Proceedings of the 33rd
Conference on Decision and Control
Lake Buena Vista, FL - December 1994

Rational Wavelets in Model Reduction and System Identification

Y. C. PATI*
Information Systems Laboratory
Department of Electrical Engineering
Stanford University, Stanford, CA 94305

P. S. KRISHNAPRASAD†
Institute for Systems Research
Department of Electrical Engineering
University of Maryland, College Park, MD 20742

Abstract

In this paper we describe the application of rational wavelet decompositions of the Hardy space $H^2(\Pi^+)$ to rational approximation of transfer functions of stable, linear, time-invariant systems. The intent of the current paper is to provide an account of earlier work [Pati and Krishnaprasad 1992] in this area and to describe some recent developments in efficient computational methods for rational wavelet approximation. It is shown by means of several examples that rational wavelets can provide 'good' approximations to some important classes of linear systems.

1 Introduction

Recently there has been a resurgence of interest in rational approximation of transfer functions using basis or basis-like decompositions, for purposes of system identification and model-reduction. The basic approach followed in these methods is to represent a transfer function G, by a series of the form

$$G(s) = \sum_k a_k(G)\Phi_k(s), \qquad (1)$$

where the Φ_k are fixed 'basis' functions (usually rational functions for finite k), and then construct rational approximations to G by taking finite truncations of the series. The two main advantages of these methods are (i) the model is linear in the parameters a_k, and (ii) it is often possible to incorporate various forms of *a priori* knowledge into the approximation problem through suitable choice of the Φ_k (see [4, 10]). The goal of rational approximation in these problems is two-fold: (i) to accurately capture the relevant behavior of the underlying dynamical system, and (ii) to keep the complexity (order) of the model as low as possible while meeting the first requirement. Some examples of basis functions that have been considered in such applications are the Laguerre bases (see *e.g.* [10]) and Kautz filters [9] and more recently, rational wavelet bases, [4, 5] (see also [11]).

In this paper we discuss rational approximation using the rational wavelet 'bases' (in $H^2(\Pi^+)$) developed in [4, 5]. Ward and Partington [11] have recently extended the idea of rational wavelets to the approximation of transfer functions (of discrete-time systems) in certain Hardy-Sobolev classes, and have shown that this approach can provide good approximations with respect to the H^∞ metric in addition to the H^2 metric. An important distinction between rational wavelets, and Laguerre and Kautz models, is that while the poles of the Laguerre and Kautz basis functions are restricted to some (usually very small) finite set, the poles of rational wavelets are distributed on an infinite lattice. In particular, the lattice of rational wavelet pole locations is generated by charting the poles of translates and dilates of a single rational *analyzing wavelet*; this provides a set of time-frequency localized rational 'building blocks' that facilitates the incorporation of both time and frequency domain *a priori* knowledge into parametric black-box models for system identification.

The rational wavelets constructed in [4, 5] do not form *bases* in a strict sense, but rather a class of generalized bases, called *frames*, in the Hardy space $H^2(\Pi^+)$. As representations of the form (1), where the Φ_k comprise a frame, are not in general unique, it is necessary to examine the problem of computing a suitable representation. We describe a recently developed

*This research was supported in part by NASA Headquarters, Center for Aeronautics and Space Information Sciences (CASIS) under Grant NAGW419,S6, and in part by the Advanced Research Projects Agency of the Department of Defense monitored by the Air Force Office of Scientific Research under Contract F49620-93-1-0085.

†This research was supported in part by the Air Force Office of Scientific Research under contract F49620-92-J-0500, the AFOSR University Research Initiative Program under Grant AFOSR-90-0105, by the Army Research Office under Smart Structures URI Contract no. DAAL03-92-G-0121, and by the National Science Foundation's Engineering Research Centers Program, NSFD CDR 8803012.

0-7803-1968-0/94$4.00©1994 IEEE

[6], efficient computational technique to construct rational wavelet approximations. The algorithm we describe constructs the "best" one-term approximation of the error at each step and is a generalization of the Matching Pursuit (MP) algorithm [3]. The generalization is necessary to accommodate the real-rationality constraint on the approximants.

