MATCHING PURSUITS AMONG SHIFTED CAUCHY KERNELS IN HIGHER-DIMENSIONAL SPACES *

Tao QIAN (钱涛)   *Jinxun WANG* (王晋勋)
Department of Mathematics, University of Macau, Macao, China
E-mail: *fsttq@umac.mo; wjxpyh@gmail.com*

Yan YANG (杨燕)
School of Mathematics and Computational Science, Sun Yat-Sen University, Guangzhou 510275,
China
E-mail: *mathyy@sina.com*

Abstract   Appealing to the Clifford analysis and matching pursuits, we study the adaptive decompositions of functions of several variables of finite energy under the dictionaries consisting of shifted Cauchy kernels. This is a realization of matching pursuits among shifted Cauchy kernels in higher-dimensional spaces. It offers a method to process signals in arbitrary dimensions.

Key words   Hardy space; monogenic; adaptive decomposition; dictionary; matching pursuit; optimal approximation by rational functions

2010 MR Subject Classification   30G35; 30H10; 41A20; 41A50; 42B30

1   Introduction

Decomposition of a given signal into certain basic components is important in signal analysis, which allows us to easily understand the structure of the signal and get hidden structure information behind the signal. Traditional methods like Fourier transform, wavelet transform, or others using fixed bases, are powerful linear expansions in most cases, but they are not flexible enough. The convergence rate can not make people satisfied in many cases, and it is hard to detect the signal patterns from the expansion coefficients. For these reasons and for the purpose of finding fast ways to decompose signals, Mallat and Zhang introduced an adaptive algorithm in 1993 (see [1]), called matching pursuit, which is a greedy algorithm that decomposes any signal into a linear combination of waveforms which belong to a redundant dictionary of functions, where by definition, a dictionary in a Hilbert space $\mathcal{H}$ is a family of unit vectors $\mathscr{D} = \{g_\gamma : \gamma \in \Gamma, \|g_\gamma\| = 1\}$, which is often assumed to be complete or even redundant. Although the algorithm itself is nonlinear, it is adaptive because at each iteration it

chooses a waveform that almost best matches the signal structures, thus it provides a flexible representation of the signal.

Note that vectors in an arbitrarily given dictionary are generally not mutually orthogonal. However, an orthonormal decomposition can offer a special case of matching pursuit (or greedy algorithm) with dictionary consisting of an orthonormal basis. In [2], the authors studied the adaptive intrinsic mono-component decomposition of functions by the orthonormal rational system, i.e., the Takenaka–Malmquist system, in the context of complex Hardy space in the unit disc. The case of upper half plane, as a counterpart, first appeared in [3]. The algorithm in both cases guarantees that the optimal projection (maximal energy) can be obtained at each recursive step, being a variation of greedy algorithm with the convergence order being $1/2$ under certain conditions (cf. [4]). As to some subsequent works along this direction, see [5, 6].

Given the work in [2] and [3] it naturally arises the question what happens for the same type of signal decomposition in the higher-dimensional spaces? Recently, the same type of adaptive decompositions of functions of two or three variables are obtained in [7] by means of quaternionic analysis as a natural extension of the complex analysis case. The method, as in the complex number case, produces an orthonormal system. Unfortunately, the method can not be directly adapted to a general $(m+1)$-dimensional space in the Clifford analysis setting, for, in general, a Clifford number is not invertible, that stands as a major technical obstacle.

For the above reasons, in this paper, under the Clifford algebra setting, we study adaptive decompositions of functions in the monogenic Hardy spaces using dictionaries consisting of shifted Cauchy kernels (the Szegö kernels for the unit ball case, see [8], also [9]). Through Hardy spaces decomposition of signals the type of decompositions is thus valid for functions of finite energy in the $(m+1)$-dimensional Euclidean space. When $m+1$ is even, what we obtain is a rational approximation, although in general not orthogonal. For the special case $m = 1$, viz., in one complex variable, shifted Cauchy kernels are referred as the Cauchy wavelet system [10]. For model reduction practice in control theory the type of questions is equivalent with best rational approximation to functions in the Hardy spaces [11].

Our algorithm is in the spirit of matching pursuit. The previous algorithms in [2], [3] and [7], however, are combinations of generalized backward shift and matching pursuit. Thus in the present context we do not result in an orthonormal system in each case. This flexibility avoids the technical difficulty mentioned above. From [4] we know that the convergence order is $1/6$ under some additional assumptions.

In practice, the signal we study may be determined by more than one factor, that is to say, the signal is of several variables (e.g., an image can be regarded as a signal of two variables). Thus our method suggests an effective way for this, and from [12] we know that each basic component (shifted Cauchy kernel) we obtain from the decomposition possesses positive instantaneous frequency.

Our writing plan is as follows. Section 2 contains some basic knowledge on Clifford analysis and monogenic Hardy spaces. In Section 3 we study matching pursuits for functions in $\mathcal{H}^2(\mathbb{B}^m)$, from which we derive the adaptive decomposition of functions in $L^2(\mathbb{S}^m)$ in Section 4. The cases $\mathcal{H}^2(\mathbb{R}^{m+1}_+)$ and $L^2(\mathbb{R}^m)$, are discussed in Section 5. In the last section, we point out that the problem in $\mathcal{H}^2(\mathbb{R}^{m+1}_+)$ can be converted to that in $\mathcal{H}^2(\mathbb{B}^m)$, and vice versa.

