ON THE CONVERGENCE OF RATIONAL RITZ VALUES*

BERNHARD BECKERMANN†, STEFAN GÜTTEL‡, AND RAF VANDEBRIL§

Abstract. Ruhe's rational Krylov method is a popular tool for approximating eigenvalues of a given matrix, though its convergence behavior is far from being fully understood. Under fairly general assumptions we characterize in an asymptotic sense which eigenvalues of a Hermitian matrix are approximated by rational Ritz values and how fast this approximation takes place. Our main tool is a constrained extremal problem from logarithmic potential theory, where an additional external field is required for taking into account the poles of the underlying rational Krylov space. Several examples illustrate our analytic results.

Key words. Rational Krylov, Ritz values, orthogonal rational functions, logarithmic potential theory.

AMS subject classifications. 15A18, 31A05, 31A15, 65F15

1. Introduction. In order to approximate parts of the spectrum $\Lambda(A)$ of a Hermitian matrix $A \in \mathbb{C}^{N\times N}$, a widely used approach is to project A onto an n-dimensional subspace of $\mathbb{C}^N$, with n being small compared to N. Given a matrix $V_n \in \mathbb{C}^{N\times n}$ with orthonormal columns, the eigenvalues of the projected counterpart $V_n^* A V_n \in \mathbb{C}^{n\times n}$ are called *Ritz values of order n*. These Ritz values are often good approximations to some of A's eigenvalues, depending on the space spanned by the columns of V_n.

A well-studied case consists of projecting A onto a (polynomial) Krylov space

$$\mathcal{K}_n(A,\mathbf{b}) = \mathrm{span}\{\mathbf{b}, A\mathbf{b}, \ldots, A^{n-1}\mathbf{b}\}$$

for a given starting vector $\mathbf{b} \in \mathbb{C}^N$. Here the so-called *polynomial Ritz values* typically approximate extremal eigenvalues of A, although counter-examples can be constructed, see, e.g., [23, Section 7].

In a more general Krylov method suggested by Ruhe [34] and further analyzed by him [35, 36] and other authors [12, 13, 19, 27, 29] the matrix A is projected onto a rational Krylov space

$$\mathcal{K}_n^{\mathrm{rat}}(A,\mathbf{b}) = q_{n-1}(A)^{-1}\mathcal{K}_n(A,\mathbf{b}), \quad q_{n-1}(z) = \prod_{\substack{j=1\\ \xi_j\neq\infty}}^{n-1} (z-\xi_j). \tag{1.1}$$

The numbers $\xi_j \in (\mathbb{R}\cup\{\infty\}) \setminus \Lambda(A)$ are referred to as the *poles* of the rational Krylov space. These poles are free parameters which can be chosen to amplify interesting parts of $\Lambda(A)$. In general, we expect that the *rational Ritz values* approximate eigenvalues in proximity of the poles.

*The work of the second author was partially supported by the Deutsche Forschungsgemeinschaft. The third author has a grant as "Postdoctoraal Onderzoeker" from the Fund for Scientific Research–Flanders (Belgium).

†Université des Sciences et Technologies de Lille, Laboratoire Painlevé UMR 8524, UFR Mathématiques, F-59655 Villeneuve d'Ascq CEDEX, France (bbecker@math.univ-lille1.fr).

‡Technische Universität Bergakademie Freiberg, Fakultät für Mathematik und Informatik, D-09596 Freiberg, Germany (stefan.guettel@math.tu-freiberg.de).

§Université des Sciences et Technologies de Lille, Laboratoire Painlevé UMR 8524, UFR Mathématiques, F-59655 Villeneuve d'Ascq CEDEX, France (raf.vandebril@cs.kuleuven.ac.be).

In what follows we assume that all eigenvalues $\lambda_1 < \cdots < \lambda_N$ of A be distinct. By Θ we denote the set of nth rational Ritz values $\theta_1 < \cdots < \theta_n$, and by Ξ we denote the multiset of poles $\xi_1, \ldots, \xi_{n-1}$. The goal of the present paper is to obtain new asymptotic upper bounds for the distance $\mathrm{dist}(\lambda_k, \Theta)$ of a given eigenvalue λ_k of A to the set of rational Ritz values, which allows us to quantify in an asymptotic sense which eigenvalues are well approached by rational Ritz values. Our results generalize recent findings of Kuijlaars [23] for the particular case of polynomial Ritz values, in particular we describe how the distribution of the rational Ritz values depends both on the spectrum and the poles.

Our statements are based on asymptotic analysis, which is not possible if only a single matrix A is considered. Therefore, all our results in §3 are formulated in terms of sequences of matrices $A_N \in \mathbb{C}^{N\times N}$ having a joint eigenvalue distribution described by the measure σ, and similarly we will presume a joint pole distribution ν. Note that sequences of matrices having a joint eigenvalue distribution occur quite frequently in applications. The most prominent examples are finite sections of a Toeplitz operator, see for instance [8]. Matrices obtained by finite difference or finite element discretization of PDEs with varying mesh width also have a joint eigenvalue distribution, see, e.g., [4] and the references therein. Even if our results are of an asymptotic nature, there is numerical evidence that the phenomena described here also occur for finite N, at least if N is sufficiently large, see §4.

1.1. Asymptotic distribution of eigenvalues and Ritz values. In order to describe our main findings on rational Ritz values, let us first recall some recent asymptotic results on polynomial Ritz values. There is a rule of thumb proposed by Trefethen and Bau [39] that $\mathrm{dist}(\lambda_k, \Theta)$ is small for eigenvalues λ_k lying in regions of the real line where there are "relatively few" eigenvalues. It was Kuijlaars [23] who first quantified this heuristic rule, and we also refer to the refinements given in an unpublished note [1] and some related work on isometric Ritz values [21]. Suppose that the asymptotic eigenvalue distribution is described by a finite positive Borel measure σ. Under mild assumptions stated explicitly in §3 below, Kuijlaars showed that the distribution of the polynomial Ritz values of order n is described by a measure μ which solves an extremal problem from logarithmic potential theory. More precisely, μ is the measure of total mass $t = n/N$ having minimal logarithmic energy

$$I(\mu) = I(\mu,\mu), \quad I(\mu_1,\mu_2) = \iint \log\frac{1}{|x-y|}\,\mathrm{d}\mu_1(x)\,\mathrm{d}\mu_2(y)$$

among all measures $\mu_1 \geq 0$ of total mass t satisfying the constraint $\mu_1 \leq \sigma$. Here the condition $\mu \leq \sigma$ comes from the fact that, in any interval, the number of Ritz values does not exceed the number of eigenvalues by more than one. This last property is an immediate consequence of the so-called *interlacing property* (cf. [31, Theorem 10.1.1]):

$$\text{In each open interval } (\theta_j, \theta_{j+1}) \text{ there is at least one eigenvalue of } A. \tag{1.2}$$

We conclude that in parts of the real line where the constraint $\mu \leq \sigma$ is active, there are asymptotically as many Ritz values as eigenvalues. Kuijlaars also showed that eigenvalues lying in a neighborhood of a point z with the logarithmic potential

$$U^\mu(z) = \int \log\frac{1}{|x-z|}\,\mathrm{d}\mu(x)$$

being strictly less than the maximum F of U^μ in the complex plane are approximated by Ritz values with a geometric rate. Typically, the set of such points z is given by $\mathbb{R} \setminus \operatorname{supp}(\sigma - \mu)$, which is outside the set where the constraint $\mu \leq \sigma$ is not active, see Remark A.2.

The goal of the present paper is to generalize the above two results to rational Ritz values. Notice that

$$q_{n-1}(A)^{-1}\mathcal{K}_n(A, \mathbf{b}) = \mathcal{K}_n(A, q_{n-1}(A)^{-1}\mathbf{b}),$$

i.e., every rational Krylov space is also a polynomial Krylov space for the modified starting vector $q_{n-1}(A)^{-1}\mathbf{b}$. This is the reason why rational Ritz values inherit many properties from polynomial Ritz values, in particular the interlacing property (1.2). We will assume that the asymptotic distribution of poles is given by some measure ν of total mass $\|\nu\| = t$, and set

$$\mathcal{M}_t^\sigma = \{\mu_1 \text{ Borel measure} : \mu_1 \geq 0, \mu_1 \leq \sigma, \|\mu_1\| = t\},$$

where again $t = n/N$. Then, under additional mild assumptions which are similar to those of Kuijlaars and stated explicitly in §3, we first show that the asymptotic distribution of the nth rational Ritz values is given by the unique measure $\mu \in \mathcal{M}_t^\sigma$ which is "closest" to ν. Here the distance is measured in terms of the energy of a signed measure

$$I(\mu - \nu) = I(\mu) - 2I(\mu, \nu) + I(\nu) \geq 0.$$

It is known from potential theory [38, Example II.4.8] that, for sufficiently large σ, μ would be just the balayage of ν onto $\operatorname{supp}(\sigma)$. In our case the situation is more complicated because of the constraint $\mu \leq \sigma$. However, we may again conclude that in real intervals I where the constraint $\mu \leq \sigma$ is active there are asymptotically as many Ritz values as eigenvalues, and this is in particular true by Lemma A.1(e) provided that $\nu - \sigma$ is positive on I, i.e., roughly speaking, the number of poles exceeds the number of eigenvalues in I.

Secondly, we will generalize Kuijlaars's findings on the rate of convergence to rational Ritz values: we show that eigenvalues close to a point z with $U^{\mu-\nu}(z)$ being strictly less than the maximum F of $U^{\mu-\nu}$ on $\mathbb{C}$ are approached by rational Ritz values with an explicit geometric rate, and this typically happens if the constraint $\mu \leq \sigma$ is active in a neighborhood of z, i.e., $z \notin \operatorname{supp}(\sigma - \mu)$. We refer the reader to §3 for the precise statements of our results and a discussion of our assumptions.

Polynomial Ritz values approaching eigenvalues may be rewritten as zeros of discrete orthogonal polynomials approaching the discrete support of the measure of orthogonality. Thus the findings of [1, 21, 23] are related to results from the late last century about weak asymptotics of discrete polynomials due to Rakhmanov [32] and Dragnev & Saff [14], Van Assche & Kuijlaars [26], Beckermann [2], and others [10, 11, 25]. In the present paper we use the well-known fact [13] that rational Ritz values may be written as zeros of discrete orthogonal rational functions [9]. To our knowledge, weak asymptotics of such discrete rational orthogonal functions have not been published elsewhere, but, as we will see, they follow by considering the numerator as a discrete orthogonal polynomial, and incorporating the denominator in a varying weight.

1.2. Structure of the paper and notation. In §2 we shortly recall some basics about polynomial and rational Krylov spaces and make the link with discrete orthogonal polynomials and discrete orthogonal rational functions. The precise statements of our results are given in §3, mainly in Theorem 3.1, Theorem 3.2, and Corollary 3.3. In §4 we illustrate our findings by discussing one analytic example and two numerical experiments. To keep the exposition easy to read we decided to present the main proofs in §5. We have added an Appendix A with some tools from potential theory, which are needed for the main proofs. Though our findings are proved using logarithmic potential theory, the reader does not need to be an expert in this field in order to understand the main statements. We refer the interested reader to the introductions [15, 28] or to the monographs [33, 38] for further details on potential theory.

Throughout this paper we assume exact arithmetic and thus effects like rounding errors and loss of orthogonality are not considered. If not otherwise stated, $\langle \cdot, \cdot \rangle$ refers to the standard scalar product in $\mathbb{C}^N$ and $\| \cdot \|$ is the induced norm.

2. Krylov spaces.

2.1. Polynomial Krylov sequences. In the polynomial Krylov approach a matrix $V_n \in \mathbb{C}^{N\times n}$ is constructed with columns $\mathbf{v}_1, \ldots, \mathbf{v}_n$ forming an orthonormal basis for the Krylov space $\mathcal{K}_n(A, \mathbf{b})$. Starting from $\mathbf{v}_1 = \mathbf{b}/\|\mathbf{b}\|$, the vectors $\mathbf{v}_{j+1}$ are constructed iteratively by the so-called *Arnoldi process*, by orthonormalizing the vector $A\mathbf{v}_j$ against the already known orthonormal vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_j$. This leads to the equations

$$A\mathbf{v}_j = h_{1,j}\mathbf{v}_1 + \cdots + h_{j,j}\mathbf{v}_j + h_{j+1,j}\mathbf{v}_{j+1},$$

which for $j = 1, \ldots, n$ may be rewritten in matrix language as an *Arnoldi decomposition*

$$AV_n = V_n H_n + h_{n+1,n}\mathbf{v}_{n+1}\mathbf{e}_n^T,$$

where $V_n = [\mathbf{v}_1, \ldots, \mathbf{v}_n]$ and $H_n = [h_{i,j}]$ is an $n \times n$ unreduced upper Hessenberg matrix. This Arnoldi decomposition provides readily the projected counterpart $V_n^* AV_n = H_n$, whose eigenvalues are the polynomial Ritz values of order n. In the present paper, A is Hermitian and hence the Arnoldi process reduces to the more economical Lanczos process.

There is a close relation between discrete orthogonal polynomials and the Lanczos process. By construction, there exist polynomials p_j of exact degree j such that

$$\mathbf{v}_{j+1} = p_j(A)\mathbf{b}/\|\mathbf{b}\|. \tag{2.1}$$

Defining the normalized eigencomponents by

$$w(\lambda_k) = |\langle \mathbf{z}_k, \mathbf{b}/\|\mathbf{b}\|\rangle| = |\langle \mathbf{z}_k, \mathbf{v}_1\rangle| \in [0, 1], \tag{2.2}$$

$\mathbf{z}_k$ being a normalized eigenvector for λ_k ($k = 1, \ldots, N$), and the scalar product

$$\langle p, q\rangle := \langle p(A)\mathbf{v}_1, q(A)\mathbf{v}_1\rangle = \sum_{k=1}^{N} w(\lambda_k)^2 p(\lambda_k)\overline{q(\lambda_k)}, \tag{2.3}$$

we find using (2.1) that $\langle p_i, p_j\rangle = \langle \mathbf{v}_i, \mathbf{v}_j\rangle = \delta_{i,j}$. In other words, p_j is the jth orthogonal polynomial with respect to the discrete scalar product (2.3). Also, one can easily prove that the zeros of p_n coincide with the polynomial Ritz values of order n.

2.2. Rational Krylov sequences. In [34], a rational Krylov method was presented as an extension of the shift-and-invert Arnoldi process allowing for varying shifts. This method recursively computes an orthonormal basis of the rational Krylov spaces defined in (1.1). Consider a multiset of poles $\Xi = \{\xi_1, \ldots, \xi_n\} \subset (\mathbb{R} \cup \{\infty\}) \setminus \Lambda(A)$. Starting from $\mathbf{v}_1 = \mathbf{b}/\|\mathbf{b}\|$, each vector $\mathbf{v}_{j+1}$ is obtained by orthonormalizing the vector $A(I - A/\xi_j)^{-1}\mathbf{v}_j$ against the already known orthonormal vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_j$. This leads to the equations

$$A(I - A/\xi_j)^{-1}\mathbf{v}_j = h_{1,j}\mathbf{v}_1 + \cdots + h_{j,j}\mathbf{v}_j + h_{j+1,j}\mathbf{v}_{j+1},$$

which for $j = 1, \ldots, n$ may be rewritten in matrix language as a *rational Arnoldi decomposition*

$$AV_n(H_nD_n + I_n) + h_{n+1,n}\xi_n^{-1}A\mathbf{v}_{n+1}\mathbf{e}_n^T = V_nH_n + h_{n+1,n}\mathbf{v}_{n+1}\mathbf{e}_n^T, \tag{2.4}$$

where $V_n = [\mathbf{v}_1, \ldots, \mathbf{v}_n]$, $H_n = [h_{i,j}]$ is an $n \times n$ unreduced upper Hessenberg matrix, $h_{n+1,n} \in \mathbb{C}$, $D_n = \operatorname{diag}(1/\xi_1, \ldots, 1/\xi_n)$ and I_n is the identity matrix of size $n \times n$. Note that if all poles $\xi_j = \infty$ then $D_n = O$ and (2.4) reduces to a standard Arnoldi decomposition.

It follows from (1.1) that we may write

$$\mathbf{v}_{j+1} = r_j(A)\mathbf{b}/\|\mathbf{b}\|, \quad r_j = p_j/q_j, \quad \deg p_j \leq j, \tag{2.5}$$

and the orthogonality of the basis vectors leads to $\langle r_k, r_\ell \rangle = \delta_{k,\ell}$ with the scalar product as in (2.3), that is, r_j is the jth orthogonal rational function with respect to a discrete scalar product. However, as pointed out, e.g., by Deckers and Bultheel [13], the nth rational Ritz values in general are no longer the zeros of r_n since the latter depend on ξ_n, but V_n and the projected matrix $V_n^*AV_n$ do not depend on the last pole ξ_n. In order to derive a simple formula for $V_n^*AV_n$, we put $\xi_n = \infty$ in (2.4) and obtain the modified rational Arnoldi decomposition

$$AV_n(\widetilde{H}_n\widetilde{D}_n + I_n) = V_n\widetilde{H}_n + \widetilde{h}_{n+1,n}\widetilde{\mathbf{v}}_{n+1}\mathbf{e}_n^T, \tag{2.6}$$

where $\widetilde{H}_n, \widetilde{D}_n$ are obtained from H_n, D_n by adapting the last column (in particular these new matrices do no longer form a nested sequence) and we have a modified vector $\widetilde{\mathbf{v}}_{n+1}$ being orthogonal to $\mathbf{v}_1, \ldots, \mathbf{v}_n$, which as in (2.5) can be written as

$$\widetilde{\mathbf{v}}_{n+1} = \widetilde{r}_n(A)\mathbf{b}/\|\mathbf{b}\|, \quad \widetilde{r}_n = \widetilde{p}_n/q_{n-1}, \quad \deg \widetilde{p}_n \leq n. \tag{2.7}$$

Note that from (2.6) we obtain the simple formula

$$V_n^*AV_n = \widetilde{H}_n(\widetilde{H}_n\widetilde{D}_n + I_n)^{-1}.$$

Moreover, by writing (2.6) in terms of rational functions

$$z[r_0(z), \ldots, r_{n-1}(z)](\widetilde{H}_n\widetilde{D}_n + I_n) = [r_0(z), \ldots, r_{n-1}(z)]\widetilde{H}_n + \widetilde{h}_{n+1,n}[0, \ldots, 0, \widetilde{r}_n(z)],$$

we observe that θ is a zero of $\widetilde{r}_n$ if and only if it is an eigenvalue of $V_n^*AV_n$. In other words, the set of nth rational Ritz values is the set of zeros of $\widetilde{r}_n$.

REMARK 2.1. *The recursive construction above does not allow for poles at zero. This is no restriction since one can consider a shifted matrix $A - \tau I$ and construct the rational Krylov space with shifted poles $\xi_j - \tau$ (see, e.g., [27]).*

The construction of the orthonormal basis vector $\mathbf{v}_{j+1}$ *from* $r_{j-1}(A)\mathbf{b}/\|\mathbf{b}\|$ *may break down if, by chance,* $r_{j-1}(\xi_j) = 0$. *Such break-downs can be avoided by using instead of* r_{j-1} *a linear combination of* $r_0, \dots, r_{j-1}$ *which has no zero at* ξ_j *(see, e.g., [34]).*

REMARK 2.2. *It is well-known that for Hermitian matrices* A *the Lanczos process generates a tridiagonal matrix* $T_n = V_n^* A V_n$. *In case of all poles being different from infinity it was proven in [19] that the matrix* T_n *is no longer tridiagonal but of semiseparable plus diagonal form, in which the diagonal consists of the poles, i.e.,* $T_n = S_n + \Delta_n^{-1}$ *with* $\Delta_n = \mathrm{diag}(1, 1/\xi_1, 1/\xi_2, \dots, 1/\xi_{n-1})$, *where* S_n *is semiseparable*[1].

