Available online at www.sciencedirect.com

ScienceDirect

Journal of Approximation Theory 206 (2016) 17–47

Journal of Approximation Theory

www.elsevier.com/locate/jat

Full length article

Minimax principle and lower bounds in H^2-rational approximation[☆]

Laurent Baratchart[a], Sylvain Chevillard[a,*], Tao Qian[b]

[a] Inria, 2004 route des Lucioles, BP 93, 06 902 Sophia-Antipolis Cedex, France
[b] Faculty of Science and Technology, University of Macau, E11 Avenida da Universidade, Taipa, Macau

Received 30 December 2013; received in revised form 9 March 2015; accepted 14 March 2015
Available online 1 April 2015

Communicated by Sergey Krushchev

Abstract

We derive lower bounds in rational approximation of given degree to functions in the Hardy space H^2 of the unit disk. We apply these to asymptotic error rates in rational approximation to Blaschke products and to Cauchy integrals on geodesic arcs. We also explain how to compute such bounds, either using Adamjan–Arov–Krein theory or linearized errors, and we present a couple of numerical experiments. We dwell on a maximin principle developed in Baratchart and Seyfert (2002).

MSC: 31B05; 35J25; 42B35; 46E20; 47B35

Keywords: Complex rational approximation; Hardy spaces; Lower bounds; Error rates

1. Introduction

Rational approximation to a given function on a curve in the complex plane is a classical topic from analysis, and a cornerstone of modeling and design in several areas of applied sciences and

☆ This work has been partly funded by Macau Government FDCT 098/2012/A3.
* Corresponding author.
E-mail addresses: Laurent.Baratchart@sophia.inria.fr (L. Baratchart), sylvain.chevillard@inria.fr (S. Chevillard), fsttq@umac.mo (T. Qian).

http://dx.doi.org/10.1016/j.jat.2015.03.004

engineering. Special interest attaches to the case where the approximated function extends holomorphically on one side of the curve. In connection with system identification and control, such issues typically arise on the line or the circle where they make contact with extremal problems in Hardy spaces [4,15,33,29,30,32,43]. Our model curve in this paper will be the circle, though everything translates easily to the line. The criterion under examination will be the L^2-norm.

From the approximation-theoretic viewpoint, much attention has been directed towards error rates, in connection with smoothness of the approximated function. Let us mention Peller's converse theorems on the speed of rational approximation [33], Glover's construction of near-best uniform rational approximants [15], Parfenov's solution of a conjecture by Gonchar on the degree of rational approximation to holomorphic functions on compact subsets of the domain of analyticity [31], the Gonchar–Rakhmanov estimates in uniform rational approximation to sectionally holomorphic functions off an S-contour, and its generalization to best L^2 and L^p approximants in [8,42].

The present paper is, in part, a sequel to [8]. In the latter reference best L^2 and L^∞ rational approximants are compared in the n-th root sense, whereas here we compare them in norm. We emphasize that the L^2 norm and weighted variants thereof are of great importance in applications, due to their interpretation as a variance in a stochastic context. Moreover, best rational H^2 approximants have the interesting property of being attained through interpolation [26]. Note also that certain functions, like Blaschke products, can be approximated in H^2-norm but not in the uniform norm by rational functions.

A key to the above-mentioned comparison is the derivation of lower bounds on the L^2 approximation error. Lower bounds in approximation are usually difficult to obtain; we dwell here on a topological machinery developed in [6] which expresses the approximation error as the solution to a max–min problem, and we rely as well on the Adamjan–Arov–Krein theory of best uniform meromorphic approximation. We prove a somewhat general result (Theorem 4) which gives a lower bound on the L^2-best rational approximation error of given degree, in terms of the ratios of L^2 and L^∞ norms of the singular vectors of the Hankel operator with symbol the approximated function. We then apply it to three cases where these ratios can be estimated: rational functions, Blaschke products, and Cauchy integrals on geodesic arcs. We use also the max–min principle to study linearized errors as a means to compute further lower bounds. We also include numerical experiments, some of which give excellent accuracy to estimate the H^2 error in rational approximation (see Table 1 in Section 7). To the author's knowledge, such results are first of their kind.

The paper is organized as follows. After some preliminaries on Hardy spaces in Section 2, we present in Section 3 the approximation problems that we consider. Section 4 is an introduction to the results of [6] and it contains a basic account of the Adamjan–Arov–Krein theory. We derive in Section 5 our main theorem giving lower bounds in L^2 rational approximation, and we apply it to cases mentioned above. Finally, in Section 6, we discuss linearized errors.

2. Notations and preliminaries

Let $\mathbb{D}$ be the unit disk in the complex plane $\mathbb{C}$, and $\mathbb{T}$ the unit circle. We denote by $C(\mathbb{T})$ the space of continuous, complex-valued functions on $\mathbb{T}$. For $1 \le p \le \infty$, we put $L^p = L^p(\mathbb{T})$ for the familiar Lebesgue space of complex measurable functions on $\mathbb{T}$ such that

$$\|f\|_p = \left(\frac{1}{2\pi}\int_0^{2\pi} |f(e^{i\theta})|^p \, d\theta\right)^{1/p} < \infty \quad \text{if } 1 \le p < \infty,$$

$$\|f\|_\infty = \text{ess.} \sup_{\theta\in[0,2\pi]} |f(e^{i\theta})| < \infty.$$

Hereafter, we let $H^2 = H^2(\mathbb{D})$ be the Hardy space of holomorphic functions in $\mathbb{D}$ whose Taylor coefficients at 0 are square summable:

$$H^2 = \{f(z) = \Sigma_{k=0}^\infty a_k z^k : \ \|f\|_{H^2} := \Sigma_{k=0}^\infty |a_k|^2 < +\infty\}.$$

We refer the reader to [14] for standard facts on Hardy spaces. By Parseval's relation

$$\|f\|_{H^2}^2 = \sup_{0\le r<1} \frac{1}{2\pi}\int_0^{2\pi} |f(re^{i\theta})|^2\, d\theta, \tag{1}$$

and the map

$$\left(f(z) = \Sigma_{k=0}^\infty a_k z^k\right) \longrightarrow \left(f^*(e^{i\theta}) := \Sigma_{k=0}^\infty a_k e^{ik\theta}\right)$$

is an isometry from H^2 onto the closed subspace of L^2 comprised of functions whose Fourier coefficients of strictly negative index do vanish. As is customary, we shall identify H^2 with this subspace so that the distinction between f and f^* as well as $\|f\|_{H^2}$ and $\|f^*\|_2$ will disappear. This conveniently allows one to regard members of the Hardy class both as functions on $\mathbb{D}$ and on $\mathbb{T}$. From the function-theoretic viewpoint, the correspondence $f \mapsto f^*$ is that $f^*(e^{i\theta})$ is almost everywhere the limit of $f(z)$ as z tends non-tangentially to $e^{i\theta}$ within $\mathbb{D}$.

We put $\bar{H}^{2,0} = \bar{H}^{2,0}(\mathbb{C}\setminus\overline{\mathbb{D}})$ for the companion Hardy space of holomorphic functions in $\mathbb{C}\setminus\overline{\mathbb{D}}$, vanishing at infinity, whose Taylor coefficients there are square summable:

$$\bar{H}^{2,0} = \{f(z) = \Sigma_{k=1}^\infty a_k z^{-k} : \ \|f\|_{\bar{H}^{2,0}} := \Sigma_{k=1}^\infty |a_k|^2 < +\infty\}.$$

The map

$$\left(f(z) = \Sigma_{k=1}^\infty a_k z^{-k}\right) \longrightarrow \left(f^*(e^{i\theta}) = \Sigma_{k=1}^\infty a_k e^{-ik\theta}\right)$$

is an isometry from $\bar{H}^{2,0}$ onto the closed subspace of L^2 comprised of functions whose Fourier coefficients of non-negative index do vanish, and as before we identify $\bar{H}^{2,0}$ with the latter. Clearly we have an orthogonal sum:

$$L^2 = H^2 \oplus \bar{H}^{2,0}. \tag{2}$$

In fact, it holds that $f \in \bar{H}^{2,0}$ if and only if the function $\check{f}$ given by

$$\check{f}(z) := z^{-1}\overline{f(1/\bar{z})} \tag{3}$$

lies in H^2, and the map $f \mapsto \check{f}$ is an involutive isometry of L^2 sending H^2 onto $\bar{H}^{2,0}$. Actually, $\check{f}$ has same modulus as f *pointwise* on $\mathbb{T}$ since $\overline{f(1/\bar{z})} = \overline{f(z)}$ when $|z| = 1$. If f is holomorphic on Ω, then $f^\sharp(z) = \overline{f(1/\bar{z})}$ is holomorphic on the reflection of Ω across $\mathbb{T}$, and if f is rational $f^\sharp$ is likewise rational. Of course, a relation like $f^\sharp = \bar{f}$ must be understood to hold on $\mathbb{T}$ only.

We let

$$\mathbf{P}_+\left(\Sigma_{k\in\mathbb{Z}}\, a_k e^{ik\theta}\right) = \Sigma_{k\ge 0}\, a_k e^{ik\theta} \quad \text{and} \quad \mathbf{P}_-\left(\Sigma_{k\in\mathbb{Z}} a_k e^{ik\theta}\right) = \Sigma_{k<0} a_k e^{ik\theta}$$

indicate the so-called Riesz projections that discard the Fourier coefficients of strictly negative and non-negative index respectively. Clearly $\mathbf{P}_+$ (resp. $\mathbf{P}_-$) contractively maps L^2 onto H^2 (resp.

$\bar{H}^{2,0}$) and $\mathbf{P}_+ + \mathbf{P}_- = I$. We call $\mathbf{P}_+$ the *analytic projection* and $\mathbf{P}_-$ the *anti-analytic projection*. Note that, by Cauchy's formula, $\mathbf{P}_\pm(f)$ can be expressed as Cauchy integrals:

$$\begin{aligned} \mathbf{P}_+(f)(z) &= \frac{1}{2i\pi}\int_{\mathbb{T}} \frac{f(\zeta)}{\zeta - z} d\zeta, \quad |z| < 1, \\ \mathbf{P}_-(f)(z) &= \frac{1}{2i\pi}\int_{\mathbb{T}} \frac{f(\zeta)}{z - \zeta} d\zeta, \quad |z| > 1. \end{aligned} \tag{4}$$

The Hardy space $H^\infty = H^\infty(\mathbb{D})$ consists of bounded holomorphic functions on $\mathbb{D}$, endowed with the *sup* norm. From (1) we see that H^∞ embeds contractively in H^2, in particular each $f \in H^\infty$ has a non-tangential limit f^* on $\mathbb{T}$. It can be shown that $\|f^*\|_\infty = \|f\|_{H^\infty}$, and that the map $f \mapsto f^*$ is an isometry from H^∞ onto the closed subspace of L^∞ comprised of functions whose Fourier coefficients of strictly negative index do vanish. Again we identify H^∞ with this subspace. Likewise, the space $\bar{H}^{\infty,0}$ of bounded holomorphic functions vanishing at infinity in $\mathbb{C}\setminus\overline{\mathbb{D}}$ identifies *via* non-tangential limits with the closed subspace of L^∞ consisting of functions whose Fourier coefficients of non-negative index do vanish. However, in contrast with the situation for L^2, the operators $\mathbf{P}_\pm$ are unbounded on L^∞. Besides the norm topology, H^∞ inherits the weak-* topology from $L^\infty(\mathbb{T})$. It is characterized by the fact that f_n tends weak-* to f if and only if $\int_{\mathbb{T}} f_n\varphi \to \int_{\mathbb{T}} f\varphi$ for every $\varphi \in L^1$. It is equivalent to require that $(\|f_n\|_\infty)_n$ is a bounded sequence and that, for each k, the k-th Fourier coefficient of f_n converges to the k-th Fourier coefficient of f.

As is well-known [14, Chapter II, Corollary 5.7], a nonzero $f \in H^2$ factors uniquely as $f = jw$ where

$$w(z) = \exp\left\{\frac{1}{2\pi}\int_0^{2\pi} \frac{e^{i\theta}+z}{e^{i\theta}-z} \log|f(e^{i\theta})|\, d\theta\right\} \tag{5}$$

belongs to H^2 and is called the *outer factor* of f, normalized so as to be positive at zero, while $j \in H^\infty$ has modulus 1 a.e. on $\mathbb{T}$ and is called the *inner factor* of f. The latter may be further decomposed as $j = bS$, where

$$b(z) = cz^k \prod_{\zeta_l \neq 0} \frac{-\bar{\zeta}_l}{|\zeta_l|} \frac{z - \zeta_l}{1 - \bar{\zeta}_l z} \tag{6}$$

is the normalized *Blaschke product*, with multiplicity $k \geq 0$ at the origin, associated to a sequence of points $\zeta_l \in \mathbb{D}\setminus\{0\}$ and to a constant $c \in \mathbb{T}$, while

$$S(z) = \exp\left\{-\frac{1}{2\pi}\int_0^{2\pi} \frac{e^{i\theta}+z}{e^{i\theta}-z}\, d\mu(\theta)\right\}$$

is the *singular inner factor* associated with a positive singular measure μ on $\mathbb{T}$. The ζ_l are of course the zeros of f in $\mathbb{D}$, counting multiplicities by repetition. The number of zeros, finite or infinite, is called the degree of the Blaschke product. Throughout, we let B_n denote the set of Blaschke products of degree at most n. If the degree is infinite, the convergence of the product in (6) is equivalent to the condition

$$\sum_l (1 - |\zeta_l|) < \infty \tag{7}$$

which holds automatically when $f \in H^2$. That $w(z)$ is well-defined rests on the fact that log $|f| \in L^1$ if $f \in H^2\backslash\{0\}$. A function $f \in H^2$ with inner–outer factorization $f = jw$ lies in H^∞ if, and only if $w \in L^\infty(\mathbb{T})$. For simplicity, we often say that a function is outer (resp. inner) if it is equal to its outer (resp. inner) factor.

We put $\mathcal{P}_n[z]$ for the space of complex algebraic polynomials of degree at most n in the variable z, or simply $\mathcal{P}_n$ if the variable is understood. Below we let $\mathcal{Z}(q)$ indicate the set of zeros of a polynomial q. For $q_n \in \mathcal{P}_n[z]$, we define its *reciprocal polynomial* to be

$$\widetilde{q}_n(z) := z^n\, \overline{q_n(1/\bar{z})}.$$

We warn the reader that this definition depends on n: if we consider $q_{n-1} \in \mathcal{P}_{n-1}$ as an element of $\mathcal{P}_n$ with zero leading coefficient, the definitions of $\widetilde{q}_{n-1}(z)$ in $\mathcal{P}_{n-1}$ and in $\mathcal{P}_n$ may be inconsistent. Therefore we always specify, *e.g.* via a subscript "n" as in "q_n", which definition is used. Clearly the "tilde" operation is an involution of $\mathcal{P}_n$ preserving modulus pointwise on $\mathbb{T}$.

We designate by $\mathcal{R}_{m,n} = \mathcal{R}_{m,n}(z)$ the set of complex rational functions of type (m, n) in L^2, namely those that can be written as p_m/q_n where p_m belongs to $\mathcal{P}_m$ and $q_n \in \mathcal{P}_n$ has no root on $\mathbb{T}$. When $r = p_m/q_n$ is in irreducible form, the integer $\max\{m, n\}$ is the (exact) degree of r. Note that $B_m \subset \mathcal{R}_{m,m}$ is comprised of rational functions of degree at most m which are analytic in $\mathbb{D}$ and have unit modulus everywhere on $\mathbb{T}$. Alternatively, B_m consists of functions $q_m/\widetilde{q}_m$ where $q_m \in \mathcal{P}_m$ has all its roots in $\mathbb{D}$. Clearly, B_m is included in the unit sphere of both H^2 and H^∞.

We further set

$$H^2_m := \left\{ \frac{g}{q_m} \,:\, g \in H^2,\ q_m \in \mathcal{P}_m \right\}.$$

Members of H^2_m identify in L^2 with non-tangential limits of meromorphic functions with at most m poles in $\mathbb{D}$ (counting multiplicities) whose L^2-means over $\{|z| = r\}$ remain eventually bounded as $r \to 1^-$. Functions in $\cup_m H^2_m$ are called meromorphic in L^2. Two equivalent descriptions of H^2_m are useful: on the one hand we get by pole-residue decomposition that $H^2_m = H^2 + (\mathcal{R}_{m-1,m} \cap \bar{H}^{2,0})$, on the other hand we have that $H^2_m = B_m^{-1} H^2$, the set of quotients of H^2-functions by Blaschke products of degree at most m. Likewise we put

$$H^\infty_m := H^2_m \cap L^\infty = B_m^{-1} H^\infty = \left\{ \frac{g}{q_m} \,:\, g \in H^\infty,\ q_m \in \mathcal{P}_m \right\}$$

for the set of meromorphic functions with at most m poles in L^∞.

