signals and SNR's of as low as -9 dB (or less), for the individual sinusoids.

As one may expect, the algorithm performance deteriorates as the noise level increases. In particular, we note that at low SNR's, the algorithm convergence slows down and the variation of the ALE parameters become more erratic, and at some stage, it may fail unless its step sizes are decreased so that a more accurate averaging of the noisy signals can be obtained. This, of course, slows down the algorithm convergence further.

Another problem that our as well as all the existing IIR ALE algorithms are suffering from is their sensitivity to the colored noise, which may result in some bias in the estimated parameters, or it may even result in unreliable behavior of the algorithms as the noise level increases [8]. This problem, in general, is more serious when the cascaded IIR ALE's are used. This is because as input noise passes through the successive stages of the IIR ALE, it becomes more colored. To minimize this effect, one has to let the s_k parameters of the ALE approach unity as close as possible so that the effect of the produced nulls on changing the noise spectrum are minimal.

REFERENCES

[1] B. Widrow *et al.*, "Adaptive noise canceling: Principles and applications," *Proc. IEEE*, vol. 63, pp. 1692–1716, Dec. 1975.

[2] B. Widrow and S. D. Stearns, *Adaptive Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1985.

[3] R. A. David, S. D. Stearns, G. R. Elliott, and D. M. Etter, "IIR algorithm for adaptive line enhancement," in *Proc. ICASSP'83*, Apr. 1983, pp. 17–20.

[4] P. A. Regalia, "An improved lattice-based adaptive IIR notch filter," *IEEE Trans. Signal Processing*, vol. 39, pp. 2124–2128, Sept. 1991.

[5] N. Ahmed, D. Hush, G. R. Elliott, and R. J. Fogler, "Detection of multiple sinusoids using an adaptive cascaded structure," in *Proc. ICASSP'84*, 1984, pp. 21.3.1–21.3.4.

[6] D. R. Hush, N. Ahmed, R. David, and S. D. Stearns, "An adaptive IIR structure for sinusoidal enhancement, frequency estimation, and detection," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-34, pp. 1380–1390, Dec. 1986.

[7] R. L. Cupo and R. D. Gitlin, "Adaptive carrier recovery systems for digital data communications receivers," *IEEE J. Select. Areas Commun.*, vol. 7, pp. 1328–1339, Dec. 1989.

[8] N. I. Cho, and S. U. Lee, "On the adaptive lattice notch filter for the detection of sinusoids," *IEEE Trans. Circuits Syst.*, vol. 40, pp. 405–416, July 1993.

[9] P. A. Regalia, *Adaptive IIR Filtering in Signal Processing and Control*. New York: Marcel Dekker, 1995.

A Parameter Estimation Scheme for Damped Sinusoidal Signals Based on Low-Rank Hankel Approximation

Ye Li, K. J. Ray Liu, and Javad Razavilar

***Abstract*—Most of the existing algorithms for parameter estimation of damped sinusoidal signals are based only on the low-rank approximation of prediction matrix and ignore the Hankel property of the prediction matrix. In this correspondence, we propose a modified KT (MKT) algorithm exploiting both rank-deficient and Hankel properties of the prediction matrix. Computer simulation results demonstrate that compared with the original KT algorithm and the matrix pencil algorithm, the MKT algorithm has lower noise threshold and can estimate the parameters of signal with larger damping factors.**

I. INTRODUCTION

The problem of parameter estimation of damped sinusoidal signals in the presence of additive noise is very important in spectral analysis and many applications, such as magnetic resonance spectroscopy and radioastronomy. The difficulty of this problem stems from the fact that the damped sinusoidal signal is nonstationary, and the correlation matrix cannot be found. Hence, many efficient traditional approaches are not applicable. There are several model-based algorithms being devised to cope with this problem. The Prony method is one of the widely used algorithms, but it is sensitive to measurement noise. The backward linear prediction algorithm (or Kumaresan–Tufts (KT) algorithm) [6] can attain the Cramer–Rao (CR) bound if the peak signal-to-noise ratio (SNR) is high, and the damping factors of signals are small. However, for the signals with lower SNR or large damping factor, the KT algorithm is unable to estimate the signal parameters effectively. Several algorithms have been proposed to improve the high noise threshold problem in the KT algorithm. Some of them are the total least square (TLS) algorithm [8], the maximum likelihood (ML) algorithm [1], and the matrix pencil algorithm [3]. Singular value decomposition (SVD)-based information theoretic criteria [9] have recently been presented to detect the number of damped/undamped sinusoids and parameter estimation.

