SCIENCE CHINA
Mathematics

• ARTICLES • doi: 10.1007/s11425-013-4710-1

Approximation of monogenic functions by higher order Szegö kernels on the unit ball and half space

WANG JinXun & QIAN Tao*

Department of Mathematics, University of Macau, Macao, China
Email: wjxpyh@gmail.com, fsttq@umac.mo

Received March 12, 2013; accepted July 13, 2013

Abstract We study the adaptive decomposition of functions in the monogenic Hardy spaces $\mathcal{H}^2$ by higher order Szegö kernels under the framework of Clifford algebra and Clifford analysis, in the context of unit ball and half space. This is a sequel and a higher-dimensional generalization of our recent study on the complex Hardy spaces.

Keywords monogenic Hardy space, Szegö kernel, adaptive decomposition, dictionary, matching pursuit

MSC(2010) 30C40, 30G35, 30H10, 41A20, 41A50, 42B30

Citation: Wang J X, Qian T. Approximation of monogenic functions by higher order Szegö kernels on the unit ball and half space. Sci China Math, 2013, 56, doi: 10.1007/s11425-013-4710-1

1 Introduction

Rational approximation has a long history. In the one-dimensional case it is naturally related to complex approximation (see [25]). Given a real-valued function $f \in L^p(\mathbb{R})$ $(1 < p < \infty)$, it has the Hardy space decomposition $f = f^+ + f^-$, where f^+ and f^- are non-tangential boundary limits of analytic functions belonging to, respectively, the Hardy space in the upper half plane and that in the lower half plane. For f^+, we use holomorphic rational functions with poles in the lower half plane to approximate. The real part of f^+, being $\frac{1}{2}f$, then is approximated by the real part of the complex approximating rational functions.

This paper is restricted to the case $p = 2$. In one dimension the criterion "best n-rational approximation" is formulated as follows. Let f be in the Hardy space $\mathcal{H}^2(\mathbb{R}^2_+)$ on the upper half plane $\mathbb{R}^2_+$, find co-prime polynomials p and q, both having degrees less or equal to n, and q does not have zero in $\mathbb{R}^2_+$, such that

$$\left\| f - \frac{p}{q} \right\|_{\mathcal{H}^2(\mathbb{R}^2_+)}$$

attains the minimum value of all the possible ones under the conditions of p and q. The existence of the solution was proved many decades ago (see [25]). A practical algorithm of it, however, has not been found. There have been several partial solutions (e.g. [1–3, 15, 20]). Solutions in the series approximation form may be found, for instance, in the literature of orthogonal rational systems. Recently, the so-called adaptive Fourier decomposition (AFD) based on Takenaka-Malmquist (TM) systems incorporating a generalized backward shift process was proposed by Qian et al. in [15] and [20]. The decomposition

*Corresponding author

© Science China Press and Springer-Verlag Berlin Heidelberg 2013

obtained through this method offers a fast rational approximation of the given signal, besides, each component has a well-defined instantaneous frequency function. The Newton gradient method based on AFD gives a solution of the best n-rational approximation under the assumption that there is only one local minimum. We note that AFD is more delicate and effective than matching pursuit (see below).

For higher-dimensional questions may be similarly formulated. Solutions, however, are not obvious. The present paper will discuss some of these problems and give solutions.

Generalization of part of the works on AFD into higher dimensional spaces with dimensions not larger than four was treated in [17], in the context of quaternionic analysis.

In general n-dimensional spaces with $n > 4$, the counterpart of the TM system has not been studied yet. A natural way to do this is to use the Clifford algebra, but usually a Clifford number is not divisible, which makes the problem still open. In spite of this, we note that the TM system can be generated from the set of shifted Cauchy kernel functions (Szegö kernel functions). From this point of view and the matching pursuit algorithm, in the Clifford algebra setting we obtained the adaptive decomposition of functions in the monogenic Hardy spaces using dictionaries consisting of Szegö kernels (see [19]). The matching pursuit algorithm was introduced in [13]. It is a powerful tool in signal analysis and decomposition of functions. In that algorithm, a dictionary $\mathcal{D}$ is given, which is a subset of a prescribed Hilbert space $\mathcal{H}$ and consists of a set of unit vectors, $\mathcal{D} = \{g_\lambda : \lambda \in \Lambda,\ \|g_\lambda\| = 1\}$ with $\overline{\mathrm{span}\mathcal{D}} = \mathcal{H}$. Let $f \in \mathcal{H}$, then in the spirit of matching pursuit,

$$f = \sum_{l=0}^{n} g_{\lambda_l} \langle R^l f, g_{\lambda_l} \rangle + R^{n+1} f, \tag{1.1}$$

where g_{λ_l} is chosen to fulfill the optimal projection principle

$$|\langle R^l f, g_{\lambda_l} \rangle| \geqslant \alpha \sup_{g \in \mathcal{D}} |\langle R^l f, g \rangle|, \tag{1.2}$$

with the optimality factor α lying between 0 and 1, and the residues are inductively defined by

$$R^0 f := f, \quad R^{n+1} f := R^n f - g_{\lambda_n} \langle R^n f, g_{\lambda_n} \rangle. \tag{1.3}$$

That is to say, in each iteration the matching pursuit algorithm aims at minimizing as much as possible the energy of the residue, so that $R^n f$ decays rapidly to zero as n tends to infinity. The algorithm thus provides a flexible and efficient way for decomposition of functions. Although (1.1) is not an orthogonal expansion, there still holds the energy conservation law

$$\|f\|^2 = \sum_{l=0}^{n} |\langle R^l f, g_{\lambda_l} \rangle|^2 + \|R^{n+1} f\|^2.$$

As for the convergence under the matching pursuit, the convergence rate, and some of its variations, we refer the reader to [5, 7, 8, 22–24].

In [19], we proved that in each iteration of decomposition the optimality factor α can attain 1. Thus our scheme results in a fast decomposition and we get a kind of nonlinear rational approximation in even dimensions.

Meanwhile, we also note that AFD may not perform well for the highly oscillatory signals or those signals of high frequencies, since each basic function is a weighted Blaschke product of an exact order. In order to fill this gap, Qian et al. [16] suggested some variations of AFD by unwinding process through factorizing out the inner function factors. In [26] the authors proposed a simpler method for adaptive decomposition of holomorphic functions by rational mono-components, which differs from those in [15], [16] and [20], is essentially the matching pursuit among higher order Szegö kernels (higher order partial derivatives of Szegö kernels) and therefore it is not necessary to get orthogonal decomposition. The method is valid for a rather large amount of signals (functions), since the Fourier system $\{z^n\}_{n=0}^{\infty}$ and the Szegö kernel functions are incorporated into the dictionary. In some sense it is an extension of AFD and the work in [19] on finding faster and more efficient ways for decomposition of functions.

In this paper, we will investigate the higher-dimensional analogue of the cases studied in [26]. We will use the collection of higher order Szegö kernels which take values in a Clifford algebra and are (left) monogenic (see Section 2 for its definition) as a dictionary for the decomposition, because this dictionary is extremely redundant, and what is more, the inner product of a function and a higher order Szegö kernel can be easily computed out by a unified closed formula (see (3.7) and (4.1)). Unlike the case for complex-valued functions, the higher order Szegö kernel functions for this setting do not have concrete or explicit expressions, since the higher order partial derivatives are involved. Therefore the difficulties rise not only in the computation, but also in proving the corresponding maximum selection criterion (1.2) with $\alpha = 1$ being optimal. By Fourier transform and changing variables via Möbius transformation, we solve these problems and thus obtain a more flexible way for nonlinear approximation of monogenic functions and signals. As a consequence, we get a constructive proof for a kind of approximation theorems (Theorems 3.6 and 4.9) that is similar to the Runge's approximation theorem (cf. [4]). However, the related convergence rate remains open.

