



# The AFD methods to compute Hilbert transform



Yan Mo<sup>a,\*</sup>, Tao Qian<sup>a</sup>, Weixiong Mai<sup>a</sup>, Qiuhui Chen<sup>b</sup>

<sup>a</sup> Department of Mathematics, University of Macau, Macao

<sup>b</sup> Cisco School of Informatics, Guangdong University of Foreign Studies, China

## ARTICLE INFO

### Article history:

Received 29 October 2014

Received in revised form 30 December 2014

Accepted 30 December 2014

Available online 19 January 2015

### Keywords:

Hilbert transform

Hardy space

Takenaka–Malmquist system

Orthogonal rational system

Adaptive Fourier decomposition

## ABSTRACT

In the literature adaptive Fourier decomposition is abbreviated as AFD that addresses adaptive rational approximation, or alternatively adaptive Takenaka–Malmquist system approximation. The AFD type approximations may be characterized as adaptive approximations by linear combinations of parameterized Szegő and higher order Szegő kernels. This note proposes two kinds of such analytic approximations of which one is called maximal-energy AFDs, including core AFD, Unwinding AFD and Cyclic AFD; and the other is again linear combinations of Szegő kernels but generated through SVM methods. The proposed methods are based on the fact that the imaginary part of an analytic signal is the Hilbert transform of its real part. As consequence, when a sequence of rational analytic functions approximates an analytic signal, then the real parts and imaginary parts of the functions in the sequence approximate, respectively, the original real-valued signals and its Hilbert transform. The two approximations have the same errors in the energy sense due to the fact that Hilbert transformation is a unitary operator in the  $L^2$  space. This paper for the first time promotes the complex analytic method for computing Hilbert transforms. Experiments show that such computational methods are as effective as the commonly used one based on FFT.

© 2015 Elsevier Ltd. All rights reserved.

## 1. Preparation

Computation of Hilbert transform is a difficult task due to the singularity of the complex Cauchy kernel at the origin (see for instance [1,2], and the references thereafter). In this paper we introduce a set of AFD type analytic methods to compute Hilbert transform. As abbreviation of adaptive Fourier decomposition, AFD, in general, addresses all types of adaptive approximations in the form of finite linear combinations of Szegő and higher order Szegő kernels. In this paper we deal with two kinds of AFDs. One can be phrased as *Maximal-Energy-AFD* that, in its core algorithm part, is based on a maximal selection principle to adaptively and optimally select the relevant parameters. Such AFD method was initiated in [3] and further developed in [4–6]. The main variations of this kind include Core AFD, Unwinding AFD and Cyclic AFD. They have been found to have effective applications in system identification, as well as in signal analysis [7–9]. The algorithm of Core AFD is studied in [10]. The other kind of AFD concerned in this paper is a complex type SVM, phrased as SVM-AFD. In this paper we also call it “complex SVM” although it is not just by using complex numbers, but complex analytic functions, and especially Szegő kernels. The complex SVM is also found to have effective applications in system identification [11,12]. The relevant algorithm codes are found in the web address <http://www.fst.umac.mo/en/staff/fsttq.html>.

\* Corresponding author.

E-mail addresses: [marvel-2008@163.com](mailto:marvel-2008@163.com) (Y. Mo), [tqian1958@gmail.com](mailto:tqian1958@gmail.com) (T. Qian), [maiweixiong@gmail.com](mailto:maiweixiong@gmail.com) (W. Mai), [chenqiuhui@hotmail.com](mailto:chenqiuhui@hotmail.com) (Q. Chen).

The importance of Hilbert transformation lies in the facts that with

$$Cs(z) = \frac{1}{2\pi i} \int_{-\infty}^{\infty} \frac{s(t)}{t - z} dt,$$

the Cauchy integral of  $s$ , one has

$$\lim_{y \rightarrow 0^+} Cs(x + iy) = \frac{1}{2\pi i} \int_{-\infty}^{\infty} \frac{s(t)}{t - (x + iy)} dt = s^+(t),$$

where

$$s^+(t) = \frac{1}{2}s(t) + \frac{1}{2}iHs(t)$$

is the analytic signal associated with the given real-valued signal  $s$  of finite energy, and  $Hs$  is its Hilbert transform. This result is known as the Plemelj formula [13].

