A Fast Adaptive Model Reduction Method Based on Takenaka-Malmquist Systems [¤]

Wen Mi[†] Tao Qian[‡] Feng Wan[§]

Abstract. An adaptive model reduction method is proposed for linear time-invariant systems based on the continuous-time rational orthogonal basis (Takenaka-Malmquist basis). The method is to find an adaptive approximation in the energy sense by selecting optimal points for the rational orthogonal basis. The stability of the reduced models holds, and the steady-state values of step responses are kept to be equal. Furthermore, the method automatically ensures the reduced system to be in the Hardy space H_2. The existence of the best approximation in the Hardy space H_2 by n Blaschke forms is proved in the proposed approach. The effectivity of this method is illustrated through three well-known examples.

Key words. Best approximation; impulse response energy; model order reduction; rational approximation; Takenaka-Malmquist basis

AMS Subject Classification. 93A30; 41A20; 93C05

1 Introduction

Modeling of complex dynamic systems is an important problem in engineering. A complicated model may lead to difficulties in both system analysis and controller design. A simple and efficient model that matches the original system well under the given criteria is always desired. In practice, different criteria are adopted for different applications. Model reduction has received considerable attentions, various methods and techniques have been proposed, including linear matrix inequalities [25], Routh approximations [32, 54], error minimization techniques [36, 39], magnitude and phase criteria [52], H_∞ model reduction [24], the Padé type model reductions [45, 22],

[¤]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.

[y]Department of Mathematics, Faculty of Science and Technology, University of Macau, Macau, China. E-mail:ya87405@umac.mo.

[z]Department of Mathematics, Faculty of Science and Technology, University of Macau, Macau, China. E-mail: fsttq@umac.mo.

[x]Department of Electrical and Electronics Engineering, Faculty of Science and Technology, University of Macau, Macau, China. E-mail: fwan@umac.mo.

balanced truncation [37, 4, 29, 46], rational interpolation [30, 23, 26] and Krylov method [6, 27, 28, 29, 30], and so on.

The model reduction problem for continuous linear time-invariant systems can be formalized either in time domain or frequency domain. In this paper, we consider the problem in the frequency domain, in which the original system is given by the following $p \times q$ transfer function matrix of the form

$$G(s) = \frac{\sum_{i=0}^{m-1} C_i s^i}{\sum_{i=0}^{m} b_i s^i}, \tag{1.1}$$

where C_i are $p \times q$ constant matrices, m is the order of the system and b_i are the scalar coefficients of the characteristic polynomial of the system, b_m is normalized to be equal to 1.

The reduced model transfer function is assumed to have the form:

$$\widetilde{G}(s) = \frac{\sum_{i=0}^{n-1} \widetilde{C}_i s^i}{\sum_{i=0}^{n-1} \widetilde{b}_i s^i}, \quad n < m, \tag{1.2}$$

where $\widetilde{C}$ are $p \times q$ constant matrices, n is the reduced order and $\widetilde{b}_i$ are the scalar coefficients of the characteristic polynomial of the reduced-order model with $\widetilde{b}_n = 1$.

The presented new method for model reduction is based on the continuous-time rational orthogonal basis (Takenaka-Malmquist (TM) basis) defined by a sequence $\{a_k\} \in \Pi$, Π is the open right-half plane, as

$$B_k \triangleq \frac{\sqrt{2\Re\{a_k\}}}{s + \overline{a_k}} \prod_{l=1}^{k-1} \frac{s - a_l}{s + \overline{a_l}}, \quad k = 2, \ldots, \tag{1.3}$$

with $B_1 \triangleq \frac{\sqrt{2\Re\{a_1\}}}{s+\overline{a_1}}$ and $\Re\{\cdot\}$ denotes the real part of a complex number. Defining the inner product by

$$\begin{aligned} \langle B_p, B_q \rangle &\triangleq \frac{1}{2\pi} \int_{-\infty}^{\infty} B_p(j\omega) \overline{B_q(j\omega)} d\omega \\ &= \begin{cases} 1, & p = q \\ 0, & p \neq q, \end{cases} \end{aligned}$$

then $\{B_k\}_{k \geq 1}$ is an orthogonal system in the Hardy-2 space $H_2(\Pi)$.

The rational orthogonal basis (1.3) is the continuous-time case of the Takenaka-Malmquist basis, which is a natural generalization of Laguerre and Kautz bases and enjoys a long history of development and applications both in pure mathematics [21, 55, 61, 11, 57] and engineering literature [1, 34, 60, 31, 58, 12].

It is proved in [1] that:

Theorem 1.1 *The model set spanned by the basis functions $\{B_k\}_{k\geq 0}$ (set $B_0 = 1$) is complete in all of the spaces $H_p(\Pi)$, $1 < p < \infty$, and $A(\Pi)$, where $A(\Pi)$ is the right half plane algebra, if and only if*

$$\sum_{k=1}^{\infty} \frac{\Re\{a_k\}}{1+|a_k|^2} = \infty. \tag{1.4}$$

In system identification and control, the rational orthogonal basis method has been explored by many researchers for a long time. Many have been working on the optimal poles for the rational bases. T. Oliveria e Silva derived the optimal pole condition for Laguerre, Kautz and general orthogonal basis function models in [41, 42, 43], respectively, by deducing the first derivative of the remainder error functions. Others attempt to estimate a good pole position of a Laguerre model are given in [51, 33, 13]. Generally in practice, the poles' locations of the rational orthogonal basis usually rely on the prior-known knowledge on the true system and are prior given. While in this paper the poles $\{-\overline{a_k}\}$ are adaptively chosen through maximizing the function which depends on the deduced recursive formula in the 2-norm sense. For different systems, according to the algorithm, there would be relevant sequences $\{a_k\}$ which generate the basis, and that is the adaptivity.

Recently, T. Qian *et al* developed an adaptive algorithm in [49] and [48]. It aims to obtain approximations through consecutively selecting optimal points for the given functions in H_2–square integrable Hardy spaces in the unit disc and the open upper half plane. We will introduce the method for functions in $H_2(\Pi)$. Further more, we improve [48] by choosing the poles simultaneously (also see [50]).

This paper is organized as follows. Section 2 provides the main mathematical theory, where the approximation algorithm based on (1.3) for functions in $H_2(\Pi)$ is introduced, it can be treated as a new realization of n-best rational approximation. This algorithm is applied in model order reduction with modifications and an added
arTJ/F610.-11lalgori431(mo)-28432optimal thlas basis54924.209D[(1)9ent(usualF610.909Tf37.096-38821T

of the best approximating rational function, but not poles. Our process is completely different. It directly reduces the criterion function to a rational function symmetric in the poles to be selected. We prove that the global maximal value of the criterion function is attainable in the interior points of the open set for the variables.

