TECHNICAL RESEARCH REPORT

Structured Low Rank Matrix Pencil for Spectral Estimation and System Identification

by J. Razavilar, Y. Li and K.J.R. Liu

T.R. 96-35

INSTITUTE FOR SYSTEMS RESEARCH

*Sponsored by
the National Science Foundation
Engineering Research Center Program,
the University of Maryland,
Harvard University,
and Industry*

A Structured Low Rank Matrix Pencil for Spectral Estimation and System Identification*

Javad Razavilar, Ye Li and K. J. Ray Liu
Electrical Engineering Department and Institute for Systems Research
University of Maryland
College Park, MD 20742
Fax:1-301-405-6707
Email: javad(liye, kjrliu)@isr.umd.edu

ABSTRACT

In this paper we propose a new parameter estimation algorithm for damped sinusoidal signals. Parameter estimation for damped sinusoidal signals with additive white noise is a problem of significant interests in many signal processing applications, such as analysis of NMR data and system identification. The proposed algorithm estimates the signal parameters using a matrix pencil constructed from the measured data. To reduce the noise effect, rank deficient Hankel approximation of prediction matrix is used. We show that the performance of the estimation can be significantly improved by structured low rank approximation of the prediction matrix. Computer simulations also show that the noise threshold of our new matrix pencil algorithm is significantly lower than those of the existing algorithms.

1 Introduction

High resolution parameter estimation for damped sinusoidal signals in the presence of additive white noise is a problem of significant interests in many signal processing applications, such as analysis of NMR data, system identification, and spectral estimation. Many approaches to high resolution spectral estimation have been proposed for pure sinusoidal signals, including linear prediction (LP) techniques [1], and signal subspace methods like MUSIC (multiple signal classification) algorithm [2], and ESPRIT (Estimation of Signal Parameters via Rotational Invariance Techniques) [3]. In comparison to earlier methods of parameter estimation, MUSIC and ESPRIT are known to provide better performance in estimation of parameters. The difficulty of parameter estimation for damped sinusoidal signals stems from the fact that these signals are nonstationary. Therefore, most of good algorithms for stationary signals can not be applied here. Kumaresan-Tufts (KT) algorithm [4] and Hua-Sarkar's matrix pencil method [5], are known to provide better estimations for frequencies and damping factors of damped sinusoidal signals.

*This work was supported in part by the NIH grant 1R01GM49707 and the NSF grants MIP9309506 and MIP9457397.

But both algorithms do not provide accurate estimates when the SNR is lower than a certain threshold. Recently, the modified KT algorithm (MKT) that has better performance than KT algorithm [4] is proposed in [7]. In this paper we propose a new algorithm that estimates the signal parameters by using a matrix pencil constructed from the noise corrupted data. To reduce the noise effect, a low rank Hankel approximation of the prediction matrix is used to preserve the Hankel structure of the prediction matrix. This structured low rank approximation of prediction matrix [6, 7] has a great effect on the performance of the new algorithm. Computer simulation results show that this new matrix pencil algorithm has lower noise threshold than that of *KT* algorithm [4], *MKT* [7], and Hua-Sarkar's matrix pencil algorithms [5].

2 Data Model and Algorithm Development

Consider a sequence $y(n)$ consisted of K damped sinusoidal signals with additive white noise $w(n)$ as

$$y(n) = \sum_{k=1}^{K} c_k e^{s_k n} + w(n), \quad n = 0, \cdots, N-1 \tag{1}$$

where K is the number of damped sinusoids, and $N \geq 2K$, $s_k = -\alpha_k + \jmath\omega_k$, $\omega_k \in [-\pi,\ \pi]$, $\alpha_k \geq 0$ which is called the damping factor. We are interested in estimating the frequencies and damping factors of the measured data $y(n)$. In order to do this, we first form $K \times K$ prediction matrices as

