expansion (13) is able to approximate the Chebyshev function very well with only the first few terms.

V. Conclusion

This correspondence proposes an array pattern synthesis method together with the associated array structure for 3-D beamforming using CRA. The proposed method decomposes the array weights into two sets: weights for the elements within each ring and weights among different rings. The former are delay-and-sum weights, while the latter are from Fourier–Bessel series expansion. The proposed method yields a 3-D array pattern that is almost invariant over a certain frequency range and is therefore suitable for broadband beamforming. The frequency range to maintain invariant pattern is determined by the structure of the CRA. The proposed design is corroborated by simulations.

Acknowledgment

The authors would like to thank W. Larkin, J. Short, M. Wicks, and R. Bolia from the U.S. Air Force for providing technical insight and support for the project. The authors would also like to thank the reviewers for providing comments and suggestions, which have resulted in improving the quality of this correspondence.

References

[1] T. B. Vu, "Side-lobe control in circular ring array," *IEEE Trans. Antennas Propag.*, vol. 41, no. 8, pp. 1143–1145, Aug. 1993.

[2] R. Vescovo, "Constrained and unconstrained synthesis of array factor for circular arrays," *IEEE Trans. Antennas Propag.*, vol. 43, no. 12, pp. 1405–1410, Dec. 1995.

[3] B. K. Lau and Y. H. Leung, "A Dolph–Chebyshev approach to the synthesis of array patterns for uniform circular arrays," in *Proc. IEEE Int. Symp. Circuits and Systems (ISCAS)*, Geneva, Switzerland, May 2000, pp. I124–I127.

[4] S. C. Chan and C. K. S. Pun, "On the design of digital broadband beamformer for uniform circular array with frequency invariant characteristics," in *Proc. IEEE Int. Symp. Circuits and Systems (ISCAS)*, vol. 1, May 2002, pp. I-693–I-696.

[5] L. Biller and G. Friedman, "Optimization of radiation patterns for an array of concentric ring sources," *IEEE Trans. Audio Electroacoust.*, vol. AU-21, no. 1, pp. 57–61, Feb. 1973.

[6] D. Huebner, "Design and optimization of small concentric ring arrays," in *Proc. IEEE Antennas Propagation Int. Symp.*, vol. 16, May 1978, pp. 455–458.

[7] B. P. Kumar and G. R. Branner, "Design of low sidelobe circular ring array by element radius optimazation," in *Proc. IEEE Antennas Propagation Int. Symp.*, Orlando, FL, Jul. 1999, pp. 2032–2035.

[8] C. O. Stearns and A. C. Stewart, "An investigation of concentric ring antennas with low sidelobes," *IEEE Trans. Antennas Propag.*, vol. AP-13, no. 6, pp. 856–863, Nov. 1965.

[9] T.-S. Lee, "Efficient wideband source localization using beamforming invariance technique," *IEEE Trans. Signal Process.*, vol. 42, no. 6, pp. 1376–1387, Jun. 1994.

[10] D. B. Ward, Z. Ding, and R. A. Kennedy, "Broadband DOA estimation using frequency invariant beamforming," *IEEE Trans. Signal Process.*, vol. 46, no. 5, pp. 1463–1469, May 1998.

[11] D. B. Ward and T. D. Abhayapala, "Range and bearing estimation of wideband sources using an orthogonal beamspace processing structure," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing (ICASSP)*, Montreal, QC, Canada, May 2004, pp. II-109–II-112.

[12] K. Nishikawa, T. Yamamoto, K. Oto, and T. Kanamori, "Wideband beamforming using fan filter," in *Proc. IEEE Int. Symp. Circuits and Systems (ISCAS)*, May 1992, pp. 533–536.

[13] D. B. Ward, R. A. Kennedy, and R. C. Williamson, "FIR filter design for frequency invariant beamformers," *IEEE Signal Process. Lett.*, vol. 3, no. 3, pp. 69–71, Mar. 1996.

[14] T. Sekiguchi and Y. Karasawa, "Wideband beamspace adaptive array utilizing FIR fan filters for multibeam forming," *IEEE Trans. Signal Process.*, vol. 48, no. 1, pp. 277–284, Jan. 2000.

[15] D. B. Ward, R. A. Kennedy, and R. C. Williamson, "Theory and design of broadband sensor arrays with frequency invariant far-field beam patterns," *J. Acoust. Soc. Amer.*, vol. 97, no. 2, pp. 1023–1034, Feb. 1995.

[16] Y. Li, K. C. Ho, and C. Kwan, "Design of broad-band circular ring microphone array for speech acquisition in 3-D," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing (ICASSP)*, Hong Kong, China, Apr. 2003, pp. 221–224.

[17] R. E. Collin and F. J. Zucker, *Antenna Theory: Part I*. New York: McGraw-Hill, 1969.

[18] E. Kreyszig, *Advanced Engineering Mathematics*, 7th ed. New York: Wiley, 1993.

[19] C. Kwan *et al.*, "An automated acoustic system to monitor and classify birds," presented at the Bird Monitoring Workshop, Toronto, ON, Canada, Aug. 2003.

A Cumulant Interference Subspace Cancellation Method for Blind SISO Channel Estimation

Jun Fang, A. Rahim Leyman, *Member, IEEE*, Yong Huat Chew, and Ying Chang Liang

***Abstract*—We present a new linear method for blind estimation of single-input single-output (SISO) finite impulse response (FIR) channel. The channel can be minimum-phase or nonminimum-phase channel. The proposed method is based on a series of fourth order cumulant matrices, where it is shown that by employing vectors chosen from the left null space of a concatenated cumulant matrix, the interference subspace of the channel convolution matrix can be cancelled and thus, channel information can be extracted. The proposed method is robust to channel order overestimation, and it has a similar computational complexity as other existing methods. Simulation results are included to validate the performance of the proposed algorithm.**

***Index Terms*—Blind channel estimation, cumulant, higher-order statistics, single-input single-output systems (SISO).**

I. Introduction

Blind identification of time-invariant FIR systems arises in a wide variety of communication and signal processing applications. Thus far, there have been a lot of research works [1]–[4] on blind channel estimation by using the second-order statistics (SOS). In this case, channel diversity should be obtained by oversampling the output data or by resorting to multiple sensors. Also, to identify this multichannel system, most SOS methods are required to meet a fundamental blind identifiability condition [2] that all subchannels do not share any common zeros, which may not be satisfied in practice [5]. In contrast, higher-order statistics (HOS) methods could be applicable in such a scenario where channel diversity is not available, i.e., channel is SISO, or blind identifiability condition for SOS methods is close to be violated.

Manuscript received August 26, 2004; revised April 13, 2005. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Franz Hlawatsch.

J. Fang is with the Department of Electrical and Computer Engineering, National University of Singapore, Singapore, 119260 (e-mail: g0202082@nus.edu.sg).

