

TABLE II
OPTIMAL NUMBER OF SCALES AND MINIMUM MEAN-SQUARE ERROR
FOR DIFFERENT VALUES OF THE PARAMETER H : THEORETICAL AND
SIMULATED VALUES CONSIDERING THE EFFECTS OF THE ALIASING

H	J^{opt}	theoretical mean-square error	simulation
0.60	4	0.5533	0.5410
0.65	4	0.5613	0.5486
0.70	3	0.5727	0.5591
0.75	3	0.5849	0.5721
0.80	3	0.5995	0.5883
0.85	3	0.6161	0.6082
0.90	2	0.6400	0.6339
0.95	2	0.6590	0.6615

We have found numerically the optimal number of scales for each value of H . The results are shown in Table II; the mean-square errors were obtained two ways: a) the expressions previously demonstrated and b) the values obtained from simulations of the filter bank with the optimal number of scales. As we can see, with the consideration of the aliasing, the mean-square error values obtained in the simulations are closer to the theoretical ones (compare with Table I).

In previous papers [5], [6], the number of scales used was $J = 11$. As we can see, with $J = 4$, the mean square-error is minimized in the cases considered. This number is smaller than that obtained in [8] using Wiener filters.

IV. CONCLUSIONS

A scheme for the estimation of fBm was developed on the basis of a bank of multiscale Kalman filters. It takes into account the correlation of the wavelet coefficients and the approximation coefficients in the wavelet expansion.

Numerical results were shown on the optimal number of scales and the minimum mean-square error for different values of the parameter $H > 1/2$ in the fBm.

It was shown that the mean-square error does not strictly decrease with respect to the number of scales J and that the number of filters required in a bank of Kalman filters is smaller than the number needed in a bank of Wiener filters. This fact represents an improvement in the design of the filter bank.

We have presented several numerical simulations. The aliasing effect was considered. The theoretical mean-square error values calculated matched the ones obtained in the numerical simulations.

REFERENCES

- [1] B. B. Mandelbrot, "Self-similar error clusters in communication systems and the concept of conditional stationarity," *IEEE Trans. Commun. Technol.*, vol. CT-13, pp. 71–90, Mar. 1965.
- [2] —, *The Fractal Geometry of Nature*. San Francisco, CA: Freeman, 1982.

- [3] M. S. Keshner, "1/f noise," *Proc. IEEE*, vol. 70, pp. 212–218, 1982.
- [4] W. E. Leland, M. S. Taqqu, W. Willinger, and D. V. Wilson, "On the self-similar nature of Ethernet traffic," in *Proc. ACM SIGComm.*, San Francisco, CA, Sept. 1993.
- [5] B. S. Chen and C. W. Lin, "Multiscale Wiener filter for the restoration of fractal signals: Wavelet filter bank approach," *IEEE Trans. Signal Processing*, vol. 42, pp. 2972–2982, 1994.
- [6] B. S. Chen and W. S. Hou, "Deconvolution filter design for fractal signal transmission systems: A multiscale Kalman filter bank approach," *IEEE Trans. Signal Processing*, vol. 45, pp. 1359–1364, May 1997.
- [7] G. A. Hirchoren and C. E. D'Attellis, "Estimation of fractal signals using wavelets and filter banks," *IEEE Trans. Signal Processing*, vol. 46, pp. 1624–1630, June 1998.
- [8] —, "On the optimal number of scales in estimation of fractal signals using wavelets and filter banks," *Signal Process.*, vol. 63, pp. 55–63, 1997.
- [9] M. F. Barnsley *et al.*, *The Science of Fractal Images*. New York: Springer-Verlag, 1988.
- [10] C. K. Chui, *An Introduction to Wavelets*. San Diego, CA: Academic, 1992.
- [11] P. Flandrin, "Wavelet analysis and synthesis of fractional brownian motion," *IEEE Trans. Inform. Theory*, vol. 38, pp. 910–917, Apr. 1992.
- [12] L. R. Rabiner and R. W. Schafer, *Digital Processing of Speech Signals*. Englewood Cliffs, NJ: Prentice-Hall, 1978.

Hybrid Nonlinear Moments Subspace Processing for Wireless Communication Systems Using Antenna Arrays

Massimiliano (Max) Martone

Abstract—It is shown that hybrid nonlinear (HNL) moments matrices can be used to identify a multichannel FIR system. The method outperforms the popular subspace method based on second-order statistics in low SNR and/or correlated noise, and it is fundamentally equivalent in terms of computational complexity. The method is applied to the channel identification problem in a cellular base-station receiver employing antenna arrays.

Index Terms—Higher order statistics, identification.

I. INTRODUCTION

Given a complex random vector \mathbf{x} whose autocorrelation matrix is $\mathbf{R}_x = E_X\{\mathbf{x}\mathbf{x}^H\}$, a hybrid nonlinear moment (HNL) matrix is defined as $\mathbf{G}_x = E_X\{\mathbf{x}\tilde{\mathbf{x}}^H\}$, where $\tilde{\mathbf{x}}$ is obtained by applying a zero-memory nonlinear (ZMNL) transformation to each component of \mathbf{x} . Array processing methods based on HNL matrices were introduced in [1], where it was shown that HNL matrices retain the same subspace of the correlation matrix in the harmonic retrieval problem and result in considerable savings of computations for particular choices of the ZMNL function. In this work, we address the more general problem of blind multichannel FIR system identification. We show that HNL-based methods outperform the popular second-order statistics based method of [3] while maintaining approximately similar computational

Manuscript received August 23, 1996; revised October 11, 1998. This work was presented in part at the IEEE International Conference on Communications, Dallas, TX, June 1996. The associate editor coordinating the review of this paper and approving it for publication was Dr. Jonathon A. Chambers.

