $\mathbf{n}_2 = (n_{2,0}, n_{2,1}, \ldots, n_{2,m_2}) \in \mathbb{Z}_+^{m_2+1}$, $|\mathbf{n}_2| = |\mathbf{n}_1| - 1$. Let $\mathbf{n} = (\mathbf{n}_1, \mathbf{n}_2)$. There exists a non-zero vector polynomial with real coefficients $(a_{\mathbf{n},0}, \ldots, a_{\mathbf{n},m_1})$, $\deg(a_{\mathbf{n},k}) \leq n_{1,k} - 1$, $k = 0, \ldots, m_1$, such that for $j = 0, \ldots, m_2$,

$$\int x^\nu \mathcal{A}_{\mathbf{n},0}(x) ds_j^2(x) = 0, \quad \nu = 0, \ldots, n_{2,j} - 1,$$

where

$$\mathcal{A}_{\mathbf{n},0} = a_{\mathbf{n},0} + \sum_{k=1}^{m_1} a_{\mathbf{n},k} \widehat{s}_{1,k}^1.$$

In other words

$$\int \left(b_{\mathbf{n},0}(x) + \sum_{j=1}^{m_2} b_{\mathbf{n},j}(x) \widehat{s}_{1,j}^2(x) \right) \mathcal{A}_{\mathbf{n},0}(x) d\sigma_0^2(x) = 0, \quad \deg b_{\mathbf{n},j} \leq n_{2,j} - 1. \quad (3.15)$$

This implies that $\mathcal{A}_{\mathbf{n},0}$ has exactly $|\mathbf{n}_2|$ zeros in $\mathbb{C} \setminus \Delta_1^1$, they are all simple and lie in $\mathring{\Delta}_0^1$ (see [10, Theorem 1.2]). Here $\Delta_0^1 = \mathrm{Co}(\mathrm{supp}(\sigma_0^1))$ and $\Delta_1^1 = \mathrm{Co}(\mathrm{supp}(\sigma_1^1))$. Therefore, $(a_{\mathbf{n},0}, \ldots, a_{\mathbf{n},m})$ is a type I multi-point Hermite–Padé approximation of $(\widehat{s}_{1,1}, \ldots, \widehat{s}_{1,m})$ with respect to $w_{\mathbf{n}}$ and the results of this paper may be applied.

References

[1] J. Bustamante, G. López Lagomasino, Hermite–Padé approximation for Nikishin systems of analytic functions, Sb. Math. 77 (1994) 367–384.
[2] T. Carleman, Les fonctions quasi-analytiques, Gauthier–Villars, Paris, 1926.
[3] K. Driver, H. Stahl, Normality in Nikishin systems, Indag. Math. (N.S.) 5 (1994) 161–187.
[4] K. Driver, H. Stahl, Simultaneous rational approximants to Nikishin systems. I, Acta Sci. Math. (Szeged) 60 (1995) 245–263.
[5] K. Driver, H. Stahl, Simultaneous rational approximants to Nikishin systems. II, Acta Sci. Math. (Szeged) 61 (1995) 261–284.
[6] U. Fidalgo, G. López Lagomasino, Rate of convergence of generalized Hermite–Padé approximants of Nikishin systems, Constr. Approx. 23 (2006) 165–196.
[7] U. Fidalgo, G. López Lagomasino, General results on the convergence of multi-point Hermite–Padé approximants of Nikishin systems, Constr. Approx. 25 (2007) 89–107.

[8] U. Fidalgo, A. López, G. López Lagomasino, V.N. Sorokin, Mixed type multiple orthogonal polynomials for two Nikishin systems, Constr. Approx. 32 (2010) 255–306.

[9] U. Fidalgo, G. López Lagomasino, Nikishin systems are perfect, Constr. Approx. 34 (2011) 297–356.

[10] U. Fidalgo, G. López Lagomasino, Nikishin systems are perfect. Case of unbounded and touching supports, J. Approx. Theory 163 (2011) 779–811.

[11] A.A. Gonchar, On the convergence of generalized Padé approximants of meromorphic functions, Sb. Math. 27 (1975) 503–514.

[12] A.A. Gonchar, E.A. Rakhmanov, V.N. Sorokin, On Hermite–Padé approximants for systems of functions of Markov type, Sb. Math. 188 (1997) 671–696.

[13] M.G. Krein, A.A. Nudel'man, The Markov Moment Problem and Extremal Problems, Transl. Math. Monogr., vol. 50, Amer. Math. Soc., Providence, RI, 1977.

[14] A.B.J. Kuijlaars, Multiple orthogonal polynomial ensembles, in: Contemp. Math., vol. 507, 2010, pp. 155–176.

[15] A.A. Markov, Deux démonstrations de la convergence de certains fractions continues, Acta Math. 19 (1895) 93–104.

[16] E.M. Nikishin, On simultaneous Padé approximants, Math. USSR Sb. 41 (1982) 409–425.

[17] E.M. Nikishin, Asymptotics of linear forms for simultaneous Padé approximants, Soviet Math. (Izv. VUZ) 30 (1986) 43–52.

[18] E.M. Nikishin, V.N. Sorokin, Rational Approximations and Orthogonality, Amer. Math. Soc., Providence, RI, 1991.

[19] E.A. Rakhmanov, S.P. Suetin, Asymptotic behaviour of the Hermite–Padé polynomials of the 1st kind for a pair of functions forming a Nikishin system, Russian Math. Surveys 67 (2012) 954–956.

[20] E.A. Rakhmanov, S.P. Suetin, The distribution of the zeros of the Hermite–Padé polynomials for a pair of functions forming a Nikishin system, Sb. Math. 204 (2013) 1347–1390, http://dx.doi.org/10.1070/SM2013v204n09ABEH004343.

[21] H. Stahl, V. Totik, General Orthogonal Polynomials, Cambridge University Press, Cambridge, 1992.

[22] T.J. Stieltjes, Recherches sur les fractions continues, Ann. Fac. Sci. Univ. Toulouse 8 (1894) J1–J122, 9 (1895) A1–A47, reprinted in his Oeuvres Complètes, Tome 2, Noordhoff, 1918, pp. 402–566.

[23] W. Van Assche, Analytic number theory and rational approximation, in: A. Branquinho, A. Foulquié (Eds.), Coimbra Lecture Notes on Orthogonal Polynomials, Nova Science Publ., New York, 2008, pp. 197–229.