A. Rahim Leyman, Y. Huat Chew, and Y. Chang Liang are with the Institute for Infocomm Research, A*STAR, Singapore, 119613 (e-mail: larahim@i2r.a-star.edu.sg; chewyh@i2r.a-star.edu.sg; ycliang@i2r.a-star.edu.sg).

Digital Object Identifier 10.1109/TSP.2005.861882

Numerous linear HOS methods [6]–[12] have been proposed over the past decade. Among them, some [6]–[9] can be considered as *subspace* methods because by using a set of cumulants, they exploit the special linear algebraic structure of the correspondingly constructed channel matrix. Compared to other methods, these *subspace* methods achieve better performance with lesser data samples. However, with the exception of the *Weighted Slices* (WS) algorithm proposed in [6], they usually require the precise knowledge of channel order and are sensitive to channel order overestimation. It is noted that a systematic generalization of [6] has been proposed in [13]. Compared to [6], [13] presents an enhanced way in exploiting the special linear algebraic (Toeplitz) structure. Recently, a matrix pencil technique [14] was adopted to estimate the channel by utilizing a series of fourth order cumulant matrices. As compared to [6]–[9], the work [14] shows an improvement in that it exploits the inherent structure relationship between a pair of constructed cumulant matrices rather than only the linear algebraic structure of the channel matrix, thus demonstrating an enhanced performance and robustness to channel order overestimation.

In this correspondence, we propose a new linear method that extracts the channel information by utilizing the derived so-called interference subspace cancellation vectors. It is noted that an implicit connection exists between our work and [14] because the nontrivial generalized eigenvectors derived in [14] can also be deemed as the interference subspace cancellation vectors. Compared to [14], we take a more direct approach to compute these information-extraction vectors. Precisely, we devise our algorithm by exploiting the *partial column space overlapping* relationship between a concatenated cumulant matrix and a target cumulant matrix to obtain these information-extraction vectors. This technique is essentially different from [14] that devises its algorithm by investigating the inherent structure of the constructed matrix pencil. As a consequence, our work induces the following advantages. First, the algorithm in [14] requires a channel identifiability condition to make sure that the constructed matrix pencil has at least one unique nontrivial generalized eigenvalue. This identifiability condition is no longer necessary for our proposed algorithm. Second, the selection of interference subspace cancellation vectors can be excluded from our algorithm without detrimental effects. However, this selection procedure is necessary in [14] in order to distinguish the trivial from the nontrivial vectors.

II. Preliminaries

A. System Model

We consider a SISO linear time-invariant FIR system given as

$$x_n \triangleq h_n * s_n + w_n \triangleq \sum_{k=0}^{L} h_k s_{n-k} + w_n \tag{1}$$

where L denotes the channel order. The above model can be easily transformed as follows:

$$\vec{\mathbf{x}}(n) = \mathbf{H}\vec{\mathbf{s}}(n) + \vec{\mathbf{w}}(n) \tag{2}$$

if we define $\vec{\mathbf{x}}(n) \triangleq [x_n \;\; x_{n-1} \;\; \ldots \;\; x_{n-N}]^T, \vec{\mathbf{s}}(n) \triangleq [s_n \;\; s_{n-1} \;\; \ldots \;\; s_{n-N-L}]^T, \vec{\mathbf{w}}(n) \triangleq [w_n \;\; w_{n-1} \;\; \ldots \;\; w_{n-N}]^T$, and the channel convolution matrix $\mathbf{H}$ is an $(N+1) \times (N+L+1)$ Toeplitz matrix written as

$$\mathbf{H} \triangleq \begin{bmatrix} h_0 & \cdots & h_L & 0 & \ldots & 0 \\ 0 & h_0 & \cdots & h_L & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & h_0 & \ldots & h_L \end{bmatrix}. \tag{3}$$

Throughout this correspondence, we define the following notations. If $\mathbf{A}$ is an $m \times n$ matrix, then $\mathbf{A}^T$ represents matrix transpose, $\mathbf{A}^H$ – matrix conjugate transpose, $\mathbf{A}^*$ – matrix conjugate, $\mathbf{A}^\dagger$-matrix pseudoinverse. $\mathcal{R}(\mathbf{A})$ is the range (column) space of matrix $\mathbf{A}$. Let $\mathbf{H}_{[k_1:k_2]} \triangleq [\mathbf{H}_{k_1} \;\; \ldots \;\; \mathbf{H}_{k_2}]$ denotes the part of $\mathbf{H}$ from its k_1th column to its k_2th column, in which $\mathbf{H}_i$ denotes the ith column of $\mathbf{H}$. The system model assumptions are: (A1) The input signal is independent and identically distributed non-Gaussian stationary process with zero mean and nonzero kurtosis $\gamma \triangleq (\mu_4/\mu_2^2)$, where μ_i denotes the ith central moment. (**A2**) Additive noise is a zero-mean Gaussian process, and is statistically independent of the input signal. (**A3**) The channel order is L. Without loss of generality, we assume $h_0 \neq 0$ and $h_L \neq 0$. Our objective is to estimate the channel impulse response $\mathbf{h}(z)$ by utilizing the fourth-order statistics of the observed data $\vec{\mathbf{x}}(n)$.

B. Cumulant Matrices

We define a series of fourth order cumulant matrices [13]–[14] of the channel output signals $\mathbf{C}[k]$ as

$$\mathbf{C}[k] \triangleq \mathrm{cum}(\vec{\mathbf{x}}(n), \vec{\mathbf{x}}(n)^H, x_{n-k}, x_{n-k}^*) \tag{4}$$

where $\mathbf{C}[k]$ is an $(N+1) \times (N+1)$ matrix with its ijth element defined as $\mathbf{C}^{ij}[k] \triangleq \mathrm{cum}(x_{n-i+1}, x_{n-j+1}^*, x_{n-k}, x_{n-k}^*)$. Invoking the cumulant properties [15] and assumptions (A1)–(A2), we have

$$\mathbf{C}[k] = \gamma \mathbf{H}\Lambda[k]\mathbf{H}^H \tag{5}$$

$$\Lambda[k] \triangleq \mathrm{diag}(\underbrace{0, \ldots, 0}_{k}, |h_0|^2, \ldots, |h_L|^2, \underbrace{0, \ldots, 0}_{(N-k)}). \tag{6}$$

Equation (5) can be further rewritten as

$$\mathbf{C}[k] = \gamma \mathbf{H}_{[k+1:k+L+1]} \bar{\Lambda} \mathbf{H}_{[k+1:k+L+1]}^H \tag{7}$$

where $\bar{\Lambda} \triangleq \mathrm{diag}(|h_0|^2, \ldots, |h_L|^2)$. In the following section, we will show that the channel information can be extracted based on a series of cumulant matrices $\mathbf{C}[k]$.

