AN APPROACH TO ARMA SYSTEM IDENTIFICATION AT A VERY LOW SIGNAL-TO-NOISE RATIO

S. A. Fattah, Student Member, IEEE, W. -P. Zhu, Senior Member, IEEE, and M. O. Ahmad, Fellow, IEEE

Centre for Signal Processing and Communications, Dept. of Electrical and Computer Engineering
Concordia University, Montreal, Quebec, Canada H3G 1M8

ABSTRACT

A new approach for the identification of minimum-phase autoregressive moving average (ARMA) systems in the presence of heavy noise is presented in this paper. A damped sinusoidal (DS) model for the autocorrelation function of a noise-free ARMA signal is proposed to estimate the AR parameters, which overcomes the failure of conventional correlation based techniques in estimating the AR parameters of an ARMA system at a very low signal-to-noise ratio (SNR). The MA parameters of the ARMA system are then estimated by using Durbin's method along with an optimum order selection criterion. Both white noise and periodic impulse train excitations are considered for the application of the proposed method to system identification as well as to speech processing. Computer simulations are carried out based on both synthetic ARMA systems and natural speech signals, showing superior identification results even at an SNR of −5 dB for which most of the existing methods would fail.

1. INTRODUCTION

Autoregressive moving average (ARMA) models have been extensively studied in various fields, such as speech processing, biomedical signal processing, control engineering, economics and others. Numerous methods have been developed for estimating the parameters of an ARMA model [1]-[5]. The maximum likelihood methods, although known to be asymptotically consistent, suffer from convergence problems [1]. As an alternative, a great deal of research effort has been conducted on separate identification of AR and MA parts of an ARMA model. The estimation of the AR parameters, if performed first, is of crucial importance for the MA parameter estimation, since a poor AR estimation result would severely affect the estimation accuracy of the MA part. The modified/extended Yule-Walker equations (MYWE) have been extensively used to identify the AR parameters of ARMA systems [1]. In the presence of observation noise, however, the performance of MYWE methods would be significantly degraded due to the non-ideal nature of the autocorrelation function (ACF) of the additive noise. Durbin's method has been widely used for MA parameter estimation [1]. But its performance in a noisy environment depends on the accuracy of the estimated intermediate AR (IAR) model (both on its order and parameters). In the ARMA-cepstrum recursion (ACR) method [2], cepstral coefficients of the ARMA system are used to estimate the MA parameters. However, these cepstral coefficients are very sensitive to noise. The lattice filter method presented in [3] gives quite a good estimation result for an AR system with noise or the AR part of an ARMA system without noise, but it fails to yield an accurate estimation for the MA part in presence of noise. The ARMA system identification methods, reported in [4]-[5], have attempted to minimize a suitably chosen cost function. Yet they are not able to provide an unbiased estimation at a low signal-to-noise (SNR).

In this paper, the identification problem of an ARMA system with heavy additive noise is addressed. The damped sinusoidal (DS) model, recently proposed in [6] for ACF of a noise-free AR signal, is extended for the modeling of a noise-free ARMA signal so that the AR parameters of the ARMA system can be estimated through a sequential approximation of the ACF of the observed noisy signal. Both finite- and infinite-duration white noise excitations are considered in the derivation of the DS model. Finite-length periodic impulse train excitation is also considered for the application of the proposed identification method to speech processing. The MA parameters of the ARMA system are estimated using Durbin's method in conjunction with an effective selection criterion for the optimum order. The objective of this paper is to present an efficient identification approach for ARMA systems with a very low SNR.

