PHARUS

International Journal of
Information Sciences and Computer Engineering

journal homepage: http://www.ijisce.org

A Joint Iterative Estimation of Noise Variance and AR Parameters

Jonah Gamba[a,*], Tetsuya Shimamura[b], Shuji Kawasaki[a], Masakazu Higuchi[a], Hitomi Murakami[a]

[a]*Department of Information and Computer Science, Seikei University, Musashino, Tokyo, 30332 Japan*
[b]*Department of Information and Computer Sciences, Saitama University, Saitama 338-8570 Japan*

***Abstract*– The problem of determining autoregressive (AR) parameters from observations corrupted by stationary white Gaussian noise without a priori knowledge of the noise variance is addressed. We propose a new approach in which the noise variance and AR parameters are jointly and iteratively estimated from low-order Yule-Walker equations. This approach avoids using unreliable high-order Yule-Walker equations (HOYWEs) or over-determined Yule-Walker equations (ODYWEs). For short observations, noise-compensated data extrapolation (NCDE) is employed. Simulation results demonstrate the effectiveness of the proposed approach.**

Keyword: Autoregressive; noise variance; Yule-Walker equations; data extrapolation.

1. Introduction

AR parameter estimation is a well-known problem [1, 2]. It finds application in many fields including speech processing [3], communication systems [4], geophysical exploration [5], and radar imaging [6, 7]. In these cases, an observed white-noise-corrupted data sequence is modeled as an AR process. Commonly used AR parameter estimation techniques that include the autocorrelation method, the least squares method and singular value decomposition can be found in the literature [8, 9, 10, 11, 12, 13, 14]. We briefly give a overview of some of these methods.

One effective autocorrelation-based method was presented in [9]. It relates the autocorrelation function (ACF) poles to those of the observed data sequence using a recursive procedure. This method is, however, only effective when the poles are very close to the unit circle. A method which requires explicit noise variance estimation by first pre-filtering the observed data sequence and then using the improved least squares (ILS) method for parameter estimation was presented in [10], but it performs best at low noise levels. A subspace-based approach [14] that employs the quadratic eigenvalue problem to estimate the noisy AR parameters gives better results than the YWE-based methods. However, this method requires long data sequences. In short, these traditional methods of AR parameter estimation require a combination of high model order, long data sequences and high signal-to-noise ratios (SNRs).

It is well-known that the most reliable AR parameter estimates can be obtained from LOYWEs [8]. However, noisy LOYWEs are known to be nonlinear, hence the need to employ HOYWEs or ODYWEs for AR parameter estimation [12, 13] . However, HOYWEs or ODYWEs are unreliable since longer correlation lags are needed [14]. We aim to overcome this noise-related shortcoming, and at the same time use LOYWEs for AR parameter estimation. This is achieved through a recently proposed iterative noise variance estimation (INVE) method [15] from which both the noise variance and the AR parameters are obtained from the noisy LOYWEs. The INVE method is motivated by the desire to use only the LOYWEs through iterative reduction of the Euclidean distance between the noisy AR parameters and noise-free AR parameters.

We also consider the case in which the available data sequence can be variable. For simplicity, we assume that the length of the data sequence can either be short or long, depending on the application. When a short data sequence is available for processing, data extrapolation can be useful to further enhance the parameter estimates [16, 17]. We refer this method to as the noise-compensated data extrapolation (NCDE) approach. In this paper, data extrapolation is defined as the process of extending an observed data sequence by linear prediction [17]. We also assumed that the order of the AR process is either known or can be computed. The main advantage of the proposed method is the ability to obtain both noise variance and AR parameters simultaneously and its adaptability to both short and long data sequences.

This paper is organized as follows. In Section 2 the noisy AR model is given. The bias in AR parameters introduced by additive noise is highlighted. This section also contains the main ideas involved in noise variance estimation and noise-compensated data extrapolation. A flowchart is given to illustrate the proposed method. In Section 3 simulation results are presented. The concluding remarks end the paper.

2. The Noisy AR Parameter Problem and the Proposed Solution

This section first presents the noisy AR model. A flowchart is then given to illustrate the steps involved in the proposed solu-

*Corresponding author:
Email address: jmgamba@gmail.com, Ph: +81 422 37 3974

tion.

