

Parametric Cumulant Based Phase Estimation of 1-D and 2-D Nonminimum Phase Systems by Allpass Filtering

Horng-Ming Chien, Huang-Lin Yang, and Chong-Yung Chi, *Senior Member, IEEE*

Abstract—This paper proposes a parametric cumulant-based phase-estimation method for one-dimensional (1-D) and two-dimensional (2-D) linear time-invariant (LTI) systems with only non-Gaussian measurements corrupted by additive Gaussian noise. The given measurements are processed by an optimum allpass filter such that a single M th-order ($M \geq 3$) cumulant of the allpass filter output is maximum in absolute value. It can be shown that the phase of the unknown system of interest is equal to the negative of the phase of the optimum allpass filter except for a linear phase term (a time delay).

For the phase estimation of 1-D LTI systems, an iterative 1-D algorithm is proposed to find the optimum allpass filter modeled either by an autoregressive moving average (ARMA) model or by a Fourier series-based model. For the phase estimation of 2-D LTI systems, an iterative 2-D algorithm is proposed that only uses the Fourier series-based allpass model. A performance analysis is then presented for the proposed cumulant-based 1-D and 2-D phase estimation algorithms followed by some simulation results and experimental results with real speech data to justify their efficacy and the analytic results on their performance. Finally, the paper concludes with a discussion and some conclusions.

I. INTRODUCTION

IDENTIFICATION of an unknown linear time-invariant (LTI) system $h(n)$ with Gaussian noise-corrupted measurements $x(n)$ of the system plays an important role in various engineering applications such as seismic deconvolution, channel equalization, speech processing, and image processing. It is known that system identification methods only using correlations of $x(n)$ are not able to recover the phase of $h(n)$ when it is nonminimum phase. In the past decade, many cumulant-based methods [1]–[5] have been reported due to two common properties of cumulants. One is that not only the amplitude but also phase of $h(n)$ can be recovered from higher order (≥ 3) cumulants of $x(n)$, and the other is that cumulant-based methods are insensitive to additive Gaussian noise because all higher order cumulants of Gaussian processes are equal to zero.

Phase estimation can be performed basically with three categories of cumulant-based methods. The first category includes cumulant-based estimation methods [1]–[5] that simultane-

ously estimate the amplitude and phase of the unknown system $h(n)$ by estimating the parameters of an assumed model for $h(n)$ such as autoregressive (AR), moving average (MA), or autoregressive moving average (ARMA) models. The second category consists of minimum-phase (MP)-allpass (AP) decomposition-based methods [6]–[12] that estimate the amplitude of $h(n)$ using a correlation-based method and then estimate the phase of $h(n)$ using a cumulant-based method. The third category includes polyspectrum phase-based methods [13]–[24] that estimate the phase of $h(n)$ from the phase of polyspectra of $x(n)$ without involving amplitude estimation of $h(n)$. Most methods of the first and second categories are parametric methods, but those of the third category are nonparametric methods. In this paper, a new cumulant-based phase-estimation method is proposed that is a parametric method, but it neither involves amplitude estimation of $h(n)$ nor involves polyspectrum phase of $x(n)$. Next, let us briefly review the third category, followed by the second category, in order to illuminate the distinctions of the new method and the methods of these two categories.

Polyspectrum phase-based methods estimate the system phase $\theta(\omega) = \arg \{H(z = e^{j\omega})\}$ from the phase of an M th-order (≥ 3) polyspectrum of $x(n)$ based on a linear relationship between the system phase $\theta(\omega)$ and the polyspectrum phase of $x(n)$. Brillinger [13] and Lii and Rosenblatt [14] estimate $\theta(\omega)$ by a recursive formula with partial phase information of bispectrum of $x(n)$. These methods are sensitive to estimation error of bispectrum phase due to recursive error propagation. Matsuoka and Ulrych [15] proposed a least squares algorithm that utilizes all bispectrum phase information to estimate $\theta(\omega)$, but it must compute a pseudo-inverse of a huge matrix. Numerous least squares methods that make use of all phase information of bispectrum or trispectrum of $x(n)$ have been reported in the open literature in the past five years such as those reported in [16]–[20]. These least squares methods are more robust to both additive noise and phase estimation error of polyspectra of $x(n)$ than those methods reported in [13] and [14] using partial phase information of polyspectra of $x(n)$, but the former must compute the pseudo-inverse of a huge matrix to solve for $\theta(\omega)$. Recently, Li and Ding [21] proposed a least squares method to estimate $\theta(\omega)$ without needing to compute the pseudo-inverse of matrices. Moreover, this method is applicable for all M th-order ($M \geq 3$) polyspectra. As mentioned above, polyspectrum phase-based methods are based on a simple linear relationship between the system phase $\theta(\omega)$ and the phase of polyspectra of $x(n)$, which is quite suitable for nonparametric estimation of $\theta(\omega)$. However, two

Manuscript received December 9, 1995; revised February 11, 1997. This work was supported by the National Science Council under Grants NSC 84-2213-E-007-074 and NSC 85-2213-E-007-012. The associate editor coordinating the review of this paper and approving it for publication was Dr. Ananthram Swami.

The authors are with the Department of Electrical Engineering, National Tsing Hua University, Hsinchu, Taiwan, R.O.C.

Publisher Item Identifier S 1053-587X(97)04953-2.

common issues are faced with these methods. One is that to obtain an accurate estimate for the phase of polyspectra of $x(n)$ requires a quite large number of data because of large variance and low resolution of nonparametric polyspectrum estimation methods, and the other is the phase unwrapping problem. On the other hand, not many 2-D polyspectrum phase-based methods [22]–[24] were reported because the two common issues always lead to extraordinary complexity in the design of phase-estimation methods. Dianat and Raghuvver [22] use a Fourier series-based parametric model for both phase and magnitude of 1-D and 2-D non-Gaussian signals with the model parameters estimated from bispectra of data. Kang *et al.* [23] proposed some recursive phase-estimation algorithms based on the recursive formulas reported in [13] and [15] for both 1-D and 2-D cases. Their algorithms estimate principal values of $\theta(\omega)$ from those of the bispectrum phase of $x(n)$. Takajo and Takahashi [24] also proposed a 2-D phase estimation algorithm that is an extension of the 1-D phase estimation algorithm reported in [19], whereas their algorithm is quite complicated with its complexity dominated by the 2-D phase unwrapping part. The crucial phase unwrapping part of the phase-estimation algorithms reported in [22]–[24] may not work well, especially when signal-to-noise ratio (SNR) is not sufficiently high or when magnitude of polyspectra of data $x(n)$ has nulls (due to zeros of $H(z)$ on the unit bicircle), in addition to the other common issue (large variance and low resolution) mentioned above. Moreover, to our knowledge, the performance analysis of both 1-D and 2-D polyspectrum phase-based methods is never reported except for limited simulation results.

Parametric MP-AP decomposition-based methods, which are free from the phase unwrapping problem, make use of the decomposition $H(z) = H_{\text{MP}}(z) \cdot H_{\text{AP}}(z)$, where $H_{\text{MP}}(z)$ is a minimum-phase system having the same amplitude spectrum with the unknown system $H(z)$, and $H_{\text{AP}}(z)$ is an allpass filter. The estimation of $H_{\text{AP}}(z)$ (phase estimation) follows the estimation of $H_{\text{MP}}(z)$ (amplitude estimation), and existing correlation-based methods are used to obtain an estimate $\hat{H}_{\text{MP}}(z)$. Tugnait [6] searches for the desired $\hat{H}(z)$ from a finite set \mathcal{S} of all the models spectrally equivalent to $\hat{H}_{\text{MP}}(z)$ such that cumulant functions associated with the desired $\hat{H}(z)$ best match the associated sample cumulant functions of $x(n)$. Note that each member of \mathcal{S} is associated with an allpass filter $H_{\text{AP}}(z) = H(z)/H_{\text{MP}}(z)$. Instead, Chi and Kung [9], [10] process $x(n)$ with the inverse filter $1/\hat{H}_{\text{MP}}(z)$ to obtain a second-order “white process” $\tilde{u}(n)$, which is further processed by each allpass filter $\hat{H}_{\text{AP}}(z)$ (associated with \mathcal{S}) such that a single M th-order cumulant of the output, which is denoted $y(n)$, of the optimum $\hat{H}_{\text{AP}}(z)$ is maximum in absolute value. Chi and Kung’s and Tugnait’s methods can be thought of as phase classification methods, whereas the former is computationally much more efficient than the latter. Two common disadvantages of these methods are as follows. One is that allpass factors get lost in the estimation of $H_{\text{MP}}(z)$ because correlation functions are not only phase blind but also allpass factor blind, and the other is that correlation-based estimators of $H_{\text{MP}}(z)$ are sensitive to additive noise. Giannakis and Mendel [7] estimate $H_{\text{AP}}(z)$ using slices of the sixth-order cumulant function of $\tilde{u}(n)$ through a quite complicated procedure. Recently, Chi and Kung [11], [12]

have identified the parameters of $H_{\text{AP}}(z)$ by maximizing a single M th-order cumulant of $y(n)$ instead of searching it from those $\hat{H}_{\text{AP}}(z)$ associated with \mathcal{S} . This method not only is able to provide an accurate estimate for $\hat{H}_{\text{AP}}(z)$ including the allpass factors of $H(z)$ but also is less sensitive to additive Gaussian noise than MP-AP decomposition-based methods mentioned above. To our knowledge, MP-AP decomposition-based methods are never used for the identification of 2-D LTI systems possibly because of difficulties in the theoretical extension of 1-D methods or extraordinary complexity.

This paper proposes a parametric cumulant-based phase-estimation method that estimates the phase response $\theta(\omega)$ of the unknown 1-D system $h(n)$ by processing $x(n)$ with an optimum allpass filter such that a single M th-order ($M \geq 3$) cumulant of the allpass filter output is maximum in absolute value. The proposed method neither involves the amplitude estimation of $h(n)$ as MP-AP decomposition-based methods do nor involves the use of polyspectra phase of $x(n)$ as do polyspectrum phase-based methods. Therefore, the proposed method is less sensitive to additive Gaussian noise than most MP-AP decomposition-based methods because, as mentioned above, the latter resort to correlation-based methods for amplitude estimation. Moreover, the proposed method is free from the phase unwrapping problem of polyspectrum phase-based methods. Furthermore, the proposed method is also extended to the case of 2-D system phase estimation.

The organization of the paper is as follows. Section II presents two parametric allpass models for the allpass filter including a well-known ARMA model and a Fourier series-based model. Section III presents the new parametric cumulant-based phase-estimation method, including one algorithm for 1-D LTI systems and one algorithm for 2-D LTI systems using the allpass models presented in Section II. Then, a performance analysis for the proposed 1-D and 2-D phase-estimation algorithms is presented in Section IV. Some simulation results as well as experimental results with real speech data are then presented to support the proposed 1-D and 2-D phase-estimation algorithms in Section V. Finally, the paper concludes with a discussion and some conclusions.

II. PARAMETRIC ALLPASS MODELS

For notational simplicity, the same notations $H_p(z)$ and $\varphi_p(\omega)$ are used in this section without confusion for the transfer function and phase of parametric allpass models with p parameters, respectively. Moreover, the frequency response of any LTI system $H(z)$ is simply denoted as $H(\omega) = H(z = e^{j\omega})$. Two parametric allpass models, which will be used for the phase estimation of 1-D and 2-D nonminimum phase LTI systems in the next section, are the well-known ARMA model [25] and a Fourier series-based model [22], which are presented, respectively, as follows.

A. ARMA Allpass Model

It is known that a real p th-order 1-D ARMA allpass filter $H_p(z)$ has the following transfer function [25]

$$H_p(z) = \frac{z^{-p}A(z^{-1})}{A(z)} \quad (1)$$

where $A(z)$ is a p th-order polynomial of z^{-1} with real coefficients, i.e.,

$$A(z) = 1 + a_1 z^{-1} + \dots + a_p z^{-p}. \quad (2)$$

It can be easily seen that $|H_p(\omega)| = 1$ and that if a is a pole of $H_p(z)$ [i.e., a root of $A(z)$], then $1/a$ must be a zero of $H_p(z)$. When $A(z)$ is minimum phase (i.e., all the roots of $A(z)$ are inside the unit circle), $H_p(z)$ is a causal stable allpass filter; when $A(z)$ is maximum phase (all the roots of $A(z)$ are outside the unit circle), $H_p(z)$ is an anticausal stable allpass filter. Note that $H_p(z)$ cannot have zeros on the unit circle; otherwise, it becomes unstable. Moreover, when $H_p(z)$ is causal (anticausal) stable, the inverse filter $1/H_p(z)$ (which is also an ARMA allpass filter) is anticausal (causal) stable.

Assume that $\Phi(\omega)$ is the phase response of an arbitrary real allpass filter and that $\Phi(\omega)$ is known *a priori*. Then, one can find an allpass filter $H_p(z)$ given by (1) such that

$$\varphi_p(\omega) = \arg \{H_p(\omega)\} = \arg \left\{ \frac{e^{-j\omega p} \cdot A^*(\omega)}{A(\omega)} \right\} \quad (3)$$

approximates $\Phi(\omega)$ by using IIR filter design techniques such as Deczky's nonlinear approximation method [25], [26].

B. Fourier Series-Based Allpass Model

Because the unwrapped phase of a real filter is a periodic odd function with period equal to 2π , one can model a 1-D allpass filter $H_p(z)$ as

$$H_p(\omega) = \exp \{j\varphi_p(\omega)\} = \exp \{-j\varphi_p(-\omega)\} \quad (4)$$

where

$$\varphi_p(\omega) = \sum_{k=1}^p a_k \cdot \sin(k\omega). \quad (5)$$

Note that the allpass model given by (4) and (5) is always stable due to $|H_p(\omega)| = 1$ for all ω (i.e., the region of convergence of $H_p(z)$ includes the unit circle). Remark that Dianat and Raghuveer [22] use the Fourier series-based model for both phase and magnitude of non-Gaussian signals, whereas we use the Fourier series-based model for an allpass system.

It is easy to see that for any arbitrary phase function $\Phi(\omega)$, when

$$a_k = \frac{1}{\pi} \int_{-\pi}^{\pi} \Phi(\omega) \sin(k\omega) d\omega \quad (6)$$

the series $\varphi_p(\omega)$ given by (5) for $p = \infty$ is exactly the Fourier series expansion of $\Phi(\omega)$. It is known [27] that $\varphi_p(\omega)$ with a_k computed by (6) converges to $\Phi(\omega)$ as $p \rightarrow \infty$ in the mean-square-error (MSE) sense.

Next, let us present a real 2-D Fourier series-based allpass model $H_{p1,p2}(z_1, z_2)$ motivated by the proposed 1-D Fourier series-based allpass model as follows [22]:

$$\begin{aligned} H_{p1,p2}(\omega_1, \omega_2) &= \exp \{j\varphi_{p1,p2}(\omega_1, \omega_2)\} \\ &= \exp \{-j\varphi_{p1,p2}(-\omega_1, -\omega_2)\} \end{aligned} \quad (7)$$

where

$$\begin{aligned} \varphi_{p1,p2}(\omega_1, \omega_2) &= \sum_{k=1}^{p1} \sum_{l=-p2}^{p2} a_{k,l} \cdot \sin(k\omega_1 + l\omega_2) \\ &+ \sum_{l=1}^{p2} a_{0,l} \cdot \sin(l\omega_2), \\ &-\pi \leq \omega_1 \leq \pi, \quad -\pi \leq \omega_2 \leq \pi \end{aligned} \quad (8)$$

which is just a direct extension of $\varphi_p(\omega)$ given by (5) with all the redundant terms removed. The 2-D parametric model $\varphi_{p1,p2}(\omega_1, \omega_2)$ given by (8) can be used as an approximation to an arbitrary 2-D phase response $\Phi(\omega_1, \omega_2)$ (of real 2-D LTI systems).

As mentioned above, both the ARMA allpass model and Fourier series-based allpass model can be used to approximate a known phase function $\Phi(\omega)$ with model parameters solved from $\Phi(\omega)$. However, when $\Phi(\omega)$ is not known, these two allpass models can still be used to approximate $\Phi(\omega)$, but model parameters cannot be obtained from $\Phi(\omega)$ any longer. In the next section, we present how these two models are used for the phase estimate of an unknown LTI system with model parameters solved from higher order cumulants.

III. PHASE ESTIMATION OF LTI SYSTEMS BY ALLPASS FILTERING

Let us define the following notations for ease of later use:

$$\mathbf{n} = (n_1, n_2)$$

$$\mathbf{k} = (k_1, k_2)$$

$$\boldsymbol{\omega} = (\omega_1, \omega_2)$$

$$\mathbf{z} = (z_1, z_2)$$

$$H(\boldsymbol{\omega}) = H(\mathbf{z} = (e^{j\omega_1}, e^{j\omega_2}))$$

$$\sum_{\mathbf{k}=-\infty}^{\infty} h(\mathbf{k}) = \sum_{k_1=-\infty}^{\infty} \sum_{k_2=-\infty}^{\infty} h(k_1, k_2)$$

$$\int_{-\pi}^{\pi} f(\boldsymbol{\omega}) d\boldsymbol{\omega} = \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} f(\omega_1, \omega_2) d\omega_1 d\omega_2.$$

Assume that $x(\mathbf{n})$ is the 2-D noisy output signal of an unknown 2-D LTI system $h(\mathbf{n})$ driven by a 2-D non-Gaussian input $u(\mathbf{n})$ as follows:

$$\begin{aligned} x(\mathbf{n}) &= u(\mathbf{n}) * h(\mathbf{n}) + w(\mathbf{n}) \\ &= \sum_{\mathbf{k}=-\infty}^{\infty} u(\mathbf{k}) \cdot h(\mathbf{n} - \mathbf{k}) + w(\mathbf{n}). \end{aligned} \quad (9)$$

Let us make the following assumptions about $u(\mathbf{n})$, $w(\mathbf{n})$ and $h(\mathbf{n})$:

- A1)** $u(\mathbf{n})$ is real, zero-mean, independent identically distributed (i.i.d.) with M th-order ($M \geq 3$) cumulant $\gamma_M \neq 0$.
- A2)** $w(\mathbf{n})$ is zero-mean Gaussian, which can be white or colored with unknown statistics.
- A3)** $u(\mathbf{n})$ and $w(\mathbf{n})$ are statistically independent.
- A4)** $h(\mathbf{n})$ is a real stable LTI system that can be nonminimum phase.

