Algorithm of Adaptive Fourier Decomposition

Tao Qian, Liming Zhang, Zhixiong Li

***Abstract*—The present paper is a continuing work on the recently established Adaptive Fourier Decomposition (AFD) mainly stressing on the algorithm aspect, including algorithm analysis and numerical examples. AFD is a variation and realization of greedy algorithm (matching pursuit) suitable for the Hardy H^2 and the L^2 spaces. Applying AFD to a given signal, one obtains a series expansion in the basic signals, called mono-components, that possess non-negative analytic phase derivatives (functions), or, equivalently, meaningful instantaneous frequencies. AFD is shown to be robust with computational complexity comparable with DFT. Consistent to the greedy algorithm principle experiments show that AFD produces (pre-) mono-component series with efficient energy decay that also leads to efficient pointwise convergence in terms of computer running time.**

Hardy spaces, Hilbert transform, rational orthogonal system, analytic signal, instantaneous frequency, mono-components, matching pursuit, greedy algorithm, adaptive decomposition.

I. Introduction

For a given signal its Fourier expansion may converge slowly, for the entries $c_k e^{ikt}$ in the expansion that build up the essential part of the total energy may arrive late. One encounters the same problem whenever expanding an element against a fixed basis in a Hilbert space. To treat this problem greedy algorithm comes into play ([11], [7], [24]).

AFD is based on the rational orthogonal system, or the Takenaka-Malmquist system, $\{B_n\}_{n=1}^{\infty}$, where

$$\begin{aligned} B_n(z) &= B_{\{a_1,...,a_n\}}(z) \\ &:= \frac{1}{\sqrt{2\pi}} \frac{\sqrt{1-|a_n|^2}}{1-\overline{a}_n z} \prod_{k=1}^{n-1} \frac{z-a_k}{1-\overline{a}_k z}, \end{aligned}$$

$a_n \in \mathbb{D}, n = 1, 2, ..., \mathbb{D} = \{z \in \mathbb{C} : |z| < 1\}$, $\mathbb{C}$ is the complex plane. Note that for any sequence $\{a_n\}$ in $\mathbb{D}$ the system $\{B_n\}$ is orthonormal. If all a_n's are taken to be 0, then the system reduces to a half of the complex trigonometric (Fourier) system, viz. $\{\frac{z^{n-1}}{\sqrt{2\pi}}\}_{n=1}^{\infty}$. Functions B_n will be called *modified Blaschke products*. Relevant studies on the theory and applications of the system have a long history with a large quantity of literature ([2], [3]). All the existing studies of the system, however, are based on the so-called hyperbolic non-separability condition

$$\sum_{k=1}^{\infty}(1-|a_k|) = \infty. \tag{1}$$

The condition (1) is sufficient and necessary for the system to be complete in each of the complex Hardy spaces $H^p(\mathbb{D}), 1 \le p < \infty$, and in the disc algebra $A(\mathbb{D})$. In the present paper we concentrate in $H^2(\mathbb{D})$.

Apart from the greedy algorithm principle, the most striking character of AFD is its relation with non-negative analytic phase derivative. It is widely agreed that the analytic phase ϕ and its derivative ϕ' (frequency) of a real-valued signal s should be defined through its associated analytic signal $As = s + iHs = \rho e^{i\phi}$ via the Hilbert transformation H. A signal is said to be a (complex) *mono-component* if its analytic phase derivative, as a measurable function, is non-negative ([15]). In the case the signal is said to possess a well defined (analytic) instantaneous frequency ϕ' ([13], [5]). An analytic signal $\rho(t)e^{i\phi(t)}$ is said to be a (complex) *pre-mono-component* if $e^{iMt}\rho(t)e^{i\phi(t)}$ is a mono-component for some $M > 0$. It is asserted that not all analytic signals have a non-negative phase derivative (instantaneous frequency). For instance, each outer function is an analytic signal but it does not have a non-negative phase derivative function. It turns out that what one can do is to express an analytic signal in H^2 as a series of complex mono-components, or, equivalently, express a general function in L^2 as a series of mono-components and conjugate mono-components ([17], [19]). Mono-component decomposition is a generalization of Fourier decomposition. A large variety of mono-components have been found. The search for mono-components also motivated a new phase of study of Bedrosian's identity ([13], [14], [15], [16], [21], [22], [23], [25], [26], [6]). The study of mono-components and their applications naturally merges with the study of signal adaptive decomposition by using $\{B_n\}$. In fact, it may be easily

The work was supported by Macao FDCT 014/2008/A1 and research grant of the University of Macau No. RG-UL/07-08s/Y1/QT/FSTR.

T. Qian is with the Department of Mathematics, Faculty of Science and Technology, University of Macau, Taipa, Macao, China (e-mail: *fsttq@umac.mo*).

L. Zhang is with the Faculty of Education, University of Macau, Taipa, Macao, China (e-mail: *lmzhang@umac.mo*).

Z. Li is with the Faculty of Science and Technology, University of Macau, Taipa, Macao, China (e-mail: *ma96584@umac.mo*).

verified that if $a_1 = 0$, then for any $a_2, a_3, ...$ (inside the disc), all B_n are mono-components; while if $a_1 \neq 0$, then the B_n's may not be mono-components, but are at least pre-mono-components ([19]). Under the energy principle AFD gives rise to the intrinsic components of the signal in the modified Blaschke product model with increasing frequencies. References for traditional studies of TM systems and applications may be found in [1], [2], [3]. Applications of AFD to control theory is referred to [12].

§II sketches the mathematical theory of AFD. §III provides detailed description of the algorithm based on discrete data. §IV is devoted to a discussion of robustness and computational complexity of AFD. §V presents experimental examples with comparisons with Fourier expansions. In §VI conclusions are drawn. §VII is an Appendix proving the H^2-robustness of AFD stated in §4.

II. Mathematical Foundation of the AFD Algorithm

For the materials presented in this section we refer to [19], [17] and [20].

Denote by L^2 the Hilbert space of signals of finite energy on the closed interval $[0, 2\pi]$ equipped with the inner product

$$\langle \tilde{F}, \tilde{G} \rangle = \int_0^{2\pi} \tilde{F}(e^{it}) \overline{\tilde{G}}(e^{it}) dt. \tag{2}$$

Below, we identify the L^2 space on the interval $[0, 2\pi]$ with that on the unit circle $\partial\mathbb{D}$ with the Lebesgue arc-length measure.