III. Channel Identification

A. Principle for Channel Identification

We consider a series of cumulant matrices $\mathbf{C}[k]$ with consecutive *delay lags* k, where $L \le k \le 2L$ and $N \ge 2L$. From (7), we have

$$\begin{aligned} \mathbf{C}[L] &= \gamma \mathbf{H}_{[L+1:2L+1]} \bar{\Lambda} \mathbf{H}_{[L+1:2L+1]}^H \\ &= \gamma |h_0|^2 \mathbf{H}_{[L+1]} \mathbf{H}_{[L+1]}^H \\ &\quad + \gamma \mathbf{H}_{[L+2:2L+1]} \mathbf{D} \mathbf{H}_{[L+2:2L+1]}^H \end{aligned} \tag{8}$$

where $\mathbf{D} \triangleq \mathrm{diag}(|h_1|^2, \ldots, |h_L|^2)$. Note that the column $\mathbf{H}_{[L+1]}$ is exactly an augmented channel vector with the desired channel vector $\mathbf{h} \triangleq [h_L \;\; h_{L-1} \;\; \ldots \;\; h_0]^T$ padded with zero entries. Thus, the column space of $\mathbf{H}_{[L+1:2L+1]}$ is constructed by a rank-1 signal subspace spanned by $\mathbf{H}_{[L+1]}$ and a rank-L interference subspace spanned by columns of $\mathbf{H}_{[L+2:2L+1]}$. Our objective here is to find an *interference subspace cancellation vector*, $\mathbf{v}_c$, which is orthogonal to the interference subspace $\mathcal{R}(\mathbf{H}_{[L+2:2L+1]})$. Such a vector can extract the signal subspace as follows:

$$\begin{aligned} \mathbf{v}_c^H \mathbf{C}[L] &= \gamma \mathbf{v}_c^H \mathbf{H}_{[L+1:2L+1]} \bar{\Lambda} \mathbf{H}_{[L+1:2L+1]}^H \\ &= \gamma |h_0|^2 \mathbf{v}_c^H \mathbf{H}_{[L+1]} \mathbf{H}_{[L+1]}^H \\ &= \alpha \mathbf{H}_{[L+1]}^H \end{aligned} \tag{9}$$

where $\alpha \triangleq \gamma |h_0|^2 \mathbf{v}_c^H \mathbf{H}_{[L+1]}$ is a complex scalar. Thus, the channel is identified up to an unknown complex scalar α. In order to find such an interference subspace cancellation vector, the interference subspace

$\mathcal{R}(\mathbf{H}_{[L+2:2L+1]})$ should be determined. Since the channel impulse response may contain zero coefficients, it is very hard for us to get an exact $\mathcal{R}(\mathbf{H}_{[L+2:2L+1]})$. However, a subspace which includes the interference subspace can be obtained by the method described in the following theorem.

Theorem 1: If we concatenate a series of cumulant matrices $\mathbf{C}[k]$ with consecutive delay lags to construct a new concatenated cumulant matrix

$$\mathbf{S} \triangleq [\mathbf{C}[k_1] \quad \mathbf{C}[k_1+1] \quad \cdots \quad \mathbf{C}[k_2]] \tag{10}$$

where $k_1 = L+1, k_2 = 2L$, then we have

$$\mathcal{R}(\mathbf{H}_{[L+2:3L+1]}) \supseteq \mathcal{R}(\mathbf{S}) \supseteq \mathcal{R}(\mathbf{H}_{[L+2:2L+1]}). \tag{11}$$

Proof: Since $\mathbf{S}$ is a concatenation of a series of cumulant matrices, it is clear that

$$\mathcal{R}(\mathbf{S}) = \mathcal{R}(\mathbf{C}[k_1]) \cup \mathcal{R}(\mathbf{C}[k_1+1]) \cup \cdots \cup \mathcal{R}(\mathbf{C}[k_2]) \tag{12}$$

where the subspace $\mathcal{R}(\mathbf{C}[k])$ is spanned by columns of $\mathbf{H}$ whose corresponding diagonal elements in $\Lambda[k]$ are nonzero. Since $h_0 \neq 0$, it is obvious that

$$\mathcal{R}(\mathbf{H}_{[k+1:k+1+L]}) \supseteq \mathcal{R}(\mathbf{C}[k]) \supseteq \mathcal{R}(\mathbf{H}_{[k+1]}). \tag{13}$$

We can easily get the results in (11) by combining (12) and (13), given that $k_1 = L+1, k_2 = 2L$. The proof is completed here. ∎

On one hand, Theorem 1 shows that the interference subspace is included in $\mathcal{R}(\mathbf{S})$; on the other hand, it indicates that $\mathrm{rank}(\mathbf{S}) \leq 2L$, which guarantees that $\mathbf{S}$ must have a nondegenerate left null space (i.e., dimension ≥ 1) when $N \geq 2L$. Here we use $\mathcal{R}(\mathbf{S}_l^{\perp})$ to denote the left null space of $\mathbf{S}$. It is clear that any vector that belongs to $\mathcal{R}(\mathbf{S}_l^{\perp})$ is orthogonal to the interference subspace. However, to be an interference subspace cancellation vector, not only the vector has to be orthogonal to the interference subspace, a hidden condition is that the vector should not be orthogonal to the signal subspace, otherwise α would be zero. Hence we are faced with a question: whether or not there exists a vector that belongs to $\mathcal{R}(\mathbf{S}_l^{\perp})$ satisfying this hidden condition. This problem is answered by the following theorem.

Theorem 2: Suppose $\mathbf{V} = [\mathbf{v}_1 \ \mathbf{v}_2 \ \ldots \ \mathbf{v}_p]$ is a basis for $\mathcal{R}(\mathbf{S}_l^{\perp})$, where p is the dimension of $\mathcal{R}(\mathbf{S}_l^{\perp})$, then we have $\mathbf{V}^H\mathbf{H}_{[L+1]} \neq \mathbf{0}$, which means that there exists at least one column vector in $\mathbf{V}$ that satisfies

$$\mathbf{v}_i^H\mathbf{H}_{[L+1]} \neq 0 \quad \exists\, 1 \leq i \leq p. \tag{14}$$

Proof: Since $\mathbf{V}$ is a basis for the orthogonal complement of $\mathcal{R}(\mathbf{S})$, it is obvious that, for any vector $\mathbf{g}, \mathbf{g}^H\mathbf{V} = 0$ if and only if $\mathbf{g}$ belongs to the left null space of $\mathbf{V}$, i.e., $\mathcal{R}(\mathbf{S})$. Therefore, to prove $\mathbf{V}^H\mathbf{H}_{[L+1]} \neq \mathbf{0}$, we only need to prove that $\mathbf{H}_{[L+1]} \nsubseteq \mathcal{R}(\mathbf{S})$. Note that $\mathbf{H}_{[L+1:N+L+1]}$ is a $(N+1)\times(N+1)$ lower triangular matrix with its main diagonal entries equal to h_L and ith sub-diagonal entries below the main diagonal equal to h_{L-i}. This structure guarantees that all columns except all-zero columns of $\mathbf{H}_{[L+1:N+L+1]}$ are linearly independent, i.e., $\mathbf{H}_{[L+1:N+L+1]}$ is full column rank after deleting all-zero columns. Therefore, $\mathbf{H}_{[L+1:3L+1]}$ is also full column rank after deleting all-zero columns. Thus, we have

$$\begin{aligned}\mathbf{H}_{[L+1]} \nsubseteq \mathcal{R}(\mathbf{H}_{[L+2:3L+1]}) &\Rightarrow \mathbf{H}_{[L+1]} \nsubseteq \mathcal{R}(\mathbf{S}) \\ &\Rightarrow \mathbf{V}^H\mathbf{H}_{[L+1]} \neq \mathbf{0}.\end{aligned} \tag{15}$$