3. Best rational and meromorphic approximation in L^2

For $n \geq 1$ an integer, the best rational approximation problem of degree n in L^2 is:

Problem R(n): *Given $h \in L^2$, to find $r^* \in \mathcal{R}_{n,n}$ such that*

$$\|h - r^*\|_2 = \min_{r \in \mathcal{R}_{n,n}} \|h - r\|_2.$$

Write $h = h_1 + h_2$ with $h_1 \in H^2$, $h_2 \in \bar{H}^{2,0}$. By partial fraction expansion, each $r \in \mathcal{R}_{n,n}$ can be decomposed as $r_1 + r_2$ where $r_1 \in H^2$, $r_2 \in \bar{H}^{2,0}$, and $\deg r_1 + \deg r_2 \leq n$. Then, by (2),

$$\|h - r\|_2^2 = \|h_1 - r_1\|_2^2 + \|h_2 - r_2\|_2^2$$

so that problem R(n) reduces, modulo optimal allocation of the degrees of r_1 and r_2 ($n + 1$ choices), to a pair of problems of the following types:

Problem RA(n): *Given* $f \in H^2$*, to find* $r^* \in \mathcal{R}_{n,n} \cap H^2$ *such that*

$$\|f - r^*\|_2 = \min_{r \in \mathcal{R}_{n,n} \cap H^2} \|f - r\|_2.$$

Problem RAB(n): *Given* $f \in \bar{H}^{2,0}$*, to find* $r^* \in \mathcal{R}_{n-1,n} \cap \bar{H}^{2,0}$ *such that*

$$\|f - r^*\|_2 = \min_{r \in \mathcal{R}_{n-1,n} \cap \bar{H}^{2,0}} \|f - r\|_2.$$

In "RA(n)" and "RAB(n)", the letter "A" is mnemonic for "analytic" and "B" stands for "bar".

Problem RA(n) is in fact equivalent to RAB(n). For we can parametrize $r \in \mathcal{R}_{n,n} \cap H^2$ as $r(0) + zr_3$ where $r(0) \in \mathbb{R}$ and $r_3 \in \mathcal{R}_{n-1,n} \cap H^2$ vary independently, and by Parseval's theorem

$$\|f - r\|_2^2 = |f(0) - r(0)|^2 + \|(f - f(0)) - zr_3\|_2^2$$

hence $r(0) = f(0)$ is the optimal choice. Thus, since multiplication by $1/z$ is an isometry, we find upon replacing f by $(f - f(0))/z$ that Problem RA(n) is equivalent to the normalized version:

Problem RAN(n): *Given* $f \in H^2$*, to find* $r^* \in \mathcal{R}_{n-1,n} \cap H^2$ *such that*

$$\|f - r^*\|_2 = \min_{r \in \mathcal{R}_{n-1,n} \cap H^2} \|f - r\|_2.$$

Now, applying the check operation defined in (3), which preserves $\mathcal{R}_{n-1,n}$ and the degree, this last problem is seen to be equivalent to RAB(n), as announced. Note that when passing from RA(n) to RAB(n), the initial $f \in H^2$ to be approximated from $\mathcal{R}_{n,n} \cap H^2$ gets transformed into the function $\overline{f(1/\bar{z})} - \overline{f(0)} \in \bar{H}^{2,0}$ to be approximated from $\mathcal{R}_{n-1,n} \cap \bar{H}^{2,0}$.

Finally, we state the best meromorphic approximation problem with at most n poles in L^2:

Problem MA(n): *Given* $f \in L^2$*, to find* $g^* \in H^2_n$ *such that*

$$\|f - g^*\|_2 = \min_{g \in H^2_n} \|f - g\|_2.$$

Problem MA(n) is also equivalent to RAB(n). Indeed, $H^2_n = H^2 + (\mathcal{R}_{n-1,n} \cap \bar{H}^{2,0})$ so that, by orthogonality of H^2 and $\bar{H}^{2,0}$, the H^2-component of a minimizer in MA(n) must be $\mathbf{P}_+(f)$ while the $\bar{H}^{2,0}$-component of this minimizer is a solution to RAB(n) with f replaced by $\mathbf{P}_-(f)$.

Let us mention that best meromorphic approximation, unlike best rational approximation, is conformally invariant. This makes it of independent interest in a broader context, see [5, Proposition 5.4] for further details.

Having reduced all previous approximation problems to RAB(n), hereafter we discuss the latter. It is known that RAB(n) has a solution which needs not be unique, and every solution has exact degree n unless f is rational of degree at most $n - 1$ [12,26,2].

We shall write $d_2(f, \mathcal{R}_{n-1,n})$ (resp. $d_2(f, \mathcal{R}_{n,n})$) for the distance from f to $\mathcal{R}_{n-1,n}$ (resp. $\mathcal{R}_{n,n}$) in L^2. For instance if $f \in \bar{H}^{2,0}$, then $d_2(f, \mathcal{R}_{n-1,n})$ is both the value of Problem RAB(n) and of Problem MA(n); and if $f \in H^2$, then $d_2(f, \mathcal{R}_{n-1,n})$ (resp. $d_2(f, \mathcal{R}_{n,n})$) is the value of problem RAN(n) (resp. RA(n)). Besides, the value of MA(n) is denoted by $d_2(f, H^2_n)$.

When $f \in L^\infty$, we let $d_\infty(f, H^\infty_n)$ indicate the distance from f to H^∞_n. This is the value of the best meromorphic approximation problem with at most n poles in L^∞, that we did not formally introduce but which stands analog to MA(n) with L^2 replaced by L^∞ and H^2_n by H^∞_n. We put also $d_\infty(f, \mathcal{R}_{n-1,n})$ (resp. $d_\infty(f, \mathcal{R}_{n,n})$) for the distance from f to $\mathcal{R}_{n-1,n}$ (resp. $\mathcal{R}_{n,n}$) in L^∞.

4. Duality in meromorphic approximation

Pick $f \in \bar{H}^{2,0}$ and let us parametrize $r \in \mathcal{R}_{n-1,n} \cap \bar{H}^{2,0}$ as $r = p_{n-1}/q_n$ where p_{n-1} ranges over P_{n-1} and q_n ranges over those polynomials in $\mathcal{P}_n$ whose roots lie in $\mathbb{D}$. Then $q_n/\widetilde{q}_n \in B_n$ and since $p_{n-1}/\widetilde{q}_n \in H^2$ we have by orthogonality of H^2 and $\bar{H}^{2,0}$ that

$$\left\| f - \frac{p_{n-1}}{q_n} \right\|_2^2 = \left\| f\frac{q_n}{\widetilde{q}_n} - \frac{p_{n-1}}{\widetilde{q}_n} \right\|_2^2 = \left\| \mathbf{P}_-\left(f\frac{q_n}{\widetilde{q}_n} \right) \right\|_2^2 + \left\| \mathbf{P}_+\left(f\frac{q_n}{\widetilde{q}_n} \right) - \frac{p_{n-1}}{\widetilde{q}_n} \right\|_2^2. \tag{8}$$

Clearly the product of a $\bar{H}^{2,0}$-function by a polynomial in $\mathcal{P}_n$ yields a member of $z^n\bar{H}^{2,0}$. Therefore

$$\widetilde{q}_n \mathbf{P}_+\left(f\frac{q_n}{\widetilde{q}_n} \right) = f q_n - \widetilde{q}_n \mathbf{P}_-\left(f\frac{q_n}{\widetilde{q}_n} \right) \in z^n \bar{H}^{2,0} \cap H^2 = \mathcal{P}_{n-1}, \tag{9}$$

entailing that $p_{n-1} = \widetilde{q}_n \mathbf{P}_+(f q_n/\widetilde{q}_n)$ is the minimizing choice in (8) for fixed q_n. Consequently

$$\min_{r\in\mathcal{R}_{n-1,n}\cap\bar{H}^{2,0}} \|f - r\|_2 = \min_{q_n\in\mathcal{P}_n,\mathcal{Z}(q_n)\subset\mathbb{D}} \left\| \mathbf{P}_-\left(f\frac{q_n}{\widetilde{q}_n} \right) \right\|_2 = \min_{b_n\in B_n} \|\mathbf{P}_-(f b_n)\|_2. \tag{10}$$

That the infimum is indeed attained in the right hand side of (10) follows from (8) and the fact that RAB(n) has a solution. Define A_f, the *Hankel operator with symbol* f, by

$$\begin{aligned} A_f &: H^\infty \longrightarrow \bar{H}^{2,0} \\ v &\mapsto \mathbf{P}_-(f v). \end{aligned} \tag{11}$$

It is evident that A_f is continuous and that $|||A_f||| = \|f\|_2$, a unit maximizing vector being $v \equiv 1$. Here and below, we let $|||.|||$ stand for the operator norm, and a maximizing vector of an operator E is a nonzero vector v such that $\|Ev\|/\|v\| = |||E|||$.

The content of the discussion leading from (8) to (10) may now be restated as follows.

Proposition 1. *For $f \in \bar{H}^{2,0}$, it holds that*

$$d_2(f, \mathcal{R}_{n-1,n}) = \min_{b_n\in B_n} \|A_f(b_n)\|_2. \tag{12}$$

A rational function $p_{n-1}/q_n \in \mathcal{R}_{n-1,n}$ is a solution to RAB(n) if, and only if $b_n = q_n/\widetilde{q}_n$ is a minimizing Blaschke product in (12) and $p_{n-1} = \widetilde{q}_n \mathbf{P}_+(f b_n)$.

Put $\mathcal{L}_k$ for the space of linear operators from H^∞ into $\bar{H}^{2,0}$ which are weak-* continuous and have rank not exceeding k. For $k = 0, 1, 2, \ldots$, we denote by $\sigma_k(A_f)$ the k-th *approximation number* of A_f defined by

$$\sigma_k(A_f) = \inf\left\{ |||A_f - \Gamma|||,\ \Gamma \in \mathcal{L}_k. \right\}. \tag{13}$$

Note that $\sigma_k(A_f) \geq \sigma_{k+1}(A_f)$ and that $\sigma_0(A_f) = |||A_f|||$.

We need also introduce the *genus* of a closed symmetric subset K in a topological vector space; here, *symmetric* means that if $v \in K$ then also $-v \in K$. By definition the genus of K, denoted by $\mathbf{gen}(K)$, is the smallest positive integer m for which there exists an odd continuous mapping

$$G : K \longrightarrow \mathbb{R}^m \backslash \{0\}, \tag{14}$$

or else $+\infty$ if no finite m meets the above requirement. By convention the genus is zero if $K = \emptyset$. When K is compact and does not contain 0, then $\mathbf{gen}(K)$ is always finite, see [44]. For instance, if $m \geq 1$, the classical Borsuk–Ulam theorem from topology [20, Chapter 2, Section 6] implies that any symmetric set in $\mathbb{R}^m$ which is homeomorphic to the (real) $(m-1)$-dimensional Euclidean sphere $\mathbb{S}^{m-1}$ through an odd map has genus m.

Below, we shall be concerned with weak-* compact subsets of $\mathcal{S}^\infty$, the unit sphere of H^∞. In this connection, we let

$$\mathcal{K}_m^\infty = \left\{ K \subset \mathcal{S}^\infty : K \text{ is a weak-* compact symmetric subset of } \mathcal{S}^\infty \text{with } \mathbf{gen}(K) \geq m \right\}.$$

Subsequently, we define the (generalized) *singular numbers* of A_f by

$$\lambda_m(A_f) = \max_{K \in \mathcal{K}_m^\infty} \min_{u \in K} \|A_f(u)\|_2, \quad m = 0, 1, 2, \ldots. \tag{15}$$

The following theorem, which was established in [6], connects approximation numbers and singular numbers of A_f with the value of Problem RAB(n):

Theorem 1 (*[6, Theorem 8.1]*). *Let $f \in \bar{H}^{2,0}$ and $A_f : H^\infty \to \bar{H}^{2,0}$ the Hankel operator with symbol f. For each integer $n \geq 0$, the following equalities hold:*

$$d_2(f, \mathcal{R}_{n-1,n}) = \sigma_n(A_f) = \lambda_{2n+1}(A_f) = \lambda_{2n+2}(A_f). \tag{16}$$

Theorem 1 is reminiscent of a famous theorem by Adamjan–Arov–Krein (in short: the AAK theorem) characterizing $d_\infty(f, H_n^\infty)$ rather than $d_2(f, \mathcal{R}_{n-1,n})$. To state the result, let us define for $f \in L^\infty$ the Hankel operator Γ_f by

$$\begin{aligned} \Gamma_f &: H^2 \longrightarrow \bar{H}^{2,0} \\ v &\mapsto \mathbf{P}_-(fv). \end{aligned} \tag{17}$$

Although the definitions of A_f and Γ_f are formally the same, observe that the domains in (11) and (17) are different. The definition of $s_k(\Gamma_f)$ is still given by (13) except that A_f is replaced by Γ_f and Γ now ranges over linear operators from H^2 into $\bar{H}^{2,0}$ having rank at most k. If in addition f is continuous on $\mathbb{T}$, then Γ_f is compact [33, Chapter 1, Theorem 5.5]. Then, if we let Γ_f^* denote the adjoint, $\Gamma_f^*\Gamma_f$ is a compact selfadjoint operator from the Hilbert space H^2 into itself and as such it has a complete orthonormal family of eigenvectors called the singular vectors of Γ_f; the associated eigenvalues are none but the squared approximation numbers of Γ_f [16, Chapter II, Theorem 2.1], and there holds the Courant max min principle [45, Section 22.11a]:

$$s_n(\Gamma_f) = \max_{V \in \mathcal{V}_{n+1}} \min_{\substack{v \in V \\ \|v\|_2 = 1}} \|\Gamma_f(v)\|_2, \tag{18}$$

where $\mathcal{V}_{n+1}$ is the collection of linear subspaces of H^2 of complex dimension at least $n+1$. In this Hilbertian context, the approximation number $s_n(\Gamma_f)$ is also called the n-th singular value of Γ_f. We say that a function v is *associated with a singular value s* when v is an eigenvector of $\Gamma_f^*\Gamma_f$ associated with the eigenvalue s^2: $v = s^2\Gamma_f^*\Gamma_f(v)$. As a particular case of Eq. (18) a maximizing vector is just a singular vector associated with $s_0(\Gamma_f)$.

Theorem 2 (*The AAK Theorem [1, Theorems 0.1 and 0.2] [33, Chapter 4, Theorem 1.2]*). *Let $f \in L^\infty$ and $\Gamma_f : H^2 \to \bar{H}^{2,0}$ be the Hankel operator with symbol f. For each integer $n \geq 0$,* *it holds that*

$$d_\infty(f, H_n^\infty) = s_n(\Gamma_f). \tag{19}$$

If in addition $f \in C(\mathbb{T})$*, then* Γ_f *is compact and the quantity* (19) *is also equal to* (18).

The case $n = 0$ of Theorem 2, *i.e.* that $|||\Gamma_f||| = d_\infty(f, H^\infty)$ was known earlier as Nehari's theorem.

If we compare (15) and (16) with (18) and (19) for $f \in \bar{H}^{2,0} \cap L^\infty$, we see that the main difference between best meromorphic approximation with at most n poles in L^2 and in L^∞ lies with the maximization step in (15), which in the L^2-case must be taken over all compact sets of genus at least[1] $2n + 2$ and not just Euclidean spheres of real dimension $2n + 1$.

It follows from [3, Theorem 1] or [23, Theorem 5.3] that B_n is homeomorphic to $\mathbb{S}^{2n+1}$ and inspection of the proof reveals that the homeomorphism is odd. Moreover B_n is weak-* compact in $\mathcal{S}^\infty$ [6, Lemma 7.3], therefore $B_n \in \mathcal{K}_{2n+2}^\infty$ and from Proposition 1 we see that it is a supremizer in (15).

We mention for completeness a companion to Theorem 1 dealing with min max (not max min):

Theorem 3 (*[34] [6, Eq. (78)]*)**.** *Let* $f \in \bar{H}^{2,0}$ *and* $A_f : H^\infty \to \bar{H}^{2,0}$ *be the Hankel operator with symbol* f*. For each integer* $n \geq 0$*, the following equality hold:*

$$d_2(f, \mathcal{R}_{n-1,n}) = \min_{W \in \mathcal{W}_n} \max_{\substack{w \in W \\ \|w\|_\infty = 1}} \|A_f(v)\|_2, \tag{20}$$

where $\mathcal{W}_n$ *is the collection of linear subspaces in* H^∞ *of (complex) codimension at most* n*.*

Note that (20) is the exact counterpart for A_f of the standard Courant min max principle for Γ_f:

$$d_\infty(f, H_n^\infty) = \min_{X \in \mathcal{X}_n} \max_{\substack{w \in X \\ \|w\|_2 = 1}} \|\Gamma_f(v)\|_2,$$

where $\mathcal{X}_n$ is the collection of linear subspaces in H^2 of (complex) codimension at most n.

Using Proposition 1 it is easy to see that if p_{n-1}/q_n is a solution to RAB(n), then the subspace $(q_n/\widetilde{q}_n)H^\infty$, comprised of multiples of $q_n/\widetilde{q}_n$ in H^∞, is a minimizing W in (20).

In the rest of the paper, we use the maximizing step in (15) together with Theorem 1 to derive lower bounds for Problems RAB(n).

5. Lower bounds

5.1. Comparing L^2 and L^∞ meromorphic approximation

Consider $f \in \bar{H}^{2,0} \cap L^\infty$ and $r, r^* \in \mathcal{R}_{n-1,n}$ with r^* a solution to RAB(n), *i.e.* a best approximant to f in L^2 from $\mathcal{R}_{n-1,n}$. Then $\|f - r^*\|_2 \leq \|f - r\|_2$. Now, for any $h \in H^\infty$, Parseval's theorem gives $\|f - r\|_2 \leq \|f - r - h\|_2$. Finally, since the L^∞-norm dominates the L^2-norm $\|f - r - h\|_2 \leq \|f - r - h\|_\infty$ and so we have

$$\|f - r^*\|_2 \leq \|f - (r + h)\|_\infty.$$

[1] That $\lambda_{2n+1}(A_f) = \lambda_{2n+2}(A_f)$ in (16) is inessential and due the fact that A_f is complex linear whereas the genus is a real notion.

Thus, minimizing over r, h, we find that $d_2(f, \mathcal{R}_{n-1,n}) \leq d_\infty(f, H_n^\infty)$. However, it is *a priori* unclear how large the gap between the two errors can be. Below, dwelling on Theorems 1 and 2, we derive when f is continuous a lower bound in terms of the ratio between L^2 and L^∞ norms of the singular vectors of the Hankel operator Γ_f.

