



HAL
open science

On the sharpness of the weighted Bernstein-Walsh inequality, with applications to the superlinear convergence of conjugate gradients

Bernhard Beckermann, Thomas Helart

► To cite this version:

Bernhard Beckermann, Thomas Helart. On the sharpness of the weighted Bernstein-Walsh inequality, with applications to the superlinear convergence of conjugate gradients. 2017. <hal-01567286>

HAL Id: hal-01567286

<https://hal.science/hal-01567286v1>

Preprint submitted on 22 Jul 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

On the sharpness of the weighted Bernstein-Walsh inequality, with applications to the superlinear convergence of conjugate gradients

by

Bernhard Beckermann¹ and Thomas Helart²

Abstract

In this paper we show that the weighted Bernstein-Walsh inequality in logarithmic potential theory is sharp up to some new universal constant, provided that the external field is given by a logarithmic potential. Our main tool for such results is a new technique of discretization of logarithmic potentials, where we take the same starting point as in earlier work of Totik and of Levin & Lubinsky, but add an important new ingredient, namely some new mean value property for the cumulative distribution function of the underlying measure.

As an application, we revisit the work of Beckermann & Kuijlaars on the superlinear convergence of conjugate gradients. These authors have determined the asymptotic convergence factor for sequences of systems of linear equations with an asymptotic eigenvalue distribution. There was some numerical evidence to let conjecture that the integral mean of Green functions occurring in their work should also allow to give inequalities for the rate of convergence if one makes a suitable link between measures and the eigenvalues of a single matrix of coefficients. We prove this conjecture, at least for a class of measures which is of particular interest for applications.

Keywords: logarithmic potential theory, Bernstein-Walsh inequality, discretization of potential, conjugate gradients, superlinear convergence.

AMS subject classification: 15A18, 31A05, 31A15, 65F10

1 Introduction

Conjugate gradients (CG) is a popular method for solving large systems of equations, with the matrix of coefficients being symmetric and positive definite [1]. However, its convergence (at least in exact arithmetic) is not yet fully understood, despite an important number of research contributions, for instance [3, 4, 5, 2, 7, 8, 9, 10, 18]. It happens quite often that there is a regime of convergence called superlinear convergence, which depends very much on the

¹Laboratoire Painlevé UMR 8524, UFR Mathématiques, Univ. Lille, F-59655 Villeneuve d'Ascq CEDEX, France. E-mail: bbecker@math.univ-lille1.fr. Supported in part by the Labex CEMPI (ANR-11-LABX-0007-01).

²Laboratoire Painlevé UMR 8524, UFR Mathématiques, Univ. Lille, F-59655 Villeneuve d'Ascq CEDEX, France. E-mail: thomas.helart@gmail.com.

eigenvalue distribution of the matrix, see §1.2 for more details. People have been aware of this phenomenon for more than 40 years, but only in [3] a general theory based on logarithmic potential theory was suggested to quantify the rate of convergence, see also [10] for a more comprehensive summary. The drawback of this theory is that all results in [3] study only the so-called asymptotic convergence factor. In addition, this theory requires to consider sequences of systems of equations with a joint eigenvalue distribution, and thus gives not so much information about the actual rate of convergence for a single matrix. Numerical evidence in [3, 4, 5] seemed to indicate that behind the asymptotic results there should be some hidden inequality valid for a single matrix, see Conjecture 1.4 below. To our knowledge, the present paper is the first which deals with this conjecture, at least for a suitable subclass of eigenvalue distributions.

This paper contains three main ingredients, all being connected with polynomial extremal problems and thus with logarithmic potential theory: we discuss in §1.1 the sharpness of the so-called weighted Bernstein-Walsh inequality for the particular case where the external field is the logarithmic potential of some measure. Here our main result in Theorem 1.3 indicates the existence of some new universal constant. Secondly, we give and discuss in §1.2 some new upper bound for the rate of convergence of conjugate gradients, and show in our Theorem 1.7 the above conjecture for a particular class of eigenvalue distributions, which is illustrated by some (academic) numerical examples.

Our main technical result stated and proved in §2 is Theorem 2.1 on a new fine discretization of logarithmic potentials for a suitable class of measures, where in contrast to preceding work of Totik, Lubinsky and others we get (large but) explicit constants. Here an essential tool is a new mean value property stated in Theorem 2.6.

1.1 The weighted Bernstein-Walsh inequality

One of the appealing aspects of CG convergence is that there is a close link with polynomial extremal problems and extremal problems in logarithmic potential theory, which we discuss now.

Given a finite union of compact intervals $\Sigma \subset \mathbb{R}$, we denote by $\mathcal{M}_1(\Sigma)$ the set of Borel measures μ with support $\text{supp}(\mu)$ in Σ and of total mass 1, and consider the logarithmic potential and energy

$$U^\mu(x) = \int \log \frac{1}{|x-t|} d\mu(t), \quad I(\mu) = \int \int \log \frac{1}{|x-t|} d\mu(t) d\mu(x).$$

Given a weight w defined on \mathbb{R} and continuous on Σ together with an external field $Q(x) = -\log(w(x))$, it is known [16, Theorem I.1.3 and Theorem I.4.8] that there is a unique minimizer $\mu \in \mathcal{M}_1(\Sigma)$ of the extremal problem

$$\inf \left\{ I(\mu) + 2 \int Q d\mu : \mu \in \mathcal{M}_1(\Sigma) \right\} \quad (1)$$

which is uniquely characterized by the existence of a constant $F \in \mathbb{R}$ such that

$$f(x) := F - U^\mu(x) - Q(x) \begin{cases} = 0 & \text{for } x \in \text{supp}(\mu), \\ \leq 0 & \text{for } x \in \Sigma. \end{cases} \quad (2)$$

Logarithmic potential theory with external fields has been applied with success for getting asymptotics for various polynomial extremal problems [16], maybe one of the most prominent results being the weighted Bernstein-Walsh inequality [16, Theorem III.2.1]

$$\forall x_0 \in \mathbb{R} \quad \forall P \in \Pi_k : \quad \frac{|w(x_0)^k P(x_0)|}{\|w^k P\|_{\text{supp}(\mu)}} \leq e^{kf(x_0)}, \quad (3)$$

and its sharpness, see, e.g., [16, Corollary III.1.10],

$$\exists P_k \in \Pi_k \quad \forall x_0 \in \mathbb{R} \setminus \text{supp}(\mu) : \quad \lim_{k \rightarrow \infty} \left(\frac{|w(x_0)^k P_k(x_0)|}{\|w^k P_k\|_{\text{supp}(\mu)}} \right)^{1/k} = e^{f(x_0)}, \quad (4)$$

where Π_k denotes the set of polynomials of degree at most k , and $\|f\|_\Sigma = \max_{x \in \Sigma} |f(x)|$. One aim of this paper is to improve (4) for a particular class of external fields, see Theorem 1.3 below, namely to show that (3) is sharp up to some constant. Before giving some more details, let us first have a look at other classes of external fields where such constants are explicitly known. In what follows we will write $g_S(\cdot, \zeta)$ to denote the Green function in $\overline{\mathbb{C}} \setminus S$ for a compact set $S \subset \mathbb{R}$ with pole at $\zeta \in \overline{\mathbb{C}} \setminus S$. We will be mainly interested in the special case of an interval S where the Green function vanishes on S and is strictly positive outside S , and where explicit formulas are available.

Example 1.1 Consider $\Sigma = [a, b]$ and $Q = 0$, then an explicit formula is known for the minimizer in (1)

$$\text{supp}(\mu) = \Sigma = [a, b], \quad \frac{d\mu}{dx}(x) = \frac{1}{\pi \sqrt{(x-a)(b-x)}},$$

also called Robin equilibrium measure of the interval $[a, b]$ and denoted by $\omega_{[a,b]}$. It is also known from, e.g., [16, Eqn. (I.4.8)] that $f(z) = g_{[a,b]}(z, \infty)$, and thus (3) becomes the classical Bernstein-Walsh inequality. Taking $P_k(x) = T_k\left(\frac{2x-a-b}{b-a}\right)$ with T_k the Chebyshev polynomial of the first kind, one may also show that (3) is sharp up to a factor $1/2$.

Example 1.2 Consider $\Sigma = [a, b]$, and $w(x)^k = 1/\sqrt{q(x)}$ with q being a polynomial of degree $\ell \leq 2k$, strictly positive on $[a, b]$, compare with [14, chap 4.4]. Thus $Q = -U^\rho$ with ρ an atomic measure of mass $\ell/(2k) \leq 1$. Here the extremal measure in (1), (2) is given in [16, Example II.4.8] in terms of balayage onto $\text{supp}(\mu) = \Sigma$, and it follows from [16, Eqn. (4.32)] that

$$f(x) = \left(1 - \frac{\ell}{2k}\right)g(x, \infty) + \int g(x, y)d\rho(y).$$

Moreover, with help of the factorization

$$\tilde{q}(y)\tilde{q}\left(\frac{1}{y}\right) = q(x), \quad \frac{2x - a - b}{b - a} = \frac{1}{2}\left(y + \frac{1}{y}\right) \in \mathbb{R},$$

$|y| \geq 1$, the polynomial \tilde{q} of degree ℓ having all its roots outside the unit circle, it is known that P_k defined by

$$w(x)^k P_k(x) = \frac{1}{2}(e^{kf(x)} + e^{-kf(x)}), \quad e^{2kf(x)} = \frac{y^{2k}\tilde{q}\left(\frac{1}{y}\right)}{\tilde{q}(y)},$$

is a polynomial of degree k , showing that again (3) is sharp up to a factor $1/2$.

We are interested in the case where the external field is a positive potential $U^{\rho/k}$ (not necessarily of an atomic measure), for instance if w^k is a (power of a) polynomial. This includes the particular case $w(x) = |x|^\theta$ on $\Sigma = [0, 1]$ for $\theta > 0$, starting point of an important research area about incomplete polynomials [16, §VI.1.1]. For external fields being a positive potential, we recall below how to solve the extremal problem, including the well-known pushing effect that the support of the equilibrium measure may be a proper subset of Σ . We then state our main result on the sharpness of the weighted Bernstein-Walsh inequality.

Theorem 1.3 *Let $k \geq 1$ be some integer, and $Q = U^{\rho/k}$ on $\Sigma = [\alpha, b]$, with the Borel measure ρ being compactly supported on $(-\infty, \alpha]$. Consider on Σ the strictly decreasing function*

$$\eta(a) := \int \sqrt{\frac{b-y}{a-y}} d\rho(y), \quad (5)$$

and set $a = \alpha$ if $k + \|\rho\| \geq \eta(\alpha)$, and else denote by a the unique solution of $k + \|\rho\| = \eta(a)$ in Σ . Then the extremal measure in (1), (2) is given by

$$\text{supp}(\mu) = [a, b], \quad kf(x) = (k + \|\rho\|)g_{[a,b]}(x, \infty) - \int g_{[a,b]}(x, y) d\rho(y). \quad (6)$$

Moreover, the weighted Bernstein-Walsh inequality (3) is sharp up to some constant, that is, there exists a universal real constant $C_{BW} > 0$ such that, for all $k \geq 2$, we may construct a polynomial P_k of degree k such that, for all $x_0 \in \mathbb{R} \setminus [a, b]$,

$$\frac{|w(x_0)^k P_k(x_0)|}{\|w^k P_k\|_{\text{supp}(\mu)}} \geq e^{-C_{BW}} e^{kf(x_0)}. \quad (7)$$

Our proof of Theorem 1.3 presented in §3.1 is based on a fine discretization of the logarithmic potential $U^{k\mu}$. We will show in this paper that $C_{BW} \leq 15383$, but this is by no means optimal. The most remarkable fact for us seems to be that such a constant does not depend on the data ρ, a, b nor on k . In particular, we do not need any further assumptions on smoothness of ρ , which is probably required by other techniques like a Riemann-Hilbert approach (which in any case would only allow to discuss asymptotics).

1.2 Superlinear convergence for conjugate gradients

Conjugate gradients is a popular method for solving large sparse linear systems $Ax = c$ with symmetric positive definite A , with spectrum $\Lambda(A) = \{\lambda_j\}$, $0 < \lambda_1 < \lambda_2 < \dots \leq b$. Here one easily obtains the error estimate³ for the n th iterate x_n^{CG}

$$\frac{\|x_n^{CG} - A^{-1}c\|_A}{\|x_0^{CG} - A^{-1}c\|_A} \leq \min_{q \in \Pi_n} \max_{x \in S} \left| \frac{q(x)}{q(0)} \right| =: E_n(S) \quad (8)$$

with the energy norm $\|y\|_A^2 = y^*Ay$, where S is any compact set containing the spectrum $\Lambda(A)$, for instance $S = \Lambda(A)$. Thus, in contrast to the polynomial extremal problem considered in §1.1, we have a trivial weight and take norms on discrete sets. One way of relating the two problems is to replace $\Lambda(A)$ by an interval $[a, b] \subset (0, \infty)$ containing all eigenvalues, leading to the classical upper bound

$$E_n(\Lambda(A)) \leq 2 \exp(-ng_{[a,b]}(0, \infty)) = 2 \left(\frac{\sqrt{b/a} - 1}{\sqrt{b/a} + 1} \right)^n, \quad (9)$$

compare with Example 1.1. It is however known for a long time that there are eigenvalue distributions which lead to convergence which is faster than the one described in (9), namely so-called superlinear convergence, see for instance Figure 1. A first attempt to quantify such a convergence behavior was suggested by Kuijlaars and Beckermann [3], see also the review [10], or the review [2] from the perspective of discrete orthogonal polynomials. The key ingredient of this theory is to dispose of a measure σ with continuous potential U^σ and compact support describing the eigenvalue distribution. In [3], this is quantified by supposing that there is a sequence of systems $A_N x_N = c_N$, with σ being the weak-star limit of normalized counting measures of the spectra of the symmetric and positive definite matrices A_N ,

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{\lambda \in \Lambda(A_N)} \delta_\lambda = \sigma. \quad (10)$$

Under some additional weak assumptions for small eigenvalues, the authors establish in [3, Theorem 2.1] for the n th iterate $x_{n,N}^{CG}$ of conjugate gradients applied to the system $A_N x_N = c_N$ the asymptotic upper bound

$$\begin{aligned} \limsup_{\substack{n, N \rightarrow \infty \\ n/N \rightarrow t}} \left(\frac{\|x_{n,N}^{CG} - A_N^{-1}c_N\|_{A_N}}{\|x_{0,N}^{CG} - A_N^{-1}c_N\|_{A_N}} \right)^{1/n} &\leq \limsup_{\substack{n, N \rightarrow \infty \\ n/N \rightarrow t}} (E_n(\Lambda(A_N)))^{1/n} \\ &\leq \exp \left(-\frac{1}{t} \int_0^t g_{S(\tau)}(0, \infty) d\tau \right), \end{aligned} \quad (11)$$

where $(S(t))_{0 < t < \|\sigma\|}$ is a decreasing family of compact subsets of the convex hull of the spectra, obtained from some constrained extremal problem in logarithmic potential theory, which we explain now.