We will treat two kinds of monogenic Hardy spaces, namely, the monogenic Hardy spaces for the unit ball and half space. The higher order Szegö kernel function in the dictionary in each context contains two parameters. One is the discrete parameter k (a multi-index) which is exactly the order of the kernel function. The other is the continuous parameter a which lies in the underlying spaces. It is interesting to note that selection criterion (1.2) with $\alpha = 1$ for the cases we study is an outcome of the fact that the bound of the linear functional $\mathcal{T}_{k,a}$ (see Sections 3 and 4) on $\mathcal{H}^2$ increases to infinity as the parameters tend to the boundary. These two parameters in fact correspond to the "frequency" and "amplitude" respectively, hence our method provides an adaptive coherent representation of the given signal, from which the implicit structure information of the signal can be easily detected.

If a scalar (real-valued) signal (or L^2 function) U is defined on the boundary of the unit ball or the half space, then one can construct a non-scalar field V which is conjugate to U (see [19]), such that the pair $U+V$ is the boundary limit of some monogenic $\mathcal{H}^2$ function F, in the sense of L^2 and almost everywhere. This enables us to apply our method to F inside the domain. The recovery of U can be obtained by taking the scalar part of the trace (boundary values) of the adaptive decomposition of F through higher order Szegö kernels. Note that another simple but not adaptive approach for rational approximation of F is to directly use the Taylor series, where the set of basic functions consists of Fueter polynomials (in complex analysis they reduce to complex power functions). But in the case of half space the Taylor series usually just converges in a smaller region. For example, $F(z) = (z+\mathrm{i})^{-1}\,(z \in \mathbb{R}^2_+)$ is an $\mathcal{H}^2$ function in the upper half plane $\mathbb{R}^2_+$, its Talyor series is convergent only in $\mathbb{D} \cap \mathbb{R}^2_+$, where $\mathbb{D}$ stands for the unit disc.

The paper is organized as follows. In Section 2, we give a quick review on Clifford analysis and monogenic Hardy spaces. In Sections 3 and 4 we discuss the problems of adaptive decomposition by higher order Szegö kernels in the monogenic Hardy spaces for the unit ball and half space respectively.

2 Preliminaries

2.1 Clifford algebra and Clifford analysis

We shall use the real Clifford algebra $\mathscr{A}_m$ of dimension 2^m, generated from $\mathbb{R}^m = \{x = x_1e_1 + \cdots + x_me_m : x_i \in \mathbb{R}, 1 \leqslant i \leqslant m\}$, where $\{e_1, \ldots, e_m\}$ forms an orthonormal basis of $\mathbb{R}^m$, with the anti-commutative relationship

$$e_ie_j + e_je_i = -2\delta_{ij}, \quad i, j = 1, \ldots, m,$$

in which δ_{ij} is the Kronecker delta function. $\mathscr{A}_m$ is an associative algebra whose element is of the form $x = \sum_T x_Te_T$, where $T = \{1 \leqslant i_1 < i_2 < \cdots < i_l \leqslant m\}$ runs over all ordered subsets of $\{1, \ldots, m\}$, $x_T \in \mathbb{R}$ with $x_\emptyset = x_0$, and $e_T = e_{i_1}e_{i_2}\cdots e_{i_l}$ with the identity element $e_\emptyset = e_0$. $\operatorname{Sc} x := x_0$ and $\operatorname{NSc} x := x - \operatorname{Sc} x$ are respectively called the scalar part and non-scalar part of x. The conjugate and the norm of x are respectively defined by $\overline{x} := \sum_T x_T\overline{e_T}$ and $|x| := \operatorname{Sc}(x\overline{x}) = (\sum_T x_T^2)^{1/2}$, where $\overline{e_T} = \overline{e_{i_l}}\cdots\overline{e_{i_2}}\,\overline{e_{i_1}}$ with $\overline{e_0} = e_0$ and $\overline{e_i} = -e_i$ for $i \neq 0$. We have for any $x, y, z \in \mathscr{A}_m$, $\overline{xy} = \overline{y}\,\overline{x}$, $(xy)z = x(yz)$ and

$|xy| \leqslant 2^{m/2}|x||y|$. $\mathscr{A}_0 = \mathbb{R}$, $\mathscr{A}_1 = \mathbb{C}$ and $\mathscr{A}_2$ coincides with the quaternion algebra $\mathbb{H}$. $\mathscr{A}_m$ is no longer a division algebra when $m \geqslant 3$, but if $x \in \mathscr{A}_m$ is of the special form like $x = \sum_{i=0}^m x_i e_i \in \mathbb{R}^{m+1}$, then obviously its inverse is given by $x^{-1} = \overline{x}/|x|^2$.

A function $f(x) = \sum_T f_T(x) e_T \in C^1(\Omega, \mathscr{A}_m)$ is said to be left (right) monogenic in the open set $\Omega \subset \mathbb{R}^{m+1}$ if and only if it satisfies the generalized Cauchy-Riemann equation

$$Df = \sum_{i=0}^m e_i \frac{\partial f}{\partial x_i} = 0 \left(fD = \sum_{i=0}^m \frac{\partial f}{\partial x_i} e_i = 0 \right),$$

where the Dirac operator D is defined by $D = \frac{\partial}{\partial x_0} + \nabla = \sum_0^m e_i \frac{\partial}{\partial x_i}$. If f is left (right) monogenic, then from $\overline{D}(Df) = (fD)\overline{D} = \triangle f$ we know that each component of f is a real-valued harmonic function. Obviously, the set of the left (right) monogenic functions in Ω is a right (left) $\mathscr{A}_m$-module.

The theory of Clifford analysis founded and developed by Brackx et al. is about the theory of monogenic functions, which is a generalization of the complex analysis into higher-dimensional spaces. Now it plays an important role in modern analysis and becomes a powerful tool for many applications. A Cauchy-type integral formula for this setting is:

Lemma 2.1 (See [4]). *Let $S \subset \Omega$ be an $(m+1)$-dimensional compact differentiable and oriented manifold with boundary. If f is left monogenic in Ω, then*

$$f(x) = \frac{1}{\omega_m} \int_{y \in \partial S} E(y-x) n(y) f(y) dS, \quad x \in \operatorname{int}(S),$$

where $E(x) = \frac{\overline{x}}{|x|^{m+1}}$ is the Cauchy kernel, $\omega_m = 2\pi^{\frac{m+1}{2}}/\Gamma(\frac{m+1}{2})$ stands for the area of the unit sphere $\mathbb{S}^m = \{x \in \mathbb{R}^{m+1} : |x| = 1\}$, $n(y)$ is the outward-pointing unit normal vector on ∂S, dS is the surface area element on ∂S and $\operatorname{int}(S)$ is the interior of S.

2.2 Monogenic Hardy spaces

Since there is no essential difference between left monogenic functions and right monogenic functions, we will work only on left monogenic functions in this paper.