Theorem 4. *Let $f \in \bar{H}^{2,0} \cap C(\mathbb{T})$ and $n \geq 0$ an integer. Consider an orthonormal family $v_0, \ldots, v_n$ of singular vectors of the Hankel operator Γ_f (cf. (17)), where v_k is associated to the singular value $s_k(\Gamma_f)$. Define $M_n(f) := \min\{d_\infty(f, H_j^\infty)/\|v_j\|_\infty, 0 \leq j \leq n\}$ if $v_j \in H^\infty$ for $0 \leq j \leq n$, and $M_n(f) := 0$ otherwise. Then*

$$\frac{M_n(f)}{\sqrt{n+1}} \leq d_2(f, \mathcal{R}_{n-1,n}). \tag{21}$$

Proof. If $M_n(f) = 0$, then (21) is trivial. Otherwise, the linear span of $\{v_0, \ldots, v_n\}$ over $\mathbb{C}$ is a real $2n+2$-dimensional vector space in $L^2 \cap L^\infty$, and we may endow it either with the L^2-norm or else with the L^∞-norm. Let S_2 and S_∞ indicate the corresponding unit spheres. Identifying a vector with its coordinates, we see that S_2 is just $\mathbb{S}^{2n+1}$, and clearly $v \mapsto v/\|v\|_\infty$ is an odd homeomorphism from S_2 onto S_∞. Therefore, by the Borsuk–Ulam theorem, S_∞ is a compact set of genus $2n+2$. Now, if we let $v \in S_\infty$ and write $v = \sum_{j=0}^n \lambda_j v_j$ while abbreviating $s_j(\Gamma_f)$ as s_j, we get using "$\langle , \rangle$" to mean Hermitian scalar product on $\mathbb{T}$ that

$$\begin{aligned} \|A_f(v)\|_2^2 &= \langle A_f(v), A_f(v)\rangle = \langle \Gamma_f(v), \Gamma_f(v)\rangle = \langle \Gamma_f^*\Gamma_f v, v\rangle = \Sigma_{j=0}^n |\lambda_j|^2 s_j^2 \\ &\geq \frac{1}{n+1}\left(\sum_{j=0}^n |\lambda_j| s_j\right)^2 \geq \frac{M_n^2(f)}{n+1}\left(\sum_{j=0}^n |\lambda_j|\|v_j\|_\infty\right)^2 \geq \frac{M_n^2(f)}{n+1}, \end{aligned} \tag{22}$$

where the second line in (22) uses the Schwarz inequality, the definition of $M_n(f)$ together with the equality $s_j(\Gamma_f) = d_\infty(f, H_j^\infty)$ from Theorem 2, the triangle inequality and the fact that $\|v\|_\infty = 1$. Inequality (21) now follows from (22) and Theorem 1. ■

The kernels $\mathrm{Ker}A_f$ and $\mathrm{Ker}\Gamma_f$ are closed subsets of H^∞ and H^2 respectively, and clearly $\mathrm{Ker}A_f = \mathrm{Ker}\Gamma_f \cap H^\infty$. (*cf.* definitions (11) and (17)). By a theorem of Beurling [14, Chapter II, Theorem 7.1], being closed and shift-invariant (*i.e.* invariant under multiplication by the variable z), $\mathrm{Ker}\Gamma_f$ is either trivial ($\{0\}$ or H^2) or else consists of all multiples of some inner function $\mathfrak{j}$, that is, $\mathrm{Ker}\Gamma_f = \mathfrak{j}H^2$. In the latter case $\mathrm{Ker}A_f = \mathfrak{j}H^\infty$, in particular $\mathrm{Ker}\Gamma_f$ and $\mathrm{Ker}A_f$ are simultaneously nontrivial. In this situation the proof of Theorem 4 quickly leads to an improvement of itself as follows. Notations and assumptions being as in the theorem, set $\|v_j\|_{H^\infty/\mathrm{Ker}A_f}$ to be $+\infty$ if $v_j \notin H^\infty$ and to be the distance from v_j to $\mathrm{Ker}A_f$ in H^∞ otherwise. Observe that if $\|v_{j_0}\|_{H^\infty/\mathrm{Ker}A_f} = 0$ for some $j_0 \in \{0, \ldots, n\}$, then $v_{j_0} \in \mathrm{Ker}\Gamma_f$ which entails that Γ_f has rank at most j_0 by definition of singular values. It is a theorem of Kronecker [33, Chapter 1, Corollary 3.2] that this happens if and only if $f \in H_{j_0}^\infty$, and since $f \in \bar{H}^{2,0} \cap C(\mathbb{T})$ by assumption we get that $f \in \mathcal{R}_{j_0-1,j_0}$. In particular it holds in this case that $d_\infty(f, H_j^\infty) = d_2(f, \mathcal{R}_{j-1,j}) = \|v_j\|_{H^\infty/\mathrm{Ker}A_f} = 0$ for all $j \geq j_0$. Keeping this observation in mind, let us define

$$Q_n(f) := \min_{0 \leq j \leq n} \left\{ \frac{d_\infty(f, H_j^\infty)}{\|v_j\|_{H^\infty/\mathrm{Ker}A_f}} \right\}, \tag{23}$$

where $Q_n(f)$ is to be interpreted as 0 if $\|v_{j_0}\|_{H^\infty/\mathrm{Ker}A_f} = 0$ for some $j_0 \in \{1, \ldots, n\}$ (in which case $d_\infty(f, H_{j_0}^\infty) = 0$ as well by what precedes).

Corollary 1. *Theorem 4 remains valid if $M_n(f)$ gets replaced by $Q_n(f)$.*

Proof. We can assume that $v_j \in H^\infty \backslash \mathrm{Ker} A_f$ for $0 \le j \le n$, otherwise $Q_n(f) = 0$ and there is nothing to prove. By the discussion before the corollary, this amounts to say that $f \notin H_n^\infty$. Next, pick $\varepsilon > 0$ and $g_j \in \mathrm{Ker} A_f$ such that $\|v_j - g_j\|_\infty < \|v_j\|_{H^\infty/\mathrm{Ker} A_f} + \varepsilon$ for each $j \in \{1, \ldots, n\}$. If we let $w_j = v_j - g_j$, then $A_f(w_j) = \Gamma_f(w_j) = \Gamma_f(v_j)$ and the w_j are linearly independent over $\mathbb{C}$. Indeed, if $\sum_{j=0}^n \lambda_j w_j = 0$ with $\lambda_{j_0} \neq 0$, applying $\Gamma_f^* \Gamma_f$ yields $\sum_{j=0}^n \lambda_j s_j^2(\Gamma_f) v_j = 0$ and since the v_j are linearly independent we have that $s_{j_0}(\Gamma_f) = 0$; thus, by the AAK theorem, we get that $f \in H_{j_0}^\infty \subset H_n^\infty$, contrary to our initial assumption. Replacing now v_j by w_j in the proof of Theorem 4 and using that $\Gamma_f(w_j) = \Gamma_f(v_j)$, we obtain instead of (22) that, whenever $w = \sum_{j=0}^n \lambda_j w_j$ is such that $\|w\|_\infty = 1$, then

$$\|A_f(w)\|_2^2 \ge \frac{1}{n+1} \min_{0 \le j \le n} \left(\frac{d_\infty(f, H_j^\infty)}{\|v_j\|_{H^\infty/\mathrm{Ker} A_f} + \varepsilon} \right)^2 .$$

Thus, letting ε go to 0, we get the desired result from Theorem 1 again. ∎

Theorem 4 is useful only if we have a fair appraisal of $M_n(f)$. The latter is delicate to estimate in general, but in the following subsections we point out three cases where this can be done in different guises. They are: the case of a general rational function which can be approached numerically; the case of a Blaschke product where estimates can be given in terms of the zeros; the case of Cauchy integrals over hyperbolic geodesic arcs in which boundedness of $M_n(f)$ can be proved *via* a careful analysis of formulas behind AAK theory, dwelling on the work in [6].

5.2. Application to rational functions

When f is rational, the bounds in Corollary 1 can be numerically computed. As explained in Section 3, the general case reduces by partial fraction extension to the special case where $f \in \bar{H}^{2,0}$, the detail of which is carried out below.

Write $f = p/q$ where $p \in \mathcal{P}_{N-1}$, $q \in \mathcal{P}_N$ is monic with all roots in $\mathbb{D}$, and p, q are coprime as polynomials. Let us write

$$q(z) = \Pi_{k=1}^N (z - \zeta_k)$$

where each $\zeta_k \in \mathbb{D}$ is repeated according to multiplicity. It is clear from definition (17) that $\mathrm{Ker}\Gamma_f$ consists of those H^2-functions vanishing at the zeros of q, hence $\mathrm{Ker}\Gamma_f = (q/\widetilde{q})H^2$. Its orthogonal complement in H^2 is $(\mathrm{Ker}\, \Gamma_f)^\perp = \mathcal{P}_{N-1}/\widetilde{q}$, an orthonormal basis of which is given according to the Malmquist–Walsh lemma by the formulas [28, Chapter V, Section 1]:

$$e_j(z) = \frac{\left(1 - |\zeta_j|^2\right)^{1/2}}{1 - \bar{\zeta}_j z} \Pi_{k=0}^{j-1} \frac{z - \zeta_k}{1 - \bar{\zeta}_k z}, \quad 1 \le j \le N, \tag{24}$$

where the empty product is understood to be 1. The effect of Γ_f on any member of $(\mathrm{Ker}\, \Gamma_f)^\perp$ is easily computed upon introducing $a \in \mathcal{P}_{N-1}$ and $b \in \mathcal{P}_{N-1}$ such that the following Bezout relation holds: $a\widetilde{q} + bq = 1$. Indeed, one has for any $u \in \mathcal{P}_{N-1}[z]$ that

$$\Gamma_f(u/\widetilde{q}) = \mathbf{P}_-\left(\frac{pu}{q\widetilde{q}}\right) = \mathbf{P}_-\left(\frac{pua}{q} + \frac{pub}{\widetilde{q}}\right) = \mathbf{P}_-\left(\frac{pua}{q}\right) = \frac{R_q(pua)}{q}, \tag{25}$$

where we used that $pub/\widetilde{q} \in H^2$ and, for any polynomial P, $R_q(P)$ indicates the remainder of Euclidean division of P by q. In particular, we get from (25) that $\mathrm{Im}\Gamma_f = \mathcal{P}_{N-1}/q$. The Hermitian scalar product on $\mathbb{T}$ can be computed in several ways for functions in $\mathcal{P}_{N-1}/q$; one which does not use partial fraction expansion is as follows. Pick $u, v \in \mathcal{P}_{N-1}$. Observing that $z^N/\widetilde{q}$ is conjugate to $1/q$ on $\mathbb{T}$ and denoting with $Q_q(P)$ the quotient of Euclidean division of the polynomial P by q (so that $P = qQ_q(P) + R_q(P)$), we get since $a\widetilde{q} + bq = 1$ that

$$\left\langle \frac{u}{q}, \frac{v}{q} \right\rangle = \left\langle \frac{z^N u}{\widetilde{q}q}, v \right\rangle = \left\langle \frac{z^N ub}{\widetilde{q}} + Q_q\left(z^N ua\right) + \frac{R_q(z^N ua)}{q}, v \right\rangle$$
$$= \left\langle Q_q\left(z^N ua\right), v \right\rangle, \tag{26}$$

where we used that $R_q(z^N ua)/q \in \bar{H}^{2,0}$ and $z^N ub/\widetilde{q} \in z^N H^2$ are both orthogonal to $v \in \mathcal{P}_{N-1}$ by Parseval's theorem. The last term in (26) is now a scalar product between polynomials which can be computed as a Euclidean one in the basis $\{z^k;\ 0 \le k \le N-1\}$.

Writing $e_j = u_j/\widetilde{q}$ where e_j was defined in (24), we can use (25), (26) to compute the Hermitian matrix $M = \langle \Gamma_f^* \Gamma_f(e_i), e_j \rangle = \langle \Gamma_f(e_i), \Gamma_f(e_j) \rangle$, and an orthonormal family of singular vectors $v_0, \ldots, v_{N-1}$ associated with the nonzero singular values of Γ_f is then obtained by diagonalization of M (of course any other orthonormal basis of $\mathcal{P}_{N-1}/\widetilde{q}$ than (e_k) could be used as well). More precisely, the k-th row of a unitary matrix U such that UMU^* is diagonal yields coordinates for v_k in the basis e_j. The diagonal terms are the squared singular values $s_k^2(\Gamma_f)$ for $0 \le k \le N-1$, which are none but the $d_\infty(f, H_k^\infty)$ by the AAK theorem. Moreover, it follows from Nehari's theorem that

$$\|v_j\|_{H^\infty/\mathrm{Ker}A_f} = d_\infty(v_j\widetilde{q}/q, H^\infty) = \|\!|\Gamma_{v_j\widetilde{q}/q}|\!\|, \tag{27}$$

and the last term in (27) is the largest singular value of a Hankel operator with rational symbol which can be computed in the same manner as indicated above to compute $s_0(\Gamma_f)$.

Thus, we can evaluate Q_n defined in (23) for all n, hence also the lower bound on $d_2(f, \mathcal{R}_{n-1,n})$ given by Theorem 4 and Corollary 1. We implemented a prototype algorithm to compute these two bounds. Numerical experiments are presented in Section 7.

5.3. Application to Blaschke products

In this section, we use Theorem 4 to derive some lower bounds for Problem RA(n) when f is a Blaschke product of finite or infinite degree. This last case is instructive to contrast rational approximation in L^2 and L^∞ norms, for on the one hand the value of Problem RA(n) tends to zero as n goes large (since rational functions are dense in H^2), while on the other hand f cannot be approximated "at all" by rational functions in H^∞, *i.e.* zero is a best uniform approximant. This follows from the lemma below which is not easy to locate in the literature.

Lemma 1. *Let b be a Blaschke product and n be a positive integer which is strictly less than the degree of b (if b has infinite degree the assumption is void). Then*

$$d_\infty(b, \mathcal{R}_{n,n}) = \|b\|_\infty = 1. \tag{28}$$

Proof. Clearly $d_\infty(b, \mathcal{R}_{n,n}) \le 1$ for zero is a candidate approximant. Moreover, if $r \in \mathcal{R}_{n,n} \cap H^\infty$ then $\bar{r} \in H_n^\infty$. Therefore, upon conjugating, we get $d_\infty(b, \mathcal{R}_{n,n}) \ge d_\infty(\bar{b}, H_n^\infty)$ and it is enough to show the latter is at least 1, hence in fact equal to 1.

Assume first that b has finite degree d, and write $b = q_d/\widetilde{q}_d$ where $q_d \in \mathcal{P}_d$ has zeros in $\mathbb{D}$ only. Then $\bar{b} = \widetilde{q}_d/q_d$, and the kernel of $\Gamma_{\bar{b}}$ is bH^2 whose orthogonal complement in H^2 is $(\operatorname{Ker}\Gamma_{\bar{b}})^\perp = \mathcal{P}_{d-1}/\widetilde{q}_d$ as pointed out in the previous section. Now, if $p_{d-1} \in \mathcal{P}_{d-1}$, then $\Gamma_{\bar{b}}(p_{d-1}/\widetilde{q}_d) = p_{d-1}/q_d$ so that $\Gamma_{\bar{b}}$ is an isometry from $(\operatorname{Ker}\Gamma_{\bar{b}})^\perp$ onto its image. Consequently the first d singular values of $\Gamma_{\bar{b}}$ are equal to 1 (the remaining ones being zero). That $d_\infty(\bar{b}, H_n^\infty) = 1$ now follows from the AAK theorem and the fact that $n \le d-1$.

Assume next that b has infinite degree. We can write $b = b_{n+1}b_\infty$ where b_{n+1} has degree $n+1$ and b_∞ has infinite degree. If $g \in H_n^\infty$ then also $b_\infty g \in H_n^\infty$, and since $|b_\infty| = 1$ a.e. on $\mathbb{T}$, we get by the first part of the proof that

$$\|\bar{b} - g\|_\infty = \|b_\infty \bar{b} - b_\infty g\|_\infty = \|\bar{b}_{n+1} - b_\infty g\|_\infty \ge 1, \quad g \in H_n^\infty, \tag{29}$$

hence $d_\infty(\bar{b}, H_n^\infty) \ge 1$, as desired. ■

We turn to the main result of this section:

Theorem 5. *Let b be a Blaschke product, of finite or infinite degree. Let us arrange its zeros into a (finite or infinite) sequence $\zeta_1, \zeta_2, \ldots,$ where each ζ_j is repeated according to its multiplicity and the corresponding sequence of moduli is nondecreasing: $|\zeta_1| \le |\zeta_2| \le \cdots$. For each positive integer n strictly less than the degree of b (if b has infinite degree the assumption is void), it holds that*

$$\frac{\left(1 - |\zeta_{n+1}|^2\right)^{1/2}}{\sqrt{n+1}} \le d_2(b, \mathcal{R}_{n,n}) \tag{30}$$

and also that

$$\left(\sum_{j=0}^{n} \frac{1}{\left(1 - |\zeta_j|^2\right)^{1/2}}\right)^{-1} \le d_2(b, \mathcal{R}_{n,n}). \tag{31}$$

Proof. Assume first that b has finite degree d, so that $b \in C(\mathbb{T})$, and write $b = q_d/\widetilde{q}_d$ where $q_d \in \mathcal{P}_d$ has zeros in $\mathbb{D}$ only. By the equivalence between Problem RA(n) and RAN(n) discussed in Section 3, we know that

$$d_2(b, \mathcal{R}_{n,n}) = d_2\Big((\bar{b} - \overline{b(0)}), \mathcal{R}_{n-1,n}\Big).$$

Now, the Hankel operators $\Gamma_{\bar{b}}$ and $\Gamma_{\bar{b}-\overline{b(0)}}$ coincide and we saw in the proof of Lemma 1 that $\Gamma_{\bar{b}}$ is an isometry from $(\operatorname{Ker}\Gamma_{\bar{b}})^\perp = \mathcal{P}_{d-1}/\widetilde{q}_d$ onto $\operatorname{Im}\Gamma_{\bar{b}} = \mathcal{P}_{d-1}/\widetilde{q}_d$. Hence the e_j given by (24) for $1 \le j \le d$ form an orthonormal family of $d \ge n+1$ singular vectors associated with the singular value 1. By the AAK theorem it follows that $d_\infty(\bar{b}, H_j^\infty) = 1$ for $0 \le j \le d-1$, and since $\|e_j\|_\infty = (1 - |\zeta_j|^2)^{-1/2}$, estimate (30) follows at once from Theorem 4 upon choosing $v_j = e_{j+1}$ for $0 \le j \le n$.

