Frequency Domain Identification: An Algorithm Based On Adaptive Rational Orthogonal System ⋆

Wen Mi [a], Tao Qian [b]

[a] *Av. Padre Toms Pereira Taipa, Macau, China*

[b] *Av. Padre Toms Pereira Taipa, Macau, China*

Abstract

This paper presents a new adaptive algorithm for frequency-domain identification. The algorithm is related to the rational orthogonal system (Takenaka-Malmquist system). This work is based on an adaptive decomposition algorithm previously proposed for decomposing the Hardy space functions, in which a greedy sequence is obtained according to the maximal selection criterion. We modify the algorithm through necessary changes for system identification.

Key words: System identification, frequency domain identification, rational approximation, adaptive algorithm, greedy algorithm.

1 Introduction

System identification concerns the modeling of physical systems that can be described by input-output measurements in the time domain or frequency domain. This paper is concerned with the problem of approximating the dynamics of single input, single output (SISO) discrete linear time-invariant (LTI) systems that are causal and stable. For the considered discrete LTI systems, let $\{h_k\}$ be the impulse response of the system. Then

$$G(z) = \sum_{l=1}^{+\infty} h_l z^{-l} \tag{1}$$

is the transfer function of the system.

A number of methods have been developed to identify a system. One widely used method is to construct a model structure with a given order, and then to estimate the parameters by the measured data. The most classical and commonly used models include the FIR model, ARX model and ARMAX model. Researchers have been using the rational orthogonal systems by making the model structure priori-linear in parameters, viz, the transfer function $G(z)$ is approximated by

$$\widetilde{G}(z) = \sum_{l=1}^{n} \theta_l \mathcal{B}_l(z), \tag{2}$$

where $\{\mathcal{B}_l(z)\}$ is a rational orthogonal system, $\{\theta_l\}$ is the n-tuple of parameters to be determined, n is the order of the model structure. Denote $\theta = [\theta_1 \ \theta_2 \ \ldots \ \theta_n]$ as a parameter vector. Let $\{E_k\}_{k=1}^N$ be the measurements in the frequency domain with

$$E_k = G(e^{j\omega_k}) + v_k, \tag{3}$$

where v_k is the noise, then associated with (2), the minimizing parameters θ^* can be determined by a least-squares criterion,

$$\theta^* = \arg\min_{\theta} \frac{1}{N} \sum_{k=1}^{N} |\widetilde{G}(e^{j\omega_k}) - E_k|^2, \tag{4}$$

which can be done almost instantaneously.

⋆ This paper was not presented at any IFAC meeting. Corresponding author T. Qian. Tel.(+853) 8397 4954. Fax: (+853) 2883 8314. The research was partially supported by Macao Sci. and Tech. Develop. Fund 014/2008/A1 and University of Macau Research Grant CG071/08-09S/QT/FST.

Email addresses: `ya87405@umac.mo` (Wen Mi), `fsttq@umac.mo` (Tao Qian).

The general setting of the rational orthogonal basis is

$$\mathcal{B}_k(z) = \frac{\sqrt{1-|a_k|^2}}{z-a_k}\prod_{l=1}^{k-1}\frac{1-\overline{a}_l z}{z-a_l},\ |a_k|<1. \qquad (5)$$

The following are the particular cases that (2) gives rise to with (5).

- all $a_k = 0$: the classical FIR model;
- all $a_k = a$, a being real-valued: the Laguerre models [21,22,39,42];
- all $a_k = a$, a being complex-valued: the Kautz models [40–42].

The general setting of the rational orthogonal system (5) was first studied in the 1920s by Takenaka (1925) and Malmquist (1926) [43], and thus was named as Takenaka-Malmquist (TM) system. It has been used in system identification and analyzed in detail since 1980s by Ninness and other researchers with an ample amount of publications, including [2–4,12,23–25,38]. The study [26] by T. Oliveira e Silva shows that small perturbations in locating the true poles do not induce much error for the approach. Nevertheless, to estimate the true poles of the original system is by no means easy. Our approach is based on a different philosophy. We do not care about where the poles of the true system are for the LTI systems. We, using a maximal selection criterion in terms of energy, find an approximation to the original system in the energy sense through adaptively selected poles defining the rational orthogonal system $\{\mathcal{B}_k\}$.

Here and after, let $\mathbb{D}$ denote the unit disc. With the necessary assumptions on the system, the transfer function (1) belongs to the space of functions, holomorphic outside the unit disk. Under the transformation: $z \longrightarrow 1/z$, the transfer function turns out to be analytic in $\mathbb{D}$. For the uniformity and clarity, the functions considered in this paper are assumed to be analytic inside the unit disc, and they belong to $H_2(\mathbb{D})$ with real-valued impulse response.

For the rational orthogonal system, a well-known and crucial result is as follows.

Theorem 1 *[3]. Consider the set of functions $\{\mathcal{B}_k(z)\}$ defined by*

$$\mathcal{B}_k(z) = \mathcal{B}_{\{\zeta_1,.\zeta_k\}}(z) \triangleq \frac{\sqrt{1-|\zeta_k|^2}}{1-\overline{\zeta}_k z}\prod_{l=1}^{k-1}\frac{z-\zeta_l}{1-\overline{\zeta}_l z}, \qquad (6)$$

where $\zeta_k \in \mathbb{D}$ and $k = 1, \ldots$. Then the set $X = span\{\mathcal{B}_k(z)\}_{k\geq 1}$ is complete in $A(\mathbb{D})$ or $H_p(\mathbb{D})$ for $1 \leq p < \infty$ if and only if

$$\sum_{l=1}^{\infty}(1-|\zeta_l|) = \infty, \qquad (7)$$

where $A(\mathbb{D})$ is the disk algebra $\{f : f$ is analytic in the unit disc $\mathbb{D}$ and continuous on $\overline{\mathbb{D}}$ $\}$, $H_p(\mathbb{D})$ is the Hardy p space of functions $f(z)$ analytic in $\mathbb{D}$.

The condition of completeness (7) is the so-called Szász condition, which has a long history [34,13].

This paper is based on the theory presented in [30,29]. We will concentrate on the frequency-domain system identification. The proposed algorithm is based on the rational orthogonal system (6) through finding $\{\zeta_k\}$ under a maximal selection criterion. We incorporate a technical treatment that makes the approximating rational functions to have real-valued coefficients which is necessary for systems with real-valued impulse responses.

This paper is arranged as follows. In section 2 we give the problem setting. A brief introduction of the adaptive decomposition algorithm for $f(z)$ in the Hardy space $H_2(\mathbb{D})$ is given in section 3. After that, our main result of this paper is given in section 4. In section 5 an example is presented. Conclusions are drawn in the last section.

2 Problem setting

Frequency-domain identification is based on a set of frequency-domain measurements. This set of data is usually obtained by the sinusoid response or, more accurately, by the correlation method [17,13].

Without loss of generality, it is also assumed that a set of frequency-domain measurements $\{E_k\}_{k=1}^N$ is available and these frequency-domain measurements are obtained from a single input, single output (SISO) discrete linear time-invariant (LTI) system $f(z)$ in $H_2(\mathbb{D})$ with real-valued impulse response. We further assume that $f(z)$ can be continuously extended to a region containing the closed unit disc. Under these assumptions, if $f(z)$ is a rational function, its coefficients are all real-valued and the poles are outside the closed unit disc.