Simulation and experimental results (see also [8]) are presented to illustrate the methods and ideas described here and a comparison is made with approximation results using Laguerre bases.

2 Background

For completeness we first review some basic definitions pertaining to Hardy spaces and wavelets. The reader is referred to [1] for an excellent introduction to wavelet transforms and to [2] for details on Hardy spaces.

2.1 The Hardy Space $H^2(\Pi^+)$

Here we consider the class of transfer functions contained in the Hardy space $H^2(\Pi^+)$, where Π^+ denotes the half-plane $\Re e\ (s) > 0$.

Definition 1 *Given a function F that is analytic in Π^+, F is said to belong to $H^2(\Pi^+)$ if*

$$\sup_{x>0} \int_{-\infty}^{\infty} |F(x+iy)|^2 \, dy < \infty. \qquad (2)$$

By the Paley-Wiener theorem, elements of $H^2(\Pi^+)$ may be identified with transfer functions of causal, input-output stable, linear time-invariant systems. The following notation is employed in the remainder of this paper:

$H^2_{\mathbb{R}}(\Pi^+)$	=	Laplace transforms of real-valued functions in $L^2(0,\infty)$
$RH^2(\Pi^+)$	=	real-rational functions in $H^2(\Pi^+)$, *i.e.* rational functions in $H^2(\Pi^+)$ with real coefficients.

Thus $RH^2(\Pi^+)$ ($\subset H^2_{\mathbb{R}}(\Pi^+)$) represents transfer functions of causal, *finite-dimensional*, linear systems with real-valued, square-integrable weighting patterns.

2.2 Frames and Wavelets

Our construction of rational wavelet decompositions of $H^2(\Pi^+)$ is based on a generalization of orthonormal bases in Hilbert spaces, which are referred to as *frames*.

Definition 2 *Given a Hilbert space $\mathcal{H}$ and a sequence of vectors $\{h_n\}_{n=-\infty}^{\infty} \subset \mathcal{H}$, $\{h_n\}_{n=-\infty}^{\infty}$ is called a frame if there exist constants (frame bounds) $A > 0$ and $B < \infty$ such that for every $f \in \mathcal{H}$,*

$$A\|f\|^2 \leq \sum_n | < f, h_n > |^2 \leq B\|f\|^2. \qquad (3)$$

A key property of frames is that the *frame operator* $\mathbf{S}$, defined by $\mathbf{S}f = \sum_n < f, h_n > h_n$, $f \in \mathcal{H}$, is invertible and therefore any $f \in \mathcal{H}$ may be represented in terms of the frame elements:

$$f = \sum < f, \mathbf{S}^{-1}h_n > h_n = \sum < f, h_n > \mathbf{S}^{-1}h_n. \qquad (4)$$

Note that in general frames can be overcomplete families of vectors, and that orthonormal bases are a special case of frames with linearly independent, normalized vectors and $A = B = 1$.

Affine, or *wavelet frames*, in the Hilbert space $L^2(\mathbb{R})$, are frames in $L^2(\mathbb{R})$ constructed from dilates and translates of a single function ψ (called the *analyzing wavelet* or *mother wavelet*) *i.e.* frames of the form

$$\left\{ \psi_{m,n}(x) = a_0^{m/2} \psi(a_0^m x - nb_0) \right\}$$

where $a_0 > 0$ and b_0 are fixed constants. For such constants a_0 and b_0 to exist[1] ψ, must also satisfy the admissibility condition $\int_{\mathbb{R}} \left|\widehat{\psi}(\omega)\right|^2 / |\omega| \, d\omega < \infty$, where $\widehat{\psi}$ is the Fourier transform of ψ. It is by now well-known that it is possible to construct wavelet frames (including orthonormal bases) for $L^2(\mathbb{R})$ from (analyzing) functions that are 'well-localized' in time-frequency. Time-frequency localization is an important property of wavelets that often allows functions exhibiting time-frequency localized behavior to be compactly represented by wavelet 'bases'.