³In general, (8) might be an important overestimation of the error, but there exist right-hand sides c with equality.

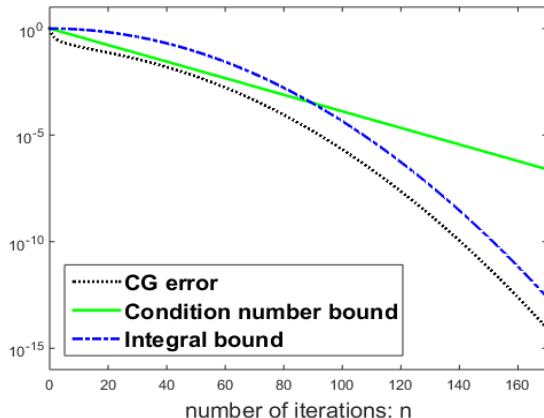


Figure 1: Lower and upper bounds for $n \mapsto E_n(\Lambda(A))$. Here $\lambda_j = (j/N)(2 - j/N)$ for $j = 1, \dots, N = 1000$. As lower bound we draw the relative CG error in energy norm, with $A = \text{diag}(\lambda_1, \dots, \lambda_N)$, $c = (1, \dots, 1)^T$, and starting vector $x_0^{CG} = 0$ (black dotted line). The upper bounds come from (9) with $b/a = \lambda_N/\lambda_1$ for the condition number bound (green solid line) and from Conjecture 1.4 with $C = 0$ (blue dash-dotted line), see Example 1.8 for further details.

For measures σ with compact support and continuous potential, and $0 < t < \|\sigma\|$, according to [6, 15] there exists a unique minimizer $\nu_{t,\sigma}$ of $I(\nu)$ under all candidates $\nu \in \mathcal{M}_1(\text{supp}(\sigma))$ with $\nu \leq \sigma/t$. This minimizer is uniquely characterized by the existence of a constant $C_{t,\sigma} \in \mathbb{R}$ such that

$$U^{\nu_{t,\sigma}}(x) = C_{t,\sigma} \text{ for } x \in \text{supp}(\sigma/t - \nu_{t,\sigma}), \quad U^{\nu_{t,\sigma}}(x) \leq C_{t,\sigma} \text{ for } x \in \text{supp}(\sigma).$$

Many Buyarov-Rakhmanov type properties are known about the measures $\nu_{t,\sigma}$ for fixed σ and varying t , we just recall here from [3, Proof of Theorem 2.1] the fact that the measures $t\nu_{t,\sigma}$ are increasing in t , and hence

$$S(t) := \text{supp}(\sigma/t - \nu_{t,\sigma}) \quad \text{decreases in } t.$$

As a consequence, the map $n \mapsto -N \int_0^{n/N} g_{S(\tau)}(0, \infty) d\tau$ is concave and describes superlinear convergence behavior. The compact sets $S(t)$ may have a quite complicated shape, and the main finding of [9] roughly says that the n th Ritz values of A_N approach well all eigenvalues in $\Lambda(A_N) \setminus S(n/N)$. There is a similar (rough) interpretation of (11): so-called "converged" eigenvalues which are already well approached by n th Ritz values should no longer contribute (in exact arithmetic) to the convergence of CG at later stages.

In many examples, numerical evidence did let to conjecture that the above upper bound (11) even holds (up to some modest constant) for a single matrix A , without limits and without taking the n -th root, see for instance [3, Eqn.(1.9)]

and Figures 1 and 4], [5, Eqn. (1.3)], or Figure 1. Of course, for a single matrix we cannot define σ through (10). This gives the following conjecture.

Conjecture 1.4 *There is a (modest) constant $C \in \mathbb{R}$ and a technique of associating a measure σ with compact support and continuous potential to the spectrum of a positive definite matrix A such that, for all n sufficiently small,*

$$E_n(\Lambda(A)) \leq \exp \left(C - N \int_0^{n/N} g_{S(t)}(0, \infty) dt \right).$$

It may be that this conjecture is wrong for measures where $S(t)$ has a complicated shape. In our proof of the conjecture, following [3, Lemma 3.1(a)], we will impose sufficient conditions on σ such that $S(t) = [a(t), b]$ for all t .

Lemma 1.5 *Suppose that σ is supported on the interval $[a(0), b]$ with density with respect to Lebesgue measure denoted by σ' , and suppose⁴ that $x \mapsto \sqrt{(x - a(0))(b - x)}\sigma'(x)$ vanishes at $x = a$, and is strictly increasing in $[a(0), b]$. Then for all $t \in (0, \|\sigma\|)$ we have $S(t) = [a(t), b]$, with $a(t)$ being the unique solution of the equation*

$$t = \int_{a(0)}^{a(t)} \sqrt{\frac{b-x}{a(t)-x}} d\sigma(x),$$

in particular $t \mapsto a(t)$ is strictly increasing.

Roughly speaking, having $S(t) = [a(t), b]$ for sufficiently small t means that there are so few eigenvalues around 0 that they are the first eigenvalues which are well approached by Ritz values of low order. One of the reasons to consider such sets $S(t)$ is that, in any case, the superlinear convergence rate is only pronounced if small eigenvalues are well approached by Ritz values, and the rate depends not as much on other "converging" eigenvalues, which in first order could be neglected. Another reason is that, if the system $Ax = c$ comes from discretizing an elliptic PDE, we might have only asymptotic knowledge on small eigenvalues of A through a so-called Weyl formula. The final reason is that in the particular case $S(t) = [a(t), b]$ the analysis becomes simpler, and also the upper bound is more explicit, since, by (9),

$$\begin{aligned} \exp \left(-N \int_0^{n/N} g_{S(t)}(0, \infty) dt \right) &= \exp \left(N \int_0^{n/N} \log \left(\frac{\sqrt{b/a(t)} - 1}{\sqrt{b/a(t)} + 1} \right) dt \right) \\ &\leq \prod_{j=0}^{n-1} \frac{\sqrt{b/a(j/N)} - 1}{\sqrt{b/a(j/N)} + 1} \end{aligned} \quad (12)$$

in terms of some "effective condition number" $b/a(j/N)$, compare with [5, Eqn. (2.27)].

⁴It follows that σ has compact support and continuous potential.

In order to proceed, we first extend our definition (8) of $E_n(S)$ to compact sets S which do not necessarily contain the spectrum of A : following [4], for a fixed matrix A , a compact set S , and sufficiently large n , let

$$E_n(S) := \min_{q \in \Pi_n} \left\{ \frac{\|q\|_S}{|q(0)|} : \forall \lambda \in \Lambda(A) \setminus S, q(\lambda) = 0 \right\},$$

and then obviously $E_n(\Lambda(A)) \leq E_n(S)$. This inequality has been used for example in [8] or [18] in order to derive a CG convergence bound taking into account few outliers represented by the set $\Lambda(A) \setminus S$, where typically S is the convex hull of the remaining eigenvalues.

In what follows we consider $S = [\lambda_{d+1}, b]$, and thus we prescribe as roots of q the smallest d eigenvalues $\lambda_1, \dots, \lambda_d$. Understanding the modulus of the product of the corresponding linear factors as a weight, and setting $\rho = \delta_{\lambda_1} + \dots + \delta_{\lambda_d}$, $\alpha = \lambda_{d+1}$ and $n = k + d = k + \|\rho\|$, Theorem 1.3 gives the following upper bounds in terms of Green functions. The sharpness follows from the weighted Bernstein-Walsh inequality (3).

Corollary 1.6 *For any integer $n > d + 1 \geq 1$, let $a = a_{d,n}$ be equal to λ_{d+1} if $n \geq \sum_{j=1}^d \sqrt{\frac{b-\lambda_j}{\lambda_{d+1}-\lambda_j}} = \eta(\lambda_{d+1})$, and else let a be the unique solution $> \lambda_{d+1}$ of the equation $n = \sum_{j=1}^d \sqrt{\frac{b-\lambda_j}{a-\lambda_j}}$. Then*

$$E_n([\lambda_{d+1}, b]) \leq \exp\left(C_{BW} - ng_{[a,b]}(0, \infty) + \sum_{j=1}^d g_{[a,b]}(0, \lambda_j)\right), \quad (13)$$

being sharp up to the factor $\exp(C_{BW})$.

Corollary 1.6 gives us for each $d < n - 1$ an upper bound for the function $n \mapsto \log E_n(\Lambda(A))$, each of them having the shape of a straight line for sufficiently large n , with the slope $-g_{[\lambda_{d+1}, b]}(0, \infty)$ of these straight lines decreasing with d , but the abscissa in general increases. We thus hope that $\log E_n(\Lambda(A))$ is close to the value of the concave lower envelope of these straight lines, which is true for the particular example of Figure 2. In fact, finding an optimal $d = d_n < n$ with minimal $E_n([\lambda_{d+1}, b])$ for given n seems to be a difficult task, we will suggest an approximate solution in order to solve the above conjecture.

Theorem 1.7 *Let σ and $S(t) = [a(t), b]$ for $0 < t < \|\sigma\|$ be as in Lemma 1.5, and A be a symmetric positive definite matrix with spectrum $\lambda_1 < \lambda_2 < \dots \leq b$.*

(a) *If the integers $n \geq 2$ and $d = d_n \in \{0, 1, \dots, n - 2\}$ are such that*

$$\text{for } j = 1, 2, \dots, d: \quad \sigma((-\infty, \lambda_j]) \geq j/N, \quad (14)$$

$$\lambda_d < a(n/N) \leq \lambda_{d+1} \quad (\text{or } a(n/N) \leq \lambda_1 \text{ in case } d = 0), \quad (15)$$

then

$$E_n(\Lambda(A)) \leq E_n([\lambda_{d+1}, b]) \leq \exp\left(C_{BW} - N \int_0^{n/N} g_{S(t)}(0, \infty) dt\right),$$

and thus Conjecture 1.4 holds.

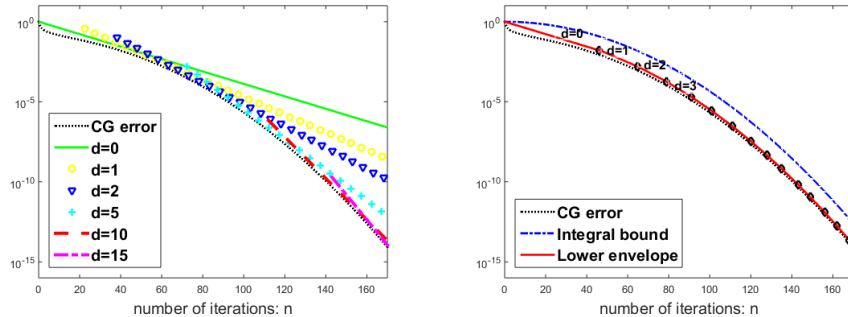


Figure 2: Illustration of Corollary 1.6. As lower bound we draw on both plots the relative CG error in energy norm, with $\lambda_j, N, A, c, x_0^{CG}$ as in Figure 1 (black dotted line). The straight lines on the left correspond to the bounds for $d \in \{0, 1, 2, 5, 10, 15\}$ given in Corollary 1.6, each time for n sufficiently large such that $a_{n,d} = \lambda_{d+1}$. Since it is difficult to see details, we have drawn on the right only the CG error and the polygon obtained from the lower envelope of all straight lines in Corollary 1.6, where we indicate in the plot the correspondence between a segment and the choice of d . To compare, we also have reproduced on the right from Figure 1 the integral bound from Conjecture 1.4 with $C = 0$ (blue dash-dotted line), verifying numerically that Corollary 1.6 is the right tool to prove the conjecture.

- (b) *The above choice (15) of d is nearly optimal in the following sense: consider diagonal A_N with eigenvalues satisfying $\sigma((-\infty, \lambda_{j,N}]) = j/N$ for $j = 1, \dots, N$. Furthermore, let $d = d_{n,N}$ with $\lambda_{d,N} < a(n/N) \leq \lambda_{d+1,N}$, then⁵*

$$\lim_{\substack{n, N \rightarrow \infty \\ n/N \rightarrow t}} E_{n,N}([\lambda_{d_{n,N}+1,N}, b])^{1/n} = \liminf_{\substack{n, N \rightarrow \infty \\ n/N \rightarrow t \\ 0 \leq d < n-1}} E_{n,N}([\lambda_{d+1,N}, b])^{1/n}.$$

It is also interesting to compare Theorem 1.7(a),(b) with [4, Theorem 3.1] which showed under the sole assumption (10) (and for quite general measures σ) that, for any fixed compact set S , the quantity $E_{n,N}(S)^{1/n}$ is asymptotically greater than or equal to the right-hand side of (11). One of the consequences of our Theorem 1.7 is that, roughly, we can achieve equality for the interval $S = S(n/N)$.

Our proof of Theorem 1.7 will be presented in §3.2, let study here some examples.

Example 1.8 *Consider the probability density*

$$\frac{d\sigma}{dx}(x) = \frac{1}{2\sqrt{1-x}} \quad \text{on} \quad [a(0), b] = [0, 1].$$

⁵We write $E_{n,N}$ instead of E_n in order to indicate that here we consider the spectrum of A_N depending on N .

For this measure we may apply Lemma 1.5, and a small computation shows for $0 < t < \|\sigma\| = 1$ that $a(t) = t^2$. We may also compute eigenvalues λ_j satisfying equality in (14):

$$\sigma([0, \lambda_j]) = \frac{j}{N} \quad \text{iff} \quad \lambda_j = \frac{j}{N} \left(2 - \frac{j}{N}\right),$$

which behave like equidistant points for $j \ll N$. These are the eigenvalues used in Figure 1 and Figure 2. In this special example we even have an explicit formula for the quantity $d = d_n$ of Theorem 1.7(a), namely

$$d_n + 1 = \lceil N(1 - \sqrt{1 - (n/N)^2}) \rceil \approx \lceil \frac{n^2}{2N} \rceil,$$

in particular $d_n = 0$ for $n \leq 45$, $d_n = 1$ for $46 \leq n \leq 64$, and $d_n = 2$ for $65 \leq n \leq 78$, in accordance with the right-hand plot of Figure 2.