2.1. The noisy AR model

A stationary AR process $x(n)$ of order p is defined by $\sum_{i=0}^{p} a(i)x(n-i) = e(n)$, where $e(n)$ is an uncorrelated driving white noise sequence of variance σ_e^2 and $a(i)$s are the noise-free AR parameters ($a(0) = 1$). The autocorrelation function (ACF) at lag k for $x(n)$ is defined by $r_{xx}(k) = E[x(n)x(n+k)]$ where E is the expectation operator. The $r_{xx}(k)$ is given in the above case by $r_{xx}(k) = -\sum_{i=1}^{p} a(i)r_{xx}(k-i) + \delta(k)\sigma_e^2$, $k \geq 0$, where $\delta(k)$ is the Kronecker delta function. In the presence of noise, the observed data sequence becomes $y(n) = x(n) + w(n)$, where $w(n)$ is assumed to the zero-mean additive white Gaussian noise of variance σ_w^2 . The ACF for $y(n)$ is similarly defined as that for $x(n)$ and denoted by $r_{yy}(k)$.

The commonly used YWE's [8] are

$$\mathbf{R}_{xx}\mathbf{a} = -\mathbf{r}_x \quad (1)$$

$$\mathbf{R}_{yy}\hat{\mathbf{a}} = -\mathbf{r}_y \quad (2)$$

where $\mathbf{R}_{xx}$ and $\mathbf{R}_{yy}$ are $p \times p$ autocorrelation matrices (ACM's) of the data sequences $x(n)$ and $y(n)$ respectively. The column vectors on the right hand sides of (1) and (2) are $\mathbf{r}_x^T = [r_{xx}(1) \cdots r_{xx}(p)]$ and $\mathbf{r}_y^T = [r_{yy}(1) \cdots r_{yy}(p)]$. The T denotes the transposition operation. The $p \times 1$ vectors $\mathbf{a}$ and $\hat{\mathbf{a}}$ are the noise-free and noisy solutions to the YWE's respectively. The AR parameter estimates from (2) are biased since $r_{yy}(k) = r_{xx}(k) + \delta(k)\sigma_w^2$. Using (1) and (2), the following relationship between ACM's is validated for noise compensation,

$$\mathbf{R}_{xx} = \mathbf{R}_{yy} - \sigma_w^2\mathbf{I} \quad (3)$$

where $\mathbf{I}$ is a $p \times p$ identity matrix.

2.2. The proposed solution

Figure 1 gives a flowchart of the proposed method. In the flowchart, the noise variance estimation step utilizes equations (1)–(3) to get the functional [15]

$$f(\alpha) = ||\hat{\mathbf{a}}|| - ||\tilde{\mathbf{a}}(\alpha)|| \quad (4)$$

where α is a parameter that gives an estimate of σ_w^2 and $\tilde{\mathbf{a}}(\alpha)$ corresponds to the solution of noise-compensated low-order Yule-Walker equations. The value of α that gives the minimum of $f(\alpha)$ results in the noise variance estimate. The functional $f(\alpha)$ is assumed to have a minimum since we consider values of α between 0 and the autocorrelation of the observations at lag 0, i.e. $r_{yy}(0)$. We also constrain the solution such that the autocorrelation function is preserved. This property of the functional $f(\alpha)$ was described analytically in [15] and forms the basis of the INVE method. In the proposed method, we use an exhaustive search method to find the minimum value of $f(\alpha)$. The functional f can also be considered as a form of generalized cross-validation where the solution is constrained by the autocorrelation function of the observed process [18]. The INVE is given below.

The INVE Method

1) Obtain an estimate of the biased noisy autocorrelation sequence $\{\hat{r}_{yy}(k)\}$ as

$$\hat{r}_{yy}(k) = \frac{1}{N}\sum_{n=0}^{N-1-|k|} y(n)y(n+|k|) \quad (5)$$

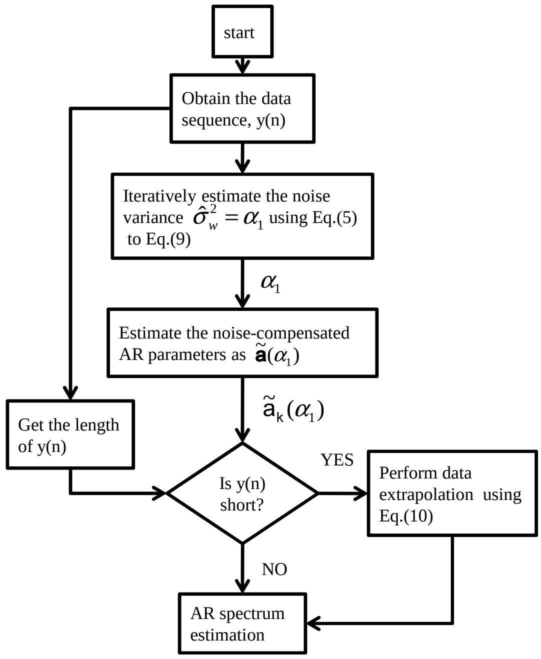


Fig. 1. A flowchart of the proposed method.

where N is the length of the observed noisy data sequence, and solve Eq.(2) by the Levinson-Durbin algorithm to obtain an $\hat{\mathbf{a}}$.

