Received: 24 November 2020

DOI: 10.1002/mma.7267

RESEARCH ARTICLE

WILEY

Pseudohyperbolic distance and n-best rational approximation in $\mathbb{H}^2$ space

Yan-Bo Wang[1] | Tao Qian[2]

[1]Faculty of Mathematics and Computer Science, Wuhan Textile University, Wuhan, China

[2]Macau Center for Mathematical Sciences, Macau University of Science and Technology, Macau, China

Correspondence
Tao Qian, Macau Center for Mathematical Sciences, Macau University of Science and Technology, Macau, China.
Email: tqian@must.edu.mo

Communicated by: W. Sprößig

Funding information
National Natural Science Foundation of China, Grant/Award Number: 11501426; Science and Technology Development Fund, Macao SAR, Grant/Award Number: FDCT0123/2018/A3

Through reducing the problem to rational orthogonal system (Takenaka–Malmquist system), this note gives a proof for existence of n-best rational approximation to functions in the Hardy $\mathbb{H}^2(\mathbb{D})$ space by using pseudohyperbolic distance.

KEYWORDS

n-best rational approximation, Schwarz lemma, T-M system

MSC CLASSIFICATION

42A50; 32A30; 32A35; 46J15

1 | INTRODUCTION

This paper concerns rational approximations in the Hardy space

$$\mathbb{H}^2(\mathbb{D}) = \left\{ f : \mathbb{D} \to \mathbb{C} \ : \ f = \sum_{k=0}^{+\infty} c_k z^k \text{ with } \sum_{k=1}^{\infty} |c_k|^2 < \infty \right\},$$

where $\mathbb{D}$ is the open unit disc. Among other equivalent definitions of the norm of the Hardy space, we adopt the one in terms of the nontangential boundary limits of the functions

$$\langle f, g \rangle = \frac{1}{2\pi} \int_0^{2\pi} f(e^{it}) \overline{g}(e^{it}) dt.$$

For a given positive integer n, an ordered pair of polynomials (p, q) is called an *n-admissible pair* if p and q are co-prime, $q \neq 0$, the zeros of q are in the exterior of $\mathbb{D}$, and the degrees of p and q are both less than $n+1$. The n-best rational approximation problem in the Hardy $\mathbb{H}^2(\mathbb{D})$ space (abbreviated as $\mathbb{H}^2$) is stated as follows: For $f \in \mathbb{H}^2(\mathbb{D})$, find an admissible pair $(\tilde{p}, \tilde{q})$ such that

$$\left\| f - \frac{\tilde{p}}{\tilde{q}} \right\|_{\mathbb{H}^2} = \min \left\{ \| f - \frac{p}{q} \|_{\mathbb{H}^2} \ : \ (p, q) \text{ is an } n\text{-admissible pair} \right\}. \tag{1}$$

For a Hardy space function f, denoting its nontangential boundary limit also by f if without confusion, there holds $\|f\|_{\mathbb{H}^2} = \|f\|_{L^2(\partial\mathbb{D})}$. In below, we will abbreviate both $\|\cdot\|_{\mathbb{H}^2}$ and $\|\cdot\|_{L^2(\partial\mathbb{D})}$ as $\|\cdot\|$. This problem and closely related ones have been long studied. The existence of a solution to (1) under Chebyshev norm with prescribed poles was first proved in Walsh,[1] and existence, uniqueness of best rational approximation under $\|.\|_p$ norms were given in Walsh.[2] Due to the great interest, researchers have given alternative proofs based on different methods, of which some were related to algorithms to find a solution.[3,4] The basic methodology before our approach introduced in Qian and Wegert[5] and Mi and Qian[6] was to parameterize the problem by the coefficients of the denominator q, and the related optimal numerator is obtained through orthogonalization.[2,7] Our approach to the problem is via Szegö kernel approximation to functions in the space. By this approach, we directly find the best suitable poles and so to find the best denominator. The Szegö kernel approach is as described now. Denote by k_a the Szegö kernel of the Hardy $\mathbb{H}^2(\mathbb{D})$ space, where

