# Subspace Methods for Blind Identification of Multichannel FIR Filters Using Space-Time Contraction of Cumulants

Massimiliano Martone, *Member, IEEE*

***Abstract*—We show that the popular subspace method for identification of multichannel systems can be implemented using particular contractions of higher-order cumulants, in the case of independent, identically distributed (i.i.d.) inputs. The algorithm appears to outperform the traditional correlation-based subspace algorithm in low signal-to-noise-ratio (SNR) and/or correlated noise.**

***Index Terms*—Higher order statistics, identification.**

## I. INTRODUCTION

THE SUBSPACE algorithm proposed in [1] uses second-order statistics to blindly identify a multichannel finite impulse response (FIR) filter. We show here how the method can be implemented using $p+q+2$th order statistics where $p+q \geq 1$. The immediate advantage is that the algorithm, being theoretically insensitive to additive Gaussian noise, improves significantly over the traditional correlation-based subspace method in terms of identification accuracy at low signal-to-noise-ratio (SNR) and/or in correlated Gaussian noise.

## II. SPACE-TIME CONTRACTION OF CUMULANTS

Using standard filtering and sampling operations, we can represent a digitally modulated oversampled signal affected by intersymbol interference [1] as

$$y_n^{(k)} = \sum_m h_k(m) x_{n-m} + \eta_n^{(k)}, \qquad k = 1, 2, \cdots, K \tag{1}$$

where $\eta_n^{(k)}$ describes Gaussian noise and $x_n$ is a complex symbol. Two assumptions are assumed to hold (see [1] for a discussion)

**AS1:** The transformation in (1) represents a stable system but possibly nonminimum phase, satisfying: 1) all channels $h_i(k), i = 1, 2, \cdots, K$, have finite support $M+1$,, and 2) $h_i(M) \neq 0$ for some $i, h_i(0) \neq 0$ for some $i$, 3) $\mathcal{H}_i(z) = \Sigma_{n=0}^{M} h_i(n) z^{-n}, i = 1, 2, \cdots, K$ have no common zeroes.

**AS2:** the sequence $\{x_n\}$ is constituted by zero-mean random variables non-Gaussian distributed, statistically independent, having nonzero $p+q+2$th order zero-lag cumulant[1], that is $\gamma_x^{[p,q]} = \mathrm{cum}[x_n\colon p+1, \overline{x}_n\colon q+1] \neq 0$ [2].

Arrange channels in vector form

$$\boldsymbol{H}^T = [\boldsymbol{H}_1^T, \boldsymbol{H}_2^T, \cdots, \boldsymbol{H}_K^T]$$

where

$$[\boldsymbol{H}_i]_{n,m} = \begin{cases} h_i(m-n), & 0 \leq m-n \leq M \\ 0, & m-n > M \text{ or } m-n < 0 \end{cases}$$

and $\boldsymbol{h}_i = [h_i(0), h_i(1), \cdots, h_i(M)]^T$. Expression (1) can be compactly expressed as

$$\boldsymbol{y}_n = \boldsymbol{H}\boldsymbol{x}_n + \boldsymbol{n}_n \tag{2}$$

where

$$\begin{aligned} \boldsymbol{y}_n &= [\boldsymbol{y}_n^{(1)^T}, \cdots, \boldsymbol{y}_n^{(K)^T}]^T, \quad \boldsymbol{y}_n^{(m)} = [y_n^{(m)}, \cdots, y_{n-N+1}^{(m)}]^T \\ \boldsymbol{x}_n &= [x_n, x_{n-1}, \cdots, x_{n-N-M+1}]^T \\ \boldsymbol{n}_n &= [\boldsymbol{n}_n^{(1)^T}, \cdots, \boldsymbol{n}_n^{(K)^T}]^T, \quad \boldsymbol{n}_n^{(m)} = [\eta_n^{(m)}, \cdots, \eta_{n-N+1}^{(m)}]^T \end{aligned}$$

with $N \geq M$ and the dimensions of $\boldsymbol{H}$ are properly defined[2]. We introduce here the following contracted-cumulant matrix:

$$\tilde{\boldsymbol{C}}_{y,p,q} = \begin{bmatrix} \boldsymbol{C}_{y,p,q}^{(1,1)} & \boldsymbol{C}_{y,p,q}^{(1,2)} & \cdots & \boldsymbol{C}_{y,p,q}^{(1,K)} \\ \boldsymbol{C}_{y,p,q}^{(2,1)} & \boldsymbol{C}_{y,p,q}^{(2,2)} & \cdots & \boldsymbol{C}_{y,p,q}^{(2,K)} \\ \vdots & \vdots & \vdots & \vdots \\ \boldsymbol{C}_{y,p,q}^{(K,1)} & \boldsymbol{C}_{y,p,q}^{(K,2)} & \cdots & \boldsymbol{C}_{y,p,q}^{(K,K)} \end{bmatrix}$$

where the generic $k, i$ element (with $1 \leq k \leq N, 1 \leq i \leq N$) of the matrix $\boldsymbol{C}_{y,p,q}^{(m_1,m_2)}$ is obtained by a space-time

---

[1] We use the notation $\overline{x}$ to denote complex conjugation of $x$ and

$$\mathrm{cum}[x, \overline{y}, z\colon p, \overline{w}\colon q] = \mathrm{cum}[x, y, \overbrace{z, \cdots, z}^{p}, \overbrace{\overline{w}, \cdots, \overline{w}}^{q}]$$

to denote the $p+q+2$th order cumulant of the joint complex random variable $\{x, y, z, w\}$.

[2] The filtering matrix $\boldsymbol{H}$ is full column rank due to **AS1** and **AS2** and to the assumption $N \geq M$ [1].

Manuscript received February 19, 1998. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. G. B. Giannakis.

The author is with the Telecommunications Group, Watkins-Johnson Company, Gaithersburg, MD 20878-1794 USA (e-mail: max.martone@wj.com).

Publisher Item Identifier S 1070-9908(98)05128-1.

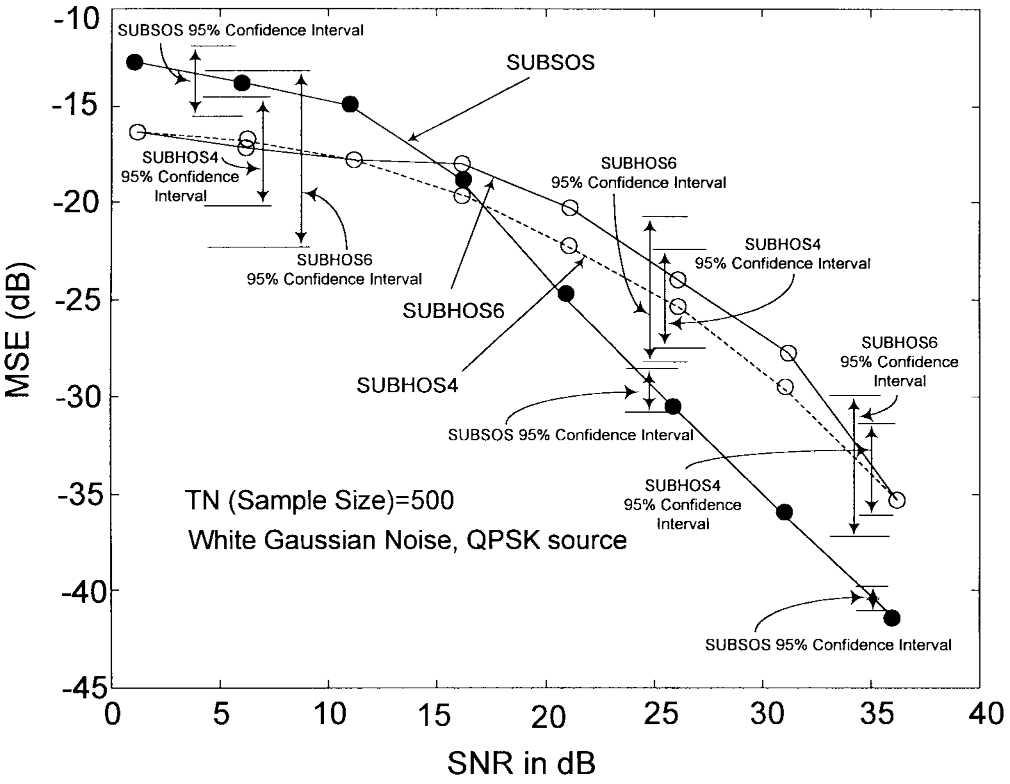

Fig. 1. Estimation MSE versus SNR for white Gaussian noise with 95% confidence intervals at SNR = 5, 25, and 35 dB.