The author is with the Telecommunications Group, Watkins-Johnson Company, Gaithersburg, MD 20878 USA.

Publisher Item Identifier S 1053-587X(99)03256-0.

complexity. We then focus on a practical channel estimation problem in a base-station receiver employing an antenna array.

II. SYSTEM MODEL

We assume a mobile transmitter communicating with a base-station employing a K -element antenna. Multipath propagation can be characterized as an N_p -path channel whose n th path ($n = 1, 2, \dots, N_p$) is represented by P_n received, delayed, and attenuated replicas of the signal. The impulse response of the n th path can be expressed as $f_n(t) = \sum_{m=1}^{P_n} \rho_{n,m} e^{j\psi_{n,m}} \delta(t - \tau_{n,m})$, where $\tau_{n,m}$, $\rho_{n,m}$, and $\psi_{n,m}$ are delay, amplitude, and phase of the m th delayed signal in the n th path, respectively, whereas $\delta(t)$ is the usual delta function [5]. The complex baseband modulated signal is $m(t) = \sum_n x_n p(t - nT)$, where

- x_n complex symbols defining the signal constellation used for the particular digital modulation scheme.
- $p(t)$ square root raised cosine shaping filter;
- T signaling interval.

Assuming reception with a uniform linear array with K sensors, filtering with a square root raised cosine filter, and sampling rate equal to $1/T$, we can compact the effect of the RF propagation channels at baseband as

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

where $\eta_n^{(k)}$ is Gaussian noise.¹

In (1), $h_k(m) = h_k(mT)$ is the sampled response $h_k(t) = \sum_{n=1}^{N_p} \sum_{m=1}^{P_n} \rho_{n,m} e^{j\phi_{n,m}} r_g(t - \tau_{n,m}) e^{-j2\pi k d (\sin \theta_n / \lambda)}$, where $\phi_{n,m} = -2\pi f_0 \tau_{n,m} + \psi_{n,m}$, d is the distance between adjacent sensors, θ_n is the angle of arrival of the n th path,² λ is the wavelength, $\omega_0 = 2\pi f_0$ is the carrier frequency, and $r_g(t)$ is a raised cosine function. The important assumptions on the discrete-time system (see [3] for a discussion) are given in the following.

AS1) The transformation in (1) represents a stable system but possibly nonminimum phase, satisfying the following.

- 1) All channels $h_i(k)$, $i = 1, 2, \dots, K$, have finite support M .
- 2) $h_i(M) \neq 0$ for some i , and $h_i(0) \neq 0$ for some i .
- 3) $\mathcal{H}_i(z) = \sum_{n=0}^M h_i(n) z^{-n}$, $i = 1, 2, \dots, K$ have no common zeroes.

AS2) The complex sequence $\{x_n\}$ is constituted by random variables that are identically non-Gaussian distributed and statistically independent. We arrange channels in vector form $\mathbf{H}^T = [\mathbf{H}_1^T, \mathbf{H}_2^T, \dots, \mathbf{H}_K^T]$, where

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

and

$$\mathbf{h}_i = [h_i(0), h_i(1), \dots, h_i(M)]^T.$$

¹This channel model can be generally described as a one-input K -output system, and it does also indeed apply to the fractionally spaced case with minor changes.

²In this model, the n th path consists of P_n delayed replicas of the signal with the same angle of arrival due to scatterers nearby the mobile. In fact, assuming the scatterers evenly spread out on a circle surrounding each mobile and assuming large distance between the mobile and the base-station simple geometric considerations [4] can lead to the simplification of a *point-source approximation* for the scattering mechanism that is local to the mobile, that is, we can assume that P_n delayed replicas of the signal are received with approximately the same angle of arrival. For a certain number N_p of reflections of the transmitted signal, particularly reflections in the vicinity of the base station, these assumptions are not reasonable, and different angles of arrival have to be considered.

Expression (1) can be compactly stated as

$$\mathbf{y}_n = \mathbf{H} \mathbf{x}_n + \mathbf{n}_n \quad (2)$$

where

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

with $N \geq M$, and the dimensions of \mathbf{H} are properly defined.