If $\tilde{G} \in L^2$, then in the L^2-norm sense,

$$\tilde{G}(e^{it}) = \sum_{k=-\infty}^{\infty} c_k e^{ikt}, \quad c_k = \frac{1}{2\pi} \int_0^{2\pi} \tilde{G}(e^{it}) e^{-ikt} dt,$$

$$\|\tilde{G}\|_2^2 = 2\pi \sum_{k=-\infty}^{\infty} |c_k|^2.$$

The function $\tilde{G}$ has the direct sum decomposition $\tilde{G} = G^+ + G^-$, where G^+ and G^- are, respectively, the non-tangential boundary values on the unit circle of the analytic functions

$$G^+(z) = \sum_{k=0}^{\infty} c_k z^k, \ |z| < 1;$$

and

$$G^-(z) = \sum_{k=-\infty}^{-1} c_k z^k, \ |z| > 1.$$

Such functions G^+ and G^- constitute, respectively, the so called Hardy spaces H^+ and H^-. For basic knowledge of Hardy spaces we refer to [9].

The Hardy space projection G^+ of $\tilde{G}$ is given by the Cauchy integral:

$$\begin{aligned} G^+(z) &= \frac{1}{2\pi i} \int_{\partial\mathbb{D}} \frac{\tilde{G}(\zeta)}{\zeta - z} d\zeta \\ &= \frac{1}{2\pi} \int_0^{2\pi} \tilde{G}(e^{it}) \sum_{k=0}^{\infty} (ze^{-it})^k dt \\ &= c_0 + c_1 z + \cdots + c_n z^n + \cdots, \quad z \in \mathbb{D}. \end{aligned} \tag{3}$$

We will deal with two types of decompositions. One is for a given function $G = G^+ \in H^2$. In some applications, including system identification, one knows that the given data is for the boundary value of an analytic function. The algorithm for this case is called the *core algorithm*. The other type is for a given real-valued function $\tilde{G} \in L^2$. For this type one needs to first work out the projection $G^+ \in H^2$ of $\tilde{G}$, and apply the core algorithm to G^+. If $\tilde{G}$ is real-valued, then

$$\tilde{G}(e^{it}) = -c_0 + 2\mathrm{Re} G^+(e^{it}), \qquad \text{a.e.} \tag{4}$$

The decomposition of $\tilde{G}$ is then obtained from the decomposition of G^+ and the relation (4). A variation of the decomposition of $\tilde{G}$ does not go through G^+ and G_k, but uses the relation $\langle \tilde{G}, B_n \rangle = \langle G_n, e_{\{a_n\}} \rangle$ (for the definition of G_n, see (7)).

In AFD we have a "dictionary" consisting of the elementary functions

$$e_{\{a\}}(z) := B_{\{a\}}(z) = \frac{1}{\sqrt{2\pi}} \frac{\sqrt{1 - |a|^2}}{1 - \overline{a} z}, \quad a \in \mathbb{D}.$$

The function $e_{\{a\}}$ is called *the evaluator at* a. Each evaluator gives rise to, essentially, an evaluating functional. In fact, for any $F \in H^2$, by using the Cauchy Formula, we have

$$\begin{aligned} \langle F, e_{\{a\}} \rangle &= \sqrt{2\pi} \sqrt{1 - |a|^2} \frac{1}{2\pi i} \int_0^{2\pi} \frac{F(e^{it})}{e^{it} - a} de^{it} \\ &= \sqrt{2\pi} \sqrt{1 - |a|^2} F(a). \end{aligned}$$

The core algorithm starts from a given $G = G^+ \in H^2$. Setting $G_1 = G = G^+$, the first step is to maximize the projection $|\langle G_1, e_{\{a\}} \rangle|^2$ among all selections of $a \in \mathbb{D}$. In the sequel, for a general function $F \in H^2$, we adopt the notation

$$\begin{aligned} A_F^2(a) &:= |\langle F, e_{\{a\}} \rangle|^2 \\ &= 2\pi (1 - |a|^2) |F(a)|^2, \quad a \in \mathbb{D}. \end{aligned} \tag{5}$$

What is crucial is the ***Maximal Projection Principle*** asserting that for any $F \in H^2$ there exists $a_1 \in \mathbb{D}$ such that

$$a_1 = \arg\max \{ A_F^2(a) \ : \ a \in \mathbb{D} \}. \tag{6}$$

For a proof of this fact we refer to [19] or [17]. Now write

$$G(z) = G_1(z) = \langle G_1, e_{\{a_1\}} \rangle e_{\{a_1\}} + R_1(z),$$

where a_1 is selected according to the Maximal Projection Principle. The *standard remainder* $R_1(z)$ is thus minimized, where we have the factorization

$$R_1(z) = G_2(z)\frac{z-a_1}{1-\overline{a}_1 z},$$

with

$$G_2(z) = \left(G_1(z) - \langle G_1, e_{\{a_1\}}\rangle e_{\{a_1\}}(z)\right)\frac{1-\overline{a}_1 z}{z-a_1}.$$

Note that the *reduced remainder* $G_2(z)$ is still in H^2, because the difference $G_1(z) - \langle G_1, e_{\{a_1\}}\rangle e_{\{a_1\}}(z)$ has zero at $z = a_1$. We call the process to get G_2 from G_1 through the Maximal Projection Principle a ***maximal sifting***, and one in the same pattern but not through the Maximal Projection Principle a ***compulsory sifting***.

Applying a maximal sifting to G_2 we obtain G_3. Repeating such process to the nth step, we obtain

$$\begin{aligned} G(z) &= \sum_{k=1}^{n}\langle G_k, e_{\{a_k\}}\rangle B_{\{a_1,...,a_k\}}(z) + R_n(z) \\ &= \sum_{k=1}^{n}\langle G_k, e_{\{a_k\}}\rangle B_{\{a_1,...,a_k\}}(z) \\ &\quad + G_{n+1}(z)\prod_{k=1}^{n}\frac{z-a_k}{1-\overline{a}_k z}, \end{aligned}$$

where the reduced remainder G_{k+1} is obtained through the recursive formula

$$G_{k+1}(z) = \left(G_k(z) - \langle G_k, e_{\{a_k\}}\rangle e_{\{a_k\}}\right)\frac{1-\overline{a}_k z}{z-a_k}, \quad (7)$$

where

$$a_k = \arg\max\{A^2_{G_k}(a) \ : \ a \in \mathbb{D}\}. \quad (8)$$

The kth standard remainder has the expression

$$R_k(z) = G_{k+1}(z)\prod_{l=1}^{k}\frac{z-a_l}{1-\overline{a}_l z}. \quad (9)$$

The orthogonal properties imply

$$\|G\|^2 = \sum_{k=1}^{n} A^2_{G_k}(a_k) + \|R_n\|^2. \quad (10)$$

Under the consecutive maximal sifting processes we can actually show that in the L^2-norm sense

$$G(z) = \sum_{k=1}^{\infty}\langle G_k, e_{\{a_k\}}\rangle B_{\{a_1,...,a_k\}}(z) \quad (11)$$