If  $s(t)$  is a given real-valued signal, then we have the easy relations

$$s = 2 \operatorname{Re} s^+, \quad Hs = 2 \operatorname{Im} s^+.$$

We note that the analytic signals form a closed subspace, denoted  $L^2(\mathbf{R})$  and called the Hardy space  $H^2(\mathbf{R})$ . The corresponding analytic functions in the upper-half plane form the complex Hardy space  $H^2(\mathbf{C}^+)$ . The last mentioned two spaces are isometrically isomorphic [13].

For periodic signals there is a parallel theory. For a periodic signal  $s$  of finite energy we have

$$s(e^{it}) = \sum_{k=-\infty}^{\infty} c_k e^{ikt}, \quad \sum_{k=-\infty}^{\infty} |c_k|^2 < \infty.$$

The analytic signal

$$s^+(z) = \sum_{k=0}^{\infty} c_k z^k = \frac{1}{2\pi i} \int_0^{2\pi} \frac{s(e^{iu})}{e^{iu} - z} de^{iu}$$

has the boundary limit

$$s^+(e^{it}) = \frac{1}{2}s(e^{it}) + \frac{1}{2}\tilde{H}s(e^{it}) + \frac{c_0}{2}, \tag{1.1}$$

where  $\tilde{H}$  is the circular Hilbert transformation, given by

$$\begin{aligned} \tilde{H}s(e^{it}) &= \sum_{k \neq 0} (-i \operatorname{sgn}(k)) c_k e^{ikt} \\ &= \frac{1}{2\pi} \text{p.v.} \int_0^{2\pi} \cot\left(\frac{t-s}{2}\right) s(e^{is}) ds. \end{aligned}$$

If  $s$  is real-valued, we have the relations

$$s = 2 \operatorname{Re} s^+ - c_0, \quad Hs = 2 \operatorname{Im} s^+. \tag{1.2}$$

As in the real line case, the periodic analytic signals of finite energy form a closed subspace of  $L^2(\partial\mathbf{D})$ , called the Hardy space  $H^2(\partial\mathbf{D})$ , that is isometrically isomorphic to  $H^2(\mathbf{D})$ , the latter consisting of the corresponding analytic functions in the open unit disc. In the sequel we will concentrate in the unit disc case corresponding to periodic signals.

Efficiency and error estimate of a discrete method to compute images of operators are usually problematic, let alone those of singular integral operators. The most commonly used method of computing Hilbert transform is the one through FFT. Some novel methods to compute Hilbert transform were recently developed [1,2]. The B splines of cardinality of order 1 are used by Yang et al. in [1] to compute Hilbert transform. Their method is for continuous functions in the  $L^2$  spaces. For functions in  $C^3 \cap L^2$  Micchelli et al. also developed a spline method to compute Hilbert transform [2]. They established error estimation in finite intervals that means locally. Referring to approximation accuracy the mentioned two methods are superior to FFT (also see [14]). The AFD methods proposed by this paper have a few advantages comparing with the existing ones. The first is that the algorithm is for general non-smooth and non-continuous signals: It is robust to changes of signal values in sets of Lebesgue measure zero that can be dense subsets in the domain of definition of the signal. The second is that the  $n$ th partial sum of a maximal-energy AFD has a global error estimation  $O(n^{-1/2})$ . This estimation is for all signals of finite energy including the non-smooth ones. This convergence rate is as coincidentally the same as the Shannon sampling. But the latter is for a very narrow class: It is for band-limited signals equivalent to those in the Paley–Wiener class consisting of entire functions, hence infinitely smooth, and with an exponential type of bounds. The third advantage is that it is a by-product of an AFD approximation to a  $L^2$  function. In such way to compute the Hilbert transform of the signal

does not require an extra effort. One obtains an approximation to the Hilbert transform at the same time as one obtains an approximation to the original signal. This shows the power of the analytic function method. Finally, the AFD methods give rise to explicit functional formulas, in fact, of the rational function type, for the approximations to the Hilbert transform, no matter the original signal is presented by a discrete or continuous (functional) formality. This is, in particular, superior to FFT.