$$k_a(z) = \frac{1}{1-\bar{a}z},$$

and by

$$\mathfrak{D} := \left\{ e_a(z) = \frac{\sqrt{1-|a|^2}}{1-\bar{a}z}, a \in \mathbb{D} \right\}, \tag{2}$$

the collection of normalized Szegö kernels. The function k_a is the reproducing kernel in $\mathbb{H}^2(\mathbb{D})$. If a sequence $\{a_n\}_{n=1}^{+\infty} \subset \mathbb{D}$ is given, where multiplicity is allowed, through the Gram–Schmidt (G-S) orthogonalization process on $\{e_{a_n}\}_{n=1}^{+\infty}$, we can get an orthonormal system, called the rational orthogonal system or Takenaka–Malmquist system (T-M system),

$$\mathfrak{E} := \left\{ E_{a_1,a_2,\ldots,a_n}(z) = \frac{\sqrt{1-|a_n|^2}}{1-\bar{a}_n z} \prod_{k=1}^{n-1} \frac{z-a_k}{1-\bar{a}_k z}, n = 1,2\ \ldots \right\}, \tag{3}$$

where $E_n = E_{a_1,a_2,\ldots,a_n} = e_{a_n} B_{a_1,a_2,\ldots,a_{n-1}}$, being the product of a normalized Szegö kernel and an $(n-1)$-order Blaschke product. For any n complex numbers $c_1, \ldots, c_n$, the form

$$\sum_{k=1}^{n} c_k E_k$$

is called an n-Blaschke form. If c_n is nonzero, then it is called an n-nondegenerated Blaschke form. For $f \in \mathbb{H}^2$, there is an associated n-Blaschke form

$$\sum_{k=1}^{n} \langle f, E_k \rangle E_k.$$

From the Hilbert space theory,

$$\left\| f - \sum_{k=1}^{n} \langle f, E_k \rangle E_k \right\| = \min_{(c_1,\ldots,c_n)} \left\{ \| f - \sum_{k=1}^{n} c_k E_k \| \right\},$$

and

$$\left\| f - \sum_{k=1}^{n} \langle f, E_k \rangle E_k \right\| = 0,$$

if and only if $f \in \mathrm{Span}\{E_k\}_{k=1}^{n}$.

There is correspondingly an n-best Blaschke form approximation problem: find a set of n parameters $a_1, \ldots, a_n$, all in $\mathbb{D}$, such that

$$\begin{aligned} &\left\| f - \sum_{k=1}^{n} \langle f, E_{a_1,\ldots,a_k} \rangle E_{a_1,\ldots,a_k} \right\| \\ &= \inf \left\{ \left\| f - \sum_{k=1}^{n} \langle f, E_{b_1,\ldots,b_k} \rangle E_{b_1,\ldots,b_k} \right\| : b_1, \ldots, b_n \text{ are all in } \mathbb{D} \right\}. \end{aligned} \tag{4}$$

Below, we will refer (4) as "n-best Blaschke form approximation." For the connection between the n-best rational and the n-best Blaschke form functions, there is the following observation: if there exists $a_k = 0$ among an $n+1$ sequence

$a_1, \dots, a_{n+1}$, where the multiplicity is allowed, then nondegenerate $(n+1)$-Blaschke form made from $a_1, \dots, a_{n+1}$ are of the form p/q, where (p, q) is an n-admissible pair.[8]

Based on the above observation, the n-best rational approximation problem and the n-best Blaschke approximation problem are essentially the same. In fact, to get an n-best rational approximation to f, one can, instead, solve the n-Blaschke form problem for $f - c_0$, where c_0 is the zeroth Fourier coefficient of f. In the wide notion of sparse representation by linear combinations of the dictionary words in a Hilbert space with a dictionary, the problem in terms of n-Blaschke form seems to be more essential and natural, as well as more general. Our main result is as follows.