The existing parameter estimation algorithms for damped sinusoidal signals use only the rank-deficient property of the prediction matrix and ignore its Hankel property. Since the parameter estimation of sinusoidal signals from noisy data is equivalent to the low-rank Hankel matrix approximation of data matrix (or prediction matrix), the performance of parameter estimation will be improved significantly if both rank deficiency and Hankel properties of the prediction matrix are exploited in matrix approximation. Based on this idea, a modified KT algorithm is proposed in this correspondence, which uses both the Hankel and the rank-deficiency properties of the prediction matrix.

This correspondence is organized as follows. In Section II, the matrix approximation in the KT algorithm is analyzed. Then, a

Manuscript received January 28, 1995; revised September 15, 1996. This work was supported in part by NIH Grant 1R01GM49707 and NSF Grants MIP9309506 and MIP9457397. The associate editor coordinating the review of this paper and approving it for publication was Dr. Geoffrey C. Orsak.

Y. Li was with the Electrical Engineering Department and Institute for Systems Research, University of Maryland, College Park, MD 20742 USA. He is now with the Wireless Systems Research Department, AT&T Bell Laboratories—Research, Holmdel, NJ 07733-0400 USA.

K. J. R. Liu and J. Razavilar are with the Electrical Engineering Department and Institute for Systems Research, University of Maryland, College Park, MD 20742 USA.

Publisher Item Identifier S 1053-587X(97)01187-2.

modified Kumaresan–Tufts (MKT) algorithm is developed in Section III. Finally, computer simulation results are presented in Section IV to demonstrate the performance of MKT algorithm.

II. Low-Rank Matrix Approximation

A sequence $x(n)$ consists of K damped sinusoidal signals can be expressed as

$$x(n) = \sum_{k=1}^{K} c_k e^{s_k n} \tag{1}$$

where c_k's are nonzero complex amplitudes, $s_k = -\alpha_k + \jmath\omega_k$, and $\alpha_k \in \mathcal{R}^+$, $\omega \in [-\pi, \pi]$ for $k = 1, 2, \cdots, K$. α_k is called the *damping factor* of the damped sinusoid with angle frequency ω_k. The larger the damping factor, the faster the amplitude of the sinusoid decays. The observed sequence $y(n)$ is obtained from $x(n)$ corrupted by additive noise $w(n)$, which is assumed to be a complex white Gaussian process. Normally, we have to make sure that $N(N > 2K)$. The observed data can be expressed as

$$y(n) = x(n) + w(n) \qquad \text{for } n = 0, 1, 2, \cdots, N-1. \tag{2}$$

The KT algorithm [6] is one of the most effective algorithms for parameter estimation of damped sinusoidal signals. To estimate the parameters of the damping sinusoidal signals using the KT algorithm, an $(N-L)\times L\,[\min(N-L, L) \geq K]$ conjugate backward prediction matrix and an $(N-L)$-component column vector are first set up as follows:

$$\mathbf{A} = \begin{bmatrix} y^*(1) & y^*(2) & \cdots & y^*(L) \\ y^*(2) & y^*(3) & \cdots & y^*(L+1) \\ \vdots & \vdots & \vdots & \vdots \\ y^*(N-L) & y^*(N-L+1) & \cdots & y^*(N-1) \end{bmatrix},$$
$$\mathbf{h} = \begin{bmatrix} y^*(0) \\ y^*(1) \\ \vdots \\ y^*(N-L-1) \end{bmatrix} \tag{3}$$

where "$*$" stands for the complex conjugate.

To find the frequencies of the damped sinusoids, an L-component prediction coefficient vector $\mathbf{c}$ should be found such that

$$\mathbf{A}\mathbf{c} \approx -\mathbf{h} \tag{4}$$

where $\mathbf{c} = (c_1, c_2, \cdots, c_L)'$ are the backward linear prediction coefficients. Then, $e^{-s_k^*}$ for $k = 1, 2, \cdots, K$ can be estimated by calculating the roots of the prediction polynomial

$$C(z) = 1 + c_1 z^{-1} + \cdots + c_L z^{-L}. \tag{5}$$

Hence, the performance of an algorithm relys on how accurate the estimation of the prediction polynomial is.