## 2. Adaptive Fourier decompositions

Various maximal-energy AFDs involve orthogonal rational systems, or Takenaka–Malmquist systems (abbreviated as TM systems). A finite or infinite TM system  $\{B_n\}$  is defined through, respectively, a finite or an infinite sequence  $a_1, \dots, a_n, \dots$  in the complex unit disc  $\mathbf{D}$ . It reads as

$$B_n(z) = \frac{\sqrt{1 - |a_n|^2}}{1 - \bar{a}_n z} \prod_{l=1}^{n-1} \frac{z - a_l}{1 - \bar{a}_l z}, \quad n = 1, 2, \dots$$

In the infinite sequence case,  $\{B_n\}$  forms a basis of  $H^p(\mathbf{D})$ ,  $1 \leq p \leq \infty$ , if and only if there holds the hyperbolic non-separable condition

$$\sum_{k=1}^{\infty} (1 - |a_k|) = \infty. \quad (2.3)$$

A rational function system  $\{B_n\}$  is called a generalized Fourier system for the reason that it becomes a half of the Fourier system,  $\{z^{n-1}\}$ , if all the parameters  $a_n$ 's are identically zero.

When we deal with signals defined on the real line (the whole time range) we involve the Hardy space  $H^2(\mathbf{C}^+)$  in which a TM system is of the form

$$B_n(z) = \sqrt{\frac{\beta_n}{\pi}} \frac{1}{z - \bar{a}_n} \prod_{l=0}^{n-1} \frac{z - a_l}{z - \bar{a}_l}, \quad (2.4)$$

$$a_n = \alpha_n + i\beta_n \in \mathbf{C}^+, \quad n = 0, 1, \dots, z \in \mathbf{C}^+.$$

Such system is a basis in  $H^p(\mathbf{C}^+)$ ,  $1 \leq p \leq \infty$ , if and only if the condition

$$\sum_{l=0}^{\infty} \frac{\beta_n}{1 + |a_n|^2} = \infty \quad (2.5)$$

is met.

Maximal-energy AFD algorithms make optimal selections of the parameters  $a_1, \dots, a_n, \dots$ , according to the given signal  $s$ . The selections do not guarantee the above condition (2.3) or (2.5) to hold. Although the resulted TM system may not be a basis, it offers fast decomposition of the given signal into the so called *mono-components*, or signals with positive analytic frequency (phase derivative).

The main maximal-energy AFD variations include Core AFD [3], Unwinding AFD [5] and Cyclic AFD [6]. The following formulation is based on Core AFD. For the other two some modifications are needed. For a real-valued  $s \in L^2(\partial\mathbf{D})$  or  $s \in L^2(\mathbf{R})$ , we first have the decomposition  $s = s^+ + s^-$ , and

$$s^+ = \sum_{k=1}^{\infty} \langle s, B_k \rangle B_k,$$

where  $B_k$  are the rational functions in the corresponding TM system with optimally selected parameters  $a_1, \dots, a_n, \dots$ . We note that the above expression is valid for both the unit circle and the real line contexts.