It is then assumed that the structure of measurements $\{E_k\}_{k=1}^N$ is set up to be

$$E_k = f(e^{-j\omega_k}) + v_k \ \ (k = 1, 2, ..., N),$$

where $\omega_k = \frac{2\pi(k-1)}{N}$, N is even and $\{v_k\}$ is a bounded sequence satisfying $|v_k| \leq \epsilon$, $\epsilon > 0$, or a zero-mean stochastic process with a bounded covariance function, and $f(z)$ is the true function to be approximated. Note that equal spacing is, in fact, unnecessary in the algorithm. We just need the measurements for $\omega \in [0, \pi)$, the rest in the interval $(\pi, 2\pi)$ will be obtained by using the conjugate symmetry of the frequency response data.

Let $X_n = span\{\mathcal{B}_1, \mathcal{B}_2, ..., \mathcal{B}_n\}$, where the orthonormal system $\{\mathcal{B}_k\}_{k=1}^n$ is defined by (6) and $\zeta_k \in \mathbb{D}$. The identification problem that we consider now can be stated as follows.

Frequency-domain identification problem: Given a set of frequency-domain measurements $\{E_k\}_{k=1}^N$ for $f \in H_2(\mathbb{D})$. Find a projection $f_n(z) \in X_n$ through finding $\{\zeta_k\} \in \mathbb{D}$, called a *greedy sequence*, such that for each k,

$$\zeta_k = \arg\max\{|\langle f, B_{\{\zeta_1,.,\zeta_{k-1},\zeta\}}\rangle|^2, \zeta \in \mathbb{D}\}. \quad (8)$$

This is a consecutive energy approximation, meaning that

$$f_{n+1} = f_n + \langle f, \mathcal{B}_{n+1}\rangle \mathcal{B}_{n+1} \quad (9)$$

and, in the H_2-norm convergence of f_n to f,

$$f = \lim_{n\to\infty} f_n. \quad (10)$$

Such function decomposition is called adaptive Fourier decomposition (AFD) in which it is expected that a sequence $\{\zeta_k\}_{k=1}^n$ is found to give rise to a TM sequence $\{\mathcal{B}_k(z)\}$ that offers efficient approximation to the given function.

3 AFD for $H_2(\mathbb{D})$ functions

In this section, we provide a brief introduction to the adaptive Fourier decomposition algorithm for $H_2(\mathbb{D})$ functions [30,29].

The inner product in $H_2(\mathbb{D})$ is defined by

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

Denote $e_{\{\zeta\}}(z) = \frac{\sqrt{1-|\zeta|^2}}{1-\bar{\zeta}z}$, called *the evaluator at* ζ, and $\mathcal{D} = \{e_\zeta, \zeta \in \mathbb{D}\}$ is the dictionary. By invoking the Cauchy integral formula, it gives rise to the evaluating functional

$$\begin{aligned}\left\langle f, e_{\{\zeta\}}\right\rangle &= \frac{\sqrt{1-|\zeta|^2}}{2\pi}\int_0^{2\pi} f(e^{j\omega})\overline{\frac{1}{1-\bar{\zeta}e^{j\omega}}}d\omega \\ &= \sqrt{1-|\zeta|^2} f(\zeta). \end{aligned} \quad (11)$$

Let $f_1 = g_1 = f$, we first have

$$\begin{aligned} f(z) &= \left(f_1(z) - \left\langle g_1, e_{\{\zeta_1\}}\right\rangle e_{\{\zeta_1\}}(z)\right) + \left\langle g_1, e_{\{\zeta_1\}}\right\rangle e_{\{\zeta_1\}}(z) \\ &= R_1(z) + \left\langle g_1, e_{\{\zeta_1\}}\right\rangle e_{\{\zeta_1\}}(z) \\ &= g_2(z)\frac{z-\zeta_1}{1-\bar{\zeta}_1 z} + \left\langle g_1, e_{\{\zeta_1\}}\right\rangle e_{\{\zeta_1\}}(z), \end{aligned}$$

where $f_2(z)$ has a zero at $z = \zeta_1$, and hence $g_2(z)$ is in $H_2(\mathbb{D})$. It is proved in [30], or alternatively [29], that there exists ζ_1 as an interior point of the unit disc $\mathbb{D}$ such that

$$\zeta_1 = \arg\max\{\left|\left\langle g_1, e_{\{\zeta\}}\right\rangle\right|^2 : \zeta \in \mathbb{D}\}. \quad (12)$$

The selection criterion for ζ_1 is named as *Maximal Selection Criterion*. The process from g_1 to g_2 through the Maximal Selection Criterion is called a "maximal sifting process". Applying the maximal sifting process to g_2, we obtain g_3, and so on. After the kth step we obtain

$$f(z) = R_k(z) + f_k(z), \quad (13)$$

where the remainder is

$$R_k(z) = g_{k+1}(z)\prod_{l=1}^{k}\frac{z-\zeta_l}{1-\bar{\zeta}_l z}, \quad (14)$$

and the kth approximating partial sum

$$f_k(z) = \sum_{l=1}^{k}\left\langle g_l, e_{\{\zeta_l\}}\right\rangle \mathcal{B}_{\{\zeta_1,\zeta_2,...,\zeta_l\}}(z). \quad (15)$$

$f_k(z)$ is a rational function $\frac{P_k(z)}{Q_k(z)}$, in which the degree of the polynomial $P_k(z)$ does not exceed $k-1$ and the degree of polynomial $Q_k(z)$ does not exceed k. For $l = 2, ..., k+1$, the recursive formula for g_l is

$$g_l(z) = \left(g_{l-1}(z) - (1-|\zeta_{l-1}|^2)\frac{g_{l-1}(\zeta_{l-1})}{1-\bar{\zeta}_{l-1}z}\right)\frac{1-\bar{\zeta}_{l-1}z}{z-\zeta_{l-1}}, \quad (16)$$

where ζ_l is selected according to the Maximal Selection Criterion

$$\zeta_l = \arg\max\{\left|\left\langle g_l, e_{\{\zeta\}}\right\rangle\right|^2 : \zeta \in \mathbb{D}\}. \quad (17)$$

The following theorem deals with the convergence.

Theorem 2 *[30,29] For a given function $f \in H_2(\mathbb{D})$, under the Maximal Selection Criterion, first applied to $g_1 = f$ and then consecutively to $g_k(z)$, there holds*

$$f(z) = \sum_{k=1}^{\infty}\left\langle g_k, e_{\{\zeta_k\}}\right\rangle \mathcal{B}_k(z) = \sum_{k=1}^{\infty}\left\langle f, \mathcal{B}_k\right\rangle \mathcal{B}_k(z), \quad (18)$$

where the convergence is in $H_2(\mathbb{D})$.

For a proof, we refer to [30] or [29].

The greedy sequence $\{\zeta_k\}$ selected under the Maximal Selection Criterion may not satisfy the condition $\sum(1-|\zeta_k|)=\infty$. In the case $\sum_{k=1}^{\infty}(1-|\zeta_k|)<\infty$, [30] shows that $H_2(\mathbb{D})$ has the following decomposition

$$H_2(\mathbb{D})=\overline{span}\{\mathcal{B}_1,...\mathcal{B}_k,...\}\bigoplus \phi H_2(\mathbb{D}), \tag{19}$$

where ϕ is the infinite Blaschke product which has and only has $\{\zeta_k\}_{k=1}^{\infty}$ as its zeros (including the multiplicities).

4 Adaptive Approximation

In this section the AFD algorithm is modified to make the approximating rational functions to have real-valued coefficients. When adopting the adaptive algorithm, we first work out a function $\widetilde{f}(z)$ approximating the true function $f(z)$ dependent on the given measurements. Then we apply the AFD algorithm to the approximating function $\widetilde{f}(z)$. We call this the two-steps algorithm.