([19] or [17]). There holds the Plancherel Theorem

$$\|G\|^2 = \sum_{k=1}^{\infty} A^2_{G_k}(a_k) = \sum_{k=1}^{\infty}|\langle G^+, B_k\rangle|^2.$$

For a given threshold $\epsilon > 0$, one sets to have the consecutive maximal sifting processes cease at the first N such that

$$\|R_N\|^2 = \|G\|^2 - \sum_{k=1}^{N} A^2_{G_k}(a_k) \le \epsilon. \quad (12)$$

The approximation by the Nth partial sum of the IMCs is

$$G^+(z) \approx \sum_{k=1}^{N}\frac{(1-|a_k|^2)G_k(a_k)}{1-\overline{a}_k z}\prod_{l=1}^{k-1}\frac{z-a_l}{1-\overline{a}_l z}. \quad (13)$$

With an error less than 2ϵ the relation (4) gives

$$\begin{aligned} \tilde{G}(e^{it}) &\approx 2\mathrm{Re}\sum_{k=1}^{N}\frac{(1-|a_k|^2)G_k(a_k)}{1-\overline{a}_k e^{it}}\prod_{l=1}^{k-1}\frac{e^{it}-a_l}{1-\overline{a}_l e^{it}} \\ &\quad - c_0. \end{aligned} \quad (14)$$

Remark 1 We can insert a compulsory sifting to change (13) to become a mono-component decomposition. In fact, denoting $B_n(e^{it}) = \rho_n(t)e^{i\theta_n(t)}$, $n = 1, 2, ...,$ we can easily show $\theta_n'(t) < \theta_{n+1}'(t)$, $t \in [0, 2\pi]$. If, in particular, we make $a_k = 0$, then all the terms B_{k+n}, $n = 0, 1, 2, ...,$ become mono-components.

Remark 2 The selected a_n's may not satisfy the relation (1), and thus may not define a complete basis $\{B_n\}_{n=1}^{\infty}$ in H^2. When $\{B_n\}_{n=1}^{\infty}$ is not complete, we have

$$\sum_{k=1}^{\infty} 1 - |a_k| < \infty.$$

In the case there is a Blaschke product, ϕ, using the a_n's as all its zeros including the multiplicities. We have the direct sum decomposition ([19])

$$H^2 = \overline{\mathrm{span}}\{B_1, ..., B_n, ...\} \oplus \phi H^2,$$

where

$$G \in \overline{\mathrm{span}}\{B_1, ..., B_n, ...\}.$$

Again, compulsory siftings can be inserted to produce complete bases.

Remark 3 In [20] we prove that the decay rate of the standard remainders of AFD is the negative square root of the partial sum order, and show that in the average sense Fourier series is the best. Some uniqueness and continuity results for the selections of a_k's under the Maximal Projection principle are proved in [19]. Some variations of AFD and the theory for a half of the complex plane are cited in [17] and [18].

III. The AFD Algorithm Based on Discrete Data

In this section we will cite the algorithm for real-valued $\tilde{G}$ given by a set of discrete data $\tilde{G}(e^{it_k}) = u_k, t_k \in [a, b], k = 1, ..., K$. To apply the core algorithm

to the projection $G^+ \in H^2$ we need to represent G^+ and its energy $\|G^+\|^2$ (12) using the data u_k of $\tilde{G}$.

We will deal with the case $[a,b] = [0, 2\pi]$. For general a and b one can proceed with the change of variable $s(t) = \frac{L}{2\pi}t + a,\ L = b - a$. For the partition $\{t_k\}$ of the interval $[0, 2\pi]$, $0 = t_0 < t_1 < \cdots < t_K = 2\pi$, define the step function

$$\tilde{G}_K(e^{it}) = \sum_{k=0}^{K-1} u_k \chi_{I_k}(t), \quad I_k = [t_k, t_{k+1}]. \tag{15}$$

Denote by G_K^+ the projection of $\tilde{G}_K$ onto H^2, we have, by (3),

$$G_K^+(z) = \frac{1}{2\pi} \sum_{k=0}^{K-1} u_k \int_{t_k}^{t_{k+1}} \frac{e^{it}}{e^{it} - z} dt. \tag{16}$$

For simplicity the formulas given below are for the equally spacing case, that is $\Delta_K = \frac{2\pi}{K} = t_{k+1} - t_k$. For the non-equally spacing cases the formulas are similar. The algorithm is divided into several steps.

(i). Computation of G_K^+ given by (16)

The H^2-function G_K^+ may be approximated by using a variety of parameterized numerical quadrature, such as the Simpson formula. We note that when using a numerical quadrature the approximating rational function of z may not be in the Hardy space. This problem shows up when an adaptively chosen a_k is close to the boundary: It gives rise to a large error. The problem may be treated through the constraint condition $|a_k| < \delta < 1, \delta \approx 1$, when choosing a_k's under the Maximal Projection Principle.

(ii). Computation of $\langle G_k, e_{\{a_1\}} \rangle$ and $\arg\max |\langle G_k, e_{\{a_1\}} \rangle|^2$

The computation of G_k is based on the recursive formula (7), where $G = G_1$ is replaced by the approximation G_K^+. The computation of the arguments giving rise to $\max |\langle G_k, e_{\{a_1\}} \rangle|^2$ is based on the formula $|\langle G_k, e_{\{a\}} \rangle|^2 = 2\pi(1 - |a|^2)|G_k(a)|^2$ that is a smooth function in two real variables represented by

$$2\pi(1 - x^2 - y^2)(u_k^2(x,y) + v_k^2(x,y)), \tag{17}$$

where u_k, v_k are two real-valued functions, $G_k = u_k + iv_k,\ z = x + iy$. The extreme problem is a usual one that can be treated by the existing numerical methods. In practice the computation may be simplified by using the Cauchy-Riemann equations $\partial_x u = \partial_y v, \partial_y u = -\partial_x v$. Primitively we do it through direct comparison of the function values given by (5) with F being replaced by G_k.

(iii). Energy in Terms of Discrete Data

To compute the energy of $G^+ = G_1$ we use the relation

$$\|\tilde{G}_K\|^2 = 2\|G_K^+\|^2 - 2\pi|c_0|^2$$

deduced from their respective energy representations of $\tilde{G}_K$ and G_K^+ in terms of their Fourier coefficients (The Plancherel Theorem). It follows

$$\begin{aligned} \|G_K^+\|^2 &= \frac{1}{2}\|\tilde{G}_K\|^2 + \pi|c_0|^2 \\ &\approx \pi\Delta_K \sum_{k=1}^{K} |u_k|^2 + \pi\Delta_K^2 |\sum_{k=1}^{K} u_k|^2. \end{aligned} \tag{18}$$

We note that the energy errors given by the left-hand-side of (12) is decreasing along with growth of N that should never become negative. A negative error, however, may occur. In that case the process should be set to cease.