In the previous example the small eigenvalues were approximately equidistant, with stepsize $2/N$, and the convex hull of the spectrum given approximately by $[2/N, 1]$. Up to correct scaling, a similar behavior is true for the eigenvalues of the finite difference discretization of the 2D Laplacian on the unit square with Dirichlet boundary conditions, and thus the convergence curves should be similar. However, this is no longer true for higher dimensions $D \geq 3$, where we expect that $\sigma'(x)$ grows like a constant times $x^{(D-2)/2}$ for small x , which motivates the following example.

Example 1.9 For a parameter $\gamma > 0$, consider the density

$$\frac{d\sigma}{dx}(x) = \frac{\gamma x^\beta}{\sqrt{b-x}} \quad \text{on} \quad [0, b].$$

In this example we only consider probability measures σ and thus $\gamma b^{\beta+1/2} B(\beta+1, 1/2) = 1$, with $B(\cdot, \cdot)$ the beta function. Notice that, for $\beta = 0$, we recover Example 1.8. A small computation using Lemma 1.5 gives

$$a(t)/b = t^{\frac{1}{\beta+1/2}}.$$

We again choose λ_j for $j = 1, 2, \dots, N = 1000$ attaining equality in (14), however, there are no longer explicit formulas, and thus the λ_j have to be computed numerically. In Figure 3 we have plotted two examples for $b = 1$, on the left for $\beta = 0.5$ and on the right for $\beta = 1$, where in both cases we have chosen the approximately optimal $d = d_n$ of Theorem 1.7(a), in accordance with the statement of Theorem 1.7. Notice also the well-known phenomena that the convergence of CG improves dramatically with β getting larger.

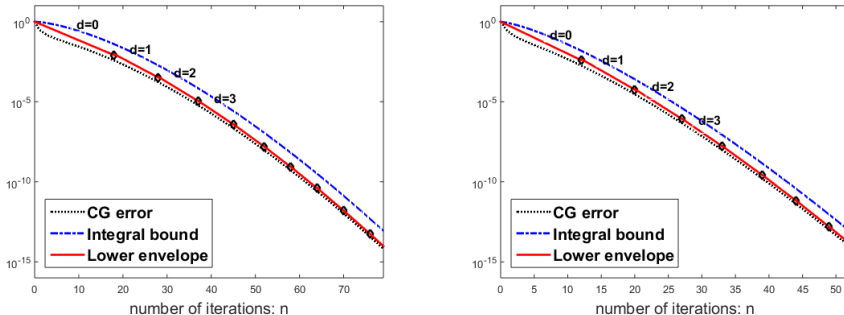


Figure 3: Illustration of Theorem 1.7(a), with λ_j for $j = 1, \dots, N = 1000$ as in Example 1.9, where on the left $\beta = 0.5$ and on the right $\beta = 1$. As lower bound we draw on both plots the relative CG error in energy norm, with A, c, x_0^{CG} as in Figure 1 (blue dotted line). The polygons are obtained from Theorem 1.7(a) (red solid line), where we indicate in the plot the correspondence between a segment and the choice of d . We also draw the integral bound (blue dash-dotted line), showing that Conjecture 1.4 holds with $C = 0$.

1.3 Structure of the paper

The reminding of the paper is organized as follows. Section 2 contains our results on discretizing the logarithmic potential of a class of measures including the extremal measure of Theorem 1.3. We first state our main Theorem 2.1, and then report in §2.1 about related results of Totik and of Lubinsky, and about the link with weighted quadrature formulas. Subsequently, we give in §2.2 the structure of the proof of Theorem 2.1, where following Totik we write the discretization error as a sum of three sums. We then state our original approach for dealing with these three sums, namely the mean value property of Theorem 2.6, and describe in §2.3 how to bound each of the three sums, with explicit constants.

In the third section we explain how to deduce Theorem 1.3 from Theorem 2.1 (§3.1), and Theorem 1.7 from Theorem 1.3 (§3.2). Subsequently, we give some concluding remarks. Our (quite technical) proof of Theorem 2.6 is postponed to Appendix A, and in Appendix B we gather some further technical results for dealing with our three sums.

2 Discretization of a potential

Our proof of Theorem 1.3 is based on the approximation of kU^μ with μ the equilibrium measure as in Theorem 1.3 by $-\log |P_k(z)|$ with P_k a suitable monic polynomial of degree k . We will show the following.

Theorem 2.1 Consider a measure $\mu \in \mathcal{M}_1([a, b])$ which has the density

$$k \frac{d\mu}{dx}(t) = g(t) \frac{k}{\pi \sqrt{(t-a)(b-t)}}$$

for a function g which is non negative, concave and increasing⁶ on (a, b) , such that $t \mapsto \frac{g(t)}{t-a}$ is convex on (a, b) . Then there exists a universal explicit constant C_{BW} such that for each $k \geq 2$ we may construct a monic polynomial P_k of degree k such that

- (a) $\forall z \in \mathbb{C}: \log |P_k(z)| + kU^\mu(z) \leq C_{BW}$,
- (b) $\forall x \in \mathbb{R} \setminus (a, b) : \log |P_k(x)| + kU^\mu(x) \geq 0$.

We will show in the proof of Theorem 1.3 that the extremal measure μ of Theorem 1.3 satisfies the assumptions of Theorem 2.1.

Example 2.2 Another class of functions g satisfying the assumptions of Theorem 2.1 for $[a, b] = [-1, 1]$ is given by

$$\begin{aligned} g(x) &= (x+1)^\theta \pi / \int_{-1}^1 (t+1)^{\theta-1/2} (1-t)^{-1/2} dt \\ &= \frac{\pi}{2^\theta} \frac{\Gamma(\theta+1)}{\Gamma(1/2)\Gamma(\theta+1/2)} (x+1)^\theta \end{aligned}$$

for $\theta \in [0, 1]$.

We will describe in §2.1 related work for discretizing potentials under various assumptions, but here the constants in general depend on μ , see for instance [16, §VI.4] for a summary. In §2.2 we give a proof of Theorem 2.1, where we initially follow the approach of Totik in [17, §2 and §9], see also the very accessible reference [13, Method 1] for the particular case $g(t) = 2t$ on $[a, b] = [0, 1]$ (up to a quadratic change of variables). Subsequently, we give in §2.3 a proof of three upper bounds we used in §2.2. Since the general case follows from a linear change of variables, we will suppose in what follows that $[a, b] = [-1, 1]$ in Theorem 2.1.

2.1 How to discretize a potential?

It is natural to approach the logarithmic potential $U^\mu(x) = \int \log(1/|x-t|)d\mu(t)$ by a quadrature rule of the form

$$\sum_{j=0}^{k-1} \log \frac{1}{|x - \xi_j|} = -\log |P_k(x)|, \quad P_k(x) = \prod_{j=0}^{k-1} (x - \xi_j), \quad (16)$$

⁶ In particular, g is continuous and bounded on (a, b) , thus we may extend g to become a continuous, non-negative, concave and increasing function in $[a, b]$.

for instance a weighted rectangular or midpoint rule, where we first cut $[-1, 1]$ into k subintervals $[t_j, t_{j+1}]$, $-1 = t_0 < t_1 < \dots < t_k = 1$, of equal mass $\mu([t_j, t_{j+1}]) = 1/k$, and chose $\xi_j \in [t_j, t_{j+1}]$ for $j = 0, \dots, k-1$. As long as $x \notin [-1, 1]$ and the density of μ does not vary too much, we may bound the error $kU^\mu(x) + \log |P_k(x)|$ above and below, and may even show convergence to 0 for $k \rightarrow \infty$ for suitable choices of ξ_j . In our case we have the additional difficulties that the density of μ may have singularities at ± 1 , showing that the interval lengths $t_{j+1} - t_j$ may strongly vary in size for $j = 0, 1, \dots, k-1$, and in addition in case $x \in [-1, 1]$ we have to deal with a logarithmic singularity of the integrand.

Totik in [13, Method 1] used the weighted midpoint rule

$$\xi_j = \int_{t_j}^{t_{j+1}} t d\mu(t) / \int_{t_j}^{t_{j+1}} d\mu(t) = k \int_{t_j}^{t_{j+1}} t d\mu(t) \quad (17)$$

for $j = 0, 1, \dots, k-1$. In the particular case $[a, b] = [0, 1]$ and $g(t) = 2t$, a proof of Theorem 2.1 can be found in [13, §2], which strongly relies on the explicit knowledge of asymptotics for the points ξ_j and t_j as a function of j and k for $k \rightarrow \infty$, and thus on the explicit knowledge of μ . In [16, Theorem VI.4.2] (see also the related result [17, Lemma 9.1] where the roots of P_k are slightly shifted into the complex plane), Totik considered probability measures μ with densities which are continuous up to a finite number of singularities of the form $|t - a_j|^{\delta_j}$ for $\delta_j > -1$. These assumptions are true in the setting of Theorem 2.1. He then shows the existence of (non explicit) constants $C_{T,1}, C_{T,2}$ depending on μ but not on k such that, for all $x \in \mathbb{R}$,

$$\begin{aligned} \log |P_k(x)| + kU^\mu(x) &\leq C_{T,1}, \\ \log |P_k(x)| + kU^\mu(x) &\geq C_{T,2} + \max\left\{0, \log(\text{dist}(x, \{\xi_0, \dots, \xi_{k-1}\}))\right\}. \end{aligned}$$

We see that the first inequality is as in Theorem 2.1(a), whereas the second one is clearly weaker than Theorem 2.1(b) for $x \in \mathbb{R} \setminus (-1, 1)$ close to $[-1, 1]$, since we get an additional term $\log(1/k^\beta)$ for some $\beta > 0$. Again, a proof of these statements uses heavily asymptotics for the points ξ_j and t_j as a function of j and k for $k \rightarrow \infty$, and thus quite a bit of information on μ .

Another technique of discretization has been considered by Lubinsky & Levin in [11] and [12], see also the very accessible reference [13, Method 2] for the particular case $g(t) = 2t$ on $[a, b] = [0, 1]$ (up to a quadratic change of variables). With t_0, \dots, t_k as before, consider intermediate abscissa $t_{j+1/2} \in (t_j, t_{j+1})$ such that all intervals $[t_{j/2}, t_{(j+1)/2}]$ have the same mass $1/(2k)$. Given $x_0 \in \mathbb{R}$, the authors then apply trapezian rule on most of the subintervals $[t_{j-1/2}, t_{j+1/2}]$ corrected with suitable rectangle rules on the remaining 2 or 3 subintervals such that $\{\xi_0, \dots, \xi_{k-1}\} \subset \{\pm 1, t_{1/2}, t_{3/2}, \dots, t_{k-1/2}\}$. Up to a (quadratic) change of variables, the authors of [11, Theorem 9.1] suppose that

$$\frac{d\mu}{dx}(t) = \frac{(t+1)h(t)}{\pi\sqrt{1-t^2}}$$

with h continuous and > 0 on $[-1, 1]$, and the modulus of continuity satisfies that $\log(1/\delta)\omega(h, \delta)$ is bounded above by some $\Gamma > 0$ for $\delta \in (0, 1)$. In this case, for all $x \in \mathbb{R}$,

$$\begin{aligned}\log |P_k(x)| + kU^\mu(x) &\leq C_{LL,1}, \\ \log |P_k(x_0)| + kU^\mu(x_0) &\geq C_{LL,2},\end{aligned}$$

where $C_{LL,1}, C_{LL,2}$ are (non explicit) constants depending only on Γ and the minimum and maximum of h on $[-1, 1]$. Note that the assumptions of [11, Theorem 9.1] and those of Theorem 2.1 are different and do not imply each other, see for instance Example 2.2 for $\theta < 1$. However, the above inequalities are quite close to those of Theorem 2.1, though our constants do not depend on μ , and our P_k does not depend on x_0 , and we only allow $x_0 \in \mathbb{R} \setminus (a, b)$.

Example 2.3 *In the particular case $[a, b] = [-1, 1]$ and $g = 1$ in Theorem 2.1, we have explicit formulas*

$$t_j = -\cos(\pi \frac{j}{k}), \quad \xi_j = -c_k \cos(\pi \frac{2j+1}{2k}), \quad c_k = \frac{2k}{\pi} \sin(\frac{\pi}{2k}).$$

Here the midpoint approach of Totik gives the monic polynomial

$$P_k(x) = 2(\frac{c_k}{2})^k T_k(x/c_k)$$

which is not optimal for the one-sided approximation of $kU^\mu(x)$ in Theorem 2.1 or the sharpness of the classical Bernstein-Walsh inequality as discussed in Example 1.1, but good enough for concluding in Theorem 2.1.

The previous example is misleading in the sense that in general there is no such sufficiently explicit formula for the t_j nor the ξ_j which will allow us to conclude in Theorem 2.1.

2.2 Structure of the proof of Theorem 2.1

We start by observing that, with the choices (16), (17),

$$\log |P_k(x)| + kU^\mu(x) = k \sum_{j=0}^{k-1} \int_{t_j}^{t_{j+1}} \log \left| \frac{x - \xi_j}{x - t} \right| d\mu(t).$$

The following classical lemma shows Theorem 2.1(b).

Lemma 2.4

$$k \int_{t_j}^{t_{j+1}} \log \left| \frac{x - \xi_j}{x - t} \right| d\mu(t) \begin{cases} \geq 0 & \text{for } x \in \mathbb{R} \setminus (t_j, t_{j+1}), \\ \leq \frac{1}{4} \frac{(t_{j+1} - t_j)^2}{(x - t_j)(x - t_{j+1})} & \text{for } x \in \mathbb{R} \setminus [t_j, t_{j+1}]. \end{cases}$$

Proof: Using the fact that $m(t) = \log \left| \frac{x - \xi_j}{x - t} \right|$ is convex on $[t_j, t_{j+1}]$ by assumption on x , we know that $m(t) \geq m(\xi_j) + m'(\xi_j)(t - \xi_j) = m'(\xi_j)(t - \xi_j)$, and thus

$$k \int_{t_j}^{t_{j+1}} \log \left| \frac{x - \xi_j}{x - t} \right| d\mu(t) \geq m'(\xi_j) \int_{t_j}^{t_{j+1}} (t - \xi_j) d\mu(t) = 0,$$

where in the last equality we have used (17). Also, using the convexity of m and the inequality $\log(x) \leq x - 1$ we obtain

$$\begin{aligned} m(t) &\leq m(t_j) \frac{t_{j+1} - t}{t_{j+1} - t_j} + m(t_{j+1}) \frac{t - t_j}{t_{j+1} - t_j} \\ &\leq \frac{t_j - \xi_j}{x - t_j} \frac{t_{j+1} - t}{t_{j+1} - t_j} + \frac{t_{j+1} - \xi_j}{x - t_{j+1}} \frac{t - t_j}{t_{j+1} - t_j}. \end{aligned}$$

Integrating and using again (17) we conclude that

$$\begin{aligned} k \int_{t_j}^{t_{j+1}} \log \left| \frac{x - \xi_j}{x - t} \right| d\mu(t) &\leq \frac{t_j - \xi_j}{x - t_j} \frac{t_{j+1} - \xi_j}{t_{j+1} - t_j} + \frac{t_{j+1} - \xi_j}{x - t_{j+1}} \frac{\xi_j - t_j}{t_{j+1} - t_j} \\ &= \frac{(t_{j+1} - \xi_j)(\xi_j - t_j)}{(x - t_j)(x - t_{j+1})} \leq \frac{1}{4} \frac{(t_{j+1} - t_j)^2}{(x - t_j)(x - t_{j+1})}. \end{aligned}$$

□

Remark 2.5 (a) The interested reader might have noticed that, by the same argument, the inequality of Theorem 2.1(b), namely $\log |P_k(x)| + kU^\mu(x) \geq 0$, also holds for $x \in \{t_0, t_1, \dots, t_k\}$.