4.1 Two-steps algorithm

In this algorithm, the first step is to construct a function $\widetilde{f}(z)\in H_2(\mathbb{D})$ as the first approximation based on the measured data $\{E_k\}_{k=1}^{N}$. In this paper, $\widetilde{f}(z)$ is constructed by the Cauchy integral,

$$\widetilde{f}(z)=\frac{1}{2\pi j}\int_0^{2\pi}\frac{\sum_k f(e^{j\omega_k})\chi_{k(\cdot,\cdot)}(\omega)}{e^{j\omega}-z}de^{j\omega}, \tag{20}$$

where $\chi(\cdot)$ is the indicator function, and, in the L_2 sense,

$$\begin{aligned}&\sum_k f(e^{j\omega_k})\chi_{k(\cdot,\cdot)}(\omega)\\ &=\sum_{k=1}^{N/2} f(e^{j\omega_k})\chi_{(\omega_k,\omega_{k+1})}(\omega)+\sum_{k=\frac{N}{2}+2}^{N+1} f(e^{j\omega_k})\chi_{(\omega_k,\omega_{k-1})}(\omega)\\ &\xrightarrow[N\to\infty]{} f(e^{j\omega}).\end{aligned}$$

Using data $\{E_k\}$ instead of the function values, then (20) gives

$$\begin{aligned}\widetilde{f}(z)=&\frac{1}{2\pi j}\sum_{k=1}^{N/2}\overline{E}_k\ln(\frac{e^{j\omega_{k+1}}-z}{e^{j\omega_k}-z})\\ &-\frac{1}{2\pi j}\sum_{k=1}^{N/2}E_k\ln(\frac{e^{-j\omega_{k+1}}-z}{e^{-j\omega_k}-z}).\end{aligned}$$

Furthermore, it is easy to show that $\widetilde{f}(\overline{z})=\overline{\widetilde{f}(z)}$.

Remark 3 *To get an approximating function $\widetilde{f}(z)$, the use of the Cauchy integral in step 1 is not essential. Other approximation methods can also work for this purpose, such as by polynomials according to [11,27].*

With an adequate sample of frequency-domain measurements, the second step is to find approximations to the function $\widetilde{f}(z)$ by using the AFD algorithm introduced in section 3. We denote by $f_n(z)$ the obtained nth approximating partial sum.

The computation of g_k in the adaptive algorithm is based on the recursive formula (16), where $f(z)=g_1(z)$ is replaced by the approximation $\widetilde{f}(z)$. The computation of the arguments giving rise to $\max|\langle g_k, e_{\{\zeta\}}\rangle|^2$ is based on the formula $|\langle g_k, e_{\{\zeta\}}\rangle|^2=(1-|\zeta|^2)|g_k(\zeta)|^2$ that is a smooth function in two real-valued variables represented by

$$(1-x^2-y^2)(u_k^2(x,y)+v_k^2(x,y))$$

where u_k, v_k are two real-valued functions and $g_k=u_k+jv_k$, $\zeta=x+jy$. The extreme problem can be treated by the existing numerical methods.

The greedy sequence $\{\zeta_k\}$ obtained by using the AFD algorithm may contain complex numbers. Similar to [24] we use the conjugate points in the greedy sequence $\zeta_{k+1}=\overline{\zeta}_k$ in case ζ_k is complex-valued. We can show that the partial sums with $\{\zeta_k\}$ being real-valued or coming by conjugate pairs have real-valued coefficients as follows. We call the sequence under the construction *modified greedy sequence.*

Lemma 4 *For function $\widetilde{f}(z)\in H_2(\mathbb{D})$ with the property $\widetilde{f}(\overline{z})=\overline{\widetilde{f}(z)}$. In the process of AFD algorithm, with $\{\zeta_k\}_{k=1}^{n}$ appearing real-valued or conjugate pairs, the approximating nth partial sum $f_n(z)$ is a rational function in $H_2(\mathbb{D})$ with real-valued coefficients.*

Proof. First, we show that if each ζ_k is either real-valued or comes by a conjugate pair, the coefficients of the recursive formula $g_{n+1}(z)$ by (16) is real-valued. We can see that if ζ_k is real-valued and $g_k(z)$ has real-valued coefficients, then $g_{k+1}(z)$ has real-valued coefficients, too. So we only need to prove that $g_3(z)$ has real-valued coefficients if ζ_1 and ζ_2 are conjugated.

By (16), denote $c=1-|\zeta_1|^2$, substituting $g_2(z)$ with its expression in terms of $g_1(z)$, we have

$$\begin{aligned}g_3(z)&=\frac{g_2(z)(1-\overline{\zeta}_2z)-(1-|\zeta_2|^2)g_2(\zeta_2)}{z-\zeta_2}\\ &=F_1(z)+F_2(z),\end{aligned}$$

where

$$F_1(z) = \frac{g_1(z)(1-\overline{\zeta}_1 z)(1-\zeta_1 z)}{(z-\overline{\zeta}_1)(z-\zeta_1)},$$

$$F_2(z) = -\frac{(1-|\zeta_1|^2)g_1(\zeta_1)(1-\zeta_1 z)}{(z-\overline{\zeta}_1)(z-\zeta_1)} - \frac{(1-|\zeta_1|^2)g_1(\overline{\zeta}_1)(1-\overline{\zeta}_1^2) - (1-|\zeta_1|^2)^2 g_1(\zeta_1)}{(z-\overline{\zeta}_1)(\overline{\zeta}_1-\zeta_1)}.$$

$F_1(z)$ satisfies $F_1(\overline{z}) = \overline{F_1(z)}$, and by computation, we have

$$\begin{aligned} F_2(z) &= c\frac{[g_1(\overline{\zeta}_1) - g_1(\zeta_1) + g_1(\zeta_1)\zeta_1^2 - g_1(\overline{\zeta}_1)\overline{\zeta}_1^2]z}{(z-\overline{\zeta}_1)(z-\zeta_1)(\zeta_1-\overline{\zeta}_1)} + \\ &\quad c\frac{g_1(\zeta_1)\overline{\zeta}_1 - g_1(\overline{\zeta}_1)\zeta_1 + |\zeta_1|^2[g_1(\overline{\zeta}_1)\overline{\zeta}_1 - g_1(\zeta_1)\zeta_1]}{(z-\overline{\zeta}_1)(z-\zeta_1)(\zeta_1-\overline{\zeta}_1)} \\ &= c\frac{(-\Im\{g_1(\zeta_1)\} + \Im\{g_1(\zeta_1)\zeta_1^2\})z}{(z-\overline{\zeta}_1)(z-\zeta_1)\Im\{\zeta_1\}} \\ &\quad + c\frac{\Im\{g_1(\zeta_1)\overline{\zeta}_1\} - |\zeta_1|^2\Im\{g_1(\zeta_1)\zeta_1\}}{(z-\overline{\zeta}_1)(z-\zeta_1)\Im\{\zeta_1\}}, \end{aligned}$$

where $\Im\{\cdot\}$ denotes the imaginary part of a complex number. Thus $F_2(z)$ has real-valued coefficients and hence $g_3(z)$ has real-valued coefficients.