(iv). AFD Series Expansions

The approximation of $\tilde{G}$ is given by (14) with $c_0 \approx \frac{1}{K}\sum_{k=1}^{K} u_k$. As a bi-product we obtain approximations to the circular Hilbert transform of $\tilde{G}$:

$$\tilde{H}(\tilde{G})(e^{it}) \approx 2\mathrm{Im} \sum_{k=1}^{N} \frac{(1 - |a_k|^2)G_k(a_k)}{1 - \overline{a}_k e^{it}} \prod_{l=1}^{k-1} \frac{e^{it} - a_l}{1 - \overline{a}_l e^{it}}. \tag{19}$$

Remark 4 In some applications, including system identification, the given complex-valued data u_k is known to be from the boundary value of an analytic function $G = G^+ \in H^2$, then G may be approximated by the same formula (16) in discrete data. The energy is given by

$$\|G\|^2 \approx 2\pi\Delta_K \sum_{k=1}^{K} |u_k|^2.$$

This is different from (18), for it is now in terms of the data u_k of the function G itself, not those of $\tilde{G}$.

IV. Algorithm Analysis: Robustness and Computational Complexity of AFD

A. Robustness

Assume that added to the real-valued boundary data $\tilde{G}(e^{it_k}) = u_k$ there is a noise data h_k from a noise $\tilde{h}, \tilde{h}(e^{it_k}) = h_k$, of small energy, viz. $\|\tilde{h}\|_2 \leq \epsilon$. Under the notation defined in (15) and (16),

$$\tilde{h}_K(e^{it}) = \sum_{k=0}^{K-1} h_k \chi_{I_k}(t), \quad I_k = [t_k, t_{k+1}],$$

and

$$h_K^+(z) = \frac{1}{2\pi} \sum_{k=0}^{K-1} h_k \int_{t_k}^{t_{k+1}} \frac{e^{it}}{e^{it} - z} dt.$$

We wish to estimate

$$\|G^+ - \hat{G}_{K,n}^+\|_{H^2},$$

where

$$\hat{G}_K = \tilde{G}_K + \tilde{h}_K, \quad \hat{G}_K^+ = G_K^+ + h_K^+,$$

and $\hat{G}^+_{K,n}$ is the n-th partial sum in the AFD of $\hat{G}^+_K$. We are to show

Theorem 4.1: Under the condition

$$\lim_{K\to\infty} \|\tilde{G} - \tilde{G}_K\|_{L^2} = 0,$$

the AFD approximation is robust, that is

$$\lim_{K\to\infty, n\to\infty, \epsilon\to 0} \|G^+ - \hat{G}^+_{K,n}\|_{H^2} = 0. \tag{20}$$

As a consequence, the decomposition for the real-valued $\tilde{G} = -c_0 + 2\mathrm{Re}G^+$ is also robust. The proof of the theorem is given in Appendix.

The robust approximation is not in the H^∞-sense as usual, but in the H^2- and L^2-sense. The L^2-convergence does not imply pointwise convergence in general. In many cases, however, besides L^2-convergence, we do have pointwise convergence as theoretically proved or experimentally observed. In the AFD case the pointwise convergence has not been proved. Some relevant information and observations are given in the following

Remark 5 In 1966 L. Carleson in his celebrating paper in Acta Math solved the long standing Lusin conjecture by disproving the conjecture. He showed that for any periodic function f of finite energy its Fourier partial sums pointwisely converge to a function that is almost everywhere identical with the function f ([4]). Subsequent to his proof there have been works devoting reproofs of the same result or extensions of the pointwise convergence to L^p with various contexts ([10], [8]). The methodologies used in Carleson's proof and the successive ones lay foundations of contemporary harmonic analysis. It is conjectured that the norm convergence should force the pointwise convergence to hold in a more general context, including Laplace-Fourier series on the higher dimensional spheres and AFD. Below we cite a result showing that in the pointwise convergence aspect TM expansions are of the same nature as Fourier expansions ([1]).

Let $a_1, ..., a_n, ...$ be any sequence in the unit disc $\mathbb{D}$, and G be analytic and bounded in $(1+\delta)\mathbb{D}, \delta > 0$. Denote by $\|\cdot\|_{\infty,C(A)}$ the supreme norm of bounded continuous functions in the set A. The n-th standard remainder R_n given by (9) satisfies

$$\|R_n\|_{\infty,C(\partial\mathbb{D})} \leq \frac{(1+\delta)\|G\|_{\infty,C((1+\delta)\mathbb{D})}}{\delta} \times \exp\left(-\frac{\delta}{2(1+\delta)}\sum_{k=1}^{n}(1-|a_k|)\right).$$

This estimate indicates that if there exist multiple choices of a_k under Maximal Projection Principle, then we should choose a_k as close as possible to the origin.

B. Computational Complexity

The statement of fastness of AFD does not mean that the algorithm is fast. It means that the convergence is fast. The computational complexity of AFD is expected to be at least as large as that of DFT, as there is an adaptivity process incorporated. Fourier series corresponds to the case where all the parameters are fixed ($a_k \equiv 0$) without a selection process.

In complexity calculations, we only worry about what happens as the data lengths increase, and take the dominant term. Based on this, AFD may be slightly modified to get an $O(K^2)$ computational complexity that is the same as for DFT. The modification is that in using the Maximal Projection Principle (8) instead of selecting a_k in $\mathbb{D}$ we select a_k in $[0, 1)$. This restricted adaptivity is adopted in our recent work in system identification dealing with rational functions of real coefficients ([12]). The calculation of the complexity is as follows.

(i) For each fixed $z \in [0, 1)$ to compute the value of $G^+_K(z)$ given by the formula (16), $O(K)$ multiplication and addition steps are necessary. Dividing $[0, 1)$ into K equal parts, there are K function values $A^2_{G_k}(z_k)$ to be worked out. Thus, $KO(K) = O(K^2)$ computational steps are necessary.

(ii) To get $\max\{A^2_{G_k}(a) : a \in [0, 1)\}$ from the K function values only $K-1$ computational steps are involved.

(iii) Both the computations for G_k based on the recursive formula (7) and the energy from the formula (18) require $O(K)$ computational steps.

Altogether the dominant term is $O(K^2)$.

Likewise, the full adaptivity costs the computational complexity $O(K^3)$. These, in the AFD case, however, are compensated by fast convergence of the partial sums: We show in our experiments (Examples A and C in §V) that to reach the same errors the AFD running times are considerably shorter than those of FFT that is of the complexity $O(K \log K)$.

