Design of Wiener filters using a cumulant based MSE criterion

Chih-Chun Feng, Chong-Yung Chi*

Department of Electrical Engineering, National Tsing Hua University, Hsinchu, Taiwan 30043, ROC

Received 7 September 1995; revised 19 April 1996

Abstract

This paper proposes a cumulant (higher-order statistics) based mean-square-error (MSE) criterion for the design of Wiener filters when both the given wide-sense stationary random signal $x(n)$ and the desired signal $d(n)$ are non-Gaussian and contaminated by Gaussian noise sources. It is theoretically shown that the designed Wiener filter associated with the proposed criterion is identical to the conventional correlation (second-order statistics) based Wiener filter as if both $x(n)$ and $d(n)$ were noise-free measurements. As the latter, the former can also be obtained by solving a cumulant-based Wiener–Hopf equation associated with a (cumulant-based) orthogonality principle. Then a generalized cumulant projection theorem is proposed which includes the projection of cumulants to correlations associated with the proposed cumulant-based MSE criterion and that associated with Delopoulos and Giannakis' cumulant-based MSE criterion as special cases. Moreover, the proposed cumulant-based MSE criterion and Delopoulos and Giannakis' cumulant-based MSE criterion are equivalent for cumulant order $M = 3$. Some simulation results for system identification and time delay estimation are then provided to demonstrate the good performance of the proposed cumulant-based Wiener filter. Finally, we draw some conclusions.

Zusammenfassung

Wir schlagen für den Entwurf von Wiener-Filtern ein Kumulanten(Statistiken höherer Ordnung)-basiertes mittleres quadratisches Fehlerkriterium (mean-square-error, MSE) für den Fall vor, daß das gegebene, im weiten Sinne stationäre Zufallssignal $x(n)$ und das gewünschte Signal $d(n)$ beide nicht Gauß-verteilt sind und durch Gauß-Rauschquellen gestört sind. Es wird theoretisch gezeigt, daß das mit dem vorgeschlagenen Kriterium entworfene Wiener-Filter identisch ist mit dem konventionellen Korrelations(Statistiken zweiter Ordnung)-basierten Wiener-Filter, wenn sowohl $x(n)$ als auch $d(n)$ rauschfreie Messungen sind. Wie im letzten Fall, kann auch der vorhergehende Fall durch Lösen einer Kumulanten-basierten Wiener–Hopf-Gleichung, verbunden mit einem (Kumulanten-basierten) Orthogonalitätsprinzip, erzielt werden. Sodann wird ein verallgemeinertes Kumulanten-Projektionstheorem vorgeschlagen, welches die Projektion von Kumulanten auf Korrelationen, sowohl auf der Basis des vorgeschlagenen Kumulanten-basierten MSE-Kriteriums, als auch in Verbindung mit dem Kumulanten-basierten MSE-Kriterium von Delopoulos und Giannakis als Spezialfälle, einschließt. Darüberhinaus sind das vorgeschlagene Kumulanten-basierte MSE-Kriterium und das Kumulanten-basierte MSE-Kriterium von Delopoulos und Giannakis äquivalent für Kumulanten der Ordnung $M = 3$. Einige Simulationsergebnisse zur Systemidentifikation und Verzögerungszeit-Schätzung werden gegeben, um das gute Verhalten des vorgeschlagenen Kumulanten-basierten Wiener-Filter zu demonstrieren. Abschließend ziehen wir einige Schlußfolgerungen.

---

*Corresponding author. Tel: 886-3-5731156; fax: 886-3-5715971; e-mail: cychi@ee.nthu.edu.tw.

Résumé

Cet article propose un critère d'erreur aux moindres carrés basé sur un cumulant (statistiques d'ordres supérieurs) pour la calibration de filtres de Wiener quand, à la fois le signal aléatoire stationaire au sens large donné $x(n)$ et le signal désiré $d(n)$ sont non-Gaussien et corrompus par des sources de bruit Gaussien. Il est théoriquement démontré que le filtre de Wiener ainsi calibré, associé avec le critère proposé, est identique à la corrélation conventionnelle (statistiques de second-ordre) basé sur un filtre de Wiener comme si à la fois $x(n)$ et $d(n)$ étaient des mesures sans bruit. Ainsi que ce dernier, le premier peut également être obtenu en résolvant un cumulant basé sur l'équation de Wiener–Hopf, associé à un principe d'orthogonalité (basé sur un cumulant). Ainsi, un théorème généralisé de projection de cumulant est proposé qui inclu la projection de cumulants à des corrélations en association avec le critère MSE proposé, basé sur un cumulant, avec comme cas particulier le cumulant de Delopoulos and Giannakis. De plus, le critère MSE proposé basé sur un cumulant et le cumulant de Delopoulos and Giannakis sont équivalents pour un ordre de cumulant $M = 3$. Quelques résultats de simulation pour l'identification de système et l'estimation de retard sont ensuite proposés afin de démontrer les bonnes performances du filtre de Wiener proposé basé sur un cumulant. Finalement, nous tirons quelques conclusions.

Keywords: Wiener filter; Mean-square-error (MSE) criterion; Cumulant

1. Introduction

The well-known Wiener filter [8, 9, 16] has widely been used in various correlation-based statistical signal processing areas such as system identification, predictive deconvolution, channel equalization, noise cancellation and suppression, echo cancellation and time delay estimation. Assuming that $x(n)$ is the given wide-sense stationary signal and $d(n)$ is the desired signal, the conventional Wiener filter is based on the mean-square-error (MSE) criterion which leads to a correlation-based orthogonality principle, and its coefficients can be solved from the well-known Wiener–Hopf equation formed of autocorrelation function $r_{xx}(i)$ as well as cross correlation function $r_{dx}(i)$. However, both $r_{xx}(i)$ and $r_{dx}(i)$ include noise correlations when $x(n)$ is corrupted by additive noise. Therefore, the performance of the correlation-based Wiener filter is sensitive to additive noise no matter whether noise is Gaussian or not.

Recently, higher-order ($\geqslant 3$) statistics (HOS) [11, 12, 14, 15], known as cumulants, have been considered in various statistical signal processing areas where signal $x(n)$ is non-Gaussian and contaminated by Gaussian noise, partly because cumulants of $x(n)$ contain not only amplitude information but also phase information of $x(n)$ and partly because all higher-order cumulants of $x(n)$ are insensitive to Gaussian noise whose Mth-order cumulants are all equal to zero for $M \geqslant 3$. As a matter of fact, Mth-order cumulants of $x(n)$ are insensitive to non-Gaussian noise as long as Mth-order cumulants of noise are equal to zero.

Chi et al. [4, 5] proposed two cumulant-based MSE criteria for the design of linear prediction error (LPE) filters. It was shown in [4, 5] that the two cumulant-based MSE criteria are equivalent to the correlation-based MSE criterion as if $x(n)$ were a noise-free non-Gaussian signal. Furthermore, the coefficients of the designed LPE filters associated with the two cumulant-based MSE criteria can be obtained by solving a set of symmetric Toeplitz linear equations to which the computationally efficient Levinson–Durbin recursion [8, 9, 16] can be applied. In this paper, we further propose a cumulant-based MSE criterion for the design of Wiener filters described in Theorem 1 below which is a generalization of one of the two cumulant-based MSE criteria reported in [4, 5] for the design of LPE filters. Similar to the correlation-based Wiener filter, the proposed cumulant-based Wiener filter also leads to a cumulant-based orthogonality principle described in Theorem 2 below. Based on the cumulant-based orthogonality principle, the optimum cumulant-based Wiener filter can be obtained from the associated cumulant-based Wiener–Hopf equation and implemented by a lattice structure [16] following the well-known Levinson–Durbin recursion.

Delopoulos and Giannakis [6] proposed a projection operator which projects a third-order cumulant function to an autocorrelation function except for a scale factor. Based on the projection concept, Delopoulos and Giannakis [6, 7] proposed some cumulant-based MSE criteria for identification of linear systems. In this paper, we further extend the projection concept to a generalized projection concept described in Theorem 3 below which states that an Mth-order cumulant function can be projected to an mth-order cumulant function except for a scale factor where $2 \leqslant m < M$. It will be shown that both Delopoulos and Giannakis' cumulant-based MSE criterion [7] and the proposed cumulant-based MSE criterion are special cases of the generalized projection described in Theorem 3. Moreover, the latter is equivalent to the former [7] for cumulant order $M = 3$ and is computationally much more practical than the former for $M \geqslant 4$.

The new cumulant-based MSE criterion for the design of Wiener filters is presented in Section 2. Section 3 presents the generalized projection concept. Then some simulation results for system identification and time delay estimation are provided in Section 4 to support the proposed cumulant-based Wiener filter. Finally, we draw some conclusions and provide a discussion in Section 5.

2. A new cumulant-based MSE criterion for the design of Wiener filters

Assume that $x(n)$ and $d(n)$, $n = 0, 1, \ldots, N - 1$, are the given non-Gaussian noisy measurements generated from the following convolutional models (see the block diagram shown in Fig. 1), respectively:

$$x(n) = x_f(n) + w_1(n), \tag{1a}$$

$$x_f(n) = u(n) * g(n), \tag{1b}$$

and

$$d(n) = d_f(n) + w_2(n), \tag{2a}$$

$$d_f(n) = x_f(n) * h(n), \tag{2b}$$

where $x_f(n)$ and $d_f(n)$ are the noise-free signals associated with $x(n)$ and $d(n)$, respectively, $w_1(n)$ and $w_2(n)$ are measurement noise sources, $g(n)$ and $h(n)$ are linear time-invariant (LTI) systems (with possibly nonminimum phase), and $u(n)$ is the driving input to the system $g(n)$. Let us make the following statistical assumptions for $u(n)$, $w_1(n)$ and $w_2(n)$:

(A1) $u(n)$ is a real, zero-mean, stationary, independent identically distributed (i.i.d.), non-Gaussian driving input sequence with variance σ_u^2 and Mth-order ($M \geqslant 3$) cumulant γ_M.

(A2) $w_1(n)$ and $w_2(n)$ are zero-mean Gaussian noise sequences which can be white or colored with unknown statistics.

(A3) The driving input $u(n)$ is statistically independent of $w_1(n)$ and $w_2(n)$.

Assume that the Wiener filter is an FIR filter, denoted $v(n)$, with $v(n) \neq 0$ for $p_1 \leqslant n \leqslant p_2$, where p_1 and p_2 are integers. The conventional correlation-based Wiener filter is designed such that the MSE, denoted $E[e^2(n)]$, of the estimation error defined as

$$e(n) = d(n) - x(n) * v(n) = d(n) - \sum_{i = p_1}^{p_2} v(i)x(n - i) \tag{3}$$

is minimum, where $d(n)$ is the desired signal which need not satisfy the convolutional model given by (2). The coefficients of the conventional Wiener filter can be solved from the following well-known Wiener–Hopf equation [8, 9, 16]:

$$\sum_{j = p_1}^{p_2} r_{xx}(i - j)v(j) = r_{dx}(i), \quad i = p_1, p_1 + 1, \ldots, p_2, \tag{4}$$
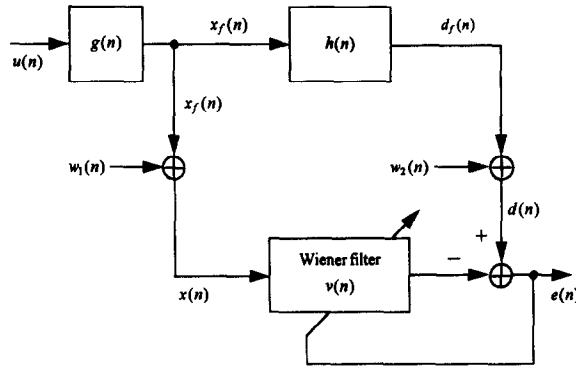
Fig. 1. Block diagram for Wiener filtering.

where $r_{xx}(i)$ denotes the autocorrelation function of $x(n)$ and $r_{dx}(i)$ denotes the cross correlation function of $d(n)$ and $x(n)$. However, both $r_{xx}(i)$ and $r_{dx}(i)$ include noise correlations due to the additive Gaussian noise sources $w_1(n)$ and $w_2(n)$.