(b) For $x > 1$ (and similarly for $x < -1$), the right-hand side of Theorem 2.1(b) cannot be improved since, by Lemma 2.4 and Lemma B.8(c),

$$\begin{aligned} \log |P_k(x)| + kU^\mu(x) &\leq \max_{\ell=0, \dots, k-1} \frac{t_{\ell+1} - t_\ell}{4} \sum_{j=0}^{k-1} \frac{t_{j+1} - t_j}{(x - t_j)(x - t_{j+1})} \\ &= \max_{\ell=0, \dots, k-1} \frac{t_{\ell+1} - t_\ell}{2(x^2 - 1)} \leq \frac{1}{(x^2 - 1)} \left(\frac{3\pi}{2k} \right)^{1/3}. \end{aligned}$$

(c) For $x \in \mathbb{C} \setminus \mathbb{R}$, it is not too difficult to show that $m(t) = \log \left| \frac{x - \xi_j}{x - t} \right|$ satisfies

$$|m(t) - m(\xi_j) - (t - \xi_j)m'(\xi_j)| \leq \frac{(t_{j+1} - t_j)^2}{2 \operatorname{dist}(x, [-1, 1])^2},$$

and hence by Lemma B.8(c)

$$|\log |P_k(x)| + kU^\mu(x)| \leq \sum_{j=0}^{k-1} \frac{(t_{j+1} - t_j)^2}{2 \operatorname{dist}(x, [-1, 1])^2} \leq \frac{1}{\operatorname{dist}(x, [-1, 1])^2} \left(\frac{12\pi}{k} \right)^{1/3}.$$

Thus, for sufficiently large k , the inequality of Theorem 2.1(b) also holds for non-real x up to some arbitrarily small constant.

Let us now turn to a proof of Theorem 2.1(a). We claim that it is sufficient to show Theorem 2.1(a) for $x \in [-1, 1] = \text{supp}(\mu)$, since then for μ -almost all x

$$kU^\mu(x) \leq C_{BW} - \log |P_k(z)| = C_{BW} + \sum_{j=0}^{k-1} U^{\delta_{\xi_j}}(x),$$

and thus this inequality holds for all $x \in \mathbb{C}$ by the principle of domination [16, Theorem II.3.2] and the finiteness of $I(\mu)$. Therefore, let $x \in [-1, 1]$ and, more precisely,

$$j_0 \in \{0, 1, \dots, k-1\} \quad \text{with} \quad x \in [t_{j_0}, t_{j_0+1}]. \quad (18)$$

According to Lemma 2.4, we get the following upper bound

$$\log |P_k(x)| + kU^\mu(x) \leq \Sigma_1 + \Sigma_2 + \Sigma_3 \quad (19)$$

with

$$\begin{aligned} \Sigma_1 &= \sum_{j=0}^{j_0-2} \int_{t_j}^{t_{j+1}} \log \left| \frac{x - \xi_j}{x - t} \right| k \, d\mu(t) \leq \frac{1}{4} \sum_{j=0}^{j_0-2} \frac{(t_{j+1} - t_j)^2}{(t_{j_0} - t_{j+1})^2}, \\ \Sigma_2 &= \sum_{j=\max\{0, j_0-1\}}^{\min\{j_0+1, k-1\}} \int_{t_j}^{t_{j+1}} \log \left| \frac{x - \xi_j}{x - t} \right| k \, d\mu(t), \\ \Sigma_3 &= \sum_{j=j_0+2}^{k-1} \int_{t_j}^{t_{j+1}} \log \left| \frac{x - \xi_j}{x - t} \right| k \, d\mu(t) \leq \frac{1}{4} \sum_{j=j_0+2}^{k-1} \frac{(t_{j+1} - t_j)^2}{(t_j - t_{j_0+1})^2}. \end{aligned}$$

Already in the particular Chebyshev case of Example 2.3 one may check that such a simple telescoping sum trick as in Remark 2.5 does not allow to conclude, since in general $|t_j - t_\ell|$ does not behave uniformly for $j, \ell \in \{0, 1, \dots, k-1\}$ like $|j - \ell|/k$, as it would be the case for equidistant points. We will discuss our upper bounds for the above three sums in the Propositions 2.7–2.9 of §2.3, which allows us to conclude the proof of Theorem 2.1, with the explicit constant

$$C_{BW} = 872 + 32 + 14479 = 15383.$$

So far we followed quite closely the reasoning in the literature, with more explicit constants. In all considerations to follow we will require precise lower and upper bounds for the ratio

$$\frac{j - \ell}{t_j - t_\ell}$$

which will follow from a new mean value property for the cumulative distribution function

$$W_g(x) = k \int_{-1}^x d\mu(t), \quad W'_g(t) = g(t)W'_1(t) = \frac{kg(t)}{\pi\sqrt{1-t^2}}, \quad (20)$$

since $W_g(t_j) = j$ for $j = 0, 1, \dots, k$.

Theorem 2.6 *Under the assumptions of Theorem 2.1 with $[a, b] = [-1, 1]$, we have for all distinct $x, t \in [-1, 1]$*

$$c_1 W'_g\left(\frac{t+x}{2}\right) \leq \frac{W_g(t) - W_g(x)}{t-x} \leq c_2 W'_g\left(\frac{t+x}{2}\right). \quad (21)$$

where $c_1 = \frac{1}{4}$ and $c_2 = \pi\sqrt{2}$.

Notice that, even for the particular case $g = 1$ and $W_1(-\cos(\alpha)) = k\alpha/\pi$, this statement is not totally obvious, but can be verified by means of elementary computations with improved constants c_1 and c_2 , see Lemma A.1 below. The proof for general g is strongly based on Jensen's inequality, we refer the reader to Appendix A for details.

2.3 Bounding three sums

For concluding our proof of Theorem 2.1, it remains to obtain upper bounds for the three terms on the right-hand side of (19), where we will proceed in order of difficulty, and apply beside Theorem 2.6 a certain number of technical results established in Appendix B, and recalled below. In the reminder of this section we will always suppose that the assumptions of Theorem 2.1 hold with $[a, b] = [-1, 1]$ and j_0 is chosen as in (18).

We start with the sum

$$\sum_3 \leq \frac{1}{4} \sum_{j=j_0+2}^{k-1} \frac{(t_{j+1} - t_j)^2}{(t_j - t_{j_0+1})^2},$$

where beside Theorem 2.6 we rely on an upper bound for the quantity

$$\left(1 + \frac{t_j + t_{j+1}}{2}\right) / (1 + t_j),$$

see Lemma B.4.

Proposition 2.7 *There holds*

$$\sum_3 \leq \frac{c_2^2 c_5 \pi^2}{12 c_1^2} \leq 872.$$

Proof: By Theorem 2.6

$$\sum_3 \leq \frac{c_2^2}{4c_1^2} \sum_{j=j_0+2}^{k-1} \frac{1}{(j - j_0 - 1)^2} \frac{W'_g\left(\frac{t_j + t_{j_0+1}}{2}\right)^2}{W'_g\left(\frac{t_j + t_{j+1}}{2}\right)^2}.$$

As g is increasing and $j > j_0 + 1$, we have that $g\left(\frac{t_j + t_{j_0+1}}{2}\right) \leq g\left(\frac{t_j + t_{j+1}}{2}\right)$, and

thus

$$\begin{aligned}
\frac{W'_g\left(\frac{t_j+t_{j_0+1}}{2}\right)^2}{W'_g\left(\frac{t_j+t_{j+1}}{2}\right)^2} &= \frac{g\left(\frac{t_j+t_{j_0+1}}{2}\right)^2}{g\left(\frac{t_j+t_{j+1}}{2}\right)^2} \frac{W'_1\left(\frac{t_j+t_{j_0+1}}{2}\right)^2}{W'_1\left(\frac{t_j+t_{j+1}}{2}\right)^2} \\
&\leq \frac{W'_1\left(\frac{t_j+t_{j_0+1}}{2}\right)^2}{W'_1\left(\frac{t_j+t_{j+1}}{2}\right)^2} \leq \frac{1 + \frac{t_j+t_{j+1}}{2}}{1 + \frac{t_j+t_{j_0+1}}{2}} \\
&\leq 2 \frac{1 + \frac{t_j+t_{j+1}}{2}}{1 + t_j} \leq 2c_5,
\end{aligned}$$

where in the last inequality we have applied Lemma B.4. Combining these two results yields the claimed upper bound. \square

Let us now turn to the sum

$$\sum_1 \leq \frac{1}{4} \sum_{j=0}^{j_0-2} \frac{(t_{j+1} - t_j)^2}{(t_{j_0} - t_{j+1})^2}.$$

Here we require beside Theorem 2.6 also upper bounds for the two ratios

$$\frac{1 + \frac{t_j+t_{j+1}}{2}}{1 + t_{j+1}}, \quad \text{and} \quad \frac{(j+1)^2}{j_0^2} \frac{1 + \frac{t_{j+1}+t_{j_0}}{2}}{1 + t_{j+1}}$$

for $j \leq j_0 - 1 \leq k - 2$, see Lemma B.5 and Lemma B.6.

Proposition 2.8 *There holds*

$$\sum_1 \leq \frac{c_2^2 \pi^2}{6c_1^2} (18 + \pi^2) \leq 14479. \tag{22}$$

Proof: Using the fact that $j < j_0 - 1$, and that $g(t) = (1+t)h(t)$ with a decreasing function h , we find that

$$\begin{aligned}
\frac{g\left(\frac{t_{j+1}+t_{j_0}}{2}\right)}{g\left(\frac{t_j+t_{j+1}}{2}\right)} &= \frac{1 + \frac{t_{j+1}+t_{j_0}}{2}}{1 + \frac{t_j+t_{j+1}}{2}} \frac{h\left(\frac{t_{j+1}+t_{j_0}}{2}\right)}{h\left(\frac{t_j+t_{j+1}}{2}\right)} \\
&\leq \frac{1 + \frac{t_{j+1}+t_{j_0}}{2}}{1 + \frac{t_j+t_{j+1}}{2}}.
\end{aligned}$$

This allows us to write

$$\begin{aligned}
\frac{W'_g\left(\frac{t_{j+1}+t_{j_0}}{2}\right)^2}{W'_g\left(\frac{t_j+t_{j+1}}{2}\right)^2} &\leq \frac{\left(1 + \frac{t_{j+1}+t_{j_0}}{2}\right)^2}{\left(1 + \frac{t_j+t_{j+1}}{2}\right)^2} \frac{W'_1\left(\frac{t_{j+1}+t_{j_0}}{2}\right)^2}{W'_1\left(\frac{t_j+t_{j+1}}{2}\right)^2} \\
&\leq \frac{1 + \frac{t_{j+1}+t_{j_0}}{2}}{1 + \frac{t_j+t_{j+1}}{2}} \frac{1 - \frac{t_j+t_{j+1}}{2}}{1 - \frac{t_{j+1}+t_{j_0}}{2}} \\
&\leq 2 \frac{1 + \frac{t_{j+1}+t_{j_0}}{2}}{1 + t_{j+1}} \frac{1 - \frac{t_j+t_{j+1}}{2}}{1 - \frac{t_{j+1}+t_{j_0}}{2}},
\end{aligned}$$

and thus, again by Theorem 2.6,

$$\sum_1 \leq \frac{c_2^2}{2c_1^2} \sum_{j=0}^{j_0-2} \frac{1}{(j_0 - j - 1)^2} \frac{1 + \frac{t_{j+1}+t_{j_0}}{2}}{1 + t_{j+1}} \frac{1 - \frac{t_j+t_{j+1}}{2}}{1 - \frac{t_{j+1}+t_{j_0}}{2}}.$$

The following arguments depend on the sign of t_{j+1} . We therefore set $j_1 = j_0 - 1$ if $t_{j_0-1} < 0$, and else chose $j_1 \in \{0, 1, \dots, j_0 - 2\}$ with $t_{j_1} < 0 \leq t_{j_1+1}$, and cut our sum into two parts $\Sigma_1 = \Sigma_{1,1} + \Sigma_{1,2}$, where in the first sum $j \in \{j_1, \dots, j_0 - 2\}$, and in the second one $j \in \{0, \dots, j_1 - 1\}$.