Now we use mathematical induction. First, without loss of generality, we assume ζ_1 is complex-valued, then we let $\zeta_2 = \overline{\zeta}_1$, so for the second partial sum, we have

$$\begin{aligned} f_2(z) &= \langle g_1, e_{\zeta_1}\rangle \mathcal{B}_1(z) + \langle g_2, e_{\zeta_2}\rangle \mathcal{B}_2(z) \\ &= c\frac{[g_2(\overline{\zeta}_1) - g_1(\zeta_1)\zeta_1]z + g_1(\zeta_1) - g_2(\overline{\zeta}_1)\zeta_1}{(1-\zeta_1 z)(1-\overline{\zeta}_1 z)}, \end{aligned}$$

according to the recursive formula (16), then the coefficient of z in the numerator in the above

$$\begin{aligned} & g_2(\overline{\zeta}_1) - g_1(\zeta_1)\zeta_1 \\ &= \frac{g_1(\overline{\zeta}_1) - g_1(\zeta_1) + g_1(\zeta_1)\zeta_1^2 - g_1(\overline{\zeta}_1)\overline{\zeta}_1^2}{\overline{\zeta}_1 - \zeta_1} \\ &= \frac{\Im\{g_1(\zeta_1)\} - \Im\{g_1(\zeta_1)\zeta_1^2\}}{\Im\{\zeta_1\}} \end{aligned}$$

is real-valued. And the constant term in the numerator

$$\begin{aligned} & g_1(\zeta_1) - g_2(\overline{\zeta}_1)\zeta_1 \\ &= \frac{g_1(\zeta_1)\overline{\zeta}_1 - g_1(\overline{\zeta}_1)\zeta_1 + |\zeta_1|^2[g_1(\overline{\zeta}_1)\overline{\zeta}_1 - g_1(\zeta_1)\zeta_1]}{\overline{\zeta}_1 - \zeta_1} \\ &= \frac{\Im\{g_1(\zeta_1)\overline{\zeta}_1\} - |\zeta_1|^2\Im\{g_1(\zeta_1)\zeta_1\}}{-\Im\{\zeta_1\}} \end{aligned}$$

is real-valued too.

Second, if to the kth time, the conclusion is correct, that is, the kth partial sum $f_k(z)$ has real-valued coefficients under the construction: $\{\zeta_l\}_{l=1}^k$ are either real-valued or come in conjugate pairs. Then, if ζ_{k+1} is real-valued, $f_{k+1}(z)$ must have real-valued coefficients; if ζ_{k+1} is complex-valued, let $\zeta_{k+2} = \overline{\zeta}_{k+1}$, the following is to show the $(k+2)$th partial sum $f_{k+2}(z)$,

$$f_{k+2}(z) = \sum_{l=1}^{k+2} \langle g_l, e_l\rangle \mathcal{B}_l(z),$$

has real-valued coefficients. For the kth partial sum already has real-valued coefficients, it just need to prove the sum of $(k+1)$th and $(k+2)$th terms has real-valued coefficients. Computation gives

$$\begin{aligned} & \langle g_{k+1}, e_{k+1}\rangle \mathcal{B}_{k+1}(z) + \langle g_{k+2}, e_{k+2}\rangle \mathcal{B}_{k+2}(z) \\ &= \frac{1-|\zeta_{k+1}|^2}{1-\overline{\zeta}_{k+1} z} g_{k+1}(\zeta_{k+1}) \prod_{i=1}^{k} \frac{z-\zeta_i}{1-\overline{\zeta}_i} \\ &+ \frac{1-|\zeta_{k+2}|^2}{1-\overline{\zeta}_{k+2} z} g_{k+2}(\zeta_{k+2}) \prod_{i=1}^{k+1} \frac{z-\zeta_i}{1-\overline{\zeta}_i} \\ &= F(z) \prod_{i=1}^{k} \frac{z-\zeta_i}{1-\overline{\zeta}_i z}, \end{aligned}$$

where

$$\begin{aligned} F(z) &= \frac{1-|\zeta_{k+1}|^2}{1-\overline{\zeta}_{k+1} z} g_{k+1}(\zeta_{k+1}) \\ &\quad + \frac{1-|\zeta_{k+2}|^2}{1-\overline{\zeta}_{k+2} z} g_{k+2}(\zeta_{k+2}) \frac{z-\zeta_{k+1}}{1-\overline{\zeta}_{k+1} z}. \end{aligned}$$

Then from the proof given in the previous step, $F(z)$ has real-valued coefficients. Consequently, $f_{k+2}(z)$ has real-valued coefficients. Mathematical induction gives the desired result. The proof is complete.

4.2 Analysis with noise

With the noise incorporated, the first approximation in our two-steps algorithm becomes

$$\widetilde{f}(z) = \widetilde{f}_N(z) + \widetilde{V}_N(z), \tag{21}$$

where $\widetilde{f}_N(z)$ is the first step approximation to the true function $f(z)$,

$$\widetilde{f}_N(z) = \frac{1}{2\pi j}\int_0^{2\pi} \frac{\sum_k f(e^{j\omega_k})\chi_{k(\cdot,\cdot)}(\omega)}{e^{j\omega} - z} de^{j\omega}, \tag{22}$$

and $\widetilde{V}_N(z)$ is the approximation to the noise, denote $v(\omega) = \sum_k v_k \chi_{k(\cdot,\cdot)}(\omega)$,

$$\widetilde{V}_N(z) = \frac{1}{2\pi j} \int_0^{2\pi} \frac{v(\omega)}{e^{j\omega} - z} de^{j\omega}. \tag{23}$$

We first assume that the noise is deterministic and bounded, $|v_k| \leq \epsilon$, $\epsilon > 0$. There is

$$\|v\|_2 = (\frac{1}{2\pi} \int_0^{2\pi} |v|^2 d\omega)^{\frac{1}{2}} \leq \epsilon.$$

Then we have the following result.

Theorem 5 *Let $f(z)$ be in $A(\mathbb{D})$, the disc algebra. Given the measurements $\{E_k\}_{k=1}^N$ with bounded noise $|v_k| \leq \epsilon$, $\epsilon > 0$. If $f_n(z)$ is the obtained nth partial sum to $\widetilde{f}(z)$ (21) by AFD method. Then the approximating sequence $f_n(z)$ has the following property:*

$$\lim_{\substack{N\to\infty, n\to\infty, \\ \epsilon \to 0}} \|f - f_n\|_{H_2} = 0. \tag{24}$$

Proof. By inserting the limit term, we have

$$\begin{aligned} \|f - f_n\|_{H_2} &\leq \|f - \widetilde{f}\|_{H_2} + \|\widetilde{f} - f_n\|_{H_2} \\ &\leq \|f - \widetilde{f}_N\|_{H_2} + \|\widetilde{V}_N\|_{H_2} + \|\widetilde{f} - f_n\|_{H_2}. \end{aligned}$$

The Hardy space theory implies that the H_2 norm is dominated by the L_2 norm of the boundary data. Denoting by $\mathbf{C}$ the Cauchy integral operator, we have

$$\begin{aligned} \|f - \widetilde{f}_N\|_{H_2} &= \|\mathbf{C}(f) - \mathbf{C}(\sum_k f(e^{j\omega_k})\chi_{k(\cdot,\cdot)}(\cdot))\|_{H_2} \\ &\leq \|f - \sum_k f(e^{j\omega_k})\chi_{k(\cdot,\cdot)}(\cdot)\|_{L_2}. \end{aligned}$$

By the same reasoning, we have

$$\|\widetilde{V}_N\|_{H_2} \leq \|v\|_{L_2} \leq \epsilon. \tag{25}$$

Meanwhile, the adaptive approximation for $\widetilde{f}$ implies

$$\lim_{n\to\infty} \|\widetilde{f} - f_n\|_{H_2} = 0.$$

By passing the limit $N, n \to \infty$ and $\epsilon \to 0$, the proof is complete.