For ease of later use, the error signal $e(n)$ defined by (3) can be further expressed as

$$e(n) = [d_f(n) + w_2(n)] - [x_f(n) + w_1(n)] * v(n) \quad \text{(see (1a) and (2a))}$$

$$= [d_f(n) - x_f(n) * v(n)] + [w_2(n) - w_1(n) * v(n)]$$

$$= e_f(n) + w(n), \tag{5}$$

where $w(n) = w_2(n) - w_1(n) * v(n)$ is also a Gaussian noise sequence since $w_1(n)$ and $w_2(n)$ are Gaussian by the assumption (A2) and

$$e_f(n) = d_f(n) - x_f(n) * v(n)$$

$$= u(n) * g(n) * h(n) - u(n) * g(n) * v(n) \quad \text{(see (1b) and (2b))}$$

$$= u(n) * f(n) \tag{6}$$

is the noise-free error signal in which

$$f(n) = g(n) * [h(n) - v(n)]. \tag{7}$$

Moreover, let $\text{Cum}^{(M)}(x_1(n + k_1), x_2(n + k_2), \ldots, x_M(n + k_M))$ denote the Mth-order joint cumulant function of real stationary random processes $\{x_i(n)\}$, $i = 1, 2, \ldots, M$. It can be shown [11, 14] that if $x_i(n) = u(n) * f_i(n)$, $i = 1, 2, \ldots, M$, where $u(n)$ is the driving input under the assumption (A1) and $f_i(n)$, $i = 1, 2, \ldots, M$, are arbitrary LTI systems, then

$$\text{Cum}^{(M)}(x_1(n + k_1), x_2(n + k_2), \ldots, x_M(n + k_M)) = \gamma_M \sum_{n = -\infty}^{\infty} f_1(n + k_1) f_2(n + k_2) \cdots f_M(n + k_M). \tag{8}$$

Note that $\text{Cum}^{(2)}(x_1(n + k_1), x_2(n + k_2)) = E[x_1(n + k_1) x_2(n + k_2)]$ and $\gamma_2 = \sigma_u^2$ for cumulant order $M = 2$.

The new cumulant-based MSE criterion for the design of Wiener filters is described in the following theorem.

Theorem 1. *Assume that $x(n)$ and $d(n)$ are the noisy signals given by (1) and (2), respectively, under the assumptions (A1)–(A3). Let $v(n)$ be the proposed Wiener filter associated with the error signal $e(n)$ defined by (3) and $\hat{v}(n)$ be the optimum Wiener filter based on minimizing the following criterion:*

$$J_M(v(n)) = \left\{ \sum_{k=-\infty}^{\infty} \mathrm{Cum}^{(M)}(e(n), e(n), x(n-k), \ldots, x(n-k)) \right\}^2 \geqslant J_M(\hat{v}(n)), \tag{9}$$

where $M \geqslant 3$. Then $\hat{v}(n)$ is identical to the conventional correlation-based Wiener filter associated with the MSE criterion $E[e_f^2(n)]$ (i.e., SNR $= \infty$), as long as $\gamma_M G_{M-2}(0) \neq 0$, where

$$G_m(\omega) \triangleq \sum_{n=-\infty}^{\infty} g^m(n) e^{-j\omega n}. \tag{10}$$

Proof. The correlation-based Wiener filter for the noise-free case is designed by minimizing

$$E[e_f^2(n)] = \sigma_u^2 \sum_{n=-\infty}^{\infty} f^2(n) \quad \text{(see (6) and (8))}. \tag{11}$$

On the other hand, the objective function J_M given by (9) can be simplified as follows:

$$J_M = \left\{ \sum_{k=-\infty}^{\infty} \mathrm{Cum}^{(M)}(e_f(n), e_f(n), x_f(n-k), \ldots, x_f(n-k)) \right\}^2 \quad \text{(see (A2))}$$

$$= \left\{ \sum_{k=-\infty}^{\infty} \left[\gamma_M \sum_{n=-\infty}^{\infty} f^2(n) g^{M-2}(n-k) \right] \right\}^2 \quad \text{(see (6), (1b) and (8))}$$

$$= \left\{ \gamma_M \sum_{k=-\infty}^{\infty} g^{M-2}(k) \right\}^2 \cdot \left\{ \sum_{n=-\infty}^{\infty} f^2(n) \right\}^2$$

$$= \left\{ \frac{\gamma_M G_{M-2}(0)}{\sigma_u^2} \right\}^2 \cdot \left\{ \sigma_u^2 \sum_{n=-\infty}^{\infty} f^2(n) \right\}^2 \quad \text{(see (10))}$$

$$= \left\{ \frac{\gamma_M G_{M-2}(0)}{\sigma_u^2} \right\}^2 \cdot \{E[e_f^2(n)]\}^2 \quad \text{(see (11))}. \tag{12}$$

One can see, from (12), that minimizing J_M is equivalent to minimizing $E[e_f^2(n)]$ when $\gamma_M \cdot G_{M-2}(0) \neq 0$. Therefore, the optimum $\hat{v}(n)$ associated with J_M is identical to the conventional Wiener filter as if SNR $= \infty$. $\quad \square$

It is well known that the Wiener–Hopf (linear) equation used to solve for correlation-based Wiener filter coefficients can be easily obtained by the orthogonality principle [8, 9, 16]. Analogously, one can also obtain the Wiener–Hopf equation associated with the proposed criterion J_M using a cumulant-based orthogonality principle described in the following theorem.

Theorem 2 (Cumulant-based orthogonality principle). *The optimum Wiener filter output $e(n)$ associated with J_M given by (9) satisfies*

$$\sum_{k=-\infty}^{\infty} \mathrm{Cum}^{(M)}(e(n), x(n-i), x(n-k), \ldots, x(n-k)) = 0, \quad i = p_1, p_1 + 1, \ldots, p_2, \tag{13}$$

with the minimum of J_M given by

$$J_{M,\min} = \left\{ \sum_{k=-\infty}^{\infty} \mathrm{Cum}^{(M)}(e(n), d(n), x(n-k), \ldots, x(n-k)) \right\}^2 . \tag{14}$$

See Appendix A for the proof of this theorem.

By the cumulant-based orthogonality principle described in Theorem 2, the Wiener–Hopf equation associated with J_M can be derived as follows:

$$\sum_{k=-\infty}^{\infty} \mathrm{Cum}^{(M)}(e(n), x(n-i), x(n-k), \ldots, x(n-k)) \quad \text{(see (13))}$$

$$= \sum_{k=-\infty}^{\infty} \mathrm{Cum}^{(M)}\left(d(n) - \sum_{j=p_1}^{p_2} \hat{v}(j) x(n-j), x(n-i), x(n-k), \ldots, x(n-k)\right)$$

$$= \sum_{k=-\infty}^{\infty} \mathrm{Cum}^{(M)}(d(n), x(n-i), x(n-k), \ldots, x(n-k))$$

$$\quad - \sum_{j=p_1}^{p_2} \hat{v}(j) \cdot \sum_{k=-\infty}^{\infty} \mathrm{Cum}^{(M)}(x(n-j), x(n-i), x(n-k), \ldots, x(n-k))$$

$$= \sum_{k=-\infty}^{\infty} \mathrm{Cum}^{(M)}(d(n), x(n-i), x(n-k), \ldots, x(n-k))$$

$$\quad - \sum_{j=p_1}^{p_2} \hat{v}(j) \cdot \sum_{k=-\infty}^{\infty} \mathrm{Cum}^{(M)}(x(n), x(n-i+j), x(n-k+j), \ldots, x(n-k+j))$$

$$= \sum_{k=-\infty}^{\infty} \mathrm{Cum}^{(M)}(d(n), x(n-i), x(n-k), \ldots, x(n-k))$$

$$\quad - \sum_{j=p_1}^{p_2} \hat{v}(j) \cdot \sum_{k=-\infty}^{\infty} \mathrm{Cum}^{(M)}(x(n), x(n-i+j), x(n-k), \ldots, x(n-k))$$

$$= c_{dx}(i) - \sum_{j=p_1}^{p_2} \hat{v}(j) c_{xx}(i-j) = 0, \quad i = p_1, p_1 + 1, \ldots, p_2, \tag{15}$$

where

$$c_{xx}(i) \triangleq \sum_{k=-\infty}^{\infty} \mathrm{Cum}^{(M)}(x(n), x(n-i), x(n-k), \ldots, x(n-k)) \tag{16a}$$

and

$$c_{dx}(i) \triangleq \sum_{k=-\infty}^{\infty} \mathrm{Cum}^{(M)}(d(n), x(n-i), x(n-k), \ldots, x(n-k)). \tag{16b}$$

From (15), one can obtain the cumulant-based Wiener–Hopf (linear) equation as follows:

$$\sum_{j=p_1}^{p_2} c_{xx}(i-j)\hat{v}(j) = c_{dx}(i), \quad i = p_1, p_1 + 1, \ldots, p_2 \tag{17}$$

or, in matrix form,

$$\mathbf{C}_x \hat{\mathbf{v}} = \mathbf{c}_{dx}, \tag{18}$$

where $\hat{\boldsymbol{v}} = [\hat{v}(p_1), \hat{v}(p_1 + 1), \ldots, \hat{v}(p_2)]^T$, $\boldsymbol{c}_{dx} = [c_{dx}(p_1), c_{dx}(p_1 + 1), \ldots, c_{dx}(p_2)]^T$, and

$$
\boldsymbol{C}_x = \begin{bmatrix}
c_{xx}(0) & c_{xx}(-1) & \cdots & c_{xx}(p_1 - p_2) \\
c_{xx}(1) & c_{xx}(0) & \cdots & c_{xx}(p_1 - p_2 + 1) \\
\vdots & & \ddots & \\
c_{xx}(-p_1 + p_2) & c_{xx}(-p_1 + p_2 - 1) & \cdots & c_{xx}(0)
\end{bmatrix}
\tag{19}
$$

is a symmetric Toeplitz matrix because $c_{xx}(i) = c_{xx}(-i)$ (see (16a)).

Moreover, the minimum value of the objective function J_M follows directly from Theorem 2 as

$$
\begin{aligned}
J_{M,\min} &= \left\{ \sum_{k=-\infty}^{\infty} \mathrm{Cum}^{(M)}(e(n), d(n), x(x - k), \ldots, x(n - k)) \right\}^2 \quad \text{(see (14))} \\
&= \left\{ \sum_{k=-\infty}^{\infty} \mathrm{Cum}^{(M)}\left(d(n) - \sum_{i=p_1}^{p_2} \hat{v}(i) x(n - i), d(n), x(n - k), \ldots, x(n - k)\right) \right\}^2 \\
&= \left\{ \sum_{k=-\infty}^{\infty} \mathrm{Cum}^{(M)}(d(n), d(n), x(n - k), \ldots, x(n - k)) \right. \\
&\qquad \left. - \sum_{i=p_1}^{p_2} \hat{v}(i) \cdot \sum_{k=-\infty}^{\infty} \mathrm{Cum}^{(M)}(x(n - i), d(n), x(n - k), \ldots, x(n - k)) \right\}^2 \\
&= \left\{ c_{dd}(0) - \sum_{i=p_1}^{p_2} \hat{v}(i) c_{dx}(i) \right\}^2 \quad \text{(see (16b))},
\end{aligned}
\tag{20}
$$

where

$$
c_{dd}(i) \triangleq \sum_{k=-\infty}^{\infty} \mathrm{Cum}^{(M)}(d(n), d(n - i), x(n - k), \ldots, x(n - k)).
\tag{21}
$$

Tables 1 and 2 summarize the correlation-based Wiener filter and the proposed cumulant-based Wiener filter, respectively. One can see, from these two tables, that all the correlation-based equations associated with the conventional Wiener filter can be mapped, respectively, to the corresponding cumulant-based counterparts associated with the proposed cumulant-based Wiener filter. This indicates that the proposed cumulant-based Wiener filter is closely related to the correlation-based Wiener filter, and the relationship between them will be further discussed in the next section.

In practice, both $c_{xx}(i)$ and $c_{dx}(i)$ given by (16a) and (16b), respectively, can be estimated from data as

$$
\hat{c}_{xx}(i) = \sum_{k=K_{xx1}}^{K_{xx2}} \hat{\mathrm{C}}\mathrm{um}^{(M)}(x(n), x(n - i), x(n - k), \ldots, x(n - k))
\tag{22a}
$$

and

$$
\hat{c}_{dx}(i) = \sum_{k=K_{dx1}}^{K_{dx2}} \hat{\mathrm{C}}\mathrm{um}^{(M)}(d(n), x(n - i), x(n - k), \ldots, x(n - k)),
\tag{22b}
$$

respectively, where $\hat{\mathrm{C}}\mathrm{um}^{(M)}(x_1(n + k_1), x_2(n + k_2), \ldots, x_M(n + k_M))$ denotes the biased Mth-order sample cumulant sequence of stationary random processes $\{x_i(n)\}$, $i = 1, 2, \ldots, M$ [11, 14] and $K_{xx1}, K_{xx2}, K_{dx1}$ as well as K_{dx2} are integers which must be chosen such that $\hat{c}_{xx}(i)$ and $\hat{c}_{dx}(i)$ approximate $c_{xx}(i)$ and $c_{dx}(i)$, respectively.