If $j \geq j_1$ and thus $t_{j+1} \geq 0$,

$$\frac{1 + \frac{t_{j+1}+t_{j_0}}{2}}{1 + t_{j+1}} \frac{1 - \frac{t_j+t_{j+1}}{2}}{1 - \frac{t_{j+1}+t_{j_0}}{2}} \leq 2 \frac{1 - \frac{t_j+t_{j+1}}{2}}{1 - \frac{t_{j+1}+t_{j_0}}{2}} \leq 4 \frac{1 - \frac{t_j+t_{j+1}}{2}}{1 - t_{j+1}} \leq 36,$$

where in the last inequality we have applied Lemma B.5. Hence,

$$\sum_{1,1} \leq \frac{18c_2^2}{c_1^2} \sum_{j=j_1}^{j_0-2} \frac{1}{(j_0 - j - 1)^2}. \quad (23)$$

If $j < j_1$ and thus $t_{j+1} < 0$,

$$\frac{1 + \frac{t_{j+1}+t_{j_0}}{2}}{1 + t_{j+1}} \frac{1 - \frac{t_j+t_{j+1}}{2}}{1 - \frac{t_{j+1}+t_{j_0}}{2}} \leq 4 \frac{1 + \frac{t_{j+1}+t_{j_0}}{2}}{1 + t_{j+1}} \leq \frac{\pi^2}{2} \frac{j_0^2}{(j+1)^2},$$

where the last inequality follows from Lemma B.6. Thus,

$$\begin{aligned} \sum_{1,2} &\leq \frac{c_2^2 \pi^2}{4c_1^2} \sum_{j=0}^{j_1-1} \frac{j_0^2}{(j_0 - j - 1)^2 (j+1)^2} \\ &\leq \frac{c_2^2 \pi^2}{4c_1^2} \left(\sum_{j=0, j+1 < j_0/2}^{j_1-1} \frac{4}{(j+1)^2} + \sum_{j=0, j+1 \geq j_0/2}^{j_1-1} \frac{4}{(j_0 - j - 1)^2} \right). \end{aligned}$$

Since $\pi^2 \leq 18$, a combination with (23) gives the upper bound for Σ_1 as claimed in Proposition 2.8. \square

We finally discuss in our third proposition the expression

$$\sum_2 = \sum_{j=\max\{0, j_0-1\}}^{\min\{j_0+1, k-1\}} \int_{t_j}^{t_{j+1}} \log \left| \frac{x - \xi_j}{x - t} \right| W_g'(t) dt,$$

where we integrate in a neighborhood of x and thus have to deal with the logarithmic singularity of the integrand. Here again Theorem 2.6 will be essential. As maybe expected from [17], our proof for $j_0 \in \{1, 2, \dots, k-2\}$ is quite different from that for x close to the endpoints and thus $j_0 \in \{0, k-1\}$: in the first case,

we require lower and upper bounds for the ratio of the lengths of two consecutive intervals $[t_j, t_{j+1}]$ established in Lemma B.9, whereas in the second case we require upper bounds for

$$k\sqrt{t_1 - t_0}, \quad \text{and} \quad k\sqrt{t_k - t_{k-1}},$$

see Lemma B.8.

Proposition 2.9 *There holds*

$$\sum_2 \leq 6c_2 + \log\left(\frac{6c_2\sqrt{c_5}}{c_1}\right) \leq 32.$$

Proof: By integration by part,

$$\begin{aligned} \int_{t_j}^{t_{j+1}} \log\left|\frac{x - \xi_j}{x - t}\right| W'_g(t) dt &= \left[\log\left|\frac{x - \xi_j}{x - t}\right| (W_g(t) - W_g(x)) \right]_{t_j}^{t_{j+1}} \\ &\quad + \int_{t_j}^{t_{j+1}} \frac{W_g(t) - W_g(x)}{t - x} dt. \end{aligned}$$

In order to make our formulas a bit easier to read, we write $j_1 = \max\{0, j_0 - 1\}$, $j_2 = \min\{k - 1, j_0 + 1\}$, and get $\Sigma_2 = \Sigma_{2,1} + \Sigma_{2,2}$, with

$$\begin{aligned} \Sigma_{2,1} &= \int_{t_{j_1}}^{t_{j_2+1}} \frac{W_g(t) - W_g(x)}{t - x} dt, \\ \Sigma_{2,2} &= \sum_{j=j_1}^{j_2} \left(\log\left|\frac{x - \xi_j}{x - t_{j+1}}\right| (W_g(t_{j+1}) - W_g(x)) \right. \\ &\quad \left. + \log\left|\frac{x - \xi_j}{x - t_j}\right| (W_g(x) - W_g(t_j)) \right). \end{aligned}$$

The first term is easily bounded. Indeed, using Theorem 2.6, we get

$$\begin{aligned} \Sigma_{2,1} &\leq c_2 \int_{t_{j_1}}^{t_{j_2+1}} W'_g\left(\frac{t+x}{2}\right) dt \\ &\leq 2c_2 \left(W_g\left(\frac{x+t_{j_2+1}}{2}\right) - W_g\left(\frac{x+t_{j_1}}{2}\right) \right) \\ &\leq 2c_2 (W_g(t_{j_2+1}) - W_g(t_{j_1})) = 2c_2(j_2 + 1 - j_1), \end{aligned}$$

and thus $\Sigma_{2,1} \leq 6c_2$ for $j_0 \in \{1, \dots, k - 2\}$, and $\Sigma_{2,1} \leq 4c_2$ for $j_0 \in \{0, k - 1\}$.

It remains to give an upper bound for $\Sigma_{2,2}$. We first study the case $j_0 \in \{1, \dots, k - 2\}$ and thus $j_1 = j_0 - 1$, $j_2 = j_0 + 1$. As $W_g(t_{j+1}) - W_g(t_j) = 1$ for every j , we notice that

$$W_g(t_{j_0+2}) - W_g(x) = 2(W_g(t_{j_0+1}) - W_g(x)) + (W_g(x) - W_g(t_{j_0})) \quad (24)$$

and

$$W_g(x) - W_g(t_{j_0-1}) = (W_g(t_{j_0+1}) - W_g(x)) + 2(W_g(x) - W_g(t_{j_0})).$$

Inserting this information into $\Sigma_{2,2}$, we obtain, after some elementary computations,

$$\begin{aligned} \Sigma_{2,2} &= \underbrace{[W_g(t_{j_0+1}) - W_g(x)]}_{\geq 0} \log \frac{|x - \xi_{j_0}|}{|x - t_{j_0+2}|} \underbrace{\frac{|x - \xi_{j_0-1}|}{|x - t_{j_0-1}|}}_{\leq 1} \underbrace{\frac{|x - \xi_{j_0+1}|}{|x - t_{j_0+2}|}}_{\leq 1} \\ &\quad + \underbrace{[W_g(x) - W_g(t_{j_0})]}_{\geq 0} \log \frac{|x - \xi_{j_0}|}{|x - t_{j_0-1}|} \underbrace{\frac{|x - \xi_{j_0-1}|}{|x - t_{j_0-1}|}}_{\leq 1} \underbrace{\frac{|x - \xi_{j_0+1}|}{|x - t_{j_0+2}|}}_{\leq 1} \\ &\leq [W_g(t_{j_0+1}) - W_g(x)] \log \left| \frac{x - \xi_{j_0}}{x - t_{j_0+2}} \right| + [W_g(x) - W_g(t_{j_0})] \log \left| \frac{x - \xi_{j_0}}{x - t_{j_0-1}} \right| \\ &\leq \log \left(\frac{6c_2\sqrt{c_5}}{c_1} \right). \end{aligned}$$

In order to justify the last inequality, we have to distinguish two cases. In case $x \in [\xi_{j_0}, t_{j_0+1}]$, we find that $\log \left| \frac{x - \xi_{j_0}}{x - t_{j_0-1}} \right| \leq 0$, implying that

$$\Sigma_{2,2} \leq [W_g(x) - W_g(t_{j_0})] \log \left| \frac{x - \xi_{j_0}}{x - t_{j_0-1}} \right| \leq \log \frac{t_{j_0+1} - t_{j_0}}{t_{j_0} - t_{j_0-1}},$$

and we conclude with help of Lemma B.9. The case $x \in [t_{j_0}, \xi_{j_0}]$ is similar, here $\Sigma_{2,2} \leq \log \frac{t_{j_0+1} - t_{j_0}}{t_{j_0+2} - t_{j_0+1}}$, and we conclude again using Lemma B.9.

Let us now consider the sum $\Sigma_{2,2}$ for the particular case $j_0 = 0$ and thus $j_1 = 0$, $j_2 = 1$. Using (24), this sum can be bounded above as before by

$$\begin{aligned} \Sigma_{2,2} &= \underbrace{[W_g(t_1) - W_g(x)]}_{\geq 0} \log \frac{|x - \xi_0|}{|x - t_2|} \underbrace{\frac{|x - \xi_1|}{|x - t_2|}}_{\leq 1} \\ &\quad + \underbrace{[W_g(x) - W_g(t_0)]}_{\geq 0} \log \frac{|x - \xi_0|}{|x - t_0|} \underbrace{\frac{|x - \xi_1|}{|x - t_2|}}_{\leq 1} \\ &\leq [W_g(t_1) - W_g(x)] \log \left| \frac{x - \xi_0}{x - t_2} \right| + [W_g(x) - W_g(t_0)] \log \left| \frac{x - \xi_0}{x - t_0} \right|. \end{aligned}$$

We have to consider three different cases: if $x \in [\frac{t_0 + \xi_0}{2}, \frac{\xi_0 + t_1}{2}]$ then $\Sigma_{2,2} \leq 0$. If $x \in [\frac{\xi_0 + t_1}{2}, t_1]$, then

$$\Sigma_{2,2} \leq [W_g(t_1) - W_g(x)] \log \left| \frac{x - \xi_0}{x - t_2} \right| \leq \log \frac{t_1 - t_0}{t_2 - t_1} \leq \log \left(\frac{6c_2\sqrt{c_5}}{c_1} \right)$$

as before. Finally, in the case $x \in [t_0, \frac{t_0 + \xi_0}{2}]$, we use the fact that $|x - \xi_0| \leq t_1 - t_0$, and apply Theorem 2.6 in order to get

$$\begin{aligned} \sum_{2,2} &\leq [W_g(x) - W_g(t_0)] \log \left| \frac{t_1 - t_0}{x - t_0} \right| \\ &\leq c_2(x - t_0) g\left(\frac{t_0 + x}{2}\right) W_1'\left(\frac{t_0 + x}{2}\right) \log \left| \frac{t_1 - t_0}{x - t_0} \right|. \end{aligned}$$

Since $\frac{t_0 + x}{2} \leq 0$ and $\frac{t_0 + x}{2} \leq t_1$, we have that

$$(x - t_0) g\left(\frac{t_0 + x}{2}\right) W_1'\left(\frac{t_0 + x}{2}\right) \leq \frac{k}{\pi} g(t_1) \frac{x - t_0}{\sqrt{1 + \frac{t_0 + x}{2}}} = \frac{k\sqrt{2}}{\pi} g(t_1) \sqrt{x - t_0}.$$

Using the fact that $\max_{y \geq 0} \sqrt{y} \log \frac{1}{y} = 2/e$, we conclude with help of Lemma B.8(a) that

$$\sum_{2,2} \leq \frac{2c_2\sqrt{2}}{\pi e} k g(t_1) \sqrt{t_1 - t_0} \leq \frac{6}{e} c_2.$$

The reasoning for $j_0 = k - 1$ is similar and allows for the same conclusion, we just have to replace Lemma B.8(a) by Lemma B.8(b) providing an upper bound for $k\sqrt{t_k - t_{k-1}}$. Thus

$$\sum_2 = \sum_{2,1} + \sum_{2,2} \leq \max \left\{ 6c_2 + \log \left(\frac{6c_2\sqrt{c_5}}{c_1} \right), 4c_2 + \frac{6}{e} c_2 \right\},$$

and the statement follows. \square

3 Proof of the main theorems

3.1 Proof of Theorem 1.3

Let us first show our claim (6) for the support of the equilibrium measure μ . We observe that the external field $Q = U^{\rho/k}$ is convex on $\Sigma = [\alpha, b]$ and hence $\text{supp}(\mu) = [a, b']$ for some $\alpha \leq a < b' \leq b$ by [16, Theorem IV.1.10(b)]. Since $U^\mu + Q$ is strictly decreasing on (b', ∞) , the equilibrium condition (2) tells us that necessarily $b = b'$. We show below the two implications

$$\text{for some } a > \alpha : \quad \text{supp}(\mu) = [a, b] \text{ implies that } \eta(a) = k + \|\rho\|, \quad (25)$$

$$\text{supp}(\mu) = [\alpha, b] \text{ implies that } \eta(\alpha) \leq k + \|\rho\|, \quad (26)$$

with the strictly decreasing η as in (5). Since there is exactly one solution $> \alpha$ of the equation $\eta(a) = k + \|\rho\|$ iff $\eta(\alpha) > k + \|\rho\|$, our statement on $\text{supp}(\mu)$ follows.

For a proof of (25), suppose that $\text{supp}(\mu) = [a, b]$ for some $a > \alpha$. Then, by [16, Theorem IV.1.11(ii)], the derivative of the F -functional of Mhaskar and Saff

$$y \mapsto \log \frac{b-y}{4} - \int Q d\omega_{[y,b]} = \left(1 + \frac{\|\rho\|}{k}\right) \log \frac{b-y}{4} + \frac{1}{k} \int g_{[y,b]}(x, \infty) d\rho(x)$$

must vanish at $y = a$, and a small calculation gives the necessary condition

$$0 = \frac{1}{k(b-a)} \left(k + \|\rho\| - \eta(a)\right)$$

and thus $\eta(a) = k + \|\rho\|$, implying (25).

In order to show (26) together with the representation (6) of f , let $\text{supp}(\mu) = [a, b]$ for some $a \in [\alpha, b)$. We denote by $Bal(\rho, [a, b])$ the measure obtained by balayage onto the interval $[a, b]$, see [16, §II.4]. Then, by construction,

$$k\mu + Bal(\rho, [a, b])$$

is a positive measure of mass $k + \|\rho\|$ having a constant potential on $[a, b]$, and thus $k\mu + Bal(\rho, [a, b]) = (k + \|\rho\|)\omega_{[a,b]}$. We apply the explicit formula for balayage onto an interval given in [16, Eqn. (II.4.47)], and get for $t \in [a, b]$

$$g(t) := \frac{d\mu}{d\omega_{[a,b]}}(t) = \frac{k + \|\rho\|}{k} - \frac{1}{k} \int_{-\infty}^a \frac{\sqrt{(b-y)(a-y)}}{t-y} d\rho(y).$$

As a consequence

$$0 \leq \lim_{t \rightarrow a+0} g(t) = \frac{k + \|\rho\| - \eta(a)}{k},$$

showing that $\eta(a) \leq k + \|\rho\|$ is finite. In particular, in case $a = \alpha$ we get (26). Moreover, by [16, Eqn. (II.5.4)], with a suitable $F \in \mathbb{R}$,

$$\begin{aligned} k(F - U^\mu(x) - Q(x)) &= kF - U^{(k+\|\rho\|)\omega_{[a,b]} + \rho - Bal(\rho, [a,b])}(x) \\ &= (k + \|\rho\|)g(x, \infty) - \int g(x, y) d\rho(y), \end{aligned}$$

the right-hand side vanishing on $[a, b]$, and thus the constant F coincides with the one in (2). Hence, the above expression equals $kf(x)$, showing (6).