The proof is completed here. ∎

Theorem 2 indicates that there exists at least one interference subspace cancellation vector from the basis for $\mathcal{R}(\mathbf{S}_l^{\perp})$. Such an interference subspace cancellation vector obtained from $\mathcal{R}(\mathbf{S}_l^{\perp})$ can help to extract the channel information through (9). In fact, it is noted that, in practice, the computed basis for $\mathcal{R}(\mathbf{S}_l^{\perp})$ can always render us *more than one* interference subspace cancellation vector. Since the channel is estimated for each interference subspace cancellation vector, these multiple interference subspace cancellation vectors obtained from the basis for $\mathcal{R}(\mathbf{S}_l^{\perp})$ provide us with an estimation diversity that can be utilized to enhance our final estimation accuracy. For simplicity, here we assume that all these p column vectors $\mathbf{v}_i$ $(1 \leq i \leq p)$ can be taken as the interference subspace cancellation vectors. This assumption has been verified by our numerous simulations. Also, such an assumption can be justified in the following sense, that is, even if there exist some column vectors in $\mathbf{V}$ satisfying $\mathbf{v}_i^H\mathbf{H}_{[L+1]} = 0$, our proposed method can still work without any detrimental effects on the channel estimation. The reason is explained as follows. Without loss of generality, suppose that we have $\mathbf{v}_i^H\mathbf{H}_{[L+1]} \neq 0$ for $i \in \{1,\ldots,m\}$ and $\mathbf{v}_i^H\mathbf{H}_{[L+1]} = 0$ for $i \in \{m+1,\ldots,p\}$. Thus the estimated channel from the former vectors $\mathbf{v}_i$ for $i \in \{1,\ldots,m\}$ is $\mathbf{v}_i^H\mathbf{C}[L] = \lambda_i\mathbf{H}_{[L+1]}^H$ and the estimated channel from the latter vectors $\mathbf{v}_i$ for $i \in \{m+1,\ldots,p\}$ is $\mathbf{v}_i^H\mathbf{C}[L] = \mathbf{0}$. Since our final channel estimate is obtained by integrating the channel information estimated from every vector $\mathbf{v}_i$ for $i \in \{1,\ldots,p\}$, theoretically, these all-zero vectors estimated from $\mathbf{v}_i$ for $i \in \{m+1,\ldots,p\}$ have no impact on the final estimation result.

B. Practical Analysis of Channel Identification

We study our channel identification method under the following two practical scenarios.

1) Channel With Small Head Taps: We study how and to what extent the proposed method is influenced when the multipath channel has small head taps. For simplicity, we assume that the multipath channel has only one small head tap, i.e., $|h_0|$ is small and $|h_0| \ll |h_1|$. Recalling (7), we have

$$\begin{aligned}\mathbf{C}[k] = \gamma|h_0|^2\mathbf{H}_{[k+1]}\mathbf{H}_{[k+1]}^H \\ + \gamma\mathbf{H}_{[k+2:k+L+1]}\mathbf{D}\mathbf{H}_{[k+2:k+L+1]}^H.\end{aligned} \tag{16}$$

It is clear that a small $|h_0|$ leads to a much smaller $\gamma|h_0|^2\mathbf{H}_{[k+1]}\mathbf{H}_{[k+1]}^H$, thus $\mathbf{H}_{[k+1]}$ has a negligible contribution in spanning the column space of $\mathbf{C}[k]$. Therefore, (13) should be modified as

$$\mathcal{R}(\mathbf{H}_{[k+2:k+1+L]}) \supseteq \mathcal{R}(\mathbf{C}[k]) \supseteq \mathcal{R}(\mathbf{H}_{[k+2]}) \tag{17}$$

and, accordingly, (11) should be modified as

$$\mathcal{R}(\mathbf{H}_{[L+3:3L+1]}) \supseteq \mathcal{R}(\mathbf{S}) \supseteq \mathcal{R}(\mathbf{H}_{[L+3:2L+2]}). \tag{18}$$

It can be seen that here $\mathcal{R}(\mathbf{S})$ only includes partial subspace of the interference subspace. Hence the interference subspace cancellation vectors obtained from $\mathcal{R}(\mathbf{S}_l^{\perp})$ are only orthogonal to partial subspace of the interference subspace $\mathcal{R}(\mathbf{H}_{[L+2:2L+1]})$. Equation (9) should be rewritten as follows:

$$\begin{aligned}\mathbf{v}_c^H\mathbf{C}[L] &= \gamma\mathbf{v}_c^H\mathbf{H}_{[L+1:2L+1]}\bar{\Lambda}\mathbf{H}_{[L+1:2L+1]}^H \\ &= \gamma|h_0|^2\mathbf{v}_c^H\mathbf{H}_{[L+1]}\mathbf{H}_{[L+1]}^H \\ &\quad + \gamma|h_1|^2\mathbf{v}_c^H\mathbf{H}_{[L+2]}\mathbf{H}_{[L+2]}^H \\ &\approx \alpha_1\mathbf{H}_{[L+2]}^H\end{aligned} \tag{19}$$

where $\alpha_1 \triangleq \gamma|h_1|^2\mathbf{v}_c^H\mathbf{H}_{[L+2]}$, the term $\gamma|h_0|^2\mathbf{v}_c^H\mathbf{H}_{[L+1]}\mathbf{H}_{[L+1]}^H$ can be omitted since $|h_0|^2$ is much smaller as compared to $|h_1|^2$. Thus the column $\mathbf{H}_{[L+2]}$ is extracted. It is noted that $\mathbf{H}_{[L+2]}$ is still an augmented channel vector surrounded by zero entries. As a generalization, we can conclude that the column $\mathbf{H}_{[L+k+1]}$ could be extracted if the channel has k small head taps, where $k < L$. Note that for every $k < L$, we can guarantee that the column $\mathbf{H}_{[L+k+1]}$ is an augmented channel vector which contains the complete channel information by choosing $N \geq 2L$. Therefore, we can see that even if the multipath channel has small head taps, it will not have a detrimental effect to our proposed method, and we can still estimate the channel vector up to a delay ambiguity.[1]