2   Clifford Analysis and Monogenic Hardy Spaces

2.1   Clifford Algebra and Clifford Analysis

Let $e_1, \ldots, e_m$ be basic elements satisfying $e_i e_j + e_j e_i = -2\delta_{ij}, i, j = 1, \ldots, m$, where δ_{ij} equals 1 if $i = j$ and 0 otherwise. Let $\mathbb{R}^{m+1} = \{x = x_0 + x_1 e_1 + \cdots + x_m e_m : x_i \in \mathbb{R}, 0 \le i \le m\}$ be identified with the usual $(m + 1)$-dimensional Euclidean space. The real Clifford algebra generated by $e_1, \ldots, e_m$, denoted by $\mathscr{A}_m$, is an associative algebra in which each element is of the form $x = \sum_T x_T e_T$, where $x_T \in \mathbb{R}$, $e_T = e_{i_1} e_{i_2} \cdots e_{i_l}$ and $T = \{1 \le i_1 < i_2 < \cdots < i_l \le m\}$ runs over all ordered subsets of $\{1, \ldots, m\}$ and $x_\emptyset = x_0$, $e_\emptyset = e_0 = 1$. $\operatorname{Sc} x = x_0$ is called the scalar part of x, and $\operatorname{NSc} x = x - \operatorname{Sc} x$ is the non-scalar part of x, the conjugate and the norm of x are defined by $\overline{x} = \sum_T x_T \overline{e_T}$ and $|x| = \operatorname{Sc}(x\overline{x}) = (\sum_T x_T^2)^{1/2}$ respectively, where $\overline{e_T} = \overline{e_{i_l}} \cdots \overline{e_{i_2}} \, \overline{e_{i_1}}$ and $\overline{e_i} = -e_i$ for $i \ne 0$, $\overline{e_0} = e_0$. We have for any $x, y, z \in \mathscr{A}_m$, $\overline{xy} = \overline{y}\,\overline{x}$, $(xy)z = x(yz)$ and $|xy| \le 2^{m/2}|x||y|$. $\mathscr{A}_1 = \mathbb{C}$ and $\mathscr{A}_2$ coincides with the quaternion algebra $\mathbb{H}$. In general $\mathscr{A}_m$ is not a division algebra.

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 each component of f is a real-valued harmonic function, and f remains left (right) monogenic after right (left) multiplying f by a constant in $\mathscr{A}_m$.

The theory of Clifford analysis developed by Brackx, Delanghe and Sommen 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 ([8])   Let $S \subset \Omega$ be an $(m + 1)$-dimensional compact differentiable and oriented manifold with boundary. If f is left monogenic in Ω, then

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

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 the theory of left monogenic functions is analogous to that of right monogenic functions, we will work on left monogenic functions in this paper.

Denote the unit ball $\{x \in \mathbb{R}^{m+1} : |x| < 1\}$ which is centered at the origin by $\mathbb{B}^m$, and let the half space $\mathbb{R}^{m+1}_+ = \{x \in \mathbb{R}^{m+1} : \operatorname{Sc} x > 0\}$. The monogenic Hardy space $\mathcal{H}^2(\mathbb{B}^m)$ (resp. $\mathcal{H}^2(\mathbb{R}^{m+1}_+)$), consists of all functions f that are left monogenic in $\mathbb{B}^m$ (resp. $\mathbb{R}^{m+1}_+$), for which

$$\sup_{0 \le r < 1} \left(\frac{1}{\omega_m} \int_{|\eta|=1} |f(r\eta)|^2 dS \right)^{1/2} < \infty$$

$$\left(\text{resp.} \ \sup_{x_0 > 0} \left(\frac{1}{\omega_m} \int_{\mathbb{R}^m} |f(x_0 + \underline{x})|^2 d\underline{x} \right)^{1/2} < \infty \right),$$

here $\underline{x} = x_1 e_1 + \cdots + x_m e_m$, $d\underline{x} = dx_1 \cdots dx_m$ is the volume element of $\mathbb{R}^m$.

For any $f \in \mathcal{H}^2(\mathbb{B}^m)$ (resp. $\mathcal{H}^2(\mathbb{R}^{m+1}_+)$), the non-tangential limit of f on $\mathbb{S}^m$ (resp. $\mathbb{R}^m$) exists and belongs to $L^2(\mathbb{S}^m)$ (resp. $L^2(\mathbb{R}^m)$). Moreover, the Cauchy's integral formula holds for f.

The monogenic Hardy space $\mathcal{H}^2(\mathbb{B}^m)$ (resp. $\mathcal{H}^2(\mathbb{R}^{m+1}_+)$) is associated with the inner product and norm which are respectively defined by

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

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

and

$$\|f\| = (\mathrm{Sc}\langle f, f \rangle)^{1/2} = \left(\frac{1}{\omega_m} \int_{\mathbb{S}^m} |f(\eta)|^2 dS \right)^{1/2}, \quad f \in \mathcal{H}^2(\mathbb{B}^m)$$

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

There holds the quasi Cauchy–Schwarz inequality

$$|\langle f, g \rangle| \leq 2^{m/2} \|f\| \|g\|.$$

For more information about the monogenic Hardy spaces please see [13, 14].

3   Matching Pursuit in $\mathcal{H}^2(\mathbb{B}^m)$

In this section, we focus our attention on the following dictionary:

$$\mathscr{D} = \{\varphi_a : a \in \mathbb{B}^m\},$$

where the function φ_a is defined by

$$\varphi_a(x) = (1 - |a|^2)^{\frac{m}{2}} \frac{\overline{1 - \overline{a}x}}{|1 - \overline{a}x|^{m+1}}, \quad x \in \mathbb{B}^m.$$

It is easy to check that φ_a is left monogenic, and obviously φ_a is continuous in the closure of $\mathbb{B}^m$ whenever $a \in \mathbb{B}^m$, hence $\varphi_a \in \mathcal{H}^2(\mathbb{B}^m)$. To see that $\|\varphi_a\| = 1$, we first show that