It has been shown in [28] and [29] that the M th-order cumulant function of $x(\mathbf{n})$, i.e., joint cumulant of M random variables $x(\mathbf{n}), x(\mathbf{n} + \mathbf{k}_1), \dots, x(\mathbf{n} + \mathbf{k}_{M-1})$, is given by

$$C_{M,x}(\mathbf{k}_1, \mathbf{k}_2, \dots, \mathbf{k}_{M-1}) = \gamma_M \sum_{\mathbf{n}=-\infty}^{\infty} h(\mathbf{n})h(\mathbf{n} + \mathbf{k}_1) \cdots h(\mathbf{n} + \mathbf{k}_{M-1}) \quad (10)$$

and the M th-order polyspectrum of $x(\mathbf{n})$ (Fourier transform of $C_{M,x}(\mathbf{k}_1, \mathbf{k}_2, \dots, \mathbf{k}_{M-1})$) is given by

$$S_{M,x}(\boldsymbol{\omega}_1, \boldsymbol{\omega}_2, \dots, \boldsymbol{\omega}_{M-1}) = \sum_{\mathbf{k}_1=-\infty}^{\infty} \cdots \sum_{\mathbf{k}_{M-1}=-\infty}^{\infty} C_{M,x}(\mathbf{k}_1, \mathbf{k}_2, \dots, \mathbf{k}_{M-1}) \cdot \exp \left\{ -j \sum_{i=1}^{M-1} \boldsymbol{\omega}_i \cdot \mathbf{k}_i^T \right\} = \gamma_M H(\boldsymbol{\omega}_1) \cdots H(\boldsymbol{\omega}_{M-1}) H^*(\boldsymbol{\omega}_1 + \cdots + \boldsymbol{\omega}_{M-1}). \quad (11)$$

Although the signal model given by (9) and the associated M th-order cumulant function given by (10) and polyspectrum given by (11) are for the 2-D case, they also apply for the 1-D case with \mathbf{n}, \mathbf{k} , and $\boldsymbol{\omega}$ replaced by scalars n, k , and ω , respectively. The parametric cumulant-based phase-estimation method to be presented below is based on the following theorem.

Theorem 1: Assume that $x(\mathbf{n})$ was generated from (9) under the assumptions **A1**) through **A4**). Let $y(\mathbf{n})$ be the output of an allpass filter $H_{AP}(\boldsymbol{\omega}) = \exp\{j\varphi(\boldsymbol{\omega})\}$ with input $x(\mathbf{n})$. Then, the absolute M th-order ($M \geq 3$) cumulant $|C_{M,y}(\mathbf{k}_1 = \mathbf{0}, \dots, \mathbf{k}_{M-1} = \mathbf{0})|$ of $y(\mathbf{n})$ is maximum if and only if

$$\varphi(\boldsymbol{\omega}) = -\theta(\boldsymbol{\omega}) + \boldsymbol{\alpha}\boldsymbol{\omega}^T \quad (12)$$

where $\theta(\boldsymbol{\omega}) = \arg\{H(\boldsymbol{\omega})\}$ is the phase of the unknown system $h(\mathbf{n})$, and $\boldsymbol{\alpha} = (\alpha_1, \alpha_2)$ is an unknown constant row vector. Moreover

$$\begin{aligned} & \text{Max} \{|C_{M,y}(\mathbf{k}_1 = \mathbf{0}, \dots, \mathbf{k}_{M-1} = \mathbf{0})|\} \\ & = \left(\frac{1}{2\pi}\right)^{2(M-1)} \cdot |\gamma_M| \cdot \int_{-\pi}^{\pi} \cdots \int_{-\pi}^{\pi} |H(\boldsymbol{\omega}_1)| \cdots \\ & \quad |H(\boldsymbol{\omega}_{M-1})| \cdot |H(\boldsymbol{\omega}_1 + \cdots + \boldsymbol{\omega}_{M-1})| d\boldsymbol{\omega}_1 \cdots \\ & \quad d\boldsymbol{\omega}_{M-1}. \end{aligned} \quad (13)$$

Proof: It is easy to see that $y(\mathbf{n})$ is the output of the overall system

$$H(\boldsymbol{\omega})H_{AP}(\boldsymbol{\omega}) = |H(\boldsymbol{\omega})| \exp\{j\Phi(\boldsymbol{\omega})\} \quad (14)$$

where

$$\Phi(\boldsymbol{\omega}) = \theta(\boldsymbol{\omega}) + \varphi(\boldsymbol{\omega}). \quad (15)$$

One can easily infer, from (11) and (14), that

$$\begin{aligned} S_{M,y}(\boldsymbol{\omega}_1, \dots, \boldsymbol{\omega}_{M-1}) & = \gamma_M \cdot |H(\boldsymbol{\omega}_1)| \cdots |H(\boldsymbol{\omega}_{M-1})| \\ & \quad \cdot |H(\boldsymbol{\omega}_1 + \cdots + \boldsymbol{\omega}_{M-1})| \\ & \quad \cdot \exp\{j(\Phi(\boldsymbol{\omega}_1) + \cdots + \Phi(\boldsymbol{\omega}_{M-1}) \\ & \quad - \Phi(\boldsymbol{\omega}_1 + \cdots + \boldsymbol{\omega}_{M-1}))\}. \end{aligned} \quad (16)$$

Then, we have

$$\begin{aligned} & |C_{M,y}(\mathbf{0}, \dots, \mathbf{0})| \\ & = \left(\frac{1}{2\pi}\right)^{2(M-1)} \cdot \left| \int_{-\pi}^{\pi} \cdots \int_{-\pi}^{\pi} S_{M,y}(\boldsymbol{\omega}_1, \dots, \boldsymbol{\omega}_{M-1}) \right. \\ & \quad \left. \cdot d\boldsymbol{\omega}_1 \cdots d\boldsymbol{\omega}_{M-1} \right| \\ & = \left(\frac{1}{2\pi}\right)^{2(M-1)} \cdot |\gamma_M| \cdot \left| \int_{-\pi}^{\pi} \cdots \int_{-\pi}^{\pi} |H(\boldsymbol{\omega}_1)| \cdots \right. \\ & \quad \cdot |H(\boldsymbol{\omega}_{M-1})| \cdot |H(\boldsymbol{\omega}_1 + \cdots + \boldsymbol{\omega}_{M-1})| \\ & \quad \cdot \exp\{j(\Phi(\boldsymbol{\omega}_1) + \cdots + \Phi(\boldsymbol{\omega}_{M-1}) \\ & \quad - \Phi(\boldsymbol{\omega}_1 + \cdots + \boldsymbol{\omega}_{M-1}))\} d\boldsymbol{\omega}_1 \cdots d\boldsymbol{\omega}_{M-1} \left| \right. \\ & \leq \left(\frac{1}{2\pi}\right)^{2(M-1)} \cdot |\gamma_M| \cdot \int_{-\pi}^{\pi} \cdots \int_{-\pi}^{\pi} |H(\boldsymbol{\omega}_1)| \cdots \\ & \quad |H(\boldsymbol{\omega}_{M-1})| \cdot |H(\boldsymbol{\omega}_1 + \cdots + \boldsymbol{\omega}_{M-1})| \\ & \quad \cdot d\boldsymbol{\omega}_1 \cdots d\boldsymbol{\omega}_{M-1}. \end{aligned} \quad (17)$$

It is trivial to show that if

$$\Phi(\boldsymbol{\omega}) = \boldsymbol{\alpha} \cdot \boldsymbol{\omega}^T \quad (18)$$

the equality in (17) holds. What remains to be proven is that the equality of (17) leads to (18).

Assume that $\Phi(\boldsymbol{\omega})$ is a continuous function of $\boldsymbol{\omega}$ and that $\Phi(\mathbf{0}) = 0$ without loss of generality. It can be inferred from (17) that the equality of (17) requires

$$\begin{aligned} & \Phi(\boldsymbol{\omega}_1) + \cdots + \Phi(\boldsymbol{\omega}_{M-1}) - \Phi(\boldsymbol{\omega}_1 + \cdots + \boldsymbol{\omega}_{M-1}) \\ & = \beta + 2\pi L, \quad \forall \boldsymbol{\omega}_1, \boldsymbol{\omega}_2, \dots, \boldsymbol{\omega}_{M-1} \end{aligned} \quad (19)$$

where $-\pi \leq \beta \leq \pi$ is a constant and L is an integer. Substituting $\boldsymbol{\omega}_1 = \cdots = \boldsymbol{\omega}_{M-1} = \mathbf{0}$ into (19) yields the result $\beta + 2\pi L = (M-2)\Phi(\mathbf{0}) = 0$, or $L = 0$ and $\beta = 0$. Thus, (19) is equivalent to

$$\Phi(\boldsymbol{\omega}_1 + \cdots + \boldsymbol{\omega}_{M-1}) = \Phi(\boldsymbol{\omega}_1) + \cdots + \Phi(\boldsymbol{\omega}_{M-1}) \quad (20)$$

which implies that $\Phi(\boldsymbol{\omega})$ is a linear function of $\boldsymbol{\omega}$ or that (18) is true. Therefore, we have completed the proof that $|C_{M,y}(\mathbf{k}_1 = \mathbf{0}, \dots, \mathbf{k}_{M-1} = \mathbf{0})|$ is maximum if and only if (12) holds. Meanwhile, the equality of (17) leads to (13). Q.E.D.

Let us emphasize that Theorem 1 is also applicable for the 1-D case with $\mathbf{n}, \mathbf{k}, \boldsymbol{\omega}$, and π replaced by scalars n, k, ω , and π , respectively, and $(1/2\pi)^{2(M-1)}$ replaced by $(1/2\pi)^{(M-1)}$ in (13). Note that Theorem 1 reduces to the corresponding theorem reported in [12] when the unknown system $H(z)$ is a 1-D allpass system.

Without confusion, let us use $x(\mathbf{n})$ to denote the given measurements for both 1-D and 2-D cases. How the new cumulant-based phase estimation method estimates the system phase $\theta(\boldsymbol{\omega})$ is shown in Fig. 1. Let $x(\mathbf{n})$ pass through an allpass model $H_p(\boldsymbol{\omega})$ given by (1) or (4) (for the 1-D case) or $H_{p1,p2}(\boldsymbol{\omega})$ given by (7) (for the 2-D case), and let $y(\mathbf{n})$ be the associated output. By Theorem 1, except for an unknown linear phase term (an unknown time delay), the phase $\theta(\boldsymbol{\omega})$ of the unknown system $h(\mathbf{n})$ can be estimated as

$$\hat{\theta}(\boldsymbol{\omega}) = -\hat{\varphi}(\boldsymbol{\omega}) \quad (21)$$

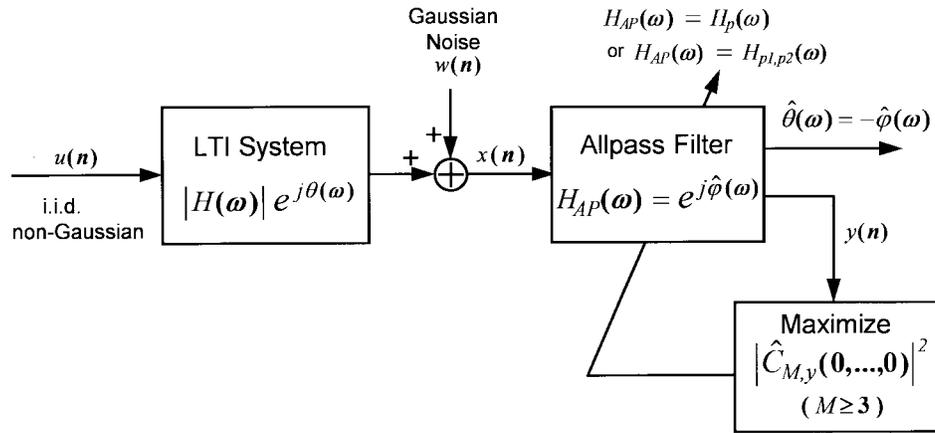


Fig. 1. New cumulant-based phase estimation method.

where $\hat{\varphi}(\omega)$ is the phase of the optimum allpass model obtained by maximizing the following objective function

$$J(\mathbf{a}) = |\hat{C}_{M,y}(\mathbf{0}, \dots, \mathbf{0})|^2 \quad (22)$$

where \mathbf{a} is a column vector containing all parameters of the allpass model used, and $\hat{C}_{M,y}(\mathbf{0}, \dots, \mathbf{0})$, which is the sample cumulant associated with $C_{M,y}(\mathbf{0}, \dots, \mathbf{0})$, can be directly calculated from the allpass model output $y(\mathbf{n})$. For instance, for $M = 3$

$$\hat{C}_{3,y}(\mathbf{0}, \mathbf{0}) = \frac{1}{N} \sum_{\mathbf{n}} y^3(\mathbf{n}) \quad (23)$$

where N is the total number of terms $y^3(\mathbf{n})$ in the summation. Note that $J(\mathbf{a})$ is a highly nonlinear function of \mathbf{a} . Therefore, we have to resort to iterative optimization algorithms for finding the optimum \mathbf{a} .

A popular gradient-type iterative optimization algorithm is considered for finding the maximum of $J(\mathbf{a})$. At the i th iteration, $\hat{\mathbf{a}}$ is updated by

$$\hat{\mathbf{a}}(i) = \hat{\mathbf{a}}(i-1) + \rho \mathbf{g}_{i-1} \quad (24)$$

where ρ is a positive constant, and \mathbf{g}_{i-1} denotes the gradient of $J(\mathbf{a})$ with respect to \mathbf{a} for $\mathbf{a} = \hat{\mathbf{a}}(i-1)$, i.e.,

$$\mathbf{g}_{i-1} = \left. \frac{\partial J(\mathbf{a})}{\partial \mathbf{a}} \right|_{\mathbf{a}=\hat{\mathbf{a}}(i-1)}. \quad (25)$$

However, a local maximum rather than a global maximum of $J(\mathbf{a})$ can be obtained as the algorithm converges. A choice for ρ is $\rho = \eta/2^l$, where $0 \leq l \leq L$ is an integer, and $0 < \eta \leq 1$ is a preassigned constant. If updating $\hat{\mathbf{a}}(i)$ by (24) with $l = 0$ results in $J(\hat{\mathbf{a}}(i)) < J(\hat{\mathbf{a}}(i-1))$, one can repeat the process for $l = 1, 2, \dots$, until $J(\hat{\mathbf{a}}(i)) > J(\hat{\mathbf{a}}(i-1))$. As to the gradient \mathbf{g}_{i-1} , for instance, it can be easily shown [12] from (22), (23), and (25) that for $M = 3$

$$\begin{aligned} \mathbf{g}_{i-1} &= 2 \cdot \hat{C}_{3,y}(\mathbf{0}, \mathbf{0}) \cdot \left. \frac{\partial \hat{C}_{3,y}(\mathbf{0}, \mathbf{0})}{\partial \mathbf{a}} \right|_{\mathbf{a}=\hat{\mathbf{a}}(i-1)} \\ &= 2 \cdot \left\{ \frac{1}{N} \sum_{\mathbf{n}} y^3(\mathbf{n}) \right\} \\ &\quad \cdot \left\{ \frac{1}{N} \sum_{\mathbf{n}} 3 \cdot y^2(\mathbf{n}) \cdot \left(\frac{\partial y(\mathbf{n})}{\partial \mathbf{a}} \right) \right\} \Bigg|_{\mathbf{a}=\hat{\mathbf{a}}(i-1)} \quad (26) \end{aligned}$$

where the computation of $y(\mathbf{n})$ and $\partial y(\mathbf{n})/\partial \mathbf{a}$, depending on the allpass model used, will be presented later. Next, let us present the new phase-estimation method using the parametric allpass models presented in Section II for 1-D case and 2-D case, respectively.

A. 1-D Phase Estimation

The new phase estimation method for the 1-D case using either the ARMA allpass model given by (1) or the Fourier series-based allpass model given by (4) is implemented by the following algorithm.

Algorithm 1:

- S1)** Set p_{\max} (the maximum of p), integer increment parameter $s \geq 1$, and convergence parameter ξ (a small positive number). Choose the causal stable or anticausal stable allpass model $H_p(z)$ given by (1) (ARMA model) or that given by (4) (Fourier series-based model).
- S2)** Set $t = 1$ (iteration number), $p = s$, and $\mathbf{a}_p = (a_1, a_2, \dots, a_p)^T$, which contains all the coefficients of the allpass model $H_p(z)$ used. Search for the maximum of $J(\mathbf{a}_p)$ by the above iterative algorithm with the initial condition $\hat{\mathbf{a}}_p(0) = \mathbf{0}_p$, where $\mathbf{0}_p$ is a $p \times 1$ column vector containing p zeros.
- S3)** Set $t = t + 1, p = s \cdot t$. Search for the maximum of $J(\mathbf{a}_p)$ by the above iterative algorithm with $\hat{\mathbf{a}}_p(0) = (\hat{\mathbf{a}}_{p-s}^T, \mathbf{0}_s^T)^T$.
- S4)** If $p \leq p_{\max}$ and $|J(\hat{\mathbf{a}}_p) - J(\hat{\mathbf{a}}_{p-s})|/J(\hat{\mathbf{a}}_p) \geq \xi$, then go to **S3**; otherwise, stop.