Theorem 1. *For any $f \in \mathbb{H}^2$ and any positive integer n, if f is not identical with an m-Blaschke form, $m < n$, then there exists a nondegenerate n-best approximation to f.*

A mathematical algorithm for finding a solution to the n-best Blaschke approximation has yet been an open problem. There, however, exist several proofs for existence of a solution that were mostly associated with the goal of finding an algorithm. In this note, we provide a new proof for the existence by surprisingly using pseudohyperbolic distance. This methodology may lead a new way to treat similar problems in general reproducing kernel Hilbert functional spaces in which the Hardy space methods are unadaptable.

The pseudohyperbolic distance on $\mathbb{D}$ is defined by

$$\rho(z_0, z) = \left| \frac{z - z_0}{1 - \bar{z}_0 z} \right|. \tag{5}$$

If analytic function $g(z)$ is defined from $\mathbb{D}$ to $\overline{\mathbb{D}}$, Schwarz lemma shows that[9]

$$\rho(g(z_0), g(z)) \le \rho(z_0, z), \;\; z_0 \neq z, \tag{6}$$

and

$$|g'(z)|(1 - |z|^2) \le 1 - |g(z)|^2. \tag{7}$$

In the sequel, we denote $f_{b_1,b_2,\dots,b_n}(z) = f(z) - \sum_{k=1}^{n} \langle f, E_{b_1,b_2,\dots,b_k} \rangle E_{b_1,b_2,\dots,b_k}(z)$. We note that

$$\left\| f - \sum_{k=1}^{n} \langle f, E_{b_1,b_2,\dots,b_k} \rangle E_{b_1,b_2,\dots,b_k} \right\|^2 = \|f\|^2 - \sum_{k=1}^{n} |\langle f, E_{b_1,b_2,\dots,b_k} \rangle|^2.$$

Hence, to attain

$$\inf \left\| f - \sum_{k=1}^{n} \langle f, E_{b_1,b_2,\dots,b_k} \rangle E_{b_1,b_2,\dots,b_k} \right\|^2$$

is equivalent with to attain

$$\sup \sum_{k=1}^{n} |\langle f, E_{b_1,b_2,\dots,b_k} \rangle|^2. \tag{8}$$

We will denote the orthogonal projection of f into the linear subspace X by $P_X f$. The projection into the subspace as orthogonal complement of X is denoted $Q_X f = (I - P_X) f$. In the case of $X = \text{Span}\{e_{a_1}, \dots, e_{a_n}\}$, they will be simply denoted as $P_{a_1,\dots,a_n}$ and $Q_{a_1,\dots,a_n}$. It is recognized that $Q_{a_1,\dots,a_k}$ is the G-S process operator, and

$$E_{a_1,a_2,\dots,a_k} = \frac{Q_{a_1,a_2,\dots,a_{k-1}}(e_{a_k})}{\|Q_{a_1,a_2,\dots,a_{k-1}}(e_{a_k})\|}.$$

We note in this notation $f_{a_1,a_2,\dots,a_k} = Q_{a_1,\dots,a_k} f$, and, owing to the self-adjoint property of projection operators and the orthogonality gained from the G-S process, for a_l among $a_1, \dots, a_k$,

$$\begin{aligned} f_{a_1,a_2,\dots,a_k}(a_l) &= \langle Q_{a_1,\dots,a_k} f, k_{a_l} \rangle \\ &= \langle f, Q_{a_1,\dots,a_k} k_{a_l} \rangle \\ &= \langle f, (I - P_{a_1,a_2,\dots,a_k}) k_{a_l} \rangle \\ &= 0. \end{aligned}$$

2 | PROOF OF THEOREM

We will first prove the following:

Lemma 1. *Let f be a Hardy space function with analytic continuation to $\overline{\mathbb{D}}$ with $\|f\|_{\mathbb{H}^\infty} \le M$; then, for any k-tuple $a_1, \dots, a_k \in \mathbb{D}$, there holds $\|f_{a_1,\dots,a_k}\|_{\mathbb{H}^\infty} \le 3^k M$.*

Proof. When $k = 1$, by using the reproducing kernel property of $k_{a_1}(z) = \frac{1}{(1-\overline{a}_1 z)}$, we have

$$\begin{aligned} |f_{a_1}(z)| &= |f(z) - \langle f, E_{a_1} \rangle E_{a_1}(z)| \\ &\le |f(z)| + |f(a_1)|(1 - |a_1|^2) \frac{1}{1 - |a_1|} \\ &\le 3M. \end{aligned}$$

Now, treat the general $k > 1$ case. For arbitrary $z \in \mathbb{D}$, due to the orthogonality and the properties of the projection operator $Q_{a_1,\dots,a_{k-1}}$,

$$\begin{aligned} f_{a_1,\dots,a_k}(z) &= f_{a_1,\dots,a_{k-1}}(z) - \langle f_{a_1,\dots,a_{k-1}}, E_{a_1,\dots,a_k} \rangle E_{a_1,\dots,a_k}(z) \\ &= f_{a_1,\dots,a_{k-1}}(z) - \left\langle f_{a_1,\dots,a_{k-1}}, B_{a_1,\dots,a_{k-1}} e_{a_k} \right\rangle B_{a_1,\dots,a_{k-1}}(z) e_{a_k}(z) \\ &= f_{a_1,\dots,a_{k-1}}(z) - \left\langle \frac{f_{a_1,\dots,a_{k-1}}}{B_{a_1,\dots,a_{k-1}}}, e_{a_k} \right\rangle B_{a_1,\dots,a_{k-1}}(z) e_{a_k}(z). \end{aligned} \tag{9}$$

The modulus of (9) is dominated by

$$|f_{a_1,\dots,a_{k-1}}(z)| + \left| \frac{f_{a_1,\dots,a_{k-1}}(a_k)}{B_{a_1,\dots,a_{k-1}}(a_k)} \right| \frac{1 - |a_k|^2}{|1 - \overline{a}_k z|}, \tag{10}$$

where the function

$$\frac{f_{a_1,\dots,a_{k-1}}(z)}{B_{a_1,\dots,a_{k-1}}(z)}$$

is analytic in $\overline{\mathbb{D}}$. By invoking the maximum modulus principle of analytic functions, it takes the maximal modulus on $\partial\mathbb{D}$ dominated by $3^{k-1}M$, according to the inductive hypothesis. On the other hand, for $|z| \le 1$,

$$\left| \frac{1 - |a_k|^2}{1 - \overline{a}_k z} \right| \le 2.$$

Altogether, the quantity in (10) is dominated by $3^k M$, as desired. □

2.1 | 1-best approximation

For $n = 1, f \in \mathbb{H}^2$, one can find $a_1 \in \mathbb{D}$ such that $|\langle f, E_{a_1}\rangle| = |f(a_1)|\sqrt{1-|a_1|^2}$ attains the maximal possible value of all the same kind. This is the so-called maximal selection principle proved through the boundary vanishing condition

$$\lim_{|a|\to 1} |\langle f, E_a\rangle| = 0 \tag{11}$$

via a Bolzano–Weierstrass compact argument.[10] For the self-containing purpose, we cite the simple proof of (11) at the point.

For $\forall \epsilon > 0$, we can find a polynomial function g such that

$$\|f - g\| < \frac{\epsilon}{2}.$$

Since g is bounded in $\overline{\mathbb{D}}$, we have

$$\begin{aligned} |\langle f, E_a\rangle| &\le |\langle g, E_a\rangle| + \epsilon/2 \\ &= \sqrt{1-|a|^2}|g(a)| + \epsilon/2 \\ &\le \epsilon, \end{aligned}$$

if $|a|$ is sufficiently close to 1.