2) Set M as a large real number and initialize the noise variance parameter and the step-size parameter as

$$\alpha = \frac{1}{M} \quad (6)$$

$$s = \frac{\alpha}{M} \quad (7)$$

respectively.

3) Substitute Eq. (3) into Eq. (1) with the setting of $\sigma_w^2 = \alpha$ and solve Eq. (1) to obtain an estimate of $\mathbf{a}$ denoted by $\tilde{\mathbf{a}}(\alpha)$.
4) Calculate Eq. (4).
5) If $\alpha \geq \hat{r}_{yy}(0)$, terminate and obtain the value of α for which $f(\alpha)$ is minimum. Otherwise, go to Step 6.
6) Calculate s according to

$$s = \begin{cases} \hat{r}_{yy}(0)/M, & \text{if } \alpha \geq 1 \\ \alpha/M, & \text{otherwise} \end{cases} \quad (8)$$

and increase α by the value of s as

7) Go to Step 3.

If $f(\alpha)$ is minimized when $\alpha = \alpha_1$, then the noise variance estimate is $\hat{\sigma}_w^2 = \alpha_1$. The noise-compensated AR parameters are $\tilde{\mathbf{a}}(\alpha_1)$. Thus, the noise variance and the AR parameters are readily available.

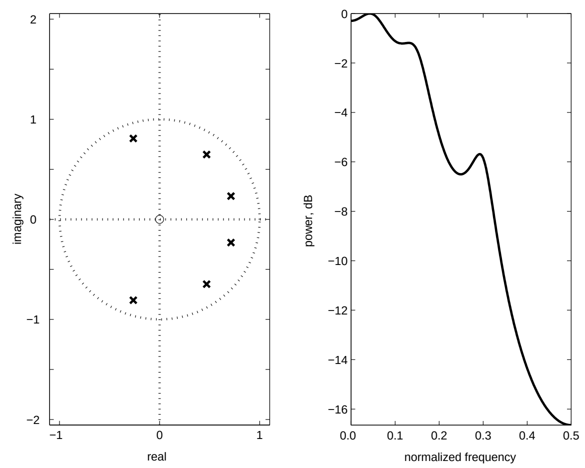
For the data extrapolation part, we can use $\tilde{\mathbf{a}}(\alpha_1)$ to extrapolate the data sequence $y(n)$ as described in [17] so that we have

$$\hat{y}(n) = -\sum_{k=1}^{p} \tilde{\alpha}_k(\alpha_1)y(n-k) \quad (9)$$

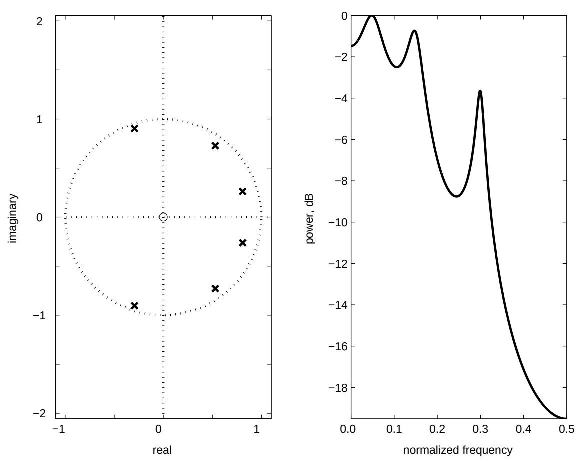
where $\hat{y}(n)$ denotes the extrapolated points, and $n \geq N$. The resulting extrapolated data sequence denoted by $y'(n)$ becomes

Table 1. Noise variance estimation RMSE performance for the SS method and the proposed method.

Variance	SS Method		Proposed Method	
α_w^2	$\hat{\alpha}_w^2$	RMSE	$\hat{\alpha}_w^2$	RMSE
0.3162	0.3559	0.0397	0.3211	0.0049
0.2512	0.2674	0.0163	0.2675	0.0163
0.1995	0.2407	0.0412	0.2163	0.0168
0.1585	0.1734	0.0149	0.1551	0.0034
0.1259	0.1453	0.0194	0.1248	0.0011
0.1000	0.1265	0.0265	0.0718	0.0282
0.0794	0.0926	0.0132	0.0688	0.0107
0.0631	0.0728	0.0097	0.0642	0.0011



(a)



(b)

Fig. 2. The pole-zero plots and the all-pole filter frequency response plots for (a) SET1 and (b) SET2.

$\{y(0), y(1), ..., y(N-1), \hat{y}(N), , ..., \hat{y}(Q-1)\}$, where Q is the total data sequence length and $Q > N$. Applying (4) to $y'(n)$ results in the desired parameter estimates.