Denote the unit ball $\{x \in \mathbb{R}^{m+1} : |x| < 1\}$ by $\mathbb{B}^m$, and the half space $\{x \in \mathbb{R}^{m+1} : \operatorname{Sc} x > 0\}$ by $\mathbb{R}^{m+1}_+$. For $0 < p \leqslant \infty$ and a function f, we say $f \in \mathcal{H}^p(\mathbb{B}^m)$ (resp. $f \in \mathcal{H}^p(\mathbb{R}^{m+1}_+)$), if f is left monogenic in $\mathbb{B}^m$ (resp. $\mathbb{R}^{m+1}_+$), and satisfies

$$\|f\|_{\mathcal{H}^p} := \sup_{0 \leqslant r < 1} \left(\frac{1}{\omega_m} \int_{\mathbb{S}^m} |f(r\eta)|^p dS(\eta) \right)^{1/p} < \infty,$$

$$\left(\text{resp. } \|f\|_{\mathcal{H}^p} := \sup_{x_0 > 0} \left(\frac{1}{\omega_m} \int_{\mathbb{R}^m} |f(x_0 + \underline{x})|^p d\underline{x} \right)^{1/p} < \infty \right)$$

with a usual modification at $p = \infty$, where $\underline{x} = x_1 e_1 + \cdots + x_m e_m$, $d\underline{x} = dx_1 \cdots dx_m$ is the volume element of $\mathbb{R}^m$.

If $\frac{m-1}{m} < p \leqslant \infty$, then for any $f \in \mathcal{H}^p(\mathbb{B}^m)$ (resp. $\mathcal{H}^p(\mathbb{R}^{m+1}_+)$), the non-tangential limit of f on $\mathbb{S}^m$ (resp. $\mathbb{R}^m$) exists and belongs to $L^p(\mathbb{S}^m)$ (resp. $L^p(\mathbb{R}^m)$). What we will be of particular interest is the Hilbert space $\mathcal{H}^2(\mathbb{B}^m)$ (resp. $\mathcal{H}^2(\mathbb{R}^{m+1}_+)$), which is associated with the inner product

$$\langle f, g \rangle := \frac{1}{\omega_m} \int_{\eta \in \mathbb{S}^m} \overline{g(\eta)} f(\eta) dS(\eta), \quad f, g \in \mathcal{H}^2(\mathbb{B}^m),$$

$$\left(\text{resp. } \langle f, g \rangle := \frac{1}{\omega_m} \int_{\mathbb{R}^m} \overline{g(\underline{y})} f(\underline{y}) d\underline{y}, \quad f, g \in \mathcal{H}^2(\mathbb{R}^{m+1}_+) \right).$$

We note that

$$\|f\|_{\mathcal{H}^2} = (\operatorname{Sc}\langle f, f \rangle)^{1/2},$$

and there holds the quasi-Cauchy-Schwarz inequality

$$|\langle f, g\rangle| \leqslant 2^{m/2}\|f\|_{\mathcal{H}^2}\|g\|_{\mathcal{H}^2}.$$

Moreover, the Cauchy's integral formula holds for $\mathcal{H}^2(\mathbb{B}^m)$ and $\mathcal{H}^2(\mathbb{R}^{m+1}_+)$.

For more information about the monogenic Hardy spaces, please see [9,14].

3 The case for the unit ball

For a multi-index $k = (k_0, k_1, \ldots, k_m) \in \mathbb{N}^{m+1}$ and a point $a = \sum_{l=0}^m a_l e_l \in \mathbb{B}^m$, the associated higher order Szegö kernel function is

$$\varphi_{k,a}(x) = \partial_a^k\left(\frac{\overline{1-\overline{a}x}}{|1-\overline{a}x|^{m+1}}\right) = \frac{\partial^{|k|}}{\partial a_0^{k_0}\partial a_1^{k_1}\cdots\partial a_m^{k_m}}\left(\frac{\overline{1-\overline{a}x}}{|1-\overline{a}x|^{m+1}}\right),$$

where $|k| = \sum_{i=0}^m k_i$. It is easy to check that $\varphi_{k,a}$ is left monogenic on the closure of $\mathbb{B}^m$, and hence $\varphi_{k,a}$ belongs to the Hardy space $\mathcal{H}^p(\mathbb{B}^m)$ $(0 < p \leqslant \infty)$. Moreover, for any $f \in \mathcal{H}^p(\mathbb{B}^m)$ $(1 \leqslant p \leqslant \infty)$, there holds

$$(\partial_x^k f)(a) = \frac{1}{\omega_m}\int_{\mathbb{S}^m}\overline{\varphi_{k,a}(\eta)}f(\eta)dS(\eta), \quad \forall a \in \mathbb{B}^m.$$

In particular, for $f \in \mathcal{H}^2(\mathbb{B}^m)$ we have

$$\langle f, \varphi_{k,a}\rangle = (\partial_x^k f)(a). \tag{3.1}$$

The key property for $\varphi_{k,a}$ that is crucial to this section is:

Proposition 3.1. *For any $0 < p \leqslant \infty$,*

$$\|\varphi_{k,a}\|_{\mathcal{H}^p} = \frac{\||x+a|^{(\frac{1}{p'}-\frac{1}{p})m}\varphi_{k,0}(x+a)\|_{L^p(\mathbb{S}^m)}}{(1-|a|^2)^{\frac{m}{p'}+|k|}}, \tag{3.2}$$

where $p' = \frac{p}{p-1}$ is the Hölder conjugate of p.

Proof. First, we assume that $p \neq \infty$, then

$$\begin{aligned}\|\varphi_{k,a}\|^p_{\mathcal{H}^p} &= \frac{1}{\omega_m}\int_{\mathbb{S}^m}\left|\partial_a^k\left(\frac{\overline{1-\overline{a}\eta}}{|1-\overline{a}\eta|^{m+1}}\right)\right|^p dS(\eta) = \frac{1}{\omega_m}\int_{\mathbb{S}^m}\left|\partial_a^k\left(\frac{\eta-a}{|\eta-a|^{m+1}}\right)\right|^p dS(\eta)\\ &= \frac{1}{\omega_m}\int_{\mathbb{S}^m}\left|\frac{P_k(\eta-a)}{|\eta-a|^{m+2|k|+1}}\right|^p dS(\eta),\end{aligned}$$

where P_k is a homogeneous polynomial of degree $|k|+1$. Changing variables by setting (cf. [10])

$$\eta = \tau_a(\omega) = (\omega+a)(1+\overline{a}\omega)^{-1} = (1-|a|^2)\frac{\omega+a}{|\omega+a|^2} + a,$$

which is in fact a special kind of Möbius transformation mapping the unit sphere $\mathbb{S}^m$ to itself, with the inverse being given by

$$\omega = \tau_a^{-1}(\eta) = \tau_{-a}(\eta) = (1-|a|^2)\frac{\eta-a}{|\eta-a|^2} - a.$$

The corresponding change of the surface area element reads

$$dS(\eta) = \left(\frac{1-|a|^2}{|\omega+a|^2}\right)^m dS(\omega).$$

So,

$$\|\varphi_{k,a}\|^p_{\mathcal{H}^p} = \frac{\frac{1}{\omega_m}\int_{\mathbb{S}^m}|P_k(\omega+a)|^p|\omega+a|^{pm-p-2m}dS(\omega)}{(1-|a|^2)^{pm-m+p|k|}}$$

$$= \frac{\frac{1}{\omega_m}\int_{\mathbb{S}^m} |\varphi_{k,0}(\omega+a)|^p |\omega+a|^{pm-2m} dS(\omega)}{(1-|a|^2)^{pm-m+p|k|}}$$
$$= \frac{\||x+a|^{(\frac{1}{p'}-\frac{1}{p})m}\varphi_{k,0}(x+a)\|^p_{L^p(\mathbb{S}^m)}}{(1-|a|^2)^{\frac{pm}{p'}+p|k|}}.$$

As to the $\mathcal{H}^\infty$ norm, it follows from

$$\|\varphi_{k,a}\|_{\mathcal{H}^\infty} = \lim_{p\to\infty}\|\varphi_{k,a}\|_{\mathcal{H}^p} = \frac{\||x+a|^m\varphi_{k,0}(x+a)\|_{L^\infty(\mathbb{S}^m)}}{(1-|a|^2)^{m+|k|}}. \qquad \square$$