Proposition 3.1   For any function $f \in \mathcal{H}^2(\mathbb{B}^m)$ and any $\varphi_a \in \mathscr{D}$, there holds

$$\langle f, \varphi_a \rangle = (1 - |a|^2)^{\frac{m}{2}} f(a).$$

Proof   By definition and the Cauchy's integral formula for $\mathcal{H}^2(\mathbb{B}^m)$,

$$\begin{aligned}
\langle f, \varphi_a \rangle &= \frac{1}{\omega_m} \int_{\mathbb{S}^m} (1 - |a|^2)^{\frac{m}{2}} \frac{1 - \overline{a}\eta}{|1 - \overline{a}\eta|^{m+1}} f(\eta) dS \\
&= (1 - |a|^2)^{\frac{m}{2}} \frac{1}{\omega_m} \int_{\mathbb{S}^m} \frac{\overline{\eta - a}}{|\eta - a|^{m+1}} (\eta dS) f(\eta) \\
&= (1 - |a|^2)^{\frac{m}{2}} f(a).
\end{aligned}$$

$\square$

The above property implies

$$\|\varphi_a\|^2 = \mathrm{Sc}(\langle \varphi_a, \varphi_a \rangle) = \mathrm{Sc}((1-|a|^2)^{\frac{m}{2}} \varphi_a(a)) = \mathrm{Sc}(1) = 1.$$

Next we will study the adaptive decomposition for functions in $\mathcal{H}^2(\mathbb{B}^m)$ with the dictionary $\mathscr{D}$. The algorithm is, in fact, the pure greedy algorithm. We, however, prove that the optimality factor can be taken as 1. That is to say, at each step we can get the projection of maximal energy that decomposes the function in a fast way. To be specific, let $f \in \mathcal{H}^2(\mathbb{B}^m)$, $\varphi_{a_0} \in \mathscr{D}$, clearly,

$$f = \varphi_{a_0} \langle f, \varphi_{a_0} \rangle + Rf,$$

where

$$Rf = f - \varphi_{a_0} \langle f, \varphi_{a_0} \rangle$$

is the residual vector. Note that φ_{a_0} is orthogonal to Rf, we have

$$\|f\|^2 = |\langle f, \varphi_{a_0} \rangle|^2 + \|Rf\|^2.$$

Doing the same procedure on Rf, one yields

$$Rf = \varphi_{a_1} \langle Rf, \varphi_{a_1} \rangle + R^2 f,$$

where

$$R^2 f = Rf - \varphi_{a_1} \langle Rf, \varphi_{a_1} \rangle,$$

and we get

$$\|Rf\|^2 = |\langle Rf, \varphi_{a_1} \rangle|^2 + \|R^2 f\|^2.$$

Let $R^0 f = f$, by induction, at the nth step we will get

$$R^n f = \varphi_{a_n} \langle R^n f, \varphi_{a_n} \rangle + R^{n+1} f,$$

and

$$\|R^n f\|^2 = |\langle R^n f, \varphi_{a_n} \rangle|^2 + \|R^{n+1} f\|^2.$$

So

$$f = \sum_{k=0}^{n} \varphi_{a_k} \langle R^k f, \varphi_{a_k} \rangle + R^{n+1} f,$$

and

$$\|f\|^2 = \sum_{k=0}^{n} |\langle R^k f, \varphi_{a_k} \rangle|^2 + \|R^{n+1} f\|^2.$$

In order to get a fast decomposition, we should select an $a_n \in \mathbb{B}^m$ at the nth iteration such that $|\langle R^n f, \varphi_{a_n} \rangle|$ takes as large as possible. The best choice is that

$$|\langle R^n f, \varphi_{a_n} \rangle| = \sup_{a \in \mathbb{B}^m} |\langle R^n f, \varphi_a \rangle|, \tag{3.1}$$

this is guaranteed by the following lemma.

Lemma 3.2 Suppose $f \in \mathcal{H}^2(\mathbb{B}^m)$, $a = |a|\xi = r\xi \in \mathbb{B}^m$, then

$$\lim_{r \to 1^-} |\langle f, \varphi_a \rangle| = 0$$

holds uniformly in $|\xi| = 1$.

Proof Let $V_r = C_m(1-r)^{m+1}$ be the volume of the ball $\mathbb{B}^m(a, 1-r) = \{x \in \mathbb{R}^{m+1} : |x - a| < 1 - r\}$, in which C_m is a constant depends only on m and it may differ at each appearance. Write $x = |x|\eta = \rho\eta$, note that

$$|x - a| \geq ||x| - |a|| = |\rho - r|$$

and

$$|x - a| = |\rho\eta - r\xi| = |r(\eta - \xi) - (r - \rho)\eta| \geq r|\eta - \xi| - |r - \rho| \geq r|\eta - \xi| - |x - a|,$$

so $x \in \mathbb{B}^m(a, 1-r)$ implies

$$\begin{cases} 2r - 1 < \rho < 1, \\ |\eta - \xi| < 2(1-r)/r. \end{cases}$$

Hence, when r is sufficiently close to 1, we have

$$|\langle f, \varphi_a \rangle| = |(1 - r^2)^{m/2} f(a)|$$

$$= (1 - r^2)^{m/2} \left| V_r^{-1} \int_{\mathbb{B}^m(a, 1-r)} f(x)dx \right|$$

$$\leq (1 - r^2)^{m/2} \left(V_r^{-1} \int_{\mathbb{B}^m(a, 1-r)} |f(x)|^2 dx \right)^{1/2}$$

$$\leq (1 - r^2)^{m/2} \left(V_r^{-1} \int_{2r-1}^1 \rho^m \int_{|\eta - \xi| < 2(1-r)/r} |f(\rho\eta)|^2 dS d\rho \right)^{1/2}$$

$$\leq (1 - r^2)^{m/2} \left(V_r^{-1} 2(1-r) \sup_{0 \leq \rho < 1} \int_{|\eta - \xi| < 2(1-r)/r} |f(\rho\eta)|^2 dS \right)^{1/2}$$

$$\leq C_m \left(\int_{|\eta - \xi| < 2(1-r)/r} \sup_{0 \leq \rho < 1} |f(\rho\eta)|^2 dS \right)^{1/2}.$$