$$\mathbf{A}_n = \begin{pmatrix} y(n) & \cdots & y(n+K-1) \\ y(n+1) & \cdots & y(n+K) \\ \vdots & \vdots & \vdots \\ y(n+K-1) & \cdots & y(n+2K-2) \end{pmatrix} \tag{2}$$

for $n = 0,\ 1, \cdots,\ N-2K+1$. From (1) it can be shown that

$$\mathbf{A}_n = \mathbf{S}^T \mathbf{C} \mathbf{\Phi}^n \mathbf{S} + \mathbf{W}_n, \tag{3}$$

where

$$\mathbf{W}_n = \begin{pmatrix} w(n) & \cdots & w(n+K-1) \\ w(n+1) & \cdots & w(n+K) \\ \vdots & \vdots & \vdots \\ w(n+K-1) & \cdots & w(n+2K-2) \end{pmatrix} \tag{4}$$

$$\begin{aligned} \mathbf{\Phi} &= \text{diag}(e^{s_1},\ e^{s_2}, \cdots,\ e^{s_K}), \\ \mathbf{C} &= \text{diag}(c_1,\ c_2, \cdots,\ c_K), \\ \mathbf{S} &= [\mathbf{r}(s_1), \mathbf{r}(s_2), \cdots, \mathbf{r}(s_K)]^T, \end{aligned} \tag{5}$$

with

$$\mathbf{r}(s_k) = [1, e^{s_k}, \cdots, e^{(K-1)s_k}]^T.$$

If there is no measurement noise, from (3)

$$\mathbf{P} = \mathbf{A}_n^{-1}\mathbf{A}_{n+1} = \mathbf{S}^{-1}\mathbf{\Phi}\mathbf{S}. \tag{6}$$

We can see in (6) that the eigenvalues of $\mathbf{A}_n^{-1}\mathbf{A}_{n+1}$ are e^{s_1}, e^{s_2},$\cdots$, e^{s_K}, i.e. *when the measured data is not corrupted by noise, signal parameters can be estimated from eigenvalues of* $\mathbf{A}_n^{-1}\mathbf{A}_{n+1}$. In the presence of measurement noise, we have to reduce the noise effect before we can apply the above matrix pencil algorithm. In other words, we have to first approximate noisy sequence $y(n)$ with another sequence $\widehat{y}(n)$ which is less noisy, and then apply the new matrix pencil algorithm to $\widehat{y}(n)$. For this purpose we first construct a prediction matrix from the measured data as

$$\mathbf{A}_y = \begin{pmatrix} y(0) & y(1) & \dots & y(L-1) \\ y(1) & y(2) & \dots & y(L) \\ \vdots & \vdots & \vdots & \vdots \\ y(L-1) & y(L) & \dots & y(2L-2) \end{pmatrix} \tag{7}$$

where $L = \lceil N/2 \rceil$ in order to have the best performance. Rank deficient Hankel approximation of the $\mathbf{A}_y$ given in [6, 7] is used to find the sequence $\widehat{y}(n)$ as follows:

Initialization: $\widehat{\mathbf{A}}_y^{[0]} = \mathbf{A}_y$ and $r = 0$ (r is the iteration index)
1) Compute $SVD(\widehat{\mathbf{A}}_y^{[r]}) = \mathbf{U}\mathbf{D}\mathbf{V}^H$
2) Obtain $\bar{\mathbf{A}}_y = [\bar{y}_{i,j}]_{i,j=0}^{L-1} = \sum_{k=1}^{K} \sigma_k \mathbf{u}_k \mathbf{v}_k^H$.
3) Find a Hankel matrix $\widehat{\mathbf{A}}_y^{[r]}$ to minimize $\| \widehat{\mathbf{A}}_y^{[r]} - \bar{\mathbf{A}}_y \|_F$, where

$$\widehat{\mathbf{A}}_y^{[r]} = [\widehat{y}^{[r]}(i+j)]_{i,j=0}^{L-1}, \tag{8}$$

where

$$\widehat{y}_{i+j}^{[r]} = \frac{1}{\Gamma_{ij}} \sum_{n+m=i+j,\ 0 \le n,m \le L-1} \bar{y}_{n,m}. \tag{9}$$

in which Γ_{ij} is the number of the elements in matrix $\bar{\mathbf{A}}_y$ satisfying $n+m = i+j$ in (9).
4) Repeat steps **1**, **2** *and* **3** till the rank of $\widehat{\mathbf{A}}_y^{[r]} = K$ (where K is the number of signals).