3 Rational Wavelets in $H^2(\Pi^+)$

In [4, 5] it was shown that frames of rational wavelets in $H^2(\Pi^+)$ may be used to construct time-frequency localized rational approximations to transfer functions. Rational wavelets in $H^2(\Pi^+)$ are constructed by taking dilates and translates of an *admissible* real-rational analyzing wavelet Ψ ($\in RH^2(\Pi^+)$) to form an affine frame $\{\Psi_{m,n}\}$ in $H^2(\Pi^+)$, where

$$\Psi_{m,n}(s) = a_0^{m/2} \Psi(a_0^m s - inb_0), \quad a_0 > 0. \qquad (5)$$

[1] Numerically, a_0 and b_0 may be determined by application of a theorem of Daubechies (see *e.g.* [1]).

The following theorem summarizes some of the main results from [4] on rational wavelet decompositions of $H^2(\Pi^+)$.

Theorem 1 ([4]) *Let $\Psi \in RH^2(\Pi^+)$, satisfy the admissibility condition*[2]

$$\int_{\mathbb{R}} \left|\widehat{\Psi}_x(t)\right|^2 / |t|\, dt < \infty, \quad x > 0, \tag{6}$$

where Ψ_x denotes the restriction of Ψ the the vertical line $\Re e\ s = x$ in Π^+. Then,

(i) There exist $a_0 > 0$ and b_0 such that (Ψ, a_0, b_0) generates an affine frame $\{\Psi_{m,n}\}$ for $H^2(\Pi^+)$.

(ii) Any F in $H^2_{\mathbb{R}}(\Pi^+)$ may be represented as,

$$F = \sum_m \sum_{n=0}^{\infty} F^{m,n}, \tag{7}$$

where each $F^{m,n}$ is a real-rational function defined by,

$$\begin{aligned} F^{m,0} &= \langle F, \mathbf{S}^{-1}\Psi_{m,0}\rangle\, \Psi_{m,0} \\ F^{m,n} &= \langle F, \mathbf{S}^{-1}\Psi_{m,n}\rangle\, \Psi_{m,n} \\ &\quad + \overline{\langle F, \mathbf{S}^{-1}\Psi_{m,n}\rangle}\Psi_{m,-n} \end{aligned}$$

($m \in \mathbb{Z}$, $n \in \mathbb{Z}^+ / \{0\}$) and $\mathbf{S}$ is the frame operator associated with $\{\Psi_{m,n}\}$. ■

The series (7) is referred to as a *wavelet system* (WS) decomposition of $F \in H^2_{\mathbb{R}}(\Pi^+)$. Given a transfer function $F \in H^2_{\mathbb{R}}(\Pi^+)$, and its WS decomposition, a real-rational approximation to F may be constructed by taking suitable truncations of the WS series *i.e.* by selecting a *finite* index set $\mathcal{J}$ ($(m,n) \in \mathcal{J}$) to use in the approximation. An important fact to note is that in the case of WS decompositions, there is no predetermined 'natural' ordering of truncations; this is in contrast to the case of Lagurre or Kautz approximation where the ordering of the truncations is fixed. This additional flexibility is useful in incorporating *a priori* knowledge, but it is also creates the need to devise suitable methods of selecting 'optimal' truncations.

3.1 Time and Frequency Domain *A Priori* Knowledge

For system identification and model reduction, it is often desirable to incorporate *a priori* knowledge of time and frequency domain properties of the system into the approximation method being used. The *a priori* knowledge that may be available includes information about dominant time-constants, key frequency 'bands', and delays, *e.g.* in many widely-used system identification methods, frequency domain information is incorporated through *frequency weighting* schemes.

[2] It may easily be shown that any rational function in $H^2(\Pi^+)$, with relative degree greater than or equal to 2, satisfies the admissibility condition (6).