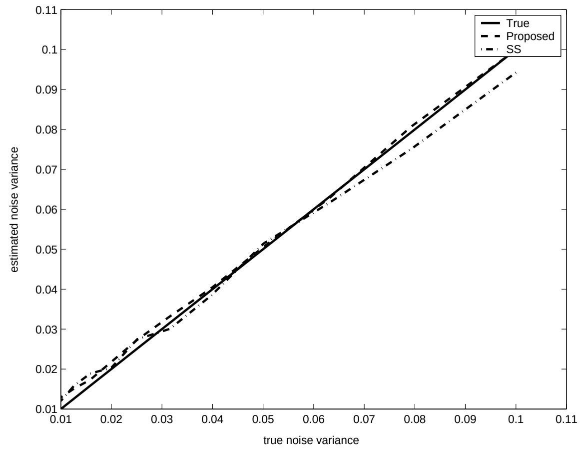


Fig. 3. Comparison of the deviation of the estimated noise variance from the true noise variance for the SS method and the proposed method using SET1. The solid line is the true noise variance, the dashed line is the proposed method and the dash-dotted line is the SS method.

Table 2. Comparison of NC only with the SS method at SNRS of 10db, 15db; averaging 100 independent runs and using set1 poles.

10 (dB)		15 (dB)	
SS method	Dev.	SS method	Dev.
$-0.3848 + 0.6449i$	0.3543	$0.1632 + 0.7578i$	0.1328
$-0.1524 - 0.5504i$	0.3940	$-0.1632 - 0.7578i$	0.1328
$0.4580 + 0.4375i$	0.2417	$0.4849 + 0.6168i$	0.0394
$0.4991 + 0.4375i$	0.1958	$0.4849 - 0.6168i$	0.0394
$0.6788 + 0.1909i$	0.1467	$0.6829 + 0.2550i$	0.0487
$0.6887 - 0.2411i$	0.0970	$0.6829 - 0.2550i$	0.0487
10 (dB)		15 (dB)	
NC Estimate	Dev.	NC	Estimate Dev.
$-0.1318 \pm 0.2668i$	0.0975	$-0.1760 \pm 0.5213i$	0.1188
$0.1914 \pm 0.2632i$	0.0289	$0.3781 \pm 0.4956i$	0.0182
$0.2879 \pm 0.1217i$	0.0443	$0.5846 \pm 0.2008i$	0.0262

Table 3. Comparison of ncde with the ss method, for a short data sequence, and settings as in table 3.

10 (dB)		15 (dB)	
SS method	Dev.	SS method	Dev.
$-0.4321 + 0.7724i$	0.2481	$-0.2666 + 0.8296i$	0.1012
$-0.3024 - 0.7560i$	0.2233	$-0.2581 - 0.8296i$	0.0982
$0.5148 + 0.6845i$	0.0554	$0.5184 + 0.7104i$	0.0222
$0.5199 - 0.6963i$	0.0440	$0.5184 - 0.7104i$	0.0222
$0.7949 + 0.2816i$	0.0410	$0.7995 + 0.2861i$	0.0262
$0.7962 - 0.2862i$	0.0366	$0.7995 - 0.2861i$	0.0262
10 (dB)		15 (dB)	
NC Estimate	Dev.	NC	Estimate Dev.
$0.1514 \pm 0.3828i$	0.1152	$-0.2127 \pm 0.6181i$	0.1008
$0.2884 \pm 0.3727i$	0.0184	$0.4169 \pm 0.5630i$	0.0185
$0.4374 \pm 0.1482i$	0.0170	$0.6374 \pm 0.2264i$	0.0238

3. Simulation Examples

In this section, simulation results are given. We show the results for noise variance estimation and then for AR parameter estimation.