Corollary 3.2.

$$\|\varphi_{k,a}\|_{\mathcal{H}^2} \geqslant \frac{\|W_k(x)\|_{L^2(\mathbb{S}^m)}}{(1-|a|^2)^{\frac{m}{2}+|k|}}, \tag{3.3}$$

where $W_k(x) = \partial_x^k E(x)$ is an outer spherical monogenic (see [4]). And for $\frac{m}{m+|k|} < p \leqslant \infty$, there exist two positive constants $c_{k,m,p}$ and $C_{k,m,p}$ depending on k, m, p, such that

$$\frac{c_{k,m,p}}{(1-|a|^2)^{\frac{m}{p'}+|k|}} \leqslant \|\varphi_{k,a}\|_{\mathcal{H}^p} \leqslant \frac{C_{k,m,p}}{(1-|a|^2)^{\frac{m}{p'}+|k|}}. \tag{3.4}$$

Proof. Observe that $\varphi_{k,0}(x+a)$ is a polynomial of degree $|k|$, so when $p > \frac{m}{m+|k|}$ we have

$$||x+a|^{(\frac{1}{p'}-\frac{1}{p})m}\varphi_{k,0}(x+a)| \leqslant C_{k,m,p}|x+a|^{(1-\frac{2}{p})m+|k|} \in L^p(\mathbb{S}^m)$$

even for $a \in \mathbb{S}^m$, which implies that $\||x+a|^{(\frac{1}{p'}-\frac{1}{p})m}\varphi_{k,0}(x+a)\|_{L^p(\mathbb{S}^m)}$ is a continuous function with respect to $a \in \overline{\mathbb{B}^m}$. Consequently,

$$0 < c_{k,m,p} = \min_{a\in\overline{\mathbb{B}^m}} \||x+a|^{(\frac{1}{p'}-\frac{1}{p})m}\varphi_{k,0}(x+a)\|_{L^p(\mathbb{S}^m)}$$
$$\leqslant \max_{a\in\overline{\mathbb{B}^m}} \||x+a|^{(\frac{1}{p'}-\frac{1}{p})m}\varphi_{k,0}(x+a)\|_{L^p(\mathbb{S}^m)} = C_{k,m,p} < \infty.$$

(3.4) is now proved. To prove (3.3), in view of (3.2) it suffices to show that

$$\|\varphi_{k,0}(x+a)\|_{L^2(\mathbb{S}^m)} = \|\varphi_{k,0}(x+a)\|_{\mathcal{H}^2} \geqslant \|W_k(x)\|_{L^2(\mathbb{S}^m)}.$$

First, from

$$\varphi_{0,a}(x) = \frac{\overline{1-\overline{a}x}}{|1-\overline{a}x|^{m+1}} = E(x)E(x^{-1}-\overline{a})$$

we know that

$$\varphi_{k,0}(x+a) = (-1)^{k_0}(\mathcal{K}(W_k(y)))|_{y=x+a},$$

where $(\mathcal{K}f)(x) := E(x)f(x^{-1})$ is the Kelvin transform (see [6]). Secondly, the Taylor expansion of $\varphi_{k,0}(x+a)$ reads

$$\varphi_{k,0}(x+a) = \sum_{\underline{\alpha}=(\alpha_1,\dots,\alpha_m)\in\mathbb{N}^m} V_{\underline{\alpha}}(x)((\partial_{\underline{x}}^{\underline{\alpha}}\varphi_{k,0}(x+a))|_{x=0}) = \sum_{l=0}^{|k|}\sum_{|\underline{\alpha}|=l} V_{\underline{\alpha}}(x)((\partial_{\underline{x}}^{\underline{\alpha}}\varphi_{k,0}(x))|_{x=a}),$$

where $V_{\underline{\alpha}}(x)$ is an inner spherical monogenic of order $|\underline{\alpha}|$ (see [4, 6]), $\partial_{\underline{x}}^{\underline{\alpha}} = \frac{\partial^{|\underline{\alpha}|}}{\partial x_1^{\alpha_1}\cdots\partial x_m^{\alpha_m}}$. It follows that (see [4, 21])

$$\|\varphi_{k,0}(x+a)\|^2_{L^2(\mathbb{S}^m)} = \sum_{l=0}^{|k|}\left\|\sum_{|\underline{\alpha}|=l} V_{\underline{\alpha}}(x)((\partial_{\underline{x}}^{\underline{\alpha}}\varphi_{k,0}(x))|_{x=a})\right\|^2_{L^2(\mathbb{S}^m)}.$$

Note that $\partial_{\underline{x}}^{\underline{\alpha}}\varphi_{k,0}(x)$ is a homogeneous polynomial of order $|k|-|\underline{\alpha}|$, so, if $l<|k|$, the norm

$$\Big\| \sum_{|\underline{\alpha}|=l} V_{\underline{\alpha}}(x)((\partial_{\underline{x}}^{\underline{\alpha}}\varphi_{k,0}(x))|_{x=a})\Big\|^2_{L^2(\mathbb{S}^m)}$$

will take the minimum value 0 at the point $a=0$. Therefore, for all $a\in\mathbb{B}^m$ we have

$$\|\varphi_{k,0}(x+a)\|_{L^2(\mathbb{S}^m)} \geqslant \|\varphi_{k,0}(x)\|_{L^2(\mathbb{S}^m)} = \|\mathcal{K}(W_k(x))\|_{L^2(\mathbb{S}^m)} = \|W_k(x)\|_{L^2(\mathbb{S}^m)}.$$

The proof of (3.3) is complete. □

Corollary 3.3. *For each $(k,a)\in\mathbb{N}^{m+1}\times\mathbb{B}^m$, define the linear functional $\mathcal{T}_{k,a}$ on $\mathcal{H}^2(\mathbb{B}^m)$ by*

$$\mathcal{T}_{k,a}f = \langle f,\varphi_{k,a}\rangle = (\partial_x^k f)(a), \quad f\in\mathcal{H}^2(\mathbb{B}^m),$$

then

$$\|\mathcal{T}_{k,a}\| \asymp (1-|a|)^{-\frac{m}{2}-|k|} \quad \text{as } |a|\to 1^-,$$

where by $A\asymp B$ we mean that $A=O(B)$ and $B=O(A)$ simultaneously hold.

Proof. Since

$$\varphi_{k,a}\overline{\varphi_{k,a}} = \overline{\varphi_{k,a}}\varphi_{k,a}\in\mathbb{R}, \tag{3.5}$$

we have the Cauchy-Schwarz inequality

$$|\langle f,\varphi_{k,a}\rangle| \leqslant \|f\|_{\mathcal{H}^2}\|\varphi_{k,a}\|_{\mathcal{H}^2},$$

and

$$\|\varphi_{k,a}\|^2_{\mathcal{H}^2} = \langle\varphi_{k,a},\varphi_{k,a}\rangle = (\partial_x^k\varphi_{k,a}(x))|_{x=a}.$$

Hence, $\|\mathcal{T}_{k,a}\| = \|\varphi_{k,a}\|_{\mathcal{H}^2}$, and (3.4) in Corollary 3.2 tells us that

$$\|\varphi_{k,a}\|_{\mathcal{H}^2} \asymp (1-|a|)^{-\frac{m}{2}-|k|} \quad \text{as } |a|\to 1^-.$$

□

Let

$$P_{k,a}(x) = \partial_a^k\left(\frac{1-|a|^2|x|^2}{|1-\bar{a}x|^{m+1}}\right), \quad (k,a)\in\mathbb{N}^{m+1}\times\mathbb{B}^m$$

be the higher order Poisson kernel function for the harmonic Hardy space $H^2(\mathbb{B}^m)$ (consisting of all functions that are harmonic on $\mathbb{B}^m$ and satisfy the same norm condition for $\mathcal{H}^2(\mathbb{B}^m)$). Define the linear functional $T_{k,a}$ on $H^2(\mathbb{B}^m)$ by

$$T_{k,a}f = \langle f,P_{k,a}\rangle = (\partial_x^k f)(a), \quad f\in H^2(\mathbb{B}^m).$$

Then we have

Corollary 3.4. $\|T_{k,a}\| \asymp (1-|a|)^{-\frac{m}{2}-|k|}$ *as* $|a|\to 1^-$.