In the general case with some of the poles being equal to infinity, it is possible to show that the the projected counterpart T_n *is a block-diagonal matrix (with the blocks overlapping the top and bottom diagonal elements) being either of semiseparable plus diagonal or tridiagonal form (see, e.g., [40, Section 1.2.5]).*

In order to compute rational Ritz values it is essential to be able to compute the eigenvalues of the projected counterpart fast enough. This coincides with computing eigenvalues of a Hermitian quasiseparable matrix. Fast $O(n^2)$*–algorithms for computing these eigenvalues can be found, e.g., in [17, 18, 41].*

2.3. Bounding the distance via a polynomial extremal problem. Classical results on the convergence of Ritz values can be found in several textbooks [20, 30, 37, 39]. Many of them are derived by exploiting the relation between polynomials and Krylov spaces, where an important ingredient for estimating the distance of an eigenvalue to the set of Ritz values is a link to some polynomial extremal problem. Typically, such procedures are used to handle extremal eigenvalues or outliers, but, as shown for instance in [1, Lemma 2.2], this approach is also useful for detecting eigenvalues in other parts of the spectrum. Let us prove here an extension of [1, Lemma 2.2] to the rational case, which will be the basic tool in establishing our main results.

LEMMA 2.3. *Consider the polynomials* $\widetilde{p}_n$ *and* q_{n-1} *defined in* (2.7), *and the eigencomponents* $w(\lambda_j)$ *defined in* (2.2).

If $\lambda_k \le \theta_1$ *then*

$$\theta_1 - \lambda_k = \min \left\{ \frac{\sum_{j=1, j\neq k}^{N} \frac{w(\lambda_j)^2}{q_{n-1}(\lambda_j)^2} (\lambda_j - \theta_1)\, s(\lambda_j)^2}{\frac{w(\lambda_k)^2}{q_{n-1}(\lambda_k)^2} s(\lambda_k)^2} \;\middle|\; \deg(s) < n,\ s(\lambda_k) \neq 0 \right\}.$$

The minimum is attained for $s(x) = \widetilde{p}_n(x)/(x - \theta_1)$.

Suppose $\lambda_k \in [\theta_{\kappa-1}, \theta_\kappa]$, *then*

$$\begin{aligned} &(\lambda_k - \theta_{\kappa-1})\,(\theta_\kappa - \lambda_k) \\ &= \min \left\{ \frac{\sum_{j=1, j\neq k}^{N} \frac{w(\lambda_j)^2}{q_{n-1}(\lambda_j)^2} (\lambda_j - \theta_{\kappa-1})\,(\lambda_j - \theta_\kappa)\, s(\lambda_j)^2}{\frac{w(\lambda_k)^2}{q_{n-1}(\lambda_k)^2} s(\lambda_k)^2} \;\middle|\; \deg(s) < n-1,\ s(\lambda_k) \neq 0 \right\}. \end{aligned}$$

The minimum is attained for $s(x) = \widetilde{p}_n(x)/\left((x - \theta_{\kappa-1})(x - \theta_\kappa)\right)$.

Proof. We only prove the second case $\lambda_k \in [\theta_{\kappa-1}, \theta_\kappa]$, a proof for the other case is similar. First recall from §2.2 that the nth rational Ritz values θ_j are the zeros of the numerator $\widetilde{p}_n$ of the rational function $\widetilde{r}_n$ defined in (2.7). We claim

[1] In case that S_n is nonsingular its inverse is a tridiagonal matrix. A semiseparable matrix is characterized by the fact that all submatrices taken out of the part below and including the diagonal are of rank at most one.

that this numerator is in fact an nth discrete orthogonal polynomial. Consider the polynomials $\widehat{q}_\ell := q_{n-1}/q_\ell$. It follows from (1.1) that the rational functions $r_\ell = p_\ell/q_\ell = (p_\ell \widehat{q}_\ell)/q_{n-1}$ span the space $\mathbb{P}_{n-1}/q_n$ ($\ell = 0, 1, \ldots, n-1$), and hence the polynomials $p_\ell \widehat{q}_\ell$ of degree $\leq n-1$ span the space $\mathbb{P}_{n-1}$. Using (2.5), (2.7) and the orthogonality of the vectors $\mathbf{v}_j$ we have

$$\begin{aligned}
0 &= \langle \widetilde{\mathbf{v}}_{n+1}, \mathbf{v}_{\ell+1} \rangle \\
&= \langle q_{n-1}(A)^{-1} \widetilde{p}_n(A)\mathbf{v}_1, q_\ell(A)^{-1} p_\ell(A)\mathbf{v}_1 \rangle \\
&= \langle q_{n-1}(A)^{-1} \widetilde{p}_n(A)\mathbf{v}_1, q_{n-1}(A)^{-1} (p_\ell(A)\widehat{q}_\ell(A))\mathbf{v}_1 \rangle \\
&= \sum_{j=1}^{N} \frac{w(\lambda_j)^2}{q_{n-1}(\lambda_j)^2} \widetilde{p}_n(\lambda_j) \left(p_\ell(\lambda_j)\widehat{q}_\ell(\lambda_j) \right),
\end{aligned}$$

hence $\widetilde{p}_n \perp \mathbb{P}_{n-1}$ for this modified discrete scalar product.

Gaussian quadrature provides us with the existence of some weights $\rho_1, \ldots, \rho_n$ such that

$$\sum_{j=1}^{N} \frac{w(\lambda_j)^2}{q_{n-1}(\lambda_j)^2} \hat{s}(\lambda_j) = \sum_{j=1}^{n} \rho_j^2 \hat{s}(\theta_j), \tag{2.8}$$

for all polynomials $\hat{s}$ of degree at most $2n-1$. Taking a polynomial s of degree less than $n-1$ with $s(\lambda_k) \neq 0$ and setting $\hat{s}(x) = (x - \theta_{\kappa-1})(x - \theta_\kappa)s(x)^2$, the right-hand side of (2.8) is positive, and hence

$$(\lambda_k - \theta_{\kappa-1})(\theta_\kappa - \lambda_k) \leq \frac{\sum_{j=1, j\neq k}^{N} \frac{w(\lambda_j)^2}{q_{n-1}(\lambda_j)^2} (\lambda_j - \theta_{\kappa-1})(\lambda_j - \theta_\kappa) s(\lambda_j)^2}{\frac{w(\lambda_k)^2}{q_{n-1}(\lambda_k)^2} s(\lambda_k)^2} .$$

For $s(x) = \widetilde{p}_n(x)/((x - \theta_{\kappa-1})(x - \theta_\kappa))$ the right-hand side of (2.8) is zero and equality is obtained in the above estimate. □

In order to give a better understanding of the potential impact of Lemma 2.3, let us have a closer look at the first part for polynomial Ritz values (i.e., $q_n = 1$). Since all Ritz values lie in the open interval (λ_1, λ_N), we may choose $k = 1$, and get for $\mathrm{dist}(\lambda_1, \Theta) = \theta_1 - \lambda_1$ the upper bound

$$\mathrm{dist}(\lambda_1, \Theta) \leq \sum_{j=2}^{N} \frac{|\lambda_N - \lambda_1| w(\lambda_j)^2}{w(\lambda_1)^2} \frac{\max_{j=2,\ldots,N} |s(\lambda_j)|^2}{|s(\lambda_1)|}$$

for any polynomial s of degree at most $n-1$. More explicit upper bounds can be obtained by choosing s taking the value 1 at λ_1 and being small on the convex hull of all other eigenvalues, leading to the well-known Kaniel-Page-Saad estimate for extremal eigenvalues [20, 30, 37, 39]. This construction is similar to the one in the proof of the classical convergence bound for the CG method, which predicts linear convergence in terms of the condition number of A: here the spectrum is also replaced by its convex hull. However, for bounding $\mathrm{dist}(\lambda_1, \Theta)$ it is only necessary that s is small on the discrete set $\{\lambda_2, \ldots, \lambda_N\}$. The optimal polynomials for both tasks can look quite different, see Figure 2.1 for a simple example. Therefore a precise upper bound of $\mathrm{dist}(\lambda_k, \Theta)$ needs to incorporate the fine structure of the spectrum, see [1,21,23]. This fine structure also explains the superlinear convergence behavior of the CG method (see [5–7] and the reviews [24] and [3]).

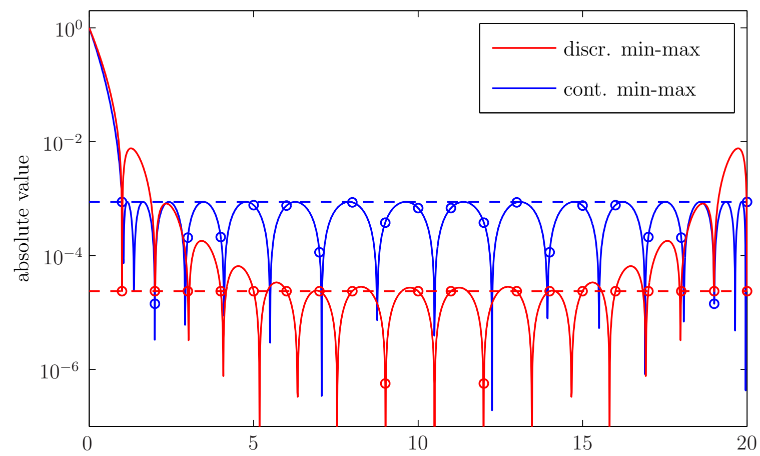

Fig. 2.1. *The absolute value of two polynomials of degree* 17 *taking value* 1 *at* 0*. One polynomial (solid red) is minimal on the discrete set* $\{1, 2, \ldots, 20\}$*, and the other polynomial (dashed blue) is minimal on the interval* $[1, 20]$.

3. Statement of the main results. Following Kuijlaars and his successors [1,5–7,21,23], we will consider a sequence of Hermitian matrices $A_N \in \mathbb{C}^{N\times N}$ having a joint eigenvalue distribution described by some measure σ: write more explicitly the set $\Lambda_N = \Lambda(A_N)$ of eigenvalues $\lambda_{1,N} < \cdots < \lambda_{N,N}$ of A_N, and consider the normalized counting measure

$$\int f(x)\, \mathrm{d}\chi_N(\Lambda_N)(x) := \frac{1}{N} \sum_{x\in\Lambda_N} f(x), \quad f \in \mathcal{C}(\mathbb{R}),$$

that is, the normalized sum of mass points δ_x. We then ask that the sequence $(\chi_N(\Lambda_N))$ has the weak star limit σ, written shorter $\chi_N(\Lambda_N) \to \sigma$, where we recall that, for a sequence of measures σ_n with supports included in some compact real interval, the relation $\sigma_n \to \sigma$ means that $\int f\, \mathrm{d}\sigma_n \to \int f\, \mathrm{d}\sigma$ for all $f \in \mathcal{C}(\mathbb{R})$. Our definition of normalized counting measures naturally extends to multisets, but here we will count each element according to its multiplicity. This will be important for the pole counting measures only since the eigenvalues and Ritz values are distinct, anyway.

From now on we adapt our notation to the following conventions:

We add an index N to all our quantities. More precisely, we consider
- Hermitian matrices $A_N \in \mathbb{C}^{N\times N}$ with distinct eigenvalues $\lambda_{1,N} < \cdots < \lambda_{N,N}$, spectra $\Lambda_N = \Lambda(A_N)$,
- starting vectors $\mathbf{b}_N \in \mathbb{C}^N$ with eigencomponents $w_N(\lambda_{j,N}) \in [0, 1]$,
- a multiset $\Xi_N \subset \mathbb{R} \setminus \Lambda_N$ of the $n-1$ poles $\xi_{1,N}, \ldots, \xi_{n-1,N}$, and
- a set Θ_N of nth rational Ritz values $\theta_{1,N} < \cdots < \theta_{n,N}$ for $(A_N, \mathbf{b}_N)$ and the poles Ξ_N.

Here $n = n(N)$ will always be chosen such that $n(N)/N \to t \in (0, \|\sigma\|)$ as $N \to \infty$, where $\|\sigma\| = \sigma(\mathbb{C})$ is the total mass of the positive measure σ.

In order to formulate the precise statements, we first specify and motivate the necessary assumptions:

(H1) The spectra and pole sets are uniformly bounded: there exist compact sets Λ and Ξ such that for all N there holds $\Lambda_N \subset \Lambda$ and $\Xi_N \subset \Xi$.

(H2) The matrices A_N have an asymptotic eigenvalue distribution described by some measure σ: we have $\chi_N(\Lambda_N) \to \sigma$ for $N \to \infty$.

(H3) We have a weak separation of eigenvalues: for any sequence $\Lambda_N \ni \lambda_{k(N),N} \to \lambda$ for $N \to \infty$ there holds

$$\limsup_{\delta\to 0+} \limsup_{N\to\infty} \frac{1}{N} \sum_{0<|\lambda_{j,N}-\lambda_{k(N),N}|\le\delta} \log \frac{1}{|\lambda_{j,N} - \lambda_{k(N),N}|} = 0.$$

It follows (see Lemma A.4 below) that $z \mapsto U^\sigma(z)$ is continuous.

(H4) The multisets of poles Ξ_N counting multiplicities have an asymptotic behavior described by some measure ν: we have $\chi_N(\Xi_N) \to \nu$ for $N \to \infty$.

(H5) The eigencomponents $w_N(\lambda_{k,N}) \in [0, 1]$ defined in (2.2) are sufficiently large

$$\liminf_{N\to\infty} \min_k w_N(\lambda_{k,N})^{1/N} = 1.$$

(H6) We have a strict separation of poles from eigenvalues: $\Lambda \cap \Xi$ is empty. It follows from Assumptions **(H1)** and **(H4)** that U^ν is continuous on Λ.

Conditions **(H1)**, **(H2)** and **(H4)** are required to define our asymptotic setting. The other conditions are essential to obtain interesting bounds for $\mathrm{dist}(\lambda_k, \Theta)$ from Lemma 2.3. For instance, in accordance with **(H5)**, one should impose that the eigencomponent $w(\lambda_k) \in [0, 1]$ is not "too small". Also, the condition **(H6)** will be convenient in order to understand the role of the denominators q_n in Lemma 2.3.

Conditions **(H1)**, **(H2)**, **(H3)** and **(H5)** were also used by Kuijlaars [23] in his study of polynomial Ritz values. The rather technical condition **(H3)**, first suggested in [14], prevents eigenvalues from clustering exponentially close for increasing N. This condition allows for equidistant eigenvalues or Chebyshev eigenvalues (i.e., the eigenvalues of the discretized 1D-Laplacian), but also more general sets of points [14]. The continuity of U^σ together with the Lemma of Rakhmanov [32] implies that U^ρ is continuous and thus $I(\rho)$ is finite for each $\rho \in \mathcal{M}_t^\sigma$.

Generalizing the work [23] of Kuijlaars, we have the following main findings.

THEOREM 3.1. *Under the Assumptions* ***(H1)–(H6)****, the* $n(N)$*th Ritz values have an asymptotic distribution described by* $\chi_N(\Theta_N) \to \mu$*, with the positive finite Borel measure* μ *being the unique minimizer of* $\mu_1 \mapsto I(\mu_1) - 2I(\nu, \mu_1)$ *within* $\mathcal{M}_t^\sigma$*.*

Define F *as the maximum of* $U^{\mu-\nu}$ *in the whole complex plane, and* $\Sigma_t^* = \{z \in \mathbb{C} : U^{\mu-\nu}(z) = F\}$*. In a closed interval* $J \subset \mathbb{R} \setminus \Sigma_t^* \subset \mathbb{R} \setminus \mathrm{supp}(\sigma - \mu)$*, all sequences* $J \ni \lambda_{k(N),N} \to \lambda$ *for* $N \to \infty$ *satisfy*

$$\lim_{N\to\infty} \mathrm{dist}(\lambda_{k(N),N}, \Theta_N)^{1/N} = \exp(2(U^{\mu-\nu}(\lambda) - F)),$$

with the possible exclusion of at most one unique "exceptional index" $k^*(N)$*.*

Under the above assumptions on σ, ν, the existence and uniqueness of a minimizer of $\mu \mapsto I(\mu) - 2I(\mu, \nu)$ within $\mathcal{M}_t^\sigma$ is shown in Lemma A.1(a). Provided that $I(\nu)$ is finite, we may write $I(\mu)-2I(\mu, \nu) = I(\mu-\nu)-I(\nu)$, and recall from [38, Lemma I.1.8] that $I(\mu-\nu) \ge 0$, with equality if and only if $\mu = \nu$. Hence, Theorem 3.1 tells us that, under the Assumptions **(H1)–(H6)**, the Ritz values are asymptotically distributed like μ, the closest element of $\mathcal{M}_t^\sigma$ to the pole measure ν, where the distance is measured in terms of the logarithmic energy $I(\mu-\nu)$. Moreover, eigenvalues close to z with $U^{\mu-\nu}(z)$ strictly less than the maximum F of $U^{\mu-\nu}$ on $\mathbb{C}$ are approached by rational Ritz values with a geometric rate, and this typically happens if the constraint $\mu \leq \sigma$ is active in a neighborhood of z, i.e., $z \notin \operatorname{supp}(\sigma-\mu)$, see Lemma A.1 and Remark A.2. At the end of §5.1 we will discuss an explicit example showing that "exceptional eigenvalues" with a different rate of convergence do indeed exist.

For a proof of Theorem 3.1 presented in §5.1, we will quote some basic results from the late last century about asymptotics of discrete polynomials due to Rakhmanov, Dragnev & Saff, Van Assche & Kuijlaars, Beckermann, and others [2,10,11,14,25,26,32]. However, Theorem 3.1 is not completely satisfactory for three reasons: first of all, what is the convergence rate for the exceptional index $\lambda_{k^*(N),N}$? Also, for the convergence of $\operatorname{dist}(\lambda_{k(N),N},\Theta_N)^{1/N}$ we should just insure that the eigencomponent $w_N(\lambda_{k(N),N})^{1/N} \to 1$, but the Assumption **(H5)** also imposes that all other eigencomponents are large. Finally, and perhaps the most important point, we expect an even better convergence rate of an $n(N)$th Ritz value towards $\lambda_{k(N),N}$ if there are poles very close to $\lambda_{k(N),N}$, which is yet forbidden[2] by **(H6)**.

By weakening our assumptions, we will no longer be able to describe the limit density of Ritz values, but it is still possible to achieve at least the same rate of convergence for the Ritz values.