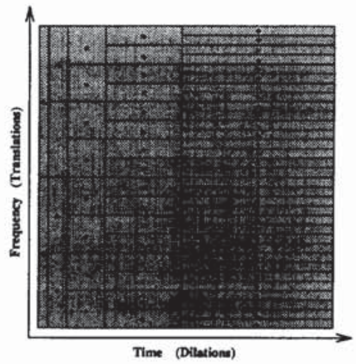

Figure 1: Time-Frequency tiling obtained through rational wavelet decompositions of $H^2(\Pi^+)$.

A useful property of rational wavelets is that the actions of translation and dilation generate a family of building blocks that are largely localized in time and frequency; this is schematically depicted in Fig. 1. An equivalent viewpoint is obtained by examining the distribution of poles of the rational wavelets (see Fig. 2). This time-frequency structure enables us to select subsets of translations and dilations that appropriately represent the *a priori* knowledge, and then use the chosen subsets to construct (rational) approximations of the transfer function of interest. Lagurre models (see [10]) enable the incorporation of only time domain knowledge, while Kautz models permit a more restricted use of both time and frequency domain information. A full discussion of incorporating time and frequency domain knowledge into rational wavelet models may be found in [4].

4 Computing WS Approximations

In [4, 5], the index set $\mathcal{J}$, was selected to include terms from the WS decomposition corresponding to the largest wavelet frame coefficients $\alpha_{m,n} = \langle F, \mathbf{S}^{-1}\Psi_{m,n}\rangle$. The drawbacks with this approach are that: (a) the frame expansion with respect to a large collection of vectors must first be computed, and (b) the frame expansion coefficients given by (4) do not necessarily generate the most parsimonious representation possible with respect to a particular collection of vectors.

The following iterative algorithm provides an efficient method to compute parsimonious rational wavelet approximations of transfer functions. The algorithm is a variant of the Matching Pursuit algorithm of Mallat and Zhang [3] and we refer to it as WRASP (WS Rational Approximation via Successive Projections) [6].

We first define the following orthogonal projection operator (in $H^2(\Pi^+)$):

$$\mathbf{P}_{m,n} = \begin{cases} \text{Proj. onto } \overline{\text{Span}\{\Psi_{m,n}, \Psi_{m,-n}\}} & \text{if } n \neq 0 \\ \text{Proj. onto } \overline{\text{Span}\{\Psi_{m,n}\}} & \text{if } n = 0 \end{cases}$$

(Note that for $n \neq 0$, $\mathbf{P}_{m,n}$ is a projection onto a two-dimensional subspace of $H^2(\Pi^+)$.) In the following, F_k and $\mathbf{R}_k F$ denote the approximation and the residual (error) respectively at the k^{th} WRASP iteration *i.e.* $F = F_k + \mathbf{R}_k F$.

The WRASP Algorithm

Initialization:

$$F_0 = 0, \ \mathbf{R}_0 F = F, \ k = 0$$

(I) Compute all projections

$$\mathbf{P}_{m,n}\mathbf{R}_k F, \quad (m,n) \in \mathcal{J} \subset \mathbb{Z} \times \mathbb{Z}^+,$$

where $\mathbb{Z}^+$ denotes nonnegative integers.

(II) Find the 'largest' projection *i.e.* find (m_k, n_k) such that,

$$\|\mathbf{P}_{m_k,n_k}\mathbf{R}_k F\| \geq \beta \sup_{m\in\mathbb{Z},\ n\in\mathbb{Z}^+} \|\mathbf{P}_{m,n}\mathbf{R}_k F\|,$$

where $0 < \beta \leq 1$.

(III) Update the model and residual:

$$\begin{aligned} F_{k+1} &= F_k + \mathbf{P}_{m_k,n_k}\mathbf{R}_k F, \\ \text{and } \mathbf{R}_{k+1}F &= \mathbf{R}_k F - \mathbf{P}_{m_k,n_k}\mathbf{R}_k F. \end{aligned}$$

(IV) Increment k (*i.e.* $k \leftarrow k+1$) and repeat Steps (I)–(IV), until some stopping criterion has been satisfied.