It remains to show that Theorem 2.1 implies (7), where we start to verify the hypotheses on

$$g(t) = \frac{d\mu}{d\omega_{[a,b]}}(t) = \frac{k + \|\rho\| - \eta(a)}{k} + \frac{t-a}{k} \int_{-\infty}^a \sqrt{\frac{b-y}{a-y}} \frac{d\rho(y)}{t-y}.$$

We first observe that g is differentiable on $(a, b]$, with derivative

$$g'(t) = \frac{1}{k} \int_{-\infty}^a \sqrt{\frac{b-y}{a-y}} \frac{a-y}{(t-y)^2} d\rho(y),$$

which is both ≥ 0 and decreasing in $t \in (a, b]$. Hence g is increasing and concave in (a, b) , and, by a similar argument, $h(t) := g(t)/(t - a)$ is convex on (a, b) . Thus the assumptions of Theorem 2.1 hold. With $P_k \in \Pi_k$ as in Theorem 2.1 we have that

$$\begin{aligned} \log \|w^k P_k\|_{[a,b]} &= \max_{x \in [a,b]} -kQ(x) + \log |P_k(x)| \\ &\leq \max_{x \in [a,b]} -kQ(x) - kU^\mu(x) + C_{BW} = -kF + C_{BW}, \end{aligned}$$

where for obtaining the inequality we have applied Theorem 2.1(a), and in the last equality we have used (2) and in particular the fact that f vanishes on $[a, b]$. Also, for $x \in \mathbb{R} \setminus (a, b)$, we deduce from Theorem 2.1(b) and (2) that

$$\log w(x)^k |P_k(x)| \geq -kQ(x) - kU^\mu(x) = kf(x) - kF.$$

Combining these two inequalities gives (7).

3.2 Proof of Theorem 1.7

For our proof of Theorem 1.7(a), we choose $n > d + 1$ as in the statement such that (14) and (15) hold. By our assumption (14) on λ_j , we may choose $\tilde{\lambda}_j \leq \lambda_j$ such that

$$\sigma((-\infty, \tilde{\lambda}_j]) = \frac{j}{N} \quad \text{for } j = 1, 2, \dots, d + 1.$$

Consider $k = n - d \geq 2$, and the two measures of mass d

$$\rho = \delta_{\lambda_1} + \dots + \delta_{\lambda_d}, \quad \tilde{\rho} = N\sigma|_{(-\infty, \tilde{\lambda}_d]}.$$

It is not too difficult to check that $U^{\tilde{\rho}}(x) - U^\rho(x)$ is ≤ 0 for $x \in [\lambda_{d+1}, b]$, and ≥ 0 for $x = 0$. Hence, going back to the definition of $E_n([\lambda_{d+1}, b])$, we get the chain of inequalities

$$\begin{aligned} E_n([\lambda_{d+1}, b]) &= \min_{p \in \Pi_k} \frac{\|e^{-U^\rho} p\|_{[\lambda_{d+1}, b]}}{e^{-U^\rho(0)} |p(0)|} \leq \min_{p \in \Pi_k} \frac{\|e^{-U^{\tilde{\rho}}} p\|_{[\lambda_{d+1}, b]}}{e^{-U^{\tilde{\rho}}(0)} |p(0)|} \\ &\leq \exp\left(C_{BW} + kU^\mu(0) + U^{\tilde{\rho}}(0) - kF\right), \end{aligned}$$

where in the last inequality we have applied Theorem 1.3 with $\alpha = \lambda_{d+1}$, and the external field $Q(x) = U^{\tilde{\rho}/k}(x)$, and where the extremal measure μ and the constant F are as in (1), (2). On the other hand, with $t = n/N$, we know from [3, Theorem 2.1] that

$$\exp\left(C_{BW} - N \int_0^{n/N} g_{S(\tau)}(0, \infty) d\tau\right) = \exp\left(C_{BW} + n(U^{\nu_{t,\sigma}}(0) - C_{t,\sigma})\right),$$

with $\nu_{t,\sigma}$ the solution of the constrained equilibrium problem mentioned in the paragraph after (11), and $C_{t,\sigma}$ the corresponding constant. Thus, for establishing Theorem 1.7(a), it only remains to show the inequality

$$kU^\mu(x) + U^{\tilde{\rho}}(x) - kF \leq n(U^{\nu_{t,\sigma}}(x) - C_{t,\sigma}) \quad (27)$$

for $x = 0$.

Let us first show that (27) holds for $x \in \text{supp}(\mu)$. Indeed, since $\text{supp}(\mu) \subset [\lambda_{d+1}, b] \subset [a(t), b] = \text{supp}(\sigma/t - \nu_{t,\sigma})$ by assumption (15), we find from the respective equilibrium conditions for both extremal problems that both expressions on the left-hand side and on the right-hand side of (27) vanish for $x \in \text{supp}(\mu)$. We also know that all measures involved in (27) have finite energy, with masses $\|k\mu + \tilde{\rho}\| = k + d = n = n \|\nu_{t,\sigma}\|$. Let us show that $\tilde{\rho} \leq n\nu_{t,\sigma}$. Indeed, $\tilde{\lambda}_d \leq \lambda_d < a(t)$ by construction and (15), and thus, by definition of $S(t) = [a(t), b] = \text{supp}(\sigma/t - \nu_{t,\sigma})$,

$$n\nu_{t,\sigma}|_{(-\infty, \tilde{\lambda}_d]} = \frac{n}{t}\sigma|_{(-\infty, \tilde{\lambda}_d]} = N\sigma|_{(-\infty, \tilde{\lambda}_d]} = \tilde{\rho}.$$

Hence, by subtracting $U^{\tilde{\rho}}(x)$ from both sides of (27), we get from the principle of domination for logarithmic potentials [16, Theorem II.3.2] that (27) holds for all $x \in \mathbb{C}$, and in particular for $x = 0$, which concludes our proof of Theorem 1.7(a).

For our proof of Theorem 1.7(b), we first observe that our assumption of Lemma 1.5 on σ and the choice of the eigenvalues $\lambda_{1,N} < \lambda_{2,N} < \dots$ of A_N , allows to show that not only (10) but also the quite technical [3, Conditions (i)–(iv)] hold, we omit details. As a consequence of [3, Theorem 2.2],

$$\lim_{\substack{n, N \rightarrow \infty \\ n/N \rightarrow t}} E_n(\Lambda(A_N))^{1/n} = \exp\left(-\frac{1}{t} \int_0^t g_{S(\tau)}(0, \infty) d\tau\right),$$

that is, we have equality in (11). Then, using Theorem 1.7(a) and the simple inequality $E_n(\Lambda(A_N)) \leq E_n([\lambda_{d+1}, b])$,

$$\begin{aligned} \liminf_{\substack{n, N \rightarrow \infty \\ n/N \rightarrow t}} E_{n,N}([\lambda_{d_{n,N}+1, N}, b])^{1/n} &\leq \limsup_{\substack{n, N \rightarrow \infty \\ n/N \rightarrow t}} E_{n,N}([\lambda_{d_{n,N}+1, N}, b])^{1/n} \\ &\leq \exp\left(-\frac{1}{t} \int_0^t g_{S(\tau)}(0, \infty) d\tau\right) = \lim_{\substack{n, N \rightarrow \infty \\ n/N \rightarrow t}} E_n(\Lambda(A_N))^{1/n} \\ &\leq \liminf_{\substack{n, N \rightarrow \infty \\ n/N \rightarrow t \\ 0 \leq d < n-1}} E_{n,N}([\lambda_{d+1, N}, b])^{1/n} \leq \liminf_{\substack{n, N \rightarrow \infty \\ n/N \rightarrow t}} E_{n,N}([\lambda_{d_{n,N}+1, N}, b])^{1/n}, \end{aligned}$$

and the statement of Theorem 1.7(b) follows.

4 Conclusions

In the particular case of an external field being given by the logarithmic potential of some positive measure supported on the left of Σ , we have shown that the weighted Bernstein-Walsh inequality is sharp up to some new universal constant C_{BW} . Our main tool is a variation of the technique of Totik of discretizing a logarithmic potential, provided that the underlying measure has a weight satisfying some monotonicity and/or convexity assumptions.

This new sharpness result for the weighted Bernstein-Walsh inequality leads to a variety of new explicit bounds for the convergence of conjugate gradients

if we fix in advance a fixed number of small eigenvalues being considered as outliers. By approximately optimizing the number of outliers, we are able to partly show a conjecture formulated by Beckermann & Kuijlaars [3] in terms of means of Green functions, and establish a new upper bound for conjugate gradients in form of an inequality for every iteration index n . Such bounds are of practical interest since the results of [3] are only of asymptotic nature. In addition, our bounds are valid for a single matrix and do no longer require to consider sequences of systems of equations with a joint eigenvalue distribution. We also give some (academic) numerical examples showing that this new bound perfectly matches the CG error (up to the choice of C_{BW}).

We believe that, with an optimal choice of C_{BW} , the quantity $e^{C_{BW}}$ is of modest size. This is clearly not true for our present explicit upper bound of C_{BW} , and remains a direction of future research, maybe asymptotic analysis could be helpful.

We also believe that our result on the discretization of a potential can be generalized to more general measures, for example without the assumption that $t \mapsto \frac{g(t)}{t-a}$ is convex on (a, b) , which is used only once. This possibly would allow us to consider both small and large eigenvalues as outliers.

Finally, the above-mentioned conjecture on the CG convergence remains open for general sets $S(t)$.

References

- [1] D. Bau and L. N. Trefethen. Numerical linear algebra, SIAM, Philadelphia PA, 1997.
- [2] B. Beckermann, Discrete orthogonal polynomials and superlinear convergence of Krylov subspace methods in numerical linear algebra, *Lecture Notes in Mathematics* **1883**, Springer-Verlag (2006) 119-185.
- [3] B. Beckermann, A.B.J. Kuijlaars, Superlinear convergence of conjugate gradients, *SIAM J. Numer. Anal.* **39** (2001) 300-329.
- [4] B. Beckermann, A.B.J. Kuijlaars, On the sharpness of an asymptotic error estimate for conjugate gradients, *BIT* **41** (2001) 856-867.
- [5] B. Beckermann, A.B.J. Kuijlaars, Superlinear CG convergence for special right-hand sides, *Electron. Trans. Numer. Anal.* **14** (2002) 1-19.
- [6] P.D. Dragnev, E.B. Saff, Constrained energy problems with applications to orthogonal polynomials of a discrete variable, *J. Anal. Math.* **72** (1997) 223-259.
- [7] T.A. Driscoll, K.-C. Toh and L.N. Trefethen, From potential theory to matrix iterations in six steps, *SIAM Review* **40** (1998) 547-578.
- [8] S.L. Campbell, I.C.F. Ipsen, C.T. Kelley, C.D. Meyer, GMRES and the minimal polynomial, *BIT* **36** (1996) 664-675.

- [9] A.B.J. Kuijlaars, Which eigenvalues are found by the Lanczos method?, *SIAM J. Matrix Anal. Appl.* **22** (2000) 306-321.
- [10] A.B.J. Kuijlaars, Convergence analysis of Krylov subspace iterations with methods from potential theory. *SIAM Rev.* **48** (2006) 3-40.
- [11] A.L. Levin, D.S. Lubinsky, Christoffel functions, orthogonal polynomials, and Nevai's conjecture for Freud weights, *Constr. Approx.* **8** (1992) 463-535.
- [12] A.L. Levin, D.S. Lubinsky, Christoffel functions and orthogonal polynomials for exponential weights on $[-1, 1]$, *Mem. Amer. Math. Soc.* **111**, AMS (1994).
- [13] D.S. Lubinsky, V. Totik, How to discretize a logarithmic potential? *Acta Sci. Math. (Szeged)* **57** (1993) 419-428.
- [14] G. Meinardus, Approximation of functions: theory and numerical methods, Springer-Verlag (1967).
- [15] E.A. Rakhmanov, Equilibrium measure and the distribution of zeros of the extremal polynomials of a discrete variable, *Sbornik: Mathematics* **187** (1996) 1213-1228.
- [16] E.B. Saff, V. Totik, Logarithmic potentials with external fields, Springer-Verlag (1997).
- [17] V. Totik, Weighted approximation with varying weight, *Lecture Notes in Mathematics* **1569**, Springer-Verlag (1994).
- [18] A. van der Sluis, H.A. van der Vorst, The rate of convergence of conjugate gradients, *Numer. Math.* **48** (1986) 543-560.

A Proof of the mean value property of Theorem 2.6

As said before, a central role in our analysis is played by the mean value property of the cumulative distribution function W_g stated in Theorem 2.6: there exist constants $c_1 = \frac{1}{4}$ and $c_2 = \pi\sqrt{2}$ such that, for all $x, t \in [-1, 1]$,

$$c_1 W_g' \left(\frac{t+x}{2} \right) \leq \frac{W_g(t) - W_g(x)}{t-x} \leq c_2 W_g' \left(\frac{t+x}{2} \right).$$

The aim of this section is to provide a proof of this mean value property. We will first consider the two particular cases $g = 1$ in Lemma A.1 and $g(t) = 1 + t$ in Lemma A.2. The general case then will follow by concavity of g and by convexity of $h(t) = g(t)/(t+1)$. In what follows it will be convenient to consider the substitution $t = -\cos(\alpha)$ and $x = -\cos(\beta)$, $\alpha, \beta \in [0, \pi]$, where we can suppose without loss of generality that $t > x$, and thus $0 \leq \beta < \alpha \leq \pi$.

Lemma A.1 For every $x, t \in [-1, 1]$, we have for $c_3 = \pi/\sqrt{2}$

$$W_1'(\frac{t+x}{2}) \leq \frac{W_1(t) - W_1(x)}{t-x} \leq c_3 W_1'(\frac{t+x}{2}). \quad (28)$$

Proof: Elementary trigonometric formulas give

$$\frac{W_1(t) - W_1(x)}{t-x} = \frac{k}{\pi} \frac{\alpha - \beta}{\cos(\beta) - \cos(\alpha)} = \frac{k}{\pi} \frac{\frac{\alpha-\beta}{2}}{\sin(\frac{\alpha-\beta}{2})} \frac{1}{\sin(\frac{\alpha+\beta}{2})}.$$

Observing that $\frac{\alpha-\beta}{2} \in [0, \frac{\pi}{2}]$ and thus

$$\sin(\frac{\alpha-\beta}{2}) \leq \frac{\alpha-\beta}{2} \leq \frac{\pi}{2} \sin(\frac{\alpha-\beta}{2}),$$

we deduce that

$$\frac{k}{\pi} \frac{1}{\sin(\frac{\alpha+\beta}{2})} \leq \frac{W_1(t) - W_1(x)}{t-x} \leq \frac{k}{2} \frac{1}{\sin(\frac{\alpha+\beta}{2})}.$$

Since

$$W_1'(\frac{t+x}{2}) = \frac{k}{\pi} \frac{1}{\sqrt{1 - \cos^2(\frac{\alpha+\beta}{2}) \cos^2(\frac{\alpha-\beta}{2})}},$$

the left-hand inequality in (28) immediately follows.