contraction[3] of cumulants as

$$[\boldsymbol{C}_{y,p,q}^{(m_1,m_2)}]_{k,i} = \sum_{r=1}^{K}\sum_{l=-\infty}^{\infty} \mathrm{cum} \cdot [y_{n-k+1}^{(m_1)}, \overline{y}_{n-i+1}^{(m_2)}, y_{n-l+1}^{(r)}: p, \overline{y}_{n-l+1}^{(r)}: q]. \tag{3}$$

Due to the linearity property of cumulants [2] and the independence of the source symbols, we have that

$$\begin{aligned}\sum_{l=-\infty}^{\infty} &\mathrm{cum}[y_{n-k+1}^{(m_1)}, \overline{y}_{n-i+1}^{(m_2)}, y_{n-l+1}^{(r)}: p, \overline{y}_{n-l+1}^{(r)}: q]\\ &= \gamma_x^{[p,q]} \sum_{l=-\infty}^{\infty}\sum_{k_1} h_{m_1}(k_1-k+1)\overline{h}_{m_2}(k_1-i+1)\\ &\quad\times [h_r(k_1-l+1)]^p[\overline{h}_r(k_1-l+1)]^q\\ &= \gamma_x^{[p,q]} \sum_{k_1} h_{m_1}(k_1-k+1)\overline{h}_{m_2}(k_1-i+1)\\ &\quad\times \sum_{l=-\infty}^{\infty} [h_r(l)]^p[\overline{h}_r(l)]^q\end{aligned} \tag{4}$$

where we have also used the fact that cumulants of a Gaussian process vanish for orders greater than two [2]. Substituting (4) into (3), we obtain

$$[\boldsymbol{C}_{y,p,q}^{(m_1,m_2)}]_{k,i} = [\boldsymbol{H}_{m_1}\boldsymbol{H}_{m_2}^H]_{k,i}\mathcal{K}^{[p,q]} \tag{5}$$

where $\mathcal{K}^{[p,q]} = \gamma_x^{[p,q]} \Sigma_{r=1}^{K} \Sigma_{l=-\infty}^{+\infty}[h_r(l)]^p[\overline{h}_r(l)]^q$. The matrix $\tilde{\boldsymbol{C}}_{y,p,q}$ can be so decomposed as

$$\tilde{\boldsymbol{C}}_{y,p,q} = \boldsymbol{H}\,\boldsymbol{K}^{[p,q]}\,\boldsymbol{H}^H \tag{6}$$

and $\boldsymbol{K}^{[p,q]} = \mathcal{K}^{[p,q]}\boldsymbol{I}_{(N+M)\times(N+M)}$ is a $(N+M)\times(N+M)$ diagonal matrix. Observe that $\tilde{\boldsymbol{C}}_{y,p,q}$ is Hermitian.

## III. PRACTICAL METHODS

The matrix $\tilde{\boldsymbol{C}}_{y,p,q}$ can be estimated using $TN$ samples as[4]

$$[\hat{\tilde{\boldsymbol{C}}}_{y,p,q}^{(m_1,m_2)}]_{k,i} = \sum_{r=1}^{K}\sum_{l=B_1^{(k-i)}}^{B_2^{(k-i)}} \times \mathrm{cum}^{(TN)}[y_n^{(m_1)}, \overline{y}_{n-i+k}^{(m_2)}, y_{n-l}^{(r)}: p, \overline{y}_{n-l}^{(r)}: q] \tag{7}$$

where $\mathrm{cum}^{(S)}[\cdots]$ is a consistent and asymptotically normal estimator [5] of the cumulant $\mathrm{cum}[\cdots]$ based on sample-statistics using $S$ samples. We limited the infinite summation of the contraction in time to $B_1^{(k-i)}$ and $B_2^{(k-i)}$. These limits are defined by the region of support of the cumulant to be estimated. In fact the generic cumulant term of (7) is of

---

[3] We use the tensor terminology *contraction* following [4] where in a direction finding context it was stated that summing fourth-order cross-cumulants with respect to the sensor index is equivalent to operating a contraction of the quadricovariance. Algorithms for system identification based on time contractions (defined *projections*) of cumulants were first presented in [3]. In the multichannel framework we sum cumulants with respect to the channel index (space contraction) and with respect to the time index (time contraction).