Theorem 3.2. *Assume that the Assumptions **(H1)**–**(H4)** hold, and let the minimizing measure $\mu \in \mathcal{M}_t^\sigma$ and $F \in \mathbb{R}$ be as in Theorem 3.1. Replace **(H6)** by*

(H6') *Consider the Jordan decomposition of the signed measure $\nu - \sigma = \nu_0 - \sigma_0$. Both $\operatorname{supp}(\sigma)$ and $\operatorname{supp}(\nu_0)$ are finite unions of closed intervals, and U^ν is continuous at each $x \in \Lambda$ with $U^\nu(x) < \infty$.*

Then we have for any sequence $\Lambda_N \ni \lambda_{k(N),N} \to \lambda$

$$\limsup_{N\to\infty} \operatorname{dist}(\lambda_{k(N),N},\Theta_N)^{1/N} \leq \exp(U^{\mu-\nu}(\lambda)-F) \limsup_{N\to\infty} w_N(\lambda_{k(N),N})^{-1/N}. \quad (3.1)$$

*If in addition **(H5)** holds, then*

$$\limsup_{N\to\infty} \operatorname{dist}(\lambda_{k(N),N},\Theta_N)^{1/N} \leq \exp(2(U^{\mu-\nu}(\lambda)-F)), \quad (3.2)$$

with $J \ni \lambda_{k(N),N} \neq \lambda_{k^*(N),N}$ as in Theorem 3.1.

A remaining drawback in Theorem 3.2 is that Assumption **(H1)** requires the spectra and poles to be uniformly bounded for all N, in particular, we do not allow for poles $\xi_{j,N} = \infty$ (excluding the case of polynomial Ritz values). One of the reasons for Assumption **(H1)** is to be able to define correctly the limits occurring in **(H2)** and **(H4)**, since for instance it is not clear what is the logarithmic potential of a mass point at ∞. This situation of unbounded spectra/poles is discussed in the following statement.

[2] In fact, by having a closer look at the above-mentioned work on discrete orthogonal polynomials, it is possible to allow for poles out of $\Lambda \setminus \Lambda_N$ in Theorem 3.1, but then for the sets $\Lambda_N \cup \Xi_N$ we would require a separation condition similar to **(H3)**. In particular, this means that poles are not exponentially close to eigenvalues. We found such a separation condition too restrictive, and suggest in this paper a new and different approach.

COROLLARY 3.3. *Suppose that there is a* $\tau \in \mathbb{R}$ *such that all spectra* Λ_N *are subsets of* $(\tau, +\infty)$*, with* $\liminf_N \operatorname{dist}(\tau, \Lambda_N \cup \Xi_N) > 0$*. Consider the transformed eigenvalues/poles*

$$\underline{\lambda}_{j,N} = \frac{1}{\lambda_{j,N} - \tau}, \quad \underline{\xi}_{j,N} = \frac{1}{\xi_{j,N} - \tau},$$

and replace the $\lambda_{j,N}$ *in* ***(H1)*–*(H4)*** *by* $\underline{\lambda}_{j,N}$*, and similarly the* $\xi_{j,N}$ *by* $\underline{\xi}_{j,N}$*. Provided that all distances are measured in the chordal metric on the Riemann sphere instead of the euclidean metric, all assertions of Theorem 3.1 and Theorem 3.2 remain true if we drop Assumption* ***(H1)****. Moreover, for a sequence of eigenvalues* $\lambda_{k(N),N} \to \lambda = \tau + 1/\underline{\lambda}$*, the claimed rate of convergence* $U^{\mu-\nu}(\underline{\lambda}) - F$ *does not depend on the actual choice of* τ*.*

In summary, these theorems allow to predict the regions of converged Ritz values for a given eigenvalue distribution σ and pole distribution ν. We illustrate this with the help of some examples in the following section.

4. Examples.

An analytic example. We consider the symmetric Toeplitz matrix

$$A_N = \begin{pmatrix} q^0 & q^1 & q^2 & \\ q^1 & q^0 & q^1 & \ddots \\ q^2 & q^1 & q^0 & \ddots \\ & \ddots & \ddots & \ddots \end{pmatrix} \in \mathbb{R}^{N \times N}$$

for $q \in (0,1)$. It is known (see [22] and also [6]) that the family of matrices $(A_N)_{N \geq 1}$ has a joint asymptotic eigenvalue distribution described by the measure σ being supported on the positive interval $[\alpha, \beta]$, with density

$$\frac{\mathrm{d}\sigma}{\mathrm{d}x}(x) = \frac{1}{\pi x \sqrt{(x-\alpha)(\beta - x)}}, \qquad \alpha = \frac{1-q}{1+q}, \qquad \beta = \frac{1+q}{1-q}.$$

One easily verifies with the help of [38, Equation (II.4.47)] that $\mathrm{d}\sigma/\,\mathrm{d}x$ coincides with the density of the balayage of the measure $\delta_0/\sqrt{\alpha\beta} = \delta_0$ onto $[\alpha, \beta]$. If we assume that the poles are only placed to the right of β, i.e., $\operatorname{supp}(\nu) \subset (\beta, +\infty)$, we can apply Lemma A.3 to compute $\operatorname{supp}(\sigma - \mu) = [\alpha, b(t)]$, outside of which we have geometric convergence of Ritz values (cf. Remark A.2). As a simple example we set $\nu = t\delta_\xi$ for some $\xi > \beta$. Then the integral equation from Lemma A.3 can be solved for $b = b(t) \in [\alpha, \beta]$ if $t \geq t_0$, where

$$t_0 := \frac{1}{\beta}\sqrt{\frac{\xi - \beta}{\xi - \alpha}}, \quad \text{and thus} \quad b(t) = \begin{cases} \beta, & \text{if } t < t_0, \\ \frac{\xi}{t^2\beta(\xi-\alpha)+1}, & \text{if } t \geq t_0. \end{cases}$$

In Figure 4.1 we illustrate the convergence of Ritz values of order $n = 1, \ldots, N$ for the matrix A_N with $N = 100$ and $q = 1/3$. In column n one finds the nth rational Ritz values (and thus the eigenvalues in the right-most column $n = N$), where in all figures we have used the color code given in Table 4.1 to display the distance of a Ritz value to the set of eigenvalues relative to the spread of the spectrum $\lambda_N - \lambda_1$.

For this example the spectral interval is $\operatorname{supp}(\sigma) = [1/2, 2]$. In the left figure we have placed all poles in $\xi = 10$, whereas in the right figure all poles are in $\xi = 2$, i.e.,

TABLE 4.1
Color code for the following figures.

Color	Relative distance of a Ritz value θ to the spectrum
Red	$\mathrm{dist}(\theta, \Lambda(A_N)) < 10^{-7.5}$
Yellow	$10^{-7.5} \leq \mathrm{dist}(\theta, \Lambda(A_N)) < 10^{-5}$
Green	$10^{-5} \leq \mathrm{dist}(\theta, \Lambda(A_N)) < 10^{-2.5}$
Blue	$10^{-2.5} \leq \mathrm{dist}(\theta, \Lambda(A_N))$

on the right-end of the support interval of σ. In both cases the starting vector was chosen to have equal components in all eigenvectors. The graph of $b(t)$ (solid black line) is a good indicator for the regions where the Ritz values begin to converge to eigenvalues of A_N. We observe that the Ritz values are attracted by the pole and this effect becomes stronger as the pole moves closer to the spectrum of A_N.

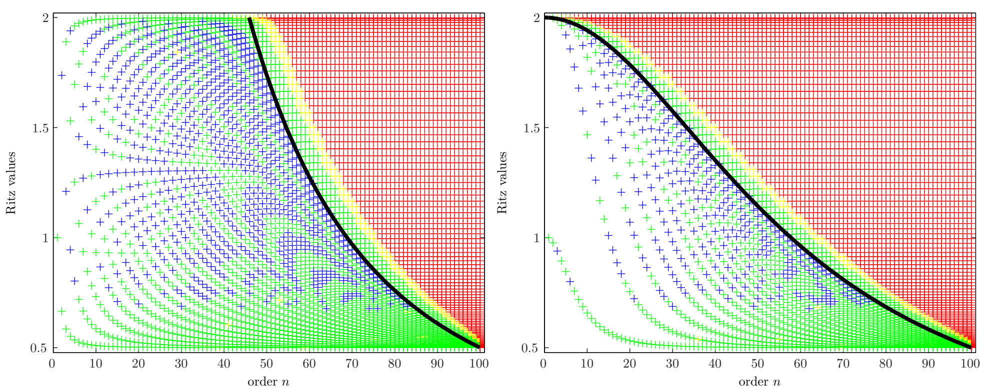

FIG. 4.1. *In these figures we plot Ritz values of order $n = 1, \ldots, N$. The colors indicate the distance of a Ritz value to a closest eigenvalue of the Toeplitz matrix A_N from the Kac-Murdock-Szegő example ($N = 100$, $q = 1/3$). The solid black line is the graph of $b(t)$, $t = n/N$. In the left figure all poles are in $\xi = 10$ and in the right figure all poles are in $\xi = 2$.*

We remark that the convergence analysis of rational Ritz values for a fixed pole ξ is closely related to the analysis of polynomial Ritz values for the shifted and inverted matrix $(A - \xi I)^{-1}$. In the following examples we also consider non-constant pole sequences. Unfortunately, this complicates the set $\mathrm{supp}(\sigma - \mu)$ and it becomes much more complicated (or even impossible) to obtain analytic expressions of its boundary points depending on t. In the following examples we have therefore approximated the extremal measure μ (depending on t) numerically, by minimizing the energy over measures having a piecewise linear density, which is a subset of $\mathcal{M}_t^\sigma$. This leads to a constrained quadratic optimization problem for the energy with a finite number of unknowns, which we solved with an active set algorithm via Matlab.

Equidistant eigenvalues. Let A_N have N equidistant eigenvalues within $[-1, 1]$. Then σ is the Lebesque measure restricted onto $[-1, 1]$. Without loss of generality we can consider $A_N = \mathrm{diag}(1 - N, 3 - N, \ldots, N - 1)/(N + 1)$, because the convergence behavior of the Ritz values depends on σ only and not on the eigenvector directions. It is known that the polynomial Lanczos method finds equidistant eigenvalues from the boundary of the interval $[-1, 1]$ to the inside. More precisely, $\mathrm{supp}(\sigma - \mu) = [-a(t), a(t)]$ with $a(t) = \sqrt{1 - t^2}$ (cf. [6]). This is different in the presence of poles as we illustrate in Figure 4.2 (in this example we set $N = 100$ and $\mathbf{b} = [1, \ldots, 1]^T$). In the left of Figure 4.2 we have placed all poles at 0, the midpoint of the spectral interval of A_N. It is clearly seen that in contrast to the polynomial Lanczos method, now the inner eigenvalues are found first by the rational Krylov method. In the right figure we have used alternating poles $\xi_{2j-1} = 0$ and $\xi_{2j} = 1$ $(j = 1, \ldots, 50)$. Now the Ritz values first converge to eigenvalues at the left boundary and close to the midpoint of the spectral interval of A_N. The black solid line indicates the boundary of $\mathrm{supp}(\sigma - \mu)$ depending on t, which was computed numerically.

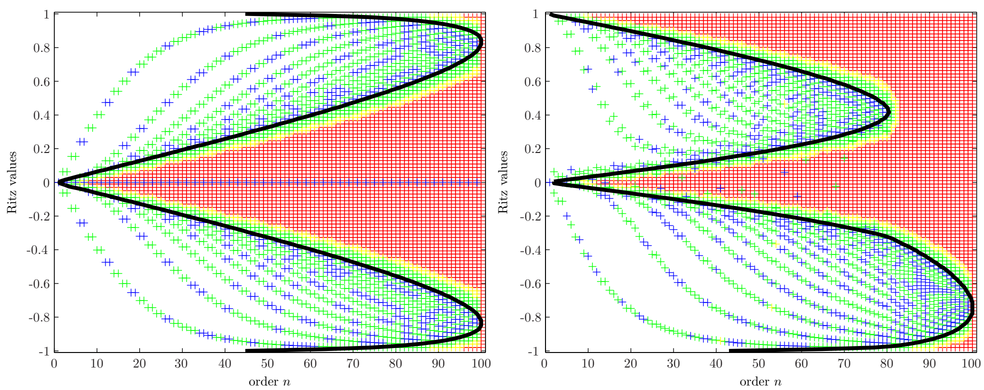

FIG. 4.2. *In these figures we plot Ritz values of order* $n = 1, \ldots, N$. *The colors indicate the distance of a Ritz value to a closest eigenvalue of* A_N *(*$N = 100$*) with equidistant eigenvalues in* $[-1, 1]$. *The solid black line is the boundary of* $\mathrm{supp}(\sigma - \mu)$ *as a function of* $t = n/N$. *In the left figure all poles are at* 0 *and in the right figure the poles are alternating* $(0, 1, 0, 1, \ldots)$.

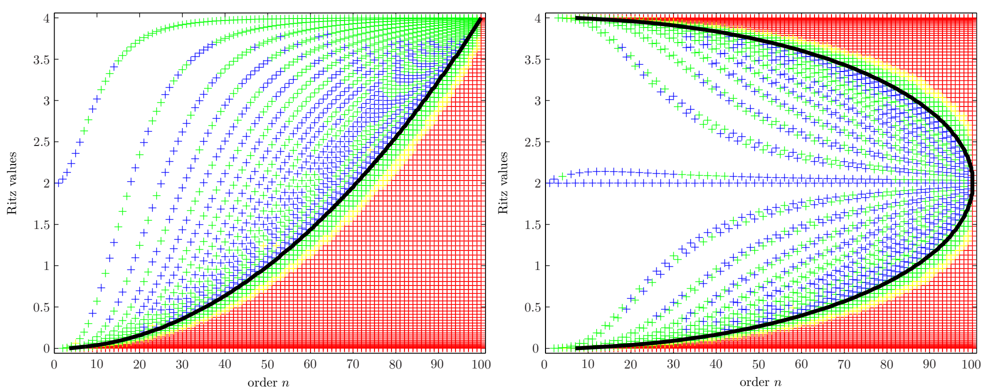

FIG. 4.3. *In these figures we plot the Ritz values of order* $n = 1, \ldots, N$. *The colors indicate the distance of a Ritz value to a closest eigenvalue of* A_N *(*$N = 100$*) with equilibrium-distributed eigenvalues in* $[0, 4]$. *The solid black line is the boundary of* $\mathrm{supp}(\sigma - \mu)$ *as a function of* $t = n/N$. *In the left figure the poles are* $(0, \infty, 0, \infty, \ldots)$ *and in the right figure the poles are* $(0, 4, 0, 4, \ldots)$.

Equilibrium-distributed eigenvalues. Let

$$A_N = \begin{pmatrix} 2 & -1 & & \\ -1 & 2 & -1 & \\ & \ddots & \ddots & \ddots \end{pmatrix} \in \mathbb{R}^{N \times N},$$

then σ is the equilibrium measure on $[0,4]$. Due to this fact actually no eigenvalue of A_N is found by the polynomial Lanczos method as long as $n < N$. This is no longer the case if a rational Krylov method is used. In Figure 4.3 (left) we illustrate the convergence of the rational Ritz values if the pole sequence $(0,\infty,0,\infty,\ldots)$ is used ($N = 100$). This sequence occurs, e.g., in the so-called extended Lanczos method for the approximation of matrix functions (see [16]). The Ritz values close to 0 are found in very early iterations. On the right of Figure 4.3 the convergence behavior for the pole sequence $(0,4,0,4,\ldots)$ is shown[3]. As before, the black solid line indicates the boundary of $\mathrm{supp}(\sigma-\mu)$ depending on t, which was computed numerically.

5. Proofs. To retain from cluttering the formulas with an overloaded subindex notation we take the following conventions in the remainder for the polynomials introduced in §2.2 and Lemma 2.3, which according to the notation of §3 have an additional index N

$$\widetilde{p}_{n(N),N} = P_N, \quad q_{n(N)-1,N} = Q_N, \tag{5.1}$$

and accordingly we write $s = S_N$ for the polynomial in Lemma 2.3. We may assume without loss of generality that both P_N and Q_N are monic. Note also that the index N *does not* indicate the degree of these polynomials.

5.1. Proof of Theorem 3.1. Our proof is divided into three parts. In Proposition 5.1 we first establish an upper bound for the quadratic polynomial of Lemma 2.3. Secondly, we show in Proposition 5.3 that this gives essentially the results claimed in Theorem 3.2. In a third step we show that these results are sharp as claimed in Theorem 3.1.

Our second step is more of a combinatorial nature and inspired by [1] (see also [21] for the case of isometric Ritz values). For the other two parts we could use well-established asymptotics for discrete orthogonal polynomials obtained by Rakhmanov [32] and Dragnev & Saff [14], followed by several other authors, since there is a link with the extremal polynomial in Lemma 2.3. However, at least for the first step, we decided to follow [6] to give the explicit proof by designing a "good" polynomial, since this construction will be generalized in our proof of Theorem 3.2.

Proposition 5.1. *Under the Assumptions **(H1)–(H4)** and **(H6)**, let $\Lambda_N \ni \lambda_{k(N),N} \to \lambda$ for $N \to \infty$, and denote by $\theta_{\kappa(N)-1,N} < \theta_{\kappa(N),N}$ the nth (rational) Ritz values out of Θ_N closest to $\lambda_{k(N),N}$ on the left and on the right of $\lambda_{k(N),N}$, respectively. Then*

$$\begin{aligned}\limsup_{N\to\infty} &|w_N(\lambda_{k(N),N})^2(\lambda_{k(N),N}-\theta_{\kappa(N)-1,N})(\theta_{\kappa(N),N}-\lambda_{k(N),N})|^{1/N} \\ &\le \exp(2(U^{\mu-\nu}(\lambda)-F)).\end{aligned} \tag{5.2}$$

If there is no such Ritz value on the left or on the right of $\lambda_{k(N),N}$, then the above bound remains valid after omitting the corresponding linear factor.

Proof. We will apply Lemma 2.3. First, note that

$$\limsup_{N\to\infty}\Big(\sum_{j=1,j\neq k(N)}^{N} w_N(\lambda_{j,N})^2|(\lambda_{j,N}-\theta_{\kappa(N)-1,N})(\theta_{\kappa(N),N}-\lambda_{j,N})|^{1/N}\Big)^{1/N} \le 1,$$

[3]It will be shown in a future publication that for the pole sequence $(0,4,0,4,\ldots)$ we have $\mathrm{supp}(\sigma-\mu) = [2-2\sqrt{1-t^2}, 2+2\sqrt{1-t^2}]$ corresponding to the circular black solid curve in Figure 4.3 (right), and that for the pole sequence $(\xi,\xi,\ldots)$ with $\xi \ge 4$ there holds $\mathrm{supp}(\sigma-\mu) = [0, \min\{4, \xi(1-t^2)\}]$ which corresponds to an incomplete parabola.

since $\sum_j w_N(\lambda_{j,N})^2 = 1$, and the Ritz values lie in the convex hull of the eigenvalues, and the latter are uniformly bounded by Assumption **(H1)**. Thus, by designing suitable monic polynomials S_N of degree $\leq n(N) - 2$ with set of zeros $Z_N \subset \Lambda_N \setminus \{\lambda_{k(N),N}\}$, we get from the second part of Lemma 2.3 and the fact $w_N(\lambda_{j,N}) \leq 1$ that

$$\begin{aligned} &\limsup_{N\to\infty} |w_N(\lambda_{k(N),N})^2(\lambda_{k(N),N} - \theta_{\kappa(N)-1,N})(\theta_{\kappa(N),N} - \lambda_{k(N),N})|^{1/(2N)} \\ &\leq \limsup_{N\to\infty} \max_{j\neq k(N)} \frac{|S_N/Q_N|^{1/N}(\lambda_{j,N})}{|S_N/Q_N|^{1/N}(\lambda_{k(N),N})} =: \limsup_{N\to\infty} \frac{|S_N/Q_N|^{1/N}(\lambda_{j(N),N})}{|S_N/Q_N|^{1/N}(\lambda_{k(N),N})} \end{aligned} \tag{5.3}$$

for some $j(N) \neq k(N)$, where we recall that the monic polynomial Q_N has the set of zeros Ξ_N. Note also that if there are no $n(N)$th Ritz values on the left of $\lambda_{k(N),N}$ we may apply the first part of Lemma 2.3 and obtain the same conclusion (5.3) with the factor $(\lambda_{k(N),N} - \theta_{\kappa(N)-1,N})$ omitted on the left.

