# LINEAR PREDICTION BASED ON HIGHER ORDER STATISTICS BY A NEW CRITERION

Chong-Yung Chi and Wu-Ton Chen

Department of Electrical Engineering
National Tsing Hua University
Hsinchu, Taiwan, Republic of China

## ABSTRACT

Chi [1-3] proposed a criterion, for obtaining a higher-order statistics (HOS) based linear prediction error (LPE) filter associated with a non-Gaussian stationary process $x(k) = y(k)+n(k)$ where $y(k)$ is a non-Gaussian signal and $n(k)$ is additive Gaussian noise, which requires all $M$th-order cumulants $C_{M,e}(k_1, k_2, ..., k_{M-1})$ of the prediction error $e(k)$. In this paper, we propose a new criterion, which requires only partial $M$th-order cumulants $C_{M,e}(0, k_1, k_1, ..., k_{M/2-1}, k_{M/2-1})$ of $e(k)$ where $M$ is even, for obtaining a new HOS based LPE filter. Theoretically, we show that the proposed HOS based LPE filter associated with $x(k)$ is the same as the conventional correlation based (minimum-phase) LPE filter associated with $y(k)$ (noise-free), and that they are equivalent to Chi's HOS based LPE filter associated with $x(k)$ for the case that $y(k)$ is an AR process as well as the case that their order is infinite. Simulation results are provided to show that when $y(k)$ is an AR process of known order, the proposed HOS based LPE filter works well.

## 1. INTRODUCTION

The linear prediction error (LPE) filter can be found in various science and engineering areas such as speech processing, seismic deconvolution and spectral estimation. The conventional LPE filter [4] is based on second-order statistics (power spectra or correlation functions). Therefore, it is phase blind no matter whether the signal of interest is Gaussian or not. Recently, a lot of higher order statistics (HOS) based methods such as [5-8] for estimating unknown parameters of non-Gaussian linear processes have been reported in the open literature. Two common characteristics of HOS based estimators are: They are insensitive to Gaussian noise since HOS of Gaussian processes are totally zero; they can recover the phase of linear processes. Chi [1-3] proposed a HOS based LPE filter which was shown to be minimum-phase, and the associated linear prediction polyspectral estimator possesses the property of maximum polyspectral flatness measure as well as the property of maximum higher-order entropy. However, the coefficients of Chi's HOS based LPE filter must be solved from all $M$th-order cumulants $C_{M,e}(k_1, k_2, ..., k_{M-1})$ of the prediction error $e(k)$. In this paper, we propose a new criterion which requires only partial $M$th-order cumulants $C_{M,e}(0, k_1, k_1, ..., k_{M/2-1})$ for obtaining a new HOS based LPE filter. Then we present the equivalencies among the proposed HOS based LPE filter, Chi's HOS based LPE filter and the conventional correlation based LPE filter. Some simulation results for the case of AR process are provided to support the proposed HOS based LPE filter. Finally, we draw some conclusions.

## 2. PREDICTION ERROR FILTER BASED ON HOS BY A NEW CRITERION

Assume that $x(k)$, $k = 0, 1, ..., N-1$ are the given real stationary non-Gaussian noisy measurements based on the following convolutional model

$$x(k) = y(k) + n(k) = u(k) * h(k) + n(k) \qquad (1)$$

where $u(k)$ is a real, zero-mean, independent identically distributed (i.i.d.) non-Gaussian process with variance $\sigma_u^2$ and $M$th-order cumulant $\gamma_M$, $n(k)$ is zero-mean Gaussian noise and $h(k)$ is the impulse response of a linear time-invariant (LTI) causal stable system. Assume that $v(k)$ is a FIR filter of order $p$ and its output $e(k)$ is given by

$$e(k) = x(k) * v(k) = x(k) + \sum_{i=1}^{p} v(i)\, x(k-i) \qquad (2)$$

Note that $v(0) = 1$. Let $\underline{\theta} = (v(1), v(2), ..., v(p))'$ and

$$w(k) = h(k) * v(k). \qquad (3)$$

Then we have, from (1) to (3), that

$$e(k) = \xi(k) + n(k) * v(k) \qquad (4)$$

where

$$\xi(k) = y(k) * v(k) = u(k) * w(k). \quad (5)$$

It is well known that the autocorrelation function of $\xi(k)$ is given by

$$\phi_{\xi\xi}(l) = E[\xi(k)\xi(k+l)] = \sigma_u^2 \phi_{ww}(l) \quad (6)$$

where

$$\phi_{ww}(l) = \sum_{n=-\infty}^{\infty} w(n)w(n+l). \quad (7)$$

Some well-known facts regarding the conventional LPE filter are summarized in (F1) as follows:

(F1) The conventional (minimum-phase) LPE filter associated with $y(k)$ is the $v(k)$ which minimizes the mean square error $E[\xi^2(k)] = \phi_{\xi\xi}(0)$. When $y(k)$ is the output of a $p$-th order AR system $1/A(z)$, the conventional LPE filter of order $p$ associated with $y(k)$ is equal to $A(z)$ and $\phi_{\xi\xi}(l) = 0, \forall l \neq 0$; when $y(k)$ is a wide-sense stationary process, it is a whitening filter ($\phi_{\xi\xi}(l) = 0,\ \forall l \neq 0$) for $p \rightarrow \infty$.

Chi [1-3] proposed a HOS based LPE filter by minimizing

$$J_M(\underline{\theta}) = \sum_{k_1=-\infty}^{\infty} \cdots \sum_{k_{M-1}=-\infty}^{\infty} C_{M,e}^2(k_1, k_2, ..., k_{M-1}) \quad (8)$$

and he showed that the optimum LPE filter is minimum-phase. Next, we show the following fact.

(F2) The optimum $p$th-order LPE filter associated with $J_M(\underline{\theta})$ is the same as the conventional $p$th-order LPE filter associated with $y(k)$ (noise-free, i.e., $e(k) = \xi(k)$) when $M$ is even and $y(k)$ is an AR process of order $p$.

Proof: Since

$$J_M(\underline{\theta}) = \sum_{k_1=-\infty}^{\infty} \cdots \sum_{k_{M-1}=-\infty}^{\infty} C_{M,e}^2(k_1, k_2, ..., k_{M-1}) = \gamma_M^2$$

$$\cdot \sum_{k_1=-\infty}^{\infty} \cdots \sum_{k_{M-1}=-\infty}^{\infty} \left\{ \sum_{n=-\infty}^{\infty} w(n)w(n+k_1)...w(n+k_{M-1}) \right\}^2$$

$$= \gamma_M^2 \sum_{n=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} w(n)w(m) \sum_{k_1=-\infty}^{\infty} w(n+k_1)w(m+k_1)$$

$$\cdots \sum_{k_{M-1}=-\infty}^{\infty} w(n+k_{M-1})w(m+k_{M-1})$$

$$= \gamma_M^2 \sum_{n=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} w(n)w(m)\phi_{ww}^{M-1}(n-m)$$

$$= \gamma_M^2 \sum_{l=-\infty}^{\infty} \left\{ \sum_{m=-\infty}^{\infty} w(m+l)w(m) \right\} \phi_{ww}^{M-1}(l)$$

$$= \gamma_M^2 \sum_{l=-\infty}^{\infty} \phi_{ww}^{M}(l) = \left( \frac{\gamma_M^2}{\sigma_u^{2M}} \right) \sum_{l=-\infty}^{\infty} \phi_{\xi\xi}^{M}(l). \quad (9)$$

When $M$ is even, it is sufficient that $J_M(\underline{\theta})$ is minimum if $|\phi_{\xi\xi}(l)|$ is minimum for all $l$. A set of sufficient conditions for $|\phi_{\xi\xi}(l)|$ to be minimum for all $l$ is that $\phi_{\xi\xi}(0)$ is minimum and $\phi_{\xi\xi}(l) = 0$ for all $l \neq 0$ in the meantime. By (F1), we conclude that the optimum $\hat{v}(k)$ is the same as the conventional LPE filter associated with $y(k)$ when $y(k)$ is an AR process of order $p$. Thus we have completed the proof.