2. PROBLEM FORMULATION

A causal, stable, and linear time-invariant ARMA($P,\ Q$) process can be characterized by

$$x(n) = -\sum_{k=1}^{P} a_k x(n-k) + \sum_{k=0}^{Q} b_k u(n-k), \quad b_o = 1 \tag{1}$$

where a_k and b_k are system parameters, P and Q the known system order, and $u(n)$ white Gaussian noise excitation with a zero-mean and variance σ_u^2. The ACF of $x(n)$ can be expressed as

$$R_x(\tau) = \frac{1}{N} \sum_{n=0}^{N-1-|\tau|} x(n)x(n+|\tau|) \tag{2}$$

where N is the number of data points. In the noise-free case, the AR parameters can be estimated using $R_x(\tau)$ [1]. In the presence of noise, the observed signal $y(n)$ can be written as

$$y(n) = x(n) + v(n) \tag{3}$$

where $v(n)$ is a white Gaussian noise with zero-mean and variance σ_v^2. It is usually assumed that $v(n)$ is independent of $u(n)$. From (2) and (3), one can obtain

$$R_x(\tau) = \begin{cases} R_y(\tau) - \sigma_v^2, & \tau = 0 \\ R_y(\tau), & \tau \neq 0 \end{cases} \tag{4}$$

where $R_y(\tau)$ is the ACF of $y(n)$. Although (4) provides a simplest way of computing $R_x(\tau)$, it is not suitable for the case of very low SNR due to the non-ideal nature of the ACF of $v(n)$. In most of the existing methods it is usually assumed that the signal and additive noise are uncorrelated. But at a very low SNR this assumption is violated and the estimate of $R_y(\tau)$ using the conventional method gives significant error at all lags other than the zero lag. Hence, many researchers have pursued a separate estimation of the AR and MA parts of the ARMA system in order to improve the identification performance in a noisy environment. In what follows, we will propose a new approach for estimating the parameters of a heavily noise-corrupted ARMA system.

3. ESTIMATION OF AR PARAMETERS

The transfer function of the ARMA system of order (P, Q) described by (1) can be expressed as

$$H(z)=\frac{B(z)}{A(z)}=\frac{1+\sum_{i=1}^{Q} b_i z^{-i}}{1+\sum_{i=1}^{P} a_i z^{-i}}=\frac{\prod_{i=1}^{Q}(1-z_i z^{-1})}{\prod_{i=1}^{P}(1-p_i z^{-1})}=\sum_{k=1}^{P}\frac{\eta_k}{1-p_k z^{-1}} \tag{5}$$

where p_k and z_k denote, respectively, the kth pole and zero of the system, and η_k is the partial fraction coefficient. It is assumed in this paper that all zeros and poles are of the first-order. Using (5), the output of the noise-free ARMA system can be written as

$$\begin{aligned} x(n) &= u(n) * h(n) \\ &= \sum_{k=1}^{P}\sum_{m=0}^{n}\eta_k u(m)\, p_k^{n-m} \end{aligned} \tag{6}$$

where $h(n)$ represents the impulse response of the ARMA system, as given by

$$h(n)=\sum_{k=1}^{P}\eta_k p_k^n \tag{7}$$

In conventional system identification problems, only infinite-duration (a relatively large duration N, denoted as N_i) white Gaussian noise excitation is considered. In speech processing applications, however, finite-duration (a relatively small duration N, denoted as N_f) excitation is also required. In order to provide a more accurate description of the speech spectra, especially for nasal or nasalized sounds, the vocal tract is represented as an ARMA model. In linear predictive coding (LPC), speech is considered as the output of a slowly time-varying ARMA filter excited by a periodic impulse train for voiced speech or a random white noise for unvoiced speech [7]. By using (6) in (2) and assuming a white Gaussian noise excitation, one can deduce ACF for both infinite- and finite-duration excitations as given by

$$R_x(\tau)=\sum_{k=1}^{P}\psi_k p_k^{\tau}, \quad (\tau=0, 1, ..., M) \tag{8}$$

where

$$\psi_k=\chi\left[\frac{\eta_k^2}{1-p_k^2}+\sum_{l=1, l\neq k}^{P}\frac{\eta_k\eta_l}{1-p_k p_l}\right], \tag{9}$$

Here, χ is a constant that depends on the variance of the input excitation. In the case of infinite duration excitation, we have assumed $M << N_\mathrm{i}$ in obtaining (8).