V. Experimental results

A. AFD on Analytic Signal

The original signal is

$$G(z) = \frac{0.0247z^4 + 0.0355z^3}{(1 - 0.9048z)(1 - 0.3679z)} \in H^2.$$

In Fig.1 the four figures are for the partial sums of the AFD expansion with the orders $1, 2, 4, 6$. Fig.2 gives the Fourier expansions of, respectively, the orders $1, 2, 6, 12$. With the relative energy errors of AFD and Fourier for analytic functions defined respectively by

$$E^A(G; N) = \frac{\|G - \sum_{k=1}^{N}\langle G, B_k\rangle B_k\|^2}{\|G\|^2}$$

and

$$E^F(G;N)=\frac{\|G-\sum_{k=0}^{N-1}\langle G,f_k\rangle f_k\|^2}{\|G\|^2},\ f_k(e^{it})=e^{ikt},$$

we have

$$E^A(G;1)\approx 0.3799,\ E^A(G;2)\approx 0.1374,$$
$$E^A(G;4)\approx 0.0430,\ E^A(G;6)\approx 0.0089,$$

and

$$E^F(G;1)\approx 0.99996,\ E^F(G;2)\approx 0.99987,$$
$$E^F(G;6)\approx 0.72533,\ E^F(G;12)\approx 0.21778.$$

Fig.3 presents the relative energy errors. The 6th AFD partial sum has the relative error $E^A(G;6)\approx 0.0089379$ that costs 2.226507 seconds as computer running time; comparatively the 27th Fourier partial sum has the error $E^F(G,27)=0.0092819$ that costs 6.591188 seconds computer running time.

B. AFD on Real signal with White Noise

The original signal $\tilde{G}$ is the real-valued function on the unit circle

$$\tilde{G}(t)=1+10\cos t+10\sin t+\cos 2t+\sin 2t+0.5\cos 5t.$$

We add Gaussian White Noise to $\tilde{G}$ with $SNR=20$ and get a signal $\tilde{G}_0$. Apply AFD to $\tilde{G}_0$. As shown in Fig.4, AFD suppresses the noise efficiently.

C. AFD on Jump Signal

The original real-valued signal to be decomposed is

$$\tilde{G}(e^{it})=-\chi_{[0,\pi)}(t)+\chi_{[\pi,2\pi)}(t),\qquad 0\le t\le 2\pi.$$

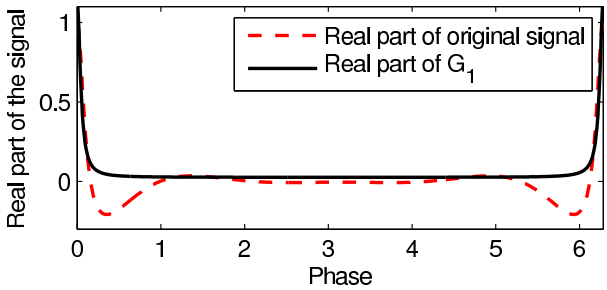
Define the relative energy errors of AFD and Fourier for real valued functions $\tilde{G}$, respectively, by

$$E^A(\tilde{G};N)=\frac{\|\tilde{G}-\left(-c_0+2\mathrm{Re}\sum_{k=1}^{N}\langle G,B_k\rangle B_k\right)\|^2}{\|\tilde{G}\|^2}$$

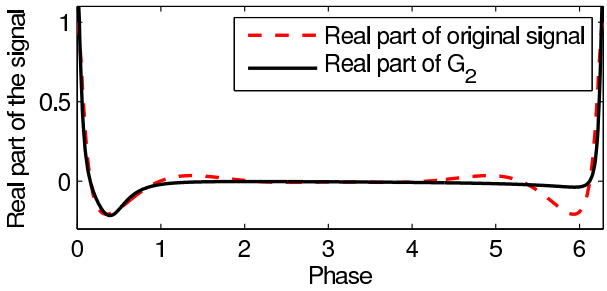
and

$$E^F(\tilde{G};N)=\frac{\|\tilde{G}-\sum_{k=-N+1}^{N-1}\left(\langle\tilde{G},B_k\rangle B_k\right)\|^2}{\|\tilde{G}\|^2}.$$

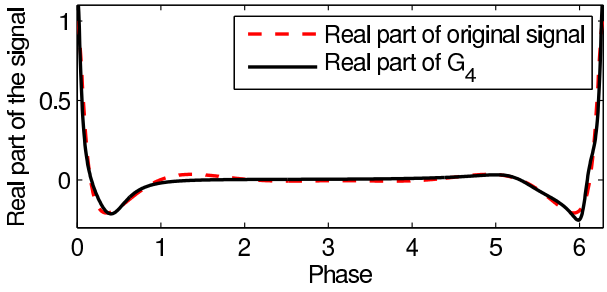
(i) Fig 5.1 is for the comparison between the original signal and the real part of the 16th AFD partial sum, $E^A(\tilde{G};16)=0.0035391$.
(ii) Fig 5.2 is for the comparison between the original signal and the real part of the 16th Fourier partial sum, where $E^F(\tilde{G};16)=0.025276$.
(iii) Fig 5.3 is the relative energy comparison between the AFD and the Fourier partial sums.
(iv) Fig 5.4.1 and Fig 5.4.2 are for the same function $\tilde{G}$, but with enlarged scale (from 0 to π). Fig 5.4.1 is for the 25th AFD partial sum with a relative energy error 0.0023339; while Fig 5.4.2 is for the 200th Fourier partial sum whose relative energy error is 0.002404. The computer running times are, respectively, 37.63599 seconds (AFD) and 152.974644 seconds (Fourier). Gibbs phenomenon is more noticeable in the Fourier expansion.