We have shown in [7] that such structured matrix approximation reduces the mean square error of the prediction error. The sequence $\widehat{y}(n)$ resulted from the above rank deficient Hankel approximation algorithm is now used to construct new $K \times K$ prediction matrices from $\widehat{y}(n)$ as follows

$$\widehat{\mathbf{A}}_n = \begin{pmatrix} \widehat{y}(n) & \cdots & \widehat{y}(n+K-1) \\ \widehat{y}(n+1) & \cdots & \widehat{y}(n+K) \\ \vdots & \vdots & \vdots \\ \widehat{y}(n+K-1) & \cdots & \widehat{y}(n+2K-2) \end{pmatrix} \tag{10}$$

for $n = 0,\ 1, \cdots,\ N-2K+1$. Then we have

$$\begin{aligned} \widehat{\mathbf{A}}_n &= \mathbf{S}^T\mathbf{C}\mathbf{\Phi}^n\mathbf{S} + \widehat{\mathbf{W}}_n \\ &= \mathbf{D}_n + \widehat{\mathbf{W}}_n \end{aligned} \tag{11}$$

where $\mathbf{S}$, $\mathbf{C}$, and $\mathbf{\Phi}$ are defined in (5) and $\mathbf{D}_n = \mathbf{S}^T\mathbf{C}\mathbf{\Phi}^n\mathbf{S}$.

To estimate the signal parameters we construct the following matrices:

$$\begin{aligned} \widehat{\mathbf{P}}_n &= \widehat{\mathbf{A}}_n^{-1}\widehat{\mathbf{A}}_{n+1} \\ &= (\mathbf{D}_n + \widehat{\mathbf{W}}_n)^{-1}(\mathbf{D}_{n+1} + \widehat{\mathbf{W}}_{n+1}). \end{aligned} \tag{12}$$

Now it can be shown that

$$\begin{aligned} \widehat{\mathbf{A}}_n^{-1} &= (\mathbf{D}_n + \widehat{\mathbf{W}}_n)^{-1} \\ &\approx \mathbf{D}_n^{-1} - \mathbf{D}_n^{-1}\widehat{\mathbf{W}}_n\mathbf{D}_n^{-1}. \end{aligned} \tag{13}$$

From (13) and noting that $\mathbf{P} = \mathbf{D}_n^{-1}\mathbf{D}_{n+1} = \mathbf{S}^{-1}\mathbf{\Phi}\mathbf{S}$, we can rewrite (12) as follows

$$\widehat{\mathbf{P}}_n \approx \mathbf{P} + \widehat{\mathbf{A}}_n^{-1}\widehat{\mathbf{W}}_{n+1} - \mathbf{D}_n^{-1}\widehat{\mathbf{W}}_n\mathbf{P}. \tag{14}$$

From (14) it is clear that $\widehat{\mathbf{P}}_n$ and $\mathbf{P}$ do not have the same eigenvalues. It was mentioned earlier that we can compute signal parameters directly from eigenvalues of $\mathbf{P}$, but in the presence of noise we can only access $\widehat{\mathbf{P}}_n$ rather than $\mathbf{P}$. Therefore we will get a poor estimate of signal parameters by only considering the eigenvalues of some $\widehat{\mathbf{P}}_n$s. To improve the accuracy of parameter estimation, we construct a new matrix, $\widetilde{\mathbf{P}}$, from the linear combination of $\widehat{\mathbf{P}}_n$s as follows