III. STRUCTURE OF HNL MATRICES

If x_1 and x_2 are two complex random variables, a HNL moment is defined as $E\{x_1 \bar{y}(x_2, \bar{x}_2)\}$, where an overbar indicates a complex conjugation, and $g(u, \bar{u})$ is a ZMNL transformation of $u = u_r + ju_i$ with u and \bar{u} as independent variables (see [1]). If $g(u, \bar{u})$ is a polynomial, HNL moments correspond to linear combinations of higher order moments and, for particular choices of the nonlinear transformation, to cumulants. Particularly, if z and y are two complex random variables, we have³ $E\{y \bar{g}(z, \bar{z})\} = \sum_{l=0}^{\infty} \sum_{n=0}^{\infty} (1/l!) (1/n!) \bar{E}_Z \{g(z, \bar{z})^{(l,n)}\} \text{cum}[y, z, l, \bar{z}; n]$ [2]. If y and z can be considered circularly complex, we can further simplify [2]

$$E_{YZ} \{y \bar{g}(z, \bar{z})\} = \sum_{l=0}^{\infty} \frac{1}{l!} \frac{1}{(l+1)!} \bar{E}_Z \{g(z, \bar{z})^{(l+1,l)}\} \cdot \text{cum}[y, z, l, \bar{z}; l+1]. \quad (3)$$

Let us define the HNL matrix

$$\mathbf{G}_y = E_Y \{\mathbf{y}_n \tilde{\mathbf{y}}_n^H\} = \begin{bmatrix} \mathbf{G}_y^{(1,1)} & \mathbf{G}_y^{(1,2)} & \dots & \mathbf{G}_y^{(1,K)} \\ \mathbf{G}_y^{(2,1)} & \mathbf{G}_y^{(2,2)} & \dots & \mathbf{G}_y^{(2,K)} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{G}_y^{(K,1)} & \mathbf{G}_y^{(K,2)} & \dots & \mathbf{G}_y^{(K,K)} \end{bmatrix}$$

and the ordinary correlation matrix $\mathbf{R}_y = E_Y \{\mathbf{y}_n \mathbf{y}_n^H\}$, where $\tilde{\mathbf{y}}_n$ is a vector defined exactly as \mathbf{y}_n , but with the generic m th component obtained from $\tilde{y}_m^{(l)} = g(y_m^{(l)}, \bar{y}_m^{(l)})$, $l = 1, 2, \dots, K$ ($g(u, \bar{u})$ is a proper ZMNL function). We show in the Appendix that the matrix \mathbf{G}_y can be expressed as

$$\begin{aligned} \mathbf{G}_y &= \mathbf{H} \mathbf{B}_{x,y} + \mathbf{R}_n \mathbf{K}_0 \\ \mathbf{B}_{x,y} &= \sum_{n=0}^{\infty} \frac{1}{n!(n+1)!} (\mathbf{H}^T)^{[n]} \mathbf{K}_n \gamma_n \end{aligned} \quad (4)$$

where $\gamma_n = \text{cum}[x_t : n+1, \bar{x}_t : n+1]$, $\mathbf{R}_n = E_N \{\mathbf{n}_n \mathbf{n}_n^H\}$, and \mathbf{K}_n is a diagonal matrix composed by diagonal matrices as

$$\begin{aligned} \mathbf{K}_n &= \begin{bmatrix} \mathbf{K}_n^{(1)} & \mathbf{0} & \dots & \mathbf{0} \\ \vdots & \ddots & \ddots & \vdots \\ \mathbf{0} & \dots & \mathbf{0} & \mathbf{K}_n^{(K)} \end{bmatrix} \\ [\mathbf{K}_n^{(m)}]_{j,j} &= \bar{E}_Y \{g(y_t^{(m)}, \bar{y}_t^{(m)})^{(n+1,n)}\} \end{aligned} \quad (5)$$

and we have defined for a complex matrix $[\mathbf{A}^{[n]}]_{i,j} = a_{i,j}^n \bar{a}_{i,j}^{n+1}$. Now, observe that we can define

$$\mathbf{F}_y \stackrel{\text{def}}{=} \mathbf{G}_y - \mathbf{R}_y \mathbf{K}_0 \Rightarrow \mathbf{F}_y = \mathbf{H} \tilde{\mathbf{B}}_{x,y} \quad (6)$$

³We use the same notation of [1] and [2] so that $\text{cum}[y: m, \bar{y}: k, z: l, \bar{z}: n]$ is the complex cumulant of order $(m; k; l; n)$ of the joint complex process $\{y, z\}$ ($\text{cum}[y, z: l, \bar{z}: n]$ is the cumulant of order $(1; 0; l; n)$), and $g(z, \bar{z})^{(l,n)} = (\partial^{(l+n)} / \partial u^l \partial \bar{u}^n) g(u, \bar{u})|_{u=\bar{z}}$.

which for a full-rank $\tilde{\mathbf{B}}_{x,y} = \sum_{n=1}^{\infty} (1/n!(n+1)!) (\mathbf{H}^T)^{[n]} \mathbf{K}_n \gamma_n$ is in the range of \mathbf{H} . The justification for the decomposition in the last equality of (6) comes from the fact that we have $\mathbf{R}_y = \mathbf{H}\mathbf{R}_x\mathbf{H}^H + \mathbf{R}_n$ [3], where \mathbf{R}_x is diagonal with elements σ_x^2 , and that (4) can be written

$$\begin{aligned} \mathbf{G}_y &= \mathbf{H}\mathbf{H}^H \mathbf{K}_0 \gamma_0 + \mathbf{H}\tilde{\mathbf{B}}_{x,y} + \mathbf{R}_n \mathbf{K}_0 \\ &= \mathbf{H}\mathbf{R}_x\mathbf{H}^H \mathbf{K}_0 + \mathbf{H}\tilde{\mathbf{B}}_{x,y} + \mathbf{R}_n \mathbf{K}_0 \end{aligned} \quad (7)$$