[4] Observe that due to the stationarity of the subprocesses $y_n^{(m)}$ we have that $\mathrm{cum}[y_{n-k+1}^{(m_1)}, \overline{y}_{n-i+1}^{(m_2)}, y_{n-l+1}^{(r)}: p, \overline{y}_{n-l+1}^{(r)}: q] = \mathrm{cum}[y_n^{(m_1)}, \overline{y}_{n-i+k}^{(m_2)}, y_{n-l+k}^{(r)}: p, \overline{y}_{n-l+k}^{(r)}: q]$. Moreover $\Sigma_{l=-\infty}^{+\infty} \mathrm{cum}[y_n^{(m_1)}, \overline{y}_{n-i+k}^{(m_2)}, y_{n-l+k}^{(r)}: p, \overline{y}_{n-l+k}^{(r)}: q] = \Sigma_{l=-\infty}^{+\infty} \mathrm{cum}[y_n^{(m_1)}, \overline{y}_{n-i+k}^{(m_2)}, y_{n-l}^{(r)}: p, \overline{y}_{n-l}^{(r)}: q]$ can be proved using a marginal modification of the derivation in (4).

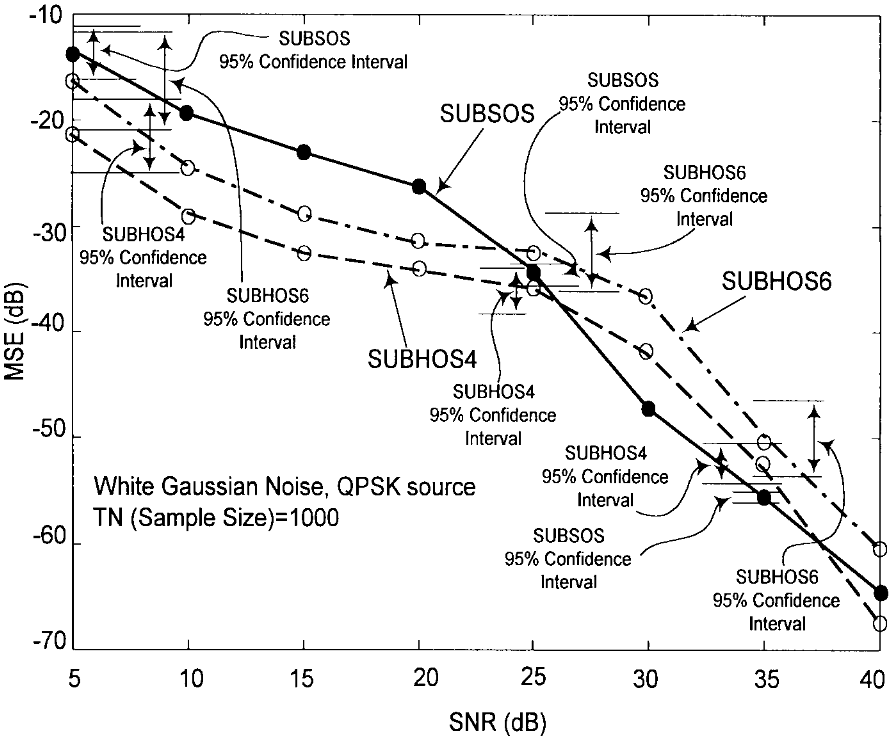

Fig. 2. Estimation MSE versus SNR for white Gaussian noise with 95% confidence intervals at SNR = 5, 25, and 35 dB.