Table 1
Summary of the correlation-based Wiener filter

Table 2
Summary of the proposed cumulant-based Wiener filter

MSE criterion

$\min E[e^2(n)]$

Cumulant-based MSE criterion

$\min J_M = \{\sum_{k=-\infty}^{\infty} \mathrm{Cum}^{(M)}(e(n), e(n), x(n-k), \ldots, x(n-k))\}^2$

Orthogonality principle

$E[e(n)x(n-i)] = 0, \quad i = p_1, p_1 + 1, \ldots, p_2$

$\min\{E[e^2(n)]\} = E[e(n)d(n)]$

Cumulant-based orthogonality principle

$\sum_{k=-\infty}^{\infty} \mathrm{Cum}^{(M)}(e(n), x(n-i), x(n-k), \ldots, x(n-k)) = 0,$
$\qquad i = p_1, p_1 + 1, \ldots, p_2$

$J_{M,\min} = \{\sum_{k=-\infty}^{\infty} \mathrm{Cum}^{(M)}(e(n), d(n), x(n-k), \ldots, x(n-k))\}^2$

Wiener–Hopf equation

$\sum_{j=p_1}^{p_2} r_{xx}(i-j)\hat{v}(j) = r_{dx}(i), \quad i = p_1, p_1 + 1, \ldots, p_2$

$r_{xx}(i) = E[x(n)x(n-i)]$

$r_{dx}(i) = E[d(n)x(n-i)]$

Cumulant-based Wiener–Hopf equation

$\sum_{j=p_1}^{p_2} c_{xx}(i-j)\hat{v}(j) = c_{dx}(i), \quad i = p_1, p_1 + 1, \ldots, p_2$

$c_{xx}(i) = \sum_{k=-\infty}^{\infty} \mathrm{Cum}^{(M)}(x(n), x(n-i), x(n-k), \ldots, x(n-k))$

$c_{dx}(i) = \sum_{k=-\infty}^{\infty} \mathrm{Cum}^{(M)}(d(n), x(n-i), x(n-k), \ldots, x(n-k))$

How to choose the values of $K_{xx1}, K_{xx2}, K_{dx1}$ and K_{dx2} for the case that both $g(n)$ and $h(n)$ are FIR filters is presented in the following fact which is shown in Appendix B.

Fact 1. *Assume that $g(n)$ is an FIR filter of length L and $h(n)$ is also an FIR filter with $h(n) \neq 0$ for $L_1 \leqslant n \leqslant L_2$. The choices of K_{xx1} and K_{xx2} and the choices of K_{dx1} and K_{dx2} are described in (F1) and (F2), respectively, as follows.*
(F1) *When $\max\{-L + 1, -L + 1 + i\} \leqslant \min\{L - 1, L - 1 + i\}$, $\hat{c}_{xx}(i)$ is a consistent estimate for $c_{xx}(i)$ if K_{xx1} and K_{xx2} are chosen such that*

$$K_{xx1} \leqslant \max\{-L + 1, -L + 1 + i\} \tag{23a}$$

and

$$K_{xx2} \geqslant \min\{L - 1, L - 1 + i\}, \tag{23b}$$

respectively; otherwise, $c_{xx}(i) = 0$ which implies the associated $\hat{c}_{xx}(i) = 0$.
(F2) *When $\max\{-L + 1 + L_1, -L + 1 + i\} \leqslant \min\{L - 1 + L_2, L - 1 + i\}$, $\hat{c}_{dx}(i)$ is a consistent estimate for $c_{dx}(i)$ if K_{dx1} and K_{dx2} are chosen such that*

$$K_{dx1} \leqslant \max\{-L + 1 + L_1, -L + 1 + i\} \tag{24a}$$

and

$$K_{dx2} \geqslant \min\{L - 1 + L_2, L - 1 + i\}, \tag{24b}$$

respectively; otherwise, $c_{dx}(i) = 0$ which implies the associated $\hat{c}_{dx}(i) = 0$.

Note that the choices of K_{xx1} and K_{xx2} and the choices of K_{dx1} and K_{dx2} can be different for computing each $\hat{c}_{xx}(i)$ and each $\hat{c}_{dx}(i)$, respectively. Notice that when L is unknown, it can be replaced with a larger value determined by prior information about $g(n)$. Similarly, when L_1 and L_2 are unknown, the former can be replaced with a smaller value and the latter can be replaced by a larger value determined by their prior information about $h(n)$ as well. Doing this will also increase the bias and variance of $\hat{c}_{xx}(i)$ and $\hat{c}_{dx}(i)$ and thus lead to some performance degradation of the proposed cumulant-based Wiener filter. Recall that the Wiener filter was assumed to be an FIR filter. Therefore the designed Wiener filter is actually an approximation to the FIR system $h(n)$, but when the length of $h(n)$ is very large, it may not be a good approximation to $h(n)$ for limited finite data. However, how to choose the values of $K_{xx1}, K_{xx2}, K_{dx1}$ and K_{dx2} when $g(n)$ is an IIR system or when the length of $g(n)$ is large will be discussed later.

Two remarks regarding the proposed cumulant-based Wiener filter are described as follows.

(R1) The Fourier transform of $c_{xx}(i)$ given by (16a) can be shown to be (see Appendix C)

$$C_{xx}(\omega) = \sum_{i=-\infty}^{\infty} c_{xx}(i) e^{-j\omega i} = \gamma_M G_{M-2}(0) \cdot |G_1(\omega)|^2,$$ (25)

which implies that the sequence $c_{xx}(i)$ is positive definite if $\gamma_M G_{M-2}(0) > 0$ and negative definite if $\gamma_M G_{M-2}(0) < 0$. Thus, $c_{xx}(i) = c_{xx}(-i)$ can be thought of as a legitimate correlation sequence if $\gamma_M G_{M-2}(0) > 0$ and the matrix C_x given by (19) is therefore a legitimate correlation matrix. Accordingly, the proposed cumulant-based Wiener filter, like the correlation-based Wiener filter, can be implemented by a lattic structure [16] associated with the well-known computationally efficient Levinson–Durbin recursion. This means that the computational complexity of the proposed cumulant-based Wiener filter is the same as that of the correlation-based Wiener filter, except for some additional computations for obtaining $\hat{c}_{xx}(i)$ and $\hat{c}_{dx}(i)$, which depend on the choices of $K_{xx1}, K_{xx2}, K_{dx1}$ as well as K_{dx2}.

(R2) When $h(n) = \delta(n+1)$, i.e., $d_f(n) = x_f(n+1)$ (see (2b)), the proposed cumulant-based Wiener filter $v(n)$ reduces to a cumulant-based linear prediction (LP) filter proposed by Chi et al. [4, 5]. Therefore, the associated prediction error signal $e(n)$ is also equivalent to the output signal of the cumulant-based LPE filter reported in [4, 5] with the input being $x(n)$.

3. Projection of higher-order cumulants

In order to provide a further insight into the proposed cumulant-based Wiener filter, let us present a generalized projection concept as follows.

Theorem 3 (Generalized projection). *Let $y_i(n) = x_f(n) * h_i(n)$, $i = 1, 2, \ldots, m$, where $x_f(n)$ is the non-Gaussian noise-free signal given by (1b) under the assumption (A1) and $h_i(n)$, $i = 1, 2, \ldots, m$, are arbitrary LTI systems. Then*

$$\sum_{k_{m+1}=-\infty}^{\infty} \mathrm{Cum}^{(M)}(y_1(n+k_1), y_2(n+k_2), \ldots, y_m(n+k_m), x_f(n+k_{m+1}), \ldots, x_f(n+k_{m+1}))$$

$$= \frac{\gamma_M G_{M-m}(0)}{\gamma_m} \cdot \mathrm{Cum}^{(m)}(y_1(n+k_1), y_2(n+k_2), \ldots, y_m(n+k_m))$$ (26)

and

$$\sum_{k_{m+1}=-\infty}^{\infty} \cdots \sum_{k_M=-\infty}^{\infty} \mathrm{Cum}^{(M)}(y_1(n+k_1), y_2(n+k_2), \ldots, y_m(n+k_m),$$

$$x_f(n+k_{m+1}), \ldots, x_f(n+k_M)) \cdot \exp\left\{ -j \sum_{i=m+1}^{M} \omega_i k_i \right\}$$

$$= \frac{\gamma_M \prod_{i=m+1}^{M} G_1(\omega_i)}{\gamma_m} \cdot \mathrm{Cum}^{(m)}(y_1(n+k_1), y_2(n+k_2), \ldots, y_m(n+k_m)),$$ (27)

with $\sum_{i=m+1}^{M} \omega_i = 0$, where $2 \leqslant m < M$ and $G_m(\omega)$ is defined by (10).

See Appendix D for the proof of this theorem.

Theorem 3 indicates that an Mth-order cumulant function can be projected to an mth-order $(2 \leqslant m < M)$ cumulant function except for a scale factor. Notice that both (26) and (27) are generalizations of the projection operator proposed by Delopoulos and Giannakis [6], which projects a third-order cumulant function to an autocorrelation function except for a scale factor.

If $m = 2$ and $y_1(n + k_1) = y_2(n + k_2) = e_f(n)$ in (26) where $e_f(n)$ is defined by (6), then the square of the left-hand side of (26) is equivalent to the proposed criterion J_M given by (9) because of the assumption (A2). Moreover, by letting $m = 2$, $y_1(n + k_1) = e_f(n)$ and $y_2(n + k_2) = x_f(n - i)$ in (26), Eq. (13) in Theorem 2 can be simplified as follows:

$$
\sum_{k=-\infty}^{\infty} \text{Cum}^{(M)}(e(n), x(n-i), x(n-k), \dots, x(n-k))
$$

$$
= \sum_{k=-\infty}^{\infty} \text{Cum}^{(M)}(e_f(n), x_f(n-i), x_f(n-k), \dots, x_f(n-k))
$$

$$
= \frac{\gamma_M G_{M-2}(0)}{\sigma_u^2} \cdot E[e_f(n) x_f(n-i)] = 0 \quad (\text{see (26)})
$$

or

$$
E[e_f(n) x_f(n-i)] = 0
$$

if $\gamma_M G_{M-2}(0) \neq 0$. This implies that the cumulant-based orthogonality principle is equivalent to the correlation-based orthogonality principle [8, 9, 16] associated with the noise-free case. Similarly, $c_{xx}(i)$ and $c_{dx}(i)$ given by (16a) and (16b), respectively, can be shown to be

$$
c_{xx}(i) = \frac{\gamma_M G_{M-2}(0)}{\sigma_u^2} \cdot E[x_f(n) x_f(n-i)] = \frac{\gamma_M G_{M-2}(0)}{\sigma_u^2} \cdot r_{x_f x_f}(i) \tag{28a}
$$

and

$$
c_{dx}(i) = \frac{\gamma_M G_{M-2}(0)}{\sigma_u^2} \cdot E[d_f(n) x_f(n-i)] = \frac{\gamma_M G_{M-2}(0)}{\sigma_u^2} \cdot r_{d_f x_f}(i), \tag{28b}
$$

respectively. Note that the square of the unknown scale factor $\gamma_M G_{M-2}(0)/\sigma_u^2$ in (28a) and (28b) is also the scale factor of the key relationship (see (12)) between the proposed cumulant-based MSE criterion J_M and the correlation-based MSE criterion. Thus, substituting (28a) and (28b) into (17) yields

$$
\sum_{j=p_1}^{p_2} r_{x_f x_f}(i-j) \hat{v}(j) = r_{d_f x_f}(i), \quad i = p_1, p_1 + 1, \dots, p_2, \tag{29}
$$

which also implies that the designed cumulant-based Wiener filter is not sensitive to the value of $\gamma_M G_{M-2}(0)/\sigma_u^2$ which turns out to disappear in (29). Once again, this states that the cumulant-based Wiener–Hopf equation given by (17) for finite SNR is equivalent to the correlation-based Wiener–Hopf equation given by (4) for SNR $= \infty$, and therefore the obtained cumulant-based Wiener filter $\hat{v}(n)$ is identical to the correlation-based Wiener filter associated with the noise-free case.