Next, if we let $w_j = (\sum_{k=0}^n e^{2i\pi kj/(n+1)} v_k)/(n+1)^{1/2}$ for $0 \le j \le n$, we get another orthonormal family of $n+1$ singular vectors associated with the singular value 1, and clearly $\|w_j\|_\infty \le (n+1)^{-1/2} \sum_{k=0}^{n+1} (1 - |\zeta_k|^2)^{-1/2}$ for all j. Estimate (31) now follows from Theorem 4 again upon replacing the previous v_j by w_j.

If now b is infinite and $k \geq 0$ is the multiplicity of the zero at the origin, we can write (6) for some constant c of unit modulus. Let us define $b_m = cz^m$ if $m \leq k$ and

$$b_m(z) = cz^k \prod_{l=k+1}^{m} \frac{-\bar{\zeta}_l}{|\zeta_l|} \frac{z-\zeta_l}{1-\bar{\zeta}_l z}, \quad m > k. \tag{32}$$

The sequence of Blaschke products $\{b_m\}$ converges to b pointwise on $\mathbb{D}$, and since it is bounded it must also converge weakly to b in H^2. Since b_m and b have norm 1, the limit of the norms is the norm of the weak limit, hence the convergence is actually strong in H^2 [9, Theorem 3.32]. Consequently

$$\lim_{m\to\infty} d_2(b_m, \mathcal{R}_{n,n}) = d_2(b, \mathcal{R}_{n,n}),$$

and since estimates (30), (31) depend only of the first $n+1$ zeros of b they remain valid in the limit. ■

In view of Corollary 1, the conclusion of Theorem 5 can be sharpened upon replacing in the proof $\|v_j\|_\infty$ and $\|w_j\|_\infty$ by $|||\Gamma_{\bar{b}v_j}|||$ and $|||\Gamma_{\bar{b}w_j}|||$. Computations become more involved but in any case cannot increase the left hand side of (30) and (31) by more than a factor 2. Incidentally, for $q_n \in \mathcal{P}_N$ having all roots in $\mathbb{D}$, it seems to be an open question which L^2-orthonormal bases of $\mathcal{P}_{n-1}/q_n$ have minimax L^∞-norm. Using such bases instead of e_j in the proof of Theorem 5 may improve on the result.

Since (7) is necessary and sufficient for $\{\zeta_l\}$ to be the zero set of a Blaschke product, an immediate corollary to Theorem 5 is:

Corollary 2. *Whenever α_n is a nonincreasing sequence in $(0, 1]$ such that $\Sigma_n\alpha_n < \infty$, there is a Blaschke product b such that*

$$\frac{\alpha_{n+1}^{1/2}}{\sqrt{n+1}} \leq \inf_{r\in\mathcal{R}_{n,n}\cap H^2} \|b-r\|_2, \quad n \in \mathbb{N}, \tag{33}$$

and also

$$\left(\sum_{j=0}^{n} \frac{1}{\alpha_j^{1/2}}\right)^{-1} \leq \inf_{r\in\mathcal{R}_{n,n}\cap H^2} \|b-r\|_2, \quad n \in \mathbb{N}. \tag{34}$$

5.4. Application to Cauchy integrals on hyperbolic geodesics

Recall that geodesic lines for the hyperbolic metric in $\mathbb{D}$ are radii and circular arcs orthogonal to $\mathbb{T}$ [14, Chapter I]. By definition, a hyperbolic geodesic segment is a compact and connected subset thereof. Alternatively, a hyperbolic geodesic segment is the image of a real segment $[a, b] \subset (0, 1)$ under an automorphism of the disk (*i.e.* a Möbius transformation of the type $z \mapsto \alpha(z - z_0)/(1 - \bar{z}_0 z)$ with $|\alpha| = 1$ and $z_0 \in \mathbb{D}$, in other words a Blaschke product of degree 1). Below is a nonstandard characterization of hyperbolic geodesic segments which is analytic in nature. We will not use the "if" part but it is interesting in itself.

Lemma 2. *A C^1-smooth, closed Jordan arc $\gamma \subset \mathbb{D}$ is a hyperbolic geodesic segment if, and only if there is a constant $C = C(\gamma) > 0$ such that, to each $g \in H^2$, there is $h \in H^2$ with $h_{|\gamma} = \bar{g}_{|\gamma}$ and $\|h\|_2 \leq C\|g\|_2$. If g is continuous on $\overline{\mathbb{D}}$, so is h.*

Proof. If γ is hyperbolic geodesic segment, then it is the image of a real segment under an automorphism φ of $\mathbb{D}$ and $h(z) = \overline{(g \circ \varphi)(\bar{z})} \circ \varphi^{-1}$ does the job. Conversely, if γ is a C^1-smooth closed Jordan arc in $\mathbb{D}$ with endpoints z_1, z_2 and if there exists a constant $C = C(\gamma)$ as in the statement of the lemma, then the proof of [6, Theorem 10.1] applies (upon trading the geodesic arc **G** for γ in that proof) to show that γ consists exactly of non-isolated points of the cluster set, as n ranges over $\mathbb{N}$, of poles of best approximants to $((z - z_1)(z - z_2))^{-1/2} \in C(\mathbb{T})$ from H_n^∞. Because this characterization depends only on z_1, z_2, it follows that γ must be the geodesic arc joining them. ■

In this section, we will consider functions of the form

$$f(z) = \frac{1}{2i\pi} \int_G \frac{h(\xi)}{z - \xi} d\xi \tag{35}$$

where:

(H1) $G \subset \mathbb{D}$ is a geodesic segment,
(H2) h is a complex-valued function on G, summable with respect to arclength, having continuous argument except possibly for finitely many jumps of amplitude π.

The prototype of such a function is one which is analytic over $\mathbb{D}$ except for two branchpoints of order strictly greater than -1. Indeed, by Cauchy formula, such a function can be written as the Cauchy integral, on any smooth cut connecting the branchpoints, of the jump of the function across the cut. This jump is locally analytic and has continuous argument on the cut (up to the branchpoints by Puiseux expansion), except at the zeros that the jump may have on this cut where the argument has left and right limits which differ by $k\pi$ if k is the order of the zero. Choosing the hyperbolic geodesic cut, we get representation (35). It may seem artificial to favor the hyperbolic geodesic segment linking the branchpoints among all possible cuts. However, this one turns out to attract almost all poles of best rational approximants (see [8] for this and generalizations to finitely many branchpoints) and also of best meromorphic approximants (see [6, Theorem 10.1] and Corollary 3 below), which makes it in some sense the natural singular set of the function.

We need additional facts from AAK theory that shed light on singular vectors of Hankel operators with continuous symbol. They apply in particular to Γ_f when f is of the form (35).

- For $f \in C(\mathbb{T})$ and $n \geq 0$, a best approximant g_n to f from H_n^∞ in L^∞ uniquely exists [1, Theorem 1.3] [33, Chapter 4, Theorem 1.3] which is given by

$$g_n = \frac{\mathbf{P}_+(f v_n)}{v_n}, \qquad f - g_n = \frac{\Gamma_f(v_n)}{v_n} = \frac{\mathbf{P}_-(f v_n)}{v_n}, \tag{36}$$

  where v_n is *any* singular vector of Γ_f associated with $s_n(\Gamma_f)$; moreover, the error function $f - g_n$ has constant modulus $s_n(\Gamma_f)$ a.e. on $\mathbb{T}$ [1, Theorem 1.3] [33, Chapter 4, Section 1, Eq. (1.12)]. In particular, (36) entails that the ratios $\mathbf{P}_\pm(f v_n)/v_n$ are independent of which singular vector v_n associated with $s_n(\Gamma_f)$ is used; this is remarkable for if $s_n(\Gamma_f)$ has multiplicity μ, then the union of $\{0\}$ and of all associated singular vectors is a vector space of complex dimension μ.
- When $f \in C(\mathbb{T})$, the inner factor of a singular vector of Γ_f is a finite Blaschke product. More precisely, keeping notations as in the previous item and letting in addition $m = m(n)$ be the smallest non-negative integer such that $s_m(\Gamma_f) = s_n(\Gamma_f)$, the singular vector v_n may be inner–outer factorized as

$$v_n = b b_m w_n \tag{37}$$

where $w_n \in H^2$ is outer and $b_m \in B_m$ is a Blaschke product of exact degree m with zeros the poles of g_n $(= g_m)$, while b is a finite Blaschke product whose zeros are also zeros of $\mathbf{P}_+(f v_n)$. Moreover, with b, b_m and w_n as in (37), it holds that

$$\Gamma_f(v_n)(z) = s_n(\Gamma_f)\, z^{-1}\overline{b_m(1/\bar{z})j(1/\bar{z})w_n(1/\bar{z})}, \quad |z| \geq 1 \tag{38}$$

where j is a finite Blaschke product such that $jb \in B_{\mu-1}$ and μ is the multiplicity of $\sigma_n(\Gamma_f)$ [1, Theorem 1.2].

- Assumptions and notations being as in the previous items, let v_n be a singular vector of Γ_f associated with $s_n(\Gamma_f)$ and (37) be its inner–outer factorization. *We claim* that $b_m w_n$ is also a singular vector of Γ_f associated with $s_n(\Gamma_f)$. Indeed, we know from the previous item that $g_n = b_m^{-1} h$ for some $h \in H^\infty$. Since $\|f - h/b_m\|_\infty = \|f b_m - h\|_\infty$, the fact that g_n is a best approximant to f from H_n^∞ entails that h is the best approximant to $f b_m \in C(\mathbb{T})$ from H^∞, hence $|||\Gamma_{f b_m}||| = \|f - h/b_m\|_\infty = s_n(\Gamma_f)$ by the AAK theorem. Taking into account that $\Gamma_{f b_m}(u) = \Gamma_f(b_m u)$ for $u \in H^2$, and also that $\Gamma_f^*(\Phi) = \mathbf{P}_+(\bar{f}\Phi)$ for $\Phi \in H^{2,0}$, while using that $\bar{b}_m \bar{H}^{2,0} \subset \bar{H}^{2,0}$ and $\mathbf{P}_+ + \mathbf{P}_- = Id$, we now compute

$$\begin{aligned}\Gamma_{f b_m}^* \Gamma_{f b_m}(b w_n) &= \Gamma_{f b_m}^* \Gamma_f(v_n) = \mathbf{P}_+\Big(\overline{f b_m}\, \Gamma_f(v_n)\Big) = \mathbf{P}_+\Big(\bar{b}_m \mathbf{P}_+\big(\bar{f}\Gamma_f(v_n)\big)\Big) \\ &= \mathbf{P}_+\left(\bar{b}_m \Gamma_f^* \Gamma_f(v_n)\right) = s_n^2(\Gamma_f)\mathbf{P}_+\left(\bar{b}_m v_n\right) = s_n^2(\Gamma_f) b w_n.\end{aligned}$$

This shows that $b w_n$ is a maximizing vector of $\Gamma_{f b_m}$. Next, we observe that

$$\|\Gamma_{f b_m}(b w_n)\|_2 = \|\mathbf{P}_-\big(b \Gamma_{f b_m}(w_n)\big)\|_2 \leq \|\Gamma_{f b_m}(w_n)\|_2 \tag{39}$$

because multiplication by b is an isometry and anti-analytic projection is a contraction in L^2. Since $\|b w_n\|_2 = \|w_n\|_2$, we conclude from (39) that w_n is in turn a maximizing vector of $\Gamma_{f b_m}$ and that equality must hold throughout in this equation. In other words $b\Gamma_{f b_m}(w_n) \in \bar{H}^{2,0}$, which implies easily that $b\Gamma_{f b_m}(w_n) = \Gamma_{f b_m}(b w_n)$. Consequently

$$\begin{aligned}\Gamma_f^* \Gamma_f(b_m w_n) &= \mathbf{P}_+\left(\overline{f b}\, b \Gamma_f(b_m w_n)\right) = \mathbf{P}_+\left(\bar{b}\mathbf{P}_+(\bar{f}\Gamma_f(b b_m w_n))\right) \\ &= \mathbf{P}_+\left(\bar{b}\Gamma_f^* \Gamma_f(v_m)\right) = s_n^2(\Gamma_f)\mathbf{P}_+\left(\bar{b} v_n\right) = s_n^2(\Gamma_f) b_m w_n.\end{aligned} \tag{40}$$

This proves the claim.

We now assume that f has the form (35). Using (4) to express definition (17) of the Hankel operator, then inserting (35) and using successively Fubini's theorem and the residue formula, we obtain:

$$\begin{aligned}\Gamma_f(v_n)(z) &= \left(\frac{1}{2i\pi}\right)^2 \int_G h(\xi) d\xi \int_{\mathbb{T}} \frac{v_n(\zeta)}{(\zeta - \xi)(z - \zeta)} d\zeta \\ &= \frac{1}{2i\pi} \int_G \frac{v_n(\xi) h(\xi)}{z - \xi} d\xi, \quad |z| > 1.\end{aligned} \tag{41}$$

In particular $\Gamma_f(v_n)$ extends analytically from $\overline{\mathbb{C}} \backslash \overline{\mathbb{D}}$ to $\overline{\mathbb{C}} \backslash G$, and Eq. (38) becomes

$$s_n(\Gamma_f)\, z^{-1}\overline{b_m(1/\bar{z})j(1/\bar{z})w_n(1/\bar{z})} = \frac{1}{2i\pi} \int_G \frac{v_n(\xi) h(\xi)}{z - \xi} d\xi, \quad |z| \geq 1. \tag{42}$$

Multiplying the restriction of (42) to $z \in \mathbb{T}$ by $b_m j$ and then taking anti-analytic projection again gives us after a similar computation:

$$s_n(\Gamma_f)\, z^{-1}\overline{w_n(1/\bar{z})} = \frac{1}{2i\pi}\int_G \frac{j(\xi)b_m^2(\xi)b(\xi)w_n(\xi)h(\xi)}{z-\xi}d\xi, \quad |z|>1, \tag{43}$$

where we took into account (37). Eq. (43) entails that in turn $\check{w}_n$ (*cf.* (3)) extends analytically from $\overline{\mathbb{C}}\backslash\overline{\mathbb{D}}$ to $\overline{\mathbb{C}}\backslash G$, or equivalently that w_n extends analytically from $\mathbb{D}$ to $\overline{\mathbb{C}}\backslash\overline{G}^{-1}$, where $\overline{G}^{-1}$ is the reflection of G across $\mathbb{T}$.

We can now establish a technical result which is the key for applying Theorem 4 to functions of the form (35). Recall that a family of analytic functions in an open set $\Omega \subset \mathbb{C}$ is said to be normal if it is uniformly bounded on every compact subset of Ω. Equivalently, a normal family of analytic functions is one which is relatively compact for the topology of locally uniform convergence in Ω.

Proposition 2. *Let f assume the form (35) where hypotheses* H1–H2 *do hold, and $\{v_n\}_{n\in\mathbb{N}}$ be a sequence of singular vectors of Γ_f such that $\|v_n\|_2 = 1$ for all n. Denote by w_n the outer factor of v_n. Then, $\{w_n\}_{n\in\mathbb{N}}$ is a normal family in $\overline{\mathbb{C}}\backslash\overline{G}^{-1}$.*

Proof. We already pointed out that w_n is analytic in $\overline{\mathbb{C}}\backslash\overline{G}^{-1}$. According to Eq. (37), the inner–outer factorization of v_n is of the form $v_n = bb_m w_n$, and we know from a previous claim (*cf.* Eq. (40)) that $b_m w_n$ is another singular vector of Γ_f associated with $s_n(\Gamma_f)$ having the same outer factor w_n. Hence we can replace v_n by $b_m w_n$ (in other words, we may – and we shall – assume that $b \equiv 1$ and write $v_n = b_m w_n$). For correctness, one should of course write $m(n)$ throughout, but we drop the dependence of m on n for simplicity.

To prove that w_n is bounded independently of n on each compact subset of $\overline{\mathbb{C}}\backslash\overline{G}^{-1}$, we parallel the argument of [6, Theorem 10.1].