We use an algorithm that comes from the Gram-Schmidt orthogonalization process based on the shifted Cauchy kernels, resulting in the rational orthogonal basis. The closely related algorithm for the adaptive Fourier decomposition for $H_2(\mathbb{C}^+)$ functions is studied in [48].

Adopting change of variable $z \to -jz$ from the upper-half complex plane to the right-half complex plane, we can covert one of the two cases to the other. Precisely, for $f \in H_2(\Pi)$, through a transformation

$$F(z) = f(-jz), z \in \mathbb{C}^+,$$

then $F(z) \in H_2(\mathbb{C}^+)$. The Cauchy integral formula for functions $f \in H_2(\Pi)$ reads

$$f(s) = \frac{1}{2\pi j}\int_{j\mathcal{R}} \frac{f(\xi)}{s-\xi} d\xi.$$

Introducing the shifted Cauchy kernel $e_{\{a\}}$, defined by

$$e_{\{a\}} = \frac{\sqrt{2\Re\{a\}}}{s+\overline{a}},$$

we have the relation

$$\begin{aligned}
\langle f, e_{\{a\}} \rangle &= \frac{1}{2\pi}\int_{-\infty}^{\infty} f(j\omega)\overline{e_{\{a\}}(j\omega)}d\omega \\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty} f(j\omega)\frac{\sqrt{2\Re\{a\}}}{-j\omega+a}d\omega \\
&= \frac{1}{2\pi j}\int_{jR} f(s)\frac{\sqrt{2\Re\{a\}}}{-s+a}ds \\
&= \sqrt{2\Re\{a\}}f(a).
\end{aligned}$$

Now assume that f is the $H_2(\Pi)$ function to be approximated by a rational function P/Q, where Q is of order n, whose zeros are in the left-half complex plane, and P is of an order less than n.

Adopt the first approximation by a shifted Cauchy kernel $e_{\{a_1\}}$, with $f_1 = f$, we have

$$\begin{aligned}
f(s) &= \langle f_1, e_{\{a_1\}}\rangle e_{\{a_1\}}(s) + \left(f(s) - \langle f_1, e_{\{a_1\}}\rangle e_{\{a_1\}}(s)\right) \\
&= \langle f_1, e_{\{a_1\}}\rangle e_{\{a_1\}}(s) + f_2(s)\frac{s-a_1}{s+\overline{a_1}},
\end{aligned}$$

where

$$f_2(s) = \left(f(s) - \langle f_1, e_{\{a_1\}}\rangle e_{\{a_1\}}(s)\right)\frac{s+\overline{a_1}}{s-a_1}.$$

Note that the analytic function in the brackets has a zero at a_1, hence f_2 is still in $H_2(\Pi)$. For f_2, do the same process, we further express

$$
\begin{aligned}
f(s) &= \langle f_1, e_{\{a_1\}}\rangle e_{\{a_1\}}(s) + \langle f_2, e_{\{a_2\}}\rangle e_{\{a_2\}}(s)\frac{s-a_1}{s+\overline{a_1}} + f_3(s)\frac{s-a_2}{s+\overline{a_2}}\frac{s-a_1}{s+\overline{a_1}} \\
&= \langle f_1, e_{\{a_1\}}\rangle B_1(s) + \langle f_2, e_{\{a_2\}}\rangle B_2(s) + f_3(s)\frac{s-a_2}{s+\overline{a_2}}\frac{s-a_1}{s+\overline{a_1}},
\end{aligned}
$$

where f_3 is a function in the H_2 space. To f_3 we perform the process again. Repeat the processes till up to the nth step, we obtain

$$
\begin{aligned}
f(s) &= \langle f_1, e_{\{a_1\}}\rangle e_{\{a_1\}}(s) + \langle f_2, e_{\{a_2\}}\rangle e_{\{a_2\}}(s)\frac{s-a_1}{s+\overline{a_1}} \cdots + f_{n+1}(s)\prod_{k=1}^{n}\frac{s-a_k}{s+\overline{a_k}} \\
&= \sum_{k=1}^{n}\langle f_k, e_{\{a_k\}}\rangle B_k(s) + R_n(s)
\end{aligned}
$$

where $f_k(s)$ $(k = 1, 2, \ldots)$ are in $H_2(\Pi)$ and recursively

$$
f_k(s) = \left(f_{k-1}(s) - \frac{2\Re\{a_{k-1}\}f_{k-1}(a_{k-1})}{s+\overline{a_{k-1}}}\right)\frac{s+\overline{a_{k-1}}}{s-a_{k-1}}, \tag{2.1}
$$

$R_n(s)$ is the remainder

$$
R_n(s) = f_{n+1}(s)\prod_{k=1}^{n}\frac{s-a_k}{s+\overline{a_k}}. \tag{2.2}
$$

There holds

$$
\langle f_k, e_k\rangle = \langle f, e_k\rangle = \langle f, B_k\rangle.
$$

So for $f(s) \in H_2(\Pi)$ by Theorem 1.1, if $\{a_k\}_{k=1}^{\infty}$ satisfies the condition (1.4), we have

$$
f(s) = \sum_{k=1}^{\infty}\langle f_k, e_k\rangle B_k(s) = \sum_{k=1}^{\infty}\langle f_k, B_k\rangle B_k(s). \tag{2.3}
$$

The nth partial sum

$$
\widetilde{f}_n(s) = \sum_{k=1}^{n}\langle f_k, e_{\{a_k\}}\rangle B_k(s), \tag{2.4}
$$

is a rational function and is called the *n-approximating partial sum*, while $\{B_{\{a_1\}}$, $B_{\{a_1,a_2\}}$,..., $B_{\{a_1,a_2,\ldots,a_n\}}\}$ is called the *n Blaschke forms.*

The above process is embodied in [49] and [48]. It corresponds to one-by-one selection of the parameters $a_1, a_2, \ldots, a_n$. On the other hand, however, simultaneous selection of the parameters should lead to a better approximation. The main technical strength of the present paper is to treat the simultaneous selection issue in the unbounded domain. We start from dealing with H_2 functions with real or conjugated poles.