The main distinction between the above algorithm and the Matching Pursuit algorithm is in the definition of the projection operators.

As we are interested in real-rational approximants, we need to verify that the approximants F_k, at each step $k < \infty$ are real-rational and that they converge to F as $k \to \infty$. These properties are readily verified and are summarized in the following theorem.

Theorem 2 *Let $\{\Psi_{m,n}\}_{m,n\in\mathbb{Z}}$ be a rational wavelet frame for $H^2(\Pi^+)$, and let $\mathcal{I}$ be a subset of $\mathbb{Z}^2$, such that $(m,n) \in \mathcal{I} \Leftrightarrow (m,-n) \in \mathcal{I}$. Also let $\mathbf{V} = \overline{\text{Span}\{\Psi_{m,n}\}_{(m,n)\in\mathcal{I}}}$ and let $\mathbf{P}_V$ the denote orthogonal projection operator onto $\mathbf{V}$. Then for any $F \in H^2_{\mathbb{R}}(\Pi^+)$.*

(1) $\|\mathbf{R}_k F - \mathbf{P}_{V^\perp} F\| \longrightarrow 0 \quad$ *as* $k \to \infty$

(2) $\mathbf{P}_V F = \sum_{k=0}^{\infty} \mathbf{P}_{m_k,n_k}\mathbf{R}_k F.$

(3) $F_N = \sum_{k=0}^{N} \mathbf{P}_{m_k,n_k}\mathbf{R}_k F \in RH^2(\Pi^+), \quad N < \infty.$

(4) $\|\mathbf{P}_V F\|^2 = \sum_{k=0}^{\infty} \|\mathbf{P}_{m_k,n_k}\mathbf{R}_k F\|^2. \quad$ ■

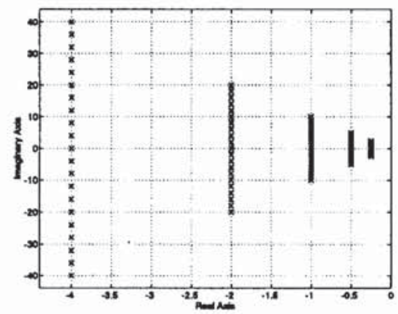

Figure 2: Distribution of poles of rational wavelets obtained by translating and dilating ($m \in \{-2,-1,0,1,2\}$, $n \in \{-10,\ldots,10\}$) a second-order rational analyzing wavelet.

Some Computational Aspects of WRASP

The WRASP algorithm proves to be a very computationally efficient procedure for constructing rational wavelet approximations. All of the required computations for WRASP may in fact be performed recursively. Here we point out a few of the factors that contribute to its efficiency.

First of all note that the inner products $\{\langle F, \Psi_{m,n}\rangle\}_{m,n}$ are required as a starting point for the algorithm. These inner products may be readily computed by convolution with $\widetilde{\Psi}_m = a_0^{m/2}\Psi(-a_0^m \cdot)$ followed by sampling at each dilation level m *i.e.*

$$\langle F, \Psi_{m,n}\rangle = \left(F * \widetilde{\Psi}_m\right)(a_0^{-m} b_0 n) \qquad (8)$$

The projections $\mathbf{P}_{m,n}F$ are also easily obtained using the following formula

$$\begin{aligned}\mathbf{P}_{m,n}F &\\ = \ & \alpha^2 \Big[\left(\langle F, \Psi_{m,n}\rangle - \overline{\langle F, \Psi_{m,n}\rangle}\,\langle \Psi_{m,-n}, \Psi_{m,n}\rangle\right)\Psi_{m,n} \\ + \ & \left(\overline{\langle F, \Psi_{m,n}\rangle} - \langle F, \Psi_{m,n}\rangle\,\overline{\langle \Psi_{m,-n}, \Psi_{m,n}\rangle}\right)\Psi_{m,-n}\Big],\end{aligned}$$

for $n \neq 0$, where, $\alpha^2 = (1 - |\langle\Psi_{m,-n}, \Psi_{m,n}\rangle|^2)^{-1}$.
For $n = 0$, $\mathbf{P}_{m,n}F = \langle F, \Psi_{m,n}\rangle\, \Psi_{m,n}$.