Based on the relation (1.2), for the unit circle case, the corresponding expansion of the originally real-valued signal  $s$  is

$$s = 2 \operatorname{Re} \left\{ \sum_{k=1}^{\infty} \langle s, B_k \rangle B_k \right\} - c_0 \quad (2.6)$$

$$= \sum_{k=1}^{\infty} \rho_k(t) \cos \theta_k(t) - c_0$$

and that for the Hilbert transform is

$$Hs = 2 \operatorname{Im} \left\{ \sum_{k=1}^{\infty} \langle s, B_k \rangle B_k \right\} = \sum_{k=1}^{\infty} \rho_k(t) \sin \theta_k(t)$$

where

$$\rho_k(t)e^{i\theta_k(t)} = 2\langle s, B_k \rangle B_k(e^{it}).$$

For the real line context we have the same expansions for  $s$  and  $Hs$ , except that the constant  $c_0$  in (2.6) should be dropped off.

The maximal-energy AFD results in a fast decomposition of  $s^+$ . In view of (1.2) and (2.6), we get a fast decomposition for  $s$ , too. Taking into account that Hilbert transformation in either of the two contexts preserves the  $L^2$ -norm, there holds  $\|s\| = \|Hs\|$ , and thus

$$\begin{aligned} & \left\| Hs - 2 \operatorname{Im} \sum_{k=1}^n \langle s, B_k \rangle B_k \right\| \\ &= \left\| s - \left( 2 \operatorname{Re} \sum_{k=1}^n \langle s, B_k \rangle B_k - c_0 \right) \right\|, \end{aligned}$$

or, in the whole time range context,

$$\begin{aligned} & \left\| Hs - 2 \operatorname{Im} \sum_{k=1}^n \langle s, B_k \rangle B_k \right\| \\ &= \left\| s - \left( 2 \operatorname{Re} \sum_{k=1}^n \langle s, B_k \rangle B_k \right) \right\|. \end{aligned}$$

Since the right-hand-sides of the above equalities converge fast to zero, the left-hand-sides do the same, too. For functions in a general enough class (see [4]) we have, in both cases,

$$\left\| Hs - \sum_{k=1}^n \rho_k(t) \sin \theta_k(t) \right\| \leq \frac{M}{\sqrt{n}},$$

where  $M$  is a constant depending on  $s$ .

The right-hand-side of the above estimation itself may not look promising. But one should not expect better results for the reason that we deal with signals that are, in general, non-smooth. However, if we restrict ourselves to functions  $s^+$  that have analytic continuation cross the whole boundary of the unit disc or the upper-half plane, then we have a convergence rate exponentially decaying along with  $n \rightarrow 0$  [3].

### 3. Complex SVM based on Szegő kernel

Complex support vector machine (SVM) based on Szegő kernels is first proposed in [12]. In the unit disc context the parameterized Szegő kernel is

$$K(z, a) = \frac{1}{1 - \bar{a}z},$$

where  $a$  is the parameter and  $z$  is the argument, both being in  $\mathbf{D}$ .

It is known that a Takenaka–Malmquist system is generated from a sequence of parameterized Szegő kernels by using the Gram–Schmidt orthogonalization process.

Complex SVM, like real SVM developed by Vapnik and his coworkers in 1995 [15], is based on statistical learning theory. Not only that our SVM algorithm uses powerful analytic function theory, but also it seeks to minimize upper bounds of the generalization error constituted by the training error and the confidence interval. This is different from the commonly used algorithms such as empirical risk minimization (ERM), the latter only minimizing the training error. The complex SVMs achieve considerably better generalization performance than the existing learning algorithms based on, for example, the ERM principle. The use of robust cost functions in complex SVMs can decrease the effect of outliers. Training SVM is equivalent to solving a linearly constrained quadratic programming problem.

With the training data  $\{(z_m, s^+(z_m))\}_{m=1}^n$ , the Szegő-kernel-based complex SVM is of the form

$$s^+(z) \approx \sum_{m=1}^n \overline{\psi_m} K(z, z_m),$$

where  $\psi_m = (\alpha_m - \alpha_m^*) - j(\beta_m - \beta_m^*)$ ,  $\alpha_m, \alpha_m^*, \beta_m$  and  $\beta_m^*$  are the solutions of a QP-problem.