To estimate $\mathbf{c}$, the optimum rank K matrix approximation of $\mathbf{A}$ is first made by

$$\hat{\mathbf{A}} = \sum_{k=1}^{K} \sigma_k \mathbf{u}_k \mathbf{v}_k^H \tag{6}$$

where σ_k's ($\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_L$) are the singular values of $\mathbf{A}$, and $\mathbf{u}_k$ and $\mathbf{v}_k$ are the left singular vector and the right singular vector of $\mathbf{A}$ corresponding to the singular value σ_k, and "H" stands for the conjugate transposition. To make the system equation

$$\hat{\mathbf{A}}\mathbf{c} = -\mathbf{h} \tag{7}$$

have a solution, either $\mathbf{h}$ must be in span$\{\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_K\}$ or $\hat{\mathbf{h}}$, which is the projection of $\mathbf{h}$ on span$\{\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_K\}$, must be used instead of $\mathbf{h}$ in (7). In either case, (7) can be written as

$$\hat{\mathbf{A}}\mathbf{c} = -\hat{\mathbf{h}} \tag{8}$$

where

$$\hat{\mathbf{h}} = \sum_{k=1}^{K} (\mathbf{u}_k^H \mathbf{h})\mathbf{u}_k. \tag{9}$$

Since rank$(\hat{\mathbf{A}}) = K \leq L$, (8) is an underdetermined system of equation about $\mathbf{c}$, and there are multiple solutions. The solution minimizing $\|\mathbf{c}\|$ is given by

$$\mathbf{c} = -\sum_{k=1}^{M} \sigma_k^{-1}(\mathbf{u}_k^H \mathbf{h})\mathbf{v}_k, \tag{10}$$

which is the $\mathbf{c}$ in the KT algorithm [6]. It has been proved in [5] and [6] that if $\mathbf{c}$ is estimated using (10), then only K of $C(z)$'s zeros are outside the unit circle, which are signal zeros $e^{-s_k^*}$ for $k = 1, 2, \cdots, K$. The rest of the $L-K$ zeros are inside the unit circle. By means of this property, the desired zeros can be easily identified to estimate the parameters. To obtain optimum performance, the L $(L \geq K)$ is chosen to be larger than $N-L$, usually $3N/4$.

From the above discussion, the KT algorithm uses the low-rank matrix approximation to reduce the noise effect. When the SNR is high and enough data are available, the rank approximation in the KT algorithm will reduce the measurement noise significantly; hence, the KT algorithm in this case will almost attain the CR bound [6]. However, if the SNR is reduced to certain degree, the rank approximation in the KT algorithm is unable to reduce the noise effect efficiently, and moreover, it may introduce an extra perturbation. In that case, the noise threshold appears. Since the noise threshold of the KT algorithm is due to the low-rank matrix approximation, to reduce the noise threshold, the matrix approximation approach employed in the KT algorithm must be improved.

III. Modified KT Algorithm

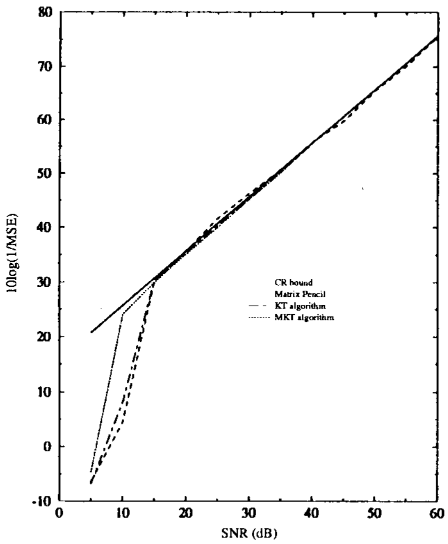
From (3), we can see that the prediction matrix of a data sequence is of Hankel form. Indeed, according to [4] there is a very interesting property that can be summarized as follows.

Lemma: If a data sequence $x(n)$ consists of K distinct sinusoids as in (1), then for any $L(L > K)$, the $L \times L$ prediction matrix $\mathbf{P}_L = [x(i+j)]_{i,\,j=0}^{L-1}$ is a singular Hankel matrix with rank K and full rank $K \times K$ principle minor $\mathbf{P}_K = [x(i+j)]_{i,\,j=0}^{K-1}$. Conversely, for any $L \times L$ singular Hankel matrix $\mathbf{P}_L = [x(i+j)]_{i,\,j=0}^{L-1}$ with rank K, if its $K \times K$ principle minor $\mathbf{P}_K = [x(i+j)]_{i,\,j=0}^{K-1}$ is full rank, then $x(n)$ for $n = 0, 1, \cdots, (2L-2)$ can be uniquely expressed as the summation of K distinct sinusoids as given by (1).