Recently, Delopoulos and Giannakis [7] proposed a cumulant-based input-output system identification method based on the following criterion:

$$
J_M^{(DG)} = \sum_{k_3=-\infty}^{\infty} \cdots \sum_{k_M=-\infty}^{\infty} \text{Cum}^{(M)}(e(n), e(n), x(n+k_3), \dots, x(n+k_M)) \cdot \exp\left\{-j \sum_{i=3}^{M} \omega_i k_i\right\}, \tag{30}
$$

where $\sum_{i=3}^{M} \omega_i = 0$. One can see, from Theorem 3, that Delopoulos and Giannakis' criterion $J_M^{(DG)}$ given by (30) is equivalent to the left-hand side of (27) with $m = 2$ and $y_1(n + k_1) = y_2(n + k_2) = e_f(n)$, and therefore is also a cumulant-based MSE criterion. Furthermore, one can observe that the proposed criterion J_3 is equivalent to Delopoulos and Giannakis' criterion $J_3^{(DG)}$ because (26) associated with the former is the same as (27) associated with the latter for $m = 2$ and $M = 3$. However, the proposed criterion J_M uses only a 'one-dimensional slice' of Mth-order cross cumulants (see (9)) while Delopoulos and Giannakis' criterion $J_M^{(DG)}$ uses an '($M - 2$)-dimensional slice' of Mth-order cross cumulants (see (30)). Therefore, the proposed criterion J_M is computationally much more practical than Delopoulos and Giannakis' criterion $J_M^{(DG)}$ for $M \geqslant 4$.

4. Simulation results

In this section, the proposed criterion for the design of Wiener filters is to be applied to the input-output moving-average (MA) system identification and time delay estimation through simulation in order to demonstrate the good performance of the proposed cumulant-based Wiener filter.

4.1. System identification

The proposed cumulant-based Wiener filter $v(n)$ was used to identify an LTI system $h(n)$. In the following two examples, the driving input $u(n)$ used was a zero-mean, exponentially distributed, i.i.d. random sequence with variance $\sigma_u^2 = 1$, skewness $\gamma_3 = 2$ and kurtosis $\gamma_4 = 6$. The system $g(n) = \delta(n)$ was assumed ($L = 1$) for which $\gamma_3 G_1(0) = \gamma_3 \neq 0$ and $\gamma_4 G_2(0) = \gamma_4 \neq 0$. A second-order MA system $h(n)$ with transfer function

$$H(z) = h(0) + h(1)z^{-1} + h(2)z^{-2} = 1 - 1.8z^{-1} + 0.4z^{-2}, \tag{31}$$

whose zeros are 0.2597 and 1.5403 (i.e., nonminimum-phase system), was used. The optimum cumulant-based Wiener filter $\hat{v}(n)$ was obtained using the proposed criterion J_3 as well as J_4, and the filter coefficients were solved from (17) in which $c_{xx}(i)$ and $c_{dx}(i)$ were replaced by $\hat{c}_{xx}(i)$ and $\hat{c}_{dx}(i)$ given by (22a) and (22b), respectively, with $K_{xx1} = \max\{-L + 1, -L + 1 + i\}$, $K_{xx2} = \min\{L - 1, L - 1 + i\}$, $K_{dx1} = \max\{-L + 1 + L_1, -L + 1 + i\}$, and $K_{dx2} = \min\{L - 1 + L_2, L - 1 + i\}$ (see (23a), (23b), (24a) and (24b)). Note that $L = 1$, $p_1 = L_1 = 0$ and $p_2 = L_2 = 2$ were used in the two examples. Thirty independent runs were performed for each simulation example with the same signal-to-noise ratio (SNR) defined as

$$SNR = \frac{E[x_f^2(n)]}{E[w_1^2(n)]} = \frac{E[d_f^2(n)]}{E[w_1^2(n)]} \quad \text{(see (1a) and (2a))} \tag{32}$$

associated with the noisy data $x(n)$ and $d(n)$. For comparison, the correlation-based Wiener filter [8, 9, 16] was also employed to estimate $h(n)$ with the same simulation data.

Example 1 (*Uncorrelated white noise sources*). The noise sources $w_1(n)$ and $w_2(n)$ were assumed to be zero-mean i.i.d. Gaussian random sequences and statistically uncorrelated. Table 3 shows mean $\pm$ standard deviation of the obtained 30 independent estimates $\hat{v}(n)$ for data length $N = 4000$, and SNR = 40, 10, 5 and 0 dB. One can see, from Table 3, that when SNR is large (SNR = 40 dB), mean values of the estimates $\hat{v}(n)$ are very close to the true MA parameters $h(n)$ for all the criteria. However, when SNR is low (SNR = 0 dB), biases of the estimates $\hat{v}(n)$ associated with the MSE criterion are quite large in spite of small standard deviations. On the other hand, the proposed cumulant-based Wiener filter keeps both bias and standard deviation small.

Table 3
Simulation results of Example 1. The noise sources are white Gaussian and uncorrelated with each other, and data length $N = 4000$.
True parameters: $h(0) = 1.0$, $h(1) = -1.8$, $h(2) = 0.4$

| Criterion | | Estimated values (mean ± standard deviation) | | | |
		SNR = 40 dB	SNR = 10 dB	SNR = 5 dB	SNR = 0 dB
MSE	$\hat{v}(0)$	1.0000 ± 0.0007	0.9129 ± 0.0168	0.7653 ± 0.0246	0.5046 ± 0.0303
	$\hat{v}(1)$	-1.7996 ± 0.0007	-1.6383 ± 0.0128	-1.3710 ± 0.0241	-0.9034 ± 0.0387
	$\hat{v}(2)$	0.3999 ± 0.0006	0.3672 ± 0.0117	0.3096 ± 0.0196	0.2056 ± 0.0289
J_3	$\hat{v}(0)$	1.0092 ± 0.0444	1.0116 ± 0.0467	1.0148 ± 0.0518	1.0258 ± 0.0660
	$\hat{v}(1)$	-1.8020 ± 0.0281	-1.8021 ± 0.0307	-1.8044 ± 0.0431	-1.8148 ± 0.0892
	$\hat{v}(2)$	0.4050 ± 0.0439	0.4099 ± 0.0453	0.4154 ± 0.0522	0.4295 ± 0.0765
J_4	$\hat{v}(0)$	1.0090 ± 0.0771	1.0145 ± 0.0807	1.0179 ± 0.0932	1.0204 ± 0.1376
	$\hat{v}(1)$	-1.7990 ± 0.0396	-1.7986 ± 0.0489	-1.7979 ± 0.0680	-1.7933 ± 0.1255
	$\hat{v}(2)$	0.4043 ± 0.0692	0.4096 ± 0.0789	0.4129 ± 0.0963	0.4147 ± 0.1457

Table 4
Simulation results of Example 2. The noise sources are colored Gaussian and uncorrelated with each other, and data length $N = 4000$.
True parameters: $h(0) = 1.0$, $h(1) = -1.8$, $h(2) = 0.4$

| Criterion | | Estimated values (mean ± standard deviation) | | | |
		SNR = 40 dB	SNR = 10 dB	SNR = 5 dB	SNR = 0 dB
MSE	$\hat{v}(0)$	1.0001 ± 0.0008	0.8465 ± 0.0179	0.6202 ± 0.0252	0.3058 ± 0.0315
	$\hat{v}(1)$	-1.7996 ± 0.0008	-1.5900 ± 0.0172	-1.2825 ± 0.0271	-0.8295 ± 0.0354
	$\hat{v}(2)$	0.3998 ± 0.0007	0.2975 ± 0.0115	0.1596 ± 0.0197	0.0017 ± 0.0300
J_3	$\hat{v}(0)$	1.0092 ± 0.0445	1.0133 ± 0.0490	1.0185 ± 0.0572	1.0324 ± 0.0790
	$\hat{v}(1)$	-1.8020 ± 0.0281	-1.8031 ± 0.0347	-1.8054 ± 0.0493	-1.8105 ± 0.0924
	$\hat{v}(2)$	0.4049 ± 0.0439	0.4078 ± 0.0460	0.4114 ± 0.0529	0.4187 ± 0.0770
J_4	$\hat{v}(0)$	1.0090 ± 0.0772	1.0155 ± 0.0841	1.0216 ± 0.0971	1.0325 ± 0.1371
	$\hat{v}(1)$	-1.7991 ± 0.0397	-1.8024 ± 0.0521	-1.8052 ± 0.0728	-1.8062 ± 0.1296
	$\hat{v}(2)$	0.4042 ± 0.0692	0.4082 ± 0.0794	0.4106 ± 0.0984	0.4096 ± 0.1618

Example 2 (*Uncorrelated colored noise sources*). The noise source $w_1(n)$ as well as $w_2(n)$ used was generated from a first-order highpass FIR filter with coefficients $\{1, -0.8\}$ driven by a white Gaussian noise sequence, respectively, and $w_1(n)$ and $w_2(n)$ were statistically uncorrelated. Table 4 shows mean $\pm$ standard deviation of the 30 independent estimates $\hat{v}(n)$ obtained for $N = 4000$, and SNR = 40, 10, 5 and 0 dB. Again, when SNR is large (SNR = 40 dB), mean values of the estimates $\hat{v}(n)$ are very close to the true MA parameters $h(n)$ for all the criteria. When SNR is low (SNR = 0 dB), biases of the estimates $\hat{v}(n)$ associated with the MSE criterion are much larger than those associated with the proposed criterion J_3 as well as J_4 although standard deviations for the former (MSE) are smaller than those for the latter (J_3 and J_4). Moreover, from Tables 3 and 4, one can observe that the performance of the correlation-based Wiener filter for the colored Gaussian noise sources is worse than that for the white Gaussian noise sources for this case. However, the performance of the proposed cumulant-based Wiener filter is insensitive to Gaussian noise sources no matter whether noise sources are white or colored.

4.2. Time delay estimation [1–3, 10, 13, 17]

Assume that $x(n)$ and $d(n)$ are two spatially separated sensor measurements that satisfy

$$x(n) = x_f(n) + w_1(n) \tag{33}$$

and

$$d(n) = x_f(n - D) + w_2(n), \tag{34}$$

respectively, where $x_f(n)$ is an unknown signal and D is an unknown time delay. Note that $d(n)$ given by (34) is a special case of (2) with $h(n) = \delta(n - D)$. The cumulant-based Wiener filter $\hat{v}(n)$ can be used to estimate the time delay D with $p_1 = -p$ and $p_2 = p$, where p is the largest possible time delay one can expect. A time delay estimate, denoted $\hat{D}$, can then be determined to be the index associated with $\hat{v}(\hat{D}) = \max\{\hat{v}(n), -p \leqslant n \leqslant p\}$ assuming that D is an integer. However, when the time delay D is not an integer, one can estimate D by applying sampling interpolation formula [2, 3] to the obtained cumulant-based Wiener filter $\hat{v}(n)$.

Example 3. The driving input $u(n)$ used was the same as that used in Example 1, and the unknown signal $x_f(n) = u(n)$ (i.e., $g(n) = \delta(n)$) was used to generate measurements $x(n)$ and $d(n)$ for data length $N = 4000$ and time delay $D = 8$. The noise source $w_1(n)$ was assumed to be a colored Gaussian sequence generated from a first-order MA system with coefficients $\{1, 0.8\}$ driven by a white Gaussian noise sequence, and the other noise source $w_2(n) = w_1(n - 3)$ (i.e., sensor noise sources were spatially coherent and colored Gaussian). The optimum cumulant-based Wiener filter $\hat{v}(n)$ was also obtained by solving (17) with $K_{xx1}, K_{xx2}, K_{dx1}$ as well as K_{dx2} chosen in the same way as in Example 1. Note that $L = 1, p_1 = L_1 = -p$, and $p_2 = L_2 = p$ were used in the example. Thirty independent runs were performed for $p = 30$, and SNR = 0 dB and SNR = −5 dB. For comparison, the conventional Wiener-filter-based method proposed by Chan et al. [2], the parametric bispectrum method proposed by Nikias and Pan [13] and the cumulant-based time delay parameter estimation (CUM-TDPE) method proposed by Tugnait [17] were also employed to estimate D with the same simulation data. Note that Nikias and Pan's parametric bispectrum method estimates the time delay D by solving a set of overdetermined linear equations [13] formed of third-order cumulants and cross cumulants, and Tugnait's CUM-TDPE method [17] is a fourth-order cumulant extension of Nikias and Pan's bispectrum method.