2) Channel Order Overestimated: In practice, it is almost impossible for us to obtain a precise channel order due to noise and estimation errors. Therefore it is very meaningful to investigate the robustness of our proposed algorithm to channel order overestimation. When channel order is overestimated, (13) should be revised as (it should be noted that here we do not consider small head taps)

$$\mathcal{R}(\mathbf{H}_{[k+1:k+1+L_e]}) \supset \mathcal{R}(\mathbf{C}[k]) \supseteq \mathcal{R}(\mathbf{H}_{[k+1]}) \tag{20}$$

where L_e is the overestimated channel order and $L_e > L$. Accordingly, (11) should be revised as

$$\mathcal{R}(\mathbf{H}_{[L_e+2:3L_e+1]}) \supset \mathcal{R}(\mathbf{S}) \supseteq \mathcal{R}(\mathbf{H}_{[L_e+2:2L_e+1]}). \tag{21}$$

Since $\mathbf{H}_{[L_e+1:3L_e+1]}$ is also full column rank after deleting all-zero columns when channel order is overestimated, Theorem 2 remains true. Therefore, (9) still holds and can be rewritten as

$$\begin{aligned}\mathbf{v}_c^H\mathbf{C}[L_e] &= \gamma\mathbf{v}_c^H\mathbf{H}_{[L_e+1:2L_e+1]}\bar{\Lambda}_{[L_e]}\mathbf{H}_{[L_e+1:2L_e+1]}^H\\ &= \gamma|h_0|^2\mathbf{v}_c^H\mathbf{H}_{[L_e+1]}\mathbf{H}_{[L_e+1]}^H\\ &= \alpha_2\mathbf{H}_{[L_e+1]}^H\end{aligned} \tag{22}$$

where $\alpha_2 \triangleq \gamma|h_0|^2\mathbf{v}_c^H\mathbf{H}_{[L_e+1]}$ and $\bar{\Lambda}_{[L_e]} \triangleq \mathrm{diag}(|h_0|^2,\ldots,|h_{L_e}|^2)$. Thus the column $\mathbf{H}_{[L_e+1]}$ is extracted and the estimated channel vector is obtained by taking the first to (L_e+1)th entries out from $\mathbf{H}_{[L_e+1]}$. We can see that the overestimated channel taps h_l $(L_e \geq l > L)$ should be zero. It means that, theoretically, channel order overestimation has no effect on our proposed method.

IV. Algorithm Development

Following the analysis, we now develop a practical algorithm for channel identification. Theoretically, the interference subspace cancellation vectors can be chosen to be the left singular vectors associated with the p smallest singular value of $\mathbf{S}$, where $p = N+1-\mathrm{rank}(\mathbf{S})$. Because of the finite sample size, the estimate $\hat{\mathbf{S}}$ would not be rank deficient in practice and in order to determine p, the rank of $\hat{\mathbf{S}}$ need to be estimated. Nevertheless, the determination of rank is always a tricky problem, especially for estimated cumulant matrix. Therefore, it is better for us to find a simple way to go around this problem. Notice that we have $L \leq \mathrm{rank}(\mathbf{S}) \leq 2L$ from (11); this implies that the number of interference subspace cancellation vectors is upper-bounded and lower-bounded by $p_u = N+1-L$ and $p_l = N+1-2L$ respectively. Since every interference subspace cancellation vector provides us with an estimated channel, we can only choose p_l interference subspace cancellation vectors from p_u candidate vectors which are the left singular vectors associated with the p_u smallest singular values of $\hat{\mathbf{S}}$. Of course, a simpler alternative is to choose the left singular vectors associated with the p_l smallest singular values of $\hat{\mathbf{S}}$ as the p_l interference subspace cancellation vectors, at the expense of mild performance degradation. For comparison purpose, the former which is more accurate is used in our work. Here we assume that the channel order L is known *a priori*. In practice, even if the channel order is overestimated, we can still determine $p_u = N+1-L_e$ and $p_l = N+1-2L_e$ since under this case, we have $L_e \leq \mathrm{rank}(\mathbf{S}) < 2L_e$ [see (21)].

Until now, we have successfully circumvented the rank determination problem by choosing p_l qualified vectors from p_u candidate vectors. It is clear that these p_u candidate vectors are not equivalent as they achieve different interference subspace cancellation effects. Hence, there are two problems faced by us. First, how to choose p_l qualified vectors from p_u candidate vectors, i.e., interference subspace cancellation vectors selection. Second, how to integrate the estimated channel information obtained from these p_l interference subspace cancellation vectors. We now enumerate the steps for our channel identification procedure.

1) Given the estimated channel order L, let $N \geq 2L$, compute a series of estimated fourth order cumulant matrices $\hat{\mathbf{C}}[k]$, where $L \leq k \leq 2L$, from the channel output samples.
2) Concatenate a series of $\hat{\mathbf{C}}[k]$ to construct a new cumulant matrix $\hat{\mathbf{S}}$ as given in (10).
3) Compute the SVD of $\hat{\mathbf{S}}$. Choose p_u left singular vectors, $\hat{\mathbf{v}}_1, \hat{\mathbf{v}}_2, \ldots, \hat{\mathbf{v}}_{p_u}$, associated with the p_u smallest singular values of $\hat{\mathbf{S}}$.
4) *Interference subspace cancellation vectors selection*: For each $\hat{\mathbf{v}}_i$, the estimated augmented channel vector can be computed as $\underline{\hat{\mathbf{h}}}_i = \hat{\mathbf{C}}[L]^H\hat{\mathbf{v}}_i$ for each $i \in \{1,\ldots,p_u\}$. And the ith estimated channel vector $\hat{\mathbf{h}}_i$ can be obtained by deleting the zero entries in $\underline{\hat{\mathbf{h}}}_i$. However, because of the possible delay ambiguity introduced by small head taps (see Section III-B), $\hat{\mathbf{h}}_i$ should be chosen from $\{\hat{\mathbf{h}}_i^j\}$, where $j \in \{1,\ldots,L\}$ and $\hat{\mathbf{h}}_i^j$ is obtained by taking the jth to $(j+L)$th entries out of $\underline{\hat{\mathbf{h}}}_i$. For each $\hat{\mathbf{h}}_i^j$, we compare the theoretical $\mathbf{C}[k]$ which is computed by using the estimated channel to $\hat{\mathbf{C}}[k]$, i.e., the estimated cumulant matrix. The distance between the theoretical $\mathbf{C}[k]$ and the estimated $\hat{\mathbf{C}}[k]$ is defined as
$$\mathrm{dis} \triangleq \min_{\beta} \|\hat{\mathbf{C}}[k] - \beta\mathbf{C}[k]\|_F^2 \tag{23}$$
where $\|\cdot\|_F$ stands for the Frobenius matrix norm, β is a scalar chosen to minimize the matrix norm. Thus, we can obtain the resulted distance, denoted by $\mathrm{dis}(i,j)$, for each estimated channel vector $\hat{\mathbf{h}}_i^j$. Finally, from the computed $\mathrm{dis}(i,j)$ for $i \in \{1,\ldots,p_u\}, j \in \{1,\ldots,L\}$, we select the best p_l vectors from $\hat{\mathbf{v}}_i, i \in \{1,\ldots,p_u\}$, as the interference subspace cancellation vectors. The criterion for choosing these p_l vectors is as follows. Let $d_i \triangleq \min\{\mathrm{dis}(i,j)\}$ for $j \in \{1,\ldots,L\}$. Then we choose the vector $\hat{\mathbf{v}}_i$ as the interference subspace cancellation vectors if d_i is among the first p_l minimum values of $\{d_1,\ldots,d_{p_u}\}$.
5) Channel information integration: Given the selected p_l vectors from above, we have p_l corresponding estimated augmented channel vectors $\{\underline{\hat{\mathbf{h}}}_i\}$. We next integrate the channel information from these multiple estimated results. This step is similar to that in [14] and thus we describe it briefly as follows. First, we select a reference vector $\underline{\hat{\mathbf{h}}}_{i_r}$ by the following criterion
$$i_r = \arg\min_{i,j} |\mathrm{dis}(i,j)|. \tag{24}$$
Given the p_l estimated augmented channel vectors, estimate delay difference $\{\tau_i\}$ relative to the selected reference vector (the estimation of the relative delay difference can be found in the counterpart of [14]), and obtain the aligned vectors $\{\underline{\hat{\mathbf{h}}}_i^{(\tau_i)}\}$ with the same delay ambiguity. Concatenate all aligned vectors $\{\underline{\hat{\mathbf{h}}}_i^{(\tau_i)}\}$ and compute the SVD of the concatenated matrix. The ultimate estimation $\underline{\hat{\mathbf{h}}}$ is obtained as the left singular vector associated with the largest singular value of the concatenated matrix.