If $\frac{\alpha+\beta}{2} \leq \frac{\pi}{2}$, then $0 \leq \frac{\alpha-\beta}{2} \leq \frac{\alpha+\beta}{2} \leq \frac{\pi}{2}$. If $\frac{\alpha+\beta}{2} \geq \frac{\pi}{2}$, then $0 \leq \frac{\alpha-\beta}{2} \leq \pi - \frac{\alpha+\beta}{2} \leq \frac{\pi}{2}$. In both cases we find that

$$1 - \cos^2(\frac{\alpha+\beta}{2}) \cos^2(\frac{\alpha-\beta}{2}) \leq 1 - \cos^4(\frac{\alpha+\beta}{2}) \leq 2 \left(1 - \cos^2(\frac{\alpha+\beta}{2})\right),$$

which implies the right-hand side of (28). \square

We now turn to the special case $g(y) = 1 + y$ where we only require one inequality for $W_g = W_{1+y}$.

Lemma A.2 For every $x, t \in [-1, 1]$, we have for $c_4 = 1/2$

$$c_4 W_{1+y}'(\frac{x+t}{2}) \leq \frac{W_{1+y}(t) - W_{1+y}(x)}{t-x}.$$

Proof: By Lemma A.1,

$$\frac{W_{1+y}(t) - W_{1+y}(x)}{t-x} \geq \frac{W_{1+y}(t) - W_{1+y}(x)}{W_1(t) - W_1(x)} W_1'(\frac{x+t}{2}).$$

Thus it is sufficient to show that

$$\frac{W_{1+y}(t) - W_{1+y}(x)}{W_1(t) - W_1(x)} \geq \frac{1}{2} \left(1 + \frac{x+t}{2}\right).$$

By definition of W_{1+y} ,

$$\begin{aligned}\frac{W_{1+y}(t) - W_{1+y}(x)}{W_1(t) - W_1(x)} &= \frac{1}{W_1(t) - W_1(x)} \int_x^t W'_{1+y}(s) ds \\ &= \frac{\alpha - \beta + \sin(\beta) - \sin(\alpha)}{\alpha - \beta} \\ &= 1 - \frac{2}{\alpha - \beta} \sin\left(\frac{\alpha - \beta}{2}\right) \cos\left(\frac{\alpha + \beta}{2}\right).\end{aligned}$$

Hence it remains to show that

$$\cos\left(\frac{\alpha + \beta}{2}\right) \left(2 \frac{\sin\left(\frac{\alpha - \beta}{2}\right)}{\frac{\alpha - \beta}{2}} - \cos\left(\frac{\alpha - \beta}{2}\right) \right) \leq 1.$$

Since $\gamma \mapsto 2 \sin(\gamma) - \gamma \cos(\gamma)$ is increasing in $[0, \pi/2]$, the factor in large brackets is ≥ 0 , and $\cos((\alpha + \beta)/2) \leq \cos((\alpha - \beta)/2)$. Thus we only have to consider the worst case $\gamma = (\alpha + \beta)/2 = (\alpha - \beta)/2 \in [0, \pi/2]$, with

$$\cos(\gamma) \left(2 \frac{\sin(\gamma)}{\gamma} - \cos(\gamma) \right) \leq 2 \cos(\gamma) - \cos^2(\gamma) \leq 1.$$

□

We are now prepared to give a proof of Theorem 2.6. For the upper bound, we use Lemma A.1 in order to conclude that

$$\begin{aligned}\frac{W_g(t) - W_g(x)}{t - x} &= \frac{W_g(t) - W_g(x)}{W_1(t) - W_1(x)} \frac{W_1(t) - W_1(x)}{t - x} \\ &\leq \frac{W_g(t) - W_g(x)}{W_1(t) - W_1(x)} c_3 W'_1\left(\frac{t+x}{2}\right).\end{aligned}$$

Recalling that g is concave, we get from the Jensen inequality

$$\begin{aligned}\frac{W_g(t) - W_g(x)}{W_1(t) - W_1(x)} &= \int_x^t \frac{W'_g(s)}{W_1(t) - W_1(x)} ds \\ &= \int_x^t g(s) \frac{W'_1(s)}{W_1(t) - W_1(x)} ds \\ &\leq g\left(\int_x^t s \frac{W'_1(s)}{W_1(t) - W_1(x)} ds\right) \\ &\leq 2g\left(\frac{t+x}{2}\right),\end{aligned}$$

the last inequality being established in Lemma A.3 below. Thus we obtain the upper bound with $c_2 = 2c_3 = \pi\sqrt{2}$. For the lower bound, our argument is similar, but now we use Lemma A.2 in order to get

$$\begin{aligned}\frac{W_g(t) - W_g(x)}{t - x} &= \frac{W_g(t) - W_g(x)}{W_{1+y}(t) - W_{1+y}(x)} \frac{W_{1+y}(t) - W_{1+y}(x)}{t - x} \\ &\geq \frac{W_g(t) - W_g(x)}{W_{1+y}(t) - W_{1+y}(x)} c_4 W'_{1+y}\left(\frac{x+t}{2}\right).\end{aligned}$$

Recalling that $h(y) = g(y)/(1+y)$ is convex, we get from the Jensen inequality

$$\begin{aligned} \frac{W_g(t) - W_g(x)}{W_{1+y}(t) - W_{1+y}(x)} &= \int_x^t h(s) \frac{W'_{1+y}(s)}{W_{1+y}(t) - W_{1+y}(x)} ds \\ &\geq h\left(\int_x^t s \frac{W'_{1+y}(s)}{W_{1+y}(t) - W_{1+y}(x)} ds\right) \\ &\geq \frac{1}{2} g\left(\frac{t+x}{2}\right), \end{aligned}$$

where for the last inequality we apply Lemma A.4 below. This gives us the lower bound with $c_1 = c_4/2 = 1/4$.

For concluding, it remains to establish two technical results.

Lemma A.3 For $-1 \leq x < t \leq 1$ we have

$$g\left(\int_x^t \frac{sW'_1(s)}{W_1(t) - W_1(x)} ds\right) \leq 2g\left(\frac{x+t}{2}\right).$$

Proof: Elementary trigonometric computations give

$$\bar{x} := \int_x^t \frac{sW'_1(s)}{W_1(t) - W_1(x)} ds = -\cos\left(\frac{\alpha+\beta}{2}\right) \frac{\sin\left(\frac{\alpha-\beta}{2}\right)}{\frac{\alpha-\beta}{2}} \in [-1, 1],$$

and

$$\frac{x+t}{2} = -\cos\left(\frac{\alpha+\beta}{2}\right) \cos\left(\frac{\alpha-\beta}{2}\right).$$

We now have to distinguish two cases. If $\frac{x+t}{2} \leq 0$ or, equivalently, $\cos\left(\frac{\alpha+\beta}{2}\right) \geq 0$ then, using the fact that $\sin(y) \geq y \cos(y) \geq 0$ for $y \in [0, \frac{\pi}{2}]$, we get $\bar{x} \leq \frac{x+t}{2}$, and the statement $g(\bar{x}) \leq g\left(\frac{x+t}{2}\right)$ follows (without a factor 2) by monotonicity of g . If however $\frac{x+t}{2} \geq 0$, then $1 \geq \bar{x} \geq \frac{x+t}{2} \geq 0$ by the same argument as in the first case. By concavity of g and (31) we get

$$\begin{aligned} g(\bar{x}) &\leq g\left(\frac{x+t}{2}\right) + g'\left(\frac{x+t}{2} - 0\right) \left(\bar{x} - \frac{x+t}{2}\right) \\ &\leq g\left(\frac{x+t}{2}\right) + g'\left(\frac{x+t}{2} - 0\right) \leq 2g\left(\frac{x+t}{2}\right), \end{aligned}$$

the last inequality being shown in Lemma B.1(b) below. Thus Lemma A.3 holds. \square

Lemma A.4 For $-1 \leq x < t \leq 1$ we have

$$h\left(\int_x^t s \frac{W'_{1+y}(s)}{W_{1+y}(t) - W_{1+y}(x)} ds\right) \geq \frac{1}{2} h\left(\frac{x+t}{2}\right).$$

Proof: Let us first show that

$$\bar{x} := \int_x^t s \frac{W'_{1+y}(s)}{W_{1+y}(t) - W_{1+y}(x)} ds \geq \frac{x+t}{2}. \quad (29)$$

We write shorter $w(s) = \frac{W'_{1+y}(s)}{W_{1+y}(t) - W_{1+y}(x)}$ being increasing in s . Hence

$$\begin{aligned} \bar{x} - \frac{x+t}{2} &= \int_x^t \left(s - \frac{x+t}{2}\right) w(s) ds \\ &= \int_x^t \left(s - \frac{x+t}{2}\right) (w(s) - w(\frac{x+t}{2})) ds \geq 0, \end{aligned}$$

as claimed in (29). Also, by definition, $\bar{x} \leq t$, and thus

$$h(\bar{x}) = \frac{g(\bar{x})}{\bar{x}+1} \geq \frac{g(\frac{x+t}{2})}{\bar{x}+1} \geq \frac{g(\frac{x+t}{2})}{t+1} \geq \frac{1}{2} \frac{g(\frac{x+t}{2})}{1 + \frac{x+t}{2}} = \frac{h(\frac{x+t}{2})}{2}.$$

□

B Some further technical lemmata

After having established the mean value property of W_g in §A, we gather in this section all the other technical properties of the abscissa $t_j = -\cos(\alpha_j)$ needed in §3.

In the sequel of this section we always suppose the conditions on k, g, h of Theorem 2.1 to be true, that is, k is some integer ≥ 2 , g is non-negative, increasing and concave, and $h(t) = g(t)/(t+1)$ is convex.

The first result summarizes some properties of the function g .

Lemma B.1 *The following properties hold:*

- (a) $h(t) = \frac{g(t)}{1+t}$ is decreasing on $(-1, 1]$;
- (b) $g'(t-0)(1+t) \leq g(t)$ for $t \in (-1, 1)$, and $g'(t-0) \leq g(t)$ for $t \in [0, 1)$;
- (c) $g(1) \leq 2$;
- (d) $g(1) \geq g(0) = h(0) \geq 1$.

Proof: We first recall that, by concavity of g on $[-1, 1]$, we have for all $-1 \leq x_1 < x_2 < x_3 \leq 1$ that

$$\frac{g(x_3) - g(x_2)}{x_3 - x_2} \leq \frac{g(x_3) - g(x_1)}{x_3 - x_1} \leq \frac{g(x_2) - g(x_1)}{x_2 - x_1}. \quad (30)$$

Since $g(-1) \geq 0$, we may therefore write

$$h(t) = \frac{g(t) - g(-1)}{t - (-1)} + \frac{g(-1)}{t+1}$$

as a sum of two decreasing functions, implying (a). Passing to the limit in (30), we also have that the directional derivatives $g'(x_2 - 0)$ and $g'(x_2 + 0)$ exist for all $x_2 \in (-1, 1)$, with

$$\frac{g(x_3) - g(x_2)}{x_3 - x_2} \leq g'(x_2 + 0) \leq g'(x_2 - 0) \leq \frac{g(x_2) - g(x_1)}{x_2 - x_1},$$

and in particular

$$g(x) \leq g(x_2) + g'(x_2 - 0)(x - x_2) \quad \text{for all } x \in [-1, 1]. \quad (31)$$

Setting $x = -1$, $x_2 = t$ in (31) leads to (b) since $g(-1) \geq 0$. Furthermore, using the concavity of g and setting $x_2 = 0$ in (31), we get for all $t \in [-1, 1]$ that

$$g(-1)\frac{1-t}{2} + g(1)\frac{1+t}{2} \leq g(t) \leq g(0) + tg'(0-).$$

Taking into account (20), multiplying by $W_1'(t)$ and integrating from -1 to 1 gives

$$k\frac{g(1)}{2} \leq k\frac{g(1) + g(-1)}{2} \leq k \leq kg(0),$$

implying parts (c) and (d). \square

The following elementary lemma will be helpful in what follows.

Lemma B.2 *For $\gamma \geq 0$ and $0 \leq \delta \leq \theta \leq \pi/2$ there holds*

$$\frac{\sin(\gamma)}{\sin(\delta)} \leq \frac{\theta}{\sin(\theta)} \frac{\gamma}{\delta} \leq \frac{\pi}{2} \frac{\gamma}{\delta}.$$

Proof: Since $x \mapsto x/\sin(x)$ is increasing in $[0, \pi/2]$, we have that

$$\frac{\sin(\gamma)}{\sin(\delta)} \leq \frac{\gamma}{\sin(\delta)} = \frac{\delta}{\sin(\delta)} \frac{\gamma}{\delta} \leq \frac{\theta}{\sin(\theta)} \frac{\gamma}{\delta} \leq \frac{\pi}{2} \frac{\gamma}{\delta}.$$

\square

The following result tells us that the angles α_j defined by $t_j = -\cos(\alpha_j)$ for $j = 0, 1, \dots, k$ have a quite regular behavior.