For the stochastic case, we define

$$\begin{aligned} &H_2(\mathcal{D}, M) \\ &= \{f \in H_2(\mathbb{D}) : f = \sum_{k=1}^{\infty} d_k e_{\zeta_k}, e_{\zeta_k} \in \mathcal{D}, \sum_k |d_k| < M\}, \end{aligned}$$

where $0 < M < \infty$, $\mathcal{D} = \{e_{\zeta_k}\}$ is the dictionary and $\{d_k\}_{k=1}^{\infty}$ are the related coefficients [10]. The following result holds.

Theorem 6 *Let $f(z) \in A(\mathbb{D})$, $\{E_k\}_{k=1}^N$ be the frequency-domain measurements corrupted by stochastic noise $\{v_k\}$ with zero-mean and $\mathbf{E}|v_k|^2 < \sigma^2 < \infty$. If $\widetilde{f}(z)$ is in $H_2(\mathcal{D}, M)$, then the obtained nth partial sum $f_n(z)$ by AFD satisfies*

$$\mathbf{E}\|f - f_n\|_{H_2}^2 \leq 4w_N^2(f) + 4\sigma^2 + 2\frac{M^2}{n}, \tag{26}$$

where

$$w_N(f) = \sup_{\substack{|\Delta\omega| \leq |\omega_{k+1} - \omega_k| \\ 1 \leq k \leq N}} |f(e^{j(\omega + \Delta\omega)}) - f(e^{j\omega})|.$$

Proof. By inserting middle terms, we have

$$\begin{aligned} &\mathbf{E}\|f - f_n\|_{H_2}^2 \\ &\leq 2\mathbf{E}\|f - \widetilde{f}\|_{H_2}^2 + 2\mathbf{E}\|\widetilde{f} - f_n\|_{H_2}^2 \\ &\leq 4\|f - \widetilde{f}_N\|_{H_2}^2 + 4\mathbf{E}\|\widetilde{V}_N\|_{H_2}^2 + 2\mathbf{E}\|\widetilde{f} - f_n\|_{H_2}^2. \end{aligned} \tag{27}$$

For the first term of (27), from the proof of Theorem 5, we have

$$\begin{aligned} \|f - \widetilde{f}_N\|_{H_2}^2 &\leq \|f - \sum_k f(e^{j\omega_k})\chi_{(\cdot,\cdot)}\|_{L_2}^2 \\ &\leq w_N^2(f). \end{aligned}$$

Noting that $|v|^2 = \sum_k |v_k|^2 \chi_{k(\cdot,\cdot)}(\omega)$, we have

$$\begin{aligned} \mathbf{E}\|\widetilde{V}_N\|_{H_2}^2 &\leq \mathbf{E}\|v\|_{L_2}^2 \\ &\leq \mathbf{E}\frac{1}{2\pi}\int_0^{2\pi} |v|^2 d\omega \\ &= \frac{1}{2\pi}\int_0^{2\pi} \mathbf{E}|v|^2 d\omega \\ &= \frac{1}{2\pi}\int_0^{2\pi} \sum_k \mathbf{E}|v_k|^2 \chi_{k(\cdot,\cdot)}(\omega) d\omega \\ &\leq= \frac{1}{2\pi}\int_0^{2\pi} \sigma^2 d\omega \\ &\leq \sigma^2. \end{aligned}$$

According to [10,31], if $\widetilde{f}(z) \in H_2(\mathcal{D}, M)$, then

$$\|\widetilde{f} - f_n\|_2 \leq \frac{M}{\sqrt{n}}. \tag{28}$$

Therefore

$$\begin{aligned}\mathbf{E}\|\widetilde{f} - f_n\|_2^2 &= \int_{-\infty}^{\infty} \|\widetilde{f}(\cdot,\xi) - f_n(\cdot,\xi)\|_2^2 d\mu(\xi) \\ &\leq \int_{-\infty}^{\infty} \frac{M^2}{n} d\mu(\xi) \\ &= \frac{M^2}{n},\end{aligned}$$

where $\mu(\cdot)$ is the probability measure. Thus, the proof is complete.

Note that on the right side of the result (26) the first term goes to zero as N goes to infinity because the maximum frequency gap tends to zero. The third term goes to zero as n goes to infinity. The middle term remains constant, that is, mean-square-error bound attributed to noise is constant.

Remark 7 *For the pointwise convergence aspect, we cite a result showing that the TM system expansion are of the same nature as Fourier expansions [3].*

Let $A(\mathbb{D}_r, K)$ denote the space of functions analytic in the disc with radius $r > 1$ and bounded by $K < \infty$ in the area, then [3] showed that the partial sums of any rational orthogonal bases (6), satisfying the Szász condition (7), pointwisely converge to the system with an exponential convergence rate.

Precisely, if $f(z) \in A(\mathbb{D}_r, K)$, the remainder of adaptive decomposition $R_k(z)$ (14) in section 3 satisfies

$$\|R_k(z)\|_\infty \leq \frac{Kr}{r-1} \exp\Big(-\frac{r-1}{2r}\sum_{l=1}^{k}(1-|\zeta_l|)\Big). \quad (29)$$

We thus observe that if $f(z)$ is continuously extendable to outside the closed unit disc and the selections of the ζ_k satisfy $1 - |\zeta_k| > \delta > 0$, then the L_∞ norm of the remainder is exponentially decaying.

It is proved in [10,31] that the decay rate of the standard remainder of AFD (14) is the negative square root of the corresponding partial sum order. The result is sharp because it deals with the worst cases without assuming smoothness of the boundary function.

Remark 8 *As guaranteed by the convergence Theorem 2, the alteration should stop after certain steps. The convergence rates proved for the general greedy algorithm are, in fact, rather modest (28). There is no theoretical result to guarantee that for a given system function in H_2 the convergence of AFD is faster than any specific but determined TM system, including Fourier series. The energy effective algorithm, however, implies a practical algorithm for the best approximation to a Hardy space function by rational functions of a certain degree or less [32].*

5 Example

In this section we give an example for the algorithm described above. We work on the system considered by [24]:

$$G(z) = \frac{z^{-2}(0.0355z + 0.0247)}{(z - 0.9048)(z - 0.3679)}. \quad (30)$$

Applying the transformation $z \longrightarrow 1/z$ to $G(z)$ we have

$$f(z) = G(1/z) = \frac{z^3(0.0247z + 0.0355)}{(1 - 0.9048z)(1 - 0.3679z)}. \quad (31)$$

Here we use the values $m = 300, 500, 600, 800, 1000$ frequency-domain measurements, respectively, in the interval $[0, \pi)$ corresponding to a half circle that, in fact, represent $N = 2m$ points on the full circle. By using the adaptive algorithm we obtain the nth partial sum $f_n(z)$ with n components, and $G_n(z) = f_n(1/z)$.

Table 1
Modified greedy sequences in the disc.