Let us first construct these polynomials S_N depending on some $\eta > 0$: we define

$$V = \{z \in \mathbb{R} : U^{\mu-\nu}(z) \geq F - \eta\}$$

being a closed neighborhood of $\mathrm{supp}(\sigma - \mu)$ by the semi-continuity of U^ν and the continuity of U^μ. By possibly making $\eta > 0$ a bit smaller we may suppose that $\sigma(\partial V) = 0$. We then apply Lemma A.5 with

$$Z_{3,N} = \Lambda_N \setminus \{\lambda_{k(N),N}\}, \quad Z_{1,N} = Z_{3,N} \setminus V, \quad \rho_2 = \mu, \quad i(N) = n(N) - 2,$$

and define $Z_N := Z_{2,N}$, obtained from $\Lambda_N \setminus \{\lambda_{k(N),N}\}$ by dropping elements from V, with $\chi_N(Z_N) \to \mu$. Indeed, from the Assumptions **(H1)** and **(H2)** we get $\rho_3 = \sigma$ and, since $\sigma(\partial V) = 0$, $\rho_1 = \sigma|_{\mathbb{R}\setminus V}$.

Then S_N vanishes on $\Lambda_N \setminus (V \cup \{\lambda_{k(N),N}\})$, implying that $\lambda_{j(N),N}$ lies in the compact $\Lambda \cap V$. By passing to subsequences if necessary, we may suppose without loss of generality that $\lambda_{j(N),N} \to \widetilde{\lambda} \in \Lambda \cap V$, and, as in the assertion of the proposition, $\lambda_{k(N),N} \to \lambda \in \Lambda$. By construction and Assumption **(H4)** we have $\chi_N(Z_N) \to \mu$ and $\chi_N(\Xi_N) \to \nu$. Then the Assumption **(H6)** together with (A.2) tells us that

$$\lim_{N\to\infty} \log(|Q_N(\lambda_{k(N),N})|^{1/N}) = -U^\nu(\lambda), \tag{5.4}$$

$$\lim_{N\to\infty} \log(|Q_N(\lambda_{j(N),N})|^{1/N}) = -U^\nu(\widetilde{\lambda}), \tag{5.5}$$

whereas from the principle of descent (A.1)

$$\limsup_{N\to\infty} \log(|S_N(\lambda_{j(N),N})|^{1/N}) \leq -U^\mu(\widetilde{\lambda}). \tag{5.6}$$

For the term $|S_N(\lambda_{k(N),N})|$, following Kuijlaars we will use the separation condition of Hypothesis **(H3)** in order to show that

$$\lim_{N\to\infty} \log(|S_N(\lambda_{k(N),N})|^{1/N}) = -U^\mu(\lambda). \tag{5.7}$$

Before giving a proof, note that (5.4), (5.5), (5.6), and (5.7) imply that the logarithm of the right-hand side of (5.3) is bounded above by

$$U^{\mu-\nu}(\lambda) - U^{\mu-\nu}(\widetilde{\lambda}) \leq U^{\mu-\nu}(\lambda) - F + \eta,$$

where we have used that $\widetilde{\lambda} \in V$. Since $\eta > 0$ was arbitrary, the assertion of Proposition 5.1 follows.

It remains to establish (5.7). By Assumption **(H3)** or its equivalent formulation of Lemma A.4, given any $\epsilon > 0$ we find $\delta \in (0, 1/4)$ satisfying (A.9). Consider $J = [\lambda - \delta, \lambda + \delta]$. By Assumption **(H3)**, both measures $\mu \leq \sigma$ do not have mass points, and hence $\chi_N(Z_N \setminus J) \to \mu|_{\mathbb{R}\setminus J}$. The principle of descent (A.2) allows to conclude that

$$\lim_{N\to\infty} \log \Big| \prod_{\lambda_{j,N} \in Z_N \setminus J} (\lambda_{k(N),N} - \lambda_{j,N}) \Big|^{1/N} = -U^{\mu|_{\mathbb{R}\setminus J}}(\lambda) = -U^{\mu}(\lambda) + U^{\mu|_J}(\lambda),$$

where $0 \leq U^{\mu|_J}(\lambda) \leq \epsilon$ by Lemma A.4. On the other hand, since by construction every term in the sum occurring in (A.9) is positive, we also get from (A.9) that

$$0 \geq \log \Big| \prod_{\lambda_{j,N} \in Z_N \cap J} (\lambda_{k(N),N} - \lambda_{j,N}) \Big|^{1/N} \geq -\epsilon.$$

Hence

$$0 \leq \limsup_{N\to\infty} \Big| \log(|S_N(\lambda_{k(N),N})|^{1/N}) + U^{\mu}(\lambda) \Big| \leq 2\epsilon,$$

and the claim (5.7) follows for $\epsilon \to 0$. ☐

Remark 5.2. *The estimate* (5.2) *of Proposition 5.1 implies that*

$$\limsup_{N\to\infty} \operatorname{dist}(\lambda_{k(N),N}, \Theta_N)^{1/N} \leq \exp(U^{\mu-\nu}(\lambda) - F) \limsup_{N\to\infty} w_N(\lambda_{k(N),N})^{-1/N}, \quad (5.8)$$

which is the statement of (3.1). *To see this, note that if* $\theta_{\kappa(N)-1,N} \leq \lambda_{k(N),N} \leq \theta_{\kappa(N),N}$ *then*

$$\begin{aligned}
&\limsup_{N\to\infty} \operatorname{dist}(\lambda_{k(N),N}, \Theta_N)^{1/N} \\
&\leq \limsup_{N\to\infty} ((\lambda_{k(N),N} - \theta_{\kappa(N)-1,N})(\theta_{\kappa(N),N} - \lambda_{k(N),N}))^{1/(2N)} \\
&= \limsup_{N\to\infty} \left| \frac{w_N(\lambda_{k(N),N})^2 (\lambda_{k(N),N} - \theta_{\kappa(N)-1,N})(\theta_{\kappa(N),N} - \lambda_{k(N),N})}{w_N(\lambda_{k(N),N})^2} \right|^{1/(2N)} \\
&\leq \exp(U^{\mu-\nu}(\lambda) - F) \limsup_{N\to\infty} w_N(\lambda_{k(N),N})^{-1/N},
\end{aligned}$$

where in the last step we have used (5.2). If there is no Ritz value on the left or on the right of $\lambda_{k(N),N}$*, then (5.8) immediately follows by recalling that* $w_N(\lambda_{k(N),N}) \in [0, 1]$.

So far we have shown in Proposition 5.1 and Remark 5.2 upper bounds for the (product of) distances of closest Ritz values on the left and on the right of a given eigenvalue, which, however, does not allow us directly to give sharp asymptotics for $\operatorname{dist}(\lambda_{k(N),N}, \Theta_N)$. We need to give a complete classification of how Ritz values interlace with eigenvalues in intervals where each eigenvalue is exponentially close to some Ritz value. Similar arguments have been used in [1, 21].

To do so, suppose that there exists a closed interval J and $\epsilon > 0$ such that for all sufficiently large N and for all $\lambda_{k,N} \in J$ there holds

$$\operatorname{dist}(\lambda_{k,N}, \Theta_N) \leq \exp(-2\epsilon N). \quad (5.9)$$

Due to **(H3)** we also know that for sufficiently large N and for all k there holds

$$|\lambda_{k+1,N} - \lambda_{k,N}| \geq \exp(-\epsilon N), \tag{5.10}$$

since by Lemma A.4 $\log(1/|\lambda_{k+1,N} - \lambda_{k,N}|) \leq \max\{N\epsilon, \log(1/(4\delta))\}$ for some $\delta > 0$ depending on ϵ. It follows from (5.9) and (5.10) that we may choose N large enough such that the $\exp(-2\epsilon N)$–neighborhoods of the $\lambda_{k,N} \in J$ do not intersect.

We say that $\lambda_{k,N} \in J$ is *(strictly) left-approached* if there is a Ritz value from Θ_N in the $\exp(-2\epsilon N)$–neighborhood of $\lambda_{k,N}$ being $\leq \lambda_{k,N}$ (and $< \lambda_{k,N}$). Similarly, we speak of *(strictly) right-approached* eigenvalues. From (5.9) we see that each $\lambda_{k,N} \in J$ is left- or right-approached, or even both. Moreover, since the above neighborhoods do not intersect, each $\lambda_{k,N} \in J$ is approached by a different Ritz value.

We now recall the interlacing property that any open interval spanned by two consecutive Ritz values does contain an eigenvalue. It follows that if $\lambda_{k,N} \in J$ is left-approached then $\lambda_{j,N} \in J$ is strictly left-approached for $j = k-1$, and thus for all $j < k$. By the same argument, if $\lambda_{k,N} \in J$ is right-approached then $\lambda_{j,N} \in J$ is strictly right-approached for $j = k+1$, and thus for all $j > k$. So we only switch at most once within the interval J from left-approached to right-approached eigenvalues.

We have enumerated the different cases in Figure 5.1 and Figure 5.2, where first we suppose that J does not contain any extremal eigenvalues. In Case 1(a) and Case 1(b) there is no switching, since either the right-most eigenvalue in J is strictly left-approached (and so are the others), or the left-most eigenvalue in J is strictly right-approached (and so are the others). By the interlacing property there is at most one Ritz value $\theta \in J$ outside of the $\exp(-2\epsilon N)$–neighborhoods of the $\lambda_{k,N} \in J$, which is described in Case 2; here the eigenvalues $\lambda_{j,N} \in J \cap (\theta, +\infty)$ must be strictly right-approached and $\lambda_{j,N} \in J \cap (-\infty, \theta)$ must be strictly left-approached. In Case 3 one neighborhood contains more than one Ritz value, and thus we must have strict approximation. In Case 5, we have an eigenvalue which is also left- and right-approached, since it is hit by a Ritz value; there cannot be any further Ritz values in the same neighborhood according to the interlacing property. Finally, in the remaining Case 4 we have $\lambda_{k,N}, \lambda_{k+1,N} \in J$ which are approached from different sides.

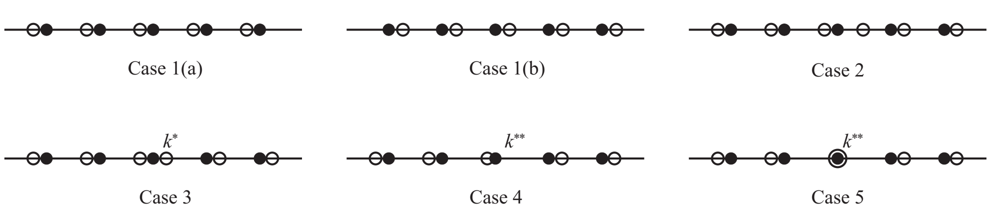

Fig. 5.1. *Examples of relative positions of the Ritz values and the eigenvalues in a closed interval J not containing the extremal eigenvalues. The Ritz values are depicted by circles, the eigenvalues are depicted by black dots. The $\lambda_{k^*(N)}$ and $\lambda_{k^{**}(N)}$ indicate the exceptional positions in Proposition 5.3 and in the proof of the lower bound in Theorem 3.1.*

In Figure 5.2 we have drawn the corresponding situations where J contains one extremal eigenvalue (it cannot contain both): since Ritz values lie in the interior of the convex hull of the eigenvalues, we have similar phenomena as in the Case 1(a) and 1(b).

Proposition 5.3. *Suppose the Assumptions **(H3)**, **(H5)**, and Equation* (5.2) *of Proposition 5.1 hold. Then in a closed interval $J \subset \{x \in \mathbb{R} : U^{\mu-\nu}(x) < F\}$, all*

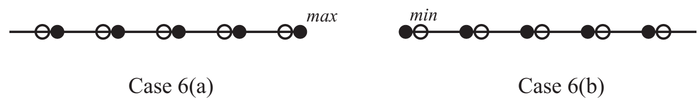

Fig. 5.2. *The relative positions of Ritz values and eigenvalues in a closed interval J containing an extremal eigenvalue. The Ritz values are depicted by circles, the eigenvalues are depicted by black dots. The eigenvalues λ_{min} and λ_{max} denote the smallest and largest eigenvalue, respectively.*

sequences $J \ni \lambda_{k(N),N} \to \lambda$ for $N \to \infty$ satisfy

$$\limsup_{N\to\infty} \operatorname{dist}(\lambda_{k(N),N}, \Theta_N)^{1/N} \le \exp(2(U^{\mu-\nu}(\lambda) - F)), \tag{5.11}$$

with the possible exclusion of at most one unique "exceptional index" $k^*(N)$.

Proof. The remainder of the proof is inspired by similar arguments in [1, 21]. We use the notation from Proposition 5.1. Equation (5.2) implies (5.8) (see Remark 5.2), and using **(H5)** we obtain for all $\lambda_{k(N),N} \in J$

$$\limsup_{N\to\infty} \min_j |\lambda_{k(N),N} - \theta_{j,N}|^{1/N} \le \exp\left(U^{\mu-\nu}(\lambda) - F\right). \tag{5.12}$$

We will tighten up this bound to obtain Equation (5.11).

Depending on the closed interval $J \subset \{x \in \mathbb{R} : U^{\mu-\nu}(x) < F\}$, we can choose $\epsilon > 0$ such that for all $\lambda_{k(N),N} \in J$, and for N large enough

$$\min_j |\lambda_{k(N),N} - \theta_{j,N}|^{1/N} \le \exp(\epsilon) \max_{x\in J} \exp(U^{\mu-\nu}(x) - F) \le \exp(-2\epsilon) < 1.$$

To continue the proof we have to exclude Case 3 of Figure 5.1. If the Ritz values and eigenvalues are positioned as in Case 3, we take $k^*(N)$ such that $\lambda_{k^*(N)}$ coincides with the middle eigenvalue as depicted in Case 3, so this eigenvalue is the "exceptional eigenvalue".

Assume as in Proposition 5.1 that $\lambda_{k(N),N} \in [\theta_{\kappa(N)-1,N}, \theta_{\kappa(N),N}]$. From the discussion about the position of the Ritz values and due to the exclusion of Case 3 we know that each eigenvalue is approached by a single Ritz value. Without loss of generality we assume that $\lambda_{k(N),N}$ is approached by $\theta_{\kappa(N),N}$ instead of $\theta_{\kappa(N)-1,N}$, implying $\min_j |\lambda_{k(N),N} - \theta_{j,N}| = |\lambda_{k(N),N} - \theta_{\kappa(N),N}| \le \exp(-2\epsilon N)$, and also $|\lambda_{k(N),N} - \theta_{\kappa(N)-1,N}| > \exp(-2\epsilon N)$, which leads to

$$\exp(-2\epsilon) |\lambda_{k(N),N} - \theta_{\kappa(N),N}|^{1/N} < \left(|\lambda_{k(N),N} - \theta_{\kappa(N)-1,N}||\lambda_{k(N),N} - \theta_{\kappa(N),N}|\right)^{1/N}.$$

Thus by (5.12) and **(H5)**,

$$\begin{aligned}
&\exp(-2\epsilon) \limsup_{N\to\infty} \min_j |\lambda_{k(N),N} - \theta_{j,N}|^{1/N} \\
&= \exp(-2\epsilon) \limsup_{N\to\infty} |\lambda_{k(N),N} - \theta_{\kappa(N),N}|^{1/N} \\
&\le \limsup_{N\to\infty} \left(|\lambda_{k(N),N} - \theta_{\kappa(N)-1,N}||\lambda_{k(N),N} - \theta_{\kappa(N),N}|\right)^{1/N} \\
&\le \exp(2(U^{\mu-\nu}(\lambda) - F)).
\end{aligned}$$

For $\epsilon \to 0$ we obtain (5.11). ∎

To prove the lower bound in Theorem 3.1, extra results related to the limiting distribution of the Ritz values are needed (see [2, Theorem 1.3] which generalizes [14, Theorem 3.3]).

THEOREM 5.4. *Under the Assumptions **(H1)**–**(H6)**, let the positive finite Borel measure μ be the unique minimizer of $\mu_1 \mapsto I(\mu_1) - 2I(\nu, \mu_1)$ within $\mathcal{M}_t^\sigma$ and define F as the maximum of $U^{\mu-\nu}$ in the whole complex plane. Then we have*

$$\chi_N(\Theta_N) \to \mu, \text{ and } \lim_{N\to\infty} \Big\| (P_N/Q_N)(A_N)\frac{\mathbf{b}_N}{\|\mathbf{b}_N\|}\Big\|^{1/N} = \exp(-F).$$

Proof. Recall from Lemma 2.3 that P_N is the $n(N)$th monic orthogonal polynomial with respect to the scalar product with varying weights

$$\langle p, q\rangle = \Big\langle \ (q/Q_N)(A_N)\frac{\mathbf{b}_N}{\|\mathbf{b}_N\|}\,,\ (p/Q_N)(A_N)\frac{\mathbf{b}_N}{\|\mathbf{b}_N\|}\ \Big\rangle = \sum_{\lambda\in\Lambda_N} \frac{w_N(\lambda)^2}{Q_N(\lambda)^2} p(\lambda)\overline{q(\lambda)}.$$

The assumptions of [2, Theorem 1.3] are readily verified. The assumption on the connectedness of $\mathrm{supp}(\sigma)$ mentioned in [2] is just to insure that $\mathrm{supp}(\sigma-\mu)\cap\mathrm{supp}(\mu) \neq \emptyset$, which is trivial in our case since $\mathrm{supp}(\mu) = \mathrm{supp}(\sigma)$ by Lemma A.1(c). ◻

We are now prepared to conclude.

Proof of Theorem 3.1. In view of Proposition 5.3 it only remains to prove the lower bound

$$\liminf_{N\to\infty} \min_j |\lambda_{k(N),N} - \theta_{j,N}|^{1/N} \geq \exp(2(U^{\mu-\nu}(\lambda) - F)).$$

Suppose that $\lambda_{k(N),N}$ is located in an interval $[\theta_{\kappa(N)-1,N}, \theta_{\kappa(N),N}]$; otherwise the reasoning becomes simpler, compare with Lemma 2.3. Let us denote by $d < \infty$ the diameter of the convex hull of the compact set Λ, which therefore is an upper bound for the distance between an arbitrary eigenvalue and a Ritz value, since both eigenvalues and Ritz values are contained in this convex hull.

We factor the above lim inf into three parts

$$\begin{aligned}
&\liminf_{N\to\infty} \min_j |\lambda_{k(N),N} - \theta_{j,N}|^{1/N} \\
&\geq \liminf_{N\to\infty} \left(\frac{|\lambda_{k(N),N} - \theta_{\kappa(N)-1,N}||\lambda_{k(N),N} - \theta_{\kappa(N),N}|}{d} \right)^{1/N} \\
&\geq \liminf_{N\to\infty} \left(\frac{1}{w_N(\lambda_{k(N),N})} \right)^{2/N} \liminf_{N\to\infty} \left(\frac{|\lambda_{k(N),N} - \theta_{\kappa(N)-1,N}||\lambda_{k(N),N} - \theta_{\kappa(N),N}|}{|P_N/Q_N|(\lambda_{k(N),N})} \right)^{2/N} \\
&\quad \liminf_{N\to\infty} \left(\frac{w_N(\lambda_{k(N),N})^2 |P_N/Q_N|^2(\lambda_{k(N),N})}{|\lambda_{k(N),N} - \theta_{\kappa(N)-1,N}||\lambda_{k(N),N} - \theta_{\kappa(N),N}|} \right)^{1/N},
\end{aligned} \tag{5.13}$$

and bound these three factors separately from below.