Finally, we compare the computational complexity of our proposed method to the other existing linear methods [14], [6]. We only consider

[1]Since, in practice, we do not know how many small head taps exist in the multipath channel, thus we also do not know how many zero entries are padded in the forepart of the estimated augmented channel vector. This can be considered as a delay ambiguity.

the linear algebraic operations involved in the algorithm implementation. It can be seen from previous part that our proposed algorithm requires to do SVD operation in step 3 and step 5, respectively. The dimension of the computed matrices are as follows

$$\begin{array}{ll}\text{Step 3} & (N+1)\times(N+1)L\\ \text{Step 5} & (N+1)\times p_l\end{array}$$

where we adopt $N = 2L+1$ in our simulations. Thus in step 3, we have to compute the SVD of a $(2L+2)\times(2L+2)L$ matrix. However, it is noted that we only need to compute the left singular vectors of the matrix $\mathbf{S}$. This is equivalent to computing the right singular vectors of the tall matrix $\mathbf{S}^H$ with dimension $(2L+2)L\times(2L+2)$. From [16] [p. 254], we know that this computation requires $2mn^2+11n^3$ flops, where $m=(2L+2)L$ and $n=(2L+2)$. Hence the total flops required for our proposed algorithm are $(2L+11)(2L+2)^3$ and of order $O(L^4)$. In the case where the channel order L is not very large and smaller than order of tens, our proposed algorithm has a similar computational complexity as the algorithm [14].[2] Also, the computational complexity of our algorithm is less than that of the WS algorithm [6] since the latter involves computing the pseudoinverse of a $(2L+1)\times(3L+1)(L+1)$ matrix and requires about $O(L^5)$ flops.

V. Simulation Results

Here, we present simulation results to illustrate the performance of our algorithm. We compare our method, namely Cumulant Interference Subspace Cancellation (CISC) algorithm, to the other two linear methods, Cumulant Weighted Overlapping Matrix Pencil algorithm (WOMP-SVD) presented in [14] and WS algorithm proposed in [6]. For comparison purposes, we will only use the same set of fourth-order cumulants as CISC and WOMP-SVD for WS. In the implementation of CISC algorithm, we choose $N = 2L+1$, also for simplicity, let $k = L$ in (23) when computing the interference subspace cancellation vectors selection criterion. In our simulations, channel outputs are added with complex white Gaussian noise with zero-mean. The performance is measured by the normalized mean square error (NMSE) of the channel estimate, which is obtained by finding the complex scalar ρ that minimizes $(\|\mathbf{h}-\rho\hat{\mathbf{h}}\|^2/\|\mathbf{h}\|^2)$, and the symbol error rate (SER) of the estimated data symbols.

A. Example A

To study the robustness of the proposed algorithm to various channel conditions, we conduct simulation tests using randomly generated wireless channels, in which $\{h_k\}$ is a complex, zero-mean Gaussian process with the channel order $L = 2$. Source signals are i.i.d QPSK signals. Results are averaged over 200 Monte Carlo runs and for each Monte Carlo run, a different FIR SISO channel is randomly generated.

In Fig. 1, we show the NMSE of the channel estimate of these three algorithms as a function of SNR, with the number of samples used to estimate the signal statistics, T_s, varying from 400 to 1600. It can be seen that, as expected, all algorithms improve consistently as SNR or number of samples T_s increases. Also the proposed algorithm CISC presents a slightly better performance than WOMP-SVD and a significant performance advantage over WS. The channel estimation NMSE of these three algorithms are compared to the Cramer-Rao bound (CRB), which provides a theoretical lower bound for all estimators. Once the channel is estimated, we can further detect the information sequences by adopting the Viterbi algorithm-based maximum likelihood detector. We present the SER performance of the algorithms in Table I, in which the SER is a function of SNR and T_s.