because $\gamma_0 = \sigma_x^2$, and $(\mathbf{H}^T)^{[n]}|_{n=0} = \mathbf{H}^H$. Substituting (7) and the expression for \mathbf{R}_y into (6), we obtain $\Gamma_y = \mathbf{H}\tilde{\mathbf{B}}_{x,y}$. Now, define $\mathbf{G1}_y$, $\mathbf{G2}_y$ as the HNL matrices obtained from two different ZMNL functions $g_1(u, \bar{u})$ and $g_2(u, \bar{u})$, respectively, and $\mathbf{K1}_n$, $\mathbf{K2}_n$ defined in (5) for $g_1(u, \bar{u})$ and $g_2(u, \bar{u})$, respectively. Another attractive decomposition is

$$\Gamma_y^{(1,2)} \stackrel{\text{def}}{=} \mathbf{G1}_y \mathbf{K2}_0 - \mathbf{G2}_y \mathbf{K1}_0 \Rightarrow \Gamma_y^{(1,2)} = \mathbf{H}\tilde{\mathbf{B}}_{x,y}^{(1,2)} \quad (8)$$

where

$$\tilde{\mathbf{B}}_{x,y}^{(1,2)} = \sum_{n=1}^{\infty} \frac{1}{n!(n+1)!} (\mathbf{H}^T)^{[n]} (\mathbf{K1}_n \mathbf{K2}_0 - \mathbf{K2}_n \mathbf{K1}_0) \gamma_n.$$

In fact, we can decompose $\mathbf{G1}_y$ and $\mathbf{G2}_y$ in (8) as in (7) and express $\Gamma_y^{(1,2)}$ as

$$\begin{aligned} \Gamma_y^{(1,2)} &= \mathbf{H}\mathbf{R}_x\mathbf{H}^H \mathbf{K1}_0 \mathbf{K2}_0 + \mathbf{H}\tilde{\mathbf{B}}_{x,y}^{(1)} \mathbf{K1}_0 \mathbf{K2}_0 \\ &\quad + \mathbf{R}_n \mathbf{K1}_0 \mathbf{K2}_0 - \mathbf{H}\mathbf{R}_x\mathbf{H}^H \mathbf{K2}_0 \mathbf{K1}_0 \\ &\quad - \mathbf{H}\tilde{\mathbf{B}}_{x,y}^{(2)} \mathbf{K2}_0 \mathbf{K1}_0 - \mathbf{R}_n \mathbf{K2}_0 \mathbf{K1}_0 \end{aligned} \quad (9)$$

where $\tilde{\mathbf{B}}_{x,y}^{(1)} = \sum_{n=1}^{\infty} (1/n!(n+1)!) (\mathbf{H}^T)^{[n]} \mathbf{K1}_n \gamma_n$, and $\tilde{\mathbf{B}}_{x,y}^{(2)} = \sum_{n=1}^{\infty} (1/n!(n+1)!) (\mathbf{H}^T)^{[n]} \mathbf{K2}_n \gamma_n$. Since $\mathbf{K2}_0$ and $\mathbf{K1}_0$ are diagonal, $\mathbf{K2}_0 \mathbf{K1}_0 = \mathbf{K1}_0 \mathbf{K2}_0$ so that $\Gamma_y^{(1,2)} = \mathbf{H}(\tilde{\mathbf{B}}_{x,y}^{(1)} - \tilde{\mathbf{B}}_{x,y}^{(2)}) = \mathbf{H}\tilde{\mathbf{B}}_{x,y}^{(1,2)}$. Γ_y (or $\Gamma_y^{(1,2)}$) is in the range of \mathbf{H} as long as the matrix $\tilde{\mathbf{B}}_{x,y}$ (or $\tilde{\mathbf{B}}_{x,y}^{(1,2)}$) satisfies a full-rank condition.

If such a full-rank condition was verified, then the subspace of Γ_y [see (6)] [or $\Gamma_y^{(1,2)}$, see (8)] would reveal the span of the columns of the filtering matrix \mathbf{H} , and as in [3], we would be able to identify the channels $h_i(l)$, $i = 1, 2, \dots, K$, $l = 0, 2, \dots, M$.

IV. DESCRIPTION OF PRACTICAL METHODS

There are two different practical approaches to verify if $\tilde{\mathbf{B}}_{x,y}$ [or $\tilde{\mathbf{B}}_{x,y}^{(1,2)}$] satisfies a full-rank condition. The first one is to choose ZMNL functions that possess only derivatives up to a certain order (for example, polynomial functions) so that the infinite series in $\tilde{\mathbf{B}}_{x,y}$ and $\tilde{\mathbf{B}}_{x,y}^{(1,2)}$ are zero beyond a certain n . Alternatively, an idea that gives more freedom and flexibility in the choice of the ZMNL function is to design a polynomial function that in a mean squared sense approximates the original $g(u, \bar{u})$. This idea was first introduced in [6]. Nonpolynomial functions that are attractive from a computational point of view are very difficult to analyze.⁴ An example is given by the following family of ZMNL functions [1]: $g_{HP}^{(L)}(x, \bar{x}) = e^{j \text{int}[(L/2\pi) \arg(x)](2\pi/L) + (\pi/L)}$, where $\text{int}[v]$ is the integer part of v . The analysis of such nonlinearities is quite complex, even for Gaussian processes [6]. Suppose we use $f_n(z, \bar{z}) = \sum_{k=1}^n a_k z^k \bar{z}^{k-1}$ as the n th-order approximation of $g(z, \bar{z})$, and we select the complex factors a_k to minimize the mean squared error (MSE) between the output of the true nonlinear function