A periodic impulse train excitation with period T and duration N_f can be represented as

$$u(n)=\sum_{m=0}^{\lambda}\delta(n-mT), \ \lambda T\le n, \lambda=\lfloor (n/T)\rfloor \tag{10}$$

where $\lfloor \zeta \rfloor$ represents the largest integer less than or equal to ζ. With this excitation, $x(n)$ can be derived as

$$x(n)=\sum_{k=1}^{P}\sum_{m=0}^{\lambda}\eta_k p_k^{n-mT}, \ \lambda T\le n, \lambda=\lfloor (n/T)\rfloor \tag{11}$$

Using (11) in (2), one can obtain an ACF of $x(n)$ for the finite-duration periodic impulse train excitation. The expression will be similar to that given by (8) with a different expression for the constant term ψ_k.

In general the ACF given by (8) can be rewritten as

$$R_x(\tau)=\sum_{l=1}^{\theta} r_l^{\tau}[\alpha_l\cos(\omega_l\tau)+\beta_l\sin(\omega_l\tau)], \ (\tau=0, 1, ..., M) \tag{12}$$

where θ represents the number of real poles plus the number of complex conjugate pole pairs, r_l and ω_l are the magnitude and argument of p_l, respectively, and α_l and β_l are the constants which can be expressed in terms of ψ_k. The ACF given by (12) is referred to as the damped sinusoidal (DS) model for the noise-free ARMA signal.

Each of the θ component functions of the DS model can be estimated sequentially from M nonzero lags of $R_y(\tau)$ [6]. The parameters of the lth component function, are chosen such that the total squared error between the ($l-1$)th residual function and the lth component function is minimized. The lth residual function is defined as

$$\Re_l(\tau)=\Re_{l-1}(\tau)-r_l^{\tau}F_l(\tau), \ (l=1,2,...,\theta-1) \tag{13}$$

with

$$\Re_0(\tau)=R_y(\tau)$$
$$F_l(\tau)=\alpha_l\cos(\omega_l\tau)+\beta_l\sin(\omega_l\tau)$$

The total squared error is then defined as

$$J_l^{(i)}=\sum_{\tau=1}^{M}\left|\Re_{l-1}(\tau)-(r_l^{(i)})^{\tau}F_l^{(i)}(\tau)\right|^2, \ (l=1,2,...,\theta-1) \tag{14}$$

The superscript "(i)" in (14) denotes the iteration index. At each iteration, $\alpha_l^{(i)}$ and $\beta_l^{(i)}$ can be determined by minimizing $J_l^{(i)}$ given the values of $r_l^{(i)}$ and $\omega_l^{(i)}$. The values of r_l and ω_l corresponding to the global minimum of J_l, are selected to be the estimate of a pair of complex conjugate poles $re^{(\pm j\omega_l)}$, if $0 < \omega_l < \pi$. If $\omega_l = 0$ or π, the estimate represents a real pole. Once all the poles have been estimated, the AR part of the ARMA system is identified.

Unlike the conventional correlation-based techniques, where $R_y(\tau)$ is directly employed for estimation, the proposed method is able to yield a better estimate of $R_x(\tau)$ at a very low SNR due to the indirect noise compensation implemented in the approximation of the DS model using M nonzero lags of $R_y(\tau)$.

4. ESTIMATION OF MA PARAMETERS

In Durbin's method, the parameters of intermediate AR (IAR) model are first estimated and then used to calculate the MA parameters. In the presence of noise, $y(n)$ is filtered by an estimated AR system in order to obtain the MA residual of the ARMA signal. It can be shown that the resulting residual is equivalent to a noise-corrupted MA signal, and a noisy IAR sequence will be available for the estimation of IAR parameters.

However, due to the rapid decay of the ACF of the MA signal, the autocorrelation-based noisy AR identification techniques exhibit poor performance at a low SNR. Therefore, a lattice filter algorithm has been applied to the identification problem of noisy AR systems, which works directly in signal domain and yields sufficient accuracy [3].