2.2 | 2-best approximation

By using the same density argument as for the $n = 1$ case, we may assume that f is a complex polynomial which is bounded, say by M, in a neighborhood of $\overline{\mathbb{D}}$. Based on the definition of supreme, one can find a sequence of 2 tuples, $(e_{a_1^{(l)}}, e_{a_2^{(l)}}), l = 1, 2, \dots$, such that the norms of the projections $P_{a_1^{(l)},a_2^{(l)}} f$ tends to the supreme (8). Owing to continuity of inner product, we may assume $a_1^{(l)} \neq a_2^{(l)}$ for every $l = 1, 2, \dots$. Since $(a_1^{(l)}, a_2^{(l)}) \in \mathbb{D}\times\mathbb{D}$, we may assume, through a Bolzano–Weierstrass compact argument, that the 2-tuple $(a_1^{(l)}, a_2^{(l)})$ converge to $(a_1, a_2) \in \overline{\mathbb{D}}\times\overline{\mathbb{D}}$. If we can show $(a_1, a_2) \in \mathbb{D}\times\mathbb{D}$, then we are done. We show this by contradiction. Assume the opposite, that is, at least one of a_1 and a_2 is on the boundary $\partial\mathbb{D}$. Since the projections $P_{a_1^{(l)},a_2^{(l)}} f$ are irrelevant with the order of $a_1^{(l)}, a_2^{(l)}$, we may assume that $a_2 \in \partial\mathbb{D}$ and will then derive a contradiction.

As a consequence of Lemma 1, for any a_1 in $\mathbb{D}$, there holds $|f_{a_1}(z)| \le 3M, z \in \mathbb{D}$. By setting $f_1^{(l)} = f_{a_1^{(l)}}/(3M)$, we have $f_1^{(l)}(\mathbb{D}) \subset \overline{\mathbb{D}}$. Since $a_1^{(l)} \neq a_2^{(l)}$ and $f_{a_1^{(l)}}(a_1^{(l)}) = 0$, a similar reasoning as in Lemma 1, we have

$$\begin{aligned} |\langle f, E_{a_1^{(l)},a_2^{(l)}}\rangle| &= |\langle f_{a_1^{(l)}}, E_{a_1^{(l)},a_2^{(l)}}\rangle| \\ &= \left|\left\langle \frac{f_{a_1^{(l)}}}{B_{a_1^{(l)}}}, e_{a_2^{(l)}} \right\rangle\right| \\ &= \left|\frac{f_{a_1^{(l)}}(a_2^{(l)})}{B_{a_1^{(l)}}(a_2^{(l)})}\right| \sqrt{1-|a_2^{(l)}|^2} \\ &= 3M \left|\frac{f_1^{(l)}(a_2^{(l)})}{B_{a_1^{(l)}}(a_2^{(l)})}\right| \sqrt{1-|a_2^{(l)}|^2} \\ &= 3M \left|\frac{\frac{f_1^{(l)}(a_2^{(l)})-f_1^{(l)}(a_1^{(l)})}{1-\overline{f_1^{(l)}(a_1^{(l)})}f_1^{(l)}(a_2^{(l)})}}{\frac{a_2^{(l)}-a_1^{(l)}}{1-\overline{a_1^{(l)}}a_2^{(l)}}}\right| \sqrt{1-|a_2^{(l)}|^2} \\ &= 3M \frac{\rho(f_1^{(l)}(a_1^{(l)}), f_1^{(l)}(a_2^{(l)}))}{\rho(a_1^{(l)}, a_2^{(l)})} \sqrt{1-|a_2^{(l)}|^2}. \end{aligned}$$