The first factor. This is covered by **(H5)**.

The second factor. Using Theorem 5.4, we obtain by the principle of descent (A.1) for P_N and by (A.2) for Q_N that

$$\limsup_{N\to\infty} \left(\frac{|P_N/Q_N|(\lambda_{k(N),N})}{|\lambda_{k(N),N} - \theta_{\kappa(N)-1,N}||\lambda_{k(N),N} - \theta_{\kappa(N),N}|} \right)^{1/N} \leq \exp(-U^{\mu-\nu}(\lambda)),$$

and thus a lower bound for the second factor in Equation (5.13).

The third factor. We use the Conventions (5.1) and Lemma 2.3 with the polynomial $S_N(x) = P_N(x)/\left((x - \theta_{\kappa(N)-1,N})(x - \theta_{\kappa(N),N})\right)$. Then the right-hand side of Equation (2.8) becomes zero and we get

$$0 = \sum_{j=1}^{N} a_j, \qquad a_j := \frac{w_N(\lambda_{j,N})^2 (P_N/Q_N)^2(\lambda_{j,N})}{\left(\lambda_{j,N} - \theta_{\kappa(N)-1,N}\right)\left(\lambda_{j,N} - \theta_{\kappa(N),N}\right)}.$$

Bringing the negative terms in the summation to the left and leaving the positive ones on the right yields

$$\sum_{\substack{\lambda_{j,N} \in \\ \left[\theta_{\kappa(N)-1,N},\theta_{\kappa(N),N}\right]}} |a_j| = \sum_{\substack{\lambda_{j,N} \notin \\ \left[\theta_{\kappa(N)-1,N},\theta_{\kappa(N),N}\right]}} |a_j|. \tag{5.14}$$

The interval $[\theta_{\kappa(N)-1,N}, \theta_{\kappa(N),N}]$ contains at most two eigenvalues (Cases 4 and 5 in Figure 5.1). In Case 5, one of the terms equals zero and we exclude $\lambda_{k^{**}(N),N}$. Hence in Cases 1, 2, 3, and 5, the above equation has only one term in the sum on the left-hand side, namely $a_{k(N)}$. Let us distinguish between one or two terms.

A single term on the left-hand side of Equation (5.14). This corresponds to Cases 1, 2, 3, and 5 (excluding $\lambda_{k^{**}(N)}$ for Case 5). We have

$$\begin{aligned}
|a_{k(N)}| &= \sum_{\substack{\lambda_{j,N} \notin \\ [\theta_{\kappa(N)-1,N},\theta_{\kappa(N),N}]}} |a_j| = \frac{1}{2}\sum_{j=1}^{N} |a_j| \\
&\geq \frac{1}{2d^2} \sum_{j=1}^{N} w_N(\lambda_{j,N})^2 (P_N/Q_N)^2(\lambda_{j,N}) \\
&= \frac{1}{2d^2} \left\| (P_N/Q_N)(A) \frac{\mathbf{b}_N}{\|\mathbf{b}_N\|} \right\|^2,
\end{aligned}$$

where the first two equalities follow from (5.14), and the inequality follows from the definition of a_j and d.

Two terms on the left-hand side of Equation (5.14). This is Case 4. Denote the two eigenvalues in $[\theta_{\kappa(N)-1,N}, \theta_{\kappa(N),N}]$ by $\lambda_{j(N)-1,N}$ and $\lambda_{j(N),N}$. Then we exclude the index $k^{**}(N) \in \{j(N)-1, j(N)\}$ corresponding to the smaller of the two terms on the left-hand side of (5.14), and get for the larger one with index $k(N) \in \{j(N)-1, j(N)\}$

$$\begin{aligned}
|a_{k(N)}| &\geq \frac{1}{2}\left(|a_{k(N)}| + |a_{k^{**}(N)}|\right) = \frac{1}{2} \sum_{\substack{\lambda_{j,N} \in \\ [\theta_{\kappa(N)-1,N},\theta_{\kappa(N),N}]}} |a_j| \\
&\geq \frac{1}{4d^2} \left\| (P_N/Q_N)(A) \frac{\mathbf{b}_N}{\|\mathbf{b}_N\|} \right\|^2.
\end{aligned}$$

Thus, in both cases we get using Theorem 5.4 that for all $k(N)$ different from $k^{**}(N)$,

$$\lim_{N\to\infty} \left(\frac{w_N(\lambda_{k(N),N})^2 (P_N/Q_N)^2(\lambda_{k(N),N})}{\left|\lambda_{k(N),N} - \theta_{\kappa(N)-1,N}\right| \left|\lambda_{k(N),N} - \theta_{\kappa(N),N}\right|} \right)^{1/N} \geq \exp(-2F).$$

Combining the estimates for all three factors leads to the lower bound

$$\liminf_{N\to\infty} \min_j |\lambda_{k(N),N} - \theta_{j,N}|^{1/N} \geq \exp(2(U^{\mu-\nu}(\lambda) - F)),$$

as claimed in the beginning of the proof.

Note that in Theorem 3.1 no distinction is made between the indices $k^*(N)$ and $k^{**}(N)$. Combining Proposition 5.3 and the above claim, and taking $k^*(N)$ from Theorem 3.1 as either $k^*(N)$ from Proposition 5.3 or $k^{**}(N)$ from this proof (which means excluding at most one eigenvalue in Cases 3, 4, and 5) proves the equality in Theorem 3.1. ◻

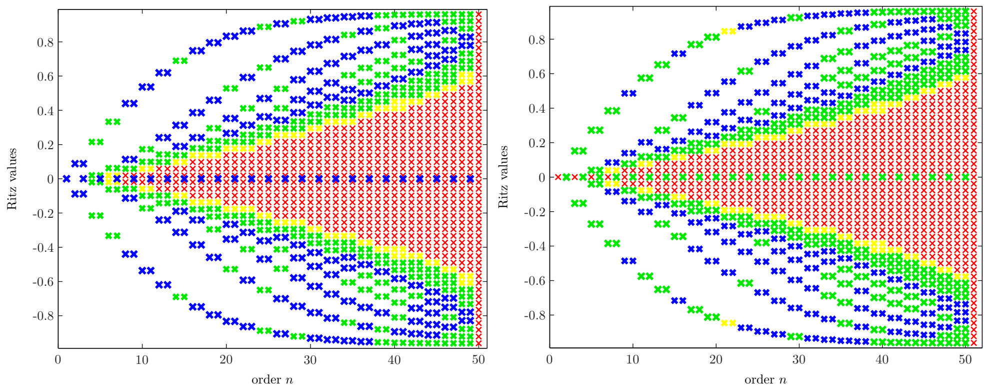

Fig. 5.3. *Ritz values of order $n = 1, \dots, N$ for equidistant eigenvalues on $[-1, 1]$, eigencomponents $1/\sqrt{N}$, and poles $(\xi, -\xi, \xi, -\xi, \dots)$, $\xi = 10^{-5}$. On the left for $N = 50$ even, we find for odd n that the Ritz value 0 is not close to any of the eigenvalues (Case 2), and for even n Case 4 occurs. On the right for $N = 51$ odd, we find for odd n that the Ritz value 0 hits the exceptional eigenvalue 0 (Case 5), and for even n there are two Ritz values not so close to the exceptional eigenvalue 0 (Case 3), that is, there is a delay in the convergence.*

In order to illustrate the phenomenon of exceptional indices, we show in Figure 5.3 some numerical experiences for equidistant eigenvalues $\lambda_{j,N} = 2j/(N+1) - 1$ on $[-1, 1]$, equal eigencomponents $w_N(\lambda_{j,N}) = 1/\sqrt{N}$, and the symmetric pole sequence $(\xi, -\xi, \xi, -\xi, \dots)$ for $\xi = 10^{-5}$. We have chosen a smaller $N \in \{50, 51\}$ in order to be able to distinguish the different Ritz values visually. For odd n, one observes by symmetry that the nth orthogonal rational function of (2.7) is odd, and hence Cases 2 and 5 must occur around the origin depending on the parity of N. For even n, this orthogonal rational function is "nearly" even, giving raise to the exceptional Cases 3 and 4.

5.2. Proof of Theorem 3.2. In view of Proposition 5.3, we have to show that the statement of Proposition 5.1 still remains valid even if we replace **(H6)** by the much weaker Assumption **(H6')**. However, if the poles in Ξ_N are allowed to approach the set of eigenvalues Λ_N then in general the limits (5.4) and (5.5) are no longer true. Here (5.4) is not essential: from the principle of descent (A.1) we obtain the weaker relation

$$\limsup_{N\to\infty} \log(|Q_N(\lambda_{k(N),N})|^{1/N}) \leq -U^{\nu}(\lambda), \tag{5.15}$$

which is sufficient for the conclusions of Proposition 5.1. But, if the distance between the set of poles Ξ_N and $\lambda_{j(N),N}$ decays exponentially in N, then (5.5) will be wrong.

By definition of $j(N)$ in the proof of Proposition 5.1 we have $\lambda_{j(N),N} \in \Lambda_N \setminus Z_N$, so that at least the eigenvalues being very close to poles should be part of the set Z_N of zeros of the polynomial S_N. Remember that an important ingredient in the proof of Proposition 5.1 was that eigenvalues in $\mathbb{R} \setminus V = \{z : U^{\mu-\nu}(z) < F - \eta\}$ are part of Z_N. Thus there is already a first question whether there are not too many such eigenvalues, since Z_N should contain at most $n(N) - 2 \approx tN$ elements. In addition, a single criterion on $\mathrm{dist}(\lambda_{j,N}, \Xi_N)$ is probably not sufficient for insuring (5.5).

In this paper we suggest to make a link with the separation condition **(H3)** of eigenvalues, but this requires to specify more precisely the way how the Nth poles cluster around the Nth eigenvalues. We show in Lemma 5.6 that if we include in Z_N all critical eigenvalues in the sense of Definition 5.5, then we may insure a limit relation as in (5.5). The number of critical eigenvalues, at least far from $\mathrm{supp}(\nu_0)$ and from the boundary of $\mathrm{supp}(\sigma)$ with ν_0 as in Lemma A.1(e), can be monitored by the pole measure ν with the help of Lemma 5.7, and thus by μ according to the first part of Lemma A.1(e).

It turns out that a major technical difficulty in our approach comes from the fact that $\mathrm{supp}(\nu_0)$ is not necessarily separated[4] from $\mathrm{supp}(\sigma - \mu)$, and hence the neighborhood V from the proof of Proposition 5.1 might contain parts of $\mathrm{supp}(\nu_0)$. We will include in Z_N all eigenvalues in a neighborhood of $\mathrm{supp}(\nu_0)$, which might lead to an overshoot, i.e., we are no longer able to discretize the equilibrium measure μ but only a measure $\mu^\eta \in \mathcal{M}_t^\sigma$ with $U^{\mu-\mu^\eta}$ being "small". The limit $\eta \to 0$ will enable us to conclude.

For a moment let us suppose that we have found

$$Z_N \subset \Lambda_N \setminus \{\lambda_{k(N),N}\} \quad \text{such that} \quad \chi_N(Z_N) \to \mu^\eta \in \mathcal{M}_t^\sigma, \tag{5.16}$$

the construction of μ^η to be specified in Lemma 5.8, and the discretization procedure in our proof of Theorem 3.2 below.

Recall that Λ_N contains the elements $\lambda_{1,N} < \cdots < \lambda_{N,N}$. In the next definition, we adapt the convention that $\lambda_{j,N} = -\infty$ if $j \leq 0$ and $\lambda_{j,N} = +\infty$ if $j > N$.

Definition 5.5. *Let $\delta \in (0, 1/4)$ be sufficiently small to be specified later. An eigenvalue $\lambda_{j,N} \in \Lambda_N$ will be called m-*critical *for some integer $m \geq 0$ if the interval*

$$\Big(\frac{\lambda_{j-m-1,N} + \lambda_{j-m,N}}{2}, \frac{\lambda_{j+m,N} + \lambda_{j+m+1,N}}{2}\Big) \cap \Big(\lambda_{j,N} - \delta, \lambda_{j,N} + \delta\Big),$$

*called m-*neighborhood of $\lambda_{j,N}$*, contains $\geq 6m+1$ poles out of Ξ_N counting multiplicities. Define $Z_N^{\mathrm{crit}} \subset \Lambda_N$ to be the set of all critical eigenvalues, that is, eigenvalues which are m-critical for some $m \geq 0$.*

Lemma 5.6. *If $\Lambda_N \setminus Z_N^{\mathrm{crit}} \ni \lambda_{\ell(N),N} \to \widetilde{\lambda}$ then*

$$\lim_{N\to\infty} \log(|Q_N(\lambda_{\ell(N),N})|^{1/N}) = -U^\nu(\widetilde{\lambda}).$$

Proof. Let $\epsilon > 0$. Note that eigenvalues being not critical for $\delta = \delta_0$ remain non-critical for all $\delta \leq \delta_0$. Hence we may suppose without loss of generality that $\delta > 0$ is as in Lemma A.4, and that in addition $\nu(\partial J) = 0$ for the interval $J = [\widetilde{\lambda} - \delta/2, \widetilde{\lambda} + \delta/2]$, since this is true for almost all $\delta > 0$.

[4]Of course, we could add in our assumptions the condition $\mathrm{supp}(\nu_0) \cap \mathrm{supp}(\sigma - \mu) = \emptyset$, however, this condition seems for us to be too restrictive and difficult to verify.

Let us show that we are able to monitor the poles from Ξ_N in an open interval $(\lambda_{\ell(N),N}, \lambda_{\ell(N),N} + \delta)$. We introduce the partition $I_{m,\ell(N),N}$ defined for $m \geq 0$ by

$$\Big[\frac{\lambda_{\ell(N)+m-1,N} + \lambda_{\ell(N)+m,N}}{2}, \frac{\lambda_{\ell(N)+m,N} + \lambda_{\ell(N)+m+1,N}}{2}\Big) \cap \Big(\lambda_{\ell(N),N}, \lambda_{\ell(N),N} + \delta\Big).$$

Note that for $m \geq 1$ there holds

$$\lambda_{\ell(N)+m,N} - \lambda_{\ell(N),N} \leq 2\Big(\frac{\lambda_{\ell(N)+m-1,N} + \lambda_{\ell(N)+m,N}}{2} - \lambda_{\ell(N),N}\Big) \leq 2\delta,$$

and hence for $\xi \in I_{m,\ell(N),N}$ we have

$$\begin{aligned} \log \frac{1}{|\lambda_{\ell(N),N} - \xi|} &\leq \log \frac{1}{|\lambda_{\ell(N),N} - \frac{\lambda_{\ell(N)+m-1,N}+\lambda_{\ell(N)+m,N}}{2}|} \\ &\leq \log \frac{2}{|\lambda_{\ell(N),N} - \lambda_{\ell(N)+m,N}|} \leq 2\log \frac{1}{|\lambda_{\ell(N),N} - \lambda_{\ell(N)+m,N}|}, \end{aligned}$$

the last inequality following from $2 \leq 1/(2\delta) \leq 1/|\lambda_{\ell(N)+m,N} - \lambda_{\ell(N),N}|$.

By Definition 5.5, there are at most $6m$ poles out of Ξ_N on the right of $\lambda_{\ell(N),N}$ in an m-neighborhood of $\lambda_{\ell(N),N}$ for any $m \geq 0$, that is, in the union $I_{0,\ell(N),N} \cup I_{1,\ell(N),N} \cup \cdots \cup I_{m,\ell(N),N}$. By shifting if necessary these poles a bit to the left, we obtain a situation where at most 6 poles lie in an interval $I_{m,\ell(N),N}$ for $m \geq 1$, and no poles in $I_{0,\ell(N),N}$. It follows that, for $N \geq 1$,

$$\begin{aligned} 0 &\leq \frac{1}{N} \sum_{\xi\in\Xi_N\cap(\lambda_{\ell(N),N},\lambda_{\ell(N),N}+\delta)} \log \frac{1}{|\lambda_{\ell(N),N} - \xi|} \\ &= \frac{1}{N} \sum_{m=1, I_{m,\ell(N),N}\neq\emptyset}^{\infty} \sum_{\xi\in\Xi_N\cap I_{m,\ell(N),N}} \log \frac{1}{|\lambda_{\ell(N),N} - \xi|} \\ &\leq \frac{12}{N} \sum_{\lambda_{\ell(N),N}<\lambda_{j,N}\leq\lambda_{\ell(N),N}+2\delta} \log \frac{1}{|\lambda_{\ell(N),N} - \lambda_{j,N}|} \leq 12\epsilon, \end{aligned}$$

where in the last step we have applied the inequality (A.9) of Lemma A.4. A similar conclusion is obtained for the poles in $\Xi_N \cap (\lambda_{\ell(N),N} - \delta, \lambda_{\ell(N),N})$ on the left of $\lambda_{\ell(N),N}$.