Note that as a function of η, $\sup\limits_{0 \leq \rho < 1} |f(\rho\eta)|$ belongs to $L^2(\mathbb{S}^m)$, and the measure of the set $\{\eta : |\eta - \xi| < 2(1-r)/r\}$ tends to zero as $r \to 1^-$. The lemma then follows from the absolute continuity of the Lebesgue integral. $\qquad \square$

The above lemma enables us to find the optimal projections that best match the residues under the condition (3.1), and thus get a fast decomposition. Here we note that there may be more than one parameter satisfying (3.1).

Theorem 3.3 Under the maximum selection criterion (3.1), we have

$$\lim_{n \to \infty} \|R^n f\| = 0. \tag{3.2}$$

Proof From [1] we know that there exists a function g such that

$$g = \sum_{k=0}^{\infty} \varphi_{a_k} \langle R^k f, \varphi_{a_k} \rangle$$

holds in the sense of energy, and

$$\|g\|^2 = \sum_{k=0}^{\infty} |\langle R^k f, \varphi_{a_k} \rangle|^2.$$

If (3.2) is not true, then $h := f - g \neq 0$, so one can find at least one point $b \in \mathbb{B}^m \setminus \bigcup_{i=0}^{\infty}\{a_i\}$ such that

$$|\langle h, \varphi_b \rangle| = \delta > 0.$$

Set

$$g_N = -\sum_{k=N}^{\infty} \varphi_{a_k} \langle R^k f, \varphi_{a_k} \rangle.$$

Choose a large N such that

$$|\langle R^N f, \varphi_{a_N} \rangle| < \delta/2$$

and

$$|\langle g_N, \varphi_b \rangle| \leq 2^{m/2} \|g_N\| \|\varphi_b\| = 2^{m/2} \|g_N\| < \delta/2.$$

Thus

$$|\langle R^N f, \varphi_b \rangle| = |\langle h - g_N, \varphi_b \rangle| \geq |\langle h, \varphi_b \rangle| - |\langle g_N, \varphi_b \rangle| > \delta - \delta/2 = \delta/2,$$

so

$$|\langle R^N f, \varphi_b \rangle| > |\langle R^N f, \varphi_{a_N} \rangle|.$$

This is contrary with our choice of a_N for $R^N f$ based on the maximum selection criterion (3.1).

$\square$

Remark 3.4   Theorem 3.3 states that after matching pursuit for $f \in \mathcal{H}^2(\mathbb{B}^m)$ we obtain

$$f = \sum_{k=0}^{\infty} \varphi_{a_k} \langle R^k f, \varphi_{a_k} \rangle, \quad \|f\|^2 = \sum_{k=0}^{\infty} |\langle R^k f, \varphi_{a_k} \rangle|^2,$$

which implies that $\mathscr{D}$ is a complete dictionary, i.e., the closed linear span $\overline{\text{span}\mathscr{D}}$ of $\mathscr{D}$ is $\mathcal{H}^2(\mathbb{B}^m)$. However, we can prove this in a simpler and more direct way. In fact, it is not hard to see from Proposition 3.1 that the orthogonal complement of $\mathscr{D}$ is $\{0\}$ (cf. [2]). Hence we conclude that $\overline{\text{span}\mathscr{D}} = \mathcal{H}^2(\mathbb{B}^m)$.

Example 3.5   Let $f(x) = \frac{(1-0.75x_0+0.25x)(1-0.5\overline{x})}{|1-0.5x|^5}$, $x = x_0 + x_1 e_1 + x_2 e_2 \in \mathbb{B}^2$, then $f \in \mathcal{H}^2(\mathbb{B}^2)$ and $\|f\|^2 = 32/9 \approx 3.5556$. The adaptive decomposition of f according to the selection principle (3.1) yields

$$f(x) = \sum_{n=0}^{4} \varphi_{a_n}(x) c_n + R^5 f = \sum_{n=0}^{4} \varphi_{a_n}(x) \langle R^n f, \varphi_{a_n} \rangle + R^5 f, \quad x \in \mathbb{B}^2,$$

where the parameters and coefficients are showed in Table 1. The energy of the residue $R^5 f$ is $\|R^5 f\|^2 = \|f\|^2 - \sum_{n=0}^{4} |c_n|^2 \approx 0.0099$.

Table 1   Values for the parameters and coefficients

i	the ith parameter a_i	the ith coefficient c_i
0	0.6458	1.8779
1	-0.1948	-0.1030
2	$0.6718 + 0.0336e_1 - 0.3231e_2$	$0.0557 + 0.0016e_1 - 0.0157e_2$
3	$0.9138 + 0.0029e_1 - 0.0077e_2$	$-0.0502 - 0.0001e_1 - 0.0045e_2$
4	$0.6913 - 0.0335e_1 + 0.3373e_2$	$0.0497 - 0.0019e_1 + 0.0202e_2 - 0.0001e_1e_2$

On $\mathbb{B}^2$, f admits the spherical monogenic expansion, i.e.,

$$f(x) = \sum_{k=0}^{\infty} P_k(f, x) = \sum_{k=0}^{5} P_k(f, x) + R_6 f, \quad x \in \mathbb{B}^2,$$

where P_k is the inner spherical monogenic of order k, $R_6 f$ is the remainder term for the expansion up to order 5. Computation shows $\|\sum_{k=0}^{5} P_k\|^2 = \sum_{k=0}^{5} \|P_k\|^2 = 7195/2048$, so $\|R_6 f\|^2 = \|f\|^2 - \sum_{k=0}^{5} \|P_k\|^2 \approx 0.0424$. It is obvious that our method works better for f than that using spherical monogenics.