As in the usual SVM framework, by letting  $\varepsilon > 0$ , we have only a subset of the Lagrange multipliers being nonzero, and thus we obtain a sparse solution.

$$s^+(z) \approx \sum_{m=1}^n \overline{\psi_m} K(z, z_m) = \sum_{m \in J} \overline{\psi_m} K(z, z_m), \tag{3.7}$$

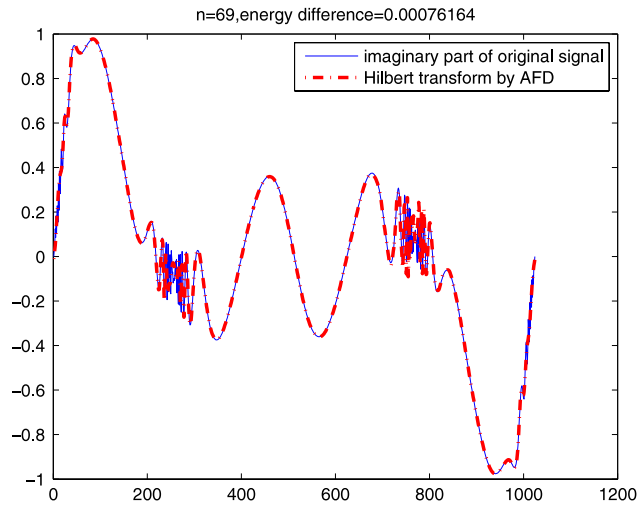


Fig. 1. Comparison of the imaginary part of  $s^+(z)$  and the Hilbert transform given by AFD.

where  $J = \{m : \psi_m \neq 0\}$ . We, in fact, have the sparsity  $|J| \ll n$ . The corresponding algorithm can be generated based on [12].

Based on relation (1.2), the corresponding expansion for the original real-valued signal  $s$  is

$$\begin{aligned} s &\approx 2 \operatorname{Re} \left\{ \sum_{m \in J} \overline{\psi}_m K(z, z_m) \right\} - c_0 \\ &= \sum_{m \in J} \rho_m(t) \cos \theta_m(t) - c_0 \end{aligned} \quad (3.8)$$

and

$$Hs \approx 2 \operatorname{Im} \left\{ \sum_{m \in J} \overline{\psi}_m K(z, z_m) \right\} = \sum_{m \in J} \rho_m(t) \sin \theta_m(t),$$

where

$$\rho_m(t) e^{i\theta_m(t)} = 2 \overline{\psi}_m K(z, z_m).$$

Note that the decomposition is not orthogonal.

#### 4. Experiment

Consider the singular inner function in the disc

$$s^+(z) = \frac{1 + 2z^2}{(z-2)(z-3)} - \frac{1}{(z+2)(z+3)} e^{\frac{z+1}{z-1} + \frac{z+i}{z-i} + \frac{z-i}{z+i}}.$$

It is known that  $s^+$  is an analytic signal, or, equivalently, a signal in  $H^2(\mathbf{D})$ . The experiments are based on the fact that the imaginary part of any analytic signal is the Hilbert transform of its real part.

It is well known that Fourier series of singular inner functions converge slowly. Figs. 1–3 show that in the high-oscillatory portions of  $s^+$  the Core AFD graph and the complex SVM graph coincide with  $s^+$  very well, but the FD graph does not. Being compared with FFT in Fig. 4, although the latter presents almost the same approximation as Core AFD and the complex SVM, it does not give a functional solution, but only a set of approximating discrete data. Theoretically, errors of FFT in computing continuous objects are hard to analyze. Comparing Core AFD and the complex SVM, on the given example, the SVM performs visibly better. The drawbacks of the SVM is that the cardinality of the support,  $|J|$ , cannot be a priori determined; and the fact that the error estimation is in terms of statistics brings uncertainty. Like SVM-AFD, the other more sophisticated AFD variations, including Unwending and Cyclic AFDs, perform better than Core AFD [16], the latter being the most fundamental in the theory, and the very basic constructive block in the algorithms for the other maximal-energy AFDs. In particular, the superiority of Unwending AFD compared with Core AFD, especially on singular inner functions, is shown in [16]. Since the formulas given in Section 2 are based on Core AFD, the experiment in this paper is only done by using Core AFD.