Proof. It is clear that

$$\|T_{k,a}\| = \|P_{k,a}\|_{H^2} = \|P_{k,a}\|_{L^2(\mathbb{S}^m)}.$$

From

$$|\langle\varphi_{k,a},P_{k,a}\rangle| \leqslant \|\varphi_{k,a}\|_{H^2}\|P_{k,a}\|_{H^2},$$

we conclude that

$$\|P_{k,a}\|_{H^2} \geqslant \frac{|\langle\varphi_{k,a},P_{k,a}\rangle|}{\|\varphi_{k,a}\|_{H^2}} = \frac{(\partial_x^k\varphi_{k,a}(x))|_{x=a}}{\|\varphi_{k,a}\|_{\mathcal{H}^2}} = \|\varphi_{k,a}\|_{\mathcal{H}^2}.$$

On the other hand, the following estimate (see [18])

$$(1-|a|)^{\frac{m}{2}+|k|}|(\partial_x^k f)(a)| \leqslant C_{k,m}\|f\|_{H^2}, \quad \forall f\in H^2(\mathbb{B}^m)$$

implies that $\|T_{k,a}\| = O((1-|a|)^{-\frac{m}{2}-|k|})$ as $|a|\to 1^-$. □

Next we will study the adaptive decomposition for functions in $\mathcal{H}^2(\mathbb{B}^m)$. We shall use the following dictionary

$$\mathcal{D}=\left\{e_{k,a}=\frac{\varphi_{k,a}}{\|\varphi_{k,a}\|_{\mathcal{H}^2}}:(k,a)\in\mathbb{N}^{m+1}\times\mathbb{B}^m\right\}.$$

Note that if each element $e_{k,a}$ is regarded as a function of a, then it satisfies $e_{k,a}\overline{D}=0$, which means that $e_{k,a}$ is a finite linear combination of the elements in $\{e_{l,a}:l_0=0,|l|=|k|\}$. Hence $\mathcal{D}$ is a redundant dictionary. In addition, $\mathcal{D}$ is complete, because

$$\mathcal{H}^2(\mathbb{B}^m)=\overline{\mathrm{span}\{e_{0,a}:a\in\mathbb{B}^m\}}\subset\overline{\mathrm{span}\mathcal{D}}.$$

For any $f\in\mathcal{H}^2(\mathbb{B}^m)$, according to the matching pursuit algorithm, there is an associated decomposition (like (1.1))

$$f=\sum_{l=0}^{n}e_{k_l,a_l}\langle R^lf,e_{k_l,a_l}\rangle+R^{n+1}f=S_nf+R^{n+1}f, \tag{3.6}$$

with the residues being similarly defined as in (1.3). From (3.1) we get the formula which will be convenient for the computation of each coefficient:

$$\langle R^lf,e_{k_l,a_l}\rangle=\frac{(\partial_x^{k_l}R^lf)(a_l)}{\|\varphi_{k_l,a_l}\|_{\mathcal{H}^2}}. \tag{3.7}$$

Although generally a Clifford number is not invertible, owing to (3.5) the energy conservation law for (3.6) still holds, i.e.,

$$\|f\|_{\mathcal{H}^2}^2=\sum_{l=0}^{n}|\langle R^lf,e_{k_l,a_l}\rangle|^2+\|R^{n+1}f\|_{\mathcal{H}^2}^2. \tag{3.8}$$

So, to make the decomposition converge as fast as possible, at every step we should select e_{k_l,a_l} from $\mathcal{D}$ such that

$$|\langle R^lf,e_{k_l,a_l}\rangle|=\sup_{g\in\mathcal{D}}|\langle R^lf,g\rangle|. \tag{3.9}$$

The reachability of the supremum in (3.9) is a corollary of the following theorem.

Theorem 3.5. *Suppose $f\in\mathcal{H}^2(\mathbb{B}^m)$, then*

$$\lim_{|k|\to\infty}|\langle f,e_{k,a}\rangle|=0 \tag{3.10}$$

holds uniformly with respect to $a\in\mathbb{B}^m$. Write $a=|a|\xi$, then

$$\lim_{|a|\to1^-}|\langle f,e_{k,a}\rangle|=0 \tag{3.11}$$

holds uniformly with respect to $(k,\xi)\in\mathbb{N}^{m+1}\times\mathbb{S}^m$.

Proof. Let

$$T_{f,N}(x)=\sum_{l=0}^{N}\sum_{|\underline{\alpha}|=l}V_{\underline{\alpha}}(x)(\partial_{\underline{x}}^{\underline{\alpha}}f)(0)$$

be the Taylor series of f up to order N. For any $\epsilon>0$, there exists an N such that

$$\|f-T_{f,N}\|_{\mathcal{H}^2}<\epsilon.$$

Therefore, if $|k|>N$, we will get

$$|\langle f,e_{k,a}\rangle|=|\langle f-T_{f,N},e_{k,a}\rangle+\langle T_{f,N},e_{k,a}\rangle|=|\langle f-T_{f,N},e_{k,a}\rangle|\leqslant\|f-T_{f,N}\|_{\mathcal{H}^2}<\epsilon,$$

where we have used the fact that $\langle T_{f,N}, \varphi_{k,a}\rangle = (\partial_x^k T_{f,N})(a) = 0$, since $T_{f,N}$ is a polynomial of order less than $|k|$. (3.10) is now proved. To show (3.11), in view of (3.10), it is enough to show the case for any fixed k. Now, for each $k \in \mathbb{N}^{m+1}$, we have

$$|\langle f, e_{k,a}\rangle| \leqslant |\langle f - T_{f,N}, e_{k,a}\rangle| + |\langle T_{f,N}, e_{k,a}\rangle| \leqslant \|f - T_{f,N}\|_{\mathcal{H}^2} + \frac{|(\partial_x^k T_{f,N})(a)|}{\|\varphi_{k,a}\|_{\mathcal{H}^2}}.$$

Then (3.11) follows by observing that $\partial_x^k T_{f,N}$ is a bounded function, and $\|\varphi_{k,a}\|_{\mathcal{H}^2} \to \infty$ as $|a| \to 1^-$ (Corollary 3.2).

The proof of the theorem is complete. □

For the adaptive decomposition (3.6) subject to the constraint (maximum selection criterion) (3.9), we have the following theorem.

Theorem 3.6.

$$\|S_n f - f\|_{\mathcal{H}^2} = \|R^{n+1} f\|_{\mathcal{H}^2} \to 0 \quad (n \to \infty), \tag{3.12}$$

and for any $k \in \mathbb{N}^{m+1}$ and any compact subset Ω of $\mathbb{B}^m$, the pointwise convergence

$$\big|(\partial_x^k S_n)(a) - (\partial_x^k f)(a)\big| = \big|(\partial_x^k R^{n+1} f)(a)\big| \to 0 \quad (n \to \infty) \tag{3.13}$$

uniformly holds with respect to $a \in \Omega$.