4   Adaptive Decompositions for Functions in $L^2(\mathbb{S}^m)$

If a function $F \in \mathcal{H}^2(\mathbb{B}^m)$, then its non-tangential boundary limit is a function f in $L^2(\mathbb{S}^m)$. According to Section 3, the adaptive decomposition for f will be obtained by taking the boundary values after performing matching pursuit on F. But conversely, if we are given a function $f \in L^2(\mathbb{S}^m)$, in general it is not the restriction of some function $F \in \mathcal{H}^2(\mathbb{B}^m)$ on $\mathbb{S}^m$. In this case, f can be split into several components such that each component corresponds to a real-valued function. Therefore, without loss of generality, we may assume that f itself is real-valued. Under this condition, there exists an $F \in \mathcal{H}^2(\mathbb{B}^m)$ with the property

$$\lim_{r \to 1^-} \mathrm{Sc}(F(r\xi)) = f(\xi), \quad \text{a.e. on } \mathbb{S}^m. \tag{4.1}$$

Hence, taking the scalar part of the boundary limit of the partial sums arising from the adaptive decomposition for F will lead to a fast approximation of f.

A solution of (4.1) is given by (cf. [8, 15])

$$F(x) = T(f)(x) = \int_{|\omega|=1} S(x, \omega) f(\omega) dS, \quad |x| < 1,$$

where $S(x, \omega) = P(x, \omega) + Q(x, \omega)$ is the monogenic Schwarz kernel, $P(x, \omega) = \frac{1}{\omega_m} \frac{1 - |x|^2}{|x - \omega|^{m+1}}$ is the Poisson kernel and

$$Q(x, \omega) = \mathrm{NSc}\left(\int_0^1 t^{m-1}(\overline{D}P)(tx, \omega)x dt\right)$$

$$= \left(\frac{1}{\omega_m} \int_0^1 \frac{(m+1)t^{m-1}(1 - t^2|x|^2)}{|tx - \omega|^{m+3}} dt\right) \mathrm{NSc}(\overline{\omega}x)$$

is the Cauchy-type harmonic conjugate of $P(x, \omega)$ on the unit sphere, which can be computed out explicitly with an expression in terms of elementary functions (cf. [16, 17]). As a consequence of boundedness of Hilbert transform on the sphere [15], T is a bounded operator from $L^2(\mathbb{S}^m)$ to $\mathcal{H}^2(\mathbb{B}^m)$. We note that F plus any non-scalar constant still satisfies (4.1).

5   Matching Pursuit in $\mathcal{H}^2(\mathbb{R}_+^{m+1})$

In this section, we will investigate a matching pursuit algorithm for functions in $\mathcal{H}^2(\mathbb{R}_+^{m+1})$ and the adaptive decomposition for functions in $L^2(\mathbb{R}^m)$.

Let $f \in L^2(\mathbb{R}^m)$. In order to get an adaptive approximation for f, let us first consider the Cauchy integral of f:

$$F(x) = C(f) = \frac{-1}{\omega_m} \int_{\mathbb{R}^m} \frac{\overline{y - x}}{|\underline{y} - x|^{m+1}} f(\underline{y}) d\underline{y}, \quad x \in \mathbb{R}_+^{m+1},$$

then $F \in \mathcal{H}^2(\mathbb{R}^{m+1}_+)$. The well known Sokhotsky–Plemelj formula reads

$$\lim_{x_0 \to 0^+} F(x_0 + \underline{x}) = \frac{1}{2}f(\underline{x}) + \frac{1}{2}H(f)(\underline{x}),$$

where $H(f) = \sum_{i=1}^m e_i R_i(f)$, and

$$R_i(f)(\underline{x}) = \frac{2}{\omega_m} \text{p.v.} \int_{\mathbb{R}^m} \frac{y_i - x_i}{|\underline{y} - \underline{x}|^{m+1}} f(\underline{y}) d\underline{y}$$

is the ith $(1 \le i \le m)$ Riesz transform of f. Consequently,

$$f(\underline{x}) = 2\text{Sc}\left(\lim_{x_0 \to 0^+} F(x_0 + \underline{x})\right).$$

Hence, the adaptive decomposition of f is now turned to the adaptive decomposition of F.

The dictionary we choose for this case is

$$\mathscr{D} = \left\{ \varphi_a = (2\text{Sc}(a))^{\frac{m}{2}} \frac{\overline{x + \overline{a}}}{|x + \overline{a}|^{m+1}} : a \in \mathbb{R}^{m+1}_+ \right\},$$

where each φ_a is a unit vector of $\mathcal{H}^2(\mathbb{R}^{m+1}_+)$. For any $f \in \mathcal{H}^2(\mathbb{R}^{m+1}_+)$, invoking Cauchy's integral formula for $\mathcal{H}^2(\mathbb{R}^{m+1}_+)$, we get

$$\begin{aligned}
\langle f, \varphi_a \rangle &= \frac{1}{\omega_m} \int_{\mathbb{R}^m} (2a_0)^{\frac{m}{2}} \frac{\underline{y} + \overline{a}}{|\underline{y} + \overline{a}|^{m+1}} f(\underline{y}) d\underline{y} \\
&= (2a_0)^{\frac{m}{2}} \frac{-1}{\omega_m} \int_{\mathbb{R}^m} \frac{\overline{\underline{y} - a}}{|\underline{y} - a|^{m+1}} f(\underline{y}) d\underline{y} \\
&= (2a_0)^{\frac{m}{2}} f(a),
\end{aligned} \tag{5.1}$$

where $a_0 = \text{Sc}(a)$.