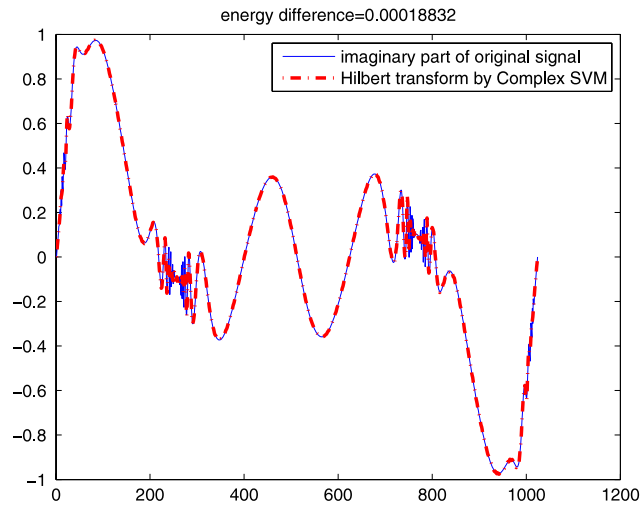


Fig. 2. Comparison of the imaginary part of  $F(z)$  and Hilbert transform given by complex SVM.

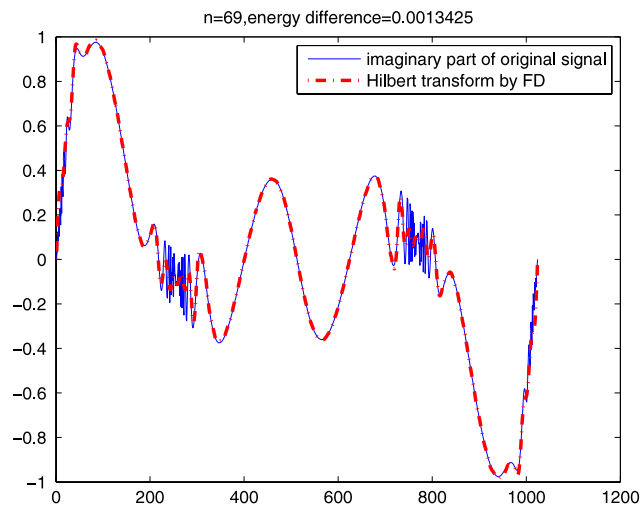


Fig. 3. Comparison of the imaginary part of  $s^+(z)$  and the Hilbert transform given by Core FD.

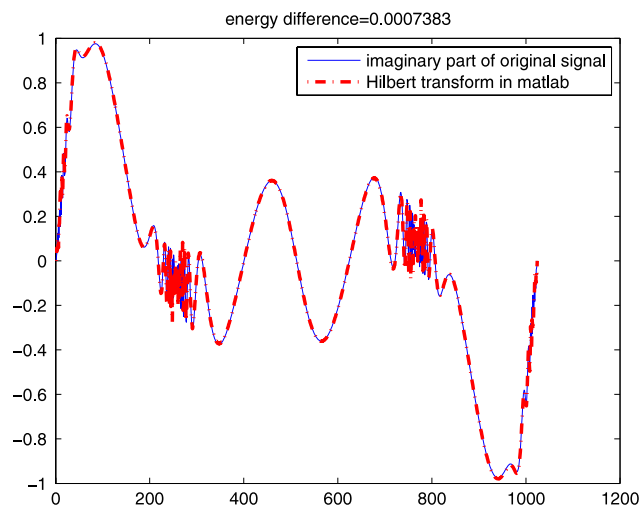


Fig. 4. Comparison of the imaginary part of  $s^+(z)$  and the Hilbert transform in Matlab through FFT.