$$\widetilde{\mathbf{P}} = \textstyle\sum_{n=0}^{N-2K} \gamma_n \widehat{\mathbf{P}}_n \tag{15}$$

where γ_n's are the weighting factors. To make the estimation unbiased we must ensure

$$\sum_{n=0}^{N-2K} \gamma_n = 1. \tag{16}$$

From (16) and (14), we can rewrite (15) as

$$\widetilde{\mathbf{P}} = \mathbf{P} + \textstyle\sum_{n=0}^{N-2K} \gamma_n \bar{\mathbf{W}}_n \tag{17}$$

where $\bar{\mathbf{W}}_n = \widehat{\mathbf{A}}_n^{-1}\widehat{\mathbf{W}}_{n+1} - \mathbf{D}_n^{-1}\widehat{\mathbf{W}}_n\mathbf{P}$.

In our simulations we have chosen γ_n's as

$$\gamma_n = \frac{|det(\mathbf{A}_n)|^{2/K}(L+1-|L-n|)}{\sum_{k=0}^{N-2K} |det(\mathbf{A}_k)|^{2/K}(L+1-|L-k|)}. \tag{18}$$

From (18) and (15) we can compute $\widetilde{\mathbf{P}}$ and the frequencies and damping factors of signalscan be estimated from the eigenvalues of $\widetilde{\mathbf{P}}$. It should be pointed out that in our algorithm the order of the model has to be determined in advance. There are various effective methods for this purpose [1, 8].

3 Computer Simulation Results

In this section, we will demonstrate the performance of the new matrix pencil algorithm by two examples drawn from [4].

Example 1:

The purpose of first example is to demonstrate the performance of the new matrix pencil algorithm for spectral estimation. The simulated data is generated as follows:

$$y(n) = e^{s_1 n} + e^{s_2 n} + w(n), \qquad for\ n = 0, 1, \cdots, 24. \tag{19}$$

where $s_i = -\alpha_i + \jmath 2\pi f_i$ with $\alpha_1 = 0.1$, $f_1 = 0.52$, $\alpha_2 = 0.2$, $f_2 = 0.42$, $w(n)$ is complex white Gaussian noise with zero-mean and variance σ^2. The signal to noise ration (SNR) is defined SNR $= 10\log(\frac{1}{2\sigma^2})$. The MSE's of ω_1, α_1, ω_2 and α_2 for KT algorithm [4], MKT [7], Hua-Sarkar's matrix pencil method [5], and the new matrix pencil algorithm, using the average of 500 trails, are shown in Figure 1. Figure 1 (a) shows that the noise threshold associated with smaller damping factor is lower than that of associated with larger damping factor as given in Figure 1 (c). For ω_1 with damping factor $\alpha_1 = 0.1$, from Figure 1 (a), the noise threshold of the new matrix pencil algorithm is about 12 dB lower than that of KT algorithm and about 8 dB is lower than those of MKT and Hua-Sarkar's matrix pencil algorithms. For ω_2 with damping factor $\alpha_2 = 0.2$, from Figure 1 (c), the noise threshold of the new matrix pencil algorithm is about 10 dB lower than that of KT algorithm and about 5 dB is lower than those of MKT and Hua-Sarkar's matrix pencil algorithms.

Example 2:

We want to estimate the poles and zeros of a linear system from its noise corrupted samples of the impulse response. The transfer function of the linear system is

$$H(z) = \frac{B(z)}{A(z)} = \frac{1 + \sum_{k=1}^{2} b_k z^{-k}}{1 + \sum_{k=1}^{10} a_k z^{-k}}, \tag{20}$$