The optimum phase estimate $\hat{\theta}(\omega)$ is then obtained [see (21)] as

$$\hat{\theta}(\omega) = -\arg \{ \hat{H}_p(\omega) \} = -\hat{\varphi}_p(\omega). \quad (27)$$

For ease of later use, *Algorithm 1* is referred to as *Algorithm 1 ARMA-Causal* and *Algorithm 1 ARMA-Anticausal* when the ARMA allpass model chosen in **S1** is causal and anticausal, respectively; *Algorithm 1* is referred to as *Algorithm 1 Fourier* when the Fourier series-based allpass model is chosen in **S1**. Note that to compute $J(\mathbf{a}_p)$ given by (22) and the gradient $\partial J(\mathbf{a}_p)/\partial \mathbf{a}_p$ (see (26) for the case $M = 3$) in **S2** and **S3** requires computation of $y(\mathbf{n})$ and $\partial y(\mathbf{n})/\partial \mathbf{a}_p$, which depend on the allpass model chosen in **S1**. Next, let us discuss how to compute $y(\mathbf{n})$ and $\partial y(\mathbf{n})/\partial \mathbf{a}_p$ for *Algorithm 1 ARMA-Causal*,

Algorithm 1 ARMA-Anticausal, and then *Algorithm 1 Fourier*, respectively.

In association with *Algorithm 1 ARMA-Causal* and *Algorithm 1 ARMA-Anticausal*, how to compute $y(n)$ and $\partial y(n)/\partial \mathbf{a}_p$ has been reported in [12]. Basically, when a causal stable allpass model is used, both $y(n)$ and $\partial y(n)/\partial \mathbf{a}_p$ are obtained through a forward processing as follows:

$$y(n) = -\sum_{k=1}^p a_k y(n-k) + x(n-p) + \sum_{k=1}^p a_k x(n+k-p) \quad (28)$$

$$\frac{\partial y(n)}{\partial a_m} = -y(n-m) - \sum_{k=1}^p a_k \frac{\partial y(n-k)}{\partial a_m} + x(n+m-p), \quad m = 1, 2, \dots, p. \quad (29)$$

On the other hand, when an anticausal stable allpass model is used, they are obtained through a backward processing. Refer to [12] for details.

Regarding the computation of $y(n)$ and $\partial y(n)/\partial \mathbf{a}_p$ required by **S2** and **S3** in *Algorithm 1 Fourier*, the former can be simply obtained either by computing $y(n) = x(n) * h_p(n)$, where $h_p(n)$ is the inverse FFT of $H_p(\omega)$, or by taking inverse FFT of $Y(\omega) = X(\omega) \cdot H_p(\omega)$, and the latter can be simply obtained by

$$\frac{\partial y(n)}{\partial a_m} = \frac{1}{2} \{y(n+m) - y(n-m)\}, \quad m = 1, 2, \dots, p. \quad (30)$$

The proof for the expression of $\partial y(n)/\partial a_m$ given by (30) is given in Appendix A.

Some worthy remarks for the proposed 1-D phase estimation algorithm are given as follows:

- R1)** The iterative search algorithm used in **S2** and **S3** guarantees the increase of $J(\mathbf{a}_p)$ whenever $\hat{\mathbf{a}}_p$ is updated. Moreover, for *Algorithm 1 ARMA-Causal* (or *ARMA-Anticausal*), the obtained allpass model $\hat{H}_p(z)$ in **S2** and **S3** must also be causal stable (or anticausal stable) as chosen in **S1** in addition to the increase of $J(\mathbf{a}_p)$. On the other hand, $J(\mathbf{a}_p)$ is bounded because $|C_{M,y}(0, \dots, 0)|$ is bounded by Theorem 1 [see (13)]. Therefore, the convergence of the proposed 1-D phase-estimation algorithm is guaranteed, but as with other nonlinear optimization algorithms, it may converge to a local optimum solution.
- R2)** The number p of the allpass model parameters is increased by s for each iteration (each t). Two reasons for this are as follows. From our experience, *Algorithm 1* often converges faster for $s > 1$ than for $s = 1$ with almost the same performance if the associated maximum values $J(\hat{\mathbf{a}}_p)$ are close to each other. The other reason is that the desired optimum solution can be chosen according to the resultant $J(\hat{\mathbf{a}}_p)$ from a set of solutions obtained by *Algorithm 1* with different values of s and in order to avoid some local optimum solutions.
- R3)** Chi and Kung [12] proposed a cumulant-based allpass system identification algorithm that estimates the phase of an unknown causal stable allpass system as

well by maximizing $J(\mathbf{a}_p)$. The obtained optimum anticausal ARMA allpass filter turns out to be the inverse filter of the unknown allpass system. However, the optimum allpass filter obtained by the proposed 1-D phase-estimation algorithm can be thought of as an optimum phase equalizer to remove the phase distortion of the unknown (nonminimum phase) LTI system $H(z)$, which itself can also be an allpass system. In other words, Chi and Kung's allpass system identification algorithm is a special case of *Algorithm 1 ARMA-Anticausal* when $H(z)$ is an allpass system.

- R4)** The proposed 1-D phase-estimation algorithm has a computationally efficient parallel structure in computing $\partial y(n)/\partial a_m$ [see (29) and (30)]. The parallel structure associated with *Algorithm 1 ARMA-Causal* is shown in Fig. 2(a), and that associated with *Algorithm 1 Fourier* is shown in Fig. 2(b). However, the latter is computationally faster than the former because $\partial y(n)/\partial a_m$ given by (30) is nothing but the output of an FIR filter with only two nonzero coefficients (1/2 and -1/2) driven by $y(n)$, whereas that given by (29) is the output of a p th-order IIR filter $1/A(z)$ driven by both $x(n)$ and $y(n)$.

B. 2-D Phase Estimation:

The new phase estimation method for the 2-D case using the proposed 2-D Fourier series-based allpass model given by (7) is implemented by the following algorithm:

Algorithm 2:

- S1)** Set p_1, p_2 , and let \mathbf{a} be a column vector containing all the coefficients of $\varphi_{p_1, p_2}(\omega)$ given by (8).
- S2)** Search for the maximum of $J(\mathbf{a})$ by the the above iterative algorithm with the initial condition $\hat{\mathbf{a}}(0) = \mathbf{0}$.

The optimum phase estimate $\hat{\theta}(\omega)$ is again obtained [see (21)] as

$$\hat{\theta}(\omega) = -\arg \{ \hat{H}_{p_1, p_2}(\omega) \} = -\hat{\varphi}_{p_1, p_2}(\omega). \quad (31)$$

Moreover, the computation of $y(\mathbf{n})$ and $\partial y(\mathbf{n})/\partial \mathbf{a}$ required by **S2**, is basically the same as that associated with *Algorithm 1 Fourier* with the partial derivative $\partial y(\mathbf{n})/\partial a_{r,s}$ given by

$$\frac{\partial y(\mathbf{n})}{\partial a_{r,s}} = \frac{1}{2} \{y(n_1+r, n_2+s) - y(n_1-r, n_2-s)\}. \quad (32)$$

The proof for (32) is similar to that for (30) and thus is omitted here.

There are also some worthy remarks regarding *Algorithm 2* described as follows:

- R5)** *Algorithm 2* can be viewed as an extension of *Algorithm 1 Fourier* for the 2-D case, whereas the 2-D version of *Algorithm 1 ARMA-Causal* (or *ARMA-Anticausal*) is not suggested due to complexity for computing the gradient $\partial J(\mathbf{a})/\partial \mathbf{a}$ and lack of efficient approaches to avoid instability of 2-D ARMA allpass models. The statements described in **R1**, **R3** and **R4** associated with *Algorithm 1* also apply to *Algorithm 2*. In summary, the convergence of *Algorithm 2* is guaranteed; *Algorithm 2* is an optimum phase equalizer to remove the phase distortion of

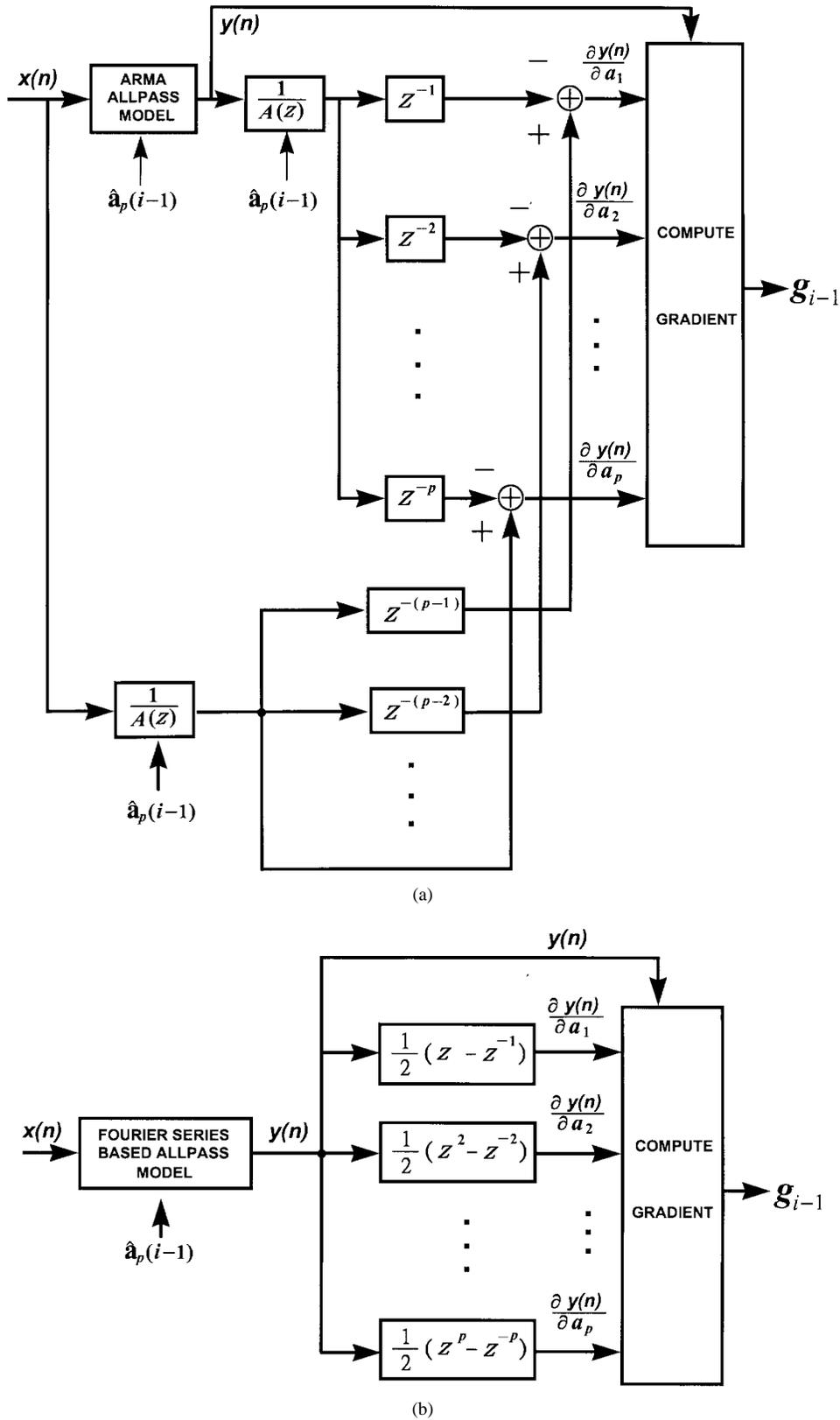


Fig. 2. Parallel structure for computing the gradient of $J(\mathbf{a}_p)$ associated with (a) *Algorithm 1 ARMA-Causal* and (b) *Algorithm 1 Fourier*, respectively.

the unknown (nonminimum phase) LTI system $H(z)$, which itself can be an allpass system; *Algorithm 2* has a parallel structure for efficiently computing $\partial y(\mathbf{n})/\partial a_{r,s}$, which is also the output of an FIR filter

with only two nonzero coefficients ($1/2$ and $-1/2$) [see (32)] driven by $y(\mathbf{n})$.

R6) Note that 2-D LTI systems are generally nonseparable. The total number of unknown coefficients in

$\varphi_{p1,p2}(\boldsymbol{\omega})$ given by (8) is

$$P_{\text{num}} = p1 \cdot (2 \cdot p2 + 1) + p2. \quad (33)$$

If it is known *a priori* that $H(z)$ is separable, i.e.,

$$H(z) = H_1(z_1) \cdot H_2(z_2) \quad (34)$$

which implies

$$\theta(\boldsymbol{\omega}) = \theta_1(\omega_1) + \theta_2(\omega_2) \quad (35)$$

where $\theta_1(\omega)$ and $\theta_2(\omega)$ are the phase of the 1-D systems $H_1(z)$ and $H_2(z)$, respectively, $\varphi_{p1,p2}(\boldsymbol{\omega})$ given by (8) can be reduced to

$$\varphi_{p1,p2}(\boldsymbol{\omega}) = \sum_{k=1}^{p1} a_{k,0} \cdot \sin(k\omega_1) + \sum_{l=1}^{p2} a_{0,l} \cdot \sin(l\omega_2),$$

$$-\pi \leq \omega_1 \leq \pi, \quad -\pi \leq \omega_2 \leq \pi. \quad (36)$$

Then, *Algorithm 1 Fourier* can be employed to estimate $\theta(\boldsymbol{\omega})$ with minor modifications for considerable computational saving because

$$P_{\text{num}} = p1 + p2 \ll p1 \cdot (2 \cdot p2 + 1) + p2$$

as computed using (33) for this case.

- R7)** The computation of $\varphi_{p1,p2}(\boldsymbol{\omega})$ given by (8) can also be performed by FFT algorithms because one can form a 2-D signal from $a_{r,s}$ such that $\varphi_{p1,p2}(\boldsymbol{\omega})$ becomes the imaginary part of the 2-D Fourier transform of the 2-D signal.

Let us conclude this section with the following remark, which summarizes major distinctions of the proposed 1-D and 2-D phase-estimation algorithms and polyspectrum phase-based methods as well as MP-AP decomposition-based methods.

- R8)** *Algorithms 1* and *2*, which estimate the system phase [see (27) and (31)] by maximizing a single absolute M th-order cumulant of the used allpass model output $y(n)$ or $y(\mathbf{n})$ with no need to perform amplitude estimation of the unknown system $h(n)$ or $h(\mathbf{n})$, as well as MP-AP decomposition-based methods are free from the phase unwrapping problem of polyspectrum-based methods since the linear relationship between the system phase and polyspectra phase of measurements $x(n)$ or $x(\mathbf{n})$ is never involved. Moreover, they are less sensitive to additive Gaussian noise than most MP-AP decomposition-based methods when SNR is low because the latter resort to the correlations of $x(n)$ or $x(\mathbf{n})$ (sum of correlations of noise-free measurements and those of additive noise) for amplitude estimation of $h(n)$ or $h(\mathbf{n})$, as mentioned in the introduction section. On the other hand, when $H_{\text{MP}}(z)$ can be accurately estimated, MP-AP decomposition-based methods can perform better than the proposed cumulant-based phase-estimation method. For instance, when SNR is high, an accurate estimate for the minimum-phase system $H_{\text{MP}}(z)$ can be obtained by correlation-based methods, thus leading to an accurate estimate for $H(z)$ obtained from the finite set of all the models spectrally equivalent

to the estimated $H_{\text{MP}}(z)$. Finally, the proposed 1-D and 2-D phase-estimation algorithms are neither linear estimators nor simple recursive estimators such as those reported in [14]–[16] and thus require larger computational load than most linear or recursive estimators.

IV. PERFORMANCE ANALYSIS

In this section, let us present a performance analysis for the proposed cumulant-based phase-estimation algorithm, *Algorithm 1 (ARMA-Causal, ARMA-Anticausal and Fourier)* used for the 1-D case. The performance analysis associated with *Algorithm 1* basically applies to *Algorithm 2* as well because they were developed based on the same philosophy except that the latter is used for the 2-D case.

It is known [25] that the phase [given by (3)] of the ARMA allpass model $H_p(z)$ used by *Algorithm 1 ARMA-Causal* and *Algorithm 1 ARMA-Anticausal* is continuous, whereas the group delay

$$\text{grd} \{H_p(\omega)\} = -\partial(\arg \{H_p(\omega)\})/\partial\omega > 0$$

for all ω when $H_p(z)$ is causal stable, and $\text{grd} \{H_p(\omega)\} < 0$ for all ω when $H_p(z)$ is anticausal stable. It is also known that the phase of the Fourier series-based allpass model [see (5)] used by *Algorithm 1 Fourier* is continuous [27]. Moreover, the absolute M th-order cumulant $|C_{M,y}(0, \dots, 0)|$ of the allpass filter output $y(n)$ is invariant for either of $y(n)$ [or $x(n)$] and $-y(n)$ [or $-x(n)$], i.e., $|C_{M,y}(0, \dots, 0)|$ is invariant for either of $H(z)$ and $-H(z)$, but $\varphi_p(\omega = 0) = 0$ or $H_p(z = 1) = 1$ no matter which allpass model is used by *Algorithm 1*. Therefore, *Algorithm 1*, which tries to maximize $|C_{M,y}(0, \dots, 0)|$ [see (22)], has the following property:

- P1)** The optimum phase estimate $\hat{\theta}(\omega)$ given by (27) is continuous, although the true system phase $\theta(\omega)$ itself can have discontinuities; meanwhile, $\hat{\theta}(\omega)$ is blind to a constant π when $\theta(\omega = 0) = \pi$.