Figs. 2 and 3 show the 30 independent estimates $\hat{v}(n)$ obtained for SNR = 0 dB and SNR = −5 dB, respectively, associated with the Chan et al. Wiener-filter-based method, Nikias and Pan's parametric bispectrum method, Tugnait's CUM-TDPE method and the proposed Wiener-filter-based method for $M = 3$ as well as $M = 4$. From Fig. 2(a), one can see that all the estimates $\hat{v}(n)$ approximate $0.43\delta(n - 3) + 0.57\delta(n - 8)$ because both $x_f(n)$ and $w_2(n) = w_1(n - 3)$ were treated as signals with unknown time delay by the correlation-based Wiener filter although the variance is small. From Figs. 2(b)–(e), one can see that all the estimates $\hat{v}(n)$ approximate $\delta(n - 8)$ except for a scale factor. The simulation results shown in Fig. 2 demonstrate that all the above HOS-based time delay estimation methods are effective for suppressing coherent and colored Gaussian noise sources for the case of SNR = 0 dB. Again, all the estimates $\hat{v}(n)$ shown in Fig. 3(a) approximate $0.63\delta(n - 3) + 0.37\delta(n - 8)$ and fail to provide reliable estimates $\hat{D}$ for SNR = −5 dB. From Figs. 3(b) and (c), one can see that all the estimates $\hat{v}(n)$ associated with Nikias and Pan's method approximate $0.32\delta(n - 3) + 0.70\delta(n - 8)$ and those associated with Tugnait's CUM-TDPE method fail to provide reliable results for SNR = −5 dB, respectively. However, all the estimates $\hat{v}(n)$ shown in Fig. 3(d) as well as Fig. 3(e) approximate $\delta(n - 8)$ except for a scale factor for this case. The time delay due to coherent Gaussian noise sources is completely suppressed by the proposed method. The simulation results shown in Fig. 3 demonstrate that the proposed method is more robust than Nikias and Pan's method and Tugnait's method for the white signal case. Moreover, it is known [14] that sample cumulants and sample cross cumulants are consistent estimates but their variance increases with cumulant order. Therefore, the variance of Nikias and Pan's method is smaller than the variance of Tugnait's CUM-TDPE method, and the variance of the proposed method for $M = 3$ is also smaller than that for $M = 4$.
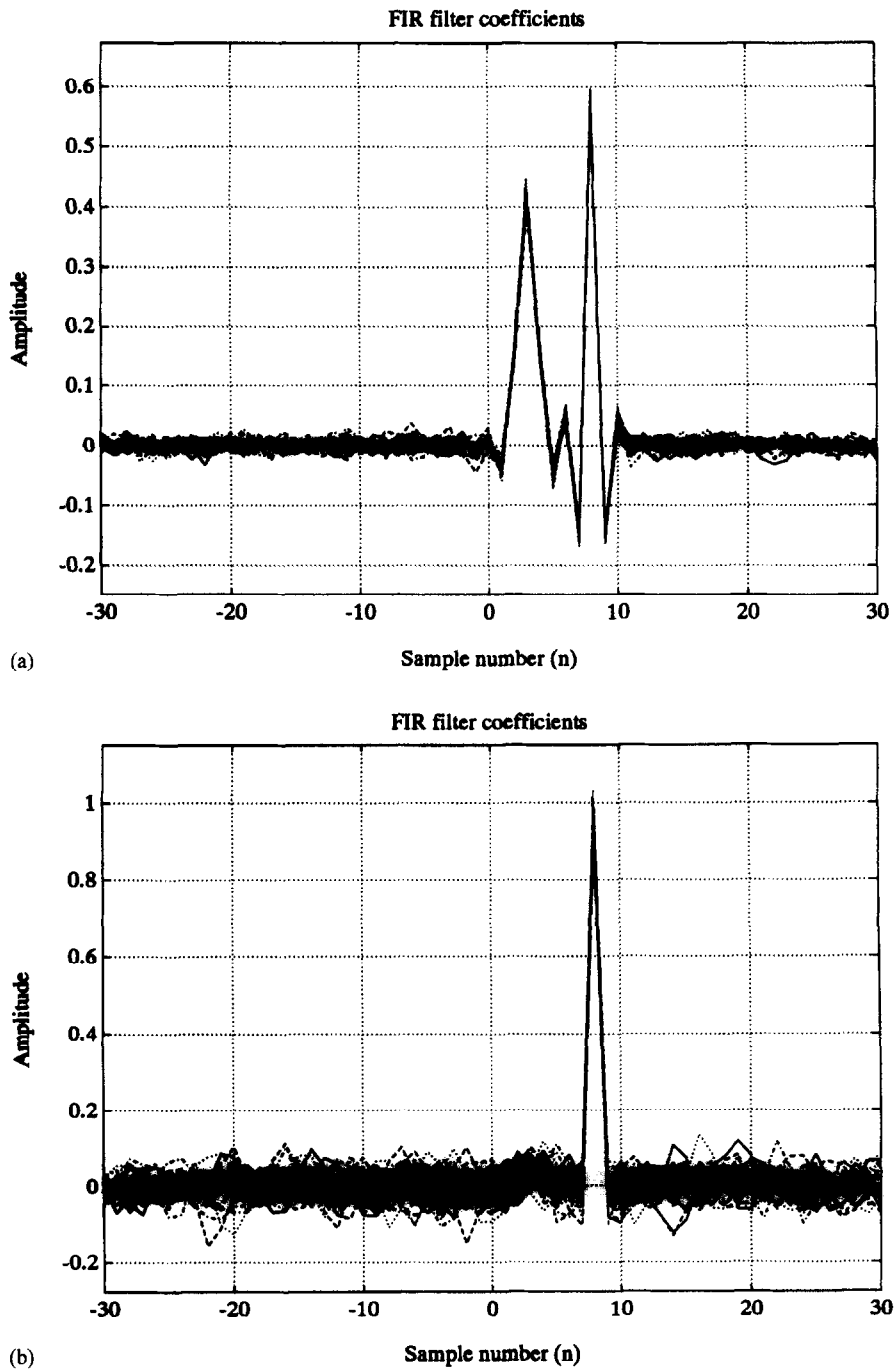
FIR filter coefficients

(a)

FIR filter coefficients

(b)

Fig. 2. Simulation results of Example 3 ($N = 4000$ and SNR $= 0$ dB). The true time delay is $D = 8$, the signal $x_f(n)$ is white, the noise sources are spatially coherent and colored Gaussian. Thirty estimates $\hat{v}(n)$ for $p = 30$ shown in the figure were obtained using (a) the Chan et al. Wiener-filter-based method, (b) Nikias and Pan's parametric bispectrum method.
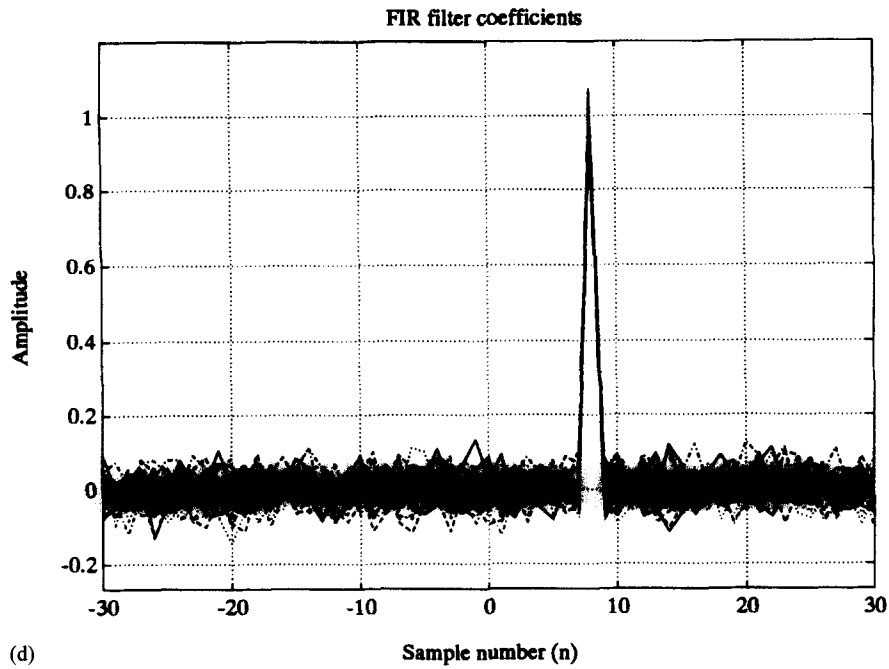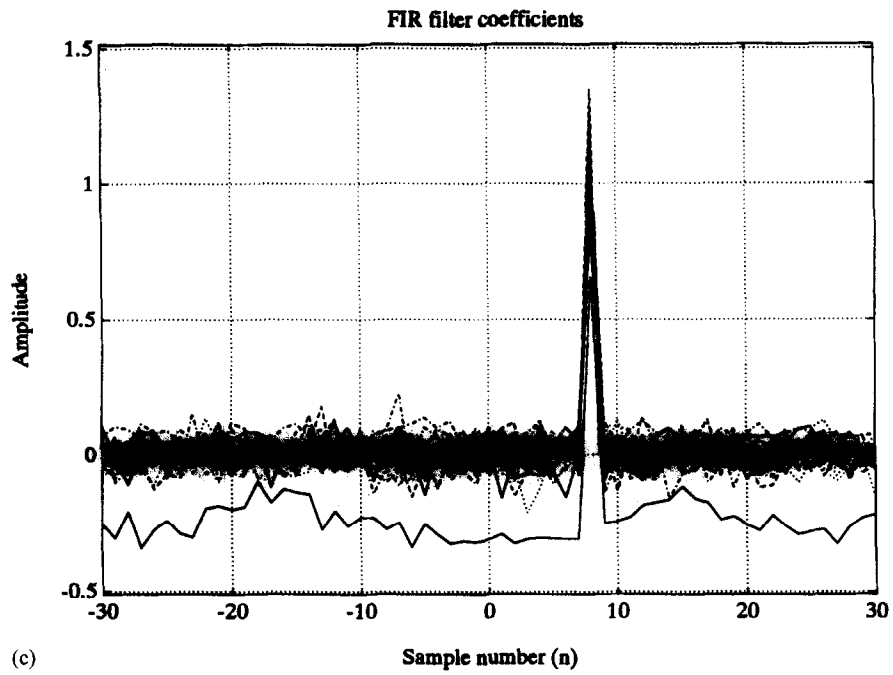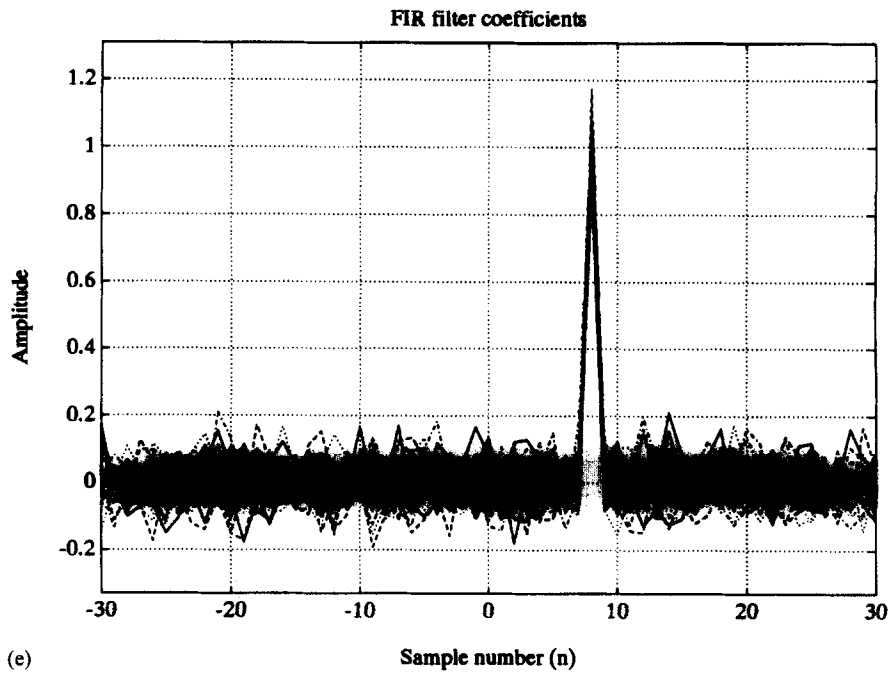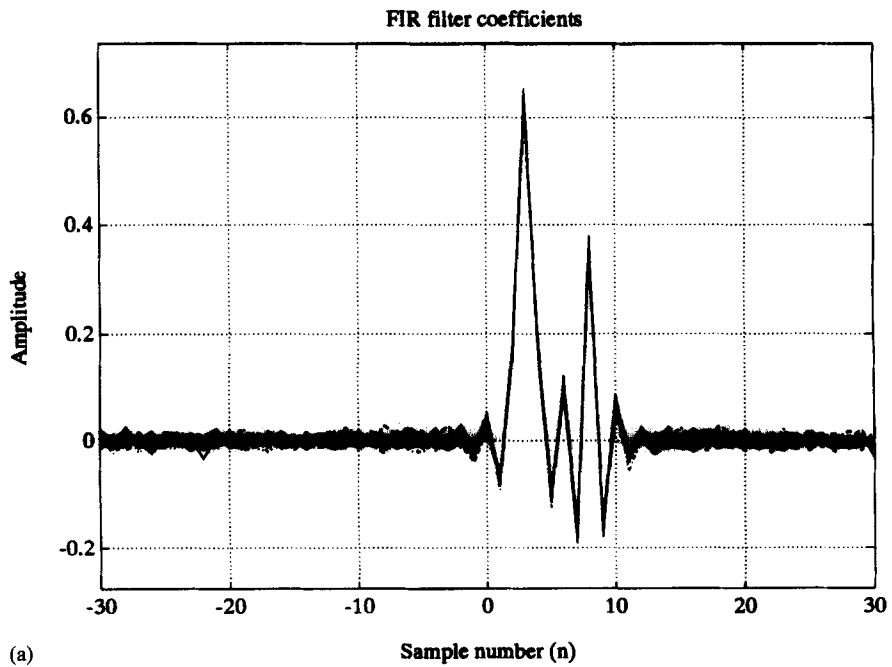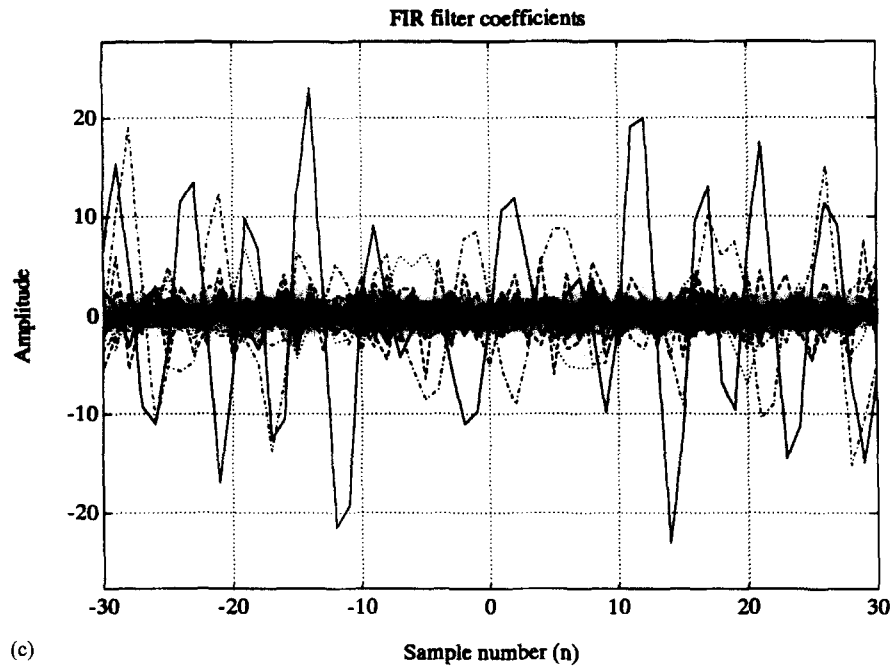
(c)