Lemma 2.1 *For functions $f(s) \in H_2(\Pi)$ with property $f(\overline{s}) = \overline{f(s)}$. In the process of the adaptive approximation, with $\{a_k\}_{k=1}^n$ appearing real or conjugated pairs, the approximating partial sum $\widetilde{f}_n(s)$ (2.4) is a rational function in $H_2(\Pi)$ with real coefficients.*

Proof. First, we show that if each a_l is either real or comes by a conjugate pair, the coefficients of the recursive formula $f_{k+1}(s)$ by (2.1) are real. We can see that if $\{a_l\}$ is real and $f_l(s)$ has real coefficients, then $f_{l+1}(s)$ also has real coefficients. So we only need to prove that $f_3(s)$ has real coefficients if a_1 and a_2 are conjugated. By substituting $f_2(s)$ with the recursive formula (2.1), we have

$$
\begin{aligned}
& f_3(s) \\
&= \frac{f_2(s)(s+\overline{a_2}) - 2\Re\{a_2\} f_2(\overline{a_2})}{s-a_2} \\
&= \frac{f(s)(s+\overline{a_1})(s+a_1) - 2\Re\{a_1\} f(a_1)(s+a_1)}{(s-\overline{a_1})(s-a_1)} - 2\Re\{a_1\}\frac{f(\overline{a_1})2\overline{a_1} - 2\Re\{a_1\} f(a_1)}{(s-\overline{a_1})(\overline{a_1}-a_1)} \\
&= F_1(s) - 2\Re\{a_1\} F_2(s),
\end{aligned}
$$

where

$$
\begin{aligned}
F_1(s) &= \frac{f(s)(s+\overline{a_1})(s+a_1)}{(s-\overline{a_1})(s-a_1)}, \\
F_2(s) &= \frac{f(a_1)(s+a_1)}{(s-\overline{a_1})(s-a_1)} + \frac{f(\overline{a_1})2\overline{a_1} - 2\Re\{a_1\} f(a_1)}{(s-\overline{a_1})(\overline{a_1}-a_1)} \\
&= \frac{2\Im\{a_1 f(a_1)\}s - 2|a_1|^2 \Im\{f(a_1)\}}{(s-a_1)(s-\overline{a_1})\Im\{a_1\}},
\end{aligned}
$$

$\Im\{\cdot\}$ denotes the imaginary part. We can see that the coefficients of $f_3(s)$ are real.

Now we use mathematical induction to prove the result. Without loss of generality, we set a_1 to be a complex number and $a_2 = \overline{a_1}$, then

$$
\begin{aligned}
\widetilde{f}_2(s) &= \langle f_1, e_1\rangle B_1(s) + \langle f_2, e_2\rangle B_2(s) \\
&= 2\Re\{a_1\}\frac{[f_1(a_1) + f_2(\overline{a_1})]s + f_1(a_1)a_1 - f_2(\overline{a_1})a_1}{(s+\overline{a_1})(s+a_1)}.
\end{aligned}
$$

Substituting the recursive formula, the coefficient of s in the numerator of the above rational function is

$$
\begin{aligned}
f_1(a_1) + f_2(\overline{a_1}) &= f(a_1) + \frac{f(\overline{a_1})(\overline{a_1}+\overline{a_1}) - 2\Re\{a_1\} f(a_1)}{\overline{a_1}-a_1} \\
&= \frac{2\overline{a_1} f(\overline{a_1}) - 2a_1 f(a_1)}{\overline{a_1}-a_1} \\
&= \frac{2\Im\{a_1 f(a_1)\}}{\Im\{a_1\}},
\end{aligned}
$$

and the constant term of the numerator is

$$
\begin{aligned}
f_1(a_1)a_1 + f_2(\overline{a_1})a_1 &= f(a_1)a_1 - a_1\frac{f(\overline{a_1})2\overline{a_1} - 2\Re\{a_1\}f(a_1)}{\overline{a_1} - a_1} \\
&= \frac{2|a_1|^2[f(a_1) - f(\overline{a_1})]}{\overline{a_1} - a_1} \\
&= -2|a_1|^2\frac{\Im\{f(a_1)\}}{\Im\{a_1\}}.
\end{aligned}
$$

They are both real. So the result holds for the first step of the induction. For the second step, assume that the kth ($k \geq 2$) partial sum $\widetilde{f}_k(s)$ has real coefficients where $a_1, a_2, \ldots, a_n$ are either real or come in conjugate pairs. In the case if a_{k+1} is real, the conclusion for $\widetilde{f}_{k+1}(s)$ holds. If a_{k+1} is complex, letting $a_{k+2} = \overline{a_{k+1}}$, we are to show that $\widetilde{f}_{k+2}(s)$ has real coefficients. Because the kth partial sum already has real coefficients, we only need to prove that the coefficients of the sum of $(k+1)$th and $(k+2)$th terms are real. Computation gives

$$
\begin{aligned}
&\langle f_{k+1}, e_{k+1}\rangle B_{k+1}(s) + \langle f_{k+2}, e_{k+2}\rangle B_{k+2}(s) \\
&= \frac{2\Re\{a_{k+1}\}f_{k+1}(a_{k+1})}{s + \overline{a_{k+1}}}\prod_{i=1}^{k}\frac{s - a_i}{s + \overline{a_i}} + \frac{2\Re\{a_{k+2}\}f_{k+2}(a_{k+2})}{s + \overline{a_{k+2}}}\prod_{i=1}^{k+1}\frac{s - a_i}{s + \overline{a_i}} \\
&= F(s)\prod_{i=1}^{k}\frac{s - a_i}{s + \overline{a_i}},
\end{aligned}
$$

where

$$
F(s) = \frac{2\Re\{a_{k+1}\}f_{k+1}(a_{k+1})}{s + \overline{a_{k+1}}} + \frac{2\Re\{a_{k+1}\}f_{k+2}(\overline{a_{k+1}})(s - a_{k+1})}{(s + \overline{a_{k+1}})(s + a_{k+1})}.
$$

Then from the proof given in the previous step, $F(s)$ has real coefficients. Consequently, $\widetilde{f}_{k+2}(s)$ has real coefficients. Mathematical induction gives the desired result. The proof is complete.

In [1, 12], there are methods to construct orthonormal basis with real coefficients when complex poles are drawn into. Based on Lemma (2.1), to have the approximations with real coefficients, it need not to make the basis functions with real coefficients.

Since all the terms on the right-hand-side of (2.4) are orthogonal to each other, we have the Plancherel theorem

$$
\|f\|^2 = \sum_{k=1}^{n}|\langle f_k, e_{\{a_k\}}\rangle|^2 + \|f_{n+1}\|^2.
$$

Denote

$$
A_f^n = |\langle f_1, e_{\{a_1\}}\rangle|^2 + |\langle f_2, e_{\{a_2\}}\rangle|^2 + \ldots + |\langle f_n, e_{\{a_n\}}\rangle|^2. \tag{2.5}
$$

Our main result is as follows.

Theorem 2.1 *Let $f(s) \in H_2(\Pi)$, and A_f^n be defined by (2.5). Then there exists an n-tuple $\{a_1, a_2, \ldots, a_n\}$ in the right-half plane that makes A_f^n to be the global maximum, i.e,*

$$\{a_1, a_2, \ldots, a_n\} = \arg\max\{\sum_{k=1}^{n} |\langle f_k, e_{\{a_k'\}}\rangle|^2, a_1', a_2', \ldots, a_n' \in \Pi\}.$$

Remark 2.1 *As shown in the literature, the existence of the n-best rational approximation is well-known. The existence is equivalent to what is claimed in Theorem 2.1. A proof of the existence is cited in [61]. A practical algorithm of finding an n-best rational function, however, is a long standing open problem and thus is a hot topic of contemporary research [8, 10]. Below we provide a new proof of the existence. The reason why we give this new proof is that it proposes an algorithm to find the n-best approximation. When n is not so large, the algorithm is practical.*

Proof: To stress on the main idea we deal with the case $n = 2.$ The same proof is valid for general n (also see [50]).