The corresponding maximum selection criterion we adopt for this setting is to find an $a_n \in \mathbb{R}^{m+1}_+$ at each stage such that

$$|\langle R^n f, \varphi_{a_n} \rangle| = \sup_{a \in \mathbb{R}^{m+1}_+} |\langle R^n f, \varphi_a \rangle|. \tag{5.2}$$

The attainability of the supremum is concluded from the following proposition whose proof is a slight modification of one in [7]. We would like to write the proof for completeness.

Proposition 5.1 Suppose $f \in \mathcal{H}^2(\mathbb{R}^{m+1}_+)$, then

$$\lim_{a_0 \to 0^+} \langle f, \varphi_a \rangle = \lim_{a_0 \to +\infty} \langle f, \varphi_a \rangle = 0$$

holds uniformly with respect to $\underline{a} \in \mathbb{R}^m$.

Proof Let $V_{a_0} = C_m a_0^{m+1}$ be the volume of the ball $\mathbb{B}^m(a, \frac{a_0}{2})$, then

$$\begin{aligned}
|\langle f, \varphi_a \rangle| &= |(2a_0)^{m/2} f(a)| \\
&= (2a_0)^{m/2} \left| V_{a_0}^{-1} \int_{\mathbb{B}^m(a, \frac{a_0}{2})} f(x_0 + \underline{x}) dx \right| \\
&\le (2a_0)^{m/2} \left(V_{a_0}^{-1} \int_{\mathbb{B}^m(a, \frac{a_0}{2})} |f(x_0 + \underline{x})|^2 dx \right)^{1/2} \tag{5.3} \\
&\le (2a_0)^{m/2} \left(V_{a_0}^{-1} \int_{\frac{a_0}{2}}^{\frac{3a_0}{2}} \int_{\mathbb{R}^m} |f(x_0 + \underline{x})|^2 d\underline{x} dx_0 \right)^{1/2} \tag{5.4}
\end{aligned}$$

$$\leq (2a_0)^{m/2} \left(a_0 V_{a_0}^{-1} \sup_{x_0 > 0} \int_{\mathbb{R}^m} |f(x_0 + \underline{x})|^2 d\underline{x} \right)^{1/2}$$

$$= C_m \|f\|. \tag{5.5}$$

On the other hand, when a_0 is small,

$$(5.3) \leq (2a_0)^{m/2} \left(V_{a_0}^{-1} \int_{\frac{a_0}{2}}^{\frac{3a_0}{2}} \int_{|\underline{x} - \underline{a}| \leq \frac{a_0}{2}} |f(x_0 + \underline{x})|^2 d\underline{x} dx_0 \right)^{1/2}$$

$$\leq (2a_0)^{m/2} \left(a_0 V_{a_0}^{-1} \sup_{x_0 \in (\frac{a_0}{2}, \frac{3a_0}{2})} \int_{|\underline{x} - \underline{a}| \leq \frac{a_0}{2}} |f(x_0 + \underline{x})|^2 d\underline{x} \right)^{1/2}$$

$$\leq C_m \left(\int_{|\underline{x} - \underline{a}| \leq \frac{a_0}{2}} \sup_{x_0 \in (\frac{a_0}{2}, \frac{3a_0}{2})} |f(x_0 + \underline{x})|^2 d\underline{x} \right)^{1/2}$$

$$\leq C_m \left(\int_{|\underline{x} - \underline{a}| \leq \frac{a_0}{2}} \sup_{x_0 > 0} |f(x_0 + \underline{x})|^2 d\underline{x} \right)^{1/2}.$$

Note that as a function of $\underline{x}$, $\sup_{x_0 > 0} |f(x_0 + \underline{x})| \in L^2(\mathbb{R}^m)$ and the measure $|\{\underline{x} : |\underline{x} - \underline{a}| \leq \frac{a_0}{2}\}| \to 0$ as $a_0 \to 0^+$, by the absolute continuity of the Lebesgue integral we have

$$\lim_{a_0 \to 0^+} \langle f, \varphi_a \rangle = 0.$$

When a_0 is large,

$$(5.4) \leq (2a_0)^{m/2} \left(a_0 V_{a_0}^{-1} \sup_{x_0 \in (\frac{a_0}{2}, \frac{3a_0}{2})} \int_{\mathbb{R}^m} |f(x_0 + \underline{x})|^2 d\underline{x} \right)^{1/2}$$

$$\leq C_m \left(\int_{\mathbb{R}^m} G(a_0, \underline{x})^2 d\underline{x} \right)^{1/2}$$

holds uniformly for $\underline{a} \in \mathbb{R}^m$, where $G(a_0, \underline{x}) = \sup_{x_0 \in (\frac{a_0}{2}, \frac{3a_0}{2})} |f(x_0 + \underline{x})|$. Obviously,

$$G(a_0, \underline{x}) \leq \sup_{x_0 > 0} |f(x_0 + \underline{x})| \in L^2(\mathbb{R}^m).$$

Also, from (5.5) we know that

$$G(a_0, \underline{x}) \leq a_0^{-m/2} C_m \|f\|,$$

which implies that

$$\lim_{a_0 \to +\infty} G(a_0, \underline{x}) = 0$$

holds uniformly for $\underline{x} \in \mathbb{R}^m$. By Lebesgue's dominated convergence theorem we have

$$\lim_{a_0 \to +\infty} \langle f, \varphi_a \rangle = 0.$$

The proof of the proposition is complete.   □

Proposition 5.2   Suppose $f \in \mathcal{H}^2(\mathbb{R}_+^{m+1})$, then

$$\lim_{|\underline{a}| \to +\infty} \langle f, \varphi_a \rangle = 0$$

holds uniformly in $a_0 > 0$.