(d)

Fig. 2. (c) Tugnait's CUM-TDPE method and the proposed Wiener-filter-based method with (d) $M = 3$.

FIR filter coefficients

(e)

Fig. 2. (e) The proposed Wiener-filter-based method with $M = 4$.

FIR filter coefficients

(a)

Fig. 3. Simulation results of Example 3 ($N = 4000$ and SNR $= -5$ dB). The true time delay is $D = 8$, the signal $x_f(n)$ is white, the noise sources are spatially coherent and colored Gaussian. Thirty estimates $\hat{v}(n)$ for $p = 30$ shown in the figure were obtained using (a) the Chan et al. Wiener-filter-based method.
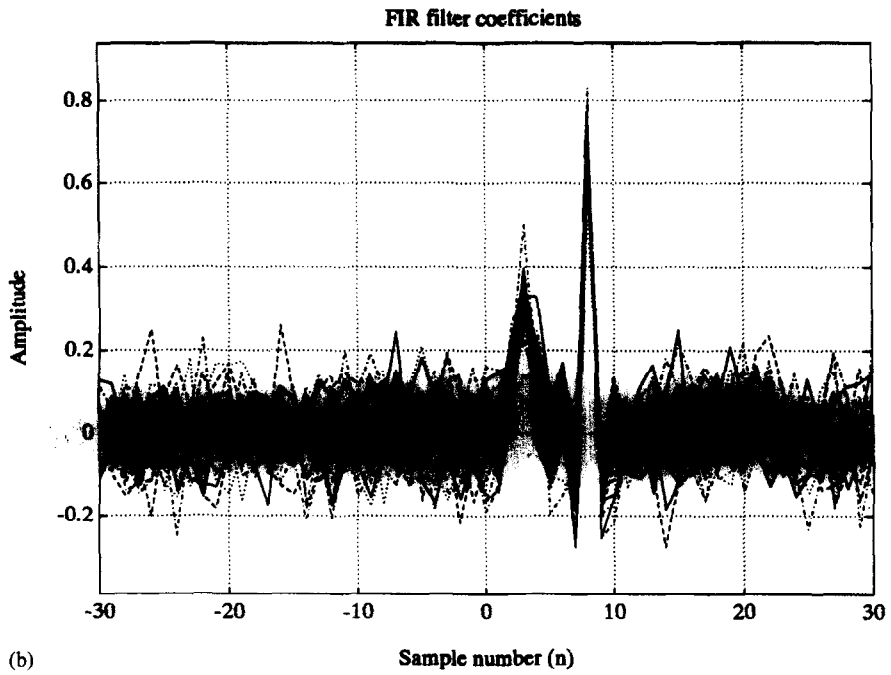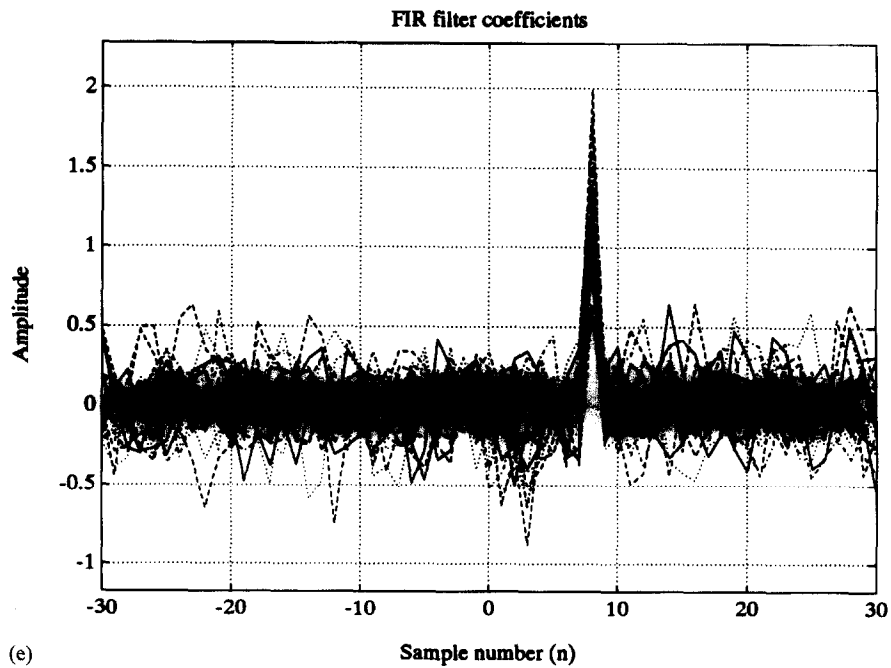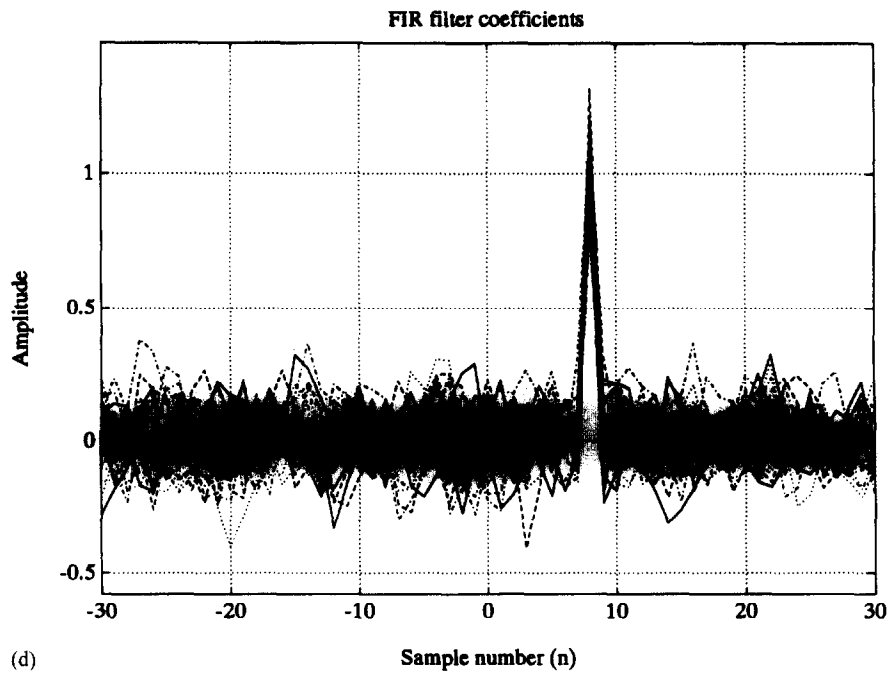
FIR filter coefficients

(b)   **Sample number (n)**

FIR filter coefficients

(c)   **Sample number (n)**

Fig. 3. (b) Nikias and Pan's parametric bispectrum method, (c) Tugnait's CUM-TDPE method and the proposed Wiener-filter-based method.

FIR filter coefficients

(d)

FIR filter coefficients

(e)

Fig. 3. (d) $M = 3$, (e) $M = 4$.

5. Conclusions

We have presented a cumulant-based MSE criterion J_M given by (9) for the design of Wiener filters. The designed cumulant-based Wiener filter with measurements corrupted by additive Gaussian noise sources was shown to be identical to the correlation-based Wiener filter with noise-free measurements (i.e., SNR $= \infty$) (see Theorem 1). Further, the proposed cumulant-based MSE criterion J_M leads to a cumulant-based orthogonality principle described in Theorem 2, and coefficients of the optimum cumulant-based Wiener filter can be solved from the associated cumulant-based Wiener–Hopf equation given by (17). Moreover, a generalized projection of Mth-order cumulants to mth-order cumulants $(2 \leqslant m < M)$ was presented in Theorem 3 to provide a further insight into the proposed cumulant-based Wiener filter. The proposed generalized projection theorem (Theorem 3) includes the projection of cumulants to correlations associated with the proposed cumulant-based MSE criterion and that associated with Delopoulos and Giannakis' cumulant-based MSE criterion as special cases. Finally, some simulation results for system identification and time delay estimation were provided to support the good performance of the proposed cumulant-based Wiener filter.

Recall that the proposed cumulant-based Wiener filter requires to compute $\hat{c}_{xx}(i)$ and $\hat{c}_{dx}(i)$ (see (22a) and (22b)) needed by the cumulant-based Wiener–Hopf equation given by (17). How to choose the values of $K_{xx1}, K_{xx2}, K_{dx1}$ and K_{dx2} for computing $\hat{c}_{xx}(i)$ and $\hat{c}_{dx}(i)$ was presented in Fact 1 in Section 2 for the case that $g(n)$ is an FIR system of length L. However, when $g(n)$ is an IIR system or when L is large, we suggest to preprocess measurements $x(n)$ and $d(n)$, respectively, by a whitening filter associated with $x(n)$ such as an LPE filter, denoted $h_w(n)$, so that $x(n)$ and $d(n)$ can be replaced by the preprocessed signals

$$\tilde{x}(n) = x(n) * h_w(n) \tag{35}$$

and

$$\tilde{d}(n) = d(n) * h_w(n), \tag{36}$$

respectively, for the design of cumulant-based Wiener filter. The reason for this is that $\tilde{x}(n)$ and $\tilde{d}(n)$ are now associated with the model shown in Fig. 1 (also see (1) and (2)) with $g(n)$ replaced by

$$\tilde{g}(n) = g(n) * h_w(n), \tag{37}$$

which usually has shorter length than $g(n)$. Then the values of $K_{xx1}, K_{xx2}, K_{dx1}$ and K_{dx} associated with $\tilde{x}(n)$ and $\tilde{d}(n)$ can be properly chosen by Fact 1. We empirically found that the designed cumulant-based Wiener filter $\hat{v}(n)$ using the prewhitened signals $\tilde{x}(n)$ and $\tilde{d}(n)$ always leads to smaller bias and variance than that without using the prewhitening filter. However, it cannot be guaranteed that the performance of the designed cumulant-based Wiener filter with the foregoing prewhitening process is satisfactory all the time, especially when $g(n)$ is a narrow-band system (its length is quite large) or when SNR is too low. We leave this as a future research topic.