Let $t \mapsto \alpha(t)$ parametrize G with an automorphism α of $\mathbb{D}$ as t ranges over a real segment $[a,b]$. Then $t \mapsto \alpha'(t)$ has continuous argument. Let β_1 be a finite Blaschke product with real coefficients vanishing precisely at the jumps of amplitude π that $t \mapsto \arg h(\alpha(t))$ may have on $[a,b]$ (if $h(\alpha(t))$ is continuous we simply put $\beta \equiv 1$). Then $t \mapsto \arg(\beta_1(t)h(\alpha(t)))$ is continuous by our assumptions on h. Thus, by Mergelyan's theorem, there is a polynomial T which is real valued on $[a,b]$ and such that $|T(t) + \arg\alpha'(t) + \arg(\beta_1(t)h(\alpha(t)))| < \pi/3$ for $t \in [a,b]$. In invariant form, this means that the function $H = P \circ \alpha^{-1} \in H^\infty \cap C(\mathbb{T})$ is real valued on G and moreover that

$$|\beta(\xi)h(\xi)d\xi| = \left|e^{iH(\xi)}\beta(\xi)h(\xi)d\xi\right| \le 2\mathrm{Re}\left(e^{iH(\xi)}\beta(\xi)h(\xi)d\xi\right), \quad \xi \in G, \tag{44}$$

where $\beta = \beta_1 \circ \alpha^{-1}$ is in turn a finite Blaschke product which is real-valued on G. Notice that H and β depend only on f and not on n.

In another connection, since w_n has no zero in $\mathbb{D}$, it has a well-defined square root $w_n^{1/2} \in H^\infty$. Note, since $\|v_n\|_2 = \|w_n\|_2 = 1$ by assumption, that $\|w_n^{1/2}\|_2 = \|w_n\|_1^{1/2} \le 1$ by the Schwarz inequality. Appealing to Lemma 2, let $H_n \in H^2$ take conjugate values to $b_m w_n^{1/2}$ on G, with $\|H_n\|_2 \le C\|b_m w_n^{1/2}\|_2 = C\|w_n^{1/2}\|_2 \le C$. Note that H_n is continuous on $\overline{\mathbb{D}}$ since b_m and $w_n^{1/2}$ are. For j as in (38) (j depends on v_n but we drop this dependence), consider the contour integral

$$\frac{s_n(\Gamma_f)}{2i\pi}\int_{\mathbb{T}} e^{iH(\xi)}\,H_n(\xi)\,\beta(\xi)\,\overline{b_m(\xi)j(\xi)}\,\frac{\overline{w_n(\xi)}}{w_n^{1/2}(\xi)}\,\frac{d\xi}{\xi} \tag{45}$$

where it should be observed that the integrand is continuous even though w_n may have zeros on $\mathbb{T}$ (of course at such points $\bar{w}_n/w_n^{1/2}$ is understood to be 0). In view of (42), this integrand extends analytically on $\mathbb{D}\backslash G$, hence we may rewrite (45) as an integral over the circle $\mathbb{T}_r = \{z : |z| = r\}$ where $r \in (0,1)$ is close enough to 1 that $\mathbb{T}_r$ encompasses G. Then, substituting (42) and using again Fubini's theorem and the Cauchy formula (which is permitted since $w_n^{1/2}(z)$ does not vanish for $|z| \le r$), the integral (45) transforms into

$$\begin{aligned}&\frac{1}{(2i\pi)^2}\int_{\mathbb{T}_r}\left(\int_G \frac{v_n(\zeta)h(\zeta)}{\xi-\zeta}d\zeta\right)\frac{e^{iH(\xi)}\,H_n(\xi)\beta(\xi)}{w_n^{1/2}(\xi)}\,d\xi\\ &\quad= \frac{1}{2i\pi}\int_G \frac{v_n(\zeta)}{w_n^{1/2}(\zeta)}h(\zeta)e^{iH(\zeta)}\,H_n(\zeta)\beta(\zeta)\,d\zeta\\ &\quad= \frac{1}{2i\pi}\int_G b_m(\zeta)w_n^{1/2}(\zeta)h(\zeta)e^{iH(\zeta)}\,H_n(\zeta)\,\beta(\zeta)\,d\zeta,\end{aligned}$$

where we took (37) into account. Altogether, by the construction of H_n, we deduce that

$$\begin{aligned}&\frac{s_n(\Gamma_f)}{2i\pi}\int_{\mathbb{T}} e^{iH(\xi)}\,H_n(\xi)\,\beta(\xi)\overline{b_m(\xi)j(\xi)}\,\frac{\overline{w_n(\xi)}}{w_n^{1/2}(\xi)}\,\frac{d\xi}{\xi}\\ &\quad= \frac{1}{2i\pi}\int_G \left|b_m^2(\zeta)\,w_n(\zeta)\right|h(\zeta)\,\beta(\zeta)e^{iH(\zeta)}d\zeta.\end{aligned} \tag{46}$$

By (44), we get on the one hand that

$$\frac{1}{4\pi}\int_G \left|b_m^2(\zeta)\,w_n(\zeta)\,\beta(\zeta)\,h(\zeta)\right|d|\zeta| \le \left|\frac{1}{2i\pi}\int_G \left|b_m^2(\zeta)\,w_n(\zeta)\right|\beta(\zeta)h(\zeta)e^{iH(\zeta)}\right|d\zeta. \tag{47}$$

On the other hand, since β, b_m, j are Blaschke products while $\|w_n^{1/2}\|_2 \le 1$ and $\|H_n\|_2 \le C$, we see from the Schwarz inequality that

$$\left|\frac{s_n(\Gamma_f)}{2i\pi}\int_{\mathbb{T}} e^{iH(\xi)}\,H_n(\xi)\,\beta(\xi)\,\overline{b_m(\xi)j(\xi)}\,\frac{\overline{w_n(\xi)}}{w_n^{1/2}(\xi)}\,\frac{d\xi}{\xi}\right| \le C\,s_n(\Gamma_f)\,\|e^{iH}\|_\infty. \tag{48}$$

Therefore, in view of (47), (46), and (48), we get that

$$\frac{1}{2\pi}\int_G \left|b_m^2(\zeta)\,w_n(\zeta)\beta(\zeta)h(\zeta)\right|d|\zeta| \le 2C\,s_n(\Gamma_f)\,\|e^{iH}\|_\infty. \tag{49}$$

Now, if we multiply (43) (where $b \equiv 1$) by β and apply $\mathbf{P}_-$ to this product, the computation based on Fubini's theorem and Cauchy formula that led us to (41) and (43) yields

$$s_n(\Gamma_f)\,\mathbf{P}_-(\beta\check{w}_n)(z) = \frac{1}{2i\pi}\int_G \frac{j(\xi)b_m^2(\xi)b(\xi)w_n(\xi)\beta(\xi)h(\xi)}{z-\xi}d\xi, \quad |z| \ge 1. \tag{50}$$

Eq. (50) entails that $\mathbf{P}_-(\beta\check{w}_n)$ extends analytically to $\overline{\mathbb{C}}\setminus G$ and, as $|j| \leq 1$ in $\mathbb{D}$ since it is a Blaschke product, it follows from (49) and (50) that

$$|\mathbf{P}_-(\beta\check{w}_n)(z)| \leq 2C\,\|e^{iH}\|_\infty \left(\inf_{\zeta\in G}|z-\zeta|\right)^{-1}, \quad z\in\overline{\mathbb{C}}\setminus G.$$

This proves that $|\mathbf{P}_-(\beta\check{w}_n)|$ is uniformly bounded with respect to n on every compact subset of $\overline{\mathbb{C}}\setminus G$. In another connection, observe from (4) that $\mathbf{P}_+(\beta\check{w}_n)$ is uniformly bounded with respect to n on compact subsets of $\mathbb{D}\setminus G$, because $\|\beta\check{w}_n\|_2 = \|\check{w}_n\|_2 = 1$. Adding up, we get that $\beta\check{w}_n$ is uniformly bounded with respect to n on compact subsets of $\mathbb{D}\setminus G$. Since $|\beta|$, which is a finite Blaschke product with all its zeros on G, is bounded from below on compact subsets of $\overline{\mathbb{C}}\setminus G$, we thus conclude that $\{\check{w}_n\}$ is normal in $\overline{\mathbb{C}}\setminus G$. By reflection across $\mathbb{T}$, normality of $\{w_n\}$ in $\overline{\mathbb{C}}\setminus\overline{G}^{-1}$ follows, as desired. ■

The following corollary to Proposition 2 is worth pointing out as it shows in a rather strong sense that most of the poles of best L^∞ meromorphic approximants to f as in (35) asymptotically cluster to G.

Corollary 3. *Let f assume the form (35) where hypotheses* H1–H2 *do hold. Denote by g_n the best approximant to f from H_n^∞ in L^∞. To each neighborhood $\mathcal{V}(G)$ of G, there are $n_0, N_0\in\mathbb{N}$ such that, if $n\geq n_0$, then g_n has at most N_0 poles outside $\mathcal{V}(G)$, counting multiplicity.*

Proof. We make notations as in the proof of Proposition 2. We noticed already before the latter that $\check{w}_n$ is analytic in $\overline{\mathbb{C}}\setminus G$. In addition, it is clear that $\overline{b_m(1/\bar{z})} = 1/b_m(z)$ (resp. $\overline{j(1/\bar{z})} = 1/j(z)$) since b_m (resp. j) is unimodular on $\mathbb{T}$. Hence $\overline{b_m(1/\bar{z})}$ (resp. $\overline{j(1/\bar{z})}$) is meromorphic in $\mathbb{C}$ with poles at the zeros of b_m (resp. of j) and no zero in $\overline{\mathbb{D}}$. Since the right hand side of (42) is analytic in $\overline{\mathbb{C}}\setminus G$, we conclude that every zero of b_m (and of j) which does not lie on G is a zero of $\check{w}_n$ with same or greater multiplicity. Now, by Proposition 2, every subsequence $\check{w}_{n_k}$, has a subsequence $\check{w}_{n_{k_\ell}}$ converging locally uniformly in $\overline{\mathbb{C}}\setminus G$ to some analytic function $\check{w}$ which is not the zero function because $\|\check{w}\|_2 = \lim_{\ell\to\infty}\|\check{w}_{n_{k_\ell}}\|_2 = 1$. In particular $\check{w}$ has only finitely many zeros $z_1,\ldots,z_N$ of respective multiplicities $\mu_1,\ldots,\mu_N$ in $\mathbb{D}\setminus\mathcal{V}_G$. Thus, by the Rouché theorem, $\check{w}_{n_{k_\ell}}$ has exactly μ_j zeros in the neighborhood of z_j for ℓ large enough, counting multiplicities, and no other zero in $\mathbb{D}\setminus\mathcal{V}_G$. Consequently every subsequence of $\{\check{w}_n\}$ has boundedly many zeros in $\mathbb{D}\setminus\mathcal{V}_G$, which implies the desired conclusion as poles of g_m which do not lie on G are zeros of $\check{w}_n$ by the first part of the proof. ■

The main result of this section is the following.

Theorem 6. *Let f assume the form (35) where hypotheses* H1–H2 *do hold. Then,*

$$C_1\frac{d_\infty(f,\mathcal{R}_{n-1,n})}{\sqrt{n+1}} \leq C_2\frac{d_\infty(f,H_n^\infty)}{\sqrt{n+1}} \leq d_2(f,\mathcal{R}_{n-1,n}) \tag{51}$$

where C_1, C_2 are strictly positive constants depending on f but not on n.

Proof. If v_n is a singular vector of Γ_f associated with $s_n(\Gamma_f)$, normalized so that $\|v_n\|_2 = 1$, and if w_n is the outer factor of v_n, we deduce from Proposition 2 that $\|w_n\|_\infty = \|v_n\|_\infty$ is bounded independently of n. Hence the second inequality in (51) follows from Theorem 4.

To prove the first inequality, we must show that

$$d_\infty(f,\mathcal{R}_{n-1,n}) \leq C\,d_\infty(f,H_n^\infty) \tag{52}$$

for some constant C independent of n. Let g_n be a best approximant to f from H_n^∞ in L^∞, and write $g_n = r_n + h_n$ where $r_n \in \mathcal{R}_{n-1,n} \cap \bar{H}^{\infty,0}$ while $h_n \in H^\infty$. Note that $f - r_n \in \bar{H}^{\infty,0}$, hence $h_n = \mathbf{P}_+(f - g_n)$. Obviously it holds that $d_\infty(f, \mathcal{R}_{n-1,n}) \leq \|f - r_n\|_\infty$, therefore, it is enough to check that $\|f - r_n\|_\infty \leq C d_\infty(f, H_n^\infty)$ in order to establish (52). Now, by the triangle inequality, we get that

$$\|f - r_n\|_\infty \leq \|f - g_n\| + \|h_n\|_\infty = d_\infty(f, H_n^\infty) + \|\mathbf{P}_+(f - g_n)\|_\infty,$$

and we are left to prove that $\|\mathbf{P}_+(f - g_n)\|_\infty \leq C d_\infty(f, H_n^\infty)$. Let v_n be a singular vector of Γ_f, associated with $s_n(\Gamma_f)$, having inner–outer factorization $v_n = b_m w_n$, where $b_m \in B_m$ vanishes exactly at the poles of g_n and w_n is outer; this is possible by a previous claim (*cf.* (40)). Here and below, we should write for correctness $m = m(n)$, but we drop the dependence of m on n for simplicity. From (36) and (38), we gather that

$$\mathbf{P}_+(f - g_n) = s_n(\Gamma_f)\,\mathbf{P}_+\left(\bar{b}_m^2 \bar{j} \check{w}_n / w_n\right),$$

where we also dropped the dependence of j on v_n, and since $s_n(\Gamma_f) = d_\infty(f, H_n^\infty)$ it remains to establish that $\|\mathbf{P}_+\left(\bar{b}_m^2 \bar{j} \check{w}_n / w_n\right)\|_\infty$ is bounded independently of n. For this, it is enough to show that from any subsequence n_k one can extract a subsequence n_{k_ℓ} for which the property holds. Appealing to Proposition 2 as in the proof of Corollary 3, we can extract from $\{w_{n_k}\}$ a subsequence $\{w_{n_{k_\ell}}\}$ converging locally uniformly to some w, analytic in $\overline{\mathbb{C}} \setminus \overline{G}^{-1}$, which is not the zero function. Let w have N zeros lying on $\mathbb{T}$, say $z_1, \ldots, z_N$, where multiplicities are accounted by repetition and it is understood if $N = 0$ that $\{z_j\}$ is the empty set. Pick $\varepsilon > 0$ small enough that the circle $\mathbb{T}_{1+\varepsilon}$ does not meet $\overline{G}^{-1}$ and w has no other zeros than $z_1, \ldots, z_N$ in the corona $\mathcal{C}_\varepsilon = \{z : 1 \leq |z| \leq 1 + \varepsilon\}$. When ℓ is large enough, by the Rouché theorem, $w_{n_{k_\ell}}$ has exactly N zeros $z_{1,\ell}, \ldots, z_{N,\ell}$ in $\mathcal{C}_\varepsilon$, counting multiplicities with repetition, and $\{z_{j,\ell}\}$ converges to $\{z_j\}$ as a set when $\ell \to +\infty$ (recall that $w_{n_{k_\ell}}$ is outer hence has no zero in $\mathbb{D}$). We label the $z_{j,\ell}$ so that, say $z_{j,\ell} \in \mathbb{T}$ for $1 \leq j \leq s_\ell$ and $z_{j,\ell} \notin \mathbb{T}$ for $s_\ell + 1 \leq j \leq N$. Define

$$P_{N,\ell}(\xi) = \Pi_{j=1}^N (\xi - z_{j,\ell}), \qquad Q_{N-s_\ell,\ell}(\xi) = \Pi_{j=s_\ell+1}^N (\xi - z_{j,\ell}),$$

and let us write $w_{n_{k_\ell}}(\xi) = u_\ell(\xi)\, P_{N,\ell}(\xi)$ where $u_\ell(\xi)$ is analytic in $\overline{\mathbb{C}} \setminus \overline{G}^{-1}$ and zero-free in $\mathcal{C}_\varepsilon$. By the maximum principle, $u_{j,\ell}$ converges to $w(z)/\Pi_{j=1}^N(z - z_j)$ locally uniformly in $\overline{\mathbb{C}} \setminus \overline{G}^{-1}$. Clearly,

$$\frac{\check{w}_n(\xi)}{w_n(\xi)} = \Pi_{l=1}^{s_\ell}\left(-\bar{z}_{l,\ell}\right)\, \xi^{-N}\, \frac{\check{u}_{j,\ell}(\xi)}{u_{j,\ell}(\xi)}\, \frac{\widetilde{Q}_{N-s_\ell,\ell}(\xi)}{Q_{N-s_\ell,\ell}(\xi)}, \tag{53}$$

where we observe that $b_{N-s_\ell} = \Pi_{l=1}^{s_\ell}\left(-\bar{z}_{l,\ell}\right)\, \widetilde{Q}_{N-s_\ell,\ell}/Q_{N-s_\ell,\ell}$ lies in B_{N-s_ℓ} and that $\check{u}_{j,\ell}/u_{j,\ell}$ is continuous and bounded independently of ℓ on $\mathcal{C}_\varepsilon$ as well as analytic in the interior of $\mathcal{C}_\varepsilon$. Now, put $\beta(\xi) = \xi$ and let us write

$$\mathbf{P}_+\left(\bar{b}_m^2 \bar{j} \check{w}_m / w_m\right) = b_{N-s_\ell}\mathbf{P}_+\left(\bar{\beta}^N \bar{b}_m^2 \bar{j} \check{u}_\ell / u_\ell\right) + \mathbf{P}_+\left(b_{N-s_\ell}\mathbf{P}_-\left(\bar{\beta}^N \bar{b}_m^2 \bar{j} \check{u}_\ell / u_\ell\right)\right). \tag{54}$$