Finally, for $k > 0$, $\langle \mathbf{R}_{k+1}F, \Psi_{m,n}\rangle$ may be computed from $\langle \mathbf{R}_k F, \Psi_{m,n}\rangle$ using the recursion,

$$\begin{aligned}&\langle \mathbf{R}_{k+1}F, \Psi_{m,n}\rangle \\ &= \ \langle \mathbf{R}_k F, \Psi_{m,n}\rangle - c_k \left\langle \Psi_{m_{k+1},n_{k+1}}, \Psi_{m,n}\right\rangle \\ &- \ \overline{c_k}\left\langle \Psi_{m_{k+1},-n_{k+1}}, \Psi_{m,n}\right\rangle,\end{aligned}$$

where c_k is the coefficient obtained from the previous projection $\mathbf{P}_{m_{k+1},n_{k+1}}\mathbf{R}_k F$, *i.e.* $\mathbf{R}_k F = c_k\Psi_{m_{k+1},n_{k+1}} + \overline{c_k}\Psi_{m_{k+1},-n_{k+1}} + \mathbf{R}_{k+1}F$.
Elements of the Gram matrix $[\langle\Psi_{m,n}, \Psi_{j,k}\rangle]$ can either be precomputed and stored or computed online using convolution and sampling (*cf.* (8)).

Using the above observations in the implementation of WRASP reduces the algorithm to a sequence of efficient recursive computations.

Remark

A further refinement of the WRASP algorithm is also possible by ensuring that the residual $\mathbf{R}_k F$ at each iteration is orthogonal to *all* previously selected vectors for each k; this is analogous to the orthogonalization techniques applied to the Matching Pursuit algorithm in [7]. Orthogonalization leads to faster convergence at the expense of greater computation, but we do not discuss the details of such methods here.

5 Examples

In the following examples, the analyzing wavelet is taken to be,

$$\Psi(s) = \frac{1}{(s+25)^2+1},$$

which generates an affine frame for $\mathrm{H}^2(\Pi^+)$ with $a_0 = 2,\ 0 < b_0 < 16.5$.

Example I: In the first example we consider the problem of approximating an unparameterized model of a single channel of a cochlear filterbank (see [4]) by a finite-dimensional system. Figure 3 compares the approximation performance of the WRASP algorithm applied to this problem with two other rational approximation methods: (i) rational wavelet approximation using the largest coefficients of a 'full' decomposition, and (ii) rational approximation using the Laguerre basis.

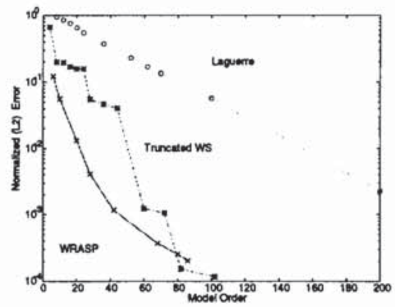

Figure 3: Example I (Cochlear Filter [5,6]): Normalized ($\mathrm{H}^2(\Pi^+)$) approximation error versus model-order: ($\times$) the WRASP algorithm, ($*$) rational wavelet approximation using the full decomposition, and ($\circ$) Laguerre basis approximation.

Example II: In this example the WRASP algorithm is used construct an approximate finite-dimensional model for a flexible beam apparatus equipped with piezoceramic sensors/actuators using measured data [8]. The experimental setup is shown in Figure 4. Figure 5 compares WRASP with Laguerre basis approximation and Figure 6 compares the identified WRASP model with the measured response.

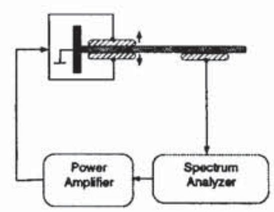

Figure 4: Experimental flexible beam setup [6]. The power amplifier drives a pair of piezoceramic actuators bonded to the beam and a spectrum analyzer measures the response through a piezoceramic sensor.