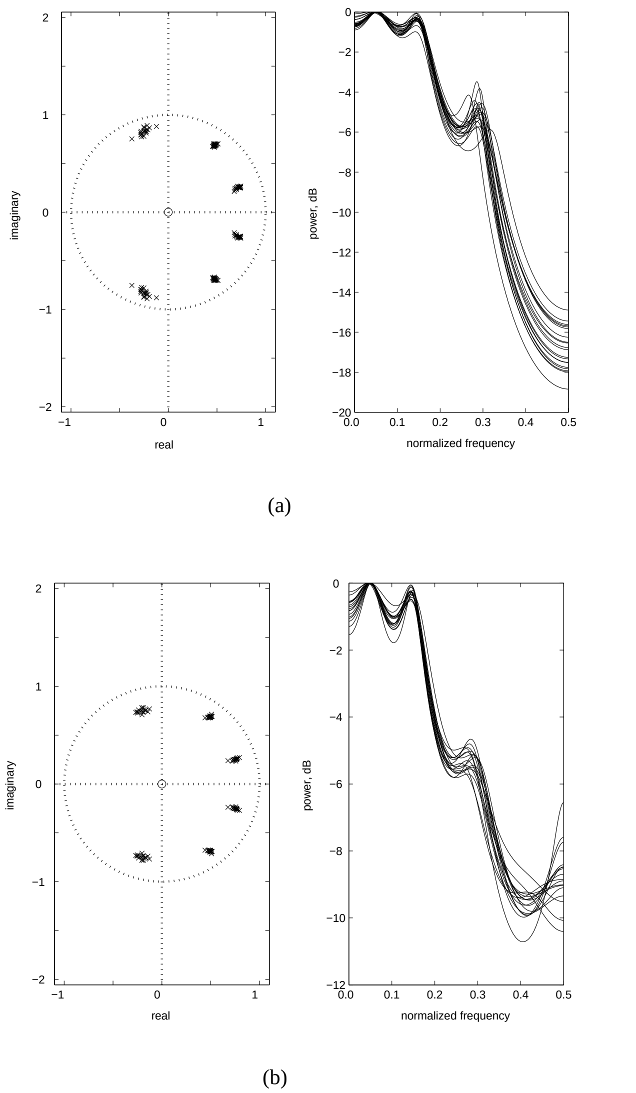


Fig. 4. The pole-zero plots and the all-pole filter frequency response plots for (a) the SS method and (b) the proposed method with noise compensation (NC) only at model order of 8 using SET1.

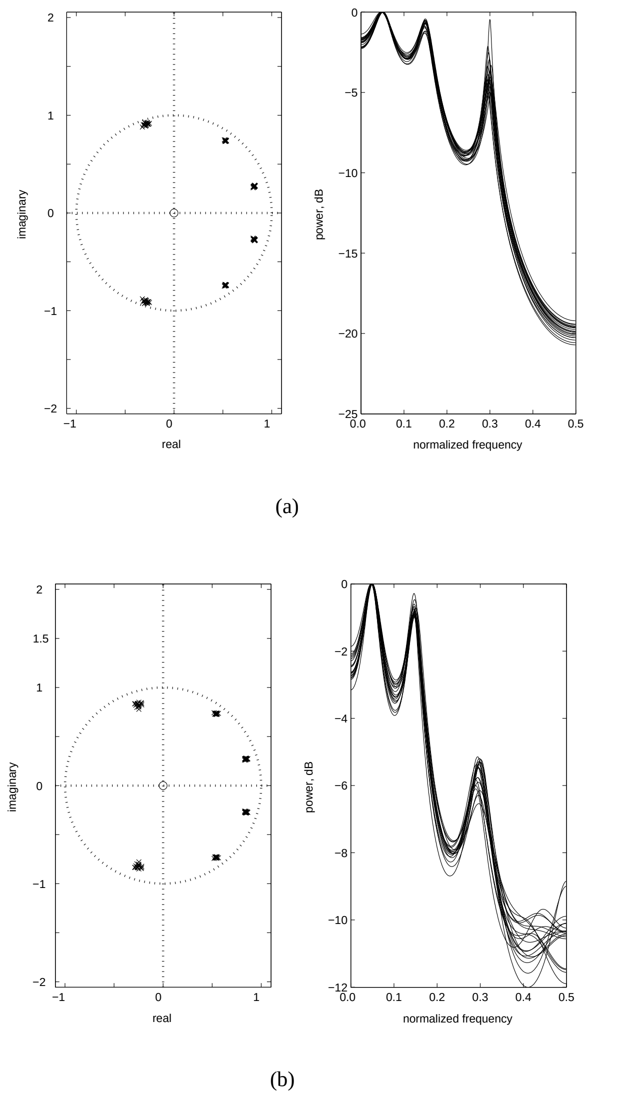


Fig. 5. The pole-zero plots and the all-pole filter frequency response plots for (a) the SS method and (b) the proposed method with NC only at model order of 8 using SET2.

3.1. Simulation environment

Two sets of poles were used as follows: SET1 = $\{-0.2627 \pm 0.8084i, 0.4702 \pm 0.6472i, 0.6657 \pm 0.2163i\}$ and SET2 = $\{-0.2936 \pm 0.9035i, 0.5290 \pm 0.7281i, 0.8084 \pm 0.2627i\}$. Fig. 2 shows the pole locations and the respective frequency response plots of the associated all-pole filters. SET2 poles are located closer the unit circle. From the all-pole filters, 5000-point white-noise excited data sequences were generated from which 500 data points were taken to represent the noise free data sequence. For the short data sequence case, the the length of the data sequence was reduced to 220. These sequences were corrupted by white Gaussian noise to give SNRs ranging from 5dB to 15dB. For NCDE an extension factor of 1.5 was used. To eliminate the effects of noise-generated poles, the threshold values of 0.70 for SET1 and 0.75 for SET2 were set. Poles with a magnitude less than the threshold values were considered to belong to the noise subspace.