Proof. (3.12) follows from [13, Theorem 1] and the fact that $\mathcal{D}$ is a complete dictionary. (3.13) is a consequence of (3.12) and the estimate

$$\big|(\partial_x^k R^{n+1} f)(a)\big| = |\langle R^{n+1} f, \varphi_{k,a}\rangle| \leqslant \|R^{n+1} f\|_{\mathcal{H}^2} \|\varphi_{k,a}\|_{\mathcal{H}^2} \leqslant \frac{C_{k,m} \|R^{n+1} f\|_{\mathcal{H}^2}}{(1-|a|^2)^{\frac{m}{2}+|k|}}.$$ □

4 The case for the half space

In this section, we will propose two methods for step by step optimal decomposition for functions in $\mathcal{H}^2(\mathbb{R}_+^{m+1})$. One is to use the isomorphism between $\mathcal{H}^2(\mathbb{B}^m)$ and $\mathcal{H}^2(\mathbb{R}_+^{m+1})$. To this end, we let

$$(\mathcal{T} f)(x) = 2^{\frac{m}{2}} \frac{\overline{1+x}}{|1+x|^{m+1}} f((1-x)(1+x)^{-1}).$$

It was showed in [19] that $\mathcal{T}$ is a unitary operator from $\mathcal{H}^2(\mathbb{B}^m)$ to $\mathcal{H}^2(\mathbb{R}_+^{m+1})$, and $\mathcal{T}^{-1} = \mathcal{T}$. So, it is natural to consider the following dictionary

$$\mathscr{D} = \left\{ e_{k,a} = \mathcal{T}\left(\partial_a^k \frac{\overline{1-\overline{a}x}}{|1-\overline{a}x|^{m+1}}\right) : (k,a) \in \mathbb{N}^{m+1} \times \mathbb{B}^m \right\}.$$

One can easily show that for any $f \in \mathcal{H}^2(\mathbb{R}_+^{m+1})$, there exists at least one $g^\star \in \mathscr{D}$ such that

$$|\langle f, g^\star\rangle| = \sup_{g \in \mathscr{D}} |\langle f, g\rangle|.$$

Under this setting, the adaptive decomposition problem for $\mathcal{H}^2(\mathbb{R}_+^{m+1})$ can be in fact converted to the same problem for $\mathcal{H}^2(\mathbb{B}^m)$. However, except for the Szegö kernel itself, $\mathscr{D}$ contains no other higher order Szegö kernel functions for $\mathbb{R}_+^{m+1}$.

The dictionary consisting of the higher order Szegö kernel functions for this case should be

$$\mathcal{D} = \left\{ e_{k,a} = \frac{\varphi_{k,a}}{\|\varphi_{k,a}\|_{\mathcal{H}^2}} : (k,a) \in \mathbb{N}^{m+1} \times \mathbb{R}_+^{m+1} \right\},$$

where

$$\varphi_{k,a}(x) = \partial_a^k \left(\frac{\overline{x+\overline{a}}}{|x+\overline{a}|^{m+1}} \right).$$

$\mathcal{D}$ is a complete and redundant dictionary. Moreover, as a consequence of the Cauchy's integral formula, for any $f \in \mathcal{H}^2(\mathbb{R}_+^{m+1})$ we have

$$\langle f, e_{k,a}\rangle = \frac{(\partial_x^k f)(a)}{\|\varphi_{k,a}\|_{\mathcal{H}^2}}, \tag{4.1}$$

and therefore

$$\|\varphi_{k,a}\|_{\mathcal{H}^2}^2 = \mathrm{Sc}((\partial_x^k \varphi_{k,a}(x))\big|_{x=a}) = \left(\partial_x^k \partial_a^k \frac{x_0 + a_0}{|x+\overline{a}|^{m+1}}\right)\bigg|_{x=a}. \tag{4.2}$$

Proposition 4.1.

$$\|\varphi_{k,a}\|_{\mathcal{H}^2}^2 = \frac{\Gamma(m+2|k|)\prod_{l=1}^m \Gamma(k_l+\frac{1}{2})}{(2\sqrt{\pi})^{m-1}\Gamma(\frac{m+1}{2})\Gamma(\frac{m}{2}+|\underline{k}|)}(2a_0)^{-m-2|k|}, \tag{4.3}$$

where $|\underline{k}| = |k| - k_0$.

Lemma 4.2 (See [11,12]). *Let* $a \in \mathbb{R}_+^{m+1}$, $f(\underline{x}) = \frac{\overline{\underline{x}+\overline{a}}}{|\underline{x}+\overline{a}|^{m+1}}$, $\underline{x} \in \mathbb{R}^m$. *Then*

$$\hat{f}(\underline{x}) = \frac{\pi^{\frac{m+1}{2}}}{\Gamma(\frac{m+1}{2})}\mathrm{e}^{-2\pi \mathrm{i}\langle \underline{x},\underline{a}\rangle}\mathrm{e}^{-2\pi a_0|\underline{x}|}\left(1 + \frac{\mathrm{i}x_1}{|\underline{x}|}e_1 + \cdots + \frac{\mathrm{i}x_m}{|\underline{x}|}e_m\right),$$

where the Fourier transform is defined by

$$\hat{f}(\underline{x}) := \int_{\mathbb{R}^m} f(\underline{\xi})\mathrm{e}^{-2\pi\mathrm{i}\langle \underline{x},\underline{\xi}\rangle}d\underline{\xi}.$$

Proof of Proposition 4.1. By Plancherel theorem and Lemma 4.2,

$$\begin{aligned}
\|\varphi_{k,a}\|_{\mathcal{H}^2}^2 &= \frac{\Gamma(\frac{m+1}{2})}{2\pi^{\frac{m+1}{2}}}\int_{\mathbb{R}^m}\left|\partial_a^k \frac{\overline{\underline{x}+\overline{a}}}{|\underline{x}+\overline{a}|^{m+1}}\right|^2 d\underline{x} \\
&= \frac{\Gamma(\frac{m+1}{2})}{2\pi^{\frac{m+1}{2}}}\int_{\mathbb{R}^m}\left|\left(\partial_a^k \frac{\overline{\cdot+\overline{a}}}{|\cdot+\overline{a}|^{m+1}}\right)^\wedge(\underline{x})\right|^2 d\underline{x} \\
&= \frac{\Gamma(\frac{m+1}{2})}{2\pi^{\frac{m+1}{2}}}\frac{2\pi^{m+1}}{\Gamma^2(\frac{m+1}{2})}\int_{\mathbb{R}^m}\left|\partial_a^k \mathrm{e}^{-2\pi\mathrm{i}\langle \underline{x},\underline{a}\rangle}\mathrm{e}^{-2\pi a_0|\underline{x}|}\right|^2 d\underline{x} \\
&= \frac{\pi^{\frac{m+1}{2}}}{\Gamma(\frac{m+1}{2})}(2\pi)^{2|k|}\int_{\mathbb{R}^m}|\underline{x}|^{2k_0}x_1^{2k_1}\cdots x_m^{2k_m}e^{-4\pi a_0|\underline{x}|}d\underline{x} \\
&= \frac{\Gamma(m+2|k|)\prod_{l=1}^m\Gamma(k_l+\frac{1}{2})}{(2\sqrt{\pi})^{m-1}\Gamma(\frac{m+1}{2})\Gamma(\frac{m}{2}+|\underline{k}|)}(2a_0)^{-m-2|k|},
\end{aligned}$$

where the last step is derived by virtue of the usual spherical coordinate transform. □

Remark 4.3. It is easy to verify from (4.3) that $(2a_0)^{m+2|k|}\|\varphi_{k,a}\|_{\mathcal{H}^2}^2$ is a rational number.