where poles of the transfer function are shown in Table 1. The magnitude of $H(e^{\jmath\omega})$ is shown in Figure 2 (a), which has two nulls at $\omega = \pm\pi/4$ respectively. The first forty real-valued samples of noise corrupted impulse response are observed. The noise is real white Gaussian with zero-mean and variance σ^2 determined by SNR defined $SNR = 10log(\frac{\sum_{n=0}^{39}|h(n)|^2}{N\sigma^2})$, KT algorithm, MKT algorithm and the new algorithm are employed to estimate the poles of the system to get $\widehat{A}(z)$. Once $\widehat{A}(z)$ is obtained, $\widehat{B}(z)$ can be estimated using Shanks' method citekt, which first generates a sequence f_n by $f(n) = \mathcal{Z}^{-1}\{\frac{1}{\widehat{A}(z)}\}$, and then estimate b_k for $k = 0,\ 1,\ 2$ by minimizing the error $E = \sum_{n=0}^{39} \mid h(n) - \sum_{k=0}^{2} \widehat{b}_k f(n-k) \mid^2$. Figure 2 (b), (c) and (d) show 10 trails of magnitudes of estimated transfer function using KT, MKT and the new matrix pencil algorithms respectively. Also Table 1 illustrates the mean and variance of the estimated poles of transfer function using KT, MKT and the new matrix pencil algorithms. From Figure 2 and Table 1, it is clear that the new matrix pencil algorithm outperfoms KT and MKT algorithms.

References

[1] S. Haykin, *Adaptive Filter Theory*, Prentice Hall, 1991.

[2] R. O. Schmidt, "Multiple Emitter Location and Signal Parameter Estimation" *IEEE Trans. on Ant. and Prop.*, vol. 34, pp. 276-280, March 1986.

[3] R. Roy, T. Kailath, "ESPRIT- Estimation of Signal Parameters Via Rotational Invariance Techniques," *IEEE Trans. on ASSP*, vol. 37, pp. 984-995, Jul. 1989.

[4] R. Kumaresan and R. W.Tufts, "Estimation the parameters of exponentially damped sinusoids and pole-zero modelling in noise," *IEEE Trans. on ASSP*, Dec. 1982.

[5] Y. Hua, T. K. Sarkar, "Matrix Pencil Method for Estimating Parameters of Exponentially Damped/Undamped Sinusoids in Noise", *IEEE Trans. on ASSP*, pp 814-824, May 1990.

[6] J. Cadzow, "Signal Enhancement: A Composite property Mapping Algorithm," *IEEE Trans. on ASSP*, vol. 36, pp. 49-62, Oct. 1988.

[7] Y. Li, K. J. R. Liu, J. Razavilar, "Improved parameter estimation schemes for damped sinusoidal signals," *Conference on Information Sciences and Systems (CISS)*, John Hopkins University, March 1995.

[8] G. W. Stewart, "Perturbation Theory for the Singular Value Decomposition," *SVD and Signal Processing, II* pp 99-109, ElsevierScience Publisher, 1991.