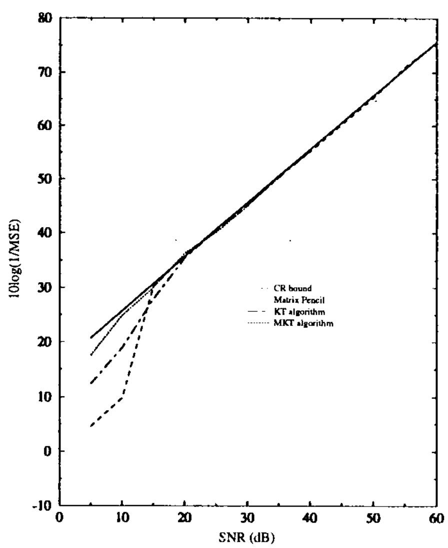
The above lemma reveals a one-to-one correspondence between a data sequence consisting of damped sinusoidal signals and rank-deficient Hankel matrix. Therefore, parameter estimation of damped sinusoidal signals from noisy data is equivalent to performing the low-rank Hankel matrix approximation. More specifically, let $\mathbf{P}_L$ be

TABLE I
MODIFIED KT ALGORITHM

Step 1	Form square prediction matrix $\mathbf{P}_{L_m}$
Step 2	Find $\bar{\mathbf{P}}_{L_m}$ by (13)
Step 3	Find $\hat{\mathbf{P}}_{L_m}$ by (14)
Step 4	Repeat Step 2 and 3 to get estimation of $\hat{y}(n)$
Step 5	Estimate parameters using the KT algorithm to $\hat{y}(n)$

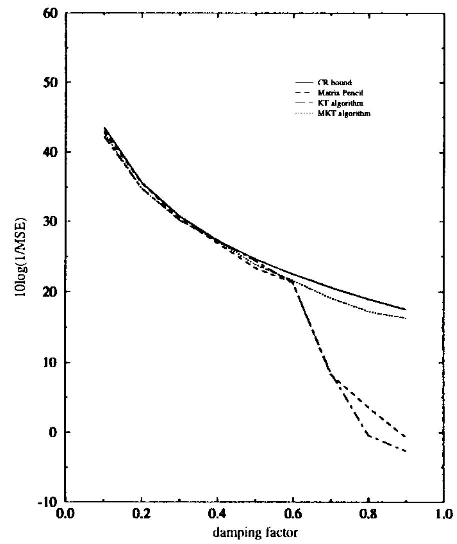


(a)

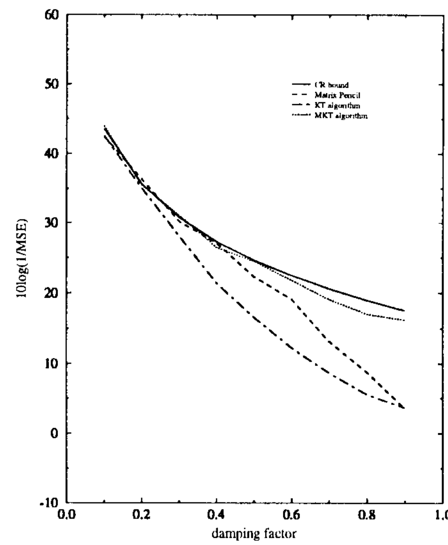


(b)

Fig. 1. MSE of (a) ω and (b) α versus SNR of the matrix pencil, KT, and MKT algorithms obtained in 200 trials when $s = -0.2 + \jmath 2\pi 0.42$ and $N = 24$



(a)



(b)

Fig. 2. MSE of (a) ω and (b) α versus α of the matrix pencil, KT, and MKT algorithms obtained in 200 trials when $s = -\alpha + \jmath 2\pi 0.42$ and $N = 24$.