Proof By Proposition 5.1 and formula (5.1), it suffices to show that

$$\lim_{|\underline{a}| \to +\infty} |f(a_0 + \underline{a})| = 0$$

holds uniformly with respect to $a_0 \in [c, d] \subset \mathbb{R}_+$.

We have

$$|f(a_0 + \underline{a})| = \frac{1}{\omega_m} \left| \int_{\mathbb{R}^m} \frac{\overline{a_0 + \underline{a} - \underline{y}}}{|a_0 + \underline{a} - \underline{y}|^{m+1}} f(\underline{y}) d\underline{y} \right|$$

$$\leq \frac{1}{\omega_m} \int_{\mathbb{R}^m} \frac{|f(\underline{y})|}{(|\underline{y} - \underline{a}|^2 + a_0^2)^{m/2}} d\underline{y}$$

$$\leq \frac{1}{\omega_m} \left(\int_{|\underline{y}|>N} \frac{|f(\underline{y})|}{(|\underline{y} - \underline{a}|^2 + c^2)^{m/2}} d\underline{y} + \int_{|\underline{y}|\leq N} \frac{|f(\underline{y})|}{(|\underline{y} - \underline{a}|^2 + c^2)^{m/2}} d\underline{y} \right)$$

$$= \frac{1}{\omega_m} (I_1 + I_2).$$

By Hölder's inequality,

$$I_1 \leq \left(\int_{|\underline{y}|>N} \frac{1}{(|\underline{y} - \underline{a}|^2 + c^2)^m} d\underline{y} \right)^{1/2} \left(\int_{|\underline{y}|>N} |f(\underline{y})|^2 d\underline{y} \right)^{1/2}$$

$$\leq \left(\int_{\mathbb{R}^m} \frac{1}{(|\underline{y}|^2 + c^2)^m} d\underline{y} \right)^{1/2} \left(\int_{|\underline{y}|>N} |f(\underline{y})|^2 d\underline{y} \right)^{1/2}$$

$$\leq C_m \left(\int_{|\underline{y}|>N} |f(\underline{y})|^2 d\underline{y} \right)^{1/2},$$

so I_1 is small provided N is large enough. With N fixed,

$$I_2 \leq \frac{C_m}{|\underline{a}|^m} \int_{|\underline{y}|\leq N} |f(\underline{y})| d\underline{y} \to 0 \ (|\underline{a}| \to +\infty),$$

which is due to the fact that $f(\underline{y})$ is integrable on $\{\underline{y} : |\underline{y}| \leq N\}$. This finishes the proof of the proposition. □

Theorem 3.3 also holds for this case.

6   Isomorphism between $\mathcal{H}^2(\mathbb{B}^m)$ and $\mathcal{H}^2(\mathbb{R}_+^{m+1})$

The matching pursuits for $\mathcal{H}^2(\mathbb{B}^m)$ and $\mathcal{H}^2(\mathbb{R}_+^{m+1})$, as we suggest, are subject to the maximum selection criteria (3.1) and (5.2) respectively. In numerical computation, to find a maximizer for (3.1) is easy. For example, we can divide the unit ball by mesh, and then evaluate the norm of the coefficient for each point in each subgrid. However, for (5.2) it is hard to implement in practice, since $\mathbb{R}_+^{m+1}$ is an unbounded domain. In this section, we solve this problem by showing that $\mathcal{H}^2(\mathbb{B}^m)$ is isometrically isomorphic to $\mathcal{H}^2(\mathbb{R}_+^{m+1})$. As a consequence, it turns out that the optimization problem (5.2) is essentially equivalent to (3.1).

We start from a Möbius transformation τ (see [18]), which is defined through

$$b = \tau(a) = (1 - a)(1 + a)^{-1} = (1 + a)^{-1}(1 - a).$$

One can verify that τ is a bijection from $\mathbb{B}^m$ to $\mathbb{R}_+^{m+1}$, and its inversion is given by

$$a = \tau^{-1}(b) = \tau(b) = (1 - b)(1 + b)^{-1} = (1 + b)^{-1}(1 - b).$$

Define the linear operator $\mathcal{T}$ acting on functions by

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

whenever it makes sense. We have

Proposition 6.1   If φ is a shifted Cauchy kernel for $\mathbb{B}^m$ (resp. $\mathbb{R}_+^{m+1}$), then $\mathcal{T}(\varphi)$ is a shifted Cauchy kernel for $\mathbb{R}_+^{m+1}$ (resp. $\mathbb{B}^m$) modulo a unimodular constant.

Proof   Let $a \in \mathbb{B}^m$, $b = \tau(a) \in \mathbb{R}_+^{m+1}$, then one can deduce by direct computations that

$$\left(\mathcal{T} \left((1-|a|^2)^{\frac{m}{2}} \frac{\overline{1-\overline{a}\mathcal{X}}}{|1-\overline{a}\mathcal{X}|^{m+1}} \right) \right)(x) = \left((2\mathrm{Sc}(b))^{\frac{m}{2}} \frac{\overline{x+\overline{b}}}{|x+\overline{b}|^{m+1}} \right) \frac{\overline{1+b}}{|1+b|},$$

and

$$\left(\mathcal{T} \left((2\mathrm{Sc}(b))^{\frac{m}{2}} \frac{\overline{x+\overline{b}}}{|x+\overline{b}|^{m+1}} \right) \right)(\mathcal{X}) = \left((1-|a|^2)^{\frac{m}{2}} \frac{\overline{1-\overline{a}\mathcal{X}}}{|1-\overline{a}\mathcal{X}|^{m+1}} \right) \frac{\overline{1+a}}{|1+a|}.$$

$\square$

Theorem 6.2 $\mathcal{T}$ is a unitary operator from $\mathcal{H}^2(\mathbb{B}^m)$ to $\mathcal{H}^2(\mathbb{R}_+^{m+1})$, and $\mathcal{T}^{-1} = \mathcal{T}$.