We show that the global maximum can be attained at a 2-tuple $\{a_1, a_2\}$ in the open right-half complex plane. If not so, then we can find a sequence $\{a_1^l, a_2^l\}, \ l = 1, 2, \ldots,$ such that

$$A_f^{2,l} = |\langle f_1, e_{\{a_1^l\}}\rangle|^2 + |\langle f_2, e_{\{a_2^l\}}\rangle|^2$$

tends to the supreme of the quantity, while at least one of cases of $\Re\{a_1^l\}$ or $\Re\{a_2^l\}$ tends to $0,$ or one of $|a_1^l|$ or $|a_2^l|$ tends to $\infty,$ $A_f^{2,l}$ becomes small. Since $A_f^{2,l}$ is the norm of the projection of f on the span$\{a_1^l, a_2^l\},$ the order does not matter. Thus we may assume, without loss of the generality, that $\mathrm{Re}\{a_2^l\} \to 0$ or $|a_2^l| \to \infty,$ regardless the behavior of $\{a_1^l\}$. We show that in either of the two cases the limit can not exceed

$$\max_{a\in\Pi} |\langle f, e_{\{a\}}\rangle|^2, \tag{2.6}$$

that is a contradiction.

First consider the case $\Re\{a_2^l\} \to 0.$ Since $\langle f_k, e_{\{a_k^l\}}\rangle = \langle f, B_k\rangle,$ for the Poisson kernel $P_y,$ owing to the properties of the Poisson kernel, when y is sufficiently close to $0,$

$$\begin{aligned}
&\|f - \langle f_1, e_{\{a_1^l\}}\rangle e_{\{a_1^l\}} - \langle f_2, e_{\{a_2^l\}}\rangle e_{\{a_2^l\}}\| \\
&= \|f - \langle f, B_1^l\rangle B_1^l - \langle f - \langle f, B_1^l\rangle B_1^l, B_2^l\rangle B_2^l\| \\
&\geq \|P_y * (f - \langle f, B_1^l\rangle B_1^l - \langle f - \langle f, B_1^l\rangle B_1^l, B_2^l\rangle B_2^l)\| \\
&\geq \|P_y * (f - \langle f, B_1^l\rangle B_1^l)\| - \|P_y * \langle f - \langle f, B_1^l\rangle B_1^l, B_2^l\rangle B_2^l)\| \\
&\geq (1 - \epsilon)\|f - \langle f, B_1^l\rangle B_1^l\| - |\langle f - \langle f, B_1^l\rangle B_1^l, B_2^l\rangle|\|P_y * B_2^l\| \\
&\geq (1 - \epsilon)\|f - \langle f, B_1^l\rangle B_1^l\| - \|f - \langle f, B_1^l\rangle B_1^l\|\|P_y * B_2^l\|.
\end{aligned}$$

Now fix $y.$ Because B_2^l is in $H_2,$ we have

$$P_y * B_2^l(s) = B_2^l(s+y).$$

It follows

$$\begin{aligned}\|P_y * B_2^l\|^2 &\le \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{\Re\{a_2^l\}}{t^2+(\Re\{a_2^l\}+y)^2}dt\\ &= \frac{\Re\{a_2^l\}}{\Re\{a_2^l\}+y}.\end{aligned}$$

Now, if $\Re\{a_2^l\} \to 0,$ then

$$\|P_y * B_2^l\| \le \epsilon.$$

Therefore,

$$\begin{aligned}&\|f - \langle f_1, e_{\{a_1^l\}}\rangle e_{\{a_1^l\}} - \langle f_2, e_{\{a_2^l\}}\rangle e_{\{a_2^l\}}\|\\ &\ge (1-\epsilon^2)\|f - \langle f_1, e_{\{a_1^l\}}\rangle e_{\{a_1^l\}}\|.\end{aligned}$$

Because of the orthogonality between $f - \langle f_1, e_{\{a_1^l\}}\rangle e_{\{a_1^l\}}$ and $\langle f_2, e_{\{a_2^l\}}\rangle e_{\{a_2^l\}},$ this shows that the selection of a_2^l has no effect, and the limit of $A_f^{2,l}$ is at most as given by (2.6). The above Poisson kernel argument is uniform for all a_1^l in the right half complex plane.

Now assume $|a_2^l| \to \infty,$ as $l \to \infty.$ In the case we show

$$|\langle f, B_2^l\rangle| \to 0. \tag{2.7}$$

Since f is square integrable, we choose a large number A so that

$$\left|\int_A^{\infty} f(it)\overline{B_2^l}(it)dt\right| + \left|\int_{-\infty}^{-A} f(it)\overline{B_2^l}(it)dt\right| \le \epsilon.$$

Now with a constant C, the Hölder inequality implies that

$$\left|\int_{-A}^{A} f(it)\overline{B_2^l}(it)dt\right| \le C\left(\int_{-\infty}^{\infty}|f(it)|^2dt\right)^{1/2}\left(\int_{\frac{-A+\Im\{a_2^l\}}{\Re\{a_2^l\}}}^{\frac{A+\Im\{a_2^l\}}{\Re\{a_2^l\}}}\frac{1}{t^2+1}dt\right)^{1/2} \tag{2.8}$$

The last integral is over an integral of length $2A/\Re\{a_2^l\}.$ We consider two possible cases here. One is that the set $\Re\{a_2^l\}$ is unbounded for $l = 1, 2, \ldots.$ In the case we select a subsequence of $\{a_2^l\}_{l=1}^{\infty}$ with $\lim \Re\{a_2^l\} = \infty$ to replace the original $\{a_2^l\}_{l=1}^{\infty}$ with the same supreme effect, while

$$\lim_{l\to\infty} 2A/\Re\{a_2^l\} = 0.$$

Because of the absolute continuity of Lebesgue integration, for large enough $\Re\{a_2^l\}$ the left hand side of (2.8) is less than $\epsilon.$ The second case is that the set for $\Re\{a_2^l\}$ is bounded. In the case $|\Im\{a_2^l\}|$ is unbounded, and, in fact, tends to ∞ as $l \to \infty.$

In the case for large enough l the integral interval $(\frac{-A+\Im\{a_2^l\}}{\Re\{a_2^l\}}, \frac{A+\Im\{a_2^l\}}{\Re\{a_2^l\}})$ shifts to either $+\infty$ or $-\infty$. In either cases the integral becomes small. To summarize, the last integral tends to zero along with $l \to \infty$ for the fixed A, if necessary for a subsequence of $\{a_2^l\}_{l=1}^{\infty}$. Thus the limit (2.7) is proved. The above Poisson kernel argument is uniform for all a_1^l in the right-half complex plane. The proof is complete.