Let $e(\omega)$ denote the phase-estimation error associated with *Algorithm 1*, i.e.,

$$e(\omega) = \theta(\omega) - \hat{\theta}(\omega) = \theta(\omega) + \hat{\varphi}_p(\omega) \quad (37)$$

where $\hat{\varphi}_p(\omega)$ is the phase of the optimum allpass filter obtained by *Algorithm 1*. Note that $e(\omega)$ is also the phase of the overall system $H(z)\hat{H}_p(z)$; therefore, $e(\omega)$ is equal to $\Phi(\boldsymbol{\omega})$ given by (15) with $\boldsymbol{\omega}$ replaced by ω and $\varphi(\boldsymbol{\omega})$ replaced by $\hat{\varphi}_p(\omega)$.

It can be easily inferred by (16) that the optimum M th-order polyspectrum $S_{M,y}(\omega_1, \dots, \omega_{M-1})$ of the output $y(n)$ of the optimum allpass filter $\hat{H}_p(z)$ is given by

$$S_{M,y}(\omega_1, \dots, \omega_{M-1}) = |S_{M,y}(\omega_1, \dots, \omega_{M-1})| \cdot \exp \{j\mathcal{E}(\omega_1, \dots, \omega_{M-1})\} \quad (38)$$

where

$$|S_{M,y}(\omega_1, \dots, \omega_{M-1})| = |\gamma_M| \cdot |H(\omega_1)| \cdots |H(\omega_{M-1})| \cdot |H(\omega_1 + \cdots + \omega_{M-1})| \quad (39)$$

and

$$\mathcal{E}(\omega_1, \dots, \omega_{M-1}) = e(\omega_1) + \dots + e(\omega_{M-1}) - e(\omega_1 + \dots + \omega_{M-1}). \quad (40)$$

As mentioned in **R3**, the optimum allpass filter obtained by *Algorithm 1* is an optimum phase equalizer that implies that $e(\omega)$ and $\mathcal{E}(\omega_1, \dots, \omega_{M-1})$ can be thought of as phase residual and M th-order polyspectrum phase residual of the overall system $H(z) \cdot \hat{H}_p(z)$, respectively. Moreover, the M th-order polyspectrum phase residual $\mathcal{E}(\omega_1, \dots, \omega_{M-1}) = 0$ when $e(\omega) = \alpha\omega$, i.e., $\mathcal{E}(\omega_1, \dots, \omega_{M-1})$ is blind to any linear phase term in $e(\omega)$. Therefore, without loss of generality, let us assume that the unknown linear phase term in $e(\omega)$ has been removed in the following performance analysis.

Because $S_{M,y}(\omega_1, \dots, \omega_{M-1})$ is the $(M-1)$ -D Fourier transform of the real M th-order cumulant function $C_{M,y}(k_1, \dots, k_{M-1})$ of $y(n)$

$$S_{M,y}(\omega_1, \dots, \omega_{M-1}) = S_{M,y}^*(-\omega_1, \dots, -\omega_{M-1})$$

(Hermitian), and thus, the M th-order cumulant $C_{M,y}(0, \dots, 0)$ can be expressed as

$$\begin{aligned} C_{M,y}(0, \dots, 0) &= \left(\frac{1}{2\pi}\right)^{M-1} \int_{-\pi}^{\pi} \dots \int_{-\pi}^{\pi} S_{M,y}(\omega_1, \dots, \omega_{M-1}) \\ &\quad \cdot d\omega_1 \dots d\omega_{M-1} \\ &= \left(\frac{1}{2\pi}\right)^{M-1} \int_{-\pi}^{\pi} \dots \int_{-\pi}^{\pi} |S_{M,y}(\omega_1, \dots, \omega_{M-1})| \\ &\quad \cdot \cos(\mathcal{E}(\omega_1, \dots, \omega_{M-1})) d\omega_1 \dots d\omega_{M-1} \end{aligned} \quad (41)$$

which further leads to the simplification of $|C_{M,y}(0, \dots, 0)|$ (maximized by *Algorithm 1*) as follows:

$$\begin{aligned} |C_{M,y}(0, \dots, 0)| &= \left(\frac{1}{2\pi}\right)^{M-1} \cdot \left| \int_{-\pi}^{\pi} \dots \int_{-\pi}^{\pi} |S_{M,y}(\omega_1, \dots, \omega_{M-1})| \right. \\ &\quad \cdot \cos(\mathcal{E}(\omega_1, \dots, \omega_{M-1})) d\omega_1 \dots d\omega_{M-1} \left. \right| \\ &\approx \left(\frac{1}{2\pi}\right)^{M-1} \cdot \left| \int_{-\pi}^{\pi} \dots \int_{-\pi}^{\pi} |S_{M,y}(\omega_1, \dots, \omega_{M-1})| \right. \\ &\quad \cdot \left\{ 1 - \frac{1}{2} \mathcal{E}^2(\omega_1, \dots, \omega_{M-1}) \right\} d\omega_1 \dots d\omega_{M-1} \left. \right| \end{aligned} \quad (42)$$

to the second-order Taylor series approximation assuming that $|\mathcal{E}(\omega_1, \dots, \omega_{M-1})| \ll 1$. Therefore, maximizing $|C_{M,y}(0, \dots, 0)|$ given by (42) is equivalent to minimizing the following cost function:

$$\mathcal{F} = \int_{-\pi}^{\pi} \dots \int_{-\pi}^{\pi} |S_{M,y}(\omega_1, \dots, \omega_{M-1})| \cdot \mathcal{E}^2(\omega_1, \dots, \omega_{M-1}) d\omega_1 \dots d\omega_{M-1} \quad (43)$$

which implies the following property of *Algorithm 1*:

P2) The phase estimate $\hat{\theta}(\omega)$ is a weighted least squares (WLS) estimate [30] by minimizing squares of polyspectrum phase residual $\mathcal{E}(\omega_1, \dots, \omega_{M-1})$ with absolute M th-order polyspectrum of $y(n)$ (which also equals absolute M th-order polyspectrum of $x(n)$) as the weighting function.

When $H(z)$ is an allpass system, i.e., $|S_{M,y}(\omega_1, \dots, \omega_{M-1})| = |\gamma_M|$, the cost function \mathcal{F} given by (43) reduces to

$$\begin{aligned} \mathcal{F} &= |\gamma_M| \int_{-\pi}^{\pi} \dots \int_{-\pi}^{\pi} \mathcal{E}^2(\omega_1, \dots, \omega_{M-1}) d\omega_1 \dots d\omega_{M-1} \\ &= |\gamma_M| \int_{-\pi}^{\pi} \dots \int_{-\pi}^{\pi} [e(\omega_1) + \dots + e(\omega_{M-1}) \\ &\quad - e(\omega_1 + \dots + \omega_{M-1})]^2 d\omega_1 \dots d\omega_{M-1} \\ &= |\gamma_M| \int_{-\pi}^{\pi} \dots \int_{-\pi}^{\pi} [e^2(\omega_1) + \dots + e^2(\omega_{M-1}) \\ &\quad + e^2(\omega_1 + \dots + \omega_{M-1})] d\omega_1 \dots d\omega_{M-1} \\ &= M \cdot |\gamma_M| \cdot \int_{-\pi}^{\pi} \dots \int_{-\pi}^{\pi} e^2(\omega_1) d\omega_1 \dots d\omega_{M-1} \\ &= M(2\pi)^{M-2} \cdot |\gamma_M| \cdot \int_{-\pi}^{\pi} e^2(\omega) d\omega \end{aligned} \quad (44)$$

where we have used the following results in the derivation of (44):

$$\begin{aligned} &\int_{-\pi}^{\pi} \dots \int_{-\pi}^{\pi} e(\omega_i) e(\omega_j) d\omega_1 \dots d\omega_{M-1} \\ &= (2\pi)^{M-3} \left\{ \int_{-\pi}^{\pi} e(\omega) d\omega \right\}^2 = 0, \quad \text{for all } i \neq j \end{aligned}$$

and

$$\begin{aligned} &\int_{-\pi}^{\pi} \dots \int_{-\pi}^{\pi} e(\omega_i) e(\omega_1 + \dots + \omega_{M-1}) d\omega_1 \dots d\omega_{M-1} \\ &= \int_{-\pi}^{\pi} \dots \int_{-\pi}^{\pi} e(\omega_i) \cdot \left\{ \int_{-\pi}^{\pi} e(\omega_1 + \dots + \omega_{M-1}) d\omega_j \right\} \\ &\quad \cdot d\omega_1 \dots d\omega_{j-1} d\omega_{j+1} \dots d\omega_{M-1} = 0, \quad \text{for all } j \neq i \end{aligned}$$

since $e(\omega)$ given by (37) is an odd function (i.e., $e(\omega) = -e(-\omega)$). The cost function \mathcal{F} given by (44) implies the following property of *Algorithm 1*:

P3) When the unknown system $H(z)$ is an allpass system, the phase estimate $\hat{\theta}(\omega)$ is a least squares (LS) estimate by minimizing squares of the overall system phase residual $e(\omega)$.

Next, let us turn back to the case that $H(z)$ is an arbitrary LTI system. It can be shown that the cost function \mathcal{F} given by (43) satisfies the following inequality:

$$\begin{aligned} \mathcal{F} &\leq |\gamma_M| \cdot M^2 \cdot \left\{ \int_{-\pi}^{\pi} |H(\omega)|^2 d\omega \right\} \\ &\quad \cdot \left\{ \int_{-\pi}^{\pi} |H(\omega)| d\omega \right\}^{M-3} \left\{ \int_{-\pi}^{\pi} |H(\omega)| \cdot e^2(\omega) d\omega \right\}. \end{aligned} \quad (45)$$

The proof for (45) is given in Appendix B. Note that the inequality of (45) qualitatively implies (rather than rigorously proves) the following property of *Algorithm 1*:

P4) The smaller $\int_{-\pi}^{\pi} |H(\omega)| e^2(\omega) d\omega$ (weighted squares of $e(\omega)$), the smaller the cost function \mathcal{F} will be.

Therefore, it can be predicted that as any WLS estimator [30], the absolute phase estimation error $|e(\omega)|$ associated with *Algorithm 1* is smaller (larger) when the weight $|H(\omega)|$ is larger (smaller). This is consistent with the fact that larger $|H(\omega)|$ leads to higher SNR for the frequency component ω in data $x(n)$ and, therefore, leads to smaller $|e(\omega)|$.

Let us discuss a limiting case of smaller $|H(\omega)|$ leading to larger phase-estimation error as mentioned in **P4**). Assume that $H(z)$ has a pair of zeros $e^{\pm j\omega_0}$ on the unit circle, i.e., $|H(\pm\omega_0)| = 0$, which results in a pair of discontinuities of magnitude equal to π in the true system phase $\theta(\omega)$ for $\omega = \pm\omega_0$. Therefore, it is impossible to recover the system phase for $\omega = \pm\omega_0$ simply because $x(n)$ does not carry any phase information about $\omega = \pm\omega_0$. In other words, the phase estimates obtained by *Algorithm 1* always show larger phase-estimation error in the vicinity of discontinuities of the true system phase $\theta(\omega)$. Therefore, *Algorithm 1* has the following property:

P5) When the system phase has discontinuities due to zeros on the unit circle, the absolute phase-estimation error $|e(\omega)|$ is always large in the vicinity of discontinuities even if $\text{SNR} = \infty$.

Next, let us discuss the consistency property of *Algorithm 1*. Assume that $H(z)$ does not have any zeros on the unit circle (i.e., the system phase $\theta(\omega)$ is continuous for all ω). It is well known [27] that the periodic continuous system phase can be expressed as a Fourier series (as given by (5) and (6) for $p \rightarrow \infty$) with uniform convergence. Moreover, it is known that the continuous system phase can also be modeled by the phase of an ARMA allpass filter given by (1) of sufficient order p except for a constant group delay [25]. Moreover, the sample cumulant $\hat{C}_{M,y}(0, \dots, 0)$ is known to be a consistent estimate of $C_{M,y}(0, \dots, 0)$ [3]. The above discussion leads to the following property of *Algorithm 1*:

P6) Assume that $H(z)$ does not have any zeros on the unit circle. The phase estimate $\hat{\theta}(\omega)$ [which is continuous as mentioned in **P1**)] is a consistent estimate except for a linear phase term as $p \rightarrow \infty$. In other words, if the continuous system phase $\theta(\omega)$ can be exactly modeled as either of (3) or (5) for a finite $p = p'$, the obtained phase estimate with $p \geq p'$ is a consistent estimate except for a linear phase term.

As mentioned above, the properties **P1**)–**P6**) also apply to *Algorithm 2* for the 2-D case since both *Algorithms 1* and *2* were developed based on the same philosophy. Next, let us present some simulation results and experimental results to support the proposed cumulant based 1-D and 2-D phase-estimation algorithms and the performance analysis presented in this section.

V. SIMULATION AND EXPERIMENTAL RESULTS

The simulation results to be presented for 1-D phase estimation using the proposed *Algorithm 1* include two examples (Examples 1 and 2). Then, some experimental results with real speech data using *Algorithm 1* are presented in Example 3. For 2-D phase estimation, two examples (Examples 4 and 5) using the proposed *Algorithm 2* are presented. Next, let us turn to the three examples for 1-D phase estimation.

A. 1-D Phase Estimation

The first simulation example is a performance test on *Algorithm 1*, and the second simulation example is a 1-D LTI system having a pair of zeros on the unit circle. Let us turn to Example 1.

Example 1—Performance Test: The noise-free data were generated by letting a zero-mean, exponentially distributed i.i.d. random sequence $u(n)$ with variance $\sigma_u^2 = 1$ and skewness $\gamma_3 = 2$ input to a nonminimum phase ARMA(3, 2) LTI system with transfer function

$$H(z) = \frac{1 - 2.95z^{-1} + 1.9z^{-2}}{1 - 1.3z^{-1} + 1.05z^{-2} - 0.32z^{-3}} \quad (46)$$

whose poles are located at 0.5 and $0.4 \pm j0.7$ and zeros are located at 0.95 and 2. Magnitude response $|H(\omega)|$ of the system is shown in Fig. 3(a). Then, the noisy data $x(n)$ of length $N = 1024$ were obtained by adding a colored Gaussian noise sequence, which was generated from the output of a second-order highpass FIR filter with coefficients $\{1, -2.333, 0.667\}$ driven by a white Gaussian noise sequence, to the noise-free data for $\text{SNR} = 10$ dB. The cumulant order used was $M = 3$. Thirty independent runs were performed using *Algorithm 1 ARMA-Causal* and *ARMA-Anticausal* with $p_{\max} = 5$ and $s = 5$ and *Algorithm 1 Fourier* with $p_{\max} = 4$ and $s = 4$, respectively, and the obtained 30 phase estimates of the system were plotted in an overlaid fashion to indicate the variability of *Algorithm 1*.

Fig. 3(b)–(d) shows 30 continuous phase estimates obtained by *Algorithm 1 ARMA-Causal*, *Algorithm 1 ARMA-Anticausal* and *Algorithm 1 Fourier*, respectively, with unknown linear phase factors artificially removed as well as constant π artificially compensated (since $\theta(\omega = 0) = \pi$ due to $H(z = 1) = -0.1163 < 0$) [see **P1**]. The respective averages of the 30 phase estimates associated with Fig. 3(b)–(d) are depicted by a dashed line, a dash-dot line, and a dotted line, respectively, in Fig. 3(e), together with the true continuous phase (solid line). One can observe from these figures that mean and standard deviation of the phase estimates obtained by the proposed 1-D phase estimation algorithm are smaller for those frequencies where the magnitude response of the system [shown in Fig. 3(a)] is larger. This is consistent with **P1**) and **P4**). This is particularly manifest in the vicinity of $\omega = 1.08$, where $|H(\omega)|$ is maximum, and in the vicinity of $\omega = 0$, where $|H(\omega)|$ is quite small. Moreover, a considerable phase estimation error $|e(\omega)|$ for $0 \leq \omega < 0.5$ can be observed in Fig. 3(e) (dash-dot line and dotted line) because the magnitude response $|H(\omega)|$ for $0 \leq \omega < 0.5$ is much smaller than that for $0.5 \leq \omega < \pi$ [see Fig. 3(a)]. The simulation results shown in Fig. 3(b) (which is associated with *Algorithm 1 ARMA-Causal*) are slightly better than those shown in Fig. 3(d) (which is associated with *Algorithm 1 Fourier*), which are also slightly better than those shown in Fig. 3(c) (which is associated with *Algorithm 1 ARMA-Anticausal*) according to their respective bias and variance. Nevertheless, the simulation results shown in Fig. 3(b)–(e) demonstrate that the proposed 1-D phase estimation algorithm can be used to estimate the system phase of 1-D LTI systems.