## 5. Conclusions

As advantage of the complex analysis method, the AFD methods, including the complex SVM, give, as a by-product, Hilbert transforms of signals to be approximated. The AFD methods, exhibiting a global estimation  $O(n^{-1/2})$  in the Core AFD case, is tolerant to non-smooth signals, being robust to changes of signal values in sets of Lebesgue measure zero. The AFD methods for computing Hilbert transforms are valid for all signals of finite energy, no matter for restricted time or unrestricted time ranges, and provide explicit rational functional approximations. In contrast, the other existing methods all require certain degrees of smoothness, depending strongly on signal pointwise values, and having at most local error estimations. We point out that the method is superior to FFT, for the latter only offering data to data approximation.

## Acknowledgments

The study was supported by University of Macau MYRG116(Y1-L3)-FST13-QT and Macao Government FDCT 098/2012/A3.

## References

- [1] C. Zhou, L. Yang, Y. Liu, Z. Yang, A novel method for computing the Hilbert transform with Haar multiresolution approximation, *J. Comput. Appl. Math.* 223 (2009) 585–579.
- [2] C.A. Micchelli, Y. Xu, B. Yu, On computing with the Hilbert spline transform, *Adv. Comput. Math.* (2011) <http://dx.doi.org/10.1007/s10444-011-9252-x>.
- [3] T. Qian, Y. Wang, Adaptive Fourier series—a variation of greedy algorithm, *Adv. Comput. Math.* 34 (3) (2011) 279–293.
- [4] T. Qian, Y. Wang, Remarks on adaptive Fourier decomposition, *Int. J. Wavelets Multiresolut. Inf. Process.* 11 (1) (2013) 1–14.
- [5] T. Qian, Intrinsic mono-components decomposition of functions: an advance of Fourier theory, *Math. Methods Appl. Sci.* 33 (2010) 880–891.
- [6] T. Qian, Cyclic AFD algorithm for best rational approximation, *Math. Methods Appl. Sci.* 37 (2013) 846–859. <http://dx.doi.org/10.1002/mma.2843>.
- [7] W. Mi, T. Qian, Frequency domain identification: an algorithm based on adaptive rational orthogonal system, *Automatica* 48 (6) (2012) 1154–1162.
- [8] W. Mi, T. Qian, On backward shift algorithm for estimating poles of systems, *Automatica* 50 (2014) 1603–1610.
- [9] W. Mi, T. Qian, F. Wan, A fast adaptive model reduction method based on Takenaka-Malmquist systems, *Systems Control Lett.* 61 (1) (2012) 223–230.
- [10] T. Qian, L. Zhang, Z. Li, Algorithm of adaptive Fourier transform, *IEEE Trans. Signal Process.* 59 (12) (2011) 5899–5906.
- [11] Y. Mo, T. Qian, Support vector machine adapted Tikhonov regularization method to solve dirichlet problem, *Appl. Math. Comput.* 245 (2014) 509–519.
- [12] Y. Mo, T. Qian, W. Mi, Szegő kernels-based complex SVM for system identification, *Int. J. Wavelets Multiresolut. Inf. Process.*, in press.
- [13] J.B. Garnett, *Bounded Analytic Functions*, Academic Press, 1987.
- [14] L. Marple Jr., Computing the discrete-time analytic signal via FFT, *IEEE Trans. Signal Process.* 47 (2001) 2600–2603.
- [15] V. Vapnik, *The Nature of Statistical Learning Theory*, Springer-Verlag, New York, 1995.
- [16] T. Qian, H. Li, M. Stessin, Comparison of adaptive mono-component decompositions, *Nonlinear Anal. RWA* 14 (2) (2013) Pages 1055V1074.