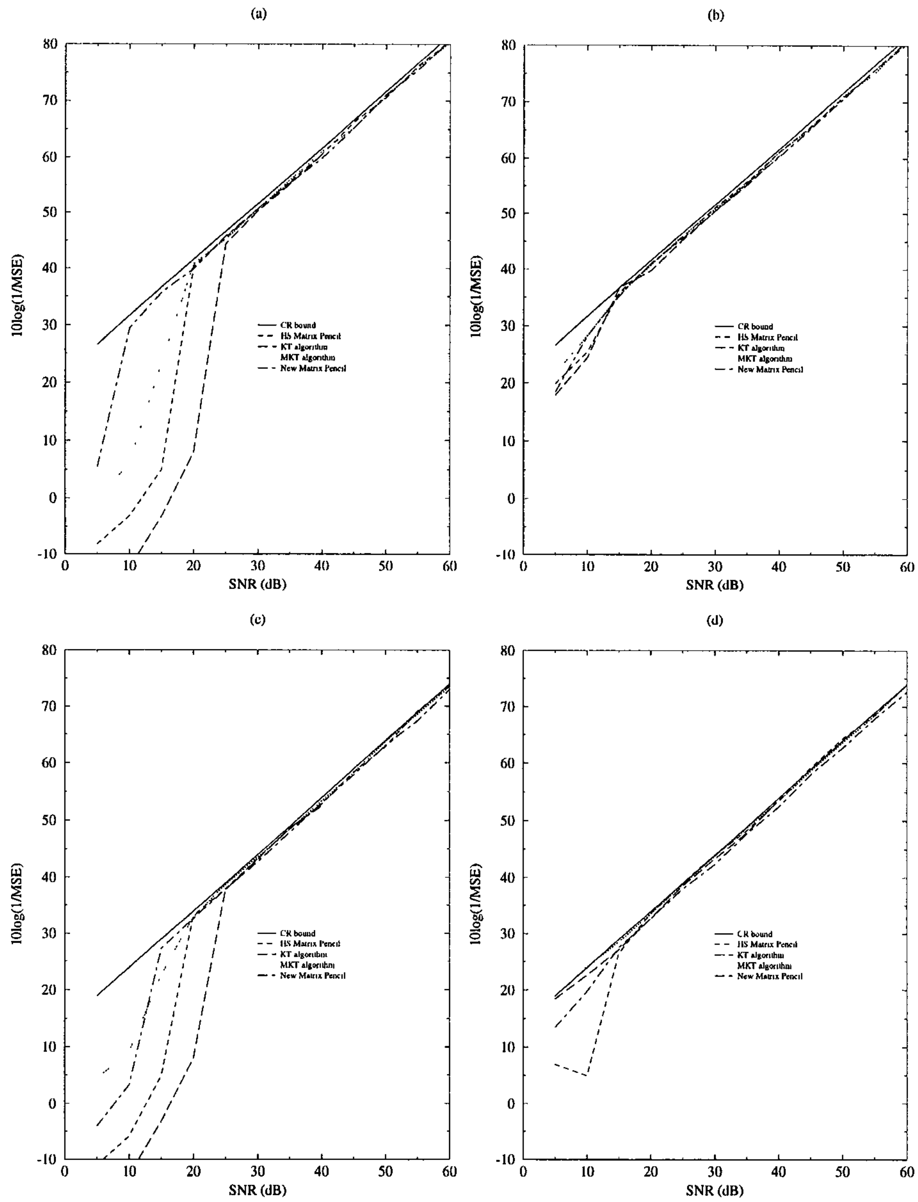

Figure 1: The MSE of (a) ω_1, (b) α_1, (c) ω_2 (d) α_2 for KT, MKT, Hua-Sarkar's matrix pencil and the new matrix pencil algorithms when $s_1 = -0.1 + \jmath 2\pi 0.52$, $s_2 = -0.2 + \jmath 2\pi 0.42$ and $N = 25$.