Example 2—System Phase with Discontinuities: This example examines the performance of the proposed 1-D phase estimation algorithm when the system phase has a pair of

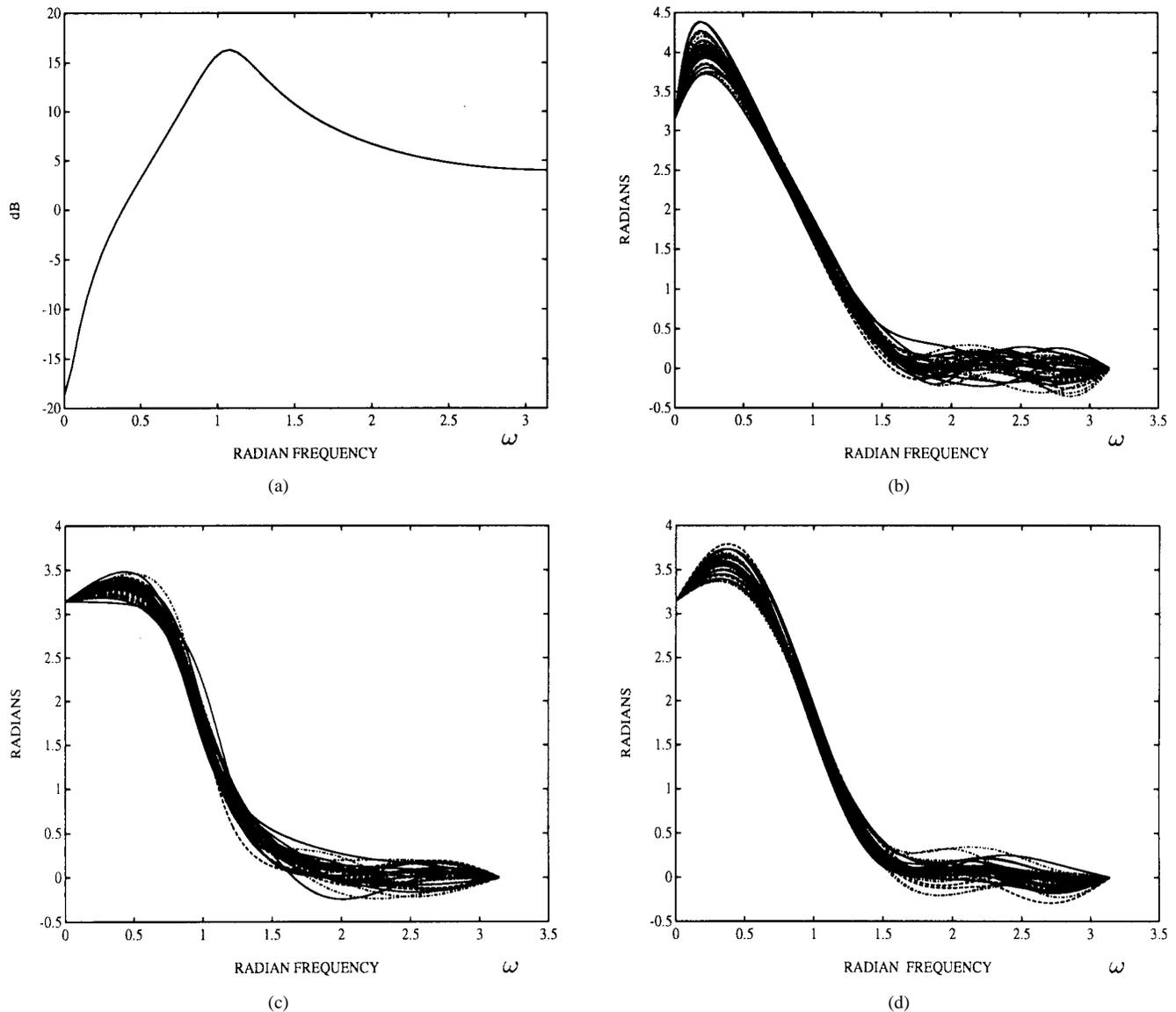


Fig. 3. 1-D simulation results for $M = 3$ (cumulant order), SNR = 10 dB and $N = 1024$ associated with Example 1. (a) Amplitude response $|H(\omega)|$ of the system; 30 continuous phase estimates $\hat{\theta}(\omega)$ obtained by (b) *Algorithm 1 ARMA-Causal* with $p_{\max} = 5$ and $s = 5$, (c) *Algorithm 1 ARMA-Anticausal* also with $p_{\max} = 5$ and $s = 5$, and (d) *Algorithm 1 Fourier* with $p_{\max} = 4$ and $s = 4$, respectively.

discontinuities of magnitude equal to π in the frequency interval $(-\pi, \pi)$ due to a pair of complex conjugate zeros on the unit circle. A nonminimum phase system ARMA(3,4) with transfer function

$$H(z) = \frac{1 - 3.4z^{-1} + 4.81z^{-2} - 3.604z^{-3} + 1.17z^{-4}}{1 - 1.9z^{-1} + 1.1525z^{-2} - 0.1625z^{-3}} \quad (47)$$

was used that has poles located at 0.2 and $0.85 \pm j0.3$ and zeros located at 1.3, 0.9, and $0.6 \pm j0.8 (= e^{\pm j0.9237})$. The magnitude response of the system is shown in Fig. 4(a) from which a spectral null at $\omega = 0.9237$ can be seen. The simulation data were generated (for $N = 1024$ and 4096, SNR = ∞ and 5 dB) following the same procedure as described in Example 1, except that only a single realization of measurements $x(n)$ was generated with measurement noise assumed to be white Gaussian. Then, *Algorithm 1 ARMA-Causal* and *Algorithm 1 Fourier* were employed to estimate

the system phase $\theta(\omega)$ with $p_{\max} = 15$, $M = 3$ or $M = 4$ and $s = 1$ or $s = 5$. Then, the optimum $\hat{\theta}(\omega)$ was chosen from the obtained solutions associated with $s = 1$ and $s = 5$ as mentioned in **R2**.

For comparison, an MP-AP decomposition-based method was also used to estimate $\theta(\omega)$ with the same simulation data. In the first step of the MP-AP decomposition-based method, the minimum-phase $H_{\text{MP}}(z) = B(z)/A(z)$ (which is an ARMA(3,4) model) was estimated using the least squares modified Yule-Walker equations (LSMYWE) [31], [32] for the AR parameters and Durbin's method [31], [33] for the MA parameters. The modified Yule-Walker equations (i.e., linear equations relating the autocorrelation function $r_{xx}(k)$ to AR parameters) for $k = 5, 6, \dots, 16$ were used to obtain $\hat{A}(z)$, and then, $x(n)$ was processed by the system $\hat{A}(z)$ to obtain the MA(4) process $v(n)$. Then, $v(n)$ was processed to obtain an AR model $A'(z)$ of order equal to $L = 100$ using Burg's method [31], [34], and $\hat{B}(z)$ was obtained using the

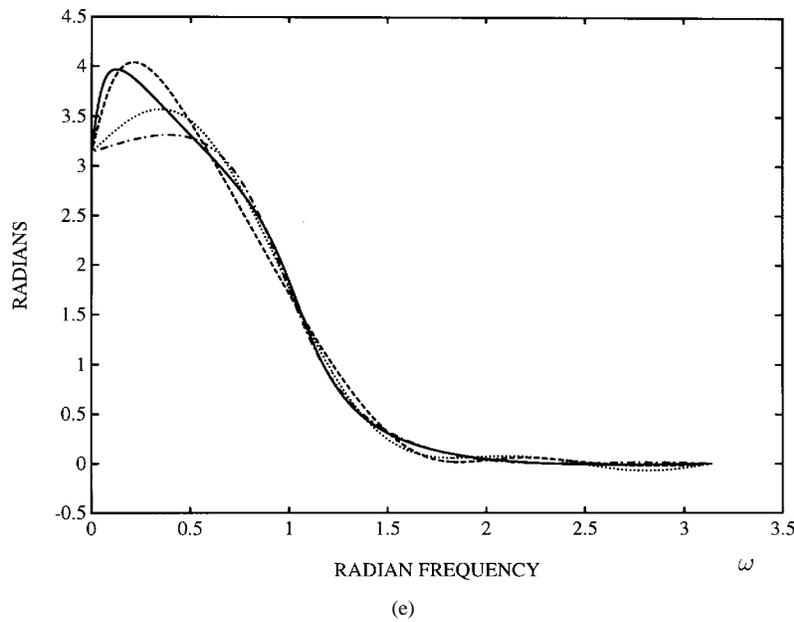


Fig. 3 (continued). 1-D simulation results for $M = 3$ (cumulant order), SNR = 10 dB and $N = 1024$ associated with Example 1. (e) Respective averages of the 30 estimates shown in parts (b)–(d) (dashed line, dash-dot line, and dotted line) together with the true system phase $\theta(\omega)$ (solid line).

autocorrelation method [31], [35] with $\{1, a'(1), \dots, a'(L)\}$ (coefficients of $A'(z)$) as the data. In the second step of the MP-AP decomposition-based method, Chi and Kung's cumulant-based allpass filter classification method [9], [10] was used to determine the optimum $\hat{H}_{AP}(z)$ associated with the set of all the ARMA(3, 4) models spectrally equivalent to $\hat{H}_{MP}(z)$. Finally, the optimum $\hat{\theta}(\omega)$ was obtained as the phase of $\hat{H}(z) = \hat{H}_{MP}(z)\hat{H}_{AP}(z)$.

The simulation results are shown in Fig. 4(b)–(d). For $N = 1024$, SNR = ∞ , the obtained continuous phase estimates $\hat{\theta}(\omega)$ associated with *Algorithm 1 ARMA-Causal* (short dash line for $M = 3$ and long dash line for $M = 4$) and those associated with *Algorithm 1 Fourier* (short dash-dot line for $M = 3$ and long dash-dot line $M = 4$) together with the true system phase $\theta(\omega)$ (solid line) are shown in Fig. 4(b), where unknown linear phase factors were artificially removed, and constant π was artificially compensated due to $\theta(\omega = 0) = \pi$ [see **P1**]. The phase estimate obtained by the MP-AP decomposition-based method is also shown in Fig. 4(b) by a dotted line. Fig. 4(c) and (d) shows the results corresponding to those shown in Fig. 4(b) for SNR = 5 dB, whereas $N = 1024$ for the former, and $N = 4096$ for the latter. One can see, from Fig. 4(b)–(d), that *Algorithm 1* performs better for larger N or higher SNR, and it also performs better for smaller M because the variance of sample cumulants is larger for larger cumulant order [3]. As predicted [see **P4**] and **P5**], the phase estimation error associated with *Algorithm 1* is smaller for all ω , where $\theta(\omega)$ is continuous with larger $|H(\omega)|$, and it is large for those frequencies in the vicinity of the phase discontinuity at $\omega = 0.9237$. On the other hand, the phase estimates [dotted lines in Fig. 4(c) and (d)] associated with the MP-AP decomposition-based method obviously are not good approximations of $\theta(\omega)$ for SNR = 5 dB (low SNR) as mentioned in **R8**). However, one can see from Fig. 4(b) that even for the case of SNR = ∞ , the performance of the MP-AP decomposition-based method is not better than that of *Algorithm 1*. This is also consistent with

R8) because $H_{MP}(z)$ was not accurately estimated though SNR = ∞ . The reason for this is as follows. The power spectrum of $x(n)$ is equal to zero for $\omega = \pm 0.9237$ due to the associated pair of zeros of $H(z)$ on the unit circle, and thus, the estimation of the minimum phase $H_{MP}(z)$ (without zeros on the unit circle) is never perfect even if SNR = ∞ . Globally speaking, these simulation results support that the proposed 1-D phase estimation algorithm can be used to estimate the phase of 1-D LTI systems, regardless of whether or not there are zeros on the unit circle.

Let us conclude this example with more simulation results, which were obtained with the same simulation data through amplitude equalization followed by phase estimation. The data $x(n)$ were preprocessed by the inverse filter $1/\hat{H}_{MP}(z)$, where $\hat{H}_{MP}(z)$ was the one obtained by the above MP-AP decomposition-based method, and thus, the inverse filter output, which is denoted $x'(n)$, can be thought of as the output of the overall system $H(z)/\hat{H}_{MP}(z)$ driven by an i.i.d. non-Gaussian input $u(n)$. Then, *Algorithm 1 ARMA-Causal* with $p_{\max} = 15$, $M = 3$, $s = 1$, or $s = 15$ was employed to process $x'(n)$ to obtain a phase estimate $\hat{\phi}(\omega)$ of $H(z)/\hat{H}_{MP}(z)$, and the phase estimate $\hat{\theta}(\omega)$ of $H(z)$ was obtained by adding the phase of $\hat{H}_{MP}(z)$ to $\hat{\phi}(\omega)$. The obtained phase estimates $\hat{\theta}(\omega)$ (short dash line for $N = 1024$, SNR = ∞ , long dash line for $N = 1024$, SNR = 5 dB, and long dash-dot line for $N = 4096$, SNR = 5 dB) are shown in Fig. 4(e) together with the true system phase $\theta(\omega)$ (solid line). Again, one can see from Fig. 4(e) that *Algorithm 1* performs better for larger N or higher SNR and that these phase estimates are also consistent with the predicted properties **P4**] and **P5**] as well. Note that inaccurate estimation of $H_{MP}(z)$ for the case of SNR = 5 dB, which leads to inaccurate phase estimates [dotted lines in Fig. 4(c) and (d)] associated with the MP-AP decomposition-based method, still ends up with good phase estimates [long dash line and long dash-dot line in Fig. 4(e)] associated with the proposed 1-D phase estimation algorithm. This implies

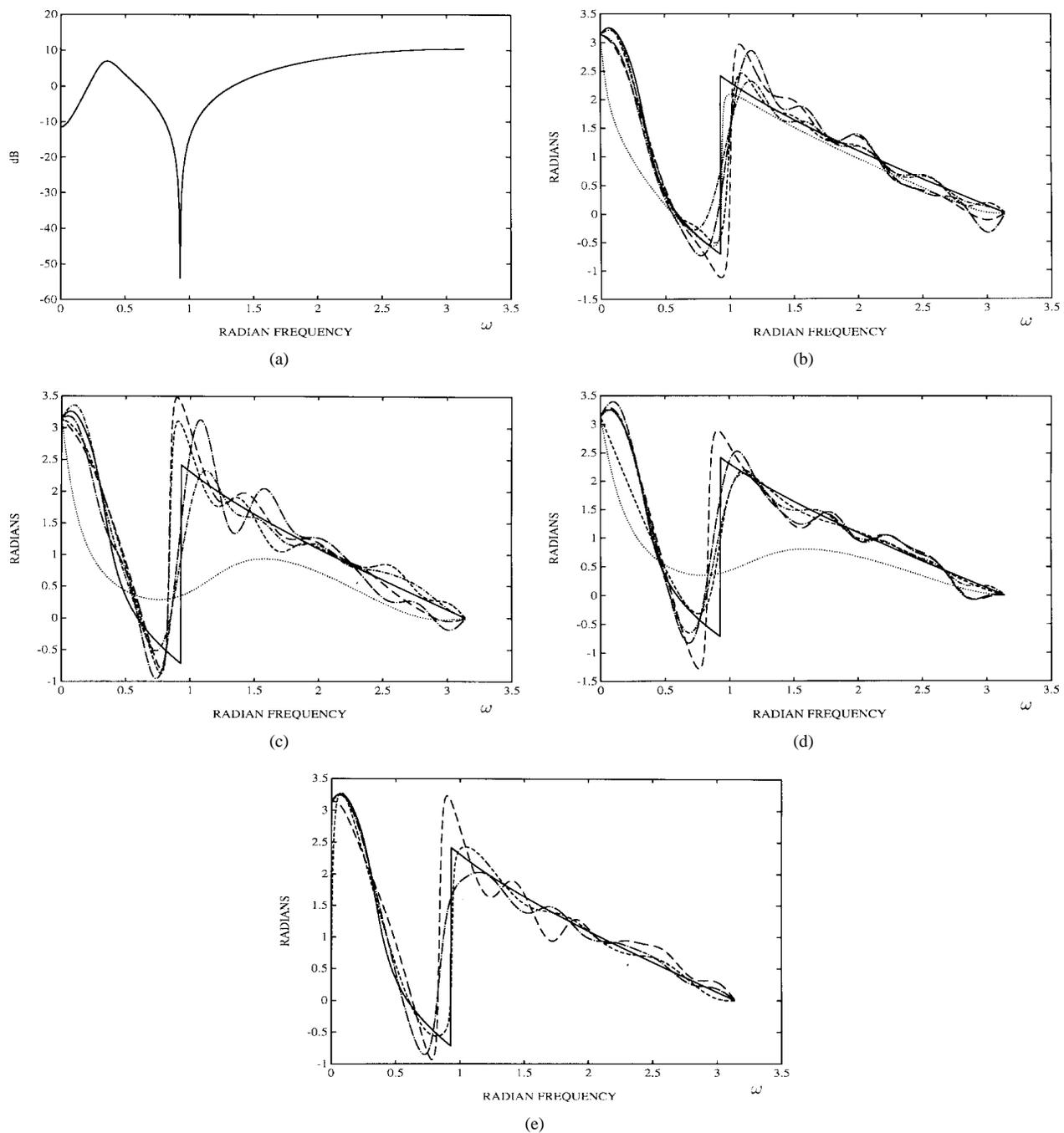


Fig. 4. 1-D simulation results associated with Example 2 obtained by *Algorithm 1 ARMA-Causal*, *Algorithm 1 Fourier* and an MP-AP decomposition-based method for a nonminimum phase system with a pair of zeros on the unit circle. (a) Amplitude response $|H(\omega)|$ of the system. (b) Phase estimates $\hat{\theta}(\omega)$ for $N = 1024$ and $\text{SNR} = \infty$ associated with *Algorithm 1 ARMA-Causal* (short dash line for $M = 3$ and long dash line for $M = 4$), *Algorithm 1 Fourier* (short dash-dot line for $M = 3$ and long dash-dot line for $M = 4$), and the MP-AP decomposition-based method (dotted line), together with the true system phase $\theta(\omega)$ (solid line). (c) Phase estimates $\hat{\theta}(\omega)$ corresponding to (b) for $N = 1024$ and $\text{SNR} = 5$ dB. (d) Phase estimates $\hat{\theta}(\omega)$ corresponding to (b) for $N = 4096$ and $\text{SNR} = 5$ dB. (e) Phase estimates $\hat{\theta}(\omega)$ for $M = 3$ (short dash line for $N = 1024$ and $\text{SNR} = \infty$, long dash line for $N = 1024$ and $\text{SNR} = 5$ dB, and long dash-dot line for $N = 4096$ and $\text{SNR} = 5$ dB) obtained by processing amplitude equalized data with *Algorithm 1 ARMA-Causal* together with the true system phase $\theta(\omega)$ (solid line).

that the proposed 1-D phase estimation algorithm is robust to the preprocessing of amplitude equalization. Moreover, these simulation results also indicate similar performance of the proposed 1-D phase estimation algorithm, regardless of whether data $x(n)$ or amplitude equalized data $x'(n)$ were processed for this example. Let us emphasize that the performance of the proposed 1-D phase estimation algorithm depends on SNR and

phase characteristics associated with the data to be processed. In practice, it may be unknown that higher SNR and simpler phase characteristics are associated with data $x(n)$ without preprocessing or associated with amplitude equalized data $x'(n)$. Therefore, whether preprocessing of the given data can improve the performance of the proposed phase estimation algorithms needs further study.