Hence, the Schwarz lemma may be used to assert the boundedness of the first factor of the last term of the above chain of inequalities. When $|a_2^{(l)}| \to 1$, for $a_1^{(l)}$ uniformly,

$$|\langle f, E_{a_1^{(l)},a_2^{(l)}}\rangle| = 3M\frac{\rho(f_1^{(l)}(a_1^{(l)}), f_1^{(l)}(a_2^{(l)}))}{\rho(a_1^{(l)}, a_2^{(l)})}\sqrt{1-|a_2^{(l)}|^2}$$
$$\leq 3M\sqrt{1-|a_2^{(l)}|^2} \to 0.$$

Referring to (8), the above argument shows that $a_2^{(l)}$ does not help to get any larger $P_{a_1^{(l)},a_2^{(l)}}f$ than $P_{a_1^{(l)}}f$, as $a_2^{(l)}$ tends to the boundary $\partial\mathbb{D}$. This happens only in the case when f is an m-Blaschke form with $m < 2$. In our case, $m = 1$ if f is nontrivial, contradictory with the assumption of the theorem.

2.3 | General *n*-best approximation

For a general n, as for the $n = 2$ case, we are assuming that f is a polynomial bounded by M in a neighborhood of $\overline{\mathbb{D}}$. An analogous argument leads to a sequence of n-tuples $(a_1^{(l)}, \dots, a_n^{(l)})$ with mutually different terms that leads to

$$\lim_{l\to\infty}\|P_{a_1^{(l)},\dots,a_n^{(l)}}f\|^2 = \sup\left\{\sum_{k=1}^{n}|\langle f, E_{b_1,b_2,\dots,b_k}\rangle|^2 \ : \ b_1, \dots, b_n \in \mathbb{D}\right\}. \tag{12}$$

Through a compact argument we may assume that the n-tuple $(a_1^{(l)}, \dots, a_n^{(l)})$ itself has a limit as an n-tuple $(a_1, \dots, a_n)$, where the a_ks are not necessarily mutually different and can be inside $\mathbb{D}$ or on the boundary of $\mathbb{D}$. If all the a_ks are inside of $\mathbb{D}$, then (12) becomes

$$\|P_{a_1,\dots,a_n}f\|^2 = \sup\left\{\sum_{k=1}^{n}|\langle f, E_{b_1,b_2,\dots,b_k}\rangle|^2 \ : \ b_1, \dots, b_n \in \mathbb{D}\right\}, \tag{13}$$

and we thus have the existence. We now show that this is indeed the case. We prove it by assuming the opposite and then derive a contradiction. Assume that at least one of the components sequences, say, $a_{k'}^{(l)}, 1 \leq k' \leq n, l = 1, 2, \dots$, tends to the boundary $\partial\mathbb{D}$. Since for each fixed l, the projections $P_{a_1^{(l)},\dots,a_n^{(l)}}f$ are irrelevant with the order of $a_1^{(l)}, \dots, a_n^{(l)}$, we may assume that $^{k'=n}$. Likewise to the $n = 2$ case, set

$$f_{n-1}^{(l)}(z) = \frac{f_{a_1^{(l)},\dots,a_{n-1}^{(l)}}(z)}{3^{n-1}MB_{a_1^{(l)},a_2^{(l)},\dots,a_{n-2}^{(l)}}(z)}.$$

Through analysis on the zeros of the denominator and the numerator functions, and invoking the maximum modulus principle, this function is analytic for $\mathbb{D} \to \overline{\mathbb{D}}$. We have, when $|a_n^{(l)}| \to 1$, for $a_1^{(l)}, a_2^{(l)}, \ldots, a_{n-1}^{(l)}$ uniformly,