	ζ_1	ζ_2	ζ_3
$m = 300$	0.9517	$0.7145 - 0.2959j$	$0.7145 + 0.2959j$
$m = 500$	0.9525	$0.7145 + 0.2916j$	$0.7145 - 0.2916j$
$m = 600$	0.9533	$0.7145 + 0.2959j$	$0.7145 - 0.2959j$
$m = 800$	0.9533	$0.7137 - 0.2956j$	$0.7137 + 0.2956j$
$m = 1000$	0.9542	$0.7152 - 0.2963j$	$0.7152 + 0.2963j$

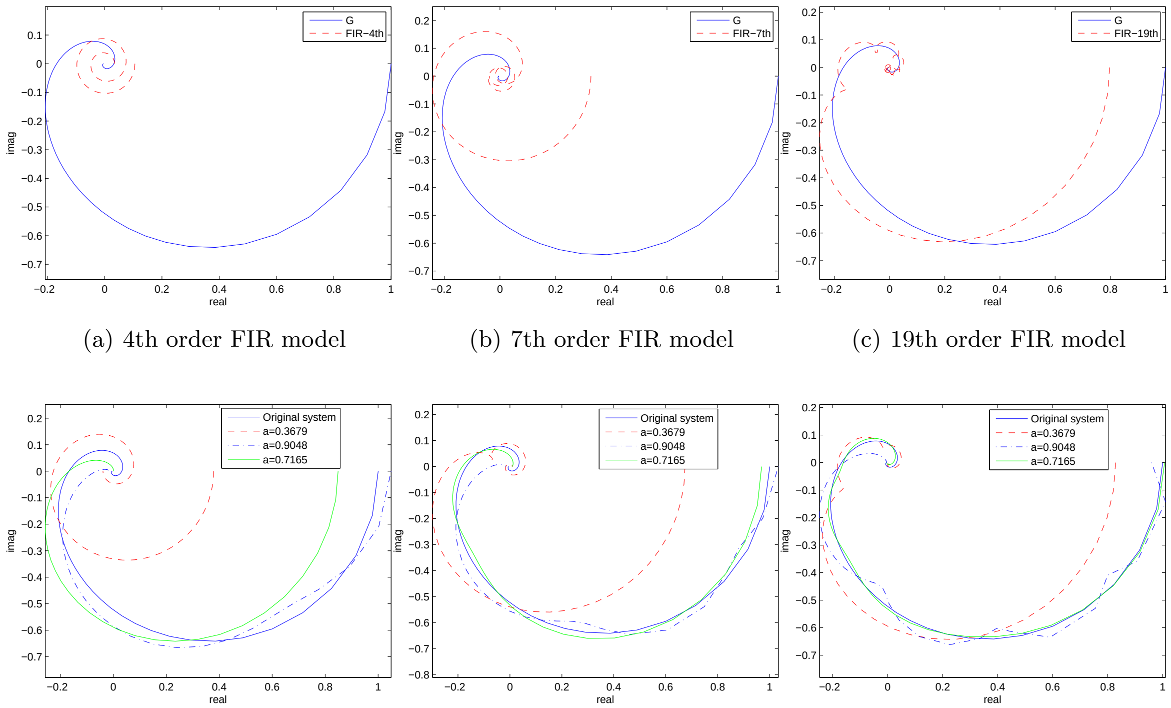
Table 2
Continued modified greedy sequences in the disc.

	ζ_4	ζ_5	ζ_6
$m = 300$	0.9258	$0.3525 - 0.6743j$	$0.3525 + 0.6743j$
$m = 500$	0.9225	$0.3537 + 0.6680j$	$0.3537 - 0.6680j$
$m = 600$	0.9233	$0.3529 - 0.6750j$	$0.3529 + 0.6750j$
$m = 800$	0.9225	$0.3564 - 0.6731j$	$0.3564 + 0.6731j$
$m = 1000$	0.9225	$0.3529 - 0.6750j$	$0.3529 + 0.6750j$

Table 3
Greedy sequences selected in $(-1, 1)$.

	ζ_1	ζ_2	ζ_3	ζ_4	ζ_5
$m = 300$	0.9519	0.5900	0.9479	0.2653	0.8940
$m = 500$	0.9529	0.5915	0.9469	0.2658	0.8945
$m = 600$	0.9534	0.5910	0.9469	0.2658	0.8955
$m = 800$	0.9534	0.5930	0.9464	0.2653	0.8950
$m = 1000$	0.9539	0.5920	0.9464	0.2653	0.8959

The frequency responses of FIR model, Laguerre model and our method are compared in figure 1 and figure 2. In all the subfigures the solid lines are the original system $G(z)$. In figure 1, the three on the first row 1(a), 1(b), 1(c) give the frequency responses of the 4th order, 7th order and 19th order FIR model, respectively. While the 1(d), 1(e), 1(f) in the second row of figure 1 give the frequency responses of the 4th order, 7th order and 10th order Laguerre model with $a = 0.3879$, $a = 0.9048$ and $a = 0.7165$, respectively. In figure 2, the three on the



(a) 4th order FIR model (b) 7th order FIR model (c) 19th order FIR model

(d) 4th order Laguerre model (e) 7th order Laguerre model (f) 10th order Laguerre model

Fig. 1. Frequency responses of fixed-pole models: FIR model and Laguerre model with priori-known knowledge $a = 0.3879, 0.9048, 0.7165$.

Table 4
Continued greedy sequences selected in $(-1, 1)$.

	ζ_6	ζ_7	ζ_8	ζ_9	ζ_{10}
$m = 300$	-0.2430	0.6955	-0.4585	0.6564	-0.4983
$m = 500$	-0.2440	0.6930	-0.4560	0.6668	-0.4924
$m = 600$	-0.2440	0.6935	-0.4560	0.6682	-0.4909
$m = 800$	-0.2450	0.6920	-0.4545	0.6737	-0.4892
$m = 1000$	-0.2445	0.6925	-0.4540	0.6742	-0.4882

Table 5
Greedy sequences selected in $(-1, 1)$ using data with added noise.

	ζ_1	ζ_2	ζ_3	ζ_4	ζ_5
$m = 300$	0.9568	0.5940	0.9440	0.2713	0.8786
$m = 500$	0.9568	0.6009	0.9499	0.2430	0.8643
$m = 1000$	0.9544	0.6014	0.9454	0.2713	0.9058

Table 6
Continued greedy sequences selected in $(-1, 1)$ using data with added noise.

	ζ_6	ζ_7	ζ_8	ζ_9	ζ_{10}
$m = 300$	-0.1880	0.7752	-0.2620	-0.9000	0.2930
$m = 500$	-0.2245	0.7158	-0.2420	0.9747	0.6831
$m = 1000$	-0.2835	0.6509	-0.4130	0.6623	-0.4410