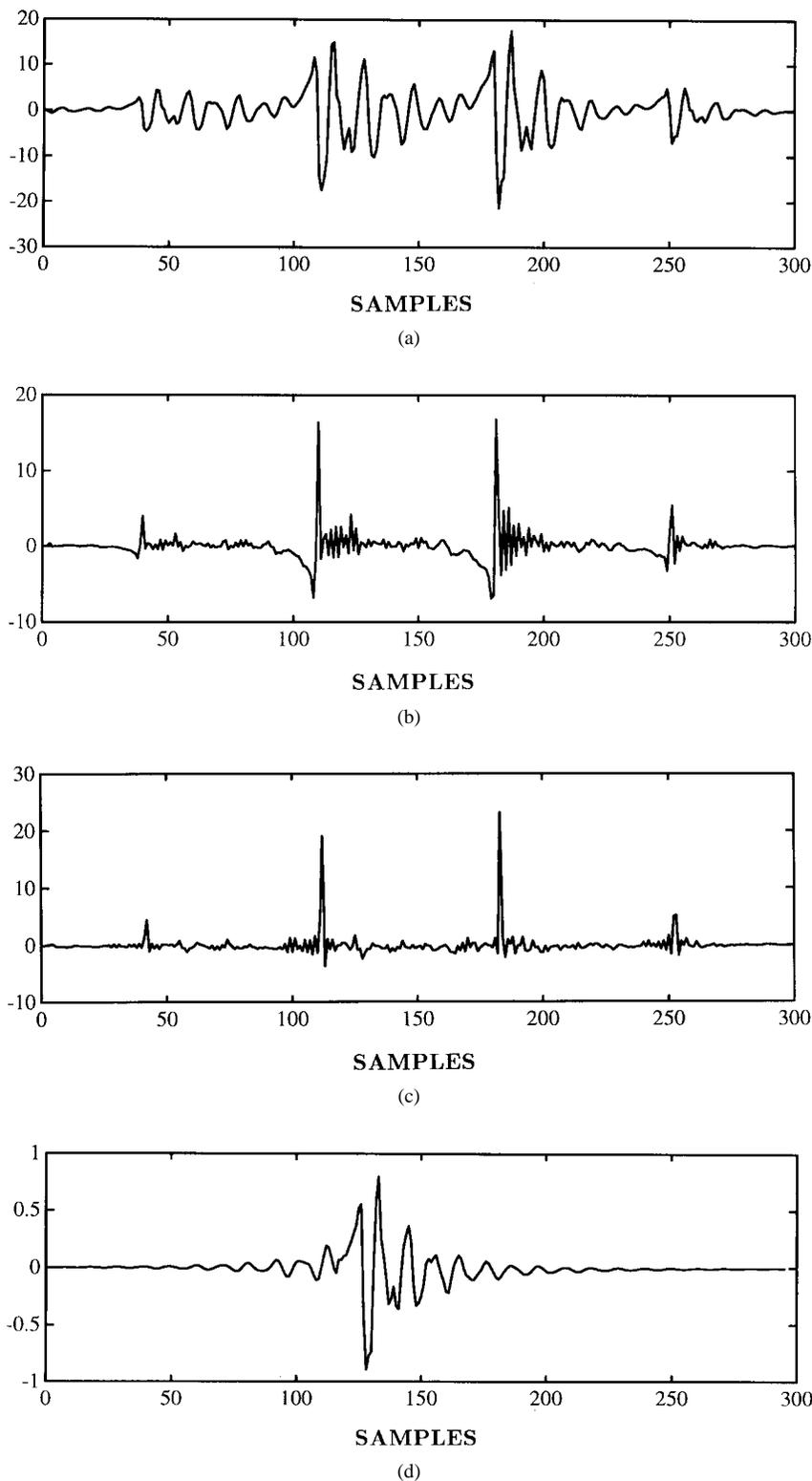


Fig. 5. Experimental results with real speech data for $M = 4$ associated with Example 3. (a) Windowed speech data (by Hamming window) of sound /a:/ uttered by a man (sampling rate equal to 8k Hz). (b) Predictive deconvolved speech signal $\tilde{u}(n)$ obtained by a correlation-based 12th-order LPE filter (obtained by Burg's algorithm). (c) Deconvolved signal $\hat{u}(n)$ obtained by the deconvolution filter $H_{INV}(\omega)$ given by (48). (d) Impulse response of the estimated vocal tract filter.

Example 3—Experimental Results with Real Speech Data: It is known [36] that a speech signal can be modeled as (9), where $h(n)$ is the vocal tract filter (possibly with nonminimum phase), and the driving input $u(n)$ is a pseudo-periodic (non-

Gaussian) impulse train for voiced speech and a white noise sequence for unvoiced speech. In this example, a set of real voiced speech data from a sound /a:/ uttered by a man was obtained through a 16-bit A/D converter with a sampling

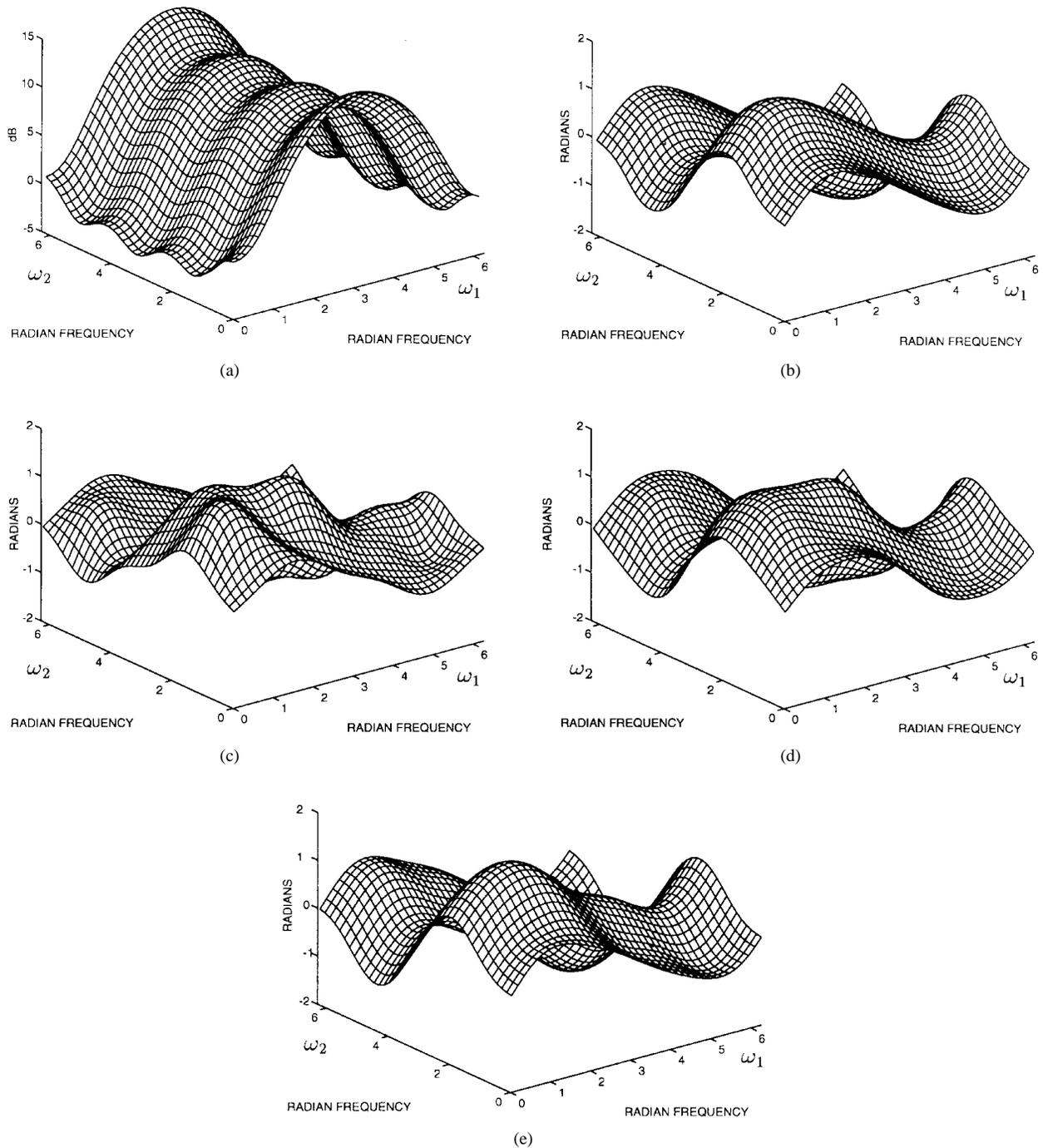


Fig. 6. 2-D simulation results for $M = 3$ associated with Example 4 obtained by *Algorithm 2* with $p_1 = p_2 = 3$ ($P_{\text{num}} = 24$) for a separable 2-D system with continuous phase response. (a) Amplitude response $|H(\omega_1, \omega_2)|$ of the system. (b) Phase response $\theta(\omega_1, \omega_2)$ of the system; phase estimates $\hat{\theta}(\omega_1, \omega_2)$ for (c) $N = 64 \times 64$ and SNR = 5 dB. (d) $N = 64 \times 64$ and SNR = ∞ . (e) $N = 128 \times 128$ and SNR = 5 dB.

frequency of 8 kHz. Then, the speech data were pre-multiplied by a Hamming window for further processing. Next, let us present some experimental results with the windowed data.

The experimental results are shown in Fig. 5. The windowed speech data shown in Fig. 5(a) were processed by a 12th-order minimum-phase linear prediction error (LPE) filter—denoted $V(\omega)$ —obtained by Burg's algorithm [31] to get the predictive deconvolved signal $\tilde{u}(n)$, which is shown in Fig. 5(b). It can be seen, from Fig. 5(b), that $\tilde{u}(n)$ approximates a pseudo-

periodic impulse train with some phase distortion because only amplitude response of the vocal tract filter can be equalized by the minimum-phase LPE filter $V(\omega)$, and the vocal tract filter is not minimum-phase for this case. This implies that phase estimation of the vocal tract filter is needed. With the windowed data shown in Fig. 5(a), the proposed *Algorithm 1 ARMA-Anticausal* with $p_{\text{max}} = 11$, $s = 1$ and cumulant order $M = 4$ was employed to obtain a phase estimate $\hat{\theta}(\omega)$ of the vocal tract filter. Then, the windowed data shown in Fig. 5(a) were processed by a deconvolution filter (corresponding to an

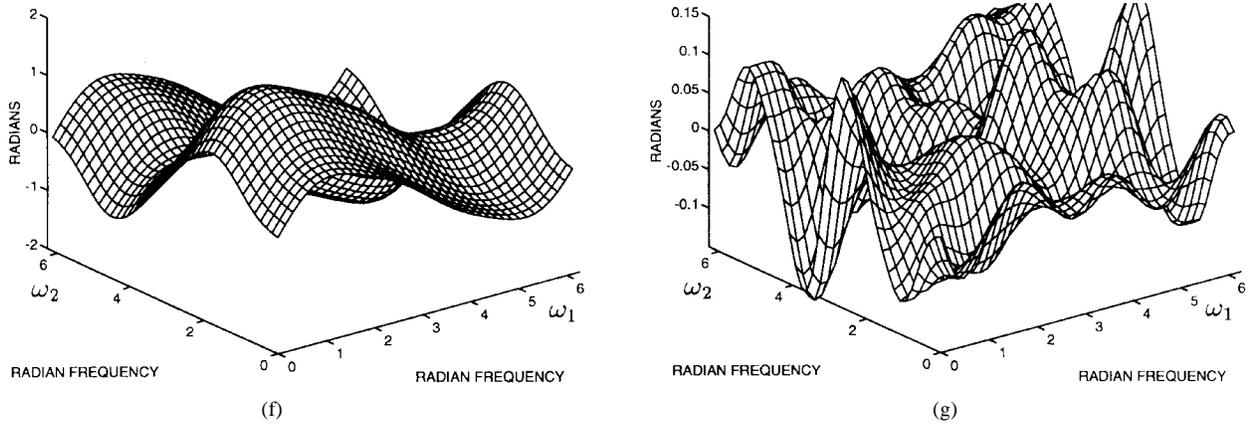


Fig. 6 (continued). 2-D simulation results for $M = 3$ associated with Example 4 obtained by *Algorithm 2* with $p_1 = p_2 = 3$ ($P_{\text{num}} = 24$) for a separable 2-D system with continuous phase response. (g) Phase estimation error $e(\omega_1, \omega_2)$ associated with (f).

inverse filter of the vocal tract filter)

$$H_{\text{INV}}(\omega) = |V(\omega)| \exp \{-j\hat{\theta}(\omega)\} \quad (48)$$

to get the estimate $\hat{u}(n)$, which is shown in Fig. 5(c). One can see, from Fig. 5(b) and (c), that $\hat{u}(n)$ approximates a pseudo-periodic impulse train much better than $\tilde{u}(n)$ with the phase distortion considerably removed by the deconvolution filter $H_{\text{INV}}(\omega)$ and that the pitch period can be easily found to be 70 samples (i.e., 8.75 ms) from the two most significant impulses in Fig. 5(c). These results justify that the proposed *Algorithm 1 ARMA-Anticausal* provides a good phase estimate of the vocal tract filter and that the LPE filter provides a good inverse amplitude estimate of the vocal tract filter; thus, $\hat{H}(\omega) = 1/H_{\text{INV}}(\omega)$ is a good estimate for the the vocal tract filter. The estimated vocal tract filter $\hat{h}(n)$ is shown in Fig. 5(d), which shows considerable resemblance to the windowed data [which is shown in Fig. 5(a)] of one pitch period and the length of the vocal tract filter approximately equal to two pitch periods. These experimental results also support the feasibility of collaboration of the proposed phase estimation method and amplitude estimation methods.

We also performed the same experiment with other speech data, which led to the same conclusions as drawn from Fig. 5, although these experimental results were not included here due to space limitations.

B. 2-D Phase Estimation

Let us present two simulation examples (Examples 4 and 5 below) to demonstrate the efficacy of *Algorithm 2* for the phase estimation of 2-D LTI systems. Example 4 is for a separable 2-D LTI system with continuous phase, and Example 5 is for the case that the system phase has discontinuities in the domain $\{(\omega_1, \omega_2), 0 \leq \omega_1 \leq 2\pi, 0 \leq \omega_2 \leq 2\pi\}$. In each of the two examples, the unknown linear phase term in the phase estimate $\hat{\theta}(\omega_1, \omega_2)$ was artificially removed for ease of comparison with the true system phase $\theta(\omega_1, \omega_2)$. Next, let us turn to Example 4.

Example 4—Separable 2-D System with Continuous Phase: A 2-D MA system $h(m, n)$ with a separable transfer function

given by

$$\begin{aligned} H(z_1, z_2) &= 1 - 0.8z_1^{-1} + 0.2z_1^{-2} \\ &\quad + 1.8z_2^{-1} - 1.44z_1^{-1}z_2^{-1} + 0.36z_1^{-2}z_2^{-1} \\ &\quad - 0.5z_2^{-2} + 0.4z_1^{-1}z_2^{-2} - 0.1z_1^{-2}z_2^{-2} \\ &\quad + 0.5z_2^{-3} - 0.4z_1^{-1}z_2^{-3} + 0.1z_1^{-2}z_2^{-3} \\ &= (1 - 0.8z_1^{-1} + 0.2z_1^{-2}) \\ &\quad \cdot (1 + 1.8z_2^{-1} - 0.5z_2^{-2} + 0.5z_2^{-3}) \end{aligned} \quad (49)$$

was used in the simulation. Magnitude response $|H(\omega_1, \omega_2)|$ and phase response $\theta(\omega_1, \omega_2)$ are shown in Fig. 6(a) and (b), respectively, where

$$\theta(\omega_1, \omega_2) = \arg \{H(\omega_1, \omega_2)\} + \alpha_1\omega_1 + \alpha_2\omega_2 \quad (50)$$

was plotted with $\alpha_1 = 0$ and $\alpha_2 = 1$ to eliminate linear phase factors. Again, a zero-mean exponentially distributed i.i.d. random field $u(m, n)$ with variance $\sigma_u^2 = 1$ and skewness $\gamma_3 = 2$ was convolved with $h(m, n)$ to obtain the noise-free synthetic data ($\text{SNR} = \infty$) to which white Gaussian noise was added to form the synthetic data for $\text{SNR} = 5$ dB. As mentioned in **R6**, when it is known *a priori* that the system is separable, the 2-D Fourier series-based phase model given by (8) reduces to the sum of two 1-D Fourier series-based phase models as given by (5). However, in this example, *Algorithm 2* with $M = 3$ and $p_1 = p_2 = 3$ was employed to estimate the system phase, assuming that the separability of the system was not known *a priori*. Thus, the total number of coefficients used in the 2-D Fourier series-based allpass model was $P_{\text{num}} = 24$ for this case.