3.2. Noise variance estimation

To assess the performance of the proposed method for noise variance estimation SET1 poles were used.

Table 1 gives the root mean square error (RMSE) for the noise variance estimation by both methods while Fig. 3 illustrates the deviation of the estimated noise variance from the true noise variance. The RMSE at each noise variance σ_w^2 was obtained as

$$RMSE_{\sigma_w^2} = \sqrt{\left(\frac{1}{100}\sum_{i=1}^{100}\left(\sigma_w^2 - \hat{\sigma}_{wi}^2\right)^2\right)} \qquad (10)$$

where $\hat{\sigma}_{wi}^2$ is the i-th independent estimate of σ_w^2. For the simulated SNR's, the proposed method results in lower RMSE than the subspace (SS) method.

3.3. AR parameter estimation

Fig. 4 shows the pole-zero plots and the all-pole filter frequency response plots for the subspace method and for noise compensation without data extrapolation at AR order 8 using

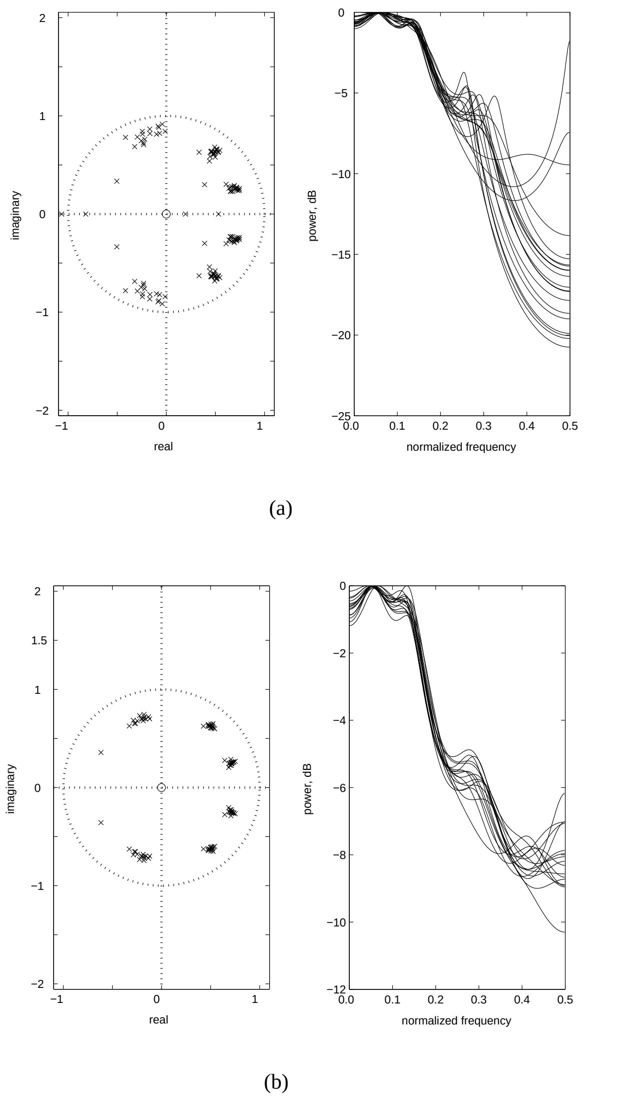


Fig. 6. The pole-zero plots and the all-pole filter frequency response plots for (a) the SS method and (b) the proposed NCDE method at model order of 8 using SET1.