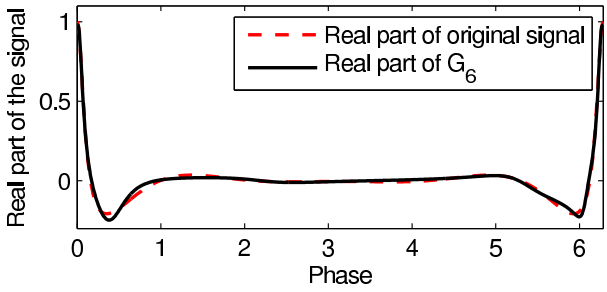
(a) The 1st AFD partial sum



(b) The 2nd AFD partial sum



(c) The 4th AFD partial sum

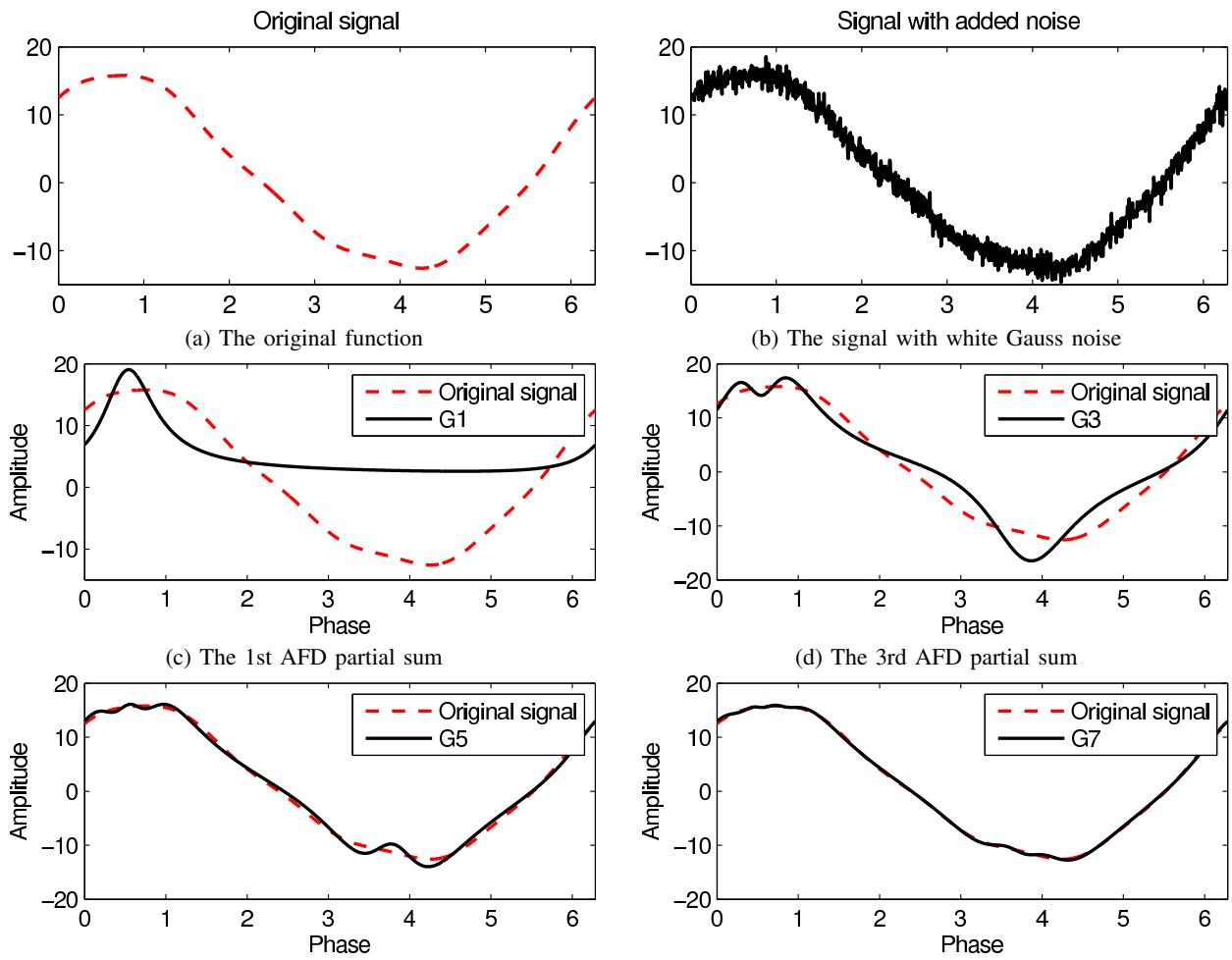


(d) The 6th AFD partial sum

Fig. 1. AFD decomposition of analytic function

VI. Conclusions

AFD is a realizable variation of greedy algorithm for functions in H^2 and L^2. The decomposition results in



(a) The original function

(b) The signal with white Gauss noise

(c) The 1st AFD partial sum

(d) The 3rd AFD partial sum

(e) The 5th AFD partial sum

(f) The 7th AFD partial sum

Fig. 4. AFD on real signal with white noise

series expansions of the so called modified Blaschke products with non-negative boundary phase derivatives, or, equivalently, meaningful instantaneous frequencies. The algorithm is practical with solid theoretical foundations. It is robust and with computational complexity comparable with that of DFT. For randomly chosen examples experiments show that in both the energy and pointwise sense AFD leads to more efficient decomposition than Fourier series. To get the same errors AFD uses considerably shorter computer running times than Fourier series. AFD has promising resolution results for filtering out noises, and exhibits less Gibbs phenomenon.

VII. Appendix

Proof of Theorem 4.1 By inserting the limit term we have

$$\begin{aligned}\|G^+ - \hat{G}^+_{K,n}\|_{H^2} &\leq \|G^+ - \hat{G}^+_K\|_{H^2} + \|\hat{G}^+_K - \hat{G}^+_{K,n}\|_{H^2} \\ &\leq \|G^+ - G^+_K\|_{H^2} + \|h^+_K\|_{H^2} \\ &\quad + \|\hat{G}^+_K - \hat{G}^+_{K,n}\|_{H^2}.\end{aligned}$$

Since the Hardy space norm is dominated by the L^2 space norm on the boundary, we have

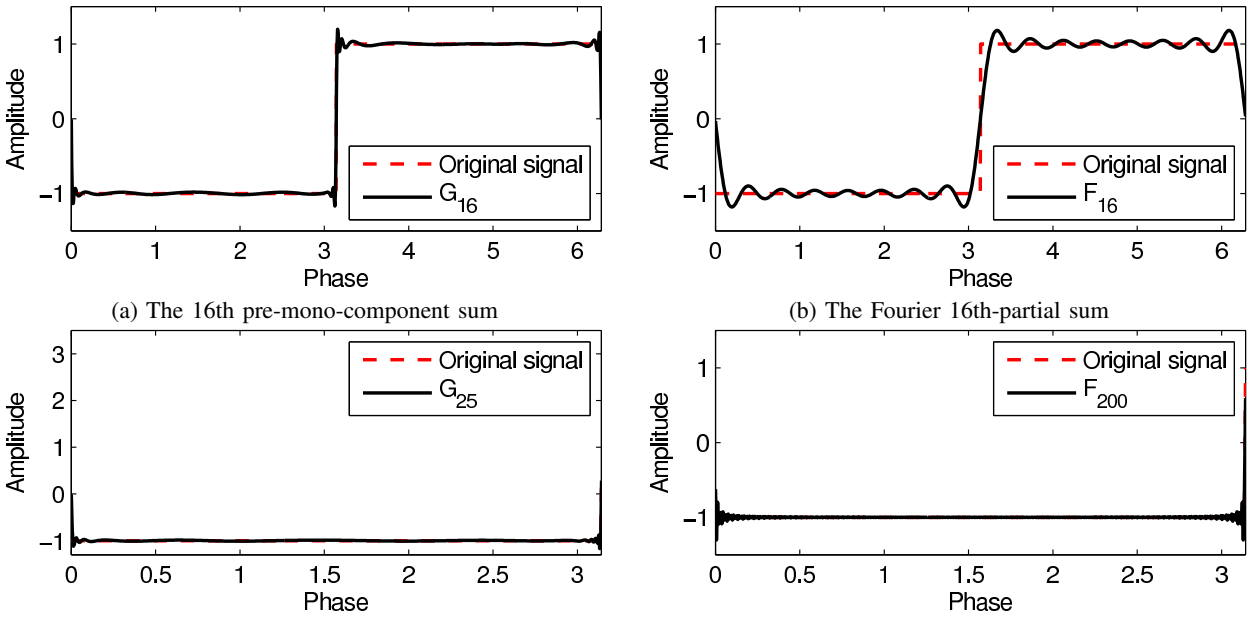
$$\|G^+ - G^+_K\|_{H^2} \leq C\|G^+ - G^+_K\|_{L^2}.$$