The obtained continuous phase estimates $\hat{\theta}(\omega_1, \omega_2)$ for $N = 64 \times 64$ are shown in Fig. 6(c) and 6(d) for $\text{SNR} = 5$ dB and ∞ , respectively. The results for $N = 128 \times 128$ corresponding to those shown in Fig. 6(c) and (d) are shown in Fig. 6(e) and (f) ($\text{SNR} = 5$ dB and ∞), respectively. One can see, from Fig. 6(b)–(f), that $\hat{\theta}(\omega_1, \omega_2)$ approximates the true continuous system phase $\theta(\omega_1, \omega_2)$ better for larger N and SNR . Moreover, Fig. 6(g) shows the phase estimation error $e(\omega_1, \omega_2)$ associated with the phase estimate $\hat{\theta}(\omega_1, \omega_2)$ shown in Fig. 6(f) ($N = 128 \times 128$ and $\text{SNR} = \infty$). From

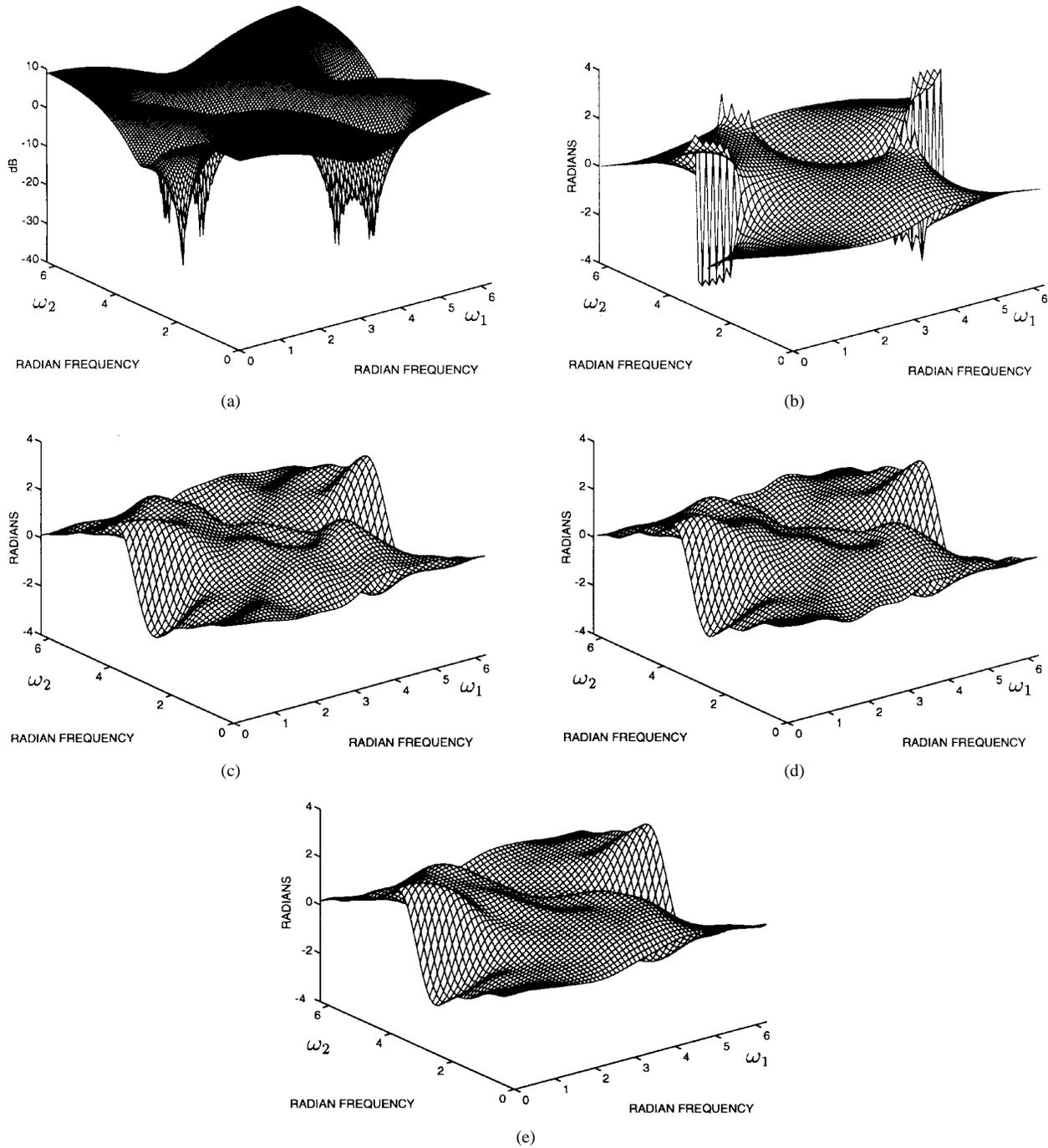


Fig. 7. 2-D simulation results for $M = 3$ associated with Example 5 obtained by *Algorithm 2* with $p_1 = p_2 = 5$ ($P_{\text{num}} = 60$) for a nonseparable 2-D system with discontinuities in phase response. (a) Amplitude response $|H(\omega_1, \omega_2)|$ of the system. (b) Phase response $\theta(\omega_1, \omega_2)$ of the system; phase estimates $\hat{\theta}(\omega_1, \omega_2)$ for (c) $N = 128 \times 128$ and SNR = 5 dB. (d) $N = 128 \times 128$ and SNR = ∞ . (e) $N = 256 \times 256$ and SNR = 5 dB.

Fig. 6(g), one can see that $\hat{\theta}(\omega_1, \omega_2)$ approximates $\theta(\omega_1, \omega_2)$ [see Fig. 6(b)] well and that $|e(\omega_1, \omega_2)|$ is smaller for those (ω_1, ω_2) where $|H(\omega_1, \omega_2)|$ [see Fig. 6(a)] is larger. These simulation results are consistent with **P1** as well as **P4** and demonstrate the phase retrieval capability of *Algorithm 2*.

Example 5—2-D System Phase with Discontinuities: A nonseparable 3×3 MA system with transfer function given by

$$\begin{aligned}
 H(z_1, z_2) = & -0.83z_1z_2 + 0.44z_2 + 0.3z_1^{-1}z_2 \\
 & + 0.3z_1 + 1 - 0.2z_1^{-1} \\
 & + 0.62z_1z_2^{-1} + 0.5z_2^{-1} + 0.7z_1^{-1}z_2^{-1} \quad (51)
 \end{aligned}$$

was used whose magnitude $|H(\omega_1, \omega_2)|$ and phase $\theta(\omega_1, \omega_2)$ responses are shown in Fig. 7(a) and (b), respectively. Spectral nulls in $|H(\omega_1, \omega_2)|$ and discontinuities (jumps of 2π or π) in $\theta(\omega_1, \omega_2)$ can be seen from these Fig. 7(a) and (b). Synthetic fields $x(m, n)$ were generated for $N = 128 \times 128$, $N = 256 \times 256$, SNR = 5 dB and SNR = ∞ , following the same procedure as Example 4. Then, each synthetic 2-D field $x(m, n)$ was processed by *Algorithm 2* with $M = 3$ and $p_1 = p_2 = 5$ (and thus $P_{\text{num}} = 60$).

The obtained continuous system phase estimates $\hat{\theta}(\omega_1, \omega_2)$ for $N = 128 \times 128$ are shown in Fig. 7(c) and (d) for SNR

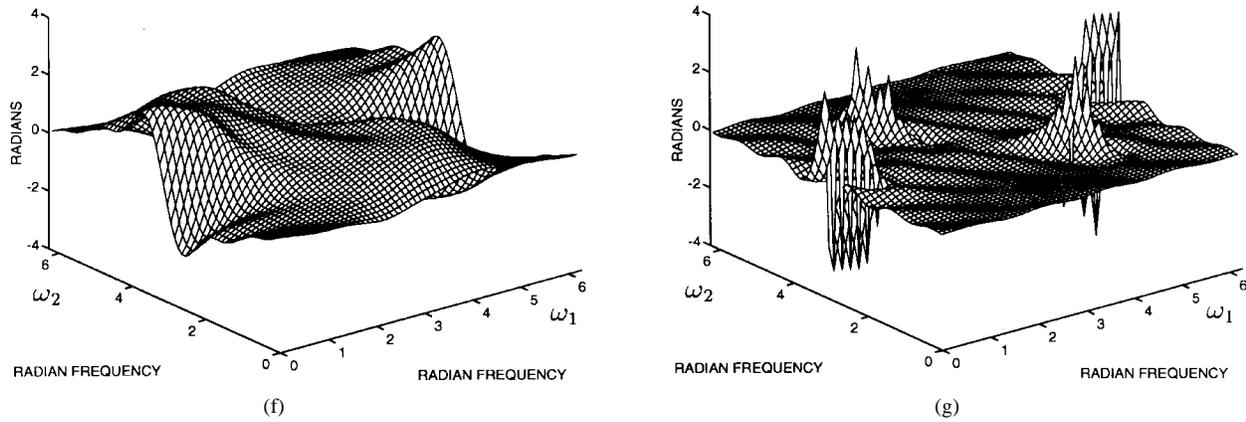


Fig. 7 (continued). 2-D simulation results for $M = 3$ associated with Example 5 obtained by *Algorithm 2* with $p_1 = p_2 = 5$ ($P_{\text{num}} = 60$) for a nonseparable 2-D system with discontinuities in phase response. (f) $N = 256 \times 256$ and $\text{SNR} = \infty$. (g) Phase estimation error $e(\omega_1, \omega_2)$ associated with (f).

$= 5$ dB and ∞ , respectively. The results for $N = 256 \times 256$ corresponding to those shown in Fig. 7(c) and (d) are shown in Fig. 7(e) and (f) ($\text{SNR} = 5$ dB and ∞), respectively. Again, one can see from Fig. 7(b)–(f) that $\hat{\theta}(\omega_1, \omega_2)$ approximates the true system phase $\theta(\omega_1, \omega_2)$ better for larger N and SNR, although $\theta(\omega_1, \omega_2)$ has discontinuities. Moreover, Fig. 7(g) shows the phase estimation error $e(\omega_1, \omega_2)$ associated with the phase estimate $\hat{\theta}(\omega_1, \omega_2)$ shown in Fig. 7(f) ($N = 256 \times 256$ and $\text{SNR} = \infty$). As predicted [see **P4** and **P5**], $|e(\omega_1, \omega_2)|$ is small for all (ω_1, ω_2) , where $\theta(\omega_1, \omega_2)$ is continuous, but large phase estimation error can happen in the vicinity of those (ω_1, ω_2) , where $\theta(\omega_1, \omega_2)$ is not continuous. In order to give a further insight of the estimated continuous phase $\hat{\theta}(\omega_1, \omega_2)$, Fig. 8 shows four slices (dashed lines) of the phase estimate $\hat{\theta}(\omega_1, \omega_2)$ shown in Fig. 7(f) for $\omega_1 = k\pi/7, k = 0, 2, 4, 6$ together with the associated four slices of the true $\theta(\omega_1, \omega_2)$ (solid lines). The slice associated with $\omega_1 = 0$ shown in Fig. 8(a) indicates that the true $\theta(\omega_1, \omega_2)$ (solid line) has a discontinuity of 2π , and the estimated continuous phase $\hat{\theta}(\omega_1, \omega_2)$ also approximates $\theta(\omega_1, \omega_2)$ well, except in the vicinity of the discontinuity. These results are consistent with the results presented in Example 2 for a 1-D system with discontinuous phase. The slice associated with $\omega_1 = 4\pi/7$ shown in Fig. 8(c) indicates that the phase-estimation error $|e(\omega_1, \omega_2)|$ can also be larger for those frequencies (ω_2) where the absolute value of group delay $(-\partial\theta(\omega_1, \omega_2)/\partial\omega_2)$ is larger, i.e., the system phase has a steeper variation. For the other slices shown in Fig. 8(b) and (d), $\hat{\theta}(\omega_1, \omega_2)$ approximates $\theta(\omega_1, \omega_2)$ well since the latter is continuous with small group delay.

The simulation results presented in this section support that the proposed 1-D and 2-D cumulant-based phase-estimation algorithms are effective for both 1-D and 2-D LTI systems.

VI. DISCUSSION AND CONCLUSIONS

Based on the ARMA allpass model, the proposed Fourier series-based allpass model, and Theorem 1, we have presented a parametric cumulant-based phase-estimation method implemented by *Algorithms 1* and *2* for estimating the phase of 1-D and 2-D nonminimum phase LTI systems with only non-Gaussian measurements. The system phase is estimated from

an optimum allpass filter that is obtained by maximizing a single absolute cumulant of the allpass filter output. *Algorithm 1 ARMA-Causal (Anticausal)* and *Algorithm 1 Fourier* are used for 1-D LTI systems and *Algorithm 2* (associated with the Fourier series-based allpass model) for 2-D LTI systems. These algorithms are iterative optimization algorithms with a parallel structure suitable for efficient implementation of both software and hardware. Moreover, optimum allpass filters obtained by the proposed phase estimation algorithms can be thought of as optimum phase equalizers to the unknown system of interest. In channel equalization (in communications), the unknown channel (or system) can be a phase distortion channel [25, p. 441], [37], although it frequently is a nonminimum phase LTI channel. The performance of the proposed phase-estimation method depends on SNR and channel phase characteristics associated with the data to be processed. Whether preprocessing of the given data such as amplitude equalization can improve the performance of the proposed phase estimation method needs further study (as mentioned in Example 2). Let us emphasize that the proposed 1-D and 2-D phase-estimation algorithms are not only applicable when the phase of the unknown LTI system is of interest but also applicable in collaboration with amplitude estimators for the identification and equalization of the system (as mentioned in Example 3).

A performance analysis for the proposed phase-estimation algorithms was also presented, and the analysis leads to six properties of the proposed phase-estimation algorithms. Then, some simulation results for different cumulant order M , data length N , and SNR, as well as some experimental results with real speech data, were offered to support that the proposed phase estimation algorithms can be used to estimate the phase of 1-D and 2-D LTI systems. The simulation results were also consistent with the performance analysis. However, we would like to make the following remarks regarding the presented simulation and experiment:

- R9)** The proposed 1-D and 2-D phase-estimation algorithms may still converge to a local optimum solution, although the use of different choices for the parameter s in **S1)** of *Algorithm 1* can reduce this possibility for the 1-D case. The values of s used in the presented simulation and experiment are good choices only for these simulation and experimental

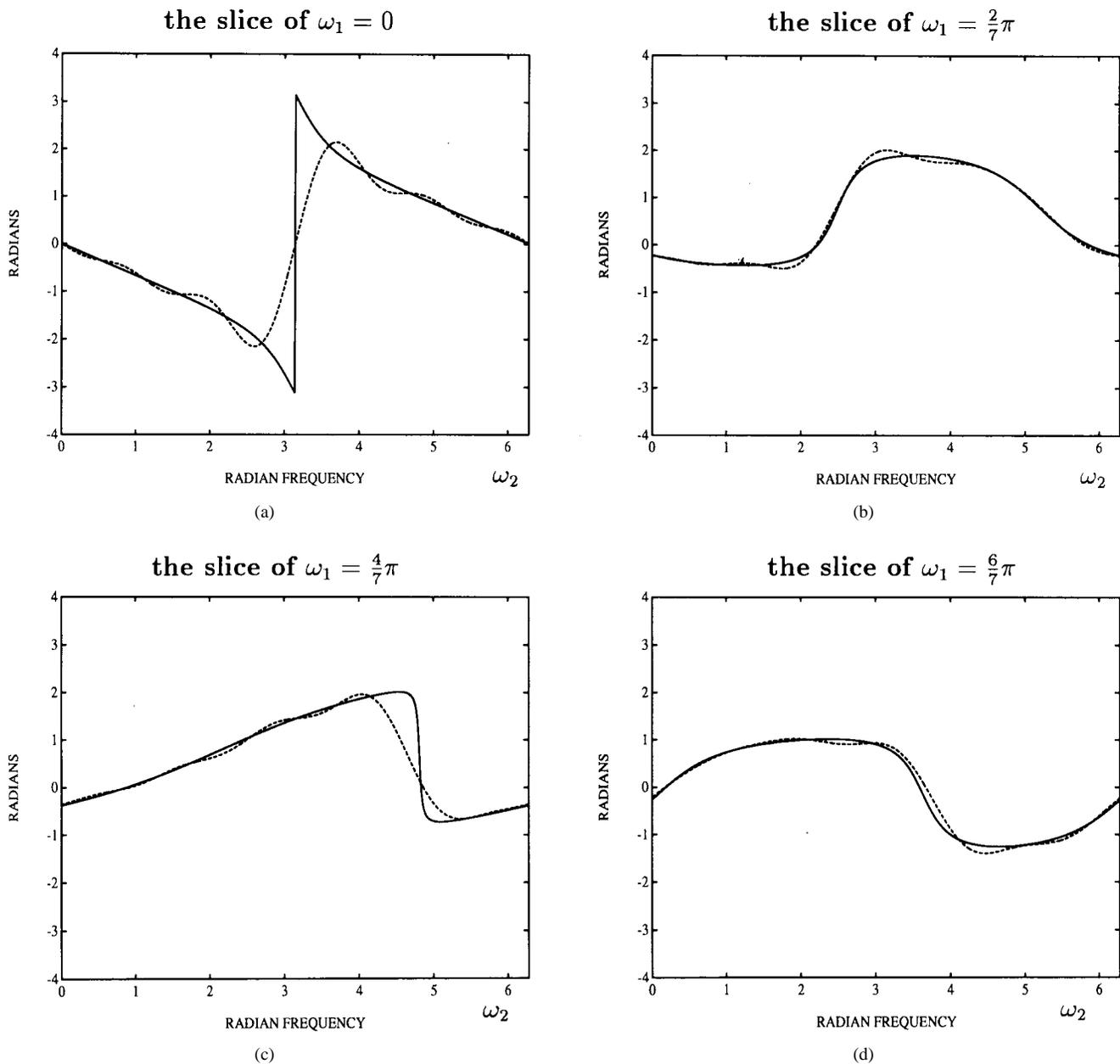


Fig. 8. Four slices of the phase estimate $\hat{\theta}(\omega_1, \omega_2)$ shown in Fig. 7(f) along (a) $\omega_1 = 0$, (b) $\omega_1 = 2\pi/7$, (c) $\omega_1 = 4\pi/7$, and (d) $\omega_1 = 6\pi/7$.

results. In other words, a good choice for s can only be obtained by comparing the resultant objective functions $J(\hat{a}_p)$ through some choices for s .