the form $\mathrm{cum}[y_n^{(m_1)}, \overline{y}_{n-k}^{(m_2)}, y_{n-l}^{(r)}: p, \overline{y}_{n-l}^{(r)}: q]$. The region of support $[B_1^{(k)}, B_1^{(k)}+1, \cdots, B_2^{(k)}]$ of this cumulant (that is the valid range of values for $l$) can be determined as follows. Since $h_i(n)$ are FIR of length $M+1$, we have that

$$\begin{aligned}&\mathrm{cum}[y_n^{(m_1)}, \overline{y}_{n-k}^{(m_2)}, y_{n-l}^{(r)}: p, \overline{y}_{n-l}^{(r)}: q]\\ &\quad= \mathrm{cum}[y_{n+l}^{(m_1)}, \overline{y}_{n+l-k}^{(m_2)}, y_n^{(r)}: p, \overline{y}_n^{(r)}: q]\\ &\quad= \gamma_x^{[p,q]} \sum_n h_{m_1}(n+l)\overline{h}_{m_2}(n+l-k)h_r(n)^p\overline{h}_r(n)^q\end{aligned}$$

is different from zero only for $\max\{-M, -M+k\} \leq l \leq \min\{M, M+k\}$. Consequently $B_1^{(k)}$ and $B_2^{(k)}$ can be selected as

$$B_1^{(k)} \leq max\{-M, -M+k\}, \quad B_2^{(k)} \geq \min\{M, M+k\}.$$

The decomposition (6) allows the use of $\tilde{\boldsymbol{C}}_{y,p,q}$ in the subspace method[5]. The span of the $M+N$ eigenvectors of $\tilde{\boldsymbol{C}}_{y,p,q}$ associated with the $M+N$ eigenvalues of largest magnitude is also the linear space spanned by the columns of the filtering matrix $\boldsymbol{H}$. Since $\boldsymbol{H}$ is full column rank and $N \geq M$, the subspace of $\tilde{\boldsymbol{C}}_{y,p,q}$ *uniquely* determines the channel coefficients up to a multiplicative constant (see [1, Th. 2]). The $KN-(M+N)$ eigenvectors associated with the $KN-(M+N)$ smallest eigenvalues (theoretically equal to zero) are orthogonal to the columns of $\boldsymbol{H}$. The orthogonality condition for $\tilde{\boldsymbol{C}}_{y,p,q}$ is

$$\boldsymbol{d}_i^H \boldsymbol{H} = 0 \quad \text{for} \quad M+N \leq i \leq kn-1$$

where $\boldsymbol{d}_i, i = 0, 1, \cdots, KN-1$ are the eigenvectors corresponding to the ordered eigenvalues in the eigenvalue decomposition of $\tilde{\boldsymbol{C}}_{y,p,q}$. In practice, only estimates of cumulants are available and the orthogonality condition is more conveniently expressed as a minimization of a quadratic form involving the channel coefficients. Defining $q(\boldsymbol{h}) = \Sigma_{i=M+N}^{KN-1} |\boldsymbol{d}_i^H \boldsymbol{H}|^2$, it is possible to show as in [1] that $q(\boldsymbol{h}) = \boldsymbol{h}^H \boldsymbol{Q} \boldsymbol{h}$, where $\boldsymbol{h} = [\boldsymbol{h}_1^T \boldsymbol{h}_2^T, \cdots, \boldsymbol{h}_K^T]^T$ and $\boldsymbol{Q} = \Sigma_{i=M+N}^{KN-1} \mathcal{D}_i \mathcal{D}_i^H$. The matrix $\mathcal{D}_i$ is a matrix (see [1]) constructed from $\boldsymbol{d}_i$ exactly like $\boldsymbol{H}$ was constructed from $\boldsymbol{h}$. The problem of finding the minimum of the function $q(\boldsymbol{h})$ can also be expressed as a constrained optimization problem with a linear/quadratic constrain (see [1] for details).

## IV. SIMULATIONS

We present results for fourth-order and sixth-order cumulants. The following labeling system applies to the results of the simulations presented in this Section.

- $SUBSOS \Rightarrow$ correlation-based method.
- $SUBHOS4 \Rightarrow \hat{\tilde{\boldsymbol{C}}}_{y,1,1}$-based method.
- $SUBHOS6 \Rightarrow \hat{\tilde{\boldsymbol{C}}}_{y,2,2}$-based method.