$$
\begin{aligned}
|\langle f, E_{a_1^{(l)},a_2^{(l)},\ldots,a_n^{(l)}}\rangle| &= \left| \left\langle \frac{f_{a_1^{(l)},a_2^{(l)},\ldots,a_{n-1}^{(l)}}}{B_{a_1^{(l)},a_2^{(l)},\ldots,a_{n-1}^{(l)}}}, e_{a_n^{(l)}} \right\rangle \right| \\
&= \left| \frac{f_{a_1^{(l)},a_2^{(l)},\ldots,a_{n-1}^{(l)}}(a_n^{(l)})}{B_{a_1^{(l)},a_2^{(l)},\ldots,a_{n-1}^{(l)}}(a_n^{(l)})} \right| \sqrt{1-|a_n^{(l)}|^2} \\
&= 3^{n-1} M \left| \frac{\frac{f_{n-1}^{(l)}(a_n^{(l)}) - f_{n-1}^{(l)}(a_{n-1}^{(l)})}{1-\overline{f_{n-1}^{(l)}(a_n^{(l)})} f_{n-1}^{(l)}(a_{n-1}^{(l)})}}{\frac{a_n^{(l)} - a_{n-1}^{(l)}}{1-\overline{a_{n-1}^{(l)}} a_n^{(l)}}} \right| \sqrt{1-|a_n^{(l)}|^2} \\
&= 3^{n-1} M \frac{\rho(f_{n-1}^{(l)}(a_{n-1}^{(l)}), f_{n-1}^{(l)}(a_n^{(l)}))}{\rho(a_{n-1}^{(l)}, a_n^{(l)})} \sqrt{1-|a_n^{(l)}|^2} \\
&\to 0.
\end{aligned}
$$

Remark 1. As mentioned in the introduction, with the general reproducing kernel Hilbert space, setting the n-best approximation question can be asked. However, this question has been fully addressed only in the classical Hardy space cases (the unit disc and the upper-half complex plane). The Hardy space methods are not easily adaptable to other types of reproducing kernel Hilbert spaces. One of the main reasons for it is unavailability of explicit orthogonal system made from the reproducing kernels of the context. It is foreseen that the geometric analysis method given above may inspire establishment of the same existence result in certain complex holomorphic Hilbert functional spaces.

ACKNOWLEDGEMENTS

This work is supported by National Natural Science Foundation of China (11501426) and Science and Technology Development Fund, Macao SAR (FDCT0123/2018/A3).

CONFLICT OF INTEREST

This work does not have any conflicts of interest.

ORCID

Yan-Bo Wang https://orcid.org/0000-0003-4754-207X
Tao Qian https://orcid.org/0000-0002-8780-9958

REFERENCES

1. Walsh JL. The existence of rational functions of best approximation. *Ibid*. 1931;33:668-689.
2. Walsh JL. Interpolation and Approximation by Rational Functions in the Complex Plane. AMS; 1965.
3. Baratchart L, Saff EB, Wielonsky F. A criterion for uniqueness of a critical point in H^2 rational approximation. *J d'Analyse Math*. 1996;70:225-266.
4. Baratchart L, Cardelli M, Olivi M. Identification and rational L2 approximation—a gradient algorithm. *Automatica*. 1991;27:413-418.
5. Qian T, Wegert E. Optimal approximation by Blaschke forms. *Compl Variables Elliptic Equat*. 2013;58:123-133.
6. Mi W, Qian T. Frequency domain identification: an algorithm based on adaptive rational orthogonal system. *Automatica*. 2012;48:1154-1162.
7. Niness B, Gustafsson F. A unified construction of orthogonal basis for system identification. *IEEE Trans Autom Control*. 1997;42:515-522.
8. Qian T. Cyclic AFD algorithm for best rational. *Math Methods Appl Sci*. 2014;37:846-859.
9. Garnett JB. *Bounded Analytic Functions*: Academic Press; 1987.
10. Qian T, Wang YB. Adaptive Fourier series—a variation of greedy algorithm. *Adv Comput Math*. 2011;34:279-293.

How to cite this article: Wang Y-B, Qian T. Pseudohyperbolic distance and n-best rational approximation in $\mathbb{H}^2$ space. *Math Meth Appl Sci*. 2021;44:8497–8504. https://doi.org/10.1002/mma.7267