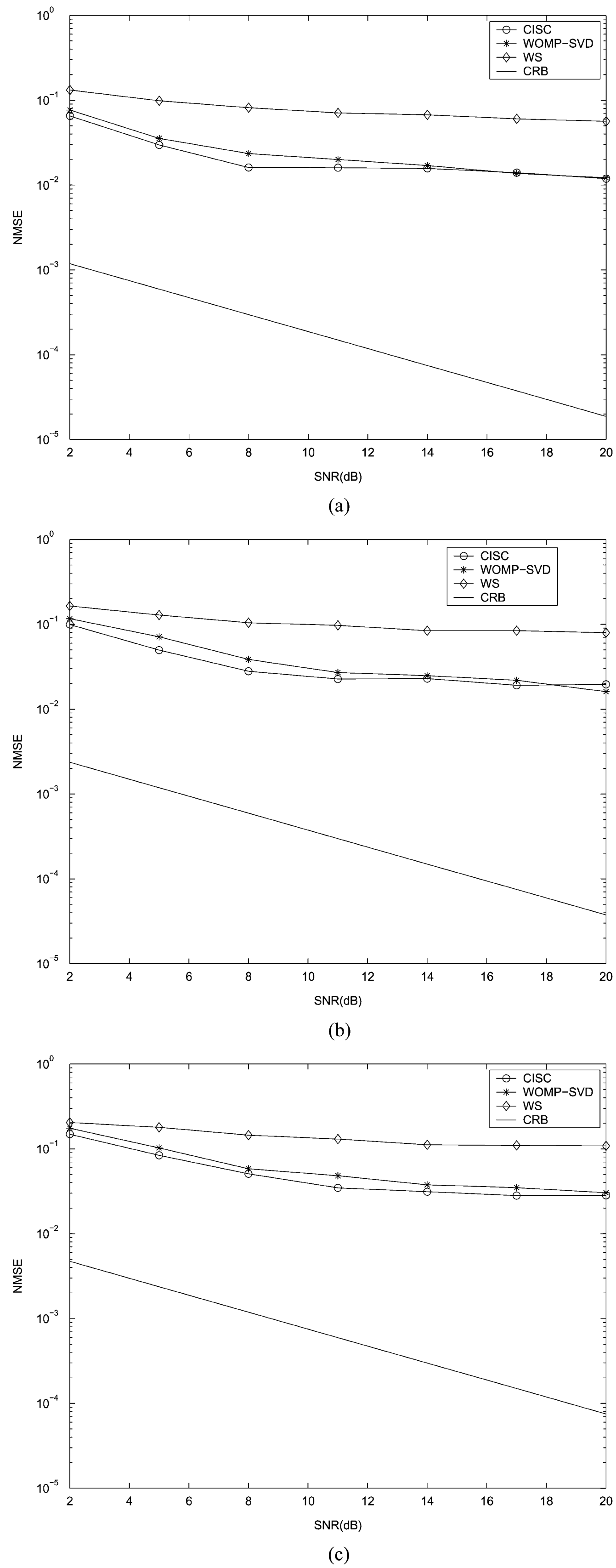

Fig. 1. The NMSE of the channel estimate versus SNR under different number of samples used. (a) $T_s = 1600$. (b) $T_s = 800$. (c) $T_s = 400$.

It can be seen that the SER performance depends on the following two parameters: SNR and NMSE of the channel estimate. On one hand,

[2]For the proposed algorithm in [16], we need to compute the generalized eigenvalue decomposition of two $(3L+1)\times(3L+1)$ matrices, which requires at least $30(3L+1)^3$ flops ([16, p. 385])

TABLE I
SER VERSUS SNR

SNR(dB)	$T_s = 1600$			$T_s = 800$			$T_s = 400$		
	CISC	WOMP-SVD	WS	CISC	WOMP-SVD	WS	CISC	WOMP-SVD	WS
20	0.0002	0.0003	0.0162	0.0005	0.0010	0.0299	0.0038	0.0028	0.0393
17	0.0005	0.0008	0.0171	0.0016	0.0015	0.0318	0.0043	0.0043	0.0447
14	0.0019	0.0014	0.0254	0.0061	0.0039	0.0346	0.0048	0.0089	0.0560
11	0.0059	0.0082	0.0349	0.0093	0.0108	0.0568	0.0212	0.0280	0.0847
8	0.0415	0.0452	0.1048	0.0528	0.0597	0.1324	0.0708	0.0801	0.1405
5	0.1558	0.1609	0.2234	0.1704	0.1794	0.2401	0.2012	0.2179	0.2744
2	0.3060	0.3126	0.3546	0.3226	0.3310	0.3789	0.3423	0.3606	0.3950

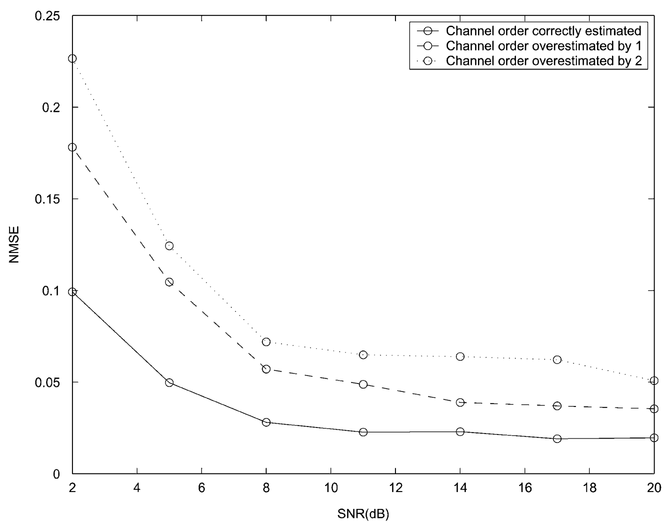

Fig. 2. The NMSE of the channel estimate versus SNR with channel order overestimated by 1 and 2, respectively.

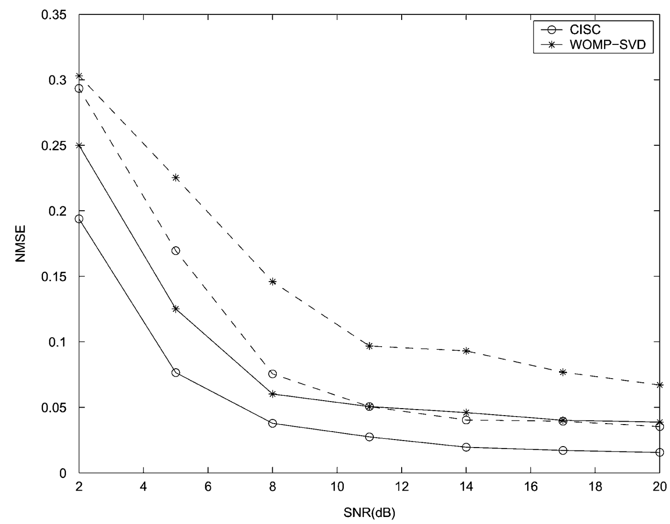

Fig. 3. NMSE of the channel estimate versus SNR. Solid lines are for $T_s = 1600$; dashed lines for $T_s = 800$.

the SER performance deteriorates as SNR decreases. On the other hand, under a certain SNR, a more accurate channel estimate yields a lower SER. Also we can see that, in general, CISC shows a lower SER than WOMP-SVD since more accurate channel estimations are used in the Viterbi detector. Also CISC and WOMP-SVD outperform WS significantly in terms of SER.

In Fig. 2, we demonstrate the performance of CISC when channel order is overestimated with $T_s = 800$. It can be observed that the performance degrades more rapidly as SNR deteriorates. The reason, we suspect, is that the interference subspace cancellation vector selection incurs more errors when SNR becomes low and channel order becomes large. While at a moderate SNR level when $\mathrm{SNR} \geq 8$ dB, the performance degradation is mild and acceptable, thus validating the theoretical analysis of the proposed method's robustness to channel order overestimation.

B. Example B

In this example, we consider the source that employs 16-ary QAM digital format. The channel transfer function is given as $\mathbf{h}(z) = -0.2039 + 0.4503z^{-1} + 0.7972z^{-2} - 0.3466z^{-3}$. In our simulations, results are averaged over 100 Monte Carlo runs. Fig. 3 shows the performance of CISC and WOMP-SVD as 1600 and 800 data samples are used, respectively. We can see that CISC owns a clear performance advantage over WOMP-SVD in both cases. It seems that, in such a channel scenario, CISC is more favorable than WOMP-SVD to obtain an accurate channel estimation. Besides, both algorithms suffer from a certain performance loss when 16-ary QAM digital modulation scheme is used. This is because, as compared to other simpler digital modulation schemes like QPSK, the source signals that employs 16-ary QAM digital modulation scheme induce a larger estimate variances between the estimated cumulants and the theoretical cumulants.