The SNR in dB is defined as $\mathrm{SNR_{dB}} = 10 \; log_{10} (\sigma_x^2 \|\boldsymbol{h}\|^2 / K\sigma_\eta^2)$ and the estimation MSE for 100 Monte Carlo runs is defined as $\mathrm{MSE_{dB}} = 20 \; log_{10} \sqrt{\Sigma_{r=1}^{100} \|\boldsymbol{h} - \hat{\boldsymbol{h}}_r\|^2}$ where $\boldsymbol{h}, \hat{\boldsymbol{h}}_r$ are the true channel and the channel estimated at the $r$th computer run. We also show the 95% MSE confidence intervals at SNR = 5, 25, and 35 dB. All the simulations are performed using the same parameters of the simulation example given in [1]: $N = 10, K = 4, M = 4$,

---

[5] It is important to observe that the idea of contracting (or projecting) cumulants to obtain generalizations of algorithms based on second-order statistics, was first used in [3].

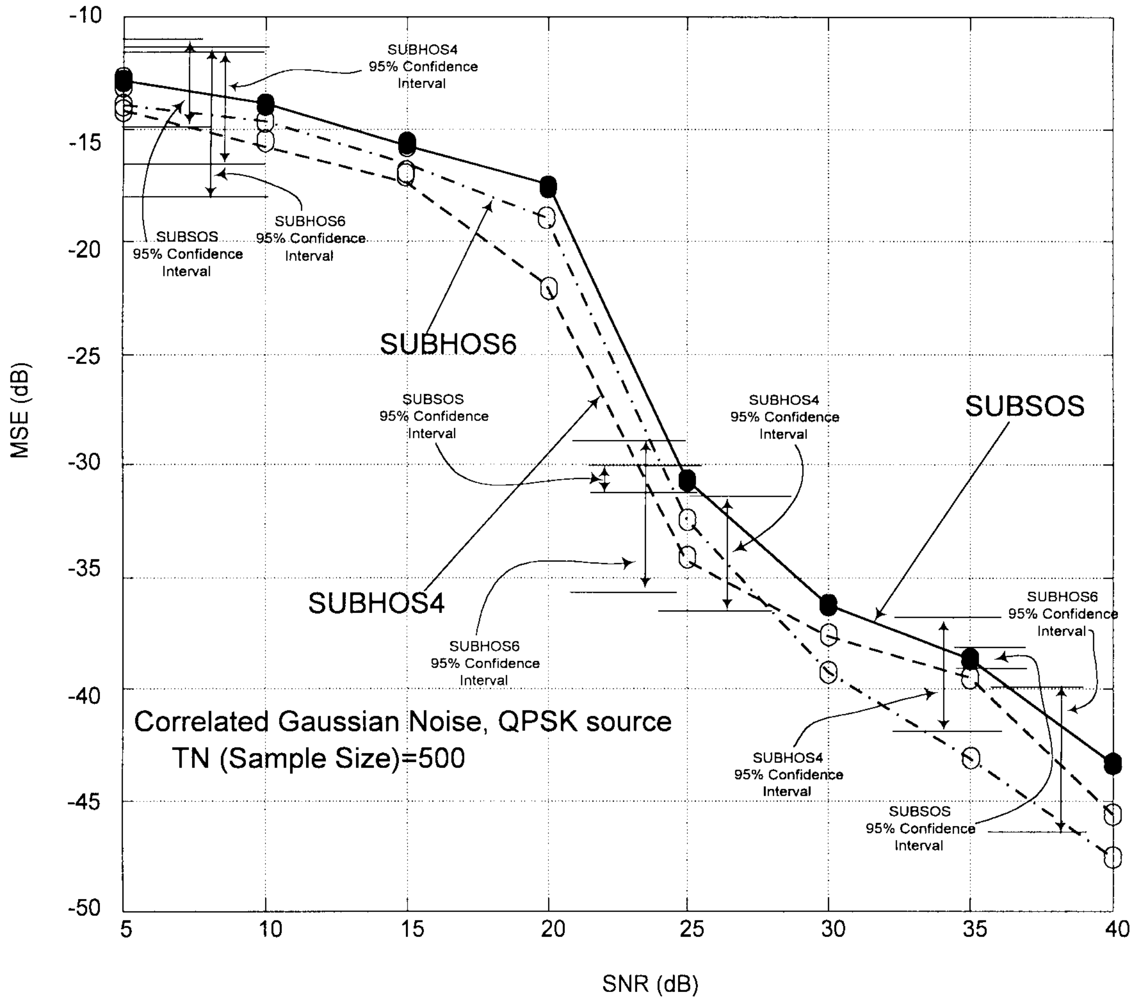

Fig. 3. Estimation MSE versus SNR for correlated Gaussian noise with 95% confidence intervals at SNR = 5, 25, and 35 dB.