Appendix A. Proof of Theorem 2

Let $v = [v(p_1), v(p_1 + 1), \ldots, v(p_2)]^T$ be any tap-weight vector, $v^{\perp}$ be the tap-weight vector satisfying (13) and $x(n) = [x(n - p_1), x(n - p_1 - 1), \ldots, x(n - p_2)]^T$. Then the estimation error $e(n)$ defined by (3) can be expressed as

$$\begin{aligned}
e(n) &= d(n) - v^T x(n) \\
&= [d(n) - (v^{\perp})^T x(n)] + (v^{\perp} - v)^T x(n) \\
&= e^{\perp}(n) + (v^{\perp} - v)^T x(n), \tag{A.1}
\end{aligned}$$

where

$$e^{\perp}(n) = d(n) - (v^{\perp})^{\mathrm{T}} x(n) \tag{A.2}$$

is the estimation error associated with $v^{\perp}$. Now observe that

$$J_M = \left\{ \sum_{k=-\infty}^{\infty} \mathrm{Cum}^{(M)}(e(n), e(n), x(n-k), \ldots, x(n-k)) \right\}^2 \quad (\text{see } (9))$$

$$= \left\{ \sum_{k=-\infty}^{\infty} \mathrm{Cum}^{(M)}(e^{\perp}(n) + (v^{\perp} - v)^{\mathrm{T}} x(n), e^{\perp}(n) + (v^{\perp} - v)^{\mathrm{T}} x(n), x(n-k), \ldots, x(n-k)) \right\}^2 \quad (\text{see } (A.1))$$

$$= \left\{ \sum_{k=-\infty}^{\infty} \mathrm{Cum}^{(M)}\bigg(e^{\perp}(n) + \sum_{i=p_1}^{p_2} [v^{\perp}(i) - v(i)] x(n-i), \right.$$
$$\left. e^{\perp}(n) + \sum_{j=p_1}^{p_2} [v^{\perp}(j) - v(j)] x(n-j), x(n-k), \ldots, x(n-k)) \right\}^2$$

$$= \left\{ \sum_{k=-\infty}^{\infty} \mathrm{Cum}^{(M)}(e^{\perp}(n), e^{\perp}(n), x(n-k), \ldots, x(n-k)) \right.$$

$$+ \sum_{i=p_1}^{p_2} [v^{\perp}(i) - v(i)] \cdot \sum_{k=-\infty}^{\infty} \mathrm{Cum}^{(M)}(x(n-i), e^{\perp}(n), x(n-k), \ldots, x(n-k))$$

$$+ \sum_{j=p_1}^{p_2} [v^{\perp}(j) - v(j)] \cdot \sum_{k=-\infty}^{\infty} \mathrm{Cum}^{(M)}(e^{\perp}(n), x(n-j), x(n-k), \ldots, x(n-k))$$

$$+ \sum_{i=p_1}^{p_2} \sum_{j=p_1}^{p_2} [v^{\perp}(i) - v(i)][v^{\perp}(j) - v(j)]$$

$$\left. \cdot \sum_{k=-\infty}^{\infty} \mathrm{Cum}^{(M)}(x(n-i), x(n-j), x(n-k), \ldots, x(n-k)) \right\}^2$$

$$= \left\{ \sum_{k=-\infty}^{\infty} \mathrm{Cum}^{(M)}(e^{\perp}(n), e^{\perp}(n), x(n-k), \ldots, x(n-k)) \right.$$

$$+ \sum_{i=p_1}^{p_2} \sum_{j=p_1}^{p_2} [v^{\perp}(i) - v(i)][v^{\perp}(j) - v(j)]$$

$$\left. \cdot \sum_{k=-\infty}^{\infty} \mathrm{Cum}^{(M)}(x(n-i), x(n-j), x(n-k), \ldots, x(n-k)) \right\}^2 \quad (\text{see } (13))$$

$$= \left\{ \sum_{k=-\infty}^{\infty} \mathrm{Cum}^{(M)}(e^{\perp}(n), e^{\perp}(n), x(n-k), \ldots, x(n-k)) \right.$$

$$+ \sum_{i=p_1}^{p_2} \sum_{j=p_1}^{p_2} [v^{\perp}(i) - v(i)][v^{\perp}(j) - v(j)]$$

$$\left. \cdot \sum_{k=-\infty}^{\infty} \left[\gamma_M \sum_{n=-\infty}^{\infty} g(n-i) g(n-j) g^{M-2}(n-k) \right] \right\}^2 \quad (\text{see } (A2), (1b) \text{ and } (8))$$

$$= \left\{ \sum_{k=-\infty}^{\infty} \mathrm{Cum}^{(M)}(e^{\perp}(n), e^{\perp}(n), x(n-k), \ldots, x(n-k)) \right.$$

$$+ \sum_{i=p_1}^{p_2} \sum_{j=p_1}^{p_2} [v^{\perp}(i) - v(i)][v^{\perp}(j) - v(j)]$$

$$\left. \cdot \left[\gamma_M \sum_{k=-\infty}^{\infty} g^{M-2}(k) \right] \cdot \sum_{n=-\infty}^{\infty} g(n-i)g(n-j) \right\}^2$$

$$= \left\{ \sum_{k=-\infty}^{\infty} \mathrm{Cum}^{(M)}(e^{\perp}(n), e^{\perp}(n), x(n-k), \ldots, x(n-k)) + \left[\gamma_M \sum_{k=-\infty}^{\infty} g^{M-2}(k) \right] \right.$$

$$\left. \cdot \sum_{n=-\infty}^{\infty} \left(\sum_{i=p_1}^{p_2} \sum_{j=p_1}^{p_2} [v^{\perp}(i) - v(i)][v^{\perp}(j) - v(j)]g(n-i)g(n-j) \right) \right\}^2$$

$$= \left\{ \sum_{k=-\infty}^{\infty} \mathrm{Cum}^{(M)}(e^{\perp}(n), e^{\perp}(n), x(n-k), \ldots, x(n-k)) \right.$$

$$\left. + \gamma_M G_{M-2}(0) \cdot \sum_{n=-\infty}^{\infty} \left(\sum_{i=p_1}^{p_2} [v^{\perp}(i) - v(i)]g(n-i) \right)^2 \right\}^2$$

$$= \left\{ \frac{\gamma_M G_{M-2}(0)}{\sigma_u^2} \cdot E[e_f^{\perp}(n)^2] + \gamma_M G_{M-2}(0) \sum_{n=-\infty}^{\infty} [(v^{\perp} - v)^{\mathrm{T}} g(n)]^2 \right\}^2 \quad \text{(see (12))}, \tag{A.3}$$

where $e_f^{\perp}(n)$ is the noise-free error signal associated with $e^{\perp}(n)$ and $g(n) = [g(n - p_1), g(n - p_1 - 1),$ $\ldots, g(n - p_2)]^{\mathrm{T}}$. Thus, J_M given by (A.3) can be expressed as

$$J_M = \left\{ \frac{\gamma_M G_{M-2}(0)}{\sigma_u^2} \right\}^2 \cdot \left\{ E[e_f^{\perp}(n)^2] + \sigma_u^2 \sum_{n=-\infty}^{\infty} [(v^{\perp} - v)^{\mathrm{T}} g(n)]^2 \right\}^2. \tag{A.4}$$

One can see, from (A.4), that J_M is minimized when $v = v^{\perp}$ and thus the optimum Wiener filter $\hat{v}(n)$ satisfies (13). Furthermore, the minimum value of J_M is given by

$$J_{M,\min} = \left\{ \frac{\gamma_M G_{M-2}(0)}{\sigma_u^2} \right\}^2 \cdot \{E[e_f^{\perp}(n)^2]\}^2$$

$$= \left\{ \sum_{k=-\infty}^{\infty} \mathrm{Cum}^{(M)}(e^{\perp}(n), e^{\perp}(n), x(n-k), \ldots, x(n-k)) \right\}^2 \quad \text{(see (12))}$$

$$= \left\{ \sum_{k=-\infty}^{\infty} \mathrm{Cum}^{(M)}\left(e^{\perp}(n), d(n) - \sum_{i=p_1}^{p_2} v^{\perp}(i)x(n-i), x(n-k), \ldots, x(n-k)\right) \right\}^2$$

$$= \left\{ \sum_{k=-\infty}^{\infty} \mathrm{Cum}^{(M)}(e^{\perp}(n), d(n), x(n-k), \ldots, x(n-k)) \right.$$

$$\left. - \sum_{i=p_1}^{p_2} v^{\perp}(i) \cdot \sum_{k=-\infty}^{\infty} \mathrm{Cum}^{(M)}(e^{\perp}(n), x(n-i), x(n-k), \ldots, x(n-k)) \right\}^2$$

$$= \left\{ \sum_{k=-\infty}^{\infty} \mathrm{Cum}^{(M)}(e^{\perp}(n), d(n), x(n-k), \ldots, x(n-k)) \right\}^2 \quad \text{(see (13))}, \tag{A.5}$$

which is (14). □

Appendix B. Proof of Fact 1

B.1. Proof of (F1)

The Mth-order cumulant $\text{Cum}^{(M)}(x(n), x(n-i), x(n-k), \ldots, x(n-k))$ used in $c_{xx}(i)$ (see (16a)) can be expressed as

$$\text{Cum}^{(M)}(x(n), x(n-i), x(n-k), \ldots, x(n-k))$$

$$= \text{Cum}^{(M)}(x_f(n+k), x_f(n+k-i), x_f(n), \ldots, x_f(n)) \quad \text{(see (A2))}$$

$$= \gamma_M \sum_{n=-\infty}^{\infty} g(n+k)g(n+k-i)g^{M-2}(n) \quad \text{(see (1b) and (8))}. \tag{B.1}$$

Let $g(n) \neq 0$ for $l_1 \leq n \leq l_2$ and $L = l_2 - l_1 + 1$ by the FIR filter assumption for $g(n)$. Then $\text{Cum}^{(M)}(x(n), x(n-i), x(n-k), \ldots, x(n-k)) \neq 0$ if

$$l_1 \leq n+k \leq l_2, \quad l_1 \leq n+k-i \leq l_2, \quad l_1 \leq n \leq l_2$$

or

$$l_1 - n \leq k \leq l_2 - n, \quad l_1 - n + i \leq k \leq l_2 - n + i, \quad l_1 \leq n \leq l_2$$

or

$$\max\{l_1 - n, l_1 - n + i\} \leq k \leq \min\{l_2 - n, l_2 - n + i\}, \quad l_1 \leq n \leq l_2$$

or

$$\max\{l_1 - l_2, l_1 - l_2 + i\} \leq k \leq \min\{l_2 - l_1, l_2 - l_1 + i\}$$

or

$$\max\{-L + 1, -L + 1 + i\} \leq k \leq \min\{L - 1, L - 1 + i\}. \tag{B.2}$$

Therefore, K_{xx1} and K_{xx2} (see (22a)) can be chosen such as

$$K_{xx1} \leq \max\{-L + 1, -L + 1 + i\} \tag{B.3a}$$

and

$$K_{xx2} \geq \min\{L - 1, L - 1 + i\}, \tag{B.3b}$$

respectively, and the associated $\hat{c}_{xx}(i)$ is a consistent estimate for $c_{xx}(i)$ when $\max\{-L + 1, -L + 1 + i\} \leq \min\{L - 1, L - 1 + i\}$. When $\max\{-L + 1, -L + 1 + i\} > \min\{L - 1, L - 1 + i\}$, $c_{xx}(i) = 0$ (since (B.2)) and the associated $\hat{c}_{xx}(i)$ does not need to be estimated.