We write $-\log(|Q_N(\lambda_{\ell(N),N})|^{1/N}) = U_{1,N} + U_{2,N}$ with

$$U_{1,N} := U^{\chi_N(\Xi_N\setminus J)}(\lambda_{\ell(N),N}), \quad U_{2,N} := U^{\chi_N(\Xi_N\cap J)}(\lambda_{\ell(N),N}),$$

and note that $U_{1,N} \to U^{\nu}(\widetilde{\lambda}) - U^{\nu|_J}(\widetilde{\lambda})$ for $N \to \infty$ by **(H4)**, (A.2), and the assumption $\nu(\partial J) = 0$. Also, from above it follows that $0 \leq U_{2,N} \leq 24\epsilon$, and, taking into account (A.1), we conclude that $0 \leq U^{\nu|_J}(\widetilde{\lambda}) \leq 24\epsilon$. Thus, for $\epsilon \to 0$ we obtain the claim of Lemma 5.6. ∎

LEMMA 5.7. *Denote by $B(\delta)$ the closed subset of* supp(σ) *of points having a distance $\geq 2\delta$ to ∂*supp(σ) *and* supp(ν_0)*, with ν_0 as in Lemma A.1(e). Then for any ρ being a weak$*$ accumulation point of the sequence of normalized counting measures $\chi_N(Z_N^{\text{crit}})$ we have $\rho|_{B(\delta)} \leq \nu|_{B(\delta)}$.*

Proof. For each $\lambda_{j,N} \in Z_N^{\text{crit}} \cap B(\delta)$ there exists a minimal $m = m(j,N) \geq 0$ such that $\lambda_{j,N}$ is m-critical. The at least $6m+1$ poles in the m-neighborhood of $\lambda_{j,N}$ are elements of m-neighborhoods of at most $4m$ other eigenvalues $\lambda_{k,N}$, since a nonempty intersection of the corresponding m-neighborhoods implies that $1 \le |j-k| \le 2m$. As a consequence, for each $\lambda_{j,N} \in Z_N^{\mathrm{crit}} \cap B(\delta)$ we may select a pole $\xi_{\ell(j,N),N} \in \Xi_N$ lying in the $m(j,N)$-neighborhood of $\lambda_{j,N}$, and these selected poles all have distinct indices. We claim that

$$\lim_{N\to\infty} \max\{|\lambda_{j,N} - \xi_{\ell(j,N),N}| : \lambda_{j,N} \in Z_N^{\mathrm{crit}} \cap B(\delta)\} = 0. \tag{5.17}$$

Suppose that (5.17) is wrong. By **(H1)**, we may pass to subsequences if necessary and obtain $\lambda_{\ell_N,N} \in Z_N^{\mathrm{crit}} \cap B(\delta)$ with $\lambda_{\ell_N,N} \to a$, and the lower and upper bounds of their $m(\ell_N,N)$-neighborhoods tending to some $[b,c] \subset [a-\delta, a+\delta] \subset \mathrm{supp}(\sigma) \setminus \mathrm{supp}(\nu_0)$, with $a \in [b,c] \cap B(\delta)$, and $c-b>0$. By construction, these neighborhoods contain at most $2m(\ell_N,N)+1$ eigenvalues. Since σ has no mass points, we conclude that

$$\begin{aligned}\sigma([b,c]) &= \limsup_{N\to\infty} \frac{\text{number of eigenvalues in } m(\ell_N,N)\text{–neighborhood of } \lambda_{\ell_N,N}}{N}\\ &\le \limsup_{N\to\infty} \frac{2m(\ell_N,N)+1}{N},\end{aligned}$$

whereas

$$\begin{aligned}\nu([b,c]) &\ge \limsup_{N\to\infty} \frac{\text{number of poles in } m(\ell_N,N)\text{–neighborhood of } \lambda_{\ell_N,N}}{N}\\ &\ge \limsup_{N\to\infty} \frac{6m(\ell_N,N)+1}{N} \ge 3\sigma([b,c]).\end{aligned}$$

By construction of $B(\delta)$ we know that $(b+c)/2 \in \mathrm{supp}(\sigma)$ and hence $(\nu-\sigma)([b,c]) \ge 2\sigma([b,c]) > 0$. Hence from the Jordan decomposition we get $\nu_0([b,c]) > 0$, in contradiction with the fact that the interval $[a-\delta, a+\delta]$ has an empty intersection with $\mathrm{supp}(\nu_0)$ by construction of $B(\delta)$. Thus (5.17) holds. Let now f be a continuous and nonnegative function, then it is uniformly continuous on the compact set Λ, and (5.17) implies that

$$\begin{aligned}\int f(x)\,\mathrm{d}\rho|_{B(\delta)}(x) &\le \limsup_{N\to\infty} \sum_{\lambda_{j,N}\in Z_N^{\mathrm{crit}}\cap B(\delta)} \frac{f(\lambda_{j,N})}{N}\\ &= \limsup_{N\to\infty} \sum_{\lambda_{j,N}\in Z_N^{\mathrm{crit}}\cap B(\delta)} \frac{f(\xi_{\ell(j,N),N})}{N}\\ &\le \lim_{N\to\infty} \sum_{\xi_{j,N}\in\Xi_N} \frac{f(\xi_{j,N})}{N} = \int f(x)\,\mathrm{d}\nu(x),\end{aligned}$$

where in the first inequality we have used the definition of ρ, and in the second inequality the facts that f is nonnegative and that our selected poles $\xi_{\ell(j,N),N}$ have distinct indices. Hence $\rho|_{B(\delta)} \le \nu$, as claimed in Lemma 5.7. ∎

Lemma 5.8. *Let ν_0 be as in Lemma A.1(e), where we suppose that $\nu_0 \ne 0$ and $\mathrm{supp}(\sigma)$ and $\mathrm{supp}(\nu_0)$ are finite unions of closed intervals. Let V be a closed neighborhood of $\mathrm{supp}(\sigma-\mu)$, and define for $\eta>0$ the sets $V(\eta) = V \cap B(\eta)$, with $B(\eta)$ as in Lemma 5.7, and $B(\eta)^c = \mathrm{supp}(\sigma)\setminus B(\eta)$, $V(\eta)^c = \mathrm{supp}(\sigma)\setminus V(\eta)$. Then, for sufficiently small η, the measure*

$$\mu^\eta := \sigma|_{V(\eta)^c} + \nu|_{V(\eta)} + \Big(1 - \frac{(\sigma-\mu)(V(\eta)^c)}{(\mu-\nu)(V(\eta))}\Big)(\mu-\nu)|_{V(\eta)}$$

is well-defined, $\mu^\eta \in \mathcal{M}_t^\sigma$*, and*

$$\limsup_{\eta\to 0} \max_{z\in\Lambda} |U^{\mu^\eta-\mu}(z)| = 0.$$

Proof. By assumption on $\mathrm{supp}(\sigma)$ and $\mathrm{supp}(\nu_0)$, the set $B(\eta)^c \setminus \mathrm{supp}(\nu_0)$ consists of a finite number of intervals, and, for sufficiently small η, these intervals are disjoint and of length $\le 2\eta$, and their number, say, k, does not depend on η.

We first show that $(\mu-\nu)(V(\eta)) > 0$ for sufficiently small $\eta > 0$. Write shorter $\mu_0 = (\mu-\nu)|_{\mathbb{R}\setminus\mathrm{supp}(\nu_0)}$, then $\mu_0 \ge 0$ by Lemma A.1(e), and $(\mu-\nu)|_{V(\eta)} = \mu_0|_{V(\eta)}$ by construction of $B(\eta)$. Since trivially $\mu_0 \le \sigma$, we find that

$$(\mu-\nu)(V(\eta)) = \mu_0(V(\eta)) \ge \mu_0(V) - \mu_0(B(\eta)^c) \ge \mu_0(V) - \sigma(B(\eta)^c \setminus \mathrm{supp}(\nu_0)).$$

Since σ has no mass points, we find according to the particular structure of $B(\eta)^c \setminus \mathrm{supp}(\nu_0)$ that $\sigma(B(\eta)^c \setminus \mathrm{supp}(\nu_0)) \to 0$ for $\eta \to 0$. Also, by Lemma A.1(e) we have that $\mathrm{supp}(\sigma-\mu) \subset \mathrm{supp}(\sigma_0) \subset \mathbb{R} \setminus \mathrm{Int}(\mathrm{supp}(\nu_0))$, and, since $\partial\,\mathrm{supp}(\nu_0)$ is finite but not $\mathrm{supp}(\sigma-\mu)$, there exists $\lambda \in \mathrm{supp}(\sigma-\mu)\setminus\mathrm{supp}(\nu_0)$. Again from Lemma A.1(e) we conclude that $\lambda \in \mathrm{supp}(\mu_0)$, and V is a neighborhood of λ, implying that $\mu_0(V) > 0$, and thus $\mu_0(V(\eta)) > 0$ for sufficiently small $\eta > 0$.

Secondly, we recall that $\mathrm{supp}(\sigma-\mu) \subset \mathbb{R} \setminus \mathrm{Int}(\mathrm{supp}(\nu_0))$, which implies that $(\sigma-\mu)(\mathrm{supp}(\nu_0)) = 0$. It follows that $0 \le (\sigma-\mu)(V(\eta)^c) = (\sigma-\mu)(V(\eta)^c\setminus\mathrm{supp}(\nu_0)) \le \sigma(B(\eta)^c \setminus \mathrm{supp}(\nu_0))$, where the right-hand term tends to 0 for $\eta \to 0$. Thus

$$\lim_{\eta\to 0} \frac{(\sigma-\mu)(V(\eta)^c)}{(\mu-\nu)(V(\eta))} = 0. \tag{5.18}$$

We therefore have shown that, for sufficiently small $\eta > 0$, the measure μ^η is well-defined, is positive (since $(\mu-\nu)|_{V(\eta)} = \mu_0|_{V(\eta)} \ge 0$), satisfies the inequality $\mu^\eta \le \sigma|_{V(\eta)^c} + \nu|_{V(\eta)} + (\mu-\nu)|_{V(\eta)} \le \sigma$, and

$$\mu^\eta(\mathbb{R}) = \sigma(V(\eta)^c) + \nu(V(\eta)) + (\mu-\nu)(V(\eta)) - (\sigma-\mu)(V(\eta)^c) = \mu(\mathbb{R}) = t,$$

and thus $\mu^\eta \in \mathcal{M}_t^\sigma$.

It remains to analyze the potential of

$$\mu^\eta - \mu = (\sigma-\mu)|_{V(\eta)^c} - \frac{(\sigma-\mu)(V(\eta)^c)}{(\mu-\nu)(V(\eta))}(\mu-\nu)|_{V(\eta)},$$
$$\text{where} \quad (\sigma-\mu)|_{V(\eta)^c} = (\sigma-\mu)|_{V(\eta)^c\setminus\mathrm{supp}(\nu_0)}.$$

We want to show that the maximum of the potential on Λ of each of the two measures in this difference tends to zero for $\eta \to 0$. For the first potential, $(\sigma-\mu)(V(\eta)^c)$ tends to 0 for $\eta \to 0$ by (5.18), and hence

$$\liminf_{\eta\to 0} \max_{z\in\Lambda} U^{(\sigma-\mu)|_{V(\eta)^c}}(z) \ge 0$$

by (A.1). On the other hand, since the set $B(\eta)^c\setminus\mathrm{supp}(\nu_0)$ larger than $V(\eta)^c\setminus\mathrm{supp}(\nu_0)$ can be written as a union $J_1 \cup \cdots \cup J_k$ of disjoint intervals J_ℓ of length $\le 2\eta \le 1$, we get from the maximum principle for logarithmic potentials [38, Corollary II.3.3]

$$\max_{z\in\Lambda} U^{(\sigma-\mu)|_{V(\eta)^c}}(z) \le \sum_{\ell=1}^{k} \sup_{x\in J_\ell} U^{(\sigma-\mu)|_{V(\eta)^c\cap J_\ell}}(x) \le \sum_{\ell=1}^{k} \sup_{x\in J_\ell} U^{\sigma|_{J_\ell}}(x),$$

where the right-hand side tends to 0 for $\eta \to 0$ by Lemma A.4. In order to discuss the second potential, we write

$$V(\eta) = V(\eta) \setminus \operatorname{supp}(\nu_0) = [V \setminus \operatorname{supp}(\nu_0)] \setminus \bigcap_{\ell=1}^{k} (J_\ell \cap V)$$

and get as before from (A.1) and Lemma A.4 that

$$\begin{aligned} &\limsup_{\eta\to 0} \max_{x\in\Lambda} |U^{(\mu-\nu)|_{V(\eta)}}(x) - U^{(\mu-\nu)|_{V\setminus \operatorname{supp}(\nu_0)}}(x)| \\ &= \limsup_{\eta\to 0} \max_{x\in\Lambda} \Big| \sum_{\ell=1}^{k} U^{(\mu-\nu)|_{J_\ell\cap V}}(x) \Big| = 0. \end{aligned}$$

By Lemma A.1(e), $0 \le (\mu-\nu)|_{V\setminus\operatorname{supp}(\nu_0)} \le \sigma$ and hence its potential is continuous. Thus $|U^{(\mu-\nu)|_{V(\eta)}}(x)|$ is uniformly bounded for $x \in \Lambda$ and sufficiently small $\eta \ge 0$, and (5.18) yields the last claim of Lemma 5.8. □

We are now prepared to conclude in our proof of Theorem 3.2. If $\nu \le \sigma$ and thus $\nu \in \mathcal{M}_t^\sigma$ then $U^{\mu-\nu}(x) - F = 0$ for all $x \in \mathbb{C}$ by Lemma A.1(b). Since all Ritz values lie in the convex hull of the eigenvalues and hence in the convex hull of Λ being compact by Assumption **(H1)**, we conclude that

$$\limsup_{N\to\infty} \operatorname{dist}(\Lambda_N, \Theta_N)^{1/N} \le 1, \quad \limsup_{N\to\infty} w_N(\lambda_{k(N),N})^{-1/N} \ge 1,$$

the latter relation following from $w_N(\lambda_{k(N),N}) \le 1$, and the assertion of Theorem 3.2 is trivial.

Suppose now that $\nu \not\le \sigma$, and thus $\nu_0 \ne 0$ in the Jordan decomposition $\nu - \sigma = \nu_0 - \sigma_0$. According to Proposition 5.3, we only have to show that the relation (5.2) of Proposition 5.1 holds. Let $\epsilon > 0$, and define

$$V = V(\epsilon) = \{x \in \mathbb{R} : U^{\mu-\nu}(x) \ge F - \epsilon\}.$$

Note that U^ν is finite on $V(2\epsilon)$, and thus $U^{\mu-\nu}$ is continuous on $V(2\epsilon)$ by the assumption of Theorem 3.2. It follows from Lemma A.1(c) that V is a closed neighborhood of $\operatorname{supp}(\sigma - \mu)$. We now choose $\eta > 0$ sufficiently small such that the measure μ^η of Lemma 5.8 is an element of $\mathcal{M}_t^\sigma$ and $|U^{\mu^\eta}(x) - U^\mu(x)| \le \epsilon$ for all $x \in \Lambda$.

We may apply the discretization procedure of Lemma A.5 with

$$Z_{3,N} = \Lambda_N \setminus \{\lambda_{k(N),N}\}, \quad Z_{1,N} = \{\lambda_{j,N} \in Z_{3,N} : \lambda_{j,N} \in V(\eta)^c \cup Z_N^{\text{crit}}\}, \quad \rho_2 = \mu^\eta,$$

and $i(N) = n(N)-2$. Indeed, with the notation of Lemma A.5, the inequality $\rho_1 \lneqq \rho_2$ follows from Lemma 5.7 with $\delta = \eta$, and $\rho_2 \lneqq \rho_3$ is a consequence of Lemma 5.8. This leads to $Z_N := Z_{2,N}$ satisfying $\chi_N(Z_N) \to \mu^\eta$, as claimed in (5.16). Defining

$$S_N(x) = \prod_{\lambda\in Z_N} (x - \lambda),$$

we note that the maximum of $|S_N/Q_N|$ over $\Lambda_N \setminus \{\lambda_{k(N),N}\}$ is attained at some $\lambda_{j(N),N} \in Z_{3,N} \setminus Z_N \subset Z_{3,N} \setminus Z_{1,N}$, that is, $\lambda_{j(N),N} \ne \lambda_{k(N),N}$ is an element of $\Lambda_N \cap V$ being non-critical for $\delta = \eta$. By passing to subsequences if necessary, suppose that $\lambda_{j(N),N} \to \widetilde{\lambda}$, and thus $\widetilde{\lambda} \in V$. Then the principle of descent (A.1) for S_N together with Lemma 5.6 for Q_N gives

$$\limsup_{N\to\infty} \log(|S_N(\lambda_{j(N),N})/Q_N(\lambda_{j(N),N})|^{1/N}) \le U^{\nu-\mu^\eta}(\widetilde{\lambda}) \le U^{\nu-\mu}(\widetilde{\lambda}) + \epsilon \le -F + 2\epsilon,$$

whereas the eigenvalue separation **(H3)** for S_N (compare with (5.7)) and the principle of descent (A.1) for Q_N (compare with (5.15)) leads to

$$\liminf_{N\to\infty} \log(|S_N(\lambda_{k(N),N})/Q_N(\lambda_{k(N),N})|^{1/N}) \ge U^{\nu-\mu^\eta}(\lambda) \ge U^{\nu-\mu}(\lambda) - \epsilon.$$

The conclusion follows as in Proposition 5.1 after $\epsilon \to 0$.

5.3. Proof of Corollary 3.3. Consider the matrix $\underline{A}_N = (A-\tau I)^{-1}$, then by assumption on τ we get that its eigenvalues $\underline{\lambda}_{k,N} = 1/(\lambda_{k,N}-\tau)$ lie in some compact interval $\underline{\Lambda} \subset (0,+\infty)$. Similarly, the new poles $\underline{\xi}_{k,N} = 1/(\xi_{k,N}-\tau)$ lie in some compact set $\underline{\Xi}$. Defining the corresponding denominator polynomials

$$\underline{q}_{j,N}(\underline{x}) = \prod_{k=1}^{j} (\underline{x} - \underline{\xi}_{k,N}),$$

we may write according to (2.5)

$$\mathbf{v}_{j+1,N} = q_{j,N}(A_N)^{-1} p_{j,N}(A_N) \frac{\mathbf{b}_N}{\|\mathbf{b}_N\|} = \underline{q}_{j,N}(\underline{A}_N)^{-1} \underline{p}_{j,N}(\underline{A}_N) \frac{\mathbf{b}_N}{\|\mathbf{b}_N\|}$$

with suitable polynomials $\underline{p}_{j,N}$ of degree $\le j$. In other words, the rational Krylov space $\mathcal{K}_n^{\mathrm{rat}}(A_N, \mathbf{b}_N)$ with poles $\xi_{1,N},\dots,\xi_{n-1,N}$ coincides with the rational Krylov space $\mathcal{K}_n^{\mathrm{rat}}(\underline{A}_N, \mathbf{b}_N)$ with poles $\underline{\xi}_{1,N},\dots,\underline{\xi}_{n-1,N}$.

However, the corresponding set $\underline{\Theta}_N$ of transformed nth rational Ritz values $\underline{\theta}_{k,N} = 1/(\theta_{k,N}-\tau) \in [a,b]$ does not give the nth rational Ritz values for $\underline{A}_N$ since, according to (2.7),

$$\widetilde{\mathbf{v}}_{n+1,N} = q_{n-1,N}(A_N)^{-1} \widetilde{p}_{n,N}(A_N) \frac{\mathbf{b}_N}{\|\mathbf{b}_N\|} = \underline{q}_{n,N}(\underline{A}_N)^{-1} \underline{\widetilde{p}}_{n,N}(\underline{A}_N) \frac{\mathbf{b}_N}{\|\mathbf{b}_N\|}$$

with suitable polynomials $\underline{\widetilde{p}}_{n,N}$ of degree $\le n$, that is, instead of the pole $\underline{\xi}_{n,N} = \infty$ we have a finite pole $\underline{\xi}_{n,N} = 0$ (like for harmonic rational Ritz values).

Observe that

$$\underline{q}_{n,N}(\underline{x})\underline{q}_{n-1,N}(\underline{x}) = \underline{x}\underline{q}_{n-1,N}(\underline{x})^2 > 0$$

for all $\underline{x} \in \underline{\Lambda}$ containing the spectrum $\underline{\Lambda}_N$ of $\underline{A}_N$. Thus our basic tool Lemma 2.3 may be generalized: $\underline{p}_{n,N}$ is a discrete orthogonal polynomial with the new weights $w_N(\lambda_{j,N})^2/(\underline{q}_{n,N}(\underline{\lambda}_{j,N})\underline{q}_{n-1,N}(\underline{\lambda}_{j,N}))$, and we obtain upper bounds for $(\underline{\lambda}_{k(N),N} - \underline{\theta}_{\kappa(N)-1,N})(\underline{\theta}_{\kappa(N),N} - \underline{\lambda}_{k(N),N})$ with a similar substitution. In particular, the results of Theorem 3.1 and Theorem 3.2 on $\mathrm{dist}(\underline{\lambda}_{k(N),N}, \underline{\Theta}_N)$ remain valid. Passing to the chordal metric we get

$$\begin{aligned}\mathrm{dist}_{\mathrm{chord}}(\lambda_{k(N),N}, \Theta_N) &= \min_j \mathrm{chord}(\lambda_{k(N),N}, \theta_{j,N}) = \min_j \frac{|\lambda_{k(N),N} - \theta_{j,N}|}{\sqrt{1+\lambda_{k(N),N}^2}\sqrt{1+\theta_{j,N}^2}} \\ &= \min_j \frac{|\underline{\lambda}_{k(N),N} - \underline{\theta}_{j,N}|}{\sqrt{\underline{\lambda}_{k(N),N}^2 + (1+\tau\underline{\lambda}_{k(N),N})^2}\sqrt{\underline{\theta}_{j,N}^2 + (1+\tau\underline{\theta}_{j,N})^2}},\end{aligned}$$

the latter expression being bounded below and above by $\mathrm{dist}(\underline{\lambda}_{k(N),N}, \underline{\Theta}_N)$ times some positive constant, since $\underline{\lambda}_{k(N),N}, \underline{\theta}_{j,N}$ both lie in the compact interval $\underline{\Lambda} \subset (0, \infty)$. Thus the first part of Corollary 3.3 follows.