On the other hand, we obtain the optimum LPE filter $\hat{v}(k)$ by minimizing a new criterion as follows:

$$\tilde{J}_M(\underline{\theta}) = \left( \sum_{k_1=-\infty}^{\infty} \cdots \sum_{k_L=-\infty}^{\infty} C_{M,e}(0, k_1, k_1, ..., k_L, k_L) \right)^2 \quad (10)$$

where $M(\geq 4)$ is even and $L = M/2-1$. Next, we show the following fact.

(F3) The optimum $\hat{v}(k)$ associated with $\tilde{J}_M(\underline{\theta})$ where $M$ is even is the same as the conventional $p$th-order LPE filter associated with $y(k)$ (noise-free, i.e., $e(k) = \xi(k)$).

Proof: Since

$$\sum_{k_1=-\infty}^{\infty} \cdots \sum_{k_L=-\infty}^{\infty} C_{M,e}(0, k_1, k_1, ..., k_L, k_L)$$

$$= \gamma_M \sum_{k_1=-\infty}^{\infty} \cdots \sum_{k_L=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} w^2(k)w^2(k+k_1)...w^2(k+k_L)$$

$$= \gamma_M \left( \sum_{k=-\infty}^{\infty} w^2(k) \right)^{L+1} = \frac{\gamma_M}{\sigma_u^{2(L+1)}} \phi_{\xi\xi}^{L+1}(0), \quad (11)$$

we obtain, by substituting (11) into (10), that

$$\tilde{J}_M(\underline{\theta}) = \{\gamma_M^2/\sigma_u^{2M}\}\, \phi_{\xi\xi}^{M}(0). \quad (12)$$

Hence, minimizing $\tilde{J}_M(\underline{\theta})$ is equivalent to minimizing the mean-square prediction error $E[\xi^2(k)]$. In other words, the optimum $\hat{v}(k)$ is the same as the conventional LPE filter associated with $y(k)$ (noise-free, i.e., $e(k) = \xi(k)$). Thus, we have completed the proof.

One can easily infer, from (9) and (12), that the proposed HOS based LPE filter is different from Chi's HOS based LPE filter. However, for the case that $y(k)$ is an AR process of order $p$, they are equivalent to the $p$th-order conventional LPE filter by (F2) and (F3) when M is even. For the case that $y(k)$ is not an AR process, we have the following fact.

(F4) The optimum $\hat{v}(k)$ associated with $\tilde{J}_M(\underline{\theta})$, the one associated with $J_M(\underline{\theta})$ and the conventional (minimum-phase) LPE filter associated with $y(k)$ (noise-free i.e., $e(k) = \xi(k)$) are equivalent when $M$ is even and their order $p$ is infinite.

Proof: That the optimum $\hat{v}(k)$ associated with $\tilde{J}_M(\underline{\theta})$ is equivalent to the conventional LPE filter associated with $y(k)$ for any $p$ has been shown in (F3) as long as $M$ is even. Hence they are also equivalent for $p \to \infty$. Next, we show that the latter is equivalent to Chi's HOS based LPE filter for $p \to \infty$. By (F1), $\phi_{\xi\xi}(0)$ is minimum for $p \to \infty$ when $\phi_{\xi\xi}(l) = 0$, $\forall\, l \neq 0$. It is sufficient that $J_M(\underline{\theta})$ given by (9) with an even M is minimum when $\phi_{\xi\xi}(0)$ is minimum and $\phi_{\xi\xi}(l) = 0$, $\forall\, l \neq 0$. Thus Chi's HOS based LPE filter is also equivalent to the conventional LPE filter for $p \to \infty$.

Now, we present how to solve for the optimum filter coefficients $\hat{\underline{\theta}}$ based on (10). For simplicity, assume that $M = 4$. We use a steepest-descent type iterative algorithm to search for a local minimum of $\tilde{J}_4$ where

$$\tilde{J}_4 = \left( \sum_{k=-K}^{K} \hat{C}_{4,e}(0,k,k) \right)^2 \tag{13}$$

in which $\hat{C}_{4,e}(0,k,k)$ is the biased sample cumulant as follows:

$$\hat{C}_{4,e}(0,k,k) = \frac{1}{N} \sum_{i=0}^{N-|k|} e^2(i)e^2(i+|k|) - \left( \frac{1}{N} \sum_{i=0}^{N} e^2(i) \right)^2 - 2 \left( \frac{1}{N} \sum_{i=0}^{N-|k|} e(i)e(i+|k|) \right)^2 \tag{14}$$