For an MA signal, the optimum IAR order based on the combined information criterion (CIC) is given as [8]

$$L = 2K + Q\,,\ \ K = \arg\min_{m}\ CIC(m) \tag{15}$$

The CIC based order selection principle is defined as

$$CIC(m) = \ln\{\varepsilon(m)\} + \max\left[\left(\prod_{i=1}^{m}\frac{1+1/(N+1-i)}{1-1/(N+1-i)} - 1\right), \left(\sum_{i=1}^{m}\frac{3}{N+1-i}\right)\right] \tag{16}$$

where $\varepsilon(m) = \varepsilon(0)\prod_{i=1}^{m}(1-k_i^2)$, $\varepsilon(0) =$ variance of MA signal. Here, k_i is the PARCOR coefficient that needs to be calculated in the lattice filter algorithm. Note that the order L determined by (15) is not sufficient in a heavy noisy situation, since the resulting residual can be viewed as a noisy IAR signal. We propose a higher order for the IAR model, namely, $L' = 2K+2P+Q$. Our extensive experimentation has shown that this choice gives sufficient estimation accuracy without requiring further increase in the IAR order.

5. SIMULATION RESULTS

In this section, the proposed method is simulated and compared with some of the existing ones including the ACR method [2] and the order selective Durbin's (OSD) method [1], showing various identification results for both synthetic and natural signals. In the OSD method, L', instead of L [8], is used as the IAR order for better estimation results. In both OSD and ACR methods, least squares MYWE [1] are used for estimating the AR parameters.

5.1. Synthetic systems

An input data sequence is generated according to (1) and (3) with N = 4000 samples and $\sigma_u^2 = 1$. To determine the DS model parameters in a computationally efficient manner, an initial estimate of $\omega_l \in [0, \pi]$ is first obtained by searching the full range at a resolution of 0.01. Then around the initial estimate, a finer search at a resolution of 0.001 is performed. In both cases, we assume that $0.5 \le r_l \le 0.999$ with a resolution of 0.001. The number of lags for the ACF is set to be $M = 10P$. Each experiment contains 20 independent runs in order to obtain the average estimate. Table 1 shows the average values of the estimated parameters of an ARMA(4,3) system and their standard deviations from mean (SDM) as well as those from true values (SDT), where three methods have been employed and two SNR values, 10 dB and −5 dB, have been considered. It is seen that the performance of all the three methods is comparable at 10 dB. However, at the SNR of −5 dB, only the proposed method is able to estimate the parameters whereas the other two have completely failed. Moreover, the proposed method exhibits a good consistency as reflected by the small values of SDT and SDM at both SNR levels.

Fig. 1 depicts the pole-zero estimates of an ARMA(6,4) system using the proposed method at SNR = −5 dB. The true parameters of the system are set as $\{a_k\}$ = [1 −0.6034 0.3245 0.3458 0.3644 −0.5167 0.8155] and $\{b_k\}$ = [1 −0.62 0.52 −0.03 0.25].

Table 1. Performance comparison of different methods (Estimated parameters with SDM and then SDT are shown)