an $L \times L$ prediction matrix of noisy data $y(n)$

$$\mathbf{P}_L = [y(i+j)]_{i,\,j=0}^{L-1}. \tag{11}$$

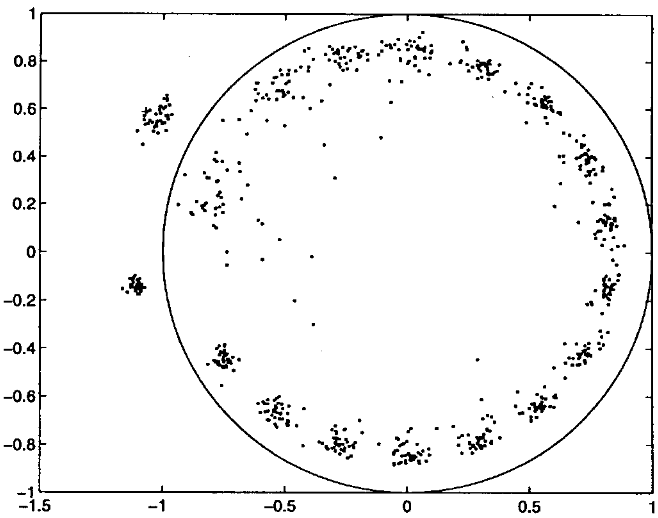
If we can find an $L \times L$ Hankel matrix $\overline{\mathbf{P}} = [\overline{y}(i+j)]_{i,\,j}^{L-1}$ with rank K and a full rank $K \times K$ principle minor, then the parameters of the signal can be uniquely determined from $\overline{\mathbf{P}}$.

For the KT algorithm, only the rank-deficiency characteristics of the prediction matrix is used in matrix approximation. The approximated matrix $\hat{\mathbf{A}}_{N-L,\,L}$ in (6) unfortunately loses the Hankel property. If both the rank and Hankel properties of the matrix are used in the matrix approximation to reduce the noise effect, the performance of the estimation will be improved significantly. The modified KT algorithm introduced here will exploit both properties.

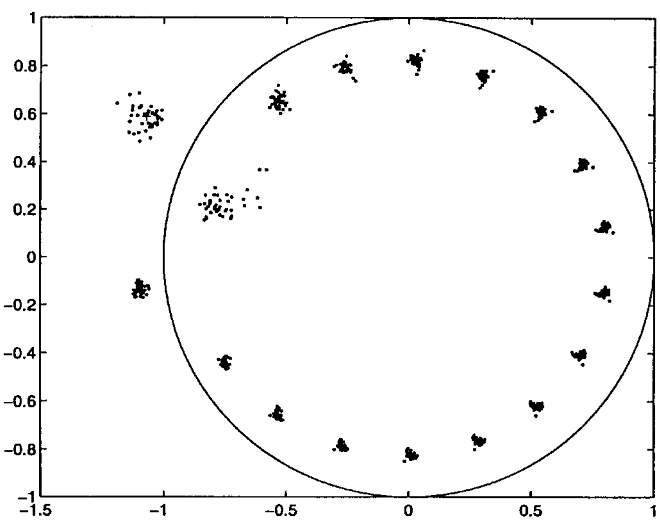
To use the low-rank Hankel matrix approximation to reduce the measurement noise, we first set up a square prediction matrix from the observed noisy data:

$$\mathbf{P}_L = [y(i+j)]_{i,\,j=0}^{L-1}. \tag{12}$$

To make full use of the given data, let $L = \lceil N/2 \rceil$ here. Since there is no analytical low-rank Hankel matrix approximation approach available, an iterative approach for low-rank Hankel matrix approximation



(a)



(b)

Fig. 3. Zeros of $C(z)$ obtained in 40 trials of (a) KT algorithm and (b) MKT algorithm when SNR $= 15$ dB.

is used here. First, an optimum rank K matrix approximation to $\mathbf{P}_L$ is made using the SDV.

$$\overline{\mathbf{P}}_L = [\overline{y}_{i,j}]_{i,j=0}^{L-1} = \sum_{k=1}^{K} \sigma_k \mathbf{u}_k \mathbf{v}_k^H \tag{13}$$

where σ_k for $k = 1, 2, \cdots, K$ are the K largest singular value of $\mathbf{P}_L$ and $\mathbf{u}_k$, and $\mathbf{v}_k$ are corresponding left and right singular vectors. $\overline{\mathbf{P}}_L$ is usually not Hankel. Then, a Hankel matrix $\hat{\mathbf{P}}_L$ is found to minimize $\|\hat{\mathbf{P}}_L - \overline{\mathbf{P}}_L\|_F$, where $\hat{\mathbf{P}}_L$ is given by

$$\hat{\mathbf{P}}_L = [\hat{y}(i+j)]_{i,j=0}^{L-1} \tag{14}$$

and

$$\hat{y}_{i+j} = \frac{1}{N} \sum_{0 \le n, m \le L-1, m+n=i+j} \overline{y}_{n,m} \tag{15}$$