Corollary 4.4. *For each* $(k,a) \in \mathbb{N}^{m+1} \times \mathbb{R}_+^{m+1}$, *define the linear functional* $\mathcal{T}_{k,a}$ *on* $\mathcal{H}^2(\mathbb{R}_+^{m+1})$ *by*

$$\mathcal{T}_{k,a}f = \langle f, \varphi_{k,a}\rangle = (\partial_x^k f)(a), \quad f \in \mathcal{H}^2(\mathbb{R}_+^{m+1}),$$

then

$$\|\mathcal{T}_{k,a}\| = \|\varphi_{k,a}\|_{\mathcal{H}^2} \asymp a_0^{-\frac{m}{2}-|k|} \quad \text{as } a_0 \to 0^+.$$

Let

$$P_{k,a}(x) = 2\partial_a^k\left(\frac{x_0+a_0}{|x+\overline{a}|^{m+1}}\right), \quad (k,a) \in \mathbb{N}^{m+1} \times \mathbb{R}_+^{m+1}$$

be the higher order Poisson kernel function for the harmonic Hardy space $H^2(\mathbb{R}_+^{m+1})$. Define the linear functional $T_{k,a}$ on $H^2(\mathbb{R}_+^{m+1})$ by

$$T_{k,a}f = \langle f, P_{k,a}\rangle = (\partial_x^k f)(a), \quad f \in H^2(\mathbb{R}_+^{m+1}),$$

then we have the following corollary.

Corollary 4.5.

$$\|T_{k,a}\| = \|P_{k,a}\|_{H^2} \asymp a_0^{-\frac{m}{2}-|k|} \quad \text{as } a_0 \to 0^+.$$

Proof. We conclude from (4.2) that

$$\|P_{k,a}\|_{H^2}^2 = \langle P_{k,a}, P_{k,a}\rangle = (\partial_x^k P_{k,a})(a) = 2\|\varphi_{k,a}\|_{\mathcal{H}^2}^2. \qquad \square$$

Theorem 4.6. *Suppose $f \in \mathcal{H}^2(\mathbb{R}_+^{m+1})$, then*

$$\lim_{|k|\to\infty} |\langle f, e_{k,a}\rangle| = 0 \tag{4.4}$$

holds uniformly with respect to $a \in \mathbb{R}_+^{m+1}$;

$$\lim_{a_0\to 0^+} |\langle f, e_{k,a}\rangle| = \lim_{a_0\to\infty} |\langle f, e_{k,a}\rangle| = 0 \tag{4.5}$$

holds uniformly with respect to $(k, \underline{a}) \in \mathbb{N}^{m+1} \times \mathbb{R}^m$, and

$$\lim_{|\underline{a}|\to\infty} |\langle f, e_{k,a}\rangle| = 0 \tag{4.6}$$

holds uniformly with respect to $(k, a_0) \in \mathbb{N}^{m+1} \times (0, \infty)$.

This theorem can be easily proved with the aid of the following two lemmas.

Lemma 4.7 (See [19]). $\overline{\mathrm{span}\{\varphi_{0,b} : b \in \mathbb{R}_+^{m+1}\}} = \mathcal{H}^2(\mathbb{R}_+^{m+1})$.

Lemma 4.8. *For any fixed $b = b_0 + \underline{b} \in \mathbb{R}_+^{m+1}$, we have*

$$\lim_{|k|\to\infty} |\langle \varphi_{0,b}, e_{k,a}\rangle| = 0 \tag{4.7}$$

holds uniformly with respect to $a \in \mathbb{R}_+^{m+1}$;

$$\lim_{a_0\to 0^+} |\langle \varphi_{0,b}, e_{k,a}\rangle| = \lim_{a_0\to\infty} |\langle \varphi_{0,b}, e_{k,a}\rangle| = 0 \tag{4.8}$$

holds uniformly with respect to $(k, \underline{a}) \in \mathbb{N}^{m+1} \times \mathbb{R}^m$, and

$$\lim_{|\underline{a}|\to\infty} |\langle \varphi_{0,b}, e_{k,a}\rangle| = 0 \tag{4.9}$$

holds uniformly with respect to $(k, a_0) \in \mathbb{N}^{m+1} \times (0, \infty)$.

Proof. By Plancherel theorem,

$$\begin{aligned}
|\langle \varphi_{0,b}, e_{k,a}\rangle| &= \frac{C_m \left| \int_{\mathbb{R}^m} (\partial_a^k \frac{\underline{x}+\overline{a}}{|\underline{x}+\overline{a}|^{m+1}})(\overline{\frac{\underline{x}+\overline{\underline{b}}}{|\underline{x}+\overline{b}|^{m+1}}}) d\underline{x} \right|}{\|\varphi_{k,a}\|_{\mathcal{H}^2}} \\
&= \frac{C_m \left| \int_{\mathbb{R}^m} (\partial_a^k \frac{\cdot+\overline{a}}{|\cdot+\overline{a}|^{m+1}})^\wedge(\underline{x}) \mathrm{Conj}((\overline{\frac{\cdot+\overline{\underline{b}}}{|\cdot+\overline{b}|^{m+1}}})^\wedge(\underline{x})) d\underline{x} \right|}{\|\varphi_{k,a}\|_{\mathcal{H}^2}},
\end{aligned}$$

where for a complexified Clifford number $x = \sum_{l=0}^m x_l e_l$ ($x_l \in \mathbb{C}$, $0 \leqslant l \leqslant m$) we define $\mathrm{Conj}(x) := \sum_{l=0}^m \overline{x_l} e_l$.

Appealing to Proposition 4.1 and Lemma 4.2, we obtain

$$\begin{aligned}
|\langle \varphi_{0,b}, e_{k,a}\rangle| &\leqslant \frac{C_m (2\pi)^{|k|} \left(\int_{\mathbb{R}^m} |\underline{x}|^{k_0} |x_1|^{k_1} \cdots |x_m|^{k_m} e^{-2\pi(a_0+b_0)|\underline{x}|} d\underline{x} \right)}{\|\varphi_{k,a}\|_{\mathcal{H}^2}} \\
&= C_m \frac{\sqrt{\Gamma(\frac{m}{2} + |\underline{k}|)} \Gamma(m + |k|) \prod_{l=1}^m \Gamma(\frac{k_l+1}{2})}{\Gamma(\frac{m+|\underline{k}|}{2}) \sqrt{\Gamma(m + 2|k|) \prod_{l=1}^m \Gamma(k_l + \frac{1}{2})}} \frac{(2a_0)^{\frac{m}{2}+|k|}}{(a_0 + b_0)^{m+|k|}}.
\end{aligned}$$