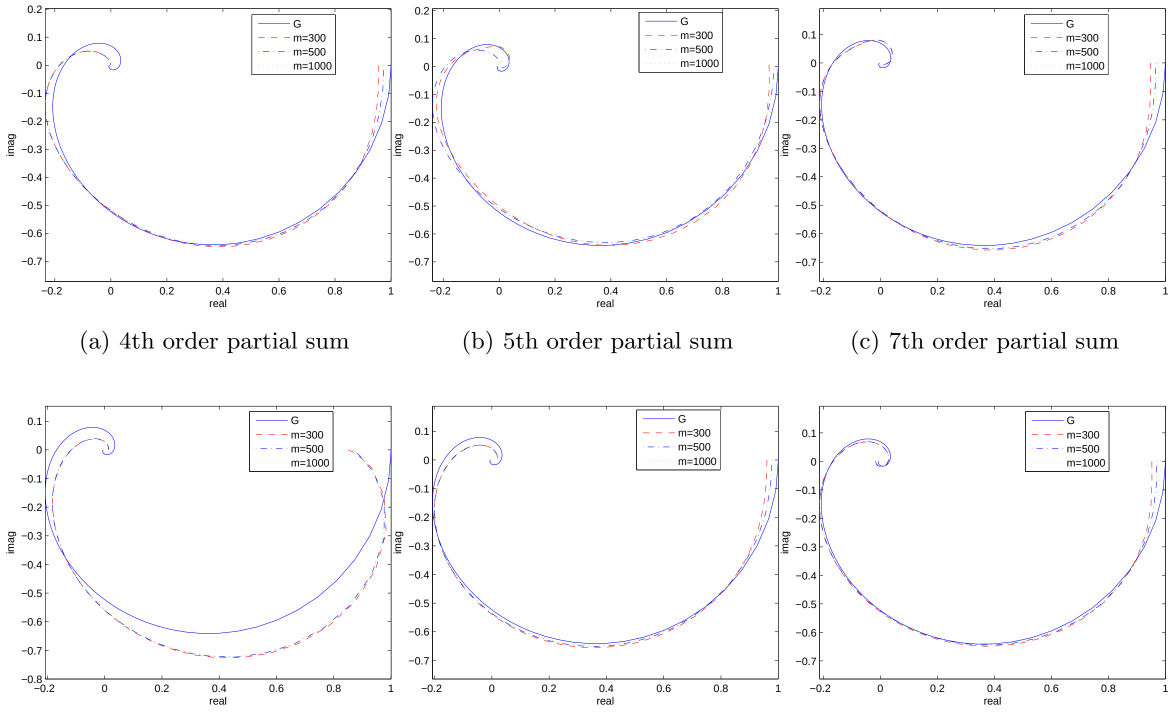
first row 2(a), 2(b), 2(c) give the frequency responses of the 4th order, 5th order and 7th order systems of our method, respectively.

The figures show that we have better approximations to the original system with the AFD algorithm. For the FIR model, even the 19th FIR model is unsatisfactory. The Laguerre model looks better than the FIR model. However, from 1(d), 1(e), 1(f) in figure 1, we can see when $a = 0.3679$, the frequency response of the Laguerre model is the furthest away from the original system. While when $a = 0.9048$, the curve keeps oscillating. This motivates adaptive selection of poles.

Table 1 and Table 2 are the listed modified greedy sequences in the whole disc. Table 3 and 4 list several of the greedy sequences selected in the interval $(-1, 1)$. The pictures in figure 2 are the 4th, 5th and 7th partial sums by the two different choosing of the optimal points, respectively.

When noise is incorporated, assume that the frequency-domain measurements are sampled with added Gaussian noise with $SNR = 20$. In the first row of figure 3, the red dots are the data with noise, while there are the 4th order, 5th order and 7th order approximating partial sums by our method in the second row of figure 3. It shows that this method is efficient in dealing with noise.

As mentioned before, the first step approximation can use other methods other than the Cauchy integral one. Approximations, for instance, may be referred to [11]. In this paper, we use the Cauchy integral approach in



(a) 4th order partial sum (b) 5th order partial sum (c) 7th order partial sum

(d) 4th order, real-valued poles (e) 5th order, real-valued poles (f) 7th order, real-valued poles

Fig. 2. Frequency responses with the AFD algorithm used: in the first row, modified greedy sequence is used, while in the second row, the greedy sequence selected in the interval $(-1, 1)$ is used, m is the number of data.

step 1 just for easy illustration. The main establishment is the adaptive use of AFD algorithm for approximation of considered systems, that gives experimental results.

6 Conclusion

So far, there can be said to have 3 usages of the rational orthogonal system $\{\mathcal{B}_k\}$ in system identification. One is the FIR model that is without any previous knowledge on the poles of the system, while the poles of $\{\mathcal{B}_k\}$ are all chosen to be zero. The second is the Lagurre model, Kautz model and the generalized model in which the poles of the true system are used as the poles of $\{\mathcal{B}_k\}$. The last one is to adaptively select the poles of $\{\mathcal{B}_k\}$ according to the frequency domain measurements.

In general, a rough choice of the poles for the basis functions will lead to a large number of basis functions than what are required. With sufficiently many sample points, by using the proposed adaptive algorithm, one can consecutively obtain a greedy sequence defining the approximating orthogonal system. Although the obtained points are not necessarily the true poles, one can obtain functions approaching to the transfer function in great efficiency.

7 Acknowledgements

The authors would like to express their sincere thankfulness to the two referees. They patiently reviewed the manuscript and the revised versions and gave many detailed constructive suggestions. Especially, to answer their challenging questions is to better understand the subject. The authors also wish to thank Brett Ninness and Feng Wan for their support and helpful discussions on this study.

References

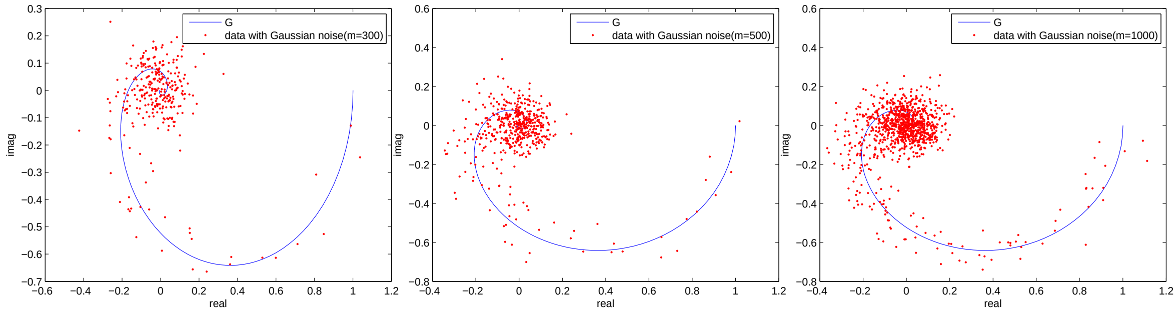
[1] Akçay, H. (2001). On the uniform approximation of discrete-time systems by generalized Fourier series. *IEEE Trans. Signal Process.*, 49, 1461-1467.

[2] Akçay, H., & Ninness, B. (1999). Orthonormal basis functions for modelling continuous-time systems, *Signal Processing*, 77, 261-274.

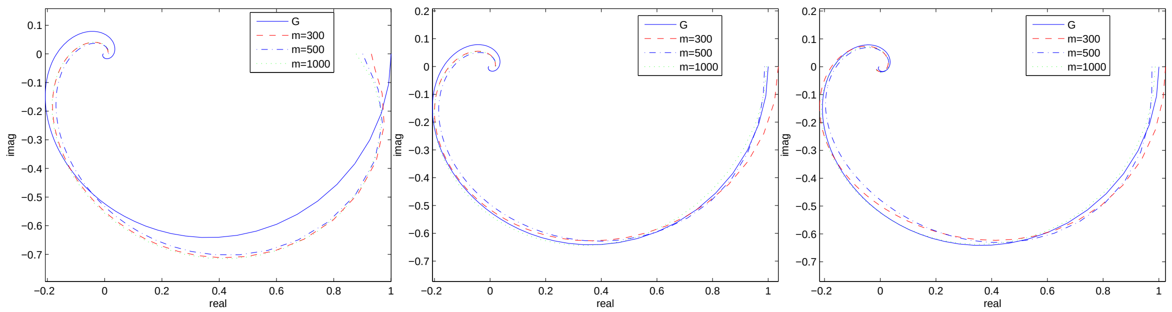
[3] Akçay, H., & Ninness, B. (1998). Rational basis functions for robust identification from frequency and time domain measuements. *Automatica*, 34, 1101-1117.