With the direct sum decomposition of the boundary L^2 into the Hardy spaces, we have

$$\|\tilde{G} - \tilde{G}_K\|^2_{L^2} = \|G^+ - G^+_K\|^2_{L^2} + \|G^- - G^-_K\|^2_{L^2},$$

and thus

$$\|G^+ - G^+_K\|_{H^2} \leq C\|\tilde{G} - \tilde{G}_K\|_{L^2}.$$

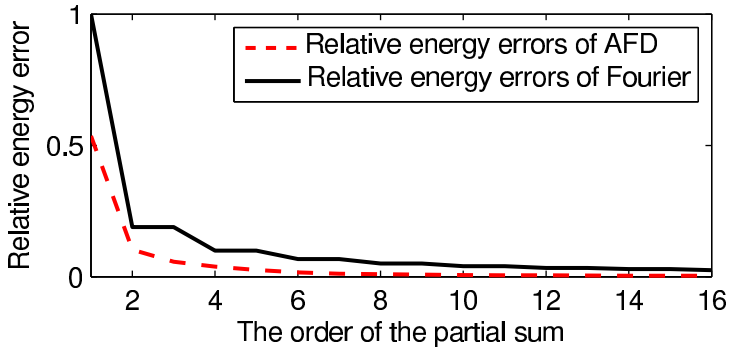


(a) The 16th pre-mono-component sum

(b) The Fourier 16th-partial sum

(c) The 25th pre-mono-component sum

(d) The Fourier 200th-partial sum



(e)Energy comparison

Fig. 5. AFD and Fourier decomposition of Jump function

Likewise,

$$\|h_K^+\|_{H^2} \leq C\|\tilde{h}\| \leq C\epsilon.$$

Consequently,

$$\|G^+ - \hat{G}_{K,n}^+\|_{H^2} \leq C\|\tilde{G} - \tilde{G}_K\|_{L^2} + C\epsilon + \|\hat{G}_K^+ - \hat{G}_{K,n}^+\|_{H^2}.$$

Since the AFD algorithm is for $\hat{G}_K^+$, the convergence of the algorithm implies

$$\lim_{n\to\infty} \|\hat{G}_K^+ - \hat{G}_{K,n}^+\|_{H^2} = 0.$$

Taking the limits $K, n \to \infty$ and $\epsilon \to 0$, we obtain (20).

REFERENCES

[1] H. Akcay and B. Niness, *Rational Basis Functions for Robust Identification from Frequency and Time-Domain Measurements,* Automatica, Vol. 34, No 9, pp. 1101-1117, 1998.

[2] H. Akcay and B. Niness, Orthonormal basis functions for modelling continuous-time systems, Signal Processing, 77 (1999) 261-274.

[3] Butzer and R. Nessel. *Fourier Analysis and Approximation,* Volume 1: One-dimensional Theory, volume 40. Academic Press, New York, 1980.

[4] L. Carleson, On convergence and grows of of partial sums of Fourier series, Acta Math, 1966, 116;135-137.

[5] L. Cohen, *Time-Frequency Analysis: Theory and Application* , Prentice Hall, Englewood Cliffs, NJ, 1995.

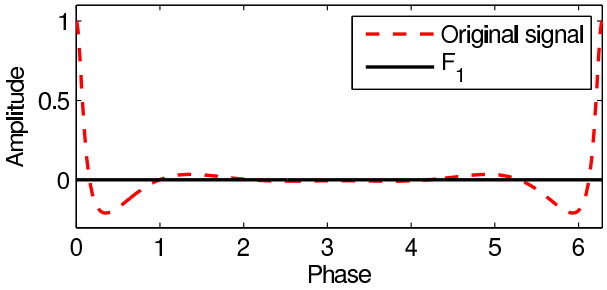
[6] P. Dang, T. Qian and Z. You, *Hardy-Sobolev Spaces Decomposition In Signal Analysis,* accepted to appear in Journal of Fourier Analysis and Applications.

[7] R.A. Devore, V.N.Temlyakov, *Some Remarks on greedy algorithm*, Advances in Computational Mathematics 5(1996) 173-187.

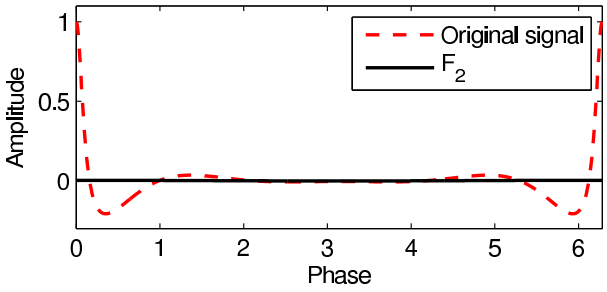
[8] C. Fefferman, *Pointwise Convergence of Fourier Series,* The Annals of Mathematics, Second Series, Vol. 98, No. 3 (Nov., 1973), pp. 551-571.

[9] J. B. Garnett, *Bounded Analytic Functions*, Academic Press, 1987.