For the $n$th iteration, $\hat{\underline{\theta}}$ is updated by

$$\hat{\underline{\theta}}_n = \hat{\underline{\theta}}_{n-1} - \rho[Tr(H_{n-1})]^{-1} \underline{g}_{n-1} \tag{15}$$

where $0 < \rho \leq p$ is a real number, $Tr(H_{n-1})$ denotes the trace of $H_{n-1}$, and $\underline{g}_{n-1}$ and $H_{n-1}$ denote the gradient and the approximate Hessian matrix for $\underline{\theta} = \hat{\underline{\theta}}_{n-1}$, respectively, as follows:

$$\underline{g}_{n-1} = \left. \frac{\partial \tilde{J}_4}{\partial \underline{\theta}} \right|_{\underline{\theta}=\hat{\underline{\theta}}_{n-1}} = 2 \left( \sum_{k=-K}^{K} \hat{C}_{4,e}(0,k,k) \right) \cdot \left. \left( \sum_{k=-K}^{K} \frac{\partial \hat{C}_{4,e}(0,k,k)}{\partial \underline{\theta}} \right) \right|_{\underline{\theta}=\hat{\underline{\theta}}_{n-1}} \tag{16}$$

$$H_{n-1} = \left. \frac{\partial^2 \tilde{J}_4}{\partial \underline{\theta}^2} \right|_{\underline{\theta}=\hat{\underline{\theta}}_{n-1}} \approx 2 \left( \sum_{k=-K}^{K} \frac{\partial \hat{C}_{4,e}(0,k,k)}{\partial \underline{\theta}} \right) \cdot \left. \left( \sum_{k=-K}^{K} \frac{\partial \hat{C}_{4,e}(0,k,k)}{\partial \underline{\theta}} \right)' \right|_{\underline{\theta}=\hat{\underline{\theta}}_{n-1}} \tag{17}$$

Computing $\underline{g}_{n-1}$ and $H_{n-1}$ requires $\partial e(k)/\partial v(i)$ which can be easily obtained by taking partial derivative of (2) with respect to $v(i)$ as follows:

$$\frac{\partial e(k)}{\partial v(i)} = x(k-i) \tag{18}$$

Remark that updating $\underline{\theta}$ by (15) with $\rho = p$ normally leads to the decrease of $\tilde{J}_4$, otherwise, a smaller $\rho$ must be considered.

## 3. SIMULATION RESULTS

A zero-mean, Exponentially i.i.d. random sequence $u(k)$ with variance $\sigma_u^2 = 1$ and 4th-order cumulant $\gamma_4 = 6$ was generated as the input to a second-order AR system with transfer function $H(z) = 1/A(z)$ where

$$A(z) = 1 + a(1)z^{-1} + a(2)z^{-2} = 1 + 0.7z^{-1} + 0.2z^{-2} \tag{19}$$

and then a white Gaussian noise sequence was added to the output of $H(z)$ to obtain the synthetic noisy data. The length of data was $N = 4096$ and the parameter $K$ (see (13)) was set to 6. Mean and standard deviation of estimated parameters were calculated from 60 independent realizations. The initial value $\hat{\underline{\theta}}_0 = (0,0)'$ was used for each realization. The simulation results associated with the conventional correlation based LPE filter obtained by the well-known Burg algorithm [4] and those associated with the proposed HOS based LPE filter obtained by the previous steepest-descent type iterative algorithm are shown in Table I. Observe, from this table, that when SNR is large (SNR = 400), mean values of estimated parameters associated with the conventional LPE filter and those associated with the proposed HOS based LPE filter are very close to the true AR parameters. When SNR is small (SNR = 10), biases of estimated parameters associated with the latter are much smaller than those associated with the former. Moreover, mean square errors (equal to the sum of variance and square of bias) of estimated parameters associated with the proposed HOS based LPE filter are also smaller than those associated with the conventional

LPE filter for SNR = 10, although standard deviations of estimated parameters associated with the latter are smaller than those associated with the former. Therefore, the simulation results support that the proposed HOS based LPE filter approximates the true AR parameters well.