Recalling that $\overline{b_m(1/\bar{z})} = 1/b_m(z)$ and $\overline{j(1/\bar{z})} = 1/j(z)$, we deduce from (4)

$$\mathbf{P}_+\left(\bar{\beta}^N \bar{b}_m^2 \bar{j} \check{u}_\ell / u_\ell\right)(z) = \frac{1}{2i\pi} \int_{\mathbb{T}} \frac{1}{b_m^2(\xi) j(\xi)} \frac{\check{u}_\ell(\xi)}{u_\ell(\xi)(\xi - z)} \frac{d\xi}{\xi^N}, \quad |z| < 1, \tag{55}$$

and by Cauchy's theorem we can deform the contour of integration to $\mathbb{T}_{1+\varepsilon}$ without changing the value of the integral:

$$\mathbf{P}_+\left(\bar\beta^N\bar b_m^2\bar j\check u_\ell/u_\ell\right)(z)=\frac{1}{2i\pi}\int_{\mathbb{T}_{1+\varepsilon}}\frac{1}{b_m^2(\xi)j(\xi)}\frac{\check u_\ell(\xi)}{u_\ell(\xi)(\xi-z)}\frac{d\xi}{\xi^N},\qquad |z|<1. \tag{56}$$

The integral in the right hand side of (56) is now bounded in modulus, independently of ℓ and $z\in\mathbb{D}$, because $|b_m|\geq 1$ and $|j|\geq 1$ on $\overline{\mathbb{C}}\setminus\mathbb{D}$, while $\varepsilon\leq|\xi-z|$ and $\check u_\ell/u_\ell$ is uniformly bounded on $\mathbb{T}_{1+\varepsilon}$. Thus, the first summand in the right hand side of (54) is bounded in L^∞, independently of ℓ, because $b_{N-s_\ell}\in B_{N-s_\ell}$. To see that the second summand is also bounded, we put $\Psi=\mathbf{P}_-\left(\bar\beta^N\bar b_m^2\bar j\check u_\ell/u_\ell\right)$ and we notice that $\|\Psi\|_\infty$ is bounded independently of ℓ because $\bar\beta^N\bar b_m^2\bar j\check u_\ell/u_\ell$ is unimodular on $\mathbb{T}$ and we just saw from (56) that $\|\mathbf{P}_+\left(\bar\beta^N\bar b_m^2\bar j\check u_\ell/u_\ell\right)\|_\infty$ is bounded independently of ℓ. Next, we observe that this second summand is $\mathbf{P}_+(b_{N-s_\ell}\Psi)$ and that it lies in $\mathcal{R}_{N-s_\ell-1,N-s_\ell}$ because for $\xi\in\mathbb{T}$ we have:

$$\frac{Q_{N-s_\ell}(\xi)}{\xi^{N-s_\ell}}\mathbf{P}_+(b_{N-s_\ell}\Psi)(\xi)=\overline{Q_{N-s_\ell}(1/\bar\xi)}\,\Psi(\xi)-\frac{Q_{N-s_\ell}(\xi)}{\xi^{N-s_\ell}}\mathbf{P}_-(b_{N-s_\ell}\Psi)(\xi),$$

and both summands on the right lie in $\bar H^{2,0}$ whence $Q_{N-s_\ell}\mathbf{P}_+(b_{N-s_\ell}\Psi)\in\mathcal{P}_{N-s_\ell-1}$. We now appeal to Grigoryan's theorem, saying that if $\mathbf{P}_-(\Phi)\in\mathcal{R}_{d-1,d}$ then $\|\mathbf{P}_-(\Phi)\|_\infty\leq cd\|\Phi\|_\infty$ for some absolute constant c, see [33, Eq. (6.1)]. As $\check{\mathbf{P}}_+(\Phi)=\mathbf{P}_-(\check\Phi)$ for any function Φ, it implies since the check operation preserves $\mathcal{R}_{N-s_\ell-1,N-s_\ell}$ and the L^∞ norm that

$$\|\mathbf{P}_+(b_{N-s_\ell}\Psi)\|_\infty=\|\mathbf{P}_-(\bar b_{N-s_\ell}\check\Psi)\|_\infty\leq c(N-s_\ell)\|\bar b_{N-s_\ell}\check\Psi\|_\infty=c(N-s_\ell)\|\Psi\|_\infty.$$

This achieves the proof. ∎

The authors conjecture that Theorem 6 carries over to Cauchy integrals of the form (35) where G is a so-called symmetric contour for the Green potential in $\mathbb{D}$ (*cf.* [40, Theorem 1] for details), the prototype of which is an analytic function with finitely many branchpoints of order greater than -1 in the disk. For such functions, more generally even if branchpoints have arbitrary order, it was proved in [17] that $\lim_{n\to\infty}d_\infty(f,\mathcal{R}_{n-1,n})^{1/n}=\exp\{-2/C\}$ where C is the condenser capacity of the pair $(\mathbb{T},G)$. The same n-th root estimate holds for $d_2(f,\mathcal{R}_{n-1,n})$ [8, Corollary 8], and more generally for the distance from f to $\mathcal{R}_{n-1,n}$ in L^p when $1\leq p\leq\infty$ [42]. Inequality (51) compares $d_2(f,\mathcal{R}_{n-1,n})$ and $d_\infty(f,\mathcal{R}_{n-1,n})$ in a much stronger sense, but still one may wonder if the factor $1/\sqrt{n+1}$ is really needed. In the special case of Markov functions, *i.e.* of Cauchy integrals of positive densities on a segment, the results in [7] show that this factor is in fact superfluous.

6. Linearized errors

Given $f\in\bar H^{2,0}$, $p_{n-1}\in\mathcal{P}_{n-1}$, $q_n\in\mathcal{P}_n$ and a (complex) weight function $w\in H^\infty$, the *linearized error* associated with p, q and w in problem RAB(n) is

$$\mathcal{L}(f,p_{n-1},q_n,w):=(q_nf-p_{n-1})w. \tag{57}$$

It is formally obtained from the error $f-p_{n-1}/q_n$ by chasing denominator q_n and multiplying by the weight. In applied sciences, problem RAB(n) and weighted variants thereof are of great importance to model time series as well as to identify linear dynamical systems,[2] *e.g.* in modal analysis of mechanical structures or in frequency analysis of microwave devices [22,24,27,35,30]. The importance of the L^2 norm in this context stems from its statistical interpretation as a variance. Because RAB(n) (or equivalently MA(n)) is a difficult non convex problem, several approaches to system identification in engineering have been based on linearization. Most popular in this connection are two closely related heuristics, namely the Steiglitz–McBride method [41,35,36] and the vector fitting method [21,11]. These are iterative procedures, first choosing $w = 1/\pi_n$ where $\pi_n \in \mathcal{P}_n$ is monic with no root on $\mathbb{T}$, then minimizing $\|\mathcal{L}(f, p_{n-1}, q_n, w)\|_2$ with respect to p_{n-1} as well as q_n, the latter being normalized so as to be monic[3] (which yields a convex problem). Subsequently, one replaces w by $1/q_n^o$, where q_n^o is the optimal q_n, and repeats the previous steps until some fixed point is reached. Such procedures are prompted by the easy observation that if $f \in R_{n-1,n}$, then the value of the problem is zero from the first iteration already. Accordingly, convergence was studied in a classical stochastic setting for system identification, where $f \in R_{n-1,n} \cap \bar{H}^{2,0}$ is perturbed by white noise (the noise is then constitutive of the model), but such heuristics do not converge in general when $f \notin R_{n-1,n}$ [37,25,36].

Our purpose here is not to discuss these techniques, nor to compare them with dedicated optimization algorithms [13,30], but rather to stress a link between the value of RAB(n) and the minimization of linearized errors.

We consider weights of the form $w = 1/\pi_n$ where $\pi_n \in \mathcal{P}_n$ is a polynomial having no root on $\mathbb{T}$. When minimizing the linearized error, Theorem 1 suggests a specific normalization for q_n: let us define

$$P_{\pi_n} := \{q_n \in \mathcal{P}_n : \ \|q_n/\pi_n\|_\infty = 1\}. \tag{58}$$

Then, the following result holds.

Theorem 7. *Let $f \in \bar{H}^{2,0}$ and $\pi_n \in \mathcal{P}_n$, with $\mathcal{Z}(\pi_n) \cap \mathbb{T} = \emptyset$. Then*

$$d_2(f, \mathcal{R}_{n-1,n}) = d_2(f, H_n^2) \geq \min_{\substack{q_n \in P_{\pi_n} \\ p_{n-1} \in \mathcal{P}_{n-1}}} \|\mathcal{L}(f, p_{n-1}, q_n, 1/\pi_n)\|_2. \tag{59}$$

Proof. We may assume without loss of generality that $\mathcal{Z}(\pi_n) \subset \overline{\mathbb{C}}\backslash\mathbb{D}$, for otherwise we can replace every linear factor $(z - a)$ of π_n for which $a \in \mathbb{D}$ by the linear factor $(1 - \bar{a}z)$ which has reflected zero across $\mathbb{T}$. This leaves $|\pi_n|$ unchanged on $\mathbb{T}$, and consequently does not affect the minimization of $\|\mathcal{L}(f, p_{n-1}, q_n, 1/\pi_n)\|_2$.

Now, arguing as we did to obtain (9), we find that $\pi_n \mathbf{P}_+(f q_n/\pi_n)$ is a polynomial of degree at most $n-1$. If we write $p_{n-1}(q_n, \pi_n)$ for this polynomial and take into account that $\mathcal{Z}(\pi_n) \cap \overline{\mathbb{D}} = \emptyset$, we see from Parseval's theorem that for fixed $q_n \in \mathcal{P}_n$ the criterion $\|\mathcal{L}(f, p_{n-1}, q_n, 1/\pi_n)\|$ gets minimized precisely when $p_{n-1} = p_{n-1}(q_n, \pi_n)$, so that

$$\min_{p_{n-1} \in \mathcal{P}_{n-1}} \|\mathcal{L}(f, p_{n-1}, q_n, 1/\pi_n)\|_2 = \left\|\mathbf{P}_-\left(f\frac{q_n}{\pi_n}\right)\right\|_2 = \|A_f(q_n/\pi_n)\|_2. \tag{60}$$

[2] For continuous time systems, rational approximation is performed on the imaginary axis rather than the circle. This, is equivalent to the present setting thanks to the isometry $f \mapsto \sqrt{2} f((z+1)/(z-1))/(z-1)$ mapping H^2 onto the Hardy space of $\{\mathrm{Re} z > 0\}$ while preserving rationality and the degree.

[3] What we describe here is the Steiglitz–McBride method, although we should mention that the criterion used is often a discretized version of $\|\mathcal{L}(f, p_{n-1}, q_n, 1/\pi_n)\|_2$ obtained from pointwise values on $\mathbb{T}$. The vector fitting method is essentially a rewriting of the Steiglitz–McBride procedure where rational functions are parametrized in pole-residue form.

Let

$$K_{\pi_n} := \{q_n/\pi_n : q_n \in P_{\pi_n}\}.$$

Since π_n has no zeros on $\overline{\mathbb{D}}$, it holds that $K_{\pi_n} \subset \mathcal{S}^\infty$, the unit sphere of H^∞. Identifying $\mathcal{P}_n$ with $\mathbb{C}^{n+1} \sim \mathbb{R}^{2n+2}$ by taking coefficients as coordinates, we see that K_{π_n} is homeomorphic to the Euclidean sphere $\mathbb{S}^{2n+1}$ via the map $q_n/\pi_n \mapsto q_n/\|q_n\|_2$ which is odd. Therefore K_{π_n} is a compact subset of $\mathcal{S}^\infty$ of genus $2n+2$, and by (16):

$$d_2(f, R_{n-1,n}) = d_2(f, H^2_n) \geq \min_{q_n/\pi_n \in K_{\pi_n}} \|A_f(q_n/\pi_n)\|_2 \tag{61}$$

which is (59) in view of (60). ■

It follows easily from a compactness argument that the minimum in the right hand side of (59) is attained. However, it not *a priori* obvious how to compute it for P_{π_n} is not convex. Numerically, this issue can be approached as follows. First, we assume without loss of generality that π_n has no roots in $\overline{\mathbb{D}}$, so that (60) holds (*cf.* proof of Theorem 7). Next, for $\xi \in \mathbb{T}$, let

$$P_{\pi_n,\xi} := \{q_n \in \mathcal{P}_n : \|q_n/\pi_n\|_\infty = 1, \ q_n(\xi) = \pi_n(\xi)\}.$$

Observe that $P_{\pi_n,\xi}$ is never empty when π_n has no zero on $\mathbb{T}$. Indeed, for small $\varepsilon > 0$, it holds that $|\pi_n(e^{i\theta})|^2 - |\varepsilon(e^{i\theta} - \xi)|^2 \geq 0$ hence, by Fejèr–Riesz factorization (see Lemma 4 to come), there is a polynomial q_n with $|q_n| \leq |\pi_n|$ on $\mathbb{T}$ and $|q_n(\xi)| = |\pi_n(\xi)|$. Thus, $q_n\pi_n(\xi)/q_n(\xi)$ lies in $P_{\pi_n,\xi}$. Clearly

$$K_{\pi_n} = \cup_{\xi,\zeta\in\mathbb{T}} \ \zeta \ P_{\pi_n,\xi}, \tag{62}$$

and multiplying q_n by $\zeta \in \mathbb{T}$ cannot change the value of $\|A_f(q_n/\pi_n)\|_2$. Therefore it holds that

$$\min_{\substack{q_n \in P_{\pi_n} \\ p_{n-1} \in \mathcal{P}_{n-1}}} \|\mathcal{L}(f, p_{n-1}, q_n, 1/\pi_n)\|_2 = \min_{\xi\in\mathbb{T}} \psi(\xi) \tag{63}$$

where the function $\psi(\xi)$ is given by (*cf.* (60))

$$\psi(\xi) = \min_{\substack{q_n \in P_{\pi_n,\xi} \\ p_{n-1} \in \mathcal{P}_{n-1}}} \|\mathcal{L}(f, p_{n-1}, q_n, 1/\pi_n)\|_2 = \min_{q_n \in P_{\pi_n,\xi}} \|A_f(q_n/\pi_n)\|_2. \tag{64}$$

Note that $\psi(\xi)$ can be computed as the solution of a convex problem for each ξ, because $P_{\pi_n,\xi}$ is a convex set and $\|A_f(q_n/\pi_n)\|_2$ a quadratic criterion. Granted this ability to evaluate ψ pointwise, we discuss below how to numerically estimate the minimum in (63).

Clearly ψ is the zero function when $f \in \mathcal{R}_{n-1,n}$, for if $f = p/q$ with $\deg q \leq n$ we may pick $q_n = q$ as minimizer in (64). The next lemma describes this minimizer in greater detail when $f \notin \mathcal{R}_{n-2,n-1}$.

Lemma 3. *Let $f \in \bar{H}^{2,0}$ and $\pi_n \in \mathcal{P}_n$, with $\mathcal{Z}(\pi_n) \cap \overline{\mathbb{D}} = \emptyset$. If $f \notin \mathcal{R}_{n-2,n-1}$, then the minimizing q_n in (64) is unique, has all its roots in $\overline{\mathbb{D}}$, and exact degree n.*

Proof. Assume that $q_{n,1}$ and $q_{n,2}$ are distinct minimizers, that is, $q_{n,1}, q_{n,2} \in P_{\pi_n,\xi}$ and $\|A_f(q_{n,1}/\pi_n)\|_2 = \|A_f(q_{n,2}/\pi_n)\|_2 = \psi(\xi)$. Put $q_{n,3} = (q_{n,1} + q_{n,2})/2 \in P_{\pi_n,\xi}$. By strict convexity of the L^2 norm, we get $A_f(q_{n,1}/\pi_n) = A_f(q_{n,2}/\pi_n)$ otherwise we would have that $\|A_f(q_{n,3}/\pi_n)\|_2 < \psi(\xi)$ which is absurd. Set $q = q_{n,1} - q_{n,2} \in \mathcal{P}_n$. Then $A_f(q/\pi_n) = 0$ implying by definition of A_f that $fq/\pi_n \in H^2$. A fortiori then $fq \in H^2$, and since $f \in \bar{H}^{2,0}$ we must have that fq is a polynomial of degree at most $n-1$, say p. Thus, $f = p/q$, and as q

has a root on $\mathbb{T}$ (namely ξ) the latter must be cancelled by a corresponding root of p. Altogether $f \in \mathcal{R}_{n-2,n-1}$, thereby showing the uniqueness part of the lemma. Let now $q_{n,\xi}$ be the unique minimizer and b a Blaschke product with poles in $\mathcal{Z}(q_{n,\xi}) \cap \overline{\mathbb{C}}\setminus\overline{\mathbb{D}}$. Then $bq_{n,\xi} \in \mathcal{P}_n$ has same modulus as q_n on $\mathbb{T}$, hence there is $\zeta \in \mathbb{T}$ such that $\zeta bq_{n,\xi} \in P_{\pi_n,\xi}$. Since ζb is a Blaschke product, reasoning as in (39) yields $\|A_f(\zeta bq_{n,\xi}/\pi_n)\|_2 \leq \|A_f(q_{n,\xi}/\pi_n)\|_2$ so that $\zeta bq_{n,\xi}$ is in turn a minimizer, hence is equal to $q_{n,\xi}$ by the uniqueness part just proved. Now, $q_{n,\xi} \not\equiv 0$ since $q_{n,\xi}(\xi) = \pi_n(\xi) \neq 0$, therefore $\zeta b = 1$. Thus, b must be a constant, that is to say there cannot be a zero of $q_{n,\xi}$ outside $\overline{\mathbb{D}}$. Finally, assume that $\deg q_{n,\xi} < n$. Then $q_{n,\xi}(z)z\bar{\xi}$ lies in $P_{\pi_n,\xi}$ and, since $z\bar{\xi}$ is a Blaschke product, it follows as before that $q_{n,\xi}z\bar{\xi}$ is a minimizer, hence it must be equal to $q_{n,\xi}$ by uniqueness. This contradiction achieves the proof. ■

We need a continuity property of the Fejèr–Riesz factorization that we could not ferret out in the literature. Write $\mathcal{T}_n$ for the space of trigonometric polynomials of degree at most n, *i.e.* sums of the form $\sum_{|k|\leq n} a_k e^{ik\theta}$. For fixed n, $\mathcal{P}_n$ and $\mathcal{T}_n$ have a natural topology induced by any norm.