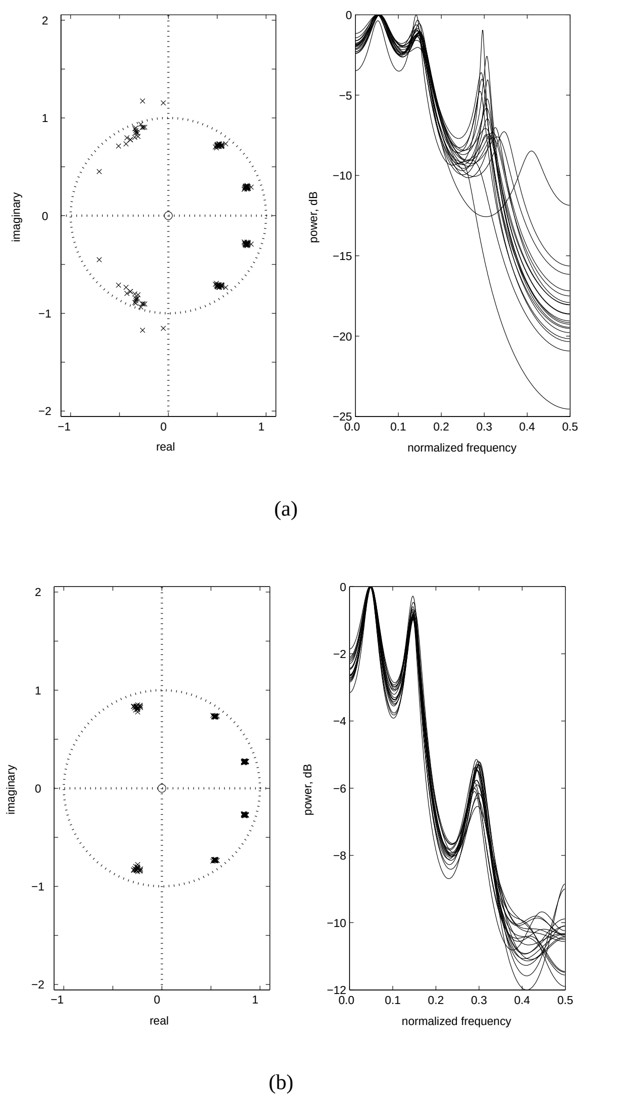


Fig. 7. The pole-zero plots and the all-pole filter frequency response plots for (a) the SS method and (b) the proposed NCDE method at model order of 8 using SET2.

SET1. The two methods both give reasonably accurate pole locations with the proposed method tending to drive the poles away from the unit circle while the subspace approach drives the poles toward the unit circle. Fig. 5 shows the effect of pushing the poles close to the unit circle (SET2) for the same settings as in Fig. 4. The subspace approach tends to produce a peaky frequency response characteristic and some of the poles even fall out of the unit circle which could raise stability problems. Fig. 6 shows the results when a short data sequence is used for both approaches with SET1. In this case NCDE performs better by more accurately locating the pole locations. Fig. 7 illustrates SET2 is used with the same settings as in Fig. 5. Again NCDE performs better. In Tables 3 and 4 the Dev. refers to the standard deviation of the estimated poles. The results show that the proposed method gives smaller deviation when compared to the subspace approach at the same SNR and data length. Therefore it is possible to obtain a better estimate of the final AR parameters using the proposed method. As previously stated, the deviations are however in different directions.

4. Concluding Remarks

In this paper we have shown that noise variance and AR parameters can be accurately extracted from observations corrupted by additive white Gaussian noise using LOYWEs. For short observations, NCDE has been shown to perform better than the SS approach. The noise variance estimates resulted in lower RMSE while the estimated poles showed smaller standard deviation. Thus, the superiority of the proposed method was confirmed.

Acknowledgment

A part of this work was supported by MEXT Grant-in-Aid for Building Strategic Research Infrastructures.

References

[1] S. M. Kay, "Noise compensation for autoregressive spectral estimates," *IEEE Trans. Acoust. Speech Signal Process*, vol. ASSP-28, no. 3, pp. 292–303, 1980.

[2] M. Pagano, "Estimation of models of autoregressive signal plus white noise," *Ann. Statist.*, vol. 2, pp. 99–108, 1974.

[3] H.-Y. Chang, S. Rahardja, and K. S. Ngee, "Autoregressive parameter estimation for kalman filtering speech enhancement," *ICASSP*, vol. 4, pp. IV–913–IV–916, 2007.

[4] S. Tsai, T. F. Wong, and J. S. Lehnert, "DS-CDMA system with joint channel estimation and map detection in time-selective fading channels," *IEEE Journal on Selected Areas in Communications*, vol. 19, no. 1, pp. 121–131, 2001.

[5] R. L. Kirlin and W. J. Done, "Covariance analysis for seismic signal processing," *SEG Digital Library*, 1999.

[6] D. Ramakrishnan and J. Krolik, "Adaptive radar detection in doubly nonstationary autoregressive doppler spread clutter," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 45, no. 2, pp. 484–501, 2009.

[7] M. Niedzwiecki, *Identification of Time-Varying Processes*. Wiley, 2000.

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

[9] D. P. McGinn and D. H. Johnson, "Estimation of all-pole model parameters from noise-corrupted sequences," *IEEE Trans. Acoust. Speech Signal Process*, vol. 37, no. 3, pp. 433–436, 1989.

[10] W. X. Zheng, "Estimation of autoregressive signals from noisy measurements," *IEE Proceedings - Vision, Image, and Signal Processing*, vol. 144, no. 1, pp. 39–45, 1997.