For the convergence rate in the energy sense, from the greedy algorithm point of view, we give a modest result of the remainder $R_n(s)$ if treating the set $D = \{e_a, a \in \Pi\}$ as a dictionary of greedy algorithm.

Define $H_2(D, A)$ as, $0 < A < \infty$,

$$H_2(D, A) = \{f \in H_2(\Pi) | f = \sum_{k=1}^{\infty} d_k e_k, e_k \in D, \sum_{k=1}^{\infty} |d_k| < A.\}$$

If $f(s) \in H_2(D, A)$, then we have

$$\begin{aligned}
\|R_n(s)\|^2 &= \|f - \sum_{k=1}^{n} \langle f_k, e_{\{a_k\}} \rangle B_k\|^2 \\
&= \|f\| - \sum_{k=1}^{n} |\langle f_k, e_{\{a_k\}} \rangle|^2 \\
&= \inf_{a_1^0, \ldots, a_n^0 \in \Pi} (\|f\| - \sum_{k=1}^{n} |\langle f_k, e_{\{a_k^0\}} \rangle|^2) \\
&\le \frac{A^2}{n}.
\end{aligned}$$

The estimation in the last step above is a result for greedy algorithm given in [14].

For the pointwise convergence, there is a result cited below. Suppose that $f(s) \in H_2(\Pi)$ has the following form

$$f(s) = \sum_{k=1}^{m} \frac{\alpha_k}{s + \beta_k}, \tag{2.9}$$

and $\{a_1, a_2, \ldots, a_n\}$ is obtained from the algorithm, then for the *n-approximating partial sum* $\widetilde{f}_n(s)$ (2.4), in [1] it shows that

$$|f(j\omega) - \widetilde{f}_n(j\omega)| \le \sum_{k=1}^{m} |\frac{\alpha_k}{j\omega + \beta_k}| \prod_{l=1}^{n} |\frac{\overline{\beta}_k - a_l}{\beta_k + \overline{a}_l}|. \tag{2.10}$$

The above result indicates that the approximation error decreases at least exponentially with increasing n, thus this method is pointwise convergent for the rational functions whose poles are on the left-hand plane.

3 Algorithm for Model reduction

In the algorithm introduced in the above section, any nth order approximating partial sum (2.4) is a rational function $\frac{P(s)}{Q(s)}$, where the degree of $P(s)$ is less than n and the degree of $Q(s)$ is equal to n, which means this algorithm automatically ensures the reduced system to be in the Hardy space H_2. This guarantees the stability of the approximating partial sums. But there is a problem that the coefficients of the partial sums may not be real. In [35], the conjugate consecutive optimal sequence is used in order to construct an approximating partial sums with real coefficients, but it seems not effective here as it also makes the order of the approximating functions higher.

So now, on one hand, the achievement of global maximum of (2.5) is expected; on the other hand, the coefficients of the reduced order models must be real. By considering both sides, if the original system has only real poles, we can just select $\{a_k\}_{k=1}^{n}$ in the right real line R^+, that may lead to a little loss of the precision of approximation but avoiding the oscillating reduced model; if the original system has complex poles, according to Lemma 2.1, conjugate complex poles can be drawn into the selection of $\{a_k\}$ to make the approximating partial sum (2.4) of real coefficients. In the rest of this section, we will use the real pole to clarify our idea, the method of incorporating conjugated pairs complex poles is the same and the number of parameters needed to be found is the same in R^+ as the real poles case.

In model order reduction, it is often important that the steady-state value of the step response of the reduced order models are kept to be equal to the original systems. An original system $G(s) \in H_2(\Pi)$, in case of single input and single output, is given by

$$G(s) = \sum_{i=0}^{m-1} c_i s^{ii}$$

Together with the relation, as a is real,

$$\langle f, e_{\{a\}}\rangle = \sqrt{2a}\, f(a),$$

the target function (2.5) can be further written as

$$A_f^n = 2a_1|f_1(a_1)|^2 + 2a_2|f_2(a_2)|^2 + \cdots + 2a_n|f_n(a_n)|^2. \tag{3.1}$$

Then the algorithm is formulated as follows.

Algorithm: Step 1. Writing out the forms of partial sum $\widetilde{f_n}(s)$ and A_f^n through the recursive formula

$$f_k(s) = \left(f_{k-1}(s) - \frac{2\Re\{a_{k-1}\} f_{k-1}(a_{k-1})}{s+\overline{a_{k-1}}} \right) \frac{s+\overline{a_{k-1}}}{s-a_{k-1}}. \tag{3.2}$$

Find out $\frac{\widetilde{c}_0}{\widetilde{b}_0}$ and it is clear that $\widetilde{f_n}(s)$ is completely determined by the sequence $\{a_1, a_2, \ldots, a_n\}$.

Step 2. Finding out a sequence $\{a_1, a_2, \ldots, a_n\}$ in the right real line $\mathbb{R}^+$ such that:

$$\begin{cases} \{a_1, a_2, \ldots, a_n\} = \arg\max\{A_f^n, a_k > 0\} \\ \dfrac{\widetilde{c}_0}{\widetilde{b}_0} = \dfrac{c_0}{b_0}. \end{cases} \tag{3.3}$$

Now in step 2, it turns to be a restrained nonlinear extremum problem, and the target function is smooth, thus it can be treated by the numerical methods. Take $n = 2$ for example, according to $f_1(s) = G(s)$ and the recursive formula

$$f_2(s) = \left(f(s) - \langle f_1, e_{\{a_1\}}\rangle e_{\{a_1\}}(s)\right) \frac{s+\overline{a}_1}{s-a_1},$$

the approximating partial sum (2.4) is written by

$$\begin{aligned} \widetilde{G}(s) &= \langle f_1, e_{\{a_1\}}\rangle B_1(s) + \langle f_2, e_{\{a_2\}}\rangle B_2(s) \\ &= \frac{(2\Re\{a_1\} f_1(a_1) + 2\Re\{a_2\} f_2(a_2))\, s}{s^2 + (\overline{a_1} + \overline{a}_2)s + \overline{a_1}\overline{a_2}} + \frac{2\Re\{a_1\} f_1(a_1) a_2 - 2\Re\{a_2\} f_2(a_2) a_1}{s^2 + (\overline{a_1} + \overline{a_2})s + \overline{a_1}\overline{a_2}}. \end{aligned}$$

When a_1, a_2 are real numbers,

$$\widetilde{G}(s) = \frac{(2a_1 f_1(a_1) + 2a_2 f_2(a_2))\, s}{s^2 + (a_1 + a_2)s + a_1 a_2} + \frac{2a_1 f_1(a_1) a_2 - 2a_2 f_2(a_2) a_1}{s^2 + (a_1 + a_2)s + a_1 a_2},$$

and then the constraint for a_1, a_2 reads

$$\frac{2a_1 f_1(a_1) a_2 - 2a_2 f_2(a_2) a_1}{a_1 a_2} = \frac{c_0}{b_0}.$$

The above constraint defines a manifold in the product space $\mathbb{R}^+ \times \mathbb{R}^+$.