Table I. Simulation results.

| a(1) = 0.7, a(2) = 0.2, N = 4096, 60 independent runs. | | | | |
|---|---|---|---|---|
| | Estimated values (mean ± one standard deviation) | | | |
| | Conventional LPE filter obtained by Burg alg. | | HOS based LPE filter obtained by the proposed criterion | |
| SNR | $\hat{v}(1)$ | $\hat{v}(2)$ | $\hat{v}(1)$ | $\hat{v}(2)$ |
| 400 | 0.6978 ± 0.0129 | 0.1988 ± 0.0138 | 0.7039 ± 0.0533 | 0.1971 ± 0.0593 |
| 40 | 0.6717 ± 0.0123 | 0.1797 ± 0.0135 | 0.7024 ± 0.0571 | 0.1961 ± 0.0629 |
| 10 | 0.5993 ± 0.0130 | 0.1291 ± 0.0138 | 0.7011 ± 0.0706 | 0.1955 ± 0.0766 |

## 4. CONCLUSIONS

We have presented a new criterion given by (10) where M is even for obtaining a new HOS based LPE filter associated with a non-Gaussian linear process $x(k) = y(k) + n(k)$ (see (1)) where $y(k)$ is a non-Gaussian signal and $n(k)$ is additive Gaussian noise. Theoretically, we have shown (see (F2) through (F4)) that the proposed HOS based LPE filter associated with $x(k)$ is the same as the conventional (minimum-phase) LPE filter associated with $y(k)$ (noise-free), and that they are equivalent to Chi's HOS based LPE filter based on (8) associated with $x(k)$ for the case that $y(k)$ is an AR process of known order as well as the case that their order is infinite. Moreover, the computational load for obtaining the proposed HOS based LPE filter is much less than that for obtaining Chi's HOS based LPE filter because the latter requires all $M$th-order cumulants of $e(k)$ and the former requires only partial $M$th-order cumulants of $e(k)$ (see (8) and (10)). Some simulation results were provided to show that when $y(k)$ is an AR process of known order and the number of data used is large enough, the smaller SNR, the more the proposed HOS based LPE filter outperforms the conventional LPE filter.

## ACKNOWLEDGEMENTS

The research described in this paper was supported by the National Science Council under Grant NSC81-0404-E-007-001 and the Telecommunication Laboratories under Grant TL-NSC-81-5201.

## REFERENCES

[1] C.-Y. Chi, "Linear prediction, maximum flatness, maximum entropy and AR polyspectral estimation," accepted for publication in *IEEE Trans. Signal Processing*.

[2] C.-Y. Chi, "Linear prediction, maximum flatness, maximum higher-order entropy on AR polyspectral estimation," *Proc. Sixth European Signal Processing Conference (EUSIPCO-92)*, Brussels, Belgium, Aug. 24-27, 1992.

[3] C.-Y. Chi and W.-T. Chen, "Linear Prediction Based on Higher Order Statistics," *Proc. 3rd International Symposium on Signal Processing and Its Applications (ISSPA-92)*, Gold Coast, Australia, Aug. 16-21, 1992.

[4] S. M. Kay, *Modern spectral estimation*, Prentice-Hall, Englewood Cliffs, New Jersey, 1988.

[5] J. M. Mendel, "Tutorial on higher-order statistics (spectra) in signal processing and system theory: theoretical results and some applications," *Proc. IEEE*, vol. 79, no. 3, pp. 278-305, March 1991.

[6] G. B. Giannakis and J. M. Mendel, "Identification of nonminimum phase systems using higher order statistics," *IEEE Trans. Acoustics, Speech, and Signal Processing*, vol. 37, pp. 360-377, March 1989.

[7] M. R. Raghuveer and C. L. Nikias, "Bispectrum estimation: A parametric approach," *IEEE Trans. Acoustics, Speech, and Signal Processing*, vol. 33, no. 5, pp. 1213-1230, Oct. 1985.

[8] M. R. Raghuveer and C. L. Nikias, "Bispectrum estimation via AR modeling," *Signal Processing (Special issue on Modern Trends in Spectral Analysis)*, vol. 9, no. 1, pp. 35-48, Jan. 1986.