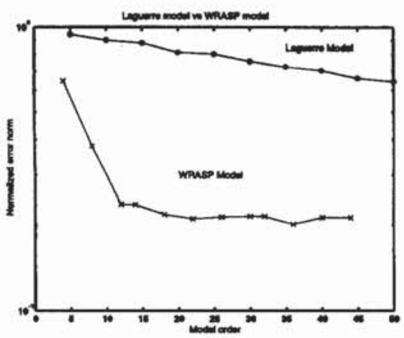

Figure 5: Example II (Flexible beam [6]): Normalized ($H^2(\Pi^+)$) approximation error versus model-order: ($\times$) the WRASP algorithm, and ($\circ$) Laguerre basis approximation.

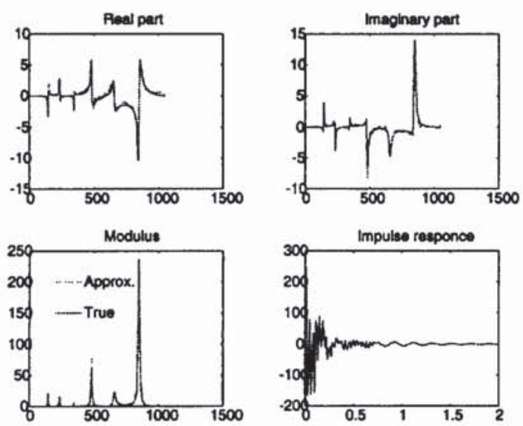

Figure 6: Example II (Flexible beam [6]): WRASP approximation results comparing measured and approximate models. Model order = 46.

6 Summary

In this paper we have shown that rational wavelet decompositions may be used to effectively capture time-frequency localized behavior of stable linear systems in a rational transfer function model. An efficient computational method to compute rational wavelet approximations was described and applied to example system identification problems.

References

[1] Ingrid Daubechies. *Ten Lectures on Wavelets.* CBMS-NSF regional conference series in applied mathematics. SIAM, Philadelphia, PA, 1992.

[2] Kenneth Hoffman. *Banach Spaces of Analytic Functions.* Dover Publications, New York, 1988.

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

[4] Y. C. Pati. *Wavelets and Time-Frequency Methods in Linear Systems and Neural Networks.* PhD thesis, University Of Maryland, College Park, MD, 1992.

[5] Y. C. Pati and P. S. Krishnaprasad. Rational wavelets in approximation and identification of stable linear systems. In *Proceedings 31st IEEE Conference on Decision and Control,* pages 1502–1507, N.Y., December 1992. IEEE.

[6] Y. C. Pati, R. Rezaiifar, W. P. Dayawansa, and P. S. Krishnaprasad. A fast recursive algorithm for system identification and model reduction using rational wavelets. In *Proceedings 27th Annual Asilomar Conference on Signals Systems and Computers,* pages 35–39, Pacific Grove, CA, November 1993.

[7] Y. C. Pati, R. Rezaiifar, and P. S. Krishnaprasad. Orthogonal matching pursuit: Recursive function approximation with application to wavelet decomposition. In *Proceedings 27th Annual Asilomar Conference on Signals Systems and Computers,* pages 40–44, Pacific Grove, CA, November 1993.

[8] R. Rezaiifar. Smart structures and wavelet based system identification. Master's thesis, University of Maryland, College Park, MD, 1993. Institute for Systems Research Technical Report No. MS 93-18.

[9] Bo Wahlberg. System identification using Kautz models. *IEEE Transactions on Automatic Control.* To appear.

[10] Bo Wahlberg. System identification using Laguerre models. *IEEE Transactions on Automatic Control,* 36(5):551–562, 1991.

[11] N. F. Dudley Ward and J. R. Partington. Rational wavelet decompositions of transfer functions in Hardy-Sobelev classes. Technical report, University of Leeds, Department of Pure Mathematics, 8, 1994.