⁴Some nonlinear functions are particularly important since the multiplications involved in the estimation based on sample statistics are trivial change-of-sign multiplications and/or swap the real and imaginary part so that the accumulation of complex samples remains the only operations to be performed.

and the approximated output. The approximation MSE for $g(z, \bar{z})$ using $f_n(z, \bar{z})$ is defined as

$$\begin{aligned} \text{MSE}_n(\mathbf{a}^z) &= E_Z \{|g(z, \bar{z}) - f_n(z, \bar{z})|^2\} \\ &= E_Z \{|g(z, \bar{z}) - \mathbf{a}^{z^T} \tilde{\mathbf{z}}|^2\} \end{aligned} \quad (10)$$

where

$$\begin{aligned} \mathbf{a}^z &= [a_1^z \quad a_2^z \quad \dots \quad a_n^z]^T, \\ \tilde{\mathbf{z}} &= [z \quad z^2 \bar{z} \quad z^3 \bar{z}^2 \quad \dots \quad z^n \bar{z}^{n-1}]^T \end{aligned}$$

and z is a generic complex random variable. The minimum $\text{MSE}_n(\mathbf{a}^z)$ with respect to \mathbf{a}^z can be obtained from

$$\hat{\mathbf{a}}^z = E_Z \{\tilde{\mathbf{z}} \tilde{\mathbf{z}}^H\}^\dagger E_Z \{\tilde{\mathbf{z}}^* g(z, \bar{z})\} \quad (11)$$

where we have used the symbol \dagger to identify the pseudoinverse of a matrix. Define $m_{z,l,m}^{(n)}$ as the l, m element of the matrix $E_Z \{\tilde{\mathbf{z}} \tilde{\mathbf{z}}^H\}^\dagger$. A quadratic approximation $f_2(z, \bar{z})$ of $g(z, \bar{z})$ gives particularly simple criteria to check the full-rank condition for $\tilde{\mathbf{B}}_{x,y}$ and $\tilde{\mathbf{B}}_{x,y}^{(1,2)}$. The following two propositions formalize the idea. The proofs of both propositions are in the Appendix.

Proposition 1: If, for the selected ZMNL function $g(u, \bar{u})$ and the input process \mathbf{x}_n , it is true that

$$\begin{aligned} m_{y(m),2,1}^{(2)} E_Y \{\bar{y}_t^{(m)} g(y_t^{(m)}, \bar{y}_t^{(m)})\} \\ + m_{y(m),2,2}^{(2)} E_Y \{\bar{y}_t^{(m)} |y_t^{(m)}|^2 g(y_t^{(m)}, \bar{y}_t^{(m)})\} \neq 0 \end{aligned} \quad (12)$$

for $m = 1, 2, \dots, K$, then the matrix $\Gamma_y = \mathbf{G}_y - \mathbf{R}_y \mathbf{K}_0$ has rank $M + N$, and the left singular vectors of its singular value decomposition perform a decomposition in orthogonal subspaces that is a basis for signal and noise subspaces of \mathbf{R}_y , provided that $(\mathbf{H}^T)^{[n]}|_{n=1}$ is full rank.

Proposition 2: If for two selected ZMNL functions $g_1(u, \bar{u})$, $g_2(u, \bar{u})$ it is true that

$$\begin{aligned} \frac{E_Y \{g_1(y_t^{(m)}, \bar{y}_t^{(m)}) |y_t^{(m)}|^2 \bar{y}_t^{(m)}\}}{E_Y \{g_1(y_t^{(m)}, \bar{y}_t^{(m)}) \bar{y}_t^{(m)}\}} \\ \neq \frac{E_Y \{g_2(y_t^{(m)}, \bar{y}_t^{(m)}) |y_t^{(m)}|^2 \bar{y}_t^{(m)}\}}{E_Y \{g_2(y_t^{(m)}, \bar{y}_t^{(m)}) \bar{y}_t^{(m)}\}} \end{aligned} \quad (13)$$

for $m = 1, 2, \dots, K$, and

$$m_{y(m),1,1}^{(2)} m_{y(m),2,2}^{(2)} \neq m_{y(m),1,2}^{(2)} m_{y(m),2,1}^{(2)} \quad (14)$$

for $m = 1, 2, \dots, K$, then the matrix $\Gamma_y^{(1,2)} = \mathbf{G1}_y \mathbf{K2}_0 - \mathbf{G2}_y \mathbf{K1}_0$ has rank $M + N$, and the left singular vectors of its singular value decomposition perform a decomposition in orthogonal subspaces, which is a basis for signal and noise subspaces of \mathbf{R}_y , provided that $(\mathbf{H}^T)^{[n]}|_{n=1}$ is full rank.