[11] H.-T. Hu, "Linear prediction analysis of speech signals in the presence of white gaussian noise with unknown variance," *IEE Proceedings - Vision, Image, and Signal Processing*, vol. 145, no. 4, pp. 303–308, 1998.

[12] Y. Chan and R. Langford, "Spectral estimation via the high-order yule-walker equations," *IEEE Trans. Acoust. Speech Signal Process*, vol. 30, no. 5, pp. 689–698, 1982.

[13] J. Cadzow, "Spectral estimation: An overdetermined rational model equation approach," *Proc. IEEE*, vol. 70, pp. 907–939, 1982.

[14] C. E. Davila, "A subspace approach to estimation of autoregressive parameters from noisy measurements," *IEEE Trans. Signal Process*, vol. 46, no. 2, pp. 531–534, 1998.

[15] J. Gamba and T. Shimamura, "A new method of noise variance estimation from low-order yule-walker equations," *IEICE Trans. Fundamentals*, vol. 1, pp. 270–274, 2004.

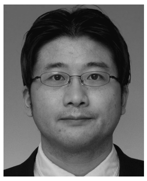
[16] T. Shimamura, T. Nagamine, and J. Suzuki, "High-resolution two-dimensional spectral estimation utilizing data extrapolation," *Trans. ISCIE*, vol. 7, no. 6, pp. 213–220, 1994.

[17] T. S. J. Gamba, "Two-dimensional spectral estimation with noise-compensated data extrapolation," *Proc. Internat. Conf. Fundamentals of Electronics, Communications and Computer Science*, pp. 17–22, March 2002.

[18] J. Chung, J. G. Nagy, and D. P. O'Leary, "A weighted-gcv method for lanczos-hybrid regularization," *Electronic Trans. Numerical Analysis*, vol. 28, pp. 149–167, 2008.

J. Gamba received the BSc.(Hons) degree in electrical and electronics engineering from the University of Zimbabwe, Harare, Zimbabwe, in 1994, the MSc. degree in computer science and engineering from Zhejiang University, Hangzhou, China, in 2000 and the PhD in Mathematical Information Science from Saitama University, Japan, in 2005. He was with Tsukuba Universitys TARA Center as a postdoctoral researcher of multimedia information systems applications from 2006 to 2008. He is presently a visiting researcher of Seikei University, Tokyo, Japan. His current research interests include time delay estimation, adaptive channel equalization, ubiquitous networks and image restoration. He is a member of the IEEE.

T. Shimamura received the B.E., M.E., and Ph.D. degrees in electrical engineering from Keio University, Yokohama, Japan, in 1986, 1988, and 1991, respectively. In 1991, he joined Saitama University, Saitama City, Japan, where he is currently a Professor. His interests are in digital signal processing and applications to speech and communication systems. He is a member of IEEE, and EURASIP.



S. Kawasaki received B.Sc. from University of Electro-Communications, Tokyo in 1991. He then received M. Sc. in electrical engineering in 1993 and Ph. D in Mathematics in 1996, respectively, of Graduate School of Science and Technology, Keio University, Japan. He joined as a research associate at Graduate School of Information Systems, University of Electro-Communications in 1996, as an assistant professor at Graduate School of Economics, Hitotsubashi University in 2003 and as a CREST researcher at TARA Center, University of Tsukuba in 2005, respectively. Presently he is a senior researcher of CATS Co. Ltd, Japan and researcher at Graduate School of Engineering, Seikei University. His research interest is in stochastic process theory.

M. Higuchi received the PhD degree in Science from the Niigata university, Japan, in 2005. He was a researcher at the university of Tsukuba until 2009. From 2005 until 2008, he was a member of the JST-CREST project. From 2008 until 2009, he was a member of the new generation network project supported by NICT. He is currently a researcher at Department of Information and Computer Science of Seikei University in Japan, and a member of a project supported by MEXT Grant-in-Aid for Building Strategic Research Infrastructures. His research interests include Acoustics related sonics, psychoacoustics, physiological acoustics, Image processing, Network traffic control, and Operations research related game theory. He has published extensively and received the best paper award in WSEAS International Conference on TELECOMMUNICATIONS AND INFORMATICS (TELE-INFO 2009).

H. Murakami received Ph. D degree from Graduate School of Engineering, Hokkaido University, Japan in 1974. Then he joined KDD corporation and since then, he engaged himself in the research of digital satelite communication, digital TV transmission, TV signal transform technology. He became an executive CTO of KDDI corporation in 1997, and since then, president of multimedia technology businesses, president of R&D and president of network technologies. He was a chairman of TTC standardization and ITUSG9 in Japan. He was also the vice president of Society of Video Media in Japan. Presently, after his retirement of KDDI corporation, he is a professor of Department of Information and Computer Science, Seikei University, Tokyo.