Table 1: The true and estimated poles of the transfer function $H(z)$

True Poles	Method		Estimated Poles $SNR = 30dB$	Estimated Poles $SNR = 20dB$
-0.2913 $\pm\jmath 0.8968$	KT alg.	Mean	$-0.2880 \pm \jmath 0.9080$	$-0.3448 \pm \jmath 0.8720$
		Std.	1.498×10^{-4}	8.583×10^{-2}
	MKT alg.	Mean	$-0.2915 \pm \jmath 0.8970$	$-0.2917 \pm \jmath 0.8976$
		Std.	3.762×10^{-5}	3.376×10^{-4}
	New alg.	Mean	$-0.2914 \pm \jmath 0.8964$	$-0.2914 \pm \jmath 0.8959$
		Std.	1.048×10^{-5}	1.088×10^{-4}
0.1014 $\pm\jmath 0.9579$	KT alg.	Mean	$0.0987 \pm \jmath 0.9510$	$0.0657 \pm \jmath 0.9392$
		Std.	1.713×10^{-5}	1.406×10^{-2}
	MKT alg.	Mean	$0.1015 \pm \jmath 0.9577$	$0.1015 \pm \jmath 0.9571$
		Std.	4.311×10^{-6}	4.556×10^{-5}
	New alg.	Mean	$0.1014 \pm \jmath 0.9580$	$0.1014 \pm \jmath 0.9578$
		Std.	5.800×10^{-7}	6.292×10^{-6}
0.2959 $\pm\jmath 0.9292$	KT alg.	Mean	$0.2979 \pm \jmath 0.9232$	$0.2845 \pm \jmath 0.9183$
		Std.	2.784×10^{-5}	3.858×10^{-3}
	MKT alg.	Mean	$0.2960 \pm \jmath 0.9291$	$0.2959 \pm \jmath 0.9287$
		Std.	1.992×10^{-6}	1.607×10^{-5}
	New alg.	Mean	$0.2959 \pm \jmath 0.9292$	$0.2960 \pm \jmath 0.9292$
		Std.	2.295×10^{-7}	2.211×10^{-6}
0.5630 $\pm\jmath 0.8019$	KT alg.	Mean	$0.5620 \pm \jmath 0.8175$	$0.5339 \pm \jmath 0.8442$
		Std.	1.827×10^{-4}	8.359×10^{-3}
	MKT alg.	Mean	$0.5629 \pm \jmath 0.8023$	$0.5626 \pm \jmath 0.8045$
		Std.	1.059×10^{-5}	9.462×10^{-5}
	New alg.	Mean	$0.5629 \pm \jmath 0.8018$	$0.5629 \pm \jmath 0.8020$
		Std.	2.224×10^{-6}	2.233×10^{-5}
0.9815 $\pm\jmath 0.1117$	KT alg.	Mean	$0.9853 \pm \jmath 0.1089$	$0.9798 \pm \jmath 0.1219$
		Std.	2.120×10^{-5}	1.763×10^{-2}
	MKT alg.	Mean	$0.9816 \pm \jmath 0.1117$	$0.9821 \pm \jmath 0.1114$
		Std.	1.894×10^{-6}	1.811×10^{-5}
	New alg.	Mean	$0.9815 \pm \jmath 0.1118$	$0.9815 \pm \jmath 0.1117$
		Std.	5.858×10^{-7}	4.988×10^{-6}

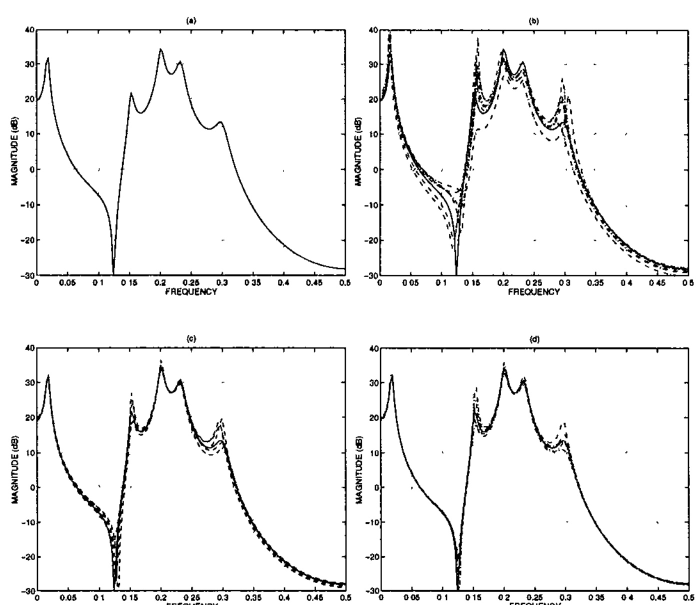

Figure 2: (a) Magnitude of $H(z)$, (b) estimated magnitude of $\widehat{H}(z)$ using KT algorithm, (c) estimated magnitude of $\widehat{H}(z)$ using MKT algorithm, (d) estimated magnitude of $\widehat{H}(z)$ using the new algorithm, SNR=30dB.