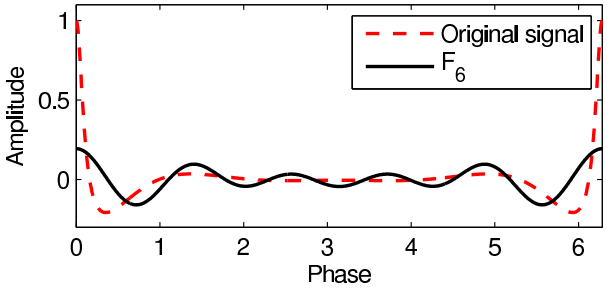
[10] R.A. Hunt, *On the convergence of Fourier series, Orthogonal Expansions and Their continuous Analogues,* Proc. Conf. Ed-



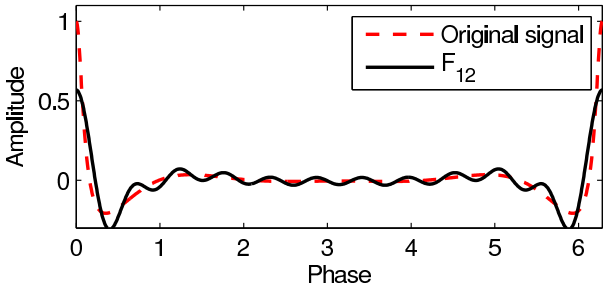
(a) The 1st Fourier partial sum



(b) The 2nd Fourier partial sum



(c) The 6th Fourier partial sum



(d) The 12nd Fourier partial sum

Fig. 2. Fourier decomposition of analytic function

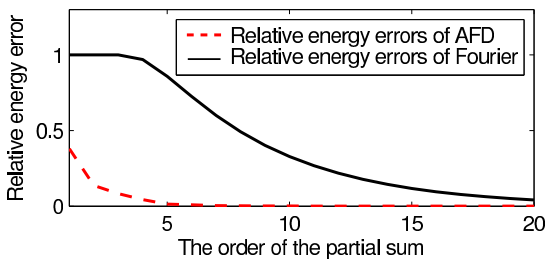


Fig. 3. Energy comparison

wardsville, I11.(1967), 235-255; Southern Illinois Univ. Press, Carbondale, I11.(1968).

[11] S.Mallat and Z.Zhang, *Matching pursuits with time-frequency dictionaries,* IEEE Transactions on Signal Processing,1993,**41**:3397-3415.

[12] W. Mi and T. Qian, *Frequency domain identification with adaptive rational orthogonal system,* accepted to appear in Proceedings of 2010 International Conference on System Science and Engineering, Taiwan.

[13] B. Picinbono, *On instantaneous amplitude and phase of signals*, IEEE Transactions on Signal Processing, vol, **45**, No. 3, March, 1997, 552-560.

[14] T. Qian, Q-H. Chen and L-Q. Li, *Analytic unit quadrature signals with non-linear phase*, Physica D: Nonlinear Phenomena, **303** (2005) 80-87.

[15] T. Qian, *Mono-components for decomposition of signals,* Math. Meth. Appl. Sci. 2006;**29**: 1187-1198.

[16] T. Qian, *Boundary derivative of the phases of inner and outer functions and applications,* Math. Meth. Appl. Sci. 2009;**32**: 253-263.

[17] T. Qian, *Mono-component Decomposition of Functions: An Advance of Fourier Theory,* Math. Meth. Appl. Sci. **2010**, 33 880-891.

[18] T. Qian, L.H. Tan and Y.B. Wang, *Adaptive decomposition by weighted inner functions: A generalization of Fourier series,* Journal of Fourier Analysis and Applications, DOI: 10.1007/s00041-010-91541.

[19] T. Qian and Y.B. Wang, *Adaptive Fourier Series - A Variation of Greedy Algorithm,* to appear in Advances in Computational Mathematics (AiCM), 2010. DOI: 10.1007/s10444-010-9153-4

[20] T. Qian and Y.B. Wang, *Convergence rate of adaptive Fourier decomposition,* preprint.

[21] T. Qian, R. Wang, Y.S. Xu, H.Z. Zhang, *Orthonormal Bases with Nonlinear Phases*, Advances in Computational Mathematics, 2010, Vol. 33, No. 1, 75-95.

[22] T. Qian, Y.S. Xu, D.Y. Yan, L.X. Yan and B. Yu, *Fourier Spectrum Characterization of Hardy Spaces and Applications,* Proc. Amer. Math. Soc. **137** (2009), 971-980.

[23] L.H. Tan, L.X. Shen and L.H. Yang, *Rational orthogonal bases satisfying the Bedrosian identity,* Advamces in Computational Mathematics, DOI: 10.1007/s10444-009-9133-8, 2009.

[24] V.N.Temlyakov, *The best m-term approximation and greedy algorithm*, Advances in Computational Mathematics, 1998, Vol 8, 249-265.

[25] L.H. Yang and H.Z. Zhang, *The Bedrosian identity for H^p functions,* J. Math. Anal. Appl. **345** (2008) 975 84.

[26] B. Yu and H.Z. Zhang, *The Bedrosian identity and homogeneous semi-convolution equations,* Journal of Integral Equations and Applications, Volume **20**, Number 4, Winter 2008, 527-568.

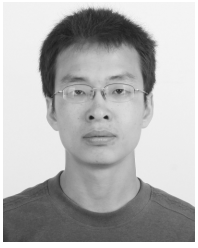
Tao Qian received the M.Sc. and Ph.D. degrees, both in harmonic analysis, from Peking University, Beijing, China, in 1981 and 1984, respectively.

From 1984 to 1986, he worked in Institute of Systems Science, the Chinese Academy of Sciences. Then he worked as Research Associate and Research Fellow in Australia till 1992 (Macquarie University, Flinders University of South Australia). He worked as Lecturer (Level B and Level C, English System) at New England University, Australia, from 1992 to 2000. He started working at University of Macau, Macao, China, from 2000 as Associate Professor. He got Full Professorship in 2003, and has been Head of Department of Mathematics from 2005 to 2011. His research interests include harmonic analysis in Euclidean spaces, complex and Clifford analysis and signal analysis. Up to now he has published more than 100 research papers of which 85 are journal papers.

Qian received various honors and awards for his research excellence, including Prize of Scientific Progress, China, 1984, 1985; and research awards of the Australian and Macao universities where he worked as regular teacher.

Liming Zhang (IEEE member) received the B.S. degree from Nankai University, China, in 1987. She received the M.S. degree from the East China Institute of Technology, China, in 1990, and the PhD in computer science (image processing) from the University of New England, Australia, 2002. Her contribution in image processing is the theory and algorithm in rotational wavelet bank for edge detection.

She worked in Tianjin Institute of Technology, China, from 1990 to 1998 as lecturer and associate professor. She participated in a number of national projects, including 863 project, major in Robert version. She joined the Faculty of Education, University of Macau in 2001 as lecturer and assistant professor till 2010 and worked in the direction of IT in education. From 2007 to 2010, she acted as the director of the Information and Communication Education Technology Center. She organized several international conferences and workshops in IT in education area and published a series of papers, book chapters, and a co-edited book. She started to work in the Faculty of Science and Technology, University of Macau as an assistant professor from 2010. Her recent research interests including signal processing and image processing.



Zhixiong Li received the B.S in Mathematics from the Zhejiang University, Zhejiang, China in 2005. And is currently pursuing a master's degree from the University of Macau, Macau. His main interests are signal processing and GNU/Linux.