Lemma B.3 (a) *The sequence $(\alpha_{j+1} - \alpha_j)_{0 \leq j < k-1}$ is decreasing.*

(b) *For $j \in \{1, \dots, k-1\}$ there holds*

$$\frac{\alpha_{j+1} - \alpha_j}{\alpha_j} \leq \frac{1}{j},$$

(c) *For $j \in \{0, \dots, k-1\}$ we have*

$$\frac{\alpha_{j+1} - \alpha_j}{\alpha_{j+1}} \leq \frac{1}{j+1}.$$

(d) *For $j \in \{0, 1, \dots, k-2\}$*

$$\frac{\alpha_k - \alpha_j}{\alpha_k - \alpha_{j+1}} \leq 4.$$

Proof: Using that g is increasing, we get for $j \in \{0, \dots, k-1\}$

$$1 = \int_{t_j}^{t_{j+1}} W'_g(t) dt = \frac{k}{\pi} \int_{t_j}^{t_{j+1}} \frac{g(t)}{\sqrt{1-t^2}} dt$$

$$\begin{cases} \geq \frac{k}{\pi} g(t_j) \int_{t_j}^{t_{j+1}} \frac{dt}{\sqrt{1-t^2}} = \frac{k}{\pi} g(t_j) (\alpha_{j+1} - \alpha_j), \\ \leq \frac{k}{\pi} g(t_{j+1}) \int_{t_j}^{t_{j+1}} \frac{dt}{\sqrt{1-t^2}} = \frac{k}{\pi} g(t_{j+1}) (\alpha_{j+1} - \alpha_j), \end{cases}$$

implying that

$$\frac{\pi}{kg(t_{j+1})} \leq \alpha_{j+1} - \alpha_j \leq \frac{\pi}{kg(t_j)}. \quad (32)$$

Thus (a) holds. For a proof of (b), we apply (a) to conclude that, for $j \in \{1, 2, \dots, k-1\}$,

$$\begin{aligned} \frac{\alpha_{j+1} - \alpha_j}{\alpha_j} &= \frac{\alpha_{j+1} - \alpha_j}{\alpha_j - \alpha_0} = \frac{\alpha_{j+1} - \alpha_j}{\sum_{p=0}^{j-1} (\alpha_{p+1} - \alpha_p)} \\ &\leq \frac{\alpha_{j+1} - \alpha_j}{j(\alpha_j - \alpha_{j-1})} \leq \frac{1}{j}. \end{aligned}$$

A proof of part (c) follows the same lines, we omit details. Let us finally show (d). In case $k=2$, we know from (32) and Lemma B.1(c) that $\alpha_2 - \alpha_1 \geq \frac{\pi}{2g(1)} \geq \pi/4$, implying (d). In case $k \geq 3$ we can write

$$\begin{aligned} \frac{\alpha_k - \alpha_j}{\alpha_k - \alpha_{j+1}} &= 1 + \frac{\alpha_{j+1} - \alpha_j}{\alpha_k - \alpha_{j+1}} \\ &= 1 + \frac{\alpha_{j+1} - \alpha_j}{\sum_{\ell=j+1}^{k-1} (\alpha_{\ell+1} - \alpha_\ell)} \\ &\leq 1 + \frac{1}{k-1-j} \frac{\alpha_{j+1} - \alpha_j}{\alpha_k - \alpha_{k-1}}, \end{aligned}$$

where in the last inequality we have applied (a). By part (c), $\alpha_{j+1} - \alpha_j \leq \alpha_{j+1}/(j+1) \leq \pi/(j+1)$, and $\alpha_k - \alpha_{k-1} \geq \pi/(kg(t_k)) \geq \pi/(2k)$ by (32) and Lemma B.1(c). Hence using that $k \geq 3$, we obtain

$$\frac{\alpha_k - \alpha_j}{\alpha_k - \alpha_{j+1}} \leq 1 + 2 \frac{k}{(j+1)(k-1-j)} \leq 1 + 2 \frac{k}{k-1} \leq 4.$$

□

The following result is used in our proof of Proposition 2.7.

Lemma B.4 For $j \in \{1, 2, \dots, k-1\}$ there holds

$$1 + \frac{t_j + t_{j+1}}{2} \leq c_5(1 + t_j)$$

where $c_5 = \frac{3\pi}{4} + 1$.

Proof: By Lemma B.3(b), $\alpha_{j+1} - \alpha_j \leq \alpha_j/j \leq \pi$, implying that

$$\sin\left(\frac{\alpha_{j+1} - \alpha_j}{2}\right) \leq \sin\left(\frac{\alpha_j}{2j}\right) \leq \sin\left(\frac{\alpha_j}{2}\right).$$

Moreover, since $\alpha_{j+1} + \alpha_j \leq (2 + 1/j)\alpha_j$, we get by Lemma B.2

$$\begin{aligned} \frac{t_{j+1} - t_j}{t_j - t_0} &= \frac{\sin\left(\frac{\alpha_{j+1} - \alpha_j}{2}\right) \sin\left(\frac{\alpha_{j+1} + \alpha_j}{2}\right)}{\sin\left(\frac{\alpha_j}{2}\right) \sin\left(\frac{\alpha_j}{2}\right)} \\ &\leq \frac{\pi}{2} \left(2 + \frac{1}{j}\right) \leq \frac{3\pi}{2}, \end{aligned}$$

and Lemma B.4 follows. \square

Let us now show the two main properties required for our proof of Proposition 2.8.

Lemma B.5 For $j \in \{0, 1, \dots, k-2\}$ we have

$$1 - \frac{t_j + t_{j+1}}{2} \leq 9(1 - t_{j+1}).$$

Proof: We will show the equivalent statement

$$t_{j+1} - t_j \leq 16(t_k - t_{j+1}).$$

If $t_{j+1} \leq 1/\sqrt{2}$, we obtain

$$\frac{t_{j+1} - t_j}{1 - t_{j+1}} \leq \frac{t_{j+1} + 1}{1 - t_{j+1}} \leq (1 + \sqrt{2})^2 \leq 6 \leq 16.$$

It remains to consider the case $t_{j+1} \geq 1/\sqrt{2}$, and thus $\alpha_{j+1} \geq 3\pi/4$, or $(\pi - \alpha_{j+1})/2 \leq \pi/8$. Using first Lemma B.2 and then Lemma B.3(d), we obtain

$$\begin{aligned} \frac{t_{j+1} - t_j}{1 - t_{j+1}} &= -1 + \frac{1 - t_j}{1 - t_{j+1}} = -1 + \frac{1 + \cos(\alpha_j)}{1 + \cos(\alpha_{j+1})} \\ &= -1 + \frac{\cos^2\left(\frac{\alpha_j}{2}\right)}{\cos^2\left(\frac{\alpha_{j+1}}{2}\right)} = -1 + \frac{\sin^2\left(\frac{\pi}{2} - \frac{\alpha_j}{2}\right)}{\sin^2\left(\frac{\pi}{2} - \frac{\alpha_{j+1}}{2}\right)} \\ &\leq -1 + \left(\frac{\pi/8}{\sin(\pi/8)} \frac{\pi - \alpha_j}{\pi - \alpha_{j+1}}\right)^2 \leq -1 + 16 \left(\frac{\pi/8}{\sin(\pi/8)}\right)^2 \leq 16. \end{aligned}$$

\square

Lemma B.6 For $j \leq j_0 - 1 \leq k-2$ we have

$$\frac{1 + \frac{t_{j+1} + t_{j_0}}{2}}{1 + t_{j+1}} \leq \frac{\pi^2}{8} \frac{j_0^2}{(j+1)^2}.$$

Proof: Notice that, by Lemma B.2,

$$\begin{aligned} \frac{1 + \frac{t_{j+1} + t_{j_0}}{2}}{1 + t_{j+1}} &= 1 + \frac{t_{j_0} - t_{j+1}}{2(1 + t_{j+1})} = 1 + \frac{\sin(\frac{\alpha_{j+1} + \alpha_{j_0}}{2}) \sin(\frac{\alpha_{j_0} - \alpha_{j+1}}{2})}{2 \sin^2(\frac{\alpha_{j+1}}{2})} \\ &\leq 1 + \frac{\pi^2 (\alpha_{j+1} + \alpha_{j_0})(\alpha_{j_0} - \alpha_{j+1})}{8 \alpha_{j+1}^2}. \end{aligned}$$

Applying Lemma B.3(a), and recalling that $j_0 \geq j + 1$, we obtain

$$\begin{aligned} \frac{\alpha_{j_0} - \alpha_{j+1}}{\alpha_{j+1}} &= \frac{\sum_{\ell=j+1}^{j_0-1} (\alpha_{\ell+1} - \alpha_\ell)}{\sum_{\ell=0}^j (\alpha_{\ell+1} - \alpha_\ell)} \\ &\leq \frac{(j_0 - j - 1)(\alpha_{j+2} - \alpha_{j+1})}{(j + 1)(\alpha_{j+1} - \alpha_j)} \leq \frac{j_0 - j - 1}{j + 1}, \end{aligned}$$

and

$$\frac{\alpha_{j_0} + \alpha_{j+1}}{\alpha_{j+1}} = \frac{\alpha_{j_0} - \alpha_{j+1}}{\alpha_{j+1}} + 2 \leq \frac{j_0 + j + 1}{j + 1}.$$

Combining the three inequalities, we deduce that

$$\frac{1 + \frac{t_{j+1} + t_{j_0}}{2}}{1 + t_{j+1}} \leq 1 + \frac{\pi^2 j_0^2 - (j + 1)^2}{8 (j + 1)^2} \leq \frac{\pi^2 j_0^2}{8 (j + 1)^2}.$$

□

The three following results are required in our proof of Proposition 2.9.

Lemma B.7 *For $k \geq 2$, we have*

$$t_{k-1} \geq 0.$$

Proof: Suppose that $t_{k-1} < 0$. Then using Lemma B.1(d), and the fact that g is increasing allows us to find a contradiction

$$\begin{aligned} 1 &= \int_{t_{k-1}}^1 W'_g(t) dt > \int_0^1 W'_g(t) dt = \int_0^1 \frac{g(t)k}{\pi} \frac{1}{\sqrt{1-t^2}} dt \\ &\geq \frac{g(0)k}{\pi} \int_0^1 \frac{1}{\sqrt{1-t^2}} dt \geq \frac{k}{2}. \end{aligned}$$

□

Lemma B.8 *There holds*

- (a) $\sqrt{t_1 - t_0} \leq \frac{3\pi}{\sqrt{2}g(t_1)k}$;
- (b) $\sqrt{t_k - t_{k-1}} \leq \frac{\pi}{kg(1)}$;
- (c) *For all $j \in \{0, 1, \dots, k-1\}$ we have $t_{j+1} - t_j \leq (\frac{12\pi}{k})^{1/3}$.*

Proof: By Lemma B.1(a), we find for $t \in (t_0, t_1]$ that

$$W'_g(t) = h(t) \frac{k \sqrt{1+t}}{\pi \sqrt{1-t}} \geq \frac{h(t_1)k}{\pi \sqrt{2}} \sqrt{1+t}.$$

Integrating over the interval $[t_0, t_1] = [-1, t_1]$ gives

$$1 \geq \frac{h(t_1)k\sqrt{2}}{3\pi} (1+t_1)^{3/2} = \frac{g(t_1)k\sqrt{2}}{3\pi} (1+t_1)^{1/2},$$

which implies part (a). By Lemma B.1(a) and Lemma B.7, there holds for $t \in [t_{k-1}, t_k] \subset [0, 1]$,

$$W'_g(t) \geq \frac{k}{\pi} \frac{h(1)}{\sqrt{1-t}},$$

and by integrating over the interval $[t_{k-1}, t_k] = [t_{k-1}, 1]$ we get

$$1 \geq \frac{2kh(1)}{\pi} (1-t_{k-1})^{1/2} = \frac{kg(1)}{\pi} (1-t_{k-1})^{1/2},$$

as required for part (b). For a proof of (c), we observe that, by Lemma B.3(a),

$$\begin{aligned} t_{j+1} - t_j &= 2 \sin\left(\frac{\alpha_{j+1} - \alpha_j}{2}\right) \sin\left(\frac{\alpha_{j+1} + \alpha_j}{2}\right) \\ &\leq 2 \sin\left(\frac{\alpha_1 - \alpha_0}{2}\right) = \sqrt{2(t_1 - t_0)}. \end{aligned}$$

By concavity and positivity of g and Lemma B.1(d),

$$g(t_1) \geq g(t_0) \frac{1-t_1}{2} + g(1) \frac{t_1-t_0}{2} \geq \frac{t_1-t_0}{2}.$$

Multiplying with $\sqrt{t_1-t_0}$ and applying part (a) we arrive at

$$(t_1 - t_0)^{3/2} \leq 2g(t_1)\sqrt{t_1 - t_0} \leq \frac{3\pi\sqrt{2}}{k},$$

which yields part (c). □

Lemma B.9 For every $j \in \{0, \dots, k-2\}$

$$\frac{2}{3\pi} \leq \frac{t_{j+1} - t_j}{t_{j+2} - t_{j+1}} \leq \frac{6c_2\sqrt{c_5}}{c_1}.$$

Proof: In order to show the left-hand inequality, we write

$$\frac{t_{j+2} - t_{j+1}}{t_{j+1} - t_j} = \frac{\sin\left(\frac{\alpha_{j+1} + \alpha_{j+2}}{2}\right) \sin\left(\frac{\alpha_{j+2} - \alpha_{j+1}}{2}\right)}{\sin\left(\frac{\alpha_j + \alpha_{j+1}}{2}\right) \sin\left(\frac{\alpha_{j+1} - \alpha_j}{2}\right)} \leq \frac{\sin\left(\frac{\alpha_{j+1} + \alpha_{j+2}}{2}\right)}{\sin\left(\frac{\alpha_j + \alpha_{j+1}}{2}\right)},$$

where we have applied Lemma B.3(a). We claim that the right-hand term is $\leq 3\pi/2$. Indeed, if $(\alpha_j + \alpha_{j+1})/2 \geq \pi/2$, then this quotient is less than one. Else, by using Lemma B.2, we obtain

$$\frac{\sin(\frac{\alpha_{j+1} + \alpha_{j+2}}{2})}{\sin(\frac{\alpha_j + \alpha_{j+1}}{2})} \leq \frac{\pi}{2} \frac{\alpha_{j+1} + \alpha_{j+2}}{\alpha_j + \alpha_{j+1}},$$

and from Lemma B.3(b) we get that $\alpha_{j+1} + \alpha_{j+2} = \alpha_{j+2} - \alpha_{j+1} + 2\alpha_{j+1} \leq 3\alpha_{j+1} \leq 3(\alpha_j + \alpha_{j+1})$.

To prove the right-hand inequality in Lemma B.9 we use Theorem 2.6 and Lemma B.1(a) in order to obtain

$$\begin{aligned} \frac{t_{j+1} - t_j}{t_{j+2} - t_{j+1}} &\leq \frac{c_2}{c_1} \frac{W'_g(\frac{t_{j+1} + t_{j+2}}{2})}{W'_g(\frac{t_j + t_{j+1}}{2})} \\ &\leq \frac{c_2}{c_1} \frac{h(\frac{t_{j+1} + t_{j+2}}{2})(1 + \frac{t_{j+1} + t_{j+2}}{2})}{h(\frac{t_j + t_{j+1}}{2})(1 + \frac{t_j + t_{j+1}}{2})} \frac{W'(\frac{t_{j+1} + t_{j+2}}{2})}{W'(\frac{t_j + t_{j+1}}{2})} \\ &\leq \frac{c_2}{c_1} \sqrt{\frac{1 + \frac{t_{j+1} + t_{j+2}}{2}}{1 + \frac{t_j + t_{j+1}}{2}} \frac{1 - \frac{t_j + t_{j+1}}{2}}{1 - \frac{t_{j+1} + t_{j+2}}{2}}}. \end{aligned}$$

With help of Lemma B.4 and Lemma B.5 we obtain

$$\frac{t_{j+1} - t_j}{t_{j+2} - t_{j+1}} \leq \frac{c_2}{c_1} 3\sqrt{c_5} \sqrt{\underbrace{\frac{1 + t_{j+1}}{1 + \frac{t_j + t_{j+1}}{2}}}_{\leq 2} \underbrace{\frac{1 - t_{j+1}}{1 - \frac{t_{j+1} + t_{j+2}}{2}}}_{\leq 2}} \leq \frac{6c_2\sqrt{c_5}}{c_1}$$

□