Observe that the full rank of $(\mathbf{H}^T)^{[n]}|_{n=1}$ is a mild assumption if \mathbf{H} satisfies **AS1**. The signal subspace corresponds to the span of the $M + N$ left singular vectors of Γ_y or $\Gamma_y^{(1,2)}$ associated with the $M + N$ singular values of the largest magnitude. Let $\mathbf{V} = [\mathbf{d}_0 \mathbf{d}_1 \dots \mathbf{d}_{KN-1}]$ be the ordered left singular vectors. If we define $\mathbf{S} = [\mathbf{d}_0 \mathbf{d}_1 \dots \mathbf{d}_{M+N-1}]$ and $\mathbf{D} = [\mathbf{d}_{M+N} \mathbf{d}_{M+N+1} \dots \mathbf{d}_{KN-1}]$, the columns of \mathbf{S} span the signal subspace, whereas the columns of \mathbf{D} span the noise subspace. The signal subspace is also the linear space spanned by the columns of the filtering matrix \mathbf{H} ; the columns of \mathbf{H} are orthogonal to any vector in the noise subspace. We can then follow [3] to turn this observation into a practical algorithm. The orthogonality condition is expressed formally as $\mathbf{d}_i^H \mathbf{H} = 0$ for $M + N \leq i \leq KN - 1$. Defining $q(\mathbf{h}) = \sum_{i=M+N}^{KN-1} |\mathbf{d}_i^H \mathbf{H}|^2$, it is easy to show, following [3], that $q(\mathbf{h}) = \mathbf{h}^H \mathbf{Q} \mathbf{h}$, where $\mathbf{h} = [\mathbf{h}_1^T \mathbf{h}_2^T \dots \mathbf{h}_K^T]^T$, with $\mathbf{Q} = \sum_{i=M+N}^{KN-1} \mathcal{D}_i \mathcal{D}_i^H$, and \mathcal{D}_i is

TABLE I
SUMMARY OF AN ALGORITHM USING PROPOSITION 1

-
- Estimate \mathbf{R}_y as $\hat{\mathbf{R}}_y = \frac{1}{T} \sum_{n=1}^T \mathbf{y}_n \mathbf{y}_n^H$.
 - Estimate \mathbf{G}_y as $\hat{\mathbf{G}}_y = \frac{1}{T} \sum_{n=1}^T \mathbf{y}_n \tilde{\mathbf{y}}_n^H$.
 - Estimate \mathbf{K}_0 as:
 - Estimate $\mathbf{K}_0^{(m)}$ as $\left[\hat{\mathbf{K}}_0^{(m)} \right]_{l,l} = \frac{1}{TN} \sum_{n=1}^{TN} \bar{g} \left(y_n^{(m)}, \bar{y}_n^{(m)} \right)^{(1,0)}$ for $l = 1, 2, \dots, N$ and $m = 1, 2, \dots, K$,
 - form $\hat{\mathbf{K}}_0$ as in (5).
 - Compute the SVD of $\hat{\Gamma}_y = \hat{\mathbf{G}}_y - \hat{\mathbf{R}}_y \hat{\mathbf{K}}_0 = \mathbf{U} \Sigma \mathbf{V}^H$ where $\mathbf{U} = [\mathbf{d}_0 \ \mathbf{d}_1 \dots \ \mathbf{d}_{KN-1}]$.
 - Form \mathcal{D}_i from \mathbf{d}_i for $i = M + N, \dots, KN - 1$ as indicated in [3].
 - Compute the estimate of the channel as the unit-norm eigenvector of $\mathbf{Q} = \sum_{i=M+N}^{KN-1} \mathcal{D}_i \mathcal{D}_i^H$ associated with the smallest eigenvalue in the EVD of \mathbf{Q} .
-

TABLE II
SUMMARY OF AN ALGORITHM USING PROPOSITION 2

-
- Estimate \mathbf{G}_{1y} as $\hat{\mathbf{G}}_{1y} = \frac{1}{T} \sum_{n=1}^T \mathbf{y}_n \tilde{\mathbf{y}}_n^H$ using $g_1(u, \bar{u})$.
 - Estimate \mathbf{G}_{2y} as $\hat{\mathbf{G}}_{2y} = \frac{1}{T} \sum_{n=1}^T \mathbf{y}_n \tilde{\mathbf{y}}_n^H$ using $g_2(u, \bar{u})$.
 - Estimate \mathbf{K}_{10} as:
 - Estimate $\mathbf{K}_{10}^{(m)}$ as $\left[\hat{\mathbf{K}}_{10}^{(m)} \right]_{l,l} = \frac{1}{TN} \sum_{n=1}^{TN} \bar{g}_1 \left(y_n^{(m)}, \bar{y}_n^{(m)} \right)^{(1,0)}$ for $l = 1, 2, \dots, N$ and $m = 1, 2, \dots, K$,
 - form $\hat{\mathbf{K}}_{10}$ as in (5).
 - Estimate \mathbf{K}_{20} as:
 - Estimate $\mathbf{K}_{20}^{(m)}$ as $\left[\hat{\mathbf{K}}_{20}^{(m)} \right]_{l,l} = \frac{1}{TN} \sum_{n=1}^{TN} \bar{g}_2 \left(y_n^{(m)}, \bar{y}_n^{(m)} \right)^{(1,0)}$ for $l = 1, 2, \dots, N$ and $m = 1, 2, \dots, K$,
 - form $\hat{\mathbf{K}}_{20}$ as in (5).
 - Compute the SVD of $\hat{\Gamma}_y^{(1,2)} = \hat{\mathbf{G}}_{1y} \hat{\mathbf{K}}_{20} - \hat{\mathbf{G}}_{2y} \hat{\mathbf{K}}_{10} = \mathbf{U} \Sigma \mathbf{V}^H$ where $\mathbf{U} = [\mathbf{d}_0 \ \mathbf{d}_1 \dots \ \mathbf{d}_{KN-1}]$.
 - Form \mathcal{D}_i from \mathbf{d}_i for $i = M + N, \dots, KN - 1$ as indicated in [3].
 - Compute the estimate of the channel as the unit-norm eigenvector of $\mathbf{Q} = \sum_{i=M+N}^{KN-1} \mathcal{D}_i \mathcal{D}_i^H$ associated with the smallest eigenvalue in the EVD of \mathbf{Q} .
-