Remark 3.1 *Theoretically the stable model generated by our proposed method may not be unique. In the case there are multi-solutions, then according to (2.10), one should choose the poles to be as close to those of the given transfer function as possible.*

4 Numerical examples

There are three numerical examples chosen from the literature in this section. Let $G(s) \in H_2(\Pi)$ be a given high order transfer function, then the inverse Laplace transform is defined by

$$g(t) = \frac{1}{2\pi j} \int_{\alpha - j\infty}^{\alpha + j\infty} e^{st} G(s) ds;$$

where α is larger than the largest real part of the poles of $G(s)$, $g(t)$ gives the impulse response of the system. The impulse response energy (IRE) [32, 53] is defined by

$$IRE = \int_0^{\infty} g^2(t) dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} |G(j\omega)|^2 d\omega;$$

and it is calculated and compared for both the original and the reduced-order models in the examples.

Just for illustrating the utilize of our method, we will only consider the second order model $G_2(s)$,

$$\widetilde{G}_2(s) = \frac{\widetilde{c}_1 s + \widetilde{c}_0}{s^2 + \widetilde{b}_1 s + \widetilde{b}_0}$$

for comparison of the chosen examples. The results obtained using various reduction methods are compared in the tables below.

In our programm, we use the algorithm *fmincon* in matlab to solve the constrained extreme problem (3.3). Since the algorithm *fmincon* can not guarantee the global optimal property in general, it is repeated with different initial points and then an optimal solution $\{a_1; a_2\}$ among the results is chosen.

Example 1 First we consider a 10th-order system previously studied in [38, 56, 45], where $G_{10}(s)$ is given by

$$G_{10}(s) = \frac{540.70748 \times 10^{17}}{\prod_{l=1}^{10}(s + b_l)}; \tag{4.1}$$

and $b_1 = 2.04; b_2 = 18.3; b_3 = 50.13; b_4 = 95.15; b_5 = 148.85; b_6 = 205.16; b_7 = 257.21; b_8 = 298.03; b_9 = 320.97; b_{10} = 404.16$. The impulse response energy (IRE) of the above system is 0.90305.

By using the proposed method, we obtain $a_1 = 2.1608$ and $a_2 = 9.7031$. Then the 2nd partial sum $\widetilde{G}_2(s)$ is

$$\widetilde{G}_2(s) = \frac{-0.5367s + 20.96}{s^2 + 11.86s + 20.97}; \tag{4.2}$$

with IRE=0.8956. The step responses of the original and reduced models are shown in figure 1, and a comparison on the impulse response energy (IRE) with the other methods is given in table 1. It can be seen that the IRE of the reduced model of our method is close to the original system, and its asymptotic behavior of the step response is better than others.

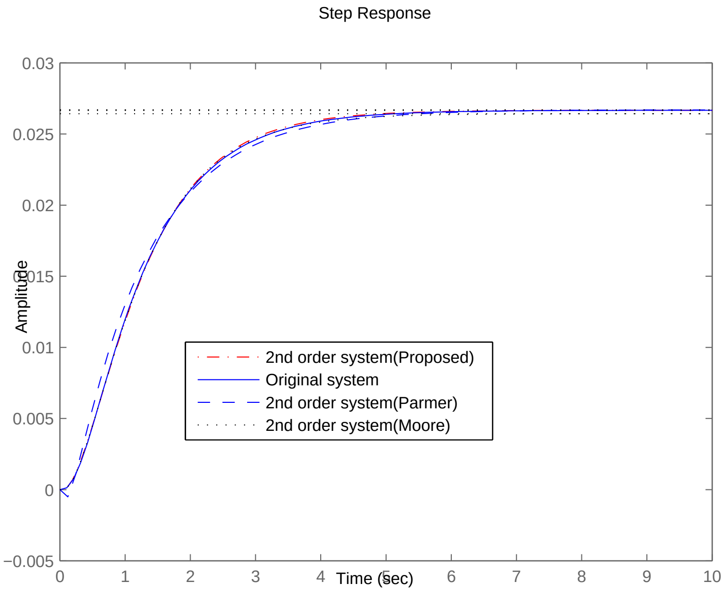

Figure 2: Step responses of the original and reduced models for example 2.

Table 2: Comparison of the reduced model for example 2.

Model reduction method	Reduced model	IRE
Original system	$G_4(s)$	2.6938×10^{-4}
Proposed method	$\frac{-0.00299s+0.072}{s^2+3.586s+2.7}$	2.6894×10^{-4}
Moore[37]	$\frac{-0.003127s+0.072358}{s^2+3.573s+2.73798}$	2.6896×10^{-4}
G. Parmar *et al.*[45]	$\frac{-0.02187s+0.19915}{s^2+9.5s+7.45868}$	3.0503×10^{-4}
Pal[44]	$\frac{-0.004855s+0.0818975}{s^2+4.066347s+3.071157}$	2.7144×10^{-4}

a comparison with the other methods on the impulse response energy (IRE) for this example is given in table 2. It can be seen that the values of IRE from our method is comparable to other methods, and the behavior of the step response is better than others. In figure 2, the asymptotic line of Parmar *et al.* [45] is different with the original system.

Example 3 The third example is an 8th-order system investigated in [32, 45, 47,

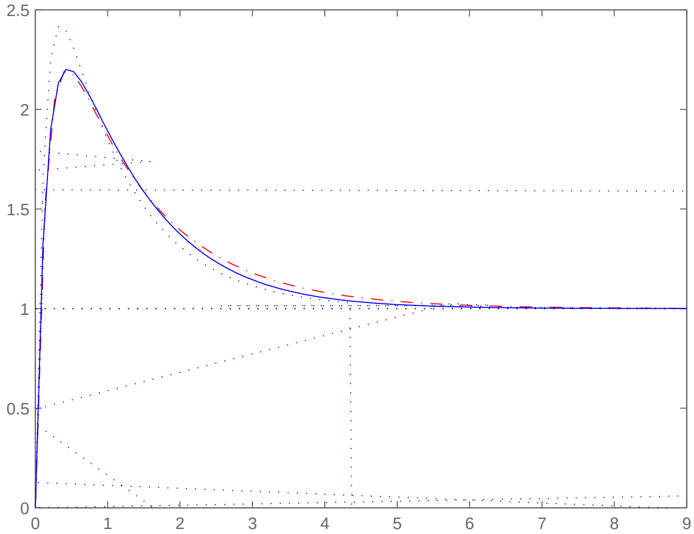

Figure 3: Step responses of the original and reduced models for example 3.