R10) *Algorithm 1 Fourier* is computationally faster than *Algorithm 1 ARMA-Causal (Anticausal)* with similar performance as discussed in **R4**, but for the latter, there are not any rules available yet for the choice of the ARMA allpass model to be causal stable or anticausal stable, which can always lead to best phase-estimation accuracy with least computational expense. This may need further study.

The proposed phase-estimation algorithms do not belong to any of the three categories of cumulant-based phase estimators mentioned in the introduction because they are based on philosophies that are different from those associated with

phase estimators of the three categories. Again, as mentioned in the introduction, there are not many 2-D cumulant-based phase-estimation algorithms [22]–[24] reported in the literature; meanwhile, their performance and efficiency are limited by the phase unwrapping problem. Moreover, MP-AP decomposition-based methods have yet to be extended to the 2-D case. Thus, *Algorithm 2* can be a quite efficient one. Extension of the proposed phase-estimation algorithms to the k -dimensional (k -D) ($k > 2$) case is also feasible by the same theorem (Theorem 1) with ω and \mathbf{n} defined as $1 \times k$ row vectors. Applications of the proposed cumulant phase-estimation algorithms such as in system identification, deconvolution, equalization, time delay estimation (estimation of the time interval that a single source signal arrives at two spatially separate sensors), signal detection, and harmonic

retrieval are under study, and the results will be reported in the future.

APPENDIX A PROOF OF (30)

Because $y(n)$ is the output of the allpass filter $H_p(\omega)$ with the input $x(n)$, $\partial y(n)/\partial a_m, m = 1, 2, \dots, p$ can be expressed as

$$\frac{\partial y(n)}{\partial a_m} = \frac{\partial}{\partial a_m} \left\{ \frac{1}{2\pi} \int_{-\pi}^{\pi} H_p(\omega) \cdot X(\omega) \cdot e^{j\omega n} d\omega \right\}. \quad (\text{A.1})$$

Substituting (4) into (A.1), we obtain

$$\begin{aligned} \frac{\partial y(n)}{\partial a_m} &= \frac{\partial}{\partial a_m} \left\{ \frac{1}{2\pi} \int_{-\pi}^{\pi} \exp \left\{ j \sum_{k=1}^p a_k \cdot \sin(k\omega) \right\} \right. \\ &\quad \left. \cdot X(\omega) \cdot e^{j\omega n} d\omega \right\} \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{\partial}{\partial a_m} \left\{ \exp \left\{ j \sum_{k=1}^p a_k \cdot \sin(k\omega) \right\} \right\} \\ &\quad \cdot X(\omega) \cdot e^{j\omega n} d\omega \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} j \cdot \sin(m\omega) \cdot \exp \left\{ j \sum_{k=1}^p a_k \cdot \sin(k\omega) \right\} \\ &\quad \cdot X(\omega) \cdot e^{j\omega n} d\omega \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} j \cdot \sin(m\omega) \cdot H_p(\omega) \cdot X(\omega) \cdot e^{j\omega n} d\omega \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} j \cdot \sin(m\omega) \cdot Y(\omega) \cdot e^{j\omega n} d\omega \end{aligned} \quad (\text{A.2})$$

which is nothing but the inverse Fourier transform of $Y(\omega) \cdot j \cdot \sin(m\omega)$. Therefore (A.2) can also be expressed as the convolution of $y(n)$ and the inverse Fourier transform of

$$j \cdot \sin(m\omega) = \frac{1}{2} \{ e^{j\omega m} - e^{-j\omega m} \}$$

as follows:

$$\begin{aligned} \frac{\partial y(n)}{\partial a_m} &= \left\{ \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{1}{2} [e^{j\omega m} - e^{-j\omega m}] \cdot e^{j\omega n} d\omega \right\} * y(n) \\ &= \frac{1}{2} \{ \delta(n+m) - \delta(n-m) \} * y(n) \\ &= \frac{1}{2} \{ y(n+m) - y(n-m) \}. \end{aligned} \quad (\text{A.3})$$

Q.E.D.

APPENDIX B

PROOF OF INEQUALITY GIVEN BY (45)

It can be easily inferred from (40) that

$$\begin{aligned} \mathcal{E}^2(\omega_1, \dots, \omega_{M-1}) &= [e(\omega_1) + \dots + e(\omega_{M-1}) - e(\omega_1 + \dots + \omega_{M-1})]^2 \\ &\leq M[e^2(\omega_1) + \dots + e^2(\omega_{M-1}) \\ &\quad + e^2(\omega_1 + \dots + \omega_{M-1})] \end{aligned} \quad (\text{B.1})$$

by the Schwartz inequality. Next, one can easily show, by (39), that

$$\begin{aligned} &\int_{-\pi}^{\pi} \dots \int_{-\pi}^{\pi} |S_{M,y}(\omega_1, \dots, \omega_{M-1})| e^2(\omega_1 + \dots + \omega_{M-1}) \\ &\quad \cdot d\omega_1 \dots d\omega_{M-1} \\ &= \int_{-\pi}^{\pi} \dots \int_{-\pi}^{\pi} |S_{M,y}(\omega_1, \dots, \omega_{M-1})| e^2(\omega_i) \\ &\quad \cdot d\omega_1 \dots d\omega_{M-1} \\ &= \int_{-\pi}^{\pi} \dots \int_{-\pi}^{\pi} |S_{M,y}(\omega_1, \dots, \omega_{M-1})| e^2(\omega_j) \\ &\quad \cdot d\omega_1 \dots d\omega_{M-1}, \quad \text{for all } j \neq i \end{aligned} \quad (\text{B.2})$$

where the first equality (the first two lines) of (B.2) is proved by changing variables of

$$\eta_1 = \omega_1 + \dots + \omega_{M-1}, \quad \eta_2 = -\omega_2, \dots, \eta_{M-1} = -\omega_{M-1}$$

and by the fact that $|H(\omega)| = |H(-\omega)|$ (which is an even function). Then we can obtain, from (39), (43), (B.1), and (B.2), that

$$\begin{aligned} \mathcal{F} &\leq |\gamma_M| M^2 \int_{-\pi}^{\pi} \dots \int_{-\pi}^{\pi} |H(\omega_1) \dots H(\omega_{M-1})| \\ &\quad \cdot H(\omega_1 + \dots + \omega_{M-1}) |e^2(\omega_1) d\omega_1 \dots d\omega_{M-1} \\ &= |\gamma_M| M^2 \int_{-\pi}^{\pi} \dots \int_{-\pi}^{\pi} \\ &\quad \cdot \left\{ \int_{-\pi}^{\pi} |H(\omega_2) H(\omega_1 + \dots + \omega_{M-1})| d\omega_2 \right\} \\ &\quad \cdot |H(\omega_1) H(\omega_3) \dots H(\omega_{M-1})| e^2(\omega_1) \\ &\quad \cdot d\omega_1 d\omega_3 \dots d\omega_{M-1} \\ &\leq |\gamma_M| M^2 \left\{ \int_{-\pi}^{\pi} |H(\omega)|^2 d\omega \right\} \\ &\quad \cdot \int_{-\pi}^{\pi} \dots \int_{-\pi}^{\pi} |H(\omega_1) H(\omega_3) \dots H(\omega_{M-1})| \\ &\quad \cdot e^2(\omega_1) d\omega_1 d\omega_3 \dots d\omega_{M-1} \end{aligned} \quad (\text{B.3})$$

where we have applied the Schwartz inequality to the integral inside the braces in the second line of (B.3). It is trivial to see that the right-hand side of (B.3) is equal to the right-hand side of (45).

Q.E.D.

ACKNOWLEDGMENT

The authors would like to thank the associate editor and the anonymous reviewers for their valuable comments and suggestions. Finally, the authors would like to thank C.-C. Feng, C.-H. Chen, C.-F. Rau, J.-K. Peng, and G.-C. Wang for their help during the revision of the paper.

REFERENCES

- [1] J. M. Mendel, "Tutorial on higher order statistics (spectra) in signal processing and system theory: Theoretical results and some applications," *Proc. IEEE*, vol. 79, pp. 278–305, Mar. 1991.
- [2] C. L. Nikias and M. Raghuveer, "Bispectrum estimation: A digital signal framework," *Proc. IEEE*, vol. 75, pp. 869–891, July 1987.
- [3] C. L. Nikias and A. P. Petropulu, *Higher Order Spectra Analysis*. Englewood Cliffs, NJ: Prentice-Hall, 1993.
- [4] C.-Y. Chi and M.-C. Wu, "Inverse filter criteria for blind deconvolution and equalization using two cumulants," *Signal Processing*, vol. 43, pp. 55–63, Apr. 1995.

- [5] O. Shalvi and E. Weinstein, "New criteria for blind deconvolution of nonminimum phase systems (channels)," *IEEE Trans. Inform. Theory*, vol. 36, pp. 312–321, Mar. 1990.
- [6] J. K. Tugnait, "Identification of nonminimum phase linear stochastic systems," *Automatica*, vol. 22, no. 4, pp. 457–464, 1986.
- [7] G. B. Giannakis and J. M. Mendel, "Stochastic realization of nonminimum phase systems," in *Proc. 1986 Amer. Contr. Conf.*, Seattle, WA, vol. 2, pp. 1254–1259.
- [8] ———, "Approximate realization and model reduction of nonminimum phase stochastic systems," in *Proc. 25th IEEE Conf. Decision Contr.*, Athens, Greece, 1986, pp. 1079–1084.
- [9] C.-Y. Chi and J.-Y. Kung, "A phase determination method for nonminimum phase ARMA systems by a single cumulant sample," *IEEE Trans. Signal Processing*, vol. 41, pp. 981–985, Feb. 1993.
- [10] ———, "A fast phase determination method by a single cumulant sample," in *Proc. Int. Signal Processing Workshop Higher Order Statist.*, Chamrousse, France, July 10–12, 1991, pp. 83–86.
- [11] ———, "A new identification algorithm for allpass systems by higher order statistics," in *Proc. Sixth Eur. Signal Processing Conf. (EUSIPCO-92)*, Brussels, Belgium, vol. 2, Aug. 24–27, 1992, pp. 779–782.
- [12] ———, "A new identification algorithm for allpass systems by higher order statistics," *Signal Processing*, vol. 41, no. 2, pp. 239–256, Jan. 1995.
- [13] L. M. Brillinger, "The identification of a particular nonlinear time series system," *Biometrika*, vol. 64, pp. 509–515, 1977.
- [14] K. S. Lii and M. Rosenblatt, "Deconvolution and estimation of transfer function phase and coefficients for non-Gaussian linear processes," *Ann. Statist.*, vol. 10, no. 4, pp. 1195–1208, 1982.
- [15] T. Matsuoka and T. J. Ulrych, "Phase estimation using the bispectrum," *Proc. IEEE*, vol. 72, pp. 1403–1411, Oct. 1984.
- [16] F. Cheng and A. N. Venetsanopoulos, "Determination of the exact phase of the bispectrum," *IEEE Trans. Circuits Syst. II*, vol. 39, pp. 155–160, Mar. 1992.
- [17] C. A. Haniff, "Least-squares Fourier phase estimation from the module 2π bispectrum phase," *J. Opt. Soc. Amer. A*, vol. 8, pp. 134–140, Jan. 1991.
- [18] M. Rangoussi and G. B. Giannakis, "FIR modeling using log-bispectra: Weighted least-squares algorithm and performance analysis," *IEEE Trans. Circuits Syst.*, vol. 38, pp. 281–296, Mar. 1991.
- [19] J. C. Marron, P. P. Sanchez, and R. C. Sullivan, "Unwrapping algorithm for least-squares phase recovery from the module 2π bispectrum phase," *J. Opt. Soc. Amer. A*, vol. 7, pp. 14–20, Jan. 1990.
- [20] R. Pan and C. L. Nikias, "Phase reconstruction in the trispectrum domain," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-35, pp. 895–897, June 1987.
- [21] Y. Li and Z. Ding, "A new nonparametric method for linear system phase recovery from bispectrum," *IEEE Trans. Circuits Syst. II*, vol. 41, pp. 415–419, June 1994.
- [22] S. A. Dianat and M. R. Raghuveer, "Fast algorithm for phase and magnitude reconstruction from bispectra," *Opt. Eng.*, vol. 29, no. 5, pp. 504–512, May 1990.
- [23] M. G. Kang, K. T. Lay, and A. K. Katsaggelos, "Phase estimation using the bispectrum and its application to image restoration," *Opt. Eng.*, vol. 30, pp. 976–985, July 1991.
- [24] H. Takajo and T. Takahashi, "Least-squares phase recovery from the bispectrum phase: An algorithm for a two-dimensional object," *J. Opt. Soc. Amer. A*, vol. 8, pp. 1038–1047, July 1991.
- [25] A. V. Oppenheim and R. W. Schaffer, *Discrete-Time Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1989.
- [26] A. G. Deczky, "Synthesis of recursive digital filters using the minimum p error criterion," *IEEE Trans. Audio Electroacoust.*, vol. AU-20, pp. 257–263, Oct. 1972.
- [27] L. B. Jackson, *Signals, Systems and Transforms*. Reading, MA: Addison Wesley, 1991.
- [28] A. Swami, G. B. Giannakis, and J. M. Mendel, "Linear modeling of multidimensional non-Gaussian processes using cumulants," *Multidimensional Syst. Signal Processing*, vol. 1, pp. 11–37, 1990.
- [29] M. A. Tekalp and A. T. Erdem, "Higher order spectrum factorization in one and two dimensions with applications in signal modeling and nonminimum phase system ID," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 37, pp. 1537–1549, Oct. 1989.
- [30] J. M. Mendel, *Lessons in Estimation Theory for Signal Processing, Communications, and Control*. Englewood Cliffs, NJ: Prentice-Hall, 1995.
- [31] S. M. Kay, *Modern Spectral Estimation*. Englewood Cliffs, NJ: Prentice-Hall, 1988.
- [32] J. Cadzow, "High performance spectral estimation—A new ARMA method," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-28, pp. 524–529, Oct. 1980.
- [33] J. Durbin, "Efficient estimation of parameters in moving-average models," *Biometrika*, vol. 46, pp. 306–316, 1959.
- [34] J. P. Burg, "Maximum entropy spectral analysis," Ph.D. dissertation, Stanford Univ., Stanford, CA, May 1975.
- [35] J. Makhoul, "Linear prediction: A tutorial review," *Proc. IEEE*, vol. 63, pp. 561–580, Apr. 1975.
- [36] L. R. Rabiner and R. M. Schafer, *Digital Processing of Speech Signals*. Englewood Cliffs, NJ: Prentice-Hall, 1978.
- [37] A. P. Clark, *Equalizers for Digital Modems*. New York: Wiley, 1985.



Horng-Ming Chien was born in Taiwan, R.O.C., on January 14, 1971. He received the B.S. and M.S. degrees from the Department of Electrical Engineering, National Tsing Hua University, Hsinchu, Taiwan, in 1993 and 1995, respectively.

From July 1995 to June 1997, he was in military service as a reserve officer of the R.O.C. Air Force, working on computer simulation programs for air traffic control. His research interests include digital signal processing, statistical signal processing, and communications.



Huang-Lin Yang was born in Taiwan, R.O.C., on October 11, 1968. He received the B.S. degree in electronic engineering from Feng-Chia University, Taichung, Taiwan, in 1991 and the M.S. degree in electrical engineering from National Tsing Hua University, Hsinchu, Taiwan, in 1993.

From July 1993 to May 1995, he was in military service as a reserve officer. Since July 1995, he has been a design engineer with the Advanced Technology Center, Computer and Communication Research Laboratories, Industrial Technology Research Institute, Hsinchu, where he works on speech signal processing and its applications including speech coding, recognition, and synthesis. His research interests include statistical signal processing and digital signal processing and their applications.



Chong-Yung Chi (S'83–M'83–SM'89) was born in Taiwan, R.O.C., on August 7, 1952. He received the B.S. degree from the Tatung Institute of Technology, Taipei, Taiwan, R.O.C., in 1975, the M.S. degree from the National Taiwan University, in 1977, and the Ph.D. degree from the University of Southern California, Los Angeles, in 1983, all in electrical engineering.

From July 1983 to September 1988, he was with the Jet Propulsion Laboratory, Pasadena, CA, where he worked on the design of various spaceborne radar remote sensing systems including radar scatterometers, SAR's, altimeters, and rain mapping radars. From October 1988 to July 1989, he was a visiting specialist at the Department of Electrical Engineering, National Taiwan University. Since August 1989, he has been Professor with the Department of Electrical Engineering, National Tsing Hua University, Hsinchu, Taiwan. He has published more than 70 technical papers in radar remote sensing, system identification and estimation theory, deconvolution and channel equalization, digital filter design, spectral estimation, and higher order statistics (HOS)-based signal processing. His research interests include statistical signal processing and digital signal processing and their applications.

Dr. Chi is an active member of New York Academy of Sciences, an associative member of Society of Exploration Geophysicists, and an active member of the Chinese Institute of Electrical Engineering. He is also a technical committee member of 1997 IEEE Signal Processing Workshop on Higher Order Statistics.