with N being the number of the elements in matrix $\overline{\mathbf{P}}_L$ satisfying $n+m = i+j$ in (13). After this step, the rank of $\hat{\mathbf{P}}_L$ is not necessarily K. A low-rank approximation is used again. The procedures are repeated until a Hankel matrix with only K dominate singular values is obtained. From the approximated Hankel matrix $\hat{\mathbf{P}}_L$, better noise-reduced data $\hat{y}(n)$ can be found. Then, by using the KT algorithm, the parameters of the signal can be obtained from $\hat{y}(n)$. The algorithm is summarized in Table I. The convergence of the above iteration can be proved using the theory of composite property mapping algorithm [2]. In [2], it has been shown that the exponential data satisfy the hypotheses of composite mapping theorem, and therefore, the composite mapping algorithm can be used to reduce the noise effect from the measured exponential data. Since the damped sinusoidal signals form a subset of exponential signals, the same results extend to these kinds of signals. Therefore, the Hankel approximation process applied to reduce the noise effect in damped sinusoidal data will converge to a solution. In what follows, we will prove that the proposed low-rank Hankel approximation can indeed achieve better performance.

Theorem 1: Let $\mathbf{P}_{true} = [x(i+j)]_{i,j=0}$ be the true prediction matrix; then,

$$\|\hat{\mathbf{P}}_L - \mathbf{P}_{true}\|_F \le \|\overline{\mathbf{P}}_L - \mathbf{P}_{true}\|_F. \tag{16}$$

The equality holds only if $\overline{\mathbf{P}}_L$ is Hankel.

Proof:

From (15), direct calculation yields that

$$\begin{aligned} &\frac{1}{N} \sum_{i+j=n,\, 0 \le i,\, j \le n} |x(n) - \overline{y}_{i,j}|^2 \\ &\quad = |x(n) - \hat{y}(n)|^2 \\ &\qquad + \frac{1}{N} \sum_{i+j=n,\, 0 \le i,\, j \le n} |\hat{y}(n) - \overline{y}_{i,j}|^2 \\ &\quad \ge |x(n) - \hat{y}(n)|^2 \end{aligned} \tag{17}$$

where N is the number of elements in matrix $\overline{\mathbf{P}}_L$ satisfying $i+j = n$.

Using the above inequality, a direct calculation yields that

$$\begin{aligned} \|\overline{\mathbf{P}}_L - \mathbf{P}_{true}\|_F^2 &= \|\overline{\mathbf{P}}_L - \hat{\mathbf{P}}_L\|_F^2 + \|\hat{\mathbf{P}}_L - \mathbf{P}_{true}\|_F^2 \\ &\ge \|\hat{\mathbf{P}}_L - \mathbf{P}_{true}\|_F^2. \end{aligned} \tag{18}$$

The above theorem demonstrates that $\overline{\mathbf{P}}_L$ is always more accurate than $\hat{\mathbf{P}}_L$. If the SVD in the iteration procedures can reduce the noise effect efficiently, a better estimation of $\mathbf{P}_{true}$ can be obtained by preserving the Hankel form after each iteration. Hence, the performance of the modified KT algorithm should be better than that of the original KT algorithm. Even though we emphasize the modified KT algorithm in this paper, similar procedures can also be used for the TLS algorithm.

It is worth mentioning that the complexity of the MKT algorithm is in the same order as the KT algorithm. Extra computations in MKT algorithm comes from the Hankel approximation part, which requires several SVD's until the algorithm converges. However, for most practical cases of interest, the Hankel approximation typically converges within a few iterations. For the results we have reported in this correspondence, we have used only two iterations of the algorithm, and we still got very good results. Since the number of