53]:

$$G_8(s) = \frac{\sum_{i=0}^{7} c_i s^i}{\sum_{i=0}^{8} b_i s^i}; \tag{4.5}$$

where $c_7 = 18; c_6 = 514; c_5 = 5982; c_4 = 36380; c_3 = 122664; c_2 = 222088; c_1 = 185760; c_0 = 40320$ and $b_8 = 1; b_7 = 36; b_6 = 546; b_5 = 4536; b_4 = 22449; b_3 = 67284; b_2 = 118124; b_1 = 109584; b_0 = 40320;$ with IRE$=21{:}739$. Using the proposed method, we obtain that $a_1 = 0{:}7910$ and $a_2 = 6{:}6071$. The second-order approximation is

$$\widetilde{G}_2(s) = \frac{17{:}7857s + 5{:}2264}{s^2 + 7{:}3981s + 5{:}2264}; \tag{4.6}$$

with the IRE$=21{:}7322$.

The step responses of the original and the reduced order models are shown in figure 3, and a comparison with the other methods on the impulse response energy (IRE) is presented in table 3. It can be seen that the proposed method gives not only a comparable step response to that of iterative rational Krylov algorithm (IRKA) [28, 30], but also a closer IRE value to the original system than the other methods.

Table 3: Comparison of the reduced model for example 3.

Model reduction method	Reduced model	IRE
Original system	$G_8(s)$	21.739
Proposed method	$\frac{17:7857s+5:2264}{s^2+7:3981s+5:2264}$	21.7322
G. Parmer *et al.*[45]	$\frac{24:1144s+8}{s^2+9s+8}$	32.75
Mukherjee *et al.*[40]	$\frac{11:39:9s+4:4357}{s^2+4:2122s+4:4357}$	15.9285
Prasad and Pal[47]	$\frac{17:98561s+500}{s^2+13:24571s+500}$	31.0849
Hutton and Fiedland[32]	$\frac{1:98955s+0:43184}{s^2+1:17368s+0:43184}$	1.8702
Shamash[53]	$\frac{6:7786s+2}{s^2+3s+2}$	7.9916
Bai[6]	$\frac{15:0990s+4:82}{s^2+5:9927s+4:82}$	19.4238
S. Gugercin *et al.*[28, 30]	$\frac{17:4119s+4:8188}{s^2+7:1315s+5:0129}$	21.5807

5 Conclusion

A fast adaptive method is proposed for model reduction of continuous systems in this paper. This method is based on the continuous-time Takenaka-Malmquist basis and aims to obtain the best approximation in the given order in the energy sense through selecting $fa_k g_{k=1}^n$ simultaneously for the Takenaka-Malmquist basis with a newly deduced recursive formula. It is adaptive for different systems. It can be seen from the tables and the figures, on considering both IRE and step response, the proposed method is considerable. Based on the proved theory and the promising experimental results, we can see the theory and the current numerical algorithm is useful in model reduction for the systems with low order. For the systems with large order, the theory has no constraint, we believe that our method is also useful under a suitable numerical realization to the second step of the algorithm.

References

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

[2] N.I Akhiezer, *The classical moment problem*, Oliver and Boyd, Edinburgh, 1969. Originally published Moscow, 1961.

[3] A.C. Antoulas, *Approximation of Large-scale Dynamical Systems*, Adv. Des. Control 6, SIAM, Philadelphia, 2005.

[4] A.C. Antoulas and D.C. Sorensen, Projection methods for balanced model reduction. *Technical Report ECE-CAAM Depts, Rice University,* September, 1999.

[5] A.C. Antoulas, D.C. sorensen and S. Gugercin, A survey of model reduction methods for large scales systems, *Contemp. math.,* AMS Publications, 280(2001),193-219.

[6] Z. Bai, Krylov sbspace techniques for reduced order modeling of large-scale dynamical systems, *Appl. Numer. Math.*, 43(2002), 9-44.

[7] L. Baratchart, *Sur l'approximation rationnelle L^2 pour les systèmes dynamiques linéaires,* Thèse de doctorat, Université de Nice, 1987.

[8] L. Baratchart, M. Olivi and F. Wielonsky, Asymptotic properties in rational l^2-approximation. *Lecture Notes in Control and Information Sciences*, 144(1990), 477-486.

[9] L. Baratchart, M. Cardelli and M. Olivi, Identification and rational L^2 approximation: a gradient algorithm, *Automatica*, no.2, 27(1991), 413-418.

[10] L. Baratchart, M. Olivi and F. Wielonsky, On a rational approximation problem in the real Hardy space H_2, *Theor. Comput. Sci.*, 94(1992), 175–197.

[11] A. Bultheel, P. González-Vera, E. Hendriksen, and O. Njåstad, *Orthogonal rational functions*, volume 5 of *Cambridge Monographs on Applied and Computational mathematics.* Cambrige University Press, 1999.

[12] A. Bultheel, P. Van gucht, and M. Van Barel, Orthogonal rational functions with real coefficients and semiseparable matrices, *J. Comput. Appl. Math.*, issue 5, 233(2010), 1192-1201.

[13] M. Casini, A. Garulli and A. Vicino, On worst-case approximation of feasible system sets via orthonomal basis functions, *IEEE Trans. Automat. Control,* no.1, 48(2003), 96-101.

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

[15] P. Dweilde, H. Dym, Lossless chain scattering matrices and optimum linear prediction: the vector case, *Internat. J. Circuit Th. Appl.*, 9(1981), 135-175.

[16] P. Dweilde, H. Dym, Schur recursions, erro formulas, and convergence of rational estimators for stationary stochastic sequences, *IEEE Trans. Inf. Th.*, IT-27(1981), 446-461.

[17] P. Dweilde, H. Dym, Lossless inverse scatting, digital filters, and estimation theory, *IEEE Trans. Inf. Th.*, IT-30(1984), 644-662.

[18] P. Dweilde, A. Viera and T. Kailath, On a generalized Szegő-Levinson realization algorithm for optimal linear predictors based on a network synthesis approach, *IEEE Trans. Circuits and Systems*, CAS-25(1978), 663-675.

[19] H. Dym, *J-contractive matrix functions, reproducing kernel Hilbert spaces and interpolation,* Volume 71 of *CBMS Regional Conf. Ser. in Math.* Amer. Math. Soc., Providence, Rhode Island, 1989.

[20] T.F. Edgar, Least squares model reduction using step response, *Int. J. Control*, 22(1975), 261-270.

[21] B. Epstein, *Orthogonal Families of Analytic Functions*, Macmillan, New York, 1965.