Finally, in order to see that the rate of convergence is independent of the choice of τ, we consider $\tau_1 \neq \tau_2$ as in Corollary 3.3, and denote by $\underline{\sigma}_j$ and $\underline{\nu}_j$ for $j = 1, 2$ the limit measures resulting from the modified conditions **(H2)** and **(H4)**, and by $\underline{\mu}_j$ the corresponding extremal measures. Then, considering the change of variables

$$x = \tau_1 + 1/\underline{x}_1 = \tau_2 + 1/\underline{x}_2,$$

elementary calculus shows that $\underline{\sigma}_1(\underline{x}_1) = \underline{\sigma}_2(\underline{x}_2) = \sigma(x)$, with $\chi_N(\Lambda_N) \to \sigma$ on the Riemann sphere. Similarly, we have $\underline{\nu}_1(\underline{x}_1) = \underline{\nu}_2(\underline{x}_2) = \nu(x)$, with $\chi_N(\Xi_N) \to \nu$ on the Riemann sphere (note that ν might have a mass point at ∞). In particular, we may define $\underline{\mu}(\underline{x}_2) = \underline{\mu}_1(\underline{x}_1)$ such that $\underline{\mu} \in \mathcal{M}_t^{\underline{\sigma}_2}$, and observe that the potential of $\underline{\mu} - \underline{\nu}_2$ at $\underline{x}_2$ differs from the potential of $\underline{\mu}_1 - \underline{\nu}_1$ at $\underline{x}_1$ only by a constant. From Lemma A.1(d) we conclude that $\underline{\mu}_2 = \underline{\mu}$, and hence the rate of convergence is in fact independent of the choice of τ.

6. Conclusions. We have given a theoretical background for designing numerical methods to compute parts of the spectrum of large sparse matrices. It was shown that rational Ritz values are indeed attracted by the poles. Moreover, our quantification of the rate of convergence via logarithmic potential theory may help to overcome a weak point of the rational Krylov method: how to choose the free parameters, namely the poles, in order to be sure to obtain excellent approximations of particular eigenvalues (like for instance the extremal ones).

Our findings are of an asymptotic nature, and they do not take into account the effect of finite precision arithmetic. Nevertheless, they might be helpful for a better understanding of related methods using, e.g., restarting. There is also some potential impact on improved Krylov subspace techniques for approximating functions of symmetric matrices, and in particular to explain superlinear convergence behavior.

REFERENCES

[1] B. Beckermann, *A note on the convergence of Ritz values for sequences of matrices*, Technical Report ANO 408, Labo Paul Painlevé, Université de Lille I, France, 2000.

[2] ———, *On a conjecture of E.A. Rakhmanov*, Constructive Approximation, 16 (2000), pp. 427–448.

[3] ———, *Discrete orthogonal polynomials and superlinear convergence of Krylov subspace methods in numerical linear algebra*, in Orthogonal Polynomials and Special Functions, F. Marcellan and W. Van Assche, eds., vol. 1883 of Lecture Notes in Mathematics, Springer, 2006, pp. 119–185.

[4] B. Beckermann and S. S. Capizzano, *On the asymptotic spectrum of finite element matrix sequences*, SIAM Journal on Numerical Analysis, 45 (2007), pp. 746–769.

[5] B. Beckermann and A. B. J. Kuijlaars, *On the sharpness of an asymptotic error estimate for conjugate gradients*, BIT, (2001), pp. 856–867.

[6] ———, *Superlinear convergence of conjugate gradients*, SIAM Journal on Numerical Analysis, 39 (2001), pp. 300–329.

[7] ———, *Superlinear CG convergence for special right-hand sides*, Electronic Transactions on Numerical Analysis, 14 (2002), pp. 1–19.

[8] A. Böttcher and B. Silbermann, *Introduction to Large Truncated Toeplitz Matrices*, Universitext, Springer-Verlag, New York, 1999.

[9] A. Bultheel, P. González-Vera, E. Hendriksen, and O. Njåstad, *Orthogonal Rational Functions*, vol. 5 of Cambridge Monographs on Applied and Computational Mathematics, Cambridge University Press, Cambridge, United Kingdom, 1999.

[10] V. Buyarov and E. A. Rakhmanov, *Families of equilibrium measures with external field on the real axis*, Sbornik: Mathematics, 190 (1999), pp. 791–802.

[11] J. Coussement and W. Van Assche, *A continuum limit of relativistic Toda lattice: asymptotic theory of discrete Laurent orthogonal polynomials with varying recurrence coefficients*, Journal of Physics. A., 38 (2005), pp. 3337–3366.

[12] G. De Samblanx, K. Meerbergen, and A. Bultheel, *The implicit application of a rational filter in the RKS method*, BIT, 37 (1997), pp. 925–947.

[13] K. Deckers and A. Bultheel, *Rational Krylov sequences and orthogonal rational functions*, Tech. Rep. TW499, Katholieke Universiteit Leuven, Departement Computerwetenschappen, 2008.

[14] P. D. Dragnev and E. B. Saff, *Constrained energy problems with applications to orthogonal polynomials of a discrete variable*, Journal d'Analyse Mathématique, 72 (1997), pp. 223–259.

[15] T. A. Driscoll, K.-C. Toh, and L. N. Trefethen, *From potential theory to matrix iterations in six steps*, SIAM Review, 40 (1998), pp. 547–578.

[16] V. Druskin and L. Knizhnerman, *Extended Krylov subspaces: Approximation of the matrix square root and related functions*, SIAM Journal on Matrix Analysis and Applications, 19 (1998), pp. 775–771.

[17] Y. Eidelman, L. Gemignani, and I. C. Gohberg, *On the fast reduction of a quasiseparable matrix to Hessenberg and tridiagonal forms*, Linear Algebra and its Applications, 420 (2007), pp. 86–101.

[18] Y. Eidelman, I. C. Gohberg, and V. Olshevsky, *The QR iteration method for Hermitian quasiseparable matrices of an arbitrary order*, Linear Algebra and its Applications, 404 (2005), pp. 305–324.

[19] D. Fasino, *Rational Krylov matrices and QR-steps on Hermitian diagonal-plus-semiseparable matrices*, Numerical Linear Algebra with Applications, 12 (2005), pp. 743–754.

[20] G. H. Golub and C. F. Van Loan, *Matrix Computations*, Johns Hopkins University Press, Baltimore, Maryland, USA, third ed., 1996.

[21] S. Helsen, A. B. J. Kuijlaars, and M. Van Barel, *Convergence of the isometric Arnoldi process*, SIAM Journal on Matrix Analysis and Applications, 26 (2005), pp. 782–809.

[22] M. Kac, W. L. Murdock, and G. Szegő, *On the eigenvalues of certain Hermitian forms*, Indiana University Mathematics Journal (formerly known as Journal of Rational Mechanics and Analysis), 2 (1953), pp. 767–800.

[23] A. B. J. Kuijlaars, *Which eigenvalues are found by the Lanczos method?*, SIAM Journal on Matrix Analysis and Applications, 22 (2000), pp. 306–321.

[24] ———, *Convergence analysis of Krylov subspace iterations with methods from potential theory*, SIAM Review, 48 (2006), pp. 3–40.

[25] A. B. J. Kuijlaars and E. A. Rakhmanov, *Zero distributions for discrete orthogonal polynomials*, Journal of Computational and Applied Mathematics, 99 (1998), pp. 255–274.

[26] A. B. J. Kuijlaars and W. Van Assche, *Extremal polynomials on discrete sets*, Proceedings of the London Mathematical Society, 79 (1999), pp. 191–221.

[27] R. Lehoucq and K. Meerbergen, *Using generalized Cayley transformations within an inexact rational Krylov sequence method*, SIAM Journal on Matrix Analysis and Applications, 20 (1998), pp. 131–148.

[28] E. Levin and E. B. Saff, *Potential theoretic tools in polynomial and rational approximation*, in Harmonic Analysis and Rational Approximation, J.-D. Fournier, ed., vol. LNCIS 327, 2006, pp. 71–94.

[29] K. Meerbergen, *Changing poles in the rational Lanczos method for the Hermitian eigenvalue problem*, Numerical Linear Algebra with Applications, 8 (2001), pp. 33–52.

[30] C. C. Paige, B. N. Parlett, and H. A. van der Vorst, *Approximate solutions and eigenvalue bounds from Krylov subspaces*, Numerical Linear Algebra with Applications, 2 (1995), pp. 115–133.

[31] B. N. Parlett, *The Symmetric Eigenvalue Problem*, Prentice-Hall, Englewood Cliffs, NJ, USA, 1980.

[32] E. A. Rakhmanov, *Equilibrium measure and the distribution of zeros on the extremal polynomials of a discrete variable*, Sbornik: Mathematics, 187 (1996), pp. 1213–1228.

[33] T. Ransford, *Potential Theory in the Complex Plane*, Cambridge University Press, Cambridge, 1995.

[34] A. Ruhe, *Rational Krylov sequence methods for eigenvalue computation*, Linear Algebra and its Applications, 58 (1984), pp. 391–405.

[35] ———, *The Rational Krylov algorithm for nonsymmetric eigenvalue problems, III: Complex shifts for real matrices*, BIT, 34 (1994), pp. 165–176.

[36] ———, *Rational Krylov algorithms for nonsymmetric eigenvalue problems*, in Recent Advances in Iterative Methods, G. H. Golub, A. Greenbaum, and M. Luskin, eds., IMA Volumes in Mathematics and its Applications, Springer-Verlag, New York, 1994, pp. 149–164.
[37] Y. Saad, *Numerical Methods for Large Eigenvalue Problems*, Halstead Press, New York, 1992.
[38] E. B. Saff and V. Totik, *Logarithmic Potentials with External Fields*, Springer-Verlag, Berlin, 1997.
[39] L. N. Trefethen and D. Bau, *Numerical Linear Algebra*, SIAM, Philadelphia, Pennsylvania, USA, 1997.
[40] R. Vandebril, M. Van Barel, and N. Mastronardi, *Matrix Computations and Semiseparable Matrices, Volume I: Linear Systems*, Johns Hopkins University Press, Baltimore, Maryland, USA, 2008.
[41] ———, *Matrix Computations and Semiseparable Matrices, Volume II: Eigenvalue and Singular Value Methods*, Johns Hopkins University Press, 2008.

Appendix A. Tools from potential theory.
With the discrete set Ξ_N and the monic polynomial Q_N as before, we find that

$$\log(|Q_N(z)|^{1/N}) = -U^{\chi_N(\Xi_N)}(z), \quad Q_N(z) = \prod_{\xi_{j,N}\in\Xi_N} (z - \xi_{j,N}),$$

which explains that logarithmic potential theory is the right tool for studying weak asymptotics of polynomials. One important tool is the so-called principle of descent proved, e.g., in [38, Theorem I.6.8]. Since this principle is heavily used in our reasoning, we shortly recall it here: if S is compact and (μ_n) is a sequence of finite positive Borel measures with $\mathrm{supp}(\mu_n) \subset S$ then

$$\text{for } z_n \to z \text{ and } \mu_n \to \mu: \qquad \liminf_{n\to\infty} U^{\mu_n}(z_n) \geq U^{\mu}(z), \tag{A.1}$$

$$\text{for } z_n \to z \notin S \text{ and } \mu_n \to \mu: \qquad \lim_{n\to\infty} U^{\mu_n}(z_n) = U^{\mu}(z). \tag{A.2}$$

Other more sophisticated aspects of potential theory will be hidden in the proofs, but the statements should be accessible also for non-experts.

If U^ν is continuous on $\mathrm{supp}(\sigma)$, our constrained energy problem is classical, see for instance [14] or [2, Theorem 1.1]. Here we discuss a more general case also allowing for ν having mass points in $\mathrm{supp}(\sigma)$.

Lemma A.1. *Assume that σ, ν are finite positive Borel measures with compact support, $0 < t = \|\nu\| < \|\sigma\|$, such that U^σ is continuous, and U^ν is continuous at each point of* $\mathrm{supp}(\sigma)$ *where it is finite. Denote as before by $\mathcal{M}_t^\sigma$ the set of positive Borel measures μ such that $\sigma - \mu \geq 0$, and $\|\mu\| = t$.*

(a) The extremal problem

$$\inf\{I(\mu) - 2I(\mu,\nu) : \mu \in \mathcal{M}_t^\sigma\} \tag{A.3}$$

has a unique minimizer $\mu \in \mathcal{M}_t^\sigma$.

(b) We have $\nu \in \mathcal{M}_t^\sigma$ if and only if $\mu = \nu$ (which in the sequel is excluded).

(c) There exists $F > 0$ such that

$$U^{\mu-\nu}(x) \begin{cases} = F & \text{for } x \in \mathrm{supp}(\sigma - \mu), \\ \leq F & \text{for } x \in \mathbb{C} \setminus \mathrm{supp}(\sigma - \mu). \end{cases} \tag{A.4}$$

Also, $\mathrm{supp}(\mu) = \mathrm{supp}(\sigma)$.

(d) Conversely, if there exist a signed measure $\mu_1 \leq \sigma$ with $\mu_1(\mathrm{supp}(\sigma)) = t$ and $F_1 \in \mathbb{R}$ such that $U^{\mu_1-\nu} \leq F_1$ quasi-everywhere on $\mathrm{supp}(\sigma)$ *and $U^{\mu_1-\nu} \geq F_1$ on* $\mathrm{supp}(\sigma - \mu_1)$*, then $\mu_1 = \mu$ and $F_1 = F$.*

(e) *With the Jordan decomposition* $\nu - \sigma = \nu_0 - \sigma_0$ *there holds* $\mathrm{supp}(\sigma - \mu) \subset \mathrm{supp}(\sigma_0)$. *Moreover, the restriction of* $\mu - \nu$ *onto* $\mathbb{R} \setminus \mathrm{supp}(\nu_0)$ *is a positive measure, with support containing* $\mathrm{supp}(\sigma_0) \setminus \mathrm{supp}(\nu_0)$.

Proof. Set $s := \|\sigma\| - t > 0$, and consider the set $\mathcal{M}$ of positive Borel measures ρ of mass s supported on $E := \mathrm{supp}(\sigma)$. We note that for each $\mu \in \mathcal{M}_t^\sigma$ we have $\sigma - \mu \in \mathcal{M}$, but in general the reciprocal is not true since we drop the constraint $\mu \geq 0$. In a first step let us show that our extremal problem (A.3) is well-posed, and that the constraint $\mu \geq 0$ is not important.

By assumption on ν, σ and the principle of descent (A.1), the external field

$$Q(x) = U^{\nu - \sigma}(x)$$

is continuous in the topology of $\mathbb{R} \cup \{+\infty\}$. Since U^σ is continuous, we get from the Fubini theorem

$$\int U^\nu(x)\, \mathrm{d}\sigma(x) = \int U^\sigma(y)\, \mathrm{d}\nu(y) < \infty.$$

Hence both U^ν and Q are finite in at least one point $x_0 \in \mathrm{supp}(\sigma)$, and thus in a neighborhood V of x_0. Since $\sigma(V) > 0$ by the definition of the support and $I(\sigma|_V) < \infty$, we may conclude that $V \cap \mathrm{supp}(\sigma)$ has positive logarithmic capacity, and thus Q is admissible in the sense of [38, Definition I.1.1] for the set E. Following [38, Theorem I.1.3], we consider the problem of minimizing the weighted energy

$$\inf_{\rho \in \mathcal{M}} I_Q(\rho), \quad \text{where} \quad I_Q(\rho) = I(\rho) + 2\int Q(x)\, \mathrm{d}\rho(x) \in (-\infty, +\infty]. \tag{A.5}$$

Note that (A.5) is not the dual problem for (A.3) in the sense of [14, Corollary 2.10] where one finds the additional constraint $\rho \leq \sigma$. As shown in [38, Theorem I.1.3], there exists a unique minimizer $\rho \in \mathcal{M}$ and a constant $F \in \mathbb{R}$ such that

$$U^\rho(x) + Q(x) = U^{\nu-\sigma+\rho}(x) \begin{cases} \geq -F & \text{for quasi every } x \in E, \\ \leq -F & \text{for } x \in \mathrm{supp}(\rho). \end{cases} \tag{A.6}$$

Here quasi-everywhere means everywhere up to an exceptional set E_1 of logarithmic capacity zero, which in our case may be dropped. This can be seen as follows: the set $M_n := \{x \in \mathbb{C} : U^{\nu-\sigma+\rho}(x) \leq -F - 1/n\}$ is closed since potentials are lower semi-continuous and U^σ is continuous. From (A.6) we know that the Borel set $E \cap M_n$ has zero capacity. Since σ has finite energy, σ is C-absolutely continuous [38, Definition II.4.5], implying that $\sigma(E \cap M_n) = \sigma(M_n) = 0$. In other words, we find that the inequality $U^{\nu-\sigma+\rho} \geq -F - 1/n$ holds σ-everywhere, and the principle of domination [38, Theorem II.3.2] tells us that this inequality is true in the whole complex plane. Taking the limit $n \to \infty$, we obtain the stronger equilibrium property

$$U^\rho(x) + Q(x) = U^{\nu-\sigma+\rho}(x) \begin{cases} \geq -F & \text{for } x \in \mathbb{C}, \\ = -F & \text{for } x \in \mathrm{supp}(\rho). \end{cases} \tag{A.7}$$

As a consequence, $S_0 = \{x \in \mathbb{C} : U^\nu(x) = +\infty\}$ has an empty intersection with $\mathrm{supp}(\rho)$, and applying the de la Vallée-Poussin Theorem [38, Theorem IV.4.5] to the set $\Omega = \mathbb{C} \setminus S_0$ we conclude that the restriction of $\rho + \nu$ onto $\mathrm{supp}(\rho)$ is $\leq \sigma$, and $\nu \geq 0$ implies that $\rho \leq \sigma$, or, in other words, $\sigma - \rho \in \mathcal{M}_t^\sigma$.

Note also that

$$I_Q(\sigma - \mu) = I(\mu) - 2I(\mu, \nu) + \int U^\sigma(x)\, \mathrm{d}(2\nu - \sigma)(x),$$

that is, the two weighted energy expressions in (A.3) for μ and (A.5) for $\rho = \sigma - \mu$ only differ by a finite constant. It follows that the unique minimizer ρ of (A.5) is such that $\sigma - \rho = \mu$ is the unique minimizer in (A.3), implying our claim (a).

If now μ_1 is as in part (d), then $\rho = \sigma - \mu_1 \in \mathcal{M}$ satisfies (A.6) with F_1 instead of F. Together with [38, Theorem I.3.1] we may conclude that ρ is the minimizer in (A.5), $F = F_1$, and thus $\sigma - \rho = \mu_1$ is the minimizer in (A.3).