Proof   Assume that $f \in \mathcal{H}^2(\mathbb{B}^m)$. Cauchy's integral formula gives

$$f(\mathcal{X}) = \frac{1}{\omega_m} \int_{\mathbb{S}^m} \frac{1-\overline{\mathcal{X}}\eta}{|1-\overline{\mathcal{X}}\eta|^{m+1}} f(\eta)dS, \quad \mathcal{X} \in \mathbb{B}^m.$$

Changing variables by setting $\mathcal{X} = \tau(x)$ and $\eta = \tau(\underline{y})$, then $\underline{y} \in \mathbb{R}^m$ and $dS(\eta) = \frac{2^m}{|1+\underline{y}|^{2m}}d\underline{y}$ (cf. [19]), from which it is not difficult to get

$$(\mathcal{T}f)(x) = \frac{-1}{\omega_m} \int_{\mathbb{R}^m} \frac{\overline{\underline{y}-x}}{|\underline{y}-x|^{m+1}} (\mathcal{T}f)(\underline{y})d\underline{y}, \quad x \in \mathbb{R}_+^{m+1}.$$

So $\mathcal{T}f$ is the Cauchy integral of its boundary limit (and hence it is monogenic on $\mathbb{R}_+^{m+1}$). Meanwhile, we obtain

$$\|f\|_{\mathcal{H}^2(\mathbb{B}^m)}^2 = \frac{1}{\omega_m} \int_{\mathbb{S}^m} |f(\underline{\eta})|^2 dS = \frac{1}{\omega_m} \int_{\mathbb{R}^m} \left| f(\tau(\underline{y})) \right|^2 \frac{2^m}{|1+\underline{y}|^{2m}} d\underline{y},$$

which means that $\mathcal{T}f|_{\mathbb{R}^m} \in L^2(\mathbb{R}^m)$ and $\|\mathcal{T}f|_{\mathbb{R}^m}\|_{L^2(\mathbb{R}^m)} = \|f\|_{\mathcal{H}^2(\mathbb{B}^m)}$. Consequently, $\mathcal{T}f \in \mathcal{H}^2(\mathbb{R}_+^{m+1})$ and $\|\mathcal{T}f\|_{\mathcal{H}^2(\mathbb{R}_+^{m+1})} = \|\mathcal{T}f|_{\mathbb{R}^m}\|_{L^2(\mathbb{R}^m)} = \|f\|_{\mathcal{H}^2(\mathbb{B}^m)}$. Moreover, $\langle \mathcal{T}f, \mathcal{T}g \rangle_{\mathcal{H}^2(\mathbb{R}_+^{m+1})} = \langle f, g \rangle_{\mathcal{H}^2(\mathbb{B}^m)}$ for any $f, g \in \mathcal{H}^2(\mathbb{B}^m)$. As to $\mathcal{T}^{-1} = \mathcal{T}$, it follows from the fact that $\mathcal{T}^2$ coincides with the identity operator. $\square$

References

[1] Mallat S G, Zhang Z. Matching pursuits with time-frequency dictionaries. IEEE Transactions on Signal Processing, 1993, **41**(12): 3397-3415

[2] Qian T, Wang Y B. Adaptive Fourier series–a variation of greedy algorithm. Advances in Computational Mathematics, 2011, **34**: 279-293

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

[4] DeVore R A, Temlyakov V N. Some remarks on greedy algorithms. Advances in Computational Mathematics, 1996, **5**: 173-187

[5] Qian T, Tan L H, Wang Y B. Adaptive decomposition by weighted inner functions: a generalization of Fourier series. J Fourier Anal Appl, 2011, **17**: 175-190

[6] Qian T, Wegert E. Optimal approximation by Blaschke forms. Complex Variables and Elliptic Equations, 2013, **58**: 123-133

[7] Qian T, Sprößig W, Wang J X. Adaptive Fourier decomposition of functions in quaternionic Hardy spaces. Math Meth Appl Sci, 2012, **35**: 43-64

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

[9] Cerejeiras P, Ferreira M, Kähler U, Sommen F. Continuous wavelet transform and wavelet frames on the sphere using Clifford analysis. Commun Pure Appl Anal, 2007, **6**: 619-641

[10] Sakaguchi F. A larger class of wavepacket eigenfunction systems which contains Cauchy wavelets and coherent states. In Proceedings of 8th IEEE Signal Processing Workshop on Statistical Signal and Array Processing, 1996, 436-439

[11] Mi W, Qian T, Wan F. A fast adaptive model reduction method based on Takenaka–Malmquist systems. Systems & Control Letters, 2012, **61**: 223-230

[12] Yang Y, Qian T, Sommen F. Phase derivative of monogenic signals in higher dimensional spaces. Complex Anal Oper Theory, 2012, **6**: 987-1010

[13] Gilbert J E, Murray A M. Clifford Algebras and Dirac Operators in Harmonic Analysis. Cambridge: Cambridge University Press, 1991

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

[15] Qian T, Yang Y. Hilbert transforms on the sphere with the Clifford algebra setting. J Fourier Anal Appl, 2009, **15**: 753-774

[16] Brackx F, De Knock B, De Schepper H, Eelbode D. On the interplay between the Hilbert transform and conjugate harmonic functions. Math Meth Appl Sci, 2006, **29**: 1435-1450

[17] Brackx F, Van Acker N. A conjugate Poisson kernel in Euclidean space. Simon Stevin, 1993, **67**: 3-14

[18] Ahlfors L V. Möbius Transformations in Several Dimensions. Minneapolis: University of Minnesota, 1981

[19] Qian T, Ryan J. Conformal transformations and Hardy spaces arising in Clifford analysis. J Operator Theory, 1996, **35**: 349-372