Since $\frac{(2a_0)^{\frac{m}{2}+|k|}}{(a_0+b_0)^{m+|k|}}$ attains its maximum at $a_0=\frac{m+2|k|}{m}b_0$, $(\frac{m+2|k|}{2m+2|k|})^{m+|k|}\leqslant \mathrm{e}^{-\frac{m}{2}}$, and Stirling's formula gives $\Gamma(z)\sim\sqrt{2\pi/z}(z/e)^z$, we have

$$\begin{aligned}
|\langle\varphi_{0,b},e_{k,a}\rangle| &\leqslant C_{b_0,m}\frac{2^{|k|}\sqrt{\Gamma(\frac{m}{2}+|\underline{k}|)}\Gamma(m+|k|)\prod_{l=1}^m\Gamma(\frac{k_l+1}{2})}{\Gamma(\frac{m+|\underline{k}|}{2})\sqrt{\Gamma(m+2|k|)\prod_{l=1}^m\Gamma(k_l+\frac{1}{2})}}\left(\frac{m}{2}+|k|\right)^{-\frac{m}{2}}\\
&\leqslant C_{b_0,m}\frac{|k|^{\frac{m}{2}-\frac{1}{4}}\sqrt{\Gamma(\frac{m}{2}+|\underline{k}|)}\prod_{l=1}^m\Gamma(\frac{k_l+1}{2})}{\Gamma(\frac{m+|\underline{k}|}{2})\sqrt{\prod_{l=1}^m\Gamma(k_l+\frac{1}{2})}}\left(\frac{m}{2}+|k|\right)^{-\frac{m}{2}}\\
&\leqslant C_{b_0,m}|k|^{\frac{m}{2}-\frac{1}{4}}\left(\frac{m}{2}+|k|\right)^{-\frac{m}{2}}2^{-\frac{|\underline{k}|}{2}}\frac{\sqrt{\Gamma(\frac{m}{2}+|\underline{k}|)}}{\Gamma(\frac{m+|\underline{k}|}{2})}\\
&\leqslant C_{b_0,m}|k|^{\frac{m}{2}-\frac{1}{4}}\left(\frac{m}{2}+|k|\right)^{-\frac{m}{2}}|\underline{k}|^{\frac{1}{4}-\frac{m}{4}}\\
&\leqslant C_{b_0,m}|k|^{-\frac{1}{4}}.
\end{aligned}$$

Hence we get (4.7), and we only need to show (4.8) and (4.9) for any fixed $k\in\mathbb{N}^{m+1}$.

By (4.1) and (4.3),

$$\begin{aligned}
|\langle\varphi_{0,b},e_{k,a}\rangle| &= C_{k,m}a_0^{\frac{m}{2}+|k|}\left|\partial_a^k\left(\frac{\overline{a+\bar{b}}}{|a+\bar{b}|^{m+1}}\right)\right| = C_{k,m}a_0^{\frac{m}{2}+|k|}\frac{|P_k(a+\bar{b})|}{|a+\bar{b}|^{m+2|k|+1}}\\
&\leqslant\begin{cases}C_{b,k,m}a_0^{\frac{m}{2}+|k|} & \text{if } a_0 \text{ is small},\\ C_{k,m}a_0^{-\frac{m}{2}} & \text{if } a_0 \text{ is large},\\ C_{b,k,m}a_0^{\frac{m}{2}+|k|}|\underline{a}|^{-m-|k|} & \text{if } |\underline{a}| \text{ is large},\end{cases}
\end{aligned}$$

where P_k is a homogeneous polynomial of degree $|k|+1$. (4.8) and (4.9) now become obvious. □

Theorem 4.6 thus enables us to adaptively decompose a function $f\in\mathcal{H}^2(\mathbb{R}_+^{m+1})$ in the form

$$f=\sum_{l=0}^n e_{k_l,a_l}\langle R^l f,e_{k_l,a_l}\rangle+R^{n+1}f, \tag{4.10}$$

subject to

$$|\langle R^l f,e_{k_l,a_l}\rangle|=\sup_{g\in\mathcal{D}}|\langle R^l f,g\rangle|. \tag{4.11}$$

Similar to Theorem 3.6, we accordingly have

Theorem 4.9.

$$\|R^{n+1}f\|_{\mathcal{H}^2}\to 0\quad(n\to\infty), \tag{4.12}$$

and for any $k\in\mathbb{N}^{m+1}$ *and any* $t>0$,

$$(\partial_x^k R^{n+1}f)(a)\to 0\ (n\to\infty) \tag{4.13}$$

uniformly holds on $\{a\in\mathbb{R}_+^{m+1}:\mathrm{Sc}(a)\geqslant t\}$.

Acknowledgements This work was supported by Macao FDCT 056/2010/A3 and research grant of the University of Macau No. UL017/08-Y4/MAT/QT01/FST. The authors would like to thank the referees for their helpful suggestions and comments.

References

1 Baratchart L, Leblond J. Hardy approximation to L^p functions on subsets of the circle with $1 \leqslant p < \infty$. Constr Approx, 1998, 14: 41–56

2 Baratchart L, Stahl H, Wielonsky F. Asymptotic uniqueness of best rational approximants of given degree to Markov functions in $\mathcal{L}^2$ of the circle. Constr Approx, 2001, 17: 103–138

3 Baratchart L, Wielonsky F. Rational approximation in the real Hardy space H_2 and Stieltjes integrals: A uniqueness theorem. Constr Approx, 1993, 9: 1–21

4 Brackx F, Delanghe R, Sommen F. Clifford Analysis. Boston: Pitman Advanced Publishing Program, 1982

5 Davis G, Mallat S, Avellaneda M. Adaptive greedy approximations. Constr Approx, 1997, 13: 57–98

6 Delanghe R, Sommen F, Souček V. Clifford Algebra and Spinor-valued Functions. A Function Theory for the Dirac Operator. Dordrecht: Kluwer Academic Publishers Group, 1992

7 DeVore R A. Nonlinear approximation. Acta Numerica, 1998, 7: 51–150

8 DeVore R A, Temlyakov V N. Some remarks on greedy algorithms. Adv Comput Math, 1996, 5: 173–187

9 Gilbert J E, Margaret A M. Clifford Algebras and Dirac Operators in Harmonic Analysis. Cambridge: Cambridge University Press, 1991

10 Hua L K. A Talk Starting with Unit Circle (in Chinese). Beijing: Science Press, 1977

11 Kou K, Qian T. The Paley-Wiener theorem in $\mathbb{R}^n$ with the Clifford analysis setting. J Funct Anal, 2002, 189: 227–241

12 Li C, McIntosh A, Qian T. Clifford algebras, Fourier transforms and singular convolution operators on Lipschitz surfaces. Rev Mat Iberoamericana, 1994, 10: 665–721

13 Mallat S, Zhang Z. Matching pursuits with time-frequency dictionaries. IEEE Trans Signal Processing, 1993, 41: 3397–3415

14 Mitrea M. Clifford Wavelets, Singular Integrals, and Hardy Spaces. Berlin: Springer-Verlag, 1994

15 Qian T. Intrinsic mono-component decomposition of functions: An advance of Fourier theory. Math Meth Appl Sci, 2010, 33: 880–891

16 Qian T, Li H, Stessin M. Comparison of adaptive mono-component decompositions. Nonlinear Anal RWA, 2013, 14: 1055–1074

17 Qian T, Sprößig W, Wang J X. Adaptive Fourier decomposition of functions in quaternionic Hardy spaces. Math Meth Appl Sci, 2012, 35: 43–64

18 Qian T, Wang J X. Some remarks on the boundary behaviors of functions in the monogenic Hardy spaces. Adv Appl Clifford Algebras, 2012, 22: 819–826

19 Qian T, Wang J X, Yang Y. Matching pursuits among shifted Cauchy kernels in higher-dimensional spaces. Preprint

20 Qian T, Wang Y B. Adaptive Fourier series–a variation of greedy algorithm. Adv Comput Math, 2011, 34: 279–293

21 Sommen F. Spherical monogenic functions and analytic functionals on the unit sphere. Tokyo J Math, 1981, 4: 427–456

22 Temlyakov V N. Weak greedy algorithms. Adv Comput Math, 2000, 12: 213–227

23 Temlyakov V N. Greedy approximation. Acta Numerica, 2008, 17: 235–409

24 Temlyakov V N. Greedy Approximation. Cambridge: Cambridge University Press, 2011

25 Walsh J L. Interpolation and Approximation by Rational Functions in the Complex Domain. 4 edition. Providence R I: Amer Math Soc, 1965

26 Wang J X, Qian T. A variation of adaptive Fourier decomposition by higher order Szegö kernels I. complex variable cases. Preprint