Lemma 4. *To each nonzero $T \in \mathcal{T}_n$ such that $T \geq 0$ on $\mathbb{T}$, one can associate continuously a unique polynomial $q \in \mathcal{P}_n$ having no zero in $\mathbb{D}$ and such that $|q(e^{i\theta})|^2 = T(e^{i\theta})$ with $q(0) > 0$.*

Proof. Let $\mathcal{T}_n^+ \subset \mathcal{T}_n$ be the closed subset of trigonometric polynomials which are non-negative on $\mathbb{T}$. For $T \in \mathcal{T}_n^+$, existence of $q \in \mathcal{P}_n$ such that $|q|^2 = T$ on $\mathbb{T}$ is a classical result known after Fejèr and Riesz [38, Section 53]. Since $|z - a| = |1 - z\bar{a}|$ for $z \in \mathbb{T}$, clearly q may be chosen zero free in $\mathbb{D}$ if $T \not\equiv 0$. Then, $q_{|\mathbb{D}}$ is outer in H^∞, for it has no zero and it extends analytically across $\mathbb{T}$ [14, Chapter II, Theorems 6.2 and 6.3]. Thus, formula (5) shows that q is uniquely defined by $\log|q| = \log T/2$, therefore also by T. Moreover, each coefficient of q is a continuous function of $\log T \in L^1$, because the k-th coefficient is just the derivative $q^{(k)}(0)/k!$ and we may differentiate (5) under the integral sign. To achieve the proof, we establish that $T \mapsto \log T$ is continuous from $\mathcal{T}_n^+\setminus\{0\}$ into L^1.

First, *we claim* that $T \mapsto \|\log T\|_1$ is continuous from $\mathcal{T}_n^+\setminus\{0\}$ into $\mathbb{R}$. To see this, it is enough to show that if $T^{\{k\}}$ tends to T in $\mathcal{T}_n^+\setminus\{0\}$ as $k \to \infty$, then $\|\log T^{\{k_\ell\}}\|_1$ tends to $\|\log T\|_1$ for some subsequence $T^{\{k_\ell\}}$. By the first part of the proof, we can write $T^{\{k\}} = |q^{\{k\}}|^2$ with

$$q^{\{k\}}(z) = q^{\{k\}}(0)\, \Pi_{l=1}^n (1 - z\, a_l^{\{k\}}), \quad a^{\{k\}} \in \overline{\mathbb{D}},$$

where multiplicities are counted by repetition and the ordering of the roots for each k is arbitrary. Note that $|q^{\{k\}}(0)|$ is bounded, since by the Schwarz inequality:

$$|q^{\{k\}}(0)| = \left|\frac{1}{2\pi}\int_0^{2\pi} q^{\{k\}}(e^{i\theta})d\theta\right| \leq \|q^{\{k\}}\|_2 = \left\|T^{\{k\}}\right\|_1^{1/2}.$$

Therefore, there is a subsequence $q^{\{k_\ell\}}$ such that $q^{\{k_\ell\}}(0)$ converges to $c \in \mathbb{C}$ and $a_j^{\{k_\ell\}}$ converges to $a_j \in \overline{\mathbb{D}}$ for each $j \in \{1, \ldots, n\}$. If we let

$$q(z) = c\, \Pi_{j=1}^n\, (1 - z\, a_j),$$

then clearly $|q^{\{k_\ell\}}|^2$ converges to $|q|^2$ almost everywhere on $\mathbb{T}$, so that necessarily $|q|^2 = T$. In particular, we have that $c \neq 0$ otherwise T would be identically zero, a contradiction. Now, since log turns products into sums, we are left to show that if $b^{\{k\}} \to b$ in $\overline{\mathbb{D}}$, then

$$\lim_{k\to+\infty}\int_0^{2\pi} \left|\log|1 - e^{i\theta}b^{\{k\}}|\right| d\theta = \int_0^{2\pi} \left|\log|1 - e^{i\theta}b|\right| d\theta. \tag{65}$$

When $|b| < 1$ relation (65) is obvious. If $|b| = 1$, we may assume by rotational symmetry that $b = 1$ and $b^{\{k\}} \in [0, 1]$, in which case (65) follows by dominated convergence from the observation that $|1 - b^{\{k\}}e^{i\theta}| \geq |\sin\theta|$ for $|\theta| \leq \pi/2$. *This proves the claim.*

Now, if $T^{\{k\}}$ tends to T in $\mathcal{T}_n^+\backslash\{0\}$, it is plain that $\log T^{\{k\}}$ converges to $\log T$ almost everywhere on $\mathbb{T}$ and by the previous claim the L^1-norm of the limit is the limit of the L^1-norms. Thus, the desired L^1-convergence of $\log T^{\{k\}}$ to $\log T$ follows from Egoroff's theorem [39, Chapter 3, Example 17]. ■

With the help of Lemmas 3 and 4, we now prove that ψ is continuous:

Lemma 5. *Let $f \in \bar{H}^{2,0}$ and $\pi_n \in \mathcal{P}_n$, with $\mathcal{Z}(\pi_n) \cap \overline{\mathbb{D}} = \emptyset$. Then the map ψ defined by (64) is continuous on $\mathbb{T}$.*

Proof. If $f \in \mathcal{R}_{n-1,n}$, we mentioned that $\psi \equiv 0$ already. Otherwise, dwelling on Lemma 3, let $q_{n,\xi}$ indicate the unique minimizer in the last term of (64). By definition of $P_{\pi_n,\xi}$ we have that $|q_{n,\xi}| \leq |\pi_n|$ on $\mathbb{T}$ and that $q_{n,\xi}(\xi) = \pi_n(\xi)$, in particular $q_{n,\xi}$ is bounded independently of ξ. Thus, from any convergent sequence $\xi_k \to \xi$ on $\mathbb{T}$, we can extract a subsequence ξ_{k_ℓ} for which $q_{n,\xi_{k_\ell}}$ converges uniformly to some $q \in \mathcal{P}_n$, and passing to the limit we see that $q \in P_{\pi_n,\xi}$. Given $\varepsilon > 0$, we can pick the sequence ξ_k so that

$$\lim_{k\to+\infty} \psi(\xi_k) = l \leq \liminf_{\zeta\to\xi} \psi(\zeta) + \varepsilon,$$

and by continuity of $q_n \mapsto \|A_f(q_n/\pi_n)\|_2$ from $\mathcal{P}_n$ into $\mathbb{R}$ we get that

$$\liminf_{\zeta\to\xi} \psi(\zeta) + \varepsilon \geq l = \lim_{\ell\to\infty} \|A_f(q_{n,\xi_{k_\ell}}/\pi_n)\|_2 = \|A_f(q/\pi_n)\|_2 \geq \psi(\xi). \tag{66}$$

Since $\varepsilon > 0$ was arbitrary, we conclude that ψ is lower semi-continuous. To see that ψ is in fact continuous, it is enough to establish the *following claim*: to each $\xi \in \mathbb{T}$ and $\varepsilon > 0$, there is $\eta > 0$ such that $|\xi - \zeta| < \eta$ implies existence of $q_\zeta \in P_{\pi_n,\zeta}$ with $\|q_{n,\xi} - q_\zeta\|_\infty < \varepsilon$. Indeed, if the claim holds, we get from (64) that when $|\xi - \zeta| < \eta$:

$$\begin{aligned}\psi(\zeta) &\leq \|A_f(q_\zeta/\pi_n)\|_2 \leq \|A_f(q_{n,\xi}/\pi_n)\|_2 + \|A_f((q_\zeta - q_{n,\xi})/\pi_n)\|_2 \\ &\leq \psi(\xi) + \varepsilon\|f/\pi_n\|_2,\end{aligned}$$

and since ε was arbitrary we conclude that $\limsup_{\zeta\to\xi} \psi(\zeta) \leq \psi(\xi)$ whence ψ is indeed continuous in view of (66).

To establish the claim, observe from Lemma 4 since $|\pi_n|^2 - |q_{n,\xi}|^2$ is a non-negative trigonometric polynomial of degree at most n on $\mathbb{T}$ that

$$|q_{n,\xi}(e^{i\theta})|^2 + |\kappa_{n,\xi}(e^{i\theta})|^2 = |\pi_n(e^{i\theta})|^2 \tag{67}$$

where $\kappa_{n,\xi} \in \mathcal{P}_n$ has no root in $\mathbb{D}$ and is uniquely defined by (67) together with the normalization $\kappa_{n,\xi}(0) > 0$. As $\pi_n(\xi) = q_{n,\xi}(\xi)$, we can write $\kappa_{n,\xi}(z) = (z - \xi)Q_{n-1}(z)$, and for $\zeta \in \mathbb{T}$ we set more generally:

$$\kappa_{n,\zeta}(z) = \frac{(z - \zeta)Q_{n-1}(z)}{\lambda_\zeta}, \qquad \lambda_\zeta = \sup_{z\in\mathbb{T}} |(z - \zeta)Q_{n-1}(z)/\pi_n(z)|. \tag{68}$$

Clearly $|\kappa_{n,\zeta}|^2 \leq |\pi_n|^2$ on $\mathbb{T}$ so that, by Lemma 4, there is a unique $P_{n,\zeta} \in \mathcal{P}_n$ having no root in $\mathbb{D}$ and meeting $P_{n,\zeta}(0) > 0$ such that

$$|P_{n,\zeta}(e^{i\theta})|^2 + |\kappa_{n,\zeta}(e^{i\theta})|^2 = |\pi_n(e^{i\theta})|^2. \tag{69}$$

Since $q_{n,\xi}$ has all its roots in $\mathbb{D}$ and exact degree n by Lemma 3, it follows from (67) and the uniqueness part of Lemma 4 that $P_{n,\xi} = \bar{c}\widetilde{q}_{n,\xi}$ where $c \in \mathbb{T}$ is such that $cq_{n,\xi}$ has positive leading coefficient. Moreover, it is easily checked from (68) that $|\kappa_{n\zeta}|^2$ is arbitrary close to $|\kappa_{n,\xi}|^2$ in $\mathcal{T}_n$ if $|\xi - \zeta|$ is sufficiently small. Therefore, from (69) and the continuity property asserted by Lemma 4, we deduce that $P_{n,\zeta}$ is arbitrary close to $\bar{c}\widetilde{q}_{n,\xi}$ in $\mathcal{P}_n$ if $|\xi - \zeta|$ is sufficiently small. Consequently $Q_\zeta = \bar{c}\widetilde{P}_{n,\zeta}$ is arbitrary close to $q_{n,\xi}$ when $|\xi - \zeta|$ is small enough. Now, by (69) and the definition of Q_ζ, it holds that $|Q_\zeta| = |P_{n,\zeta}| \le |\pi_n|$ on $\mathbb{T}$, and also that $|Q_\zeta(\zeta)| = |P_{n,\zeta}(\zeta)| = |\pi_n(\zeta)|$ because $\kappa_{n,\zeta}(\zeta) = 0$ by construction. Hence, $q_\zeta = (\pi_n(\zeta)/Q_\zeta(\zeta))Q_\zeta$ lies in $P_{\pi_n,\zeta}$, and since $Q_\zeta(\zeta) \to q_{n,\xi}(\xi) = \pi_n(\xi)$ as $\zeta \to \xi$ we have that $(\pi_n(\zeta)/Q_\zeta(\zeta)) \to 1$ when $\zeta \to \xi$. Thus, just like Q_ζ, the polynomial q_ζ is arbitrary close to $q_{n,\xi}$ when $|\xi - \zeta|$ is small enough, which *proves the claim*. ■

To estimate the right hand side (63), it remains to minimize $\psi(\xi)$ over $\xi \in \mathbb{T}$, which can be numerically performed by dichotomy because $\mathbb{T}$ is compact and 1-dimensional while ψ is continuous by Lemma 5.

A natural question is whether the lower bound (59) can be sharp. The answer is no except in the trivial case where $f \in \mathcal{R}_{n-1,n}$:

Proposition 3. *Assumptions and notations as in Theorem 7, it holds if $f \notin \mathcal{R}_{n-1,n}$ that*

$$d_2(f, \mathcal{R}_{n-1,n}) > \min_{\substack{q_n \in P_{\pi_n} \\ p_{n-1} \in \mathcal{P}_{n-1}}} \|\mathcal{L}(f, p_{n-1}, q_n, 1/\pi_n)\|_2 \tag{70}$$

Proof. As in the proof of Theorem 7, we may assume that $\mathcal{Z}(\pi_n) \subset \overline{\mathbb{C}}\backslash\mathbb{D}$ and then, by (60), the right hand side of (70) is equal to $\min_{q_n \in P_{\pi_n}} \|A_f(q_n/\pi_n)\|_2$. Let $q_{n,0}$ be a minimizer of the latter, and $b_{n,1} = q_{n,1}/\widetilde{q}_{n,1}$ a minimizing Blaschke product in (12). Multiplying $q_{n,1}$ and $\widetilde{q}_{n,1}$ by a real constant, we may assume $q_{n,1} \in P_{\pi_n}$. We can also multiply $q_{n,0}$ by a unimodular constant so that, using (62), there is $\xi_0 \in \mathbb{T}$ for which $q_{n,0} \in P_{\pi_n,\xi_0}$. Now, if (70) is an equality, we get by definition of $q_{n,0}$, $q_{n,1}$ that

$$\begin{aligned} d_2(f, \mathcal{R}_{n-1,n}) &= \|\mathbf{P}_-(f q_{n,0}/\pi_n)\|_2 \le \|\mathbf{P}_-(f q_{n,1}/\pi_n)\|_2 \\ &= \|\mathbf{P}_-\Big(f(q_{n,1}/\widetilde{q}_{n,1})(\widetilde{q}_{n,1}/\pi_n)\Big)\|_2 \le \|\mathbf{P}_-(f q_{n,1}/\widetilde{q}_{n,1})\|_2 \\ &= d_2(f, \mathcal{R}_{n-1,n}) \end{aligned} \tag{71}$$

where we used in the second inequality that $\widetilde{q}_{n,1}/\pi_n \in \mathcal{S}^\infty$. Consequently equality holds throughout (71), implying in particular that $|\widetilde{q}_{n,1}| = |\pi_n|$ on $\mathbb{T}$. Thus, as both polynomial have no root in $\mathbb{D}$ their ratio is a unimodular constant, and renormalizing $q_{n,1}$ if necessary we may assume that $\widetilde{q}_{n,1} = \pi_n$. Then $c = \widetilde{q}_{n,1}(\xi_0)/q_{n,1}(\xi_0)$ is a unimodular constant such that $cq_{n,1} \in P_{\pi_n,\xi_0}$, and by the uniqueness part in Lemma 3 we see that the first inequality in (71) can be an equality only if $q_{n,0} = cq_{n,1}$. Altogether $q_{n,0}/\pi_n = cb_{n,1}$ is in turn an optimal Blaschke product in (12). This optimality entails that [6, Theorem 8.2]

$$A_f^* A_f(q_{n,0}/\pi_n) = \mathbf{P}_+\Big(|A_f(q_{n,0}/\pi_n)|^2 q_{n,0}/\pi_n\Big).$$

Consequently, letting $\langle u, v\rangle_{L^2} = \mathrm{Re}\langle u, v\rangle$, it holds for all $q_n \in P_{\pi_n,\xi_0}$ that

$$\begin{aligned} &\langle A_f(q_{n,0}/\pi_n), A_f(q_{n,0}/\pi_n) - A_f(q_n/\pi_n)\rangle_{L^2} \\ &\quad = \langle A_f^* A_f(q_{n,0}/\pi_n), q_{n,0}/\pi_n - q_n/\pi_n\rangle_{L^2} \end{aligned}$$

$$= \langle \mathbf{P}_+\Big(|A_f(q_{n,0}/\pi_n)|^2 q_{n,0}/\pi_n\Big), q_{n,0}/\pi_n - q_n/\pi_n\rangle_{L^2}$$
$$= \langle |A_f(q_{n,0}/\pi_n)|^2 q_{n,0}/\pi_n, q_{n,0}/\pi_n - q_n/\pi_n\rangle_{L^2}$$
$$= \langle |A_f(q_{n,0}/\pi_n)|^2, 1 - q_n/q_{n,0}\rangle_{L^2}, \tag{72}$$

where we used in the third line that $(q_{n,0} - q_n)/\pi_n \in H^2$ to get rid of $\mathbf{P}_+$ and in the last line that $\overline{q_{n,0}/\pi_n} = \pi_n/q_{n,0}$ on $\mathbb{T}$.

On the one hand, we see that (72) is non-negative for all $q_n \in P_{\pi_n,\xi_0}$ because $|q_n/q_{n,0}| = |q_n/\pi_n| \le 1$ on $\mathbb{T}$. In fact, it can be made strictly positive: indeed, $A_f(q_{n,0}/\pi_n)$ is not the zero function for $f \notin \mathcal{R}_{n-1,n}$, and as discussed before (62) one can pick $q_n \in P_{\pi_n,\xi_0}$ such that $|q_n/q_{n,0}|(\xi) < 1$ for $\xi \ne \xi_0$ on $\mathbb{T}$. On the other hand, the fact that $q_{n,0}$ is a minimizer in the convex problem (64) (where ξ is set to ξ_0) implies that (72) is nonpositive for all $q_n \in P_{\pi_n,\xi_0}$ [10, Proposition 5.23], a contradiction which concludes the proof. ■

7. Numerical results

In order to study how effective the bounds given by Theorem 4, Corollary 1 and Theorem 7, we wrote a prototype implementation in each case and ran it on a few examples. We report in this section the results obtained on the following set of functions $f \in \bar{H}^{2,0}$. For each of them, we consider the problem of best H^2 approximation by a rational function in $\mathcal{R}_{n-1,n}$ with $n = 4$.