the filtering matrix associated with \mathbf{d}_i (see [3]) defined similarly to \mathbf{H} . The problem of finding the minimum of the function $q(\mathbf{h})$ can also be expressed as a constrained optimization problem with a linear/quadratic constrain (see [3] for details). The algorithms are summarized in Tables I and II. The complexity of the algorithms is approximately equivalent to the subspace method of [3], but the EVD of the autocorrelation matrix is replaced by the SVD of Γ_y or $\Gamma_y^{(1,2)}$.

V. SIMULATIONS

Two different sets of simulations were performed. The first set is a test case that uses the same discrete-time channels as in [3] to allow a comparison. In the second set of simulations, we consider instead a TDMA system where the mobile-to-base frequency is in the 824–849 MHz band, and the transmitter employs QPSK at a data rate of 13 Kbps. The labeling system of Table III applies to the results of the simulations presented in this section. Methods SUBNL1, SUBNL2, and SUBNL3 satisfy the inequalities of Proposition 1, whereas SUBNL4 and SUBNL5 satisfy the inequalities of Proposition 2.

A. Comparison with the Subspace Method of [3]

The SNR in decibels is defined as $\text{SNR}_{\text{dB}} = 10 \log_{10}(\sigma_x^2 \|\mathbf{h}\|^2 / K \sigma_y^2)$, and the estimation MSE for 100 Monte

Carlo runs is defined as $\text{MSE}_{\text{dB}} = 20 \log_{10} \sqrt{\sum_{r=1}^{100} \|\mathbf{h} - \hat{\mathbf{h}}_r\|^2}$, where \mathbf{h} and $\hat{\mathbf{h}}_r$ are the true channel and the channel estimated at the r th computer run, respectively. All the simulations are performed using the same parameters of the simulation example given in [3]: $N = 10$, $K = 4$, and $M = 4$ are the discrete-time complex channels given in [3]. Fig. 1 shows results for white Gaussian noise with a sample size of $TN = 500$. It is evident that there is an advantage of 3–4 dB at low SNR using SUBNL4 or SUBNL5. Fig. 2 shows results for white Gaussian noise with a sample size of $TN = 1000$. 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 four-input four-output moving average model shown in Table IV. In colored noise, the SUBSOS method uses the whitening approach of [3], which assumes a very unrealistic knowledge of the noise correlation properties.

B. Cellular TDMA System Simulations

In this set of simulations, we evaluate the method in terms of bit error rate (BER) at the output of a maximum likelihood sequence estimation (MLSE) detector that uses the channel estimates obtained from the subspace methods. Antenna spacing is $\lambda/2$. We assume that five independent rays are received at the antenna ($N_p = 3$),

TABLE III
LABELING SYSTEM FOR THE FIGURES

- *SUBSOS* \Rightarrow Autocorrelation – based algorithm from [3]
- *SUBNL1* \Rightarrow Proposition 1 with $g(z, \bar{z}) = |z|^2 z$
- *SUBNL2* \Rightarrow Proposition 1 with $g(z, \bar{z}) = g_{HP}^{(4)}(z, \bar{z}) = csign(z)$
- *SUBNL3* \Rightarrow Proposition 1 with $g(z, \bar{z}) = g_{HP}^{(8)}(z, \bar{z}) = \frac{1}{\sqrt{2+\sqrt{2}}} (csign(z) + e^{j\frac{\pi}{4}} csign(z e^{-j\frac{\pi}{4}}))$
- *SUBNL4* \Rightarrow Proposition 2 with :
 - $g_1(z, \bar{z}) = g(z, \bar{z}) = |z|^2 z$
 - $g_2(z, \bar{z}) = g_{HP}^{(4)}(z, \bar{z}) = csign(z)$
- *SUBNL5* \Rightarrow Proposition 2 with :
 - $g_1(z, \bar{z}) = g(z, \bar{z}) = |z|^2 z$
 - $g_2(z, \bar{z}) = g_{HP}^{(8)}(z, \bar{z}) = \frac{1}{\sqrt{2+\sqrt{2}}} (csign(z) + e^{j\frac{\pi}{4}} csign(z e^{-j\frac{\pi}{4}}))$

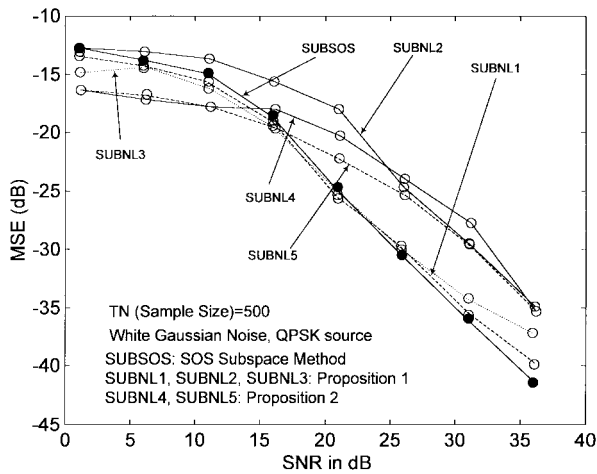


Fig. 1. Results of simulations. Mean squared error (MSE) versus SNR for white Gaussian noise. Black dots indicate performance of the subspace method based on the correlation matrix of [3]. The estimation of the statistics is made using 500 samples.