VI. CONCLUSION

In this correspondence, we present a new linear HOS method for blind SISO FIR channel estimation. The proposed method is robust to channel order overestimation. And, with a similar computational complexity, the new algorithm performs favorably with other existing linear HOS methods WS [6] and WOMP [14].

ACKNOWLEDGMENT

The authors would like to acknowledge the anonymous reviewers for their helpful comments and suggestions on this correspondence.

REFERENCES

[1] E. Moulines, P. Duhamel, J. F. Cardoso, and S. Mayrargue, "Subspace methods for the blind identification of multichannel FIR filters," *IEEE Trans. Signal Process.*, vol. 43, pp. 516–525, Feb. 1995.

[2] L. Tong, G. Xu, and T. Kailath, "Blind identification and equalization based on second-order statistics: A time domain approach," *IEEE Trans. Inf. Theory*, vol. 40, pp. 340–349, Mar. 1994.

[3] K. A. Meraim, P. Loubaton, and E. Moulines, "Prediction error method for second-order blind identification," *IEEE Trans. Signal Process.*, vol. 45, pp. 694–705, Mar. 1997.

[4] Z. Ding, "Matrix outer product decomposition method for blind multiple channel identification," *IEEE Trans. Signal Process.*, vol. 45, pp. 3053–3061, Dec. 1997.

[5] J. K. Tugnait, "On blind identifiability of multipath channels using fractional sampling and second-order cyclostationary statistics," *IEEE Trans. Inf. Theory*, vol. 41, pp. 308–311, Jan. 1995.

[6] J. A. R. Fonollosa and J. Vidal, "System identification using a linear combination of cumulant slices," *IEEE Trans. Signal Process.*, vol. 41, pp. 2405–2411, Jul. 1993.

[7] L. Srinivas and K. V. S. Hari, "FIR system identification using higher-order cumulant—A generalized approach," *IEEE Trans. Commun.*, vol. 43, pp. 3061–3065, Dec. 1995.

[8] ——, "FIR system identification based on subspaces of a higher order cumulant matrix," *IEEE Trans. Signal Process.*, vol. 44, pp. 1485–1491, Jun. 1996.

[9] Z. Ding and J. Liang, "A cumulant matrix subspace algorithm for blind single FIR channel identification," *IEEE Trans. Signal Process.*, vol. 49, pp. 325–333, Feb. 2001.

[10] G. B. Giannakis and J. M. Mendal, "Identification of nonminimum-phase system using higher order statistics," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, pp. 360–377, Mar. 1989.

[11] A. Swami and J. M. Mendel, "Arma parameter estimation using only output cumulants," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 38, pp. 1257–1265, Jul. 1990.

[12] J. K. Tugnait, "Approaches to FIR system identification with noisy data using higher-order statistics," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 38, pp. 1307–1317, Jul. 1990.

[13] J. Liang and Z. Ding, "FIR channel estimation through generalized cumulant slice weighting," *IEEE Trans. Signal Process.*, vol. 52, pp. 657–667, Mar. 2004.

[14] ——, "Nonminimum-phase FIR channel estimation using cumulant matrix pencils," *IEEE Trans. Signal Process.*, vol. 51, pp. 2310–2319, Sep. 2003.

[15] C. L. Nikias and A. P. Petropulu, *Higher-Order Spectra Analysis: A Nonlinear Signal Processing Framework*. Eaglewood Cliffs, NJ: Prentice-Hall, 1993.

[16] G. H. Golub and C. F. V. Loan, *Matrix Computations*. Baltimore, MD: The Johns Hopkins University Press, 1996.

Roundoff Noise Analysis of Two Efficient Digital Filter Structures

Zixue Zhao and Gang Li

***Abstract*—In this correspondence, two efficient structures are analyzed for the implementation scheme of *rounding before multiplication* (RBM). The first one is the direct-form II transposed structure in ρ-operator (ρDFIIt), recently proposed in [11]. The expression of roundoff noise gain is derived, based on which a revised ρDFIIt structure, denoted as ρRDFIIt, is obtained. It is shown that the ρRDFIIt structure, having the same implementation complexity, yields a smaller roundoff noise gain than the ρDFIIt. The roundoff noise gain for the ρRDFIIt structure with error feedbacks is also derived. Each structure can be optimized with regard to the free parameters it possesses with a genetic algorithm. Numerical examples are presented, which show that the optimized structures are very competitive because they can even outperform the traditional optimal state–space realization in terms of the roundoff noise performance as well as the implementation simplicity.**

***Index Terms*—Direct-form II transposed (DFIIt) structure, error feedback, polynomial parametrization, roundoff noise, state–space realization.**

I. INTRODUCTION

Roundoff noise has been considered as one of the most serious issues in digital filter implementation for more than three decades. There are two rounding schemes: *rounding after multiplication* (RAM) and *rounding before multiplication* (RBM). It is well known that roundoff noise gain can be reduced considerably by an appropriate selection of filter structures. The basic idea behind this approach is that for a given filter there exist a number of different structures. They are theoretically equivalent since they represent the same system transfer function. However, different structures have different numerical properties and, for a given application (measure or criterion), one structure can be better than another. Another important issue is the implementation complexity. For practical considerations, it is desired that the actually implemented filters have a nice performance as well as a simple structure that possesses many trivial parameters,[1] which can be implemented exactly and produce no rounding errors.

The optimal state–space realization design [1]–[4] has been known as one of the effective methods to reduce the roundoff noise. For a digital filter of order p, such an optimal realization has $(p+1)^2$ nontrivial parameters, which is obviously not efficient to implement. A lot of efforts have been made to search for the realizations that not only yield a nice performance against roundoff noise but also have a sparse structure [5]–[7].

It is well known that although having poor numerical properties, the conventional-shift-operator-based direct forms such as direct-form II (denoted as zDFII) and direct-form II transposed (zDFIIt) structures are the simplest structures. Recently, the direct forms in delta operator have been studied by researchers [8]–[10]. In [8], the δDFIIt structure was investigated for an arbitrary order infinite-impulse-response (IIR) filter, but it was found that it has a very good performance just for the

Manuscript received September 30, 2004; revised April 9, 2005. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Yuan-Pei Lin.

The authors are with the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore 639798 (e-mail: ZHAOZX@pmail.ntu.edu.sg; egli@ntu.edu.sg).

Digital Object Identifier 10.1109/TSP.2005.861895

[1]By *trivial parameters*, we mean those that are 0 and ±1. Other parameters are, therefore, referred to *nontrivial parameters*.