SNR dB	True Values	Proposed method	OSD method	ACR method
10	a_1 = −2.5950	−2.5813 ±0.0697 ±0.0711	−2.4978 ±0.0938 ±0.1351	−2.4978 ±0.0938 ±0.1351
	a_2 = 3.3390	3.3286 ±0.1368 ±0.1372	3.1262 ±0.2113 ±0.2999	3.1262 ±0.2113 ±0.2999
	a_3 = −2.2000	−2.2096 ±0.1419 ±0.1422	−2.0169 ±0.1817 ±0.2579	−2.0169 ±0.1817 ±0.2579
	a_4 = 0.7310	0.7281 ±0.0793 ±0.0793	0.6605 ±0.0688 ±0.0985	0.6605 ±0.0688 ±0.0985
	b_1 = −2.0922	−2.0557 ±0.0706 ±0.0794	−1.8542 ±0.3254 ±0.4031	−2.0593 ±0.0951 ±0.1006
	b_2 = 1.8438	1.8467 ±0.1010 ±0.1010	1.4638 ±0.5151 ±0.6401	1.9128 ±0.1755 ±0.1886
	b_3 = −0.6480	−0.6764 ±0.0535 ±0.0605	−0.4645 ±0.2871 ±0.3408	−0.8338 ±0.1070 ±0.2144
−5	a_1 = −2.5950	−2.5701 ±0.1077 ±0.1106	−0.6582 ±0.1374 ±1.9416	−0.6582 ±0.1374 ±1.9416
	a_2 = 3.3390	3.2629 ±0.1986 ±0.2126	0.1663 ±0.1785 ±3.1777	0.1663 ±0.1785 ±3.1777
	a_3 = −2.2000	−2.1185 ±0.2198 ±0.2344	0.1212 ±0.1546 ±2.3263	0.1212 ±0.1546 ±2.3263
	a_4 = 0.7310	0.6793 ±0.1234 ±0.1338	0.1279 ±0.1091 ±0.6129	0.1279 ±0.1091 ±0.6129
	b_1 = −2.0922	−2.0674 ±0.2007 ±0.2022	−0.5900 ±0.5572 ±1.6022	−0.5560 ±0.1384 ±1.5424
	b_2 = 1.8438	1.9093 ±0.2133 ±0.2231	0.1747 ±0.4603 ±1.7314	0.0720 ±0.1645 ±1.7794
	b_3 = −0.6480	−0.6991 ±0.1223 ±0.1325	−0.0296 ±0.2646 ±0.6726	0.0506 ±0.1249 ±0.7097

Obviously, the estimation accuracy for both poles and zeros are very good.

5.2. Natural speech

To investigate the performance of the proposed identification technique in a natural speech signal, we have considered some English phonemes from TIMIT standard database with a sampling frequency of 16 KHz. Note that the pre-filtering is not performed in order to see the accuracy of the frequency estimation over the entire range. Fig. 2 shows a comparison of the spectral analysis results for different SNR values obtained from different methods

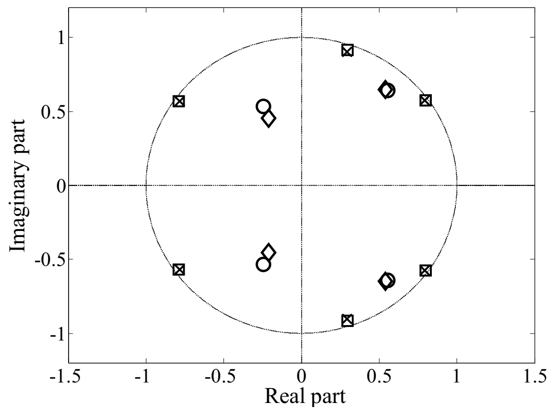

Fig. 1. Estimated poles and zeros of an ARMA(6,4) system at SNR = −5 dB (“×”: true poles, “○”: true zeros, “□”: estimated poles, and “◊”: estimated zeros).

considering an ARMA(8,4) model, where a naturally spoken nasal sound /*m*/ of the word “*him*”, uttered by a female speaker was taken. In this experiment, the synthesized speech is produced from the average estimate of the vocal tract parameters. The excitation period or pitch (T) is estimated according to the scheme given in [9] and the excitation gain is adjusted based on the RMS power level and the peak power spectral density (PSD). The ACF lags are considered up to $T/2$. According to the general behavior of the vocal tract parameters, r_l is searched within $0.8 \leq r_l \leq 0.999$ and search range for ω_l can be decreased from the fundamental knowledge of formant bands. It is seen from Fig. 2 that only the proposed method is able to identify the formant location quite accurately both at high and at very low SNRs.