We pursue with a proof of part (b): obviously, $\mu = \nu$ implies that $\nu \in \mathcal{M}_t^\sigma$. Conversely, $\nu \in \mathcal{M}_t^\sigma$ together with the Lemma of Rakhmanov [32] implies that U^ν is continuous, and $I(\nu) < \infty$. Then

$$I(\mu) - 2I(\mu, \nu) = I(\mu - \nu) - I(\nu).$$

Since elements in $\mathcal{M}_t^\sigma$ have the same mass, we may apply [38, Lemma I.1.8] to conclude that $I(\mu - \nu) \geq 0$, with equality if and only if $\mu = \nu$. Hence the minimizer of (A.3) must be $\mu = \nu$.

For a proof of (c) we exclude the trivial case $\mu = \nu$, and thus, again by [38, Lemma I.1.8], $U^{\nu-\mu}$ is not the constant 0 on $\mathbb{C}$. Recalling that $E = \operatorname{supp}(\sigma)$, we get from (A.7) that the minimum of $U^{\nu-\mu} = U^{\nu+\rho-\sigma}$ on $\operatorname{supp}(\mu)$ is $\geq -F$, and $U^{\nu-\mu}$ is a non-constant and superharmonic function in the domain $\mathbb{C} \setminus \operatorname{supp}(\mu)$ including infinity, taking the value 0 at infinity. From the minimum principle for superharmonic functions we may conclude that $-F < 0$. The same argument shows that for all $x \in E \setminus \operatorname{supp}(\mu)$ we must have $U^{\nu-\mu}(x) > -F$, but $E \setminus \operatorname{supp}(\mu) = \operatorname{supp}(\sigma) \setminus \operatorname{supp}(\sigma - \rho) \subset \operatorname{supp}(\rho)$ and thus $U^{\nu-\mu}(x) = -F$ by (A.7), which is a contradiction. Thus $\operatorname{supp}(\mu) = \operatorname{supp}(\sigma)$, and (A.4) follows from (A.7).

For a proof of (e), note first that by construction $\nu - \nu_0 = \sigma - \sigma_0 \geq 0$. We have $\|\nu_0\| \neq 0$ since otherwise $\nu \in \mathcal{M}_t^\sigma$. Also, if $\nu_0 = \nu$ then the restriction of ν onto $\mathbb{R} \setminus \operatorname{supp}(\nu_0)$ is the zero measure, and assertion (e) becomes trivial. Therefore it remains to consider the case $t_0 := t - \|\nu - \nu_0\| = \|\nu_0\| \in (0, \|\sigma_0\|)$. Thus the assumptions of Lemma A.1 are true for the triple (t_0, ν_0, σ_0) instead of the triple (t, ν, σ). Denoting by $\mu_0 \in \mathcal{M}_{t_0}^{\sigma_0}$ the corresponding extremal measure, we find from the equilibrium conditions (A.4) for μ_0 and the triple (t_0, ν_0, σ_0) that the measure $\mu_1 := \mu_0 + \nu - \nu_0 = \mu_0 + \sigma - \sigma_0 \in \mathcal{M}_t^\sigma$ with $\operatorname{supp}(\sigma - \mu_1) = \operatorname{supp}(\sigma_0 - \mu_0)$ satisfies also (A.4) for the triple (t, ν, σ) with the same constant F. Hence $\mu = \mu_1$ by part (d). Writing $J := \mathbb{R} \setminus \operatorname{supp}(\nu_0)$ it follows that $(\mu - \nu)|_J = (\mu_0 - \nu_0)|_J = \mu_0|_J \geq 0$. The last claim of part (e) is a consequence of part (c) since $\operatorname{supp}(\mu_0) = \operatorname{supp}(\sigma_0)$. ∎

Since $I(\cdot)$ corresponds to the physical electrostatic energy in two dimensions, it is possible to give an electrostatic interpretation of the construction of our extremal measure μ: we have a condenser with an isolating plate with negative charge ν, and a plate with positive charge μ, the second plate being conducting but subject to the maximum charge-per-unit restriction $\mu \leq \sigma$. This situation of continuous charges is the limit of an equilibrium between $n - 1$ negative unit charges (the poles) and n positive unit charges (the Ritz values) on the real line, with the constraint that two Ritz values are separated by an eigenvalue.

Remark A.2. *In our examples studied in §4 we have chosen $N - 1$ poles $\xi_{1,N}, \ldots, \xi_{N-1,N}$, and displayed the nth rational Ritz values for poles $\xi_{1,N}, \ldots, \xi_{n-1,N}$ for $n = 1, \ldots, N$. In other words, here we have a family of pole measures*

$$\chi_N(\{\xi_{1,N}, \ldots, \xi_{n-1,N}\}) \to \nu_t \qquad \text{for } n, N \to \infty \text{ such that } n/N \to t,$$

and ν_t of total mass $\|\nu_t\| = t$ is increasing in $t \in (0, \|\sigma\|)$. To each t we may associate an extremal measure $\mu_t \in \mathcal{M}_t^\sigma$. We claim that μ_t is also increasing, and, more *precisely, for* $0 < t_1 < t_2 < \|\sigma\|$ *there holds*

$$\operatorname{supp}(\sigma - \mu_{t_2}) \subset \Sigma_{t_2}^* := \{x \in \mathbb{C} : U^{\mu_{t_2} - \nu_{t_2}}(x) = F_{t_2}\} \subset \operatorname{supp}(\sigma - \mu_{t_1}).$$

As a consequence, for almost all t *we have* $\operatorname{supp}(\sigma - \mu_t) = \Sigma_t^*$*, and this is true for all* t *if* $t \mapsto \operatorname{supp}(\sigma - \mu_t)$ *is continuous.*

We should mention that, for a fixed external field and no upper constraint, the above claim is part of what is now known as the Buyarov-Rakhmanov formula [10], which has been extended to the case of an upper constraint and no external field in [6] (see also [7] for an additional fixed external field). The above claims have been established by Coussement & Van Assche [11, Appendix A] in our setting for the particular family $\nu_t = t\delta_0$*, but the arguments used by these authors remain fully valid in our more general context, we omit details.*

There is a subclass of constraints where μ can be computed explicitly.

Lemma A.3. *Let* $E = \operatorname{supp}(\sigma) = [\alpha, \beta]$*, and suppose that* $\operatorname{supp}(\nu) \subset (\beta, +\infty)$*, and* $\sigma = \operatorname{Bal}(\check{\sigma}, E)$*, where* $\operatorname{supp}(\check{\sigma}) \subset (-\infty, \alpha)$*.*

If the integral equation

$$\int \sqrt{\frac{\alpha - y}{b - y}}\, \mathrm{d}\check{\sigma}(y) = \int \sqrt{\frac{y - \alpha}{y - b}}\, \mathrm{d}\nu(y) \tag{A.8}$$

does not have a solution $b \in [\alpha, \beta]$*, then* $\operatorname{supp}(\sigma - \mu) = [\alpha, \beta]$*, and* $\mu = \operatorname{Bal}(\nu, [\alpha, \beta])$*, the balayage of* ν *onto* $[\alpha, \beta]$*. Furthermore,* $U^{\mu - \nu}$ *is equal to the constant* F *on* E*.*

Otherwise, the integral equation has exactly one solution $b \in [\alpha, \beta]$*, with* $\operatorname{supp}(\sigma - \mu) = [\alpha, b]$*, and* $\mu = \sigma + \operatorname{Bal}(\nu - \check{\sigma}, [\alpha, b])$*. Furthermore,* $U^{\mu - \nu}$ *is equal to the constant* F *on* $[\alpha, b]$*, and strictly less than* F *on* $(b, \beta]$*.*

Proof. We have already seen in the proof of Lemma A.1 that μ is the minimizer of the constrained energy problem in Theorem 3.1 if and only if $\rho = \sigma - \mu$ is the minimizing measure supported on E with mass $s = \|\sigma\| - t$ for the unconstrained energy problem (A.5) with external field $Q(z) = U^{\nu - \sigma}(z)$. We define $\check{Q}(z) = U^{\nu - \check{\sigma}}(z)$. Since $\sigma = \operatorname{Bal}(\check{\sigma}, E)$ and Q, $\check{Q}$ are continuous on E there holds $Q(z) = \check{Q}(z) + c$ for all $z \in E$ and some finite constant c (cf. [38, Theorem II.4.7]). Hence $\sigma - \mu$ remains the minimizer of (A.5) if the external field is $\check{Q}$ instead of Q. Let us show that, for suitable b, the solution of our extremal problem is obtained by balayage of the measure $\check{\sigma} - \nu$ onto the interval $[\alpha, b]$. We write more explicitly $\rho = \operatorname{Bal}(\check{\sigma} - \nu, [\alpha, b])$ for our candidate, and obtain from [38, Formula II.4.47] for its density $\rho'(x)$ for $x \in [\alpha, b]$ the formula

$$\rho'(x) = \frac{f(x, b)}{\pi\sqrt{(x - \alpha)(b - x)}}, \qquad f(x, b) := \int_{[\alpha, b]^c} \frac{\sqrt{(y - \alpha)(y - b)}}{|y - x|}\, \mathrm{d}(\check{\sigma} - \nu)(y).$$

By assumption on $\check{\sigma}$ and ν, the function $x \mapsto f(x, b)$ is strictly decreasing on $[\alpha, \beta]$ for each $b \in (\alpha, \beta]$, and thus positive on $[\alpha, b]$ if and only if $f(b, b) \geq 0$. By a similar argument, $b \mapsto f(b, b)$ strictly decreases on $[\alpha, \beta]$, and $f(\alpha, \alpha) = \|\check{\sigma}\| - \|\nu\| = s > 0$. By comparing with (A.8), we see that this integral equation does not have a solution in $[\alpha, \beta]$ if and only if $f(\beta, \beta) > 0$ and $b = \beta$, and otherwise there exists a unique $b \in [\alpha, \beta]$ with $f(b, b) = 0$. Thus, in both cases, we find that ρ is a positive measure supported on $[\alpha, b]$ with total mass $\|\rho\| = s$. According to Lemma A.1(d), it remains to show that the equilibrium conditions (A.7) hold. From [38, Formula II.5.4 and II.5.7],

$$g(x) := U^\rho(x) + Q(x) - U^\rho(\alpha) - Q(\alpha) = -\int g(x, y)\, \mathrm{d}\check{\sigma}(y) + \int g(x, y)\, \mathrm{d}\nu(y),$$

where $x \mapsto g(x,y)$ denotes the Green function of the domain $\mathbb{C} \setminus [\alpha, b]$ with pole y. Using the explicit expression given, e.g., in [38, §II.4], we may find the derivative of g as

$$g'(x) = \begin{cases} 0, & x \in [\alpha, b), \\ -\frac{f(x,b)}{\sqrt{(x-\alpha)(x-b)}}, & x \in (b, \beta], \end{cases}$$

where we recall from above that $-f(x,b) > 0$ for $b < x \le \beta$. Hence g is indeed zero on $[\alpha, b]$ and nonnegative on $[\alpha, \beta]$, as required for (A.7). □

In our proofs we use Assumption **(H3)** in a slightly different form, namely the one used in [24].

LEMMA A.4. *Suppose that **(H1)**–**(H3)** hold. For all $\epsilon > 0$ there exist $\delta \in (0, 1/4)$ such that for all integers N and for all $\lambda_{k,N} \in \Lambda_N$*

$$0 \le \frac{1}{N} \sum_{\substack{\lambda_{j,N} \in \Lambda_N \\ 0 < |\lambda_{j,N} - \lambda_{k,N}| \le 4\delta}} \log \frac{1}{|\lambda_{j,N} - \lambda_{k,N}|} \le \epsilon. \tag{A.9}$$

In addition, for each interval J of length $\le 2\delta$ and $\lambda \in J$ we have $0 \le U^{\sigma|_J}(\lambda) \le \epsilon$. Finally, the function $z \mapsto U^\sigma(z)$ is continuous.

Proof. By assumption on δ, each term in the sum occurring in (A.9) is ≥ 0, and so is the sum. If (A.9) is false, then there exists $\epsilon_0 > 0$ such that for all $\delta(m) = 1/m$ there exists an integer $N(m)$ and $\lambda_{k(m),N(m)} \in \Lambda_{N(m)}$ satisfying

$$\frac{1}{N(m)} \sum_{\substack{\lambda_{j,N(m)} \in \Lambda_{N(m)} \\ 0 < |\lambda_{j,N(m)} - \lambda_{k(m),N(m)}| \le 4\delta(m)}} \log \frac{1}{|\lambda_{j,N(m)} - \lambda_{k(m),N(m)}|} \ge \epsilon_0 > 0.$$

Note that $\limsup_m N(m) = \infty$, since otherwise, for sufficiently large m, the above sum would be empty. Hence, by possibly extracting subsequences by taking into account **(H1)**, we may suppose that $N(m) \to \infty$, and that $\lambda_{k(m),N(m)} \to \lambda$ for $m \to \infty$, and we obtain a contradiction to Assumption **(H3)**.

For a proof of the second statement, we note that the first inequality $0 \le U^{\sigma|_J}(\lambda)$ immediately follows from the fact that $2\delta \le 1$. By the maximum principle for logarithmic potentials [38, Corollary II.3.3], we may suppose that $\lambda \in \mathrm{supp}(\sigma|_J)$. Then **(H2)** implies that there exist $\Lambda_N \ni \lambda_{k(N),N} \to \lambda$ for $N \to \infty$. By making J slightly larger if necessary we may suppose that $\sigma(\partial J) = 0$. It follows from **(H2)** that $\sigma_N = \chi_N(J \cap \Lambda_N \setminus \{\lambda_{k(N),N}\}) \to \sigma|_J$. Since eigenvalues $\lambda_{j(N),N} \in J$ satisfy $|\lambda_{j(N),N} - \lambda_{k(N),N}| \le 2\delta + |\lambda - \lambda_{k(N),N}|$, we get from (A.9) for sufficiently large N that $U^{\sigma_N}(\lambda_{k(N),N}) \le \epsilon$, and (A.1) implies that $U^{\sigma|_J}(\lambda) \le \epsilon$.

Finally, by the Continuity Theorem [38, Theorem II.3.5], it is sufficient to show the continuity of U^σ on $\mathrm{supp}(\sigma)$. Let $x \in \mathrm{supp}(\sigma)$, and $\epsilon, \delta > 0$ as above. Denote $J = [x - \delta, x + \delta]$, then $U^{\sigma|_{\mathbb{R}\setminus J}}$ is continuous in x, and for all y with $|y - x| \le \delta$ we obtain from above that $|U^{\sigma|_J}(x) - U^{\sigma|_J}(y)| \le 2\epsilon$, implying the continuity of U^σ in x. □

Another very important ingredient in our reasoning is that we may discretize measures supported on the real line in an appropriate manner, compare with [2, Lemma 2.1(d)] or [6, Lemma A.1] for a similar discretization procedure in the complex plane.

LEMMA A.5. *Let $Z_{1,N} \subset Z_{3,N} \subset \mathbb{R}$ be discrete sets with asymptotic behavior $\chi_N(Z_{1,N}) \to \rho_1$, $\chi_N(Z_{3,N}) \to \rho_3$, and suppose that ρ_3 has no mass points.*

Then for each measure ρ_2 satisfying $\rho_1 \lneqq \rho_2 \lneqq \rho_3$ and for integers $i(N)$ with $i(N)/N \to \|\rho_2\|$ we find for sufficiently large N discrete sets $Z_{2,N}$ containing $i(N)$ elements such that

$$\chi_N(Z_{2,N}) \to \rho_2, \quad \forall N : \quad Z_{1,N} \subset Z_{2,N} \subset Z_{3,N}.$$

Proof. Denote by $[a, b)$ an interval containing $\mathrm{supp}(\rho_2)$, and consider the intervals $I_{j,k} = [a + (j-1)/k, a + j/k)$ for $j = 1, \dots, j(k)$, where $j(k) - 1$ is the integer part of $k(b-a)$ such that $\mathrm{supp}(\rho_2)$ is contained in the partition $I_{1,k} \cup \dots \cup I_{j(k),k}$. Since ρ_3 and thus ρ_1 and ρ_2 have no mass points, we find that

$$\lim_{N\to\infty} \frac{\#(Z_{1,N} \cap I_{j,k})}{N} = \rho_1(I_{j,k}) \le \rho_2(I_{j,k}) \le \rho_3(I_{j,k}) = \lim_{N\to\infty} \frac{\#(Z_{3,N} \cap I_{j,k})}{N}$$

for all $j = 1, \dots, j(k)$. Thus there exists $N(k) \ge 0$ such that for all $N \ge N(k)$ we can construct $Z_{2,k,N}$ contained in $Z_{3,N}$ and containing $Z_{1,N}$ with

$$\Big| \frac{\#(Z_{2,k,N} \cap I_{j,k})}{N} - \rho_2(I_{j,k}) \Big| \le \frac{1}{k^2} \quad \text{for all } j = 1, \dots, j(k).$$

Clearly, we may choose these integers $N(k)$ to be strictly increasing in k. Define now $Z_{2,N} = Z_{2,k,N}$ for $N(k) \le N < N(k+1)$. Then for $N(k) \le N < N(k+1)$ by construction of our partition,

$$\Big| \frac{\#Z_{2,N}}{N} - \|\rho_2\| \Big| \le \sum_{j=1}^{j(k)} \Big| \frac{\#(Z_{2,N} \cap I_{j,k})}{N} - \rho_2(I_{j,k}) \Big| \le \frac{j(k)}{k^2}, \tag{A.10}$$

the right-hand side tending to zero for $k \to \infty$, or, what amounts to the same, for $N \to \infty$. Similarly, any $f \in \mathcal{C}([a, b])$ is bounded by some $M > 0$, and, with ω denoting the modulus of continuity of f,

$$\begin{aligned} \Big| \int f d\chi_N(Z_{2,N}) - \int f \,\mathrm{d}\rho_2 \Big| &\le \sum_{j=1}^{j(k)} \Big| \int f \,\mathrm{d}\chi_N(Z_{2,N} \cap I_{j,k}) - \int_{I_{j,k}} f \,\mathrm{d}\rho_2 \Big| \\ &\le M \frac{j(k)}{k^2} + \omega(1/k) \Big(2\, \|\rho_2\| + \frac{j(k)}{k^2} \Big), \end{aligned}$$

the right-hand side tending to zero for $k \to \infty$, implying that $\chi_N(Z_{2,N}) \to \rho_2$ for $N \to \infty$. Also, estimate (A.10) implies that $\#Z_{2,N} = N\|\rho_2\| + o(N)_{N\to\infty} = i(N) + o(N)_{N\to\infty}$. By slightly changing our construction we can drop the $o(N)$ term: if $\#Z_{2,N} < i(N)$ then there exists a j such that $\rho_3(I_{j,k}) > \rho_2(I_{j,k})$, and we can add $i(N) - \#Z_{2,N}$ new elements from $Z_{3,N} \cap I_{j,k}$ to $Z_{2,N}$. Similarly, if $\#Z_{2,N} > i(N)$ then there exists a j such that $\rho_2(I_{j,k}) > \rho_1(I_{j,k})$, and we can drop $\#Z_{2,N} - i(N)$ elements from $Z_{2,N} \cap I_{j,k}$ which are not in $Z_{1,N}$. In both cases, the modification concerns only $o(N)$ points, and hence the conclusion $\chi_N(Z_{2,N}) \to \rho_2$ remains valid. ∎