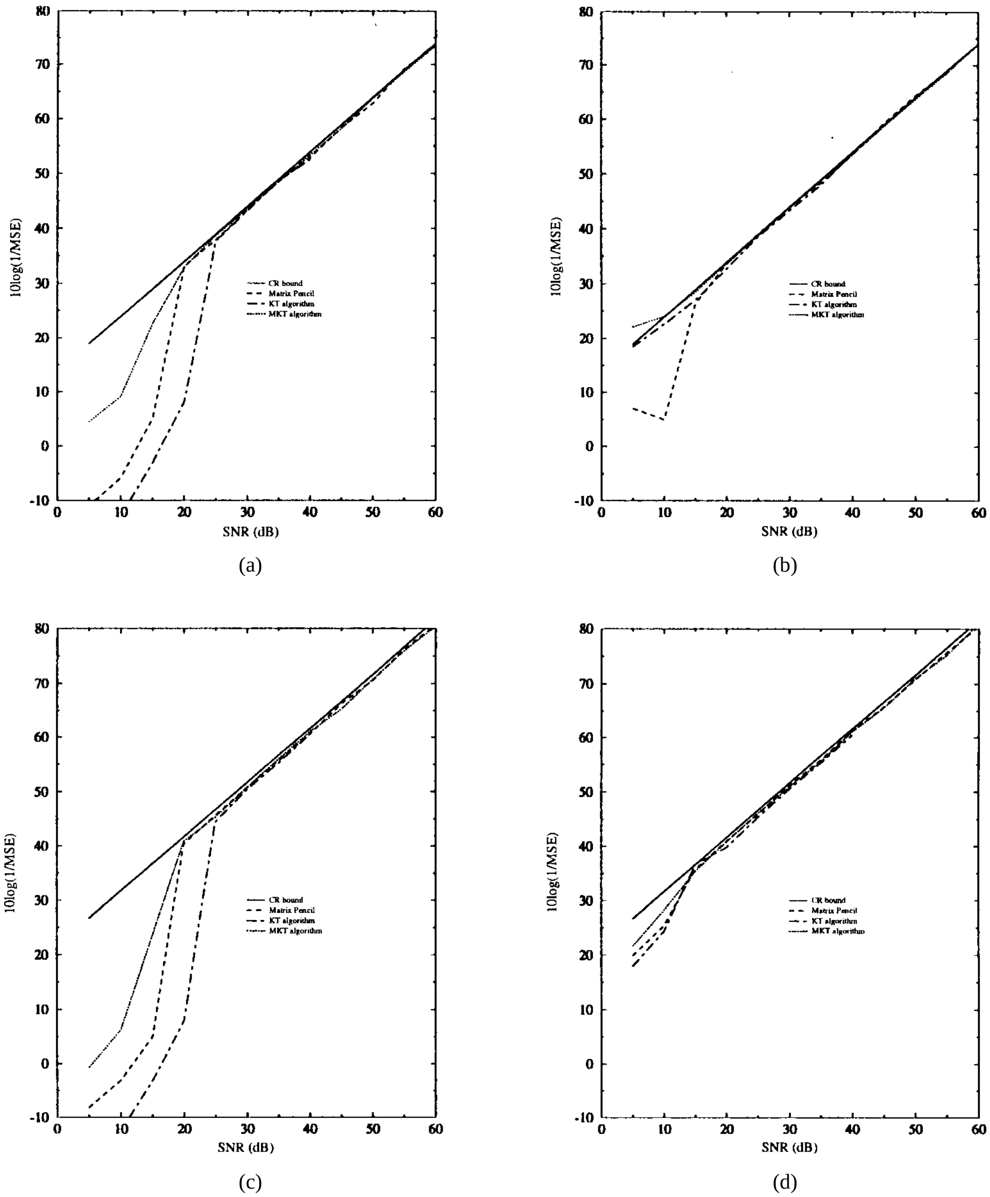


Fig. 4. MSE of (a) ω_1, (b) α_1, (c) ω_2, and (d) α_2 obtained in 200 trials of the matrix pencil, KT, and MKT algorithms when $s_1 = -0.2 + j2\pi 0.42$, $s_2 = -0.1 + j2\pi 0.52$, and $N = 25$.

additional SVD's required in MKT is only two, the complexity of the MKT algorithm remains almost the same as that of the original KT [6] algorithm and the matrix pencil method [3].

IV. Computer Simulation Examples

In this section, we will test the performance of the MKT algorithm and compare it with the KT algorithm and the matrix pencil algorithm [3] by two computer simulation examples.

In our examples, the damped sinusoid is corrupted by complex white Gaussian noise with zero mean and variance σ^2. The SNR used in the examples is the peak signal-to-noise ratio defined as

$$\mathrm{SNR} = 10 \log \left(\frac{1}{2\sigma^2}\right). \tag{19}$$

The performance of the algorithms is measured by the mean square error (MSE). For comparison, we also simulate the performance of Kumaresan–Tufts (KT) algorithm [6], the matrix pencil algorithm [3], and calculate the CR bound using the formula in [6]. In our simulation, $N = 24$ and $L = 18$ for both the KT algorithm and the MKT algorithm.

Example 1: The simulated data are given by

$$x(n) = e^{sn} + w(n) \tag{20}$$

where $s = -\alpha + j\omega$, and $w(n)$ is complex white Gaussian noise and with variance σ^2.

When we fix $s = -0.20 + j2\pi(0.52)$ and change the SNR, the simulation results are shown in Fig. 1(a) and (b) for the MSE of damping factor α and frequency ω, respectively. From this figure, the MSE's of the matrix pencil algorithm, the KT algorithm, and the MKT algorithm are all near CR bound if the SNR is high. We can see that the performance of MKT algorithm follows closely to the CR bound in estimating damping factor α, and the noise threshold in estimating the frequency is about 3–4 dB below those of the KT algorithm and the matrix pencil algorithm.