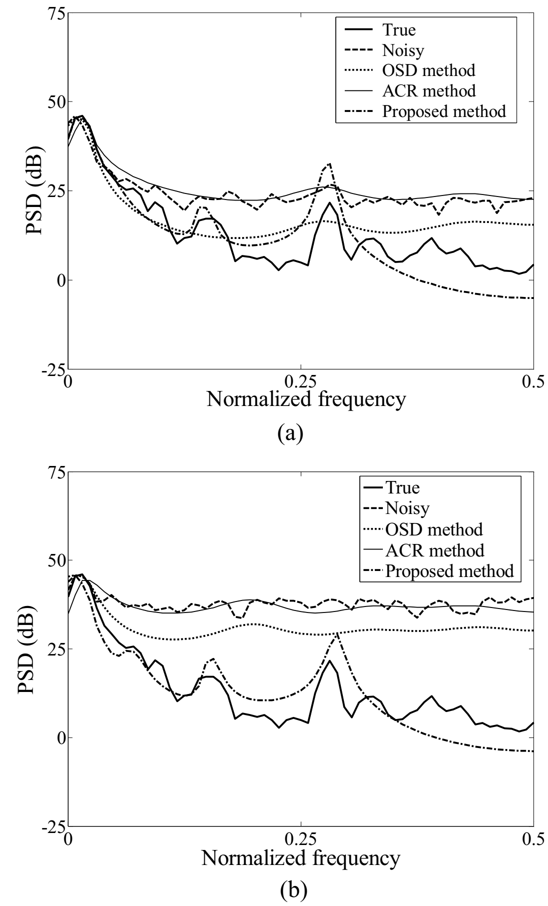

Fig. 2. Spectrum comparison of a natural nasal sound /*m*/ taken from a female utterance “*him*”. (a) SNR = 10 dB. (b) SNR = −5 dB.

6. CONCLUSION

A two-step identification approach for noisy ARMA systems has been proposed. The AR parameters have been estimated using the proposed damped sinusoidal model for the ACF of a noise-free ARMA signal along with a sequential approximation of the ACF of the observed noisy signal. The estimation of MA part has been carried out by employing Durbin’s method with an optimum IAR order selection. It has been shown that the proposed approach is much superior to some of the existing methods and is able to identify accurately a noisy ARMA system at a very low SNR level.

7. REFERENCES

[1] S. M. Kay, *Modern Spectral Estimation: Theory and Application*, Englewood Cliffs, NJ: Prentice-Hall, 1988.

[2] A. Kaderli and A. S. Kayhan, “Spectral estimation of ARMA processes using ARMA-cepstrum recursion,” *IEEE Signal Processing Lett.*, vol. 7, no. 9, pp. 259 – 261, Sept. 2000.

[3] S. Li and B. W. Dickinson, “Application of the lattice filter to robust estimation of AR and ARMA models,” *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 36, no. 4, pp. 502 – 512, April 1988.

[4] J. L. Rojo-Alvarez, M. Martinez-Ramon, M. Prado-Cumplido, A. Artes-Rodriguez, and A. R. Figueiras-Vidal, “Support vector method for robust ARMA system identification,” *IEEE Trans. Signal Processing*, vol. 52, no. 1, pp. 155 – 164, Jan. 2004.

[5] S. Lu, K. H. Ju, and K. H. Chon, “A new algorithm for linear and nonlinear ARMA model parameter estimation using affine geometry,” *IEEE Trans. Biomed. Eng.*, vol. 48, no. 10, pp. 1116 – 1124, Oct. 2001.

[6] M. K. Hasan, S. A. Fattah, and M. R. Khan, “Identification of noisy AR systems using damped sinusoidal model of autocorrelation function,” *IEEE Signal Processing Lett.*, vol. 10, no. 6, pp. 157 – 160, June 2003.

[7] D. O’Shaughnessy, *Speech Communications: Human and Machine*, IEEE Press, NY, second ed., 2000.

[8] P. M. T. Broersen, “Facts and fiction in spectral analysis,” *IEEE Trans. Instrum. Meas.*, vol. 49, no. 4, pp. 766 – 772, Aug. 2000.

[9] C. Shahnaz, M. Rokanuzzaman, M. M. Adnan, B. P. Shrestha, and S. A. Fattah, “A time-frequency domain approach for pitch-peak picking of noisy speech,” in *Proc. Eur. Signal Processing Conf. (EUSIPCO 04)*, Viena, Austria, Sept. 2004, pp. 1015 – 1018.