discrete-time complex channels given in [1]. Fig. 1 shows results for white Gaussian noise with a sample-size of $TN = 500$. Fig. 2 shows results for white Gaussian noise with a sample-size of $TN = 1000$. The sample MSE results show significant advantage at low SNR but the important well-known disadvantage of the cumulant-based methods, that is the large variance of their estimates, is evident from the confidence intervals. Fig. 3 shows results for correlated Gaussian noise with a sample-size of $TN = 500$. The Gaussian noise multivariate process $[\eta_n^{(1)}, \eta_n^{(2)}, \eta_n^{(3)}, \eta_n^{(4)}]^T$ was correlated filtering white Gaussian noise with a 4-input 4-output moving average (MA) model. The MA model used

---

$$\boldsymbol{B}(0) = \begin{bmatrix} -0.7883-0.6440i & -0.9700+0.1970i & -0.9373-0.6676i & 0.3650+0.6101i \\ -1.6017+0.8028i & 2.1416+0.3649i & 0.3868-0.0762i & -0.1479+0.1063i \\ 0.0499+0.5263i & -0.1103-1.5429i & 1.7134-0.3482i & 1.7822-0.1681i \\ 0.2352+0.9788i & 0.6543-1.6864i & 1.1562-0.4805i & 0.7681-0.0680i \end{bmatrix}$$

$$\boldsymbol{B}(1) = \boldsymbol{0}$$

$$\boldsymbol{B}(2) = \begin{bmatrix} -1.4617-0.2601i & 1.6100+1.3369i & 0.6082+0.1064i & -0.7617+0.9979i \\ -0.3692+0.1155i & 0.7293-0.2901i & 1.6078-0.8970i & -0.2730+1.0206i \\ 1.1310-1.2038i & 0.9472+0.8051i & 0.2063+1.0812i & -2.3280-0.3077i \\ 1.5015+0.9021i & -0.2532-0.0132i & 1.6959+0.2936i & 0.9663-0.3939i \end{bmatrix}$$

$$\boldsymbol{B}(3) = \boldsymbol{0}$$

$$\boldsymbol{B}(4) = \begin{bmatrix} 0.7910+0.7907i & -0.0893+1.3425i & -0.7716-0.2655i & 1.9886+0.8873i \\ 0.6541+0.3361i & 1.1789-0.3488i & 0.1720+0.2985i & 0.3374+0.8204i \\ 1.5183+0.8142i & 0.6823+0.1308i & 1.1864-0.4979i & 0.9673+0.3080i \\ -0.1206+0.9063i & -1.5915-0.7726i & -0.5155+0.2168i & 0.3650-0.8783i \end{bmatrix}$$

to correlate noise is shown at the bottom of the previous page. In colored noise, the SUBSOS method uses the whitening approach of [1], which assumes knowledge of the noise correlation properties.

## 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 Processing*, vol. 43, pp. 516–525, Feb. 1995.

[2] J. M. Mendel, "Tutorial on higher order statistics (spectra) in signal processing and system theory: theoretical results and some applications," *Proc. IEEE*, vol. 79, pp. 278–305, Mar. 1991.

[3] A. N. Delopoulos and G. B. Giannakis, "Strongly consistent identification algorithms and noise insensitive MSE criteria," *IEEE Trans. Signal Processing*, vol. 30, pp. 1955–1970, Aug. 1992.

[4] J. Cardoso and E. Moulines, "Asymptotic performance analysis of direction-finding algorithms based on fourth-order cumulants," *IEEE Trans. Signal Processing*, vol. 43, pp. 214–224, Jan. 1995.

[5] A. V. Dandawate and G. B. Giannakis, "Asymptotic theory of mixed time averages and $k$th order cyclic-moment and cumulant statistics," *IEEE Trans. Inform. Theory*, vol. 41, pp. 216–232, Jan. 1995.