[22] Roland W. Freund, Padé-Type model reduction of second-order and higher-order linear dynamical systems, *Lecture Notes in Computational Science and Engineering*, 45(2005), 191-223.

[23] K. Gallivan, G. Grimme and P. Dooren, A rational Lanczos algorithm for model reduction, *Numer. Algorithms*, 12(1996), No. 1, 33-63.

[24] J.C. Geromel, R.G. Egas and F.R.R. Kawaoka, H_∞ model reduction with application to Flexible systems, *IEEE Trans. Automat. Control*, No.3, 50(2005), 402-406.

[25] J.C. Geromel, F.R.R. Kawaoka and R.G. Egas, Model reduction of discrete time systems through linear matrix inequalities, *Int. J. Control*, no.10, 77(2004), 978-984.

[26] E. Grimme, K. Gallivan, A rational Lanczos algorithm for model reduction II: Interpolation point selection, *Numer. Algorithms,* 12(1998), 33-63.

[27] S. Gugercin, An iterative SVD-Krylov based method for model reduction of large-scaled dynamical systems, *Linear Alg. Appl.*, 428(2008), issues 8-9, 1964-1986.

[28] S. Gugercin, A.C. Antoulas and C.A. Beattie, A rational Krylov iteration for optimal H_2 Model reduction, *Proceedings of the 17th International Symposium on Mathematical Theory of Networks and Systems,* kyoto, Japan, July 2006, 1665-1667.

[29] S. Gugercin, D.C. Sorensen and A.C. Antoulas, A modified low-rank Smith method for large-scale Lyapunov equations, *Numer. Algorithms,* 32(2003), issue 1, 27-55.

[30] S. Gugercin, A.C. Antoulas and C.A. Beattie, H_2 model reduction for large-scale linear dynamical systems, *SIAM J. Matrix Anal. Appl.*, 3(2008), issue.2, 609-638.

[31] Peter S.C. Heuberger, Paul M.J. Van den Hof and Bo Wahlberg, *Modelling and Identi¯cation with Rational Orthogonal Basis Functions*, 2005, Springer Verlag, ISBN:1-85233-956-x.

[32] M.F. Hutton and B. Fridland, Routh approximations for reducing order of linear time invariant systems, *IEEE Trans. Automat. Control*, AC-20(1975), 329-337.

[33] R. Malti, S.B. Ekongolo and J. Ragot. Dynamic SISO and MISO system approximation based on optimal Laguerre Models, *IEEE Trans. Automat. Control*, no.9, 43(1998), 1317-1322.

[34] J. Mendel, A unified approach to the synthesis of orthonormal exponential functions useful in systems analysis, *IEEE Trans. Systems Sci. Cybernet*. 2(1966), 54-62.

[35] Wen Mi and Tao Qian, Frequency domain identification with adaptive rational orthogonal system , *Proceedings of IEEE International Conference on System Science and Engineering (ICSSE)*, 2010, 41-45.

[36] A.K. Mittla, R. Prasad and S.P. Sharma, Reduction of linear dynamic systems using an error minimization technique, *J. Inst. Engi.(India)*, 84(2004), 201-206.

[37] B. Moore, Principal component analysis in linear systems:controllibility, observability and system reudction, *IEEE Trans. Automat. Control*, no.1, AC-26(1981), 17-32.

[38] S. Mukherjee, Order reduction of linear system using eigen spectrum analysis, *J. Inst. Engi.(India)*, part El, 77(1996), 76-79.

[39] S. Mukherjee and R.N Mishra, Order reduction of linear systems using an error minimization technique, *J. Frankl. Inst.*, 323(1987), 23-32.

[40] S. Mukherjee, Satakshi and R.C. Mittal, Model order reduction using response matching technique, *J. Frankl. Inst.*, 342(2005), 503-519.

[41] T. Oliveria e Silval, Optimality conditions for Laguerre lattice filters, *IEEE Signal Process. Lett.,* no.5, 2(1995), 97-98.

[42] T. Oliveria e Silva1, Optimality conditions for truncated Kautz networks with two periodically repeating complex conjugated poles, *IEEE Trans. Automat. Control,* no.2, 40(1995), 342-346.

[43] T. Oliveria e Silva, Stationarity conditions for L_2 error surface of generalized orthonormal basis functions lattice filter, *Signal Processing*, no.3, 56(1997), 233-253.

[44] J. Pal, An algorithmic method for the simplification of linear dynamic scalar system, *Int. J. Control*, 43(1986), 257-269.

[45] G. Parmar, S Mukherjee and R. Prasad, System reduction using eigen spectrum analysis and Padé approximation technique, *Int. J. Comput. Math.*, 84(2007), 261-274.

[46] T. Penzl, Algorithms for model reduction of large dynamical systems, *Linear Alg. Appl.*, 415(2006), 322-343.

[47] R. Prasad and J. Pal, Stable reduction of linear systems by continued fractions, *J. Insi. Engi.(India)*, Part El, 72(1991), 113-116.

[48] T. Qian, Intrinsic mono-component decomposition of functions: An advance of Fourier theory, *Math. Meth. Appl. Sci.*, 10.1002/mma.1214(DOI).

[49] T. Qian, Y.B. Wang, Adaptive fourier series - A variation of greedy algorithm, *Adv. Comput. Math.*, 34(2011), 279-293.

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

[51] A.M. Sabatini, A hybrid genetic algorithm for estimating the optimal time scale of linear systems approximations using Laguerre models, *IEEE Trans. Automat. Control,* no.5, 45(2000), 1007-1011.

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

[53] Y. Shamash, Linear system reduction using Padé approximation to allow retention of dominant modes, *Int. J. Control*, 21(1975), 257-272.

[54] V. Singh, D. Chandra and H. Kar, Improved Routh-Páde approximatants: A computer-aided approach, *IEEE Trans. Automat. Control*, no.2, 49(2004), 292-296.

[55] S. Takenaka, On the orthogonal functions and a new formula of interpolation, Jpn. J. Math. II (1925) 129-145.

[56] C.P. Therapos and J.E. Diamessis, A new method for linear system reduction, *J. Frankl. Inst.*, 316(1983), 87-92.

[57] J. Van Deun, A. Bultheel, An interpolation algorithm for orthogonal rational functions, *J. Comput. Appl. Math.,* 164-165(2004), 749-762.

[58] P. Van gucht, A, Bultheel, Using orthogonal rational functions for system identification: numerical aspects, *IEEE Trans. Automat. Control*, no.4, 48(2003), 705-709.

[59] C.De Villemagne, R. Skelton, Model reduction using a projection formulation, *Internat. J. Control.*, 40(1987), 2141-2169.

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

[61] J.L Walsh, *Interpolation and approximation by rational functions in the complex domain*, American Mathematical Society, Fourth edition, 1965.