B.2. Proof of (F2)

The Mth-order cross cumulant $\text{Cum}^{(M)}(d(n), x(n-i), x(n-k), \ldots, x(n-k))$ used in $c_{dx}(i)$ (see (16b)) can be expressed as

$$\text{Cum}^{(M)}(d(n), x(n-i), x(n-k), \ldots, x(n-k))$$

$$= \text{Cum}^{(M)}(\sum_{j=-\infty}^{\infty} h(j)x_f(n-j), x_f(n-i), x_f(n-k), \ldots, x_f(n-k)) \quad \text{(see (A2) and (2b))}$$

$$= \sum_{j=-\infty}^{\infty} h(j) \cdot \text{Cum}^{(M)}(x_f(n+k-j), x_f(n+k-i), x_f(n), \dots, x_f(n))$$

$$= \sum_{j=-\infty}^{\infty} h(j) \cdot \left[\gamma_M \sum_{n=-\infty}^{\infty} g(n+k-j)g(n+k-i)g^{M-2}(n) \right] \quad \text{(see (1b) and (8))}. \tag{B.4}$$

Then, $\text{Cum}^{(M)}(d(n), x(n-i), x(n-k), \dots, x(n-k)) \neq 0$ if

$$L_1 \leqslant j \leqslant L_2, \quad l_1 \leqslant n+k-j \leqslant l_2, \quad l_1 \leqslant n+k-i \leqslant l_2, \quad l_1 \leqslant n \leqslant l_2$$

or

$$L_1 \leqslant j \leqslant L_2, \quad l_1 - n + j \leqslant k \leqslant l_2 - n + j, \quad l_1 - n + i \leqslant k \leqslant l_2 - n + i, \quad l_1 \leqslant n \leqslant l_2$$

or

$$L_1 \leqslant j \leqslant L_2, \quad \max\{l_1 - n + j, l_1 - n + i\} \leqslant k \leqslant \min\{l_2 - n + j, l_2 - n + i\}, \quad l_1 \leqslant n \leqslant l_2$$

or

$$L_1 \leqslant j \leqslant L_2, \quad \max\{l_1 - l_2 + j, l_1 - l_2 + i\} \leqslant k \leqslant \min\{l_2 - l_1 + j, l_2 - l_1 + i\}$$

or

$$\max\{l_1 - l_2 + L_1, l_1 - l_2 + i\} \leqslant k \leqslant \min\{l_2 - l_1 + L_2, l_2 - l_1 + i\}$$

or

$$\max\{-L + 1 + L_1, -L + 1 + i\} \leqslant k \leqslant \min\{L - 1 + L_2, L - 1 + i\}. \tag{B.5}$$

In other words, K_{dx1} and K_{dx2} (see (22b)) can be chosen such as

$$K_{dx1} \leqslant \max\{-L + 1 + L_1, -L + 1 + i\} \tag{B.6a}$$

and

$$K_{dx2} \geqslant \min\{L - 1 + L_2, L - 1 + i\}, \tag{B.6b}$$

respectively, and the associated $\hat{c}_{dx}(i)$ is a consistent estimate for $c_{dx}(i)$ when $\max\{-L + 1 + L_1, -L + 1 + i\} \leqslant \min\{L - 1 + L_2, L - 1 + i\}$. Again, the associated $\hat{c}_{dx}(i)$ does not need to be estimated since $c_{dx}(i) = 0$ when $\max\{-L + 1 + L_1, -L + 1 + i\} > \min\{L - 1 + L_2, L - 1 + i\}$ (see (B.5)). $\quad \square$

Appendix C. Proof of Eq. (25)

$$C_{xx}(\omega) = \sum_{i=-\infty}^{\infty} c_{xx}(i) e^{-j\omega i}$$

$$= \sum_{i=-\infty}^{\infty} \left[\sum_{k=-\infty}^{\infty} \text{Cum}^{(M)}(x(n), x(n-i), x(n-k), \dots, x(n-k)) \right] e^{-j\omega i} \quad \text{(see (16a))}$$

$$= \sum_{i=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} \left[\gamma_M \sum_{n=-\infty}^{\infty} g(n)g(n-i)g^{M-2}(n-k) \right] e^{-j\omega i} \quad \text{(see (A2), (1b) and (8))}$$

$$= \left[\gamma_M \sum_{k=-\infty}^{\infty} g^{M-2}(k) \right] \cdot \sum_{n=-\infty}^{\infty} g(n) \left[\sum_{i=-\infty}^{\infty} g(n-i) e^{-j\omega i} \right]$$

$$= \gamma_M G_{M-2}(0) \cdot \sum_{n=-\infty}^{\infty} g(n) \left[\sum_{l=-\infty}^{\infty} g(l) e^{-j\omega(n-l)} \right] \quad \text{(by letting } l = n - i)$$

$$= \gamma_M G_{M-2}(0) \cdot \left[\sum_{n=-\infty}^{\infty} g(n) e^{-j\omega n} \right] \cdot \left[\sum_{l=-\infty}^{\infty} g(l) e^{j\omega l} \right]$$

$$= \gamma_M G_{M-2}(0) \cdot G_1(\omega) \cdot G_1^*(\omega) \quad (\text{'*' denotes complex conjugation})$$

$$= \gamma_M G_{M-2}(0) \cdot |G_1(\omega)|^2. \quad \square$$

Appendix D. Proof of Theorem 3

Let

$$y_i(n) = x_f(n) * h_i(n) = u(n) * \tilde{h}_i(n) \tag{D.1}$$

for $i = 1, 2, \ldots, m$, where

$$\tilde{h}_i(n) \triangleq g(n) * h_i(n) \quad \text{(see (1b))} \tag{D.2}$$

D.1. Proof of Eq. (26)

$$\sum_{k_{m+1}=-\infty}^{\infty} \text{Cum}^{(M)}(y_1(n+k_1), y_2(n+k_2), \ldots, y_m(n+k_m), x_f(n+k_{m+1}), \ldots, x_f(n+k_{m+1}))$$

$$= \sum_{k_{m+1}=-\infty}^{\infty} \left[\gamma_M \sum_{n=-\infty}^{\infty} \tilde{h}_1(n+k_1) \tilde{h}_2(n+k_2) \cdots \tilde{h}_m(n+k_m) g^{M-m}(n+k_{m+1}) \right] \quad \text{(see (D.1), (1b) and (8))}$$

$$= \left[\gamma_M \sum_{k_{m+1}=-\infty}^{\infty} g^{M-m}(k_{m+1}) \right] \cdot \sum_{n=-\infty}^{\infty} \tilde{h}_1(n+k_1) \tilde{h}_2(n+k_2) \cdots \tilde{h}_m(n+k_m)$$

$$= \frac{\gamma_M G_{M-m}(0)}{\gamma_m} \cdot \left[\gamma_m \sum_{n=-\infty}^{\infty} \tilde{h}_1(n+k_1) \tilde{h}_2(n+k_2) \cdots \tilde{h}_m(n+k_m) \right]$$

$$= \frac{\gamma_M G_{M-m}(0)}{\gamma_m} \cdot \text{Cum}^{(m)}(y_1(n+k_1), y_2(n+k_2), \ldots, y_m(n+k_m)) \quad \text{(see (D.1) and (8))}.$$

Therefore, we have completed the proof of Eq. (26).

D.2. Proof of Eq. (27)

$$\sum_{k_{m+1}=-\infty}^{\infty} \cdots \sum_{k_M=-\infty}^{\infty} \text{Cum}^{(M)}(y_1(n+k_1), y_2(n+k_2), \ldots, y_m(n+k_m), x_f(n+k_{m+1}), \ldots, x_f(n+k_M))$$

$$\cdot \exp \left\{ -j \sum_{i=m+1}^{M} \omega_i k_i \right\}$$

$$= \sum_{k_{m+1}=-\infty}^{\infty} \cdots \sum_{k_M=-\infty}^{\infty} \left[\gamma_M \sum_{n=-\infty}^{\infty} \tilde{h}_1(n+k_1)\tilde{h}_2(n+k_2)\cdots\tilde{h}_m(n+k_m) \right.$$

$$\left. \cdot g(n+k_{m+1})\cdots g(n+k_M) \right] \cdot \exp\left\{ -j \sum_{i=m+1}^{M} \omega_i k_i \right\} \quad \text{(see (D.1), (1b) and (8))}$$

$$= \gamma_M \sum_{n=-\infty}^{\infty} \tilde{h}_1(n+k_1)\tilde{h}_2(n+k_2)\cdots\tilde{h}_m(n+k_m) \cdot \prod_{i=m+1}^{M} \left[\sum_{k_i=-\infty}^{\infty} g(n+k_i)e^{-j\omega_i k_i} \right]$$

$$= \gamma_M \sum_{n=-\infty}^{\infty} \tilde{h}_1(n+k_1)\tilde{h}_2(n+k_2)\cdots\tilde{h}_m(n+k_m) \cdot \prod_{i=m+1}^{M} G_1(\omega_i)e^{j\omega_i n}$$

$$= \gamma_M \sum_{n=-\infty}^{\infty} \tilde{h}_1(n+k_1)\tilde{h}_2(n+k_2)\cdots\tilde{h}_m(n+k_m) \cdot \prod_{i=m+1}^{M} G_1(\omega_i) \quad \left(\text{see } \sum_{i=m+1}^{M} \omega_i = 0 \right)$$

$$= \frac{\gamma_M \prod_{i=m+1}^{M} G_1(\omega_i)}{\gamma_m} \cdot \left[\gamma_m \sum_{n=-\infty}^{\infty} \tilde{h}_1(n+k_1)\tilde{h}_2(n+k_2)\cdots\tilde{h}_m(n+k_m) \right]$$

$$= \frac{\gamma_M \prod_{i=m+1}^{M} G_1(\omega_i)}{\gamma_m} \cdot \text{Cum}^{(m)}(y_1(n+k_1), y_2(n+k_2), \ldots, y_m(n+k_m)) \quad \text{(see (D.1) and (8))}. \quad \square$$

Acknowledgements

References

[1] G.C. Carter, "Coherence and time delay estimation", *Proc. IEEE*, Vol. 75, No. 2, February 1987, pp. 236–255.

[2] Y.T. Chan, J.M.F. Riley and J.B. Plant, "A parameter estimation approach to time-delay estimation and signal detection", *IEEE Trans. Acoust. Speech Signal Process.*, Vol. ASSP-28, No. 1, February 1980, pp. 8–16.

[3] Y.T. Chan, J.M.F. Riley and J.B. Plant, "Modeling of time delay and its application to estimation of nonstationary delays", *IEEE Trans. Acoust. Speech Signal Process.*, Vol. ASSP-29, No. 3, June 1981, pp. 577–581.

[4] C.-Y. Chi and W.-J. Chang, "New higher-order statistics based criterion for the design of linear prediction error filters", *Proc. IEEE Signal Processing Workshop on Higher-Order Statistics*, South Lake Tahoe, California, 7–9 June 1993, pp. 106–110.

[5] C.-Y. Chi, W.-J. Chang and C.-C. Feng, "A new algorithm for the design of linear prediction error filters using cumulant-based MSE criteria", *IEEE Trans. Signal Process.*, Vol. SP-42, No. 10, October 1994, pp. 2876–2880.

[6] A.N. Delopoulos and G.B. Giannakis, "Strongly consistent identification algorithms and noise insensitive MSE criteria", *IEEE Trans. Signal Process.*, Vol. SP-40, No. 8, August 1992, pp. 1955–1970.

[7] A.N. Delopoulos and G.B. Giannakis, "Consistent identification of stochastic linear systems with noisy input-output data", *Automatica*, Vol. 30, No. 8, August 1994, pp. 1271–1294.

[8] S. Haykin, *Modern Filters*, Macmillan, New York, 1989.

[9] S.M. Kay, *Modern Spectral Estimation*, Prentice-Hall, Englewood Cliffs, NJ, 1988.

[10] C.H. Knapp and G.C. Carter, "The generalized correlation method for estimation of time delay", *IEEE Trans. Acoust. Speech Signal Process.*, Vol. ASSP-24, August 1976, pp. 320–327.

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

[12] C.L. Nikias and F. Liu, "Bispectrum computation based on second- and third-order statistics with applications", *Proc. Internat. Conf. Acoust. Speech Signal Process. '90*, Albuquerque, New Mexico, April 1990, pp. 2381–2386.

[13] C.L. Nikias and R. Pan, "Time delay estimation in unknown Gaussian spatially correlated noise", *IEEE Trans. Acoust. Speech Signal Process.*, Vol. ASSP-36, No. 11, November 1988, pp. 1706–1714.

[14] C.L. Nikias and A.P. Petropulu, *Higher-Order Spectra Analysis*, Prentice-Hall, Englewood Cliffs, NJ, 1993.

[15] C.L. Nikias and M.R. Raghuveer, "Bispectrum estimation: A digital signal processing framework", *Proc. IEEE*, Vol. 75, No. 7, July 1987, pp. 869–891.

[16] C.W. Therrien, *Discrete Random Signals and Statistical Signal Processing*, Prentice-Hall, Englewood Cliffs, NJ, 1992.

[17] J.K. Tugnait, "On time delay estimation with unknown spatially correlated Gaussian noise using fourth-order cumulants and cross cumulants", *IEEE Trans. Signal Process.*, Vol. SP-39, No. 6, June 1991, pp. 1258–1267.