- **Example 1**: $f : z \mapsto \log((10z - 9)/(10z + 9))$.
- **Example 2**: f is a rational function of degree 5.
- **Example 3**: f is a rational function with 20 poles that have been randomly and uniformly chosen inside the unit disk.
- **Example 4**: f is a rational function with 20 poles that have been randomly and uniformly chosen inside the disk of radius 0.2 centered at the origin.
- **Example 5**: f is a rational function with 20 poles that have been randomly and uniformly chosen inside the annulus of radii 18/20 and 19/20 centered at the origin.
- **Example 6**: f is fairly close to a rational function of degree 4. Namely, we chose a rational function g of degree 4 and then we obtained f by perturbing each Fourier coefficients of g with a small noise of relative error bounded by 0.01.
- **Example 7**: $f : z \mapsto \exp(-\mathrm{i}/(z - 0.9\mathrm{i})) - 1$.

These examples have been chosen so as to exhibit different kind of singularities inside the disk, which may or may not be close to the unit circle, in order to cover various situations.

Before discussing the results, let us say a few words on the implementation. We are using Matlab R2011b. For each example, f is actually approximated by a truncated Fourier series $\widehat{f}$. The order of truncation is chosen so as to ensure that f and $\widehat{f}$ agree to at least 40 bits on the unit circle. The bounds of Theorem 4 and Corollary 1 are computed as described in Section 5.2: indeed, since $\widehat{f}$ is a truncated Fourier series, it can be written $\widehat{f} = p/q$ where $p \in \mathcal{P}_{N-1}$ and $q = z^N$. It turns out that, when q is a power of z, the construction in Section 5.2 gets simpler since $\widetilde{q} = 1$, whence Bezout relation is just $a\widetilde{q} + bq = 1$ with $a = 1$ and $b = 0$. We handled all examples using this technique, even though in examples 2 to 5 the number N becomes quite large (up to 1500) and it would have been more efficient (but would also have required more implementation) to forget about $\widehat{f}$ and to apply the construction in Section 5.2 to the original f. Anyway, even for fairly large N, we obtained our results in a few seconds on an Intel Xeon at 2.67 GHz, with 4 GB of memory. When computing the bound of Theorem 4, the norms $\|v_j\|_\infty$ are estimated by sampling v_j at 8000 points evenly distributed on the unit circle.

Regarding the bound of Theorem 7, its computation reduces to finding the minimum of $\psi(\xi)$, for $\xi \in \mathbb{T}$ (see (63) and (64)) as explained in Section 6. For a given ξ, we evaluate $\psi(\xi)$ by solving a convex optimization probLemma For this purpose we use CVX, a package for specifying and solving convex programs within Matlab [19,18]. More precisely, we precompute $A_f(z^{j-1}/\pi_n)$ $(j = 1 \ldots n+1)$ in the Fourier basis, *i.e.* we compute a $N \times (n+1)$ matrix $M = (m_{ij})$ such that $A_f(z^{j-1}/\pi_n) = \sum_{i=1}^{N} m_{ij}\, z^{-i}$. Therefore, if $q_n = \sum_{j=1}^{n+1} a_j\, z^{j-1}$, we have that

$$\|A_f(q_n/\pi_n)\|_2^2 = \left\| M \begin{pmatrix} a_1 \\ \vdots \\ a_{n+1} \end{pmatrix} \right\|_2^2 .$$

Since N is much larger than n, it is convenient to compute a decomposition $M = QR$, where Q is orthogonal and R is upper-triangular. Since $\|Mv\|_2 = \|Rv\|_2$ for all vectors v and only the first $n+1$ rows of R are non-zero, we end up handling a $(n+1) \times (n+1)$ matrix instead of M. Now,

$$\psi(\xi) = \min_{q_n \in P_{\pi_n,\xi}} \left\| R \begin{pmatrix} a_1 \\ \vdots \\ a_{n+1} \end{pmatrix} \right\|_2 .$$

The set $P_{\pi_n,\xi}$ is convex, but described by infinitely many constraints. Therefore, we consider a set $\mathbb{T}_1$ of 50 points regularly spaced on the unit circle and we define

$$P^{(1)}_{\pi_n,\xi} = \left\{ q_n = \sum_{j=1}^{n+1} a_j\, z^{j-1} : q_n(\xi) = \pi_n(\xi), \text{ and } \forall \zeta \in \mathbb{T}_1, |q_n(\zeta)| \le \pi_n(\zeta) \right\}.$$

We first use CVX to solve our minimization problem subject to $q_n \in P^{(1)}_{\pi_n,\xi}$. This gives an optimal polynomial $q_n^{(1)}$. Next, we construct a set $\mathbb{T}_2$ by adding to $\mathbb{T}_1$ the points of $\mathbb{T}$ where $|q_n^{(1)}/\pi_n|$ reaches a local maximum. We then use CVX to solve our minimization problem subject to $q_n \in P^{(2)}_{\pi_n,\xi}$, where

$$P^{(2)}_{\pi_n,\xi} = \left\{ q_n = \sum_{j=1}^{n+1} a_j\, z^{j-1} : q_n(\xi) = \pi_n(\xi), \forall \zeta \in \mathbb{T}_2, |q_n(\zeta)| \le \pi_n(\zeta) \right\}.$$

This gives a new optimal polynomial $q_n^{(2)}$, and we repeat the process until we reach a step k where $\max_{\zeta \in \mathbb{T}} |q_n^{(k)}(\zeta)/\pi_n(\zeta)| - 1$ falls below the level of numerical errors produced by CVX.

It is worth pointing out that, although sufficient in most cases to get an idea of the numerical value of $\psi(\xi)$, this procedure yields no certified estimate of the bound in Theorem 7. Actually, CVX is a user-friendly generic software, able to tackle many types of convex optimization problems with a powerful syntax. However, it offers little control on the difference between the true mathematical solution and the numerical estimate thereof. Moreover, when too many constraints enter the game, it quickly yields no solution at all. In addition, it is probably much slower than would be a dedicated tool to solve that particular convex problem. The point we want here to make is that accurately estimating the bound in Theorem 7 (*i.e.* aiming at more than a prototypical illustration of the content of the paper) requires further work.

The numerical results proper are reported in Table 1 where the second column is $\frac{M_n(f)}{\sqrt{n+1}}$ (lower bound given by Theorem 4), the third column is $\frac{Q_n(f)}{\sqrt{n+1}}$ (lower bound given by Corollary 1), the fourth and fifth columns are $\min_{q_n \in P_{\pi_n}, p_{n-1} \in \mathcal{P}_{n-1}} \|\mathcal{L}(f, p_{n-1}, q_n, 1/\pi_n)\|_2$ (lower bound given by Theorem 7) for two different choices of π_n.

For an appraisal of the sharpness of our results, we also ran RARL2,[4] a software tool that tries to compute a solution to problem RAB(n). RARL2 looks for local minima of the criterion in a fairly systematic way, and returns the best approximant it could find. As a consequence, it gives an upper bound for the value of problem RAB(n) which is likely to be tight and therefore interesting to compare with our lower bounds. The error $\|f - r\|_2$ generated by the candidate best approximant r computed by RARL2 is reported in the last column.

The bound given by Theorem 7 has the advantage of allowing the user to choose a weight π_n which offers extra-flexibility to try to improve the estimate. Yet, it is not obvious how to pick π_n in general. The simplest choice is $\pi_n \equiv 1$ (reported in the fourth column of the table). Another, appealing possibility is to put $\pi_n = \widetilde{q}_n^*$ where ${q_n}^*$ is the denominator of the rational function computed by RARL2 (since ${q_n}^*$ has all its poles inside the disk, π_n has all its poles outside, as required). The corresponding results are reported in the fifth column of the table.

Table 1
Numerical results.

Example	Bound of Theorem 4	Bound of Corollary 1	Bound of Theorem 7 with $\pi_n = 1$	Bound of Theorem 7 with $\pi_n = \widetilde{q}_n^*$	RARL2
1	2.884744e−3	2.887532e−3	4.04e−3	10.8e−3	11.5e−3
2	7.731880e−2	7.732037e−2	12.4e−2	24.3e−2	24.72e−2
3	2.459346	2.470149	2.286	0.258	16.6907
4	1.234503	1.234861	1.94	1.8	6.5721
5	47.26312	47.30424	2.14	N/A	178.3152
6	2.894007e−3	2.894380e−3	9.62e−3	12.46e−3	12.5e−3
7	1.780707e−4	1.782276e−4	0.7977e−4	6.3409e−4	6.3742e−4

As can be seen from the table, the refinement of Corollary 1 with respect to Theorem 4 is almost negligible on all examples. The bound given by Theorem 7 with $\pi_n = 1$ is better than Theorem 4 and Corollary 1 in 4 cases out of 7, but not overly so. Considering that the computation time is generally much longer for this bound, this improvement can be considered as rather expensive. In contrast, when $\pi_n = \widetilde{q}_n^*$, the bound of Theorem 7 becomes fairly sharp in cases 1,2,6, and 7, which is encouraging. This better bound comes at the cost of a longer computation time though, mostly because the convex optimization problems involved with this choice of π_n seem more difficult to solve. In Example 5, for instance, we were not able to obtain reliable results from CVX. However, it also appears that choosing $\pi_n = \widetilde{q}_n^*$ is not *always* best, and it would be quite interesting to further understand which π_n are efficient in this respect.

References

[1] V.M. Adamjan, D.Z. Arov, M.G. Krein, Analytic properties of Schmidt pairs for a Hankel operator and the generalized Schur–Takagi problem, Math. USSR-Sb. 15 (1971) 31–73.

[2] L. Baratchart, Existence and generic properties for L^2 approximants of linear systems, IMA J. Math. Control Identif. 3 (1986) 89–101.

[3] L. Baratchart, On the topological structure of inner functions and its use in identification of linear systems, in: B. Bonnard, B. Bride, J.P. Gauthier, I. Kupka (Eds.), Analysis of Controlled Dynamical Systems, in: Progress in Systems and Control Theory, vol. 8, Birkhäuser, Boston, 1991, pp. 51–59.

[4] http://www-sop.inria.fr/apics/RARL2/rarl2.html.

[4] L. Baratchart, Rational and meromorphic approximation in L^p of the circle: system-theoretic motivations, critical points and error rates, in: N. Papamichael, S. Ruscheweyh, E.B. Saff (Eds.), Proceedings of the Third CMFT Conference—Computational Methods and Function Theory 1997, in: Series in Approximations and Decompositions, vol. 11, World Scientific Publishing Company, 1999, pp. 45–78.

[5] L. Baratchart, F. Mandrèa, E.B. Saff, F. Wielonsky, 2-D inverse problems for the Laplacian: a meromorphic approximation approach, J. Math. Pures Appl. 86 (2006) 1–41.

[6] L. Baratchart, F. Seyfert, An L^p analog of AAK theory for $p \geq 2$, J. Funct. Anal. 191 (2002) 52–122.

[7] L. Baratchart, H. Stahl, F. Wielonsky, Asymptotic uniqueness of best rational approximants of given degree to Markov functions in L^2 of the circle, Constr. Approx. 17 (2001) 103–138.

[8] L. Baratchart, M. Yattselev, H. Stahl, Weighted extremal domains and best rational approximation, Adv. Math. 229 (2012) 357–407.

[9] H. Brezis, Functional Analysis, Sobolev Spaces and Partial Differential Equations, Springer, 2011.

[10] F. Demengel, G. Demengel, Espaces Fonctionnels, Utilisation dans la Résolution des Équations aux Dérivées Partielles, EDP Sciences, 2007.

[11] D. Deschrijver, B. Haegeman, T. Dhaene, Orthonormal vector fitting: a robust macromodeling tool for rational approximation of frequency domain responses, IEEE Trans. Adv. Packag. 30 (2007) 216–225.

[12] V.D. Erohin, On the best approximation of analytic functions by rational functions with free poles, Dokl. Akad. Nauk SSSR 128 (1959) 29–32. (in Russian).

[13] P. Fulcheri, M. Olivi, Matrix rational H^2 approximation: a gradient algorithm based on Schur analysis, SIAM J. Control Optim. 36 (1998) 2103–2127.

[14] J. Garnett, Bounded Analytic Functions, in: Pure and Applied Mathematics, vol. 96, Academic Press, 1981.

[15] K. Glover, All optimal Hankel-norm approximations of linear multivariable systems and their L^∞—error bounds, Internat. J. Control 39 (1984) 1115–1193.

[16] I.C. Gohberg, M.G. Krein, Introduction to the Theory of Linear Nonselfadjoint Operators in Hilbert Space, in: Translations of Mathematical Monographs, vol. 18, American Mathematical Society, 1969.

[17] A.A. Gonchar, E.A. Rakhmanov, Equilibrium distributions and the degree of rational approximation of analytic functions, Mat. Sb. 134 (176) (1987) 306–352. English translation in Mathematics of the USSR-Sbornik 62 (1989) 305–348.

[18] M. Grant, S. Boyd, Graph implementations for nonsmooth convex programs, in: V. Blondel, S. Boyd, H. Kimura (Eds.), Recent Advances in Learning and Control, in: Lecture Notes in Control and Information Sciences, vol. 371, Springer, 2008, pp. 95–110. http://stanford.edu/~boyd/graph_dcp.html.

[19] M. Grant, S. Boyd, CVX: MATLAB software for disciplined convex programming, version 2.0 beta, 2013. http://cvxr.com/cvx.

[20] V. Guillemin, A. Pollack, Differential Topology, Prentice-Hall, 1974.

[21] B. Gustavsen, A. Semlyen, Rational approximation of frequency domain responses by vector fitting, IEEE Trans. Power Deliv. 14 (1999) 1052–1061.

[22] E.J. Hannan, M. Deistler, The Statistical Theory of Linear Systems, Wiley, 1988.

[23] B. Hanzon, R. Peeters, Balanced parametrizations of stable SISO all-pass systems in discrete time, Math. Control Signals Systems 13 (2000) 240–276.

[24] L. Hermans, H. Van der Auweraer, Modal testing and analysis of structures under operational conditions: industrial applications, Mech. Syst. Signal Process. 13 (1999) 193–216.

[25] S. Lefteriu, A.C. Antoulas, On the convergence of the vector fitting algorithm, IEEE Trans. Microw. Theory Tech. 61 (4) (2013) 1435–1443.

[26] A.L. Levin, The distribution of poles of rational functions of best approximation and related questions, Math. USSR-Sb. 9 (1969) 267–274.

[27] L. Ljung, System Identification: Theory for the User, Prentice-Hall, 1987.

[28] N.K. Nikolskii, Treatise on the Shift Operator, in: Grundlehren der Mathematischen Wissenschaften, vol. 273, Springer, 1986.

[29] N.K. Nikolskii, Operators, Functions, and Systems: An Easy Reading. Volumes I & II, in: Mathematical Surveys and Monographs, vol. 92–93, American Mathematical Society, 2002.

[30] M. Olivi, F. Seyfert, J.P. Marmorat, Identification of microwave filters by analytic and rational H^2 approximation, Automatica 49 (2013) 317–325.

[31] O.G. Parfenov, Estimates of the singular numbers of the Carleson imbedding operator, Mat. Sb. 131 (173) (1986) 501–518. English translation in Mathematics of the USSR-Sbornik 59 (1988) 497–514.

[32] J.R. Partington, Interpolation, Identification, and Sampling, in: London Mathematical Society Monographs (New Series), vol. 17, Oxford University Press, 1997.

[33] V.V. Peller, Hankel Operators and their Applications, Springer, 2003.

[34] V.A. Prokhorov, On L^p-generalization of a theorem of Adamjan, Arov, and Krein, J. Approx. Theory 116 (2002) 380–396.
[35] P.A. Regalia, Adaptive IIR Filtering in Signal Processing and Control, Marcel Dekker, 1995.
[36] P.A. Regalia, M. Mboup, Undermodeled adaptive filtering: an a priori error bound for the Steiglitz–McBride method, IEEE Trans. Circuits Syst. II 43 (1996) 105–116.
[37] P.A. Regalia, M. Mboup, M. Ashari, On the existence of stationary points for the Steiglitz–McBride algorithm, IEEE Trans. Automat. Control 42 (1997) 1592–1596.
[38] F. Riesz, B. Sz.-Nagy, Functional Analysis, Dover, 1990.
[39] W. Rudin, Real and Complex Analysis, Mc-Graw-Hill, 1982.
[40] H. Stahl, The structure of extremal domains associated with an analytic function, Complex Variables Theory Appl. 4 (1985) 339–356.
[41] K. Steiglitz, L.E. McBride, A technique for the identification of linear systems, IEEE Trans. Automat. Control 10 (1965) 461–464.
[42] V. Totik, A note on rational L^p approximation on Jordan curves, Comput. Methods Funct. Theory 13 (2013) 425–431.
[43] M. Yu, Y. Wang, Synthesis and beyond, IEEE Microw. Mag. 12 (2011) 62–76.
[44] E. Zeidler, Nonlinear Functional Analysis and its Applications III, Springer-Verlag, 1985.
[45] E. Zeidler, Nonlinear Functional Analysis and its Applications II/A, Springer-Verlag, 1990.