[4] Akçay, H., & Ninness, B. (1999). Orthonormal basis functions for continuous-time systems and L_p convergence. *Math. Control, Signals Systems*, 295–305.

[5] Billings, S.A., & Wei, H.L. (2008). An adaptive orthogonal search algorithm for model subset selection and non-linear system identification. *Internat. J. Control,* 81, 714-724.



(a) Data with noise (b) Data with noise (c) Data with noise



(d) 4th order partial sum (e) 5th order partial sum (f) 7th order partial sum

Fig. 3. 3(a),3(b),3(c) show the data added noise, while 3(d),3(e),3(f) give the relevant 4th,5th and 7th order adaptive approximation with AFD algorithm using the noised data, m is the number of data.

[6] Bultheel, A., & Gucht, P.V. (2000). Boundary asymptotics for orthogonal rational functions on the unit circle. *Acta Appl. Math.*, 61, 333-349.

[7] Bultheel, A., & Carrette, P. (2003). Fourier analysis and the Takenaka-Malmquist basis, *Proceedings of the 42nd IEEE Conference on Decision and Control*, 09/01/2004; 1:486- 491 Vol.1.

[8] Cluett, W., & Wang, L. (1992). Frequency smoothing using Laguerre model. *Proc. IEE-D*, 139, 88-96.

[9] Davis, G., Mallat, S., & Avellaneda, M. (1997). Adaptive greedy approximations. *Constr. Approx.*, 13, 57-98.

[10] Devore, R.A., & Temlyakov, V.N. (1996). Some remarks on greedy algorithm, *Adv. Comput. Math.,* 5, 173-187.

[11] Gu, G., & Khargonekar, P.P. (1992). A class of algorithms for identification in H_∞, *Automatica*, no.2, 28, 299-312.

[12] Gucht, P.V., & Bultheel, A. (2003). Orthonormal rational functions for system identification: numerical aspects. *IEEE Trans. Automat. Control*, 48, 705-709.

[13] Heuberger, P.S.C., Van den Hof, P.M.J., & Wahlberg, B.(Eds.) (2005). *Modelling and Identification with Rational Orthogonal Basis Functions.* Springer, London.

[14] Hunt, R.A. (1968). On the convergence of Fourier series, orthogonal expansions and their continuous analogues. *Proc. conf. Edwardsville*, I11. pp. 235-255; Southern Illinois Univ. Press, Carbondale, I11.

[15] King, R., & Paraskevopoulos, P. (1979). Parametric identification of discrete time SISO systems. *Internat. J. of Control*, 30, 1023-1029.

[16] Li, C.J., & Andersen, S.V. (2007). Efficient blind system identification of non-Gaussian autoregressive models with HMM modeling of the excitation. *IEEE Trans. on Signal Processing*, 55, 2432-2445.

[17] Ljung, L. (1999). *System Identification: Theory for the User, Second Edition.* Prentice-Hall, Inc., New Jersey.

[18] Longman, R., & Phan, M.Q. (2006). Iterative learning control as a method of experiment design for improved system identification. *Optimization Methods and Software*, 21, 919-941.

[19] Mäkilä, P. (1990). Approximation of stable systems by Laguerre filters. *Automatica*, 26, 333-345.

[20] Mäkilä, P. (1990). Laguerre series approximation of infinite dimensional systems. *Automatica*, 26, 985-995.

[21] Mäkilä, P. (1991). Laguerre methods and H^∞ identification of continuous-time systems. *Internat. J. Control*, 53, 689-707.

[22] Nurges, Ü. (1987). Laguerre methods in problems of approximation and identification. *In Adaptive systems, Plenum,* 346-352. Translated from Avtomatica i Telemekhanika, 3(March 1987), 88-96.

[23] Ninness, B. (1996). Frequency domain estimation using orthonormal bases. *In Proceedings of the 13th IFAC World Congress,* 381-386.

[24] Ninness, B., & Gustafsson, F. (1997). A unifying construction of orthonormal bases for system identification. *IEEE Trans. Automat. Control,* 42, 515-512.

[25] Ninness, B., Hjalmarsson, H., & Gustafsson, F. (1997). The fundamental role of general orthonormal bases in system identification. *IEEE Trans. Automat. Control,* 44, 1384-1406.

[26] Oliveira e Silva, T. (1996). A n-width result for the generalized orthonormal basis function model. *In Preprints*

of the 13-th IFAC World Congress, volume I, 375-380, San Francisco, USA.

[27] Partington, J.R. (1997). *Interpolation, Identification, and Sampling*, Oxford University Press Inc., New York.

[28] Pintelon, R., Guillaume, P., Rolain, Y., & Schoukens, J., Hamme, H.V. (1994). Parametric identification of transfer functions in the frequency domain-a survey. *IEEE Trans. Automat. Control*, 39, 2245-2260.

[29] Qian, T. (2009). Intrinsic mono-component decomposition of functions: An advance of Fourier theory. *Mathematical Methods in Applied Sciences*, DOI:10.1002/mma.1214.

[30] Qian, T., & Wang, Y.B. (2010). Adaptive decomposition into basic signals of non-negative instantaneous frequencies - a variation and realization of greedy algorithm. *Adv. Comput. Math.,* DOI:10.1007/s10444-010-9153-4.

[31] Qian,T., & Wang, Y.B (2011). Convergence rate of Adaptive Fourier Decomposition. Submitted to J. Math. Anal. Appl..

[32] Qian T. & Wegert E. (2010), Best Approximation by Blaschke Forms and Rational Functions of a Finite Degree, accepted to appear in *Complex Variables and Elliptic Equations* .

[33] Sandberg, H., Lanzon, A., & Anderson, B.D.O. (2007). Model approximation using magnitude and phase criteria: Implications for model reduction and system identification. *Int. J. Robust Nonlinear Control,* 17, 435-461.

[34] Szász, O. (1953). On closed sets of rational functions. *Annali di Matematica Puru ed Applicata,* Serie Quarta-Tomo XXXXIV:195-218, 1953.

[35] Tan, L.H., & Zhou, C.Y. (2010). The point-wise convergence of general rational Fourier series. accepted by *Math. Meth. Appl. Sci.*.

[36] Troutt, M.D., Tadisina, S.K., & Sohn, C., Brandyberry, A.A. (2005). Linear programming system identification. *Eur. J. Oper. Res.,* 161, 663-672.

[37] Troutt, M.D., Brandyberry, A.A., & Sohn, C., Tadisina, S.K. (2008). Linear programming system identification: The general nonnegative parameters case. *Eur. J. Oper. Res.*, 185, 63-75.

[38] Vries, D.K.de, & Van den Hof, P.M.J. (1998). Frequency domain identification with generalized orthonomal basis functions. *IEEE Trans. Automat. Control,* v.43 No.5, 656-669.

[39] Wahlberg, B. (1991). System identification using Laguerre models. *IEEE Trans. Automat. Control,* AC-36, 551-562.

[40] Wahlberg, B. (1991). Identification of resonant systems using Kautz filters. *In proceedings of the 30th IEEE Conference on Decision and Control,* 2002-2010.

[41] Wahlberg, B. (1994). System identification using Kautz models. *IEEE Trans. Automat. Control,* AC-39, 1276-1282.

[42] Wahlberg, B., & Mäkilä, P. (1996). On approximation of stable linear dynamical systems using Laguerre and Kautz functions. *Automatica,* 32, 693-708.

[43] Walsh, J.L. (1965). *Interpolation and approximation by rational functions in the complex domain.* American Mathematical Scienty, Fourth edition.