When we fix SNR $= 20$ dB, $\omega = 2\pi(0.52)$, and we change the damping factor α. The MSE's of α and ω are shown in Fig. 2(a) and (b). From the figure, we can see that when the damping factor is less than 0.2, the performance of KT, MKT, and the matrix pencil algorithms is near the CR bound. If $\alpha \geq 0.55$, the KT algorithm and the matrix pencil algorithm are unable to estimate the parameters, whereas the MKT algorithm can still estimate the parameters effectively.

Example 2: The simulated data are generated by

$$y(n) = e^{s_1 n} + e^{s_2 n} + w(n) \tag{21}$$

where $s_1 = -0.2 + \jmath 2\pi(0.42)$, $s_2 = -0.1 + \jmath 2\pi(0.52)$, and $w(n)$ is complex white Gaussian noise with variance σ^2. This example is from [6].

The zeros of prediction polynomials are shown in Fig. 3. From the figure, it is clear that the MKT algorithm has less bias and smaller variance than the KT algorithm. The MSE's of α_1, ω_1, α_2, and ω_2 for the matrix pencil algorithm, the KT algorithm, and the MKT algorithm are shown in Fig. 4(a)–(d). From these figures, the noise threshold of the MKT algorithm is about 3–4 dB lower than that of KT algorithm.

V. Conclusions

The reduced-rank matrix approximation has been an effective tool in many branches of signal processing. In this paper, we demonstrate that if we can also preserve the matrix structure, such as the Hankel structure for the case of parameter estimation of damped sinusoidal signals, the performance can be further improved. Specifically, we presented the MKT algorithm to estimate the parameters of damped sinusoidal signals. The MKT algorithm exploits both reduced rank and Hankel properties of the prediction matrix. Compared with the original KT algorithm and the matrix pencil algorithm, it has lower noise threshold and is able to estimate the parameters of signal with large damping factors. Hence, preserving the Hankel structure in reduced-rank matrix approximation improves the performance significantly. The proposed approach and concept presented in this article can also be extended to the general area of reduced rank signal processing [10], where structural low-rank approximation can be very effective in performance improvement.

References

[1] Y. Bresler and A. Macovski, "Exact maximum likelihood parameter estimation of superimposed exponential signals in noise," *IEEE Trans. Acoust., Speech, Signal Processing,* vol. ASSP-34, pp. 1081–1089, Oct. 1986.

[2] J. A. Cadzow "Signal enhancement—A composite property mapping algorithm," *IEEE Trans. Acoust., Speech, Signal Processing,* vol. 36, pp. 49–62, Jan. 1988.

[3] Y. Hua and T. K. Sarkar "Matrix pencil method for estimation parameters of exponentially damped/undamped sinusoids in noise," *IEEE Trans. Acoust., Speech, Signal Processing,* vol. 38, pp. 814–824, May 1990.

[4] I. S. Iohvidov, *Hankel and Toeplitz Matrices and Forms.* Boston, MA: Birkhäuser, 1982.

[5] R. Kumaresan, "On the zeros of the linear prediction-error filter for deterministic signals," *IEEE Trans. Acoust., Speech, Signal Processing,* vol. ASSP-31, pp. 217–220, Feb. 1983.

[6] R. Kumaresan and R. W. Tufts, "Estimation the parameters of exponentially damped sinusoids and pole-zero modeling in noise," *IEEE Trans. Acoust., Speech, Signal Processing,* vol. ASSP-30, pp. 833–840, Dec. 1982.

[7] B. Porat and B. Friedlander,"On the accuracy of the Kumaresan–Tufts method for estimating complex damped exponentials,"*IEEE Trans. Acoust., Speech, Signal Processing,* vol. ASSP-35, pp. 231–235, Feb. 1987.

[8] M. A. Rahman and K. B. Yu, "Total least squares approach for frequency estimation using linear prediction," *IEEE Trans. Acoust., Speech, Signal Processing,* vol. ASSP-35, pp. 1440–1454, Oct. 1987.

[9] V. U. Reddy and L. B. Biradar, "SVD-based information theoretic criteria for detection of the number of damped/undamped sinusoids and their performance analysis," *IEEE Trans. Signal Processing,* vol. 41, pp. 2872–2881, Sept. 1993.

[10] L. L. Scharf "The SVD and reduced rank signal processing," *Signal Processing,* vol. 41, pp. 113–133, 1991.