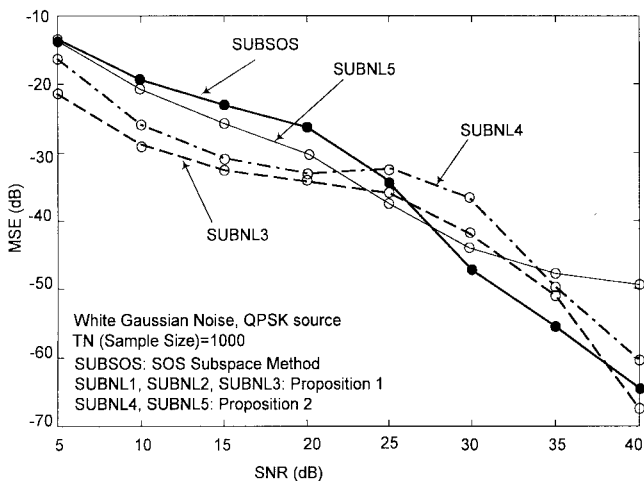


Fig. 2. Mean squared error versus SNR of the channel estimates for white Gaussian noise. Black dots indicate performance of the subspace method based on the correlation matrix of [3]. The estimation of the statistics is made using 1000 samples.

where each ray is characterized by $P_1 = 6$, $P_2 = 20$, $P_3 = 18$, $P_4 = 15$, and $P_5 = 20$ paths. The powers of the delayed paths, that is, $E\{|\rho_{i,m}|^2\}$ and the delays $\tau_{i,m}$ for $m = 1, 2, \dots, P_i$ and

$i = 1, 2, 3, 4, 5$, are distributed according to the power delay profile, which is constituted of two clusters with one-sided exponential delay.⁵ Each channel is specified uniquely by the delay interval D_i between the two clusters, which we normalize with respect to the symbol period T . The angles of arrival of the five rays are $\theta_1 = -10^\circ$, $\theta_2 = 10^\circ$, $\theta_3 = 15^\circ$, $\theta_4 = 65^\circ$, and $\theta_5 = -45^\circ$. The normalized (to symbol period) delay interval of the first path is 0.15, whereas the other paths are specified in each test case. The number of elements of the antenna is four. The support of each (FIR) channel (the impulse response of each path) is truncated to five. Speed specifies time varying characteristics of the channels.⁶ The detector is an MLSE (see [5]) operated on a vector output (vector-MLSE). Particularly, this last algorithm was implemented by means of the well-known M -algorithm with $\bar{M} = 8$ and $\bar{M} = 16$ [7]. The M -algorithm in the case of heavily dispersive channels and multichannel signals is less computationally intensive than the full search of the Viterbi algorithm (for more details on the use of the M -algorithm in a multisensor antenna system, see [7]). In Fig. 4, the BER versus SNR is shown at different delay spread values. Observe that at low SNR, the advantage of the nonlinear methods is evident.

VI. CONCLUSIONS

New methods have been presented for the blind identification of multichannel FIR models. The method exploits some of the ideas in [1] and [3], outlying a new approach to the problem. Particularly, the theoretical results of [1] have been extended from the simple case of exponentials in noise to the multichannel FIR model to show that subspace information is retained by HNL matrices also in this

⁵The power delay profile is

$$\Phi_i(\tau) = \begin{cases} e^{-\tau}, & 0 \leq \tau \leq D_i \\ 0.5e^{-(D_i-\tau)}, & D_i \leq \tau < 2D_i \end{cases}$$

where τ and D_i are expressed in microseconds. The values of the actual delays for $i = 1, 2, 3, 4$, and 5 are obtained by uniform sampling of $\Phi_i(\tau)$, that is, $\tau_{i,m} = m(2D_i/P_i)$; then, the power of the rays are $E\{|\rho_{i,m}|^2\} = \Phi_i(\tau_{i,m})$.

⁶Doppler frequency is related through wavelength λ to vehicle motion. The model used is based on the wide sense stationary uncorrelated scattering (WSSUS) assumption. According to this model, we generated the time evolution in 1 s for the discrete-time channels obtained by symbol sampling for a vehicle moving at the speed of 60 m/hr and for delay intervals D_1 , D_2 , and D_3 specified in each test. Then, we obtained 100 channel configurations by taking the channel realization obtained at time instants $t_n = n/100$ s with $n = 1, 2, \dots, 100$. For each of these channel realizations, we simulated the performance of the proposed algorithm by running input data to the particular obtained channel. The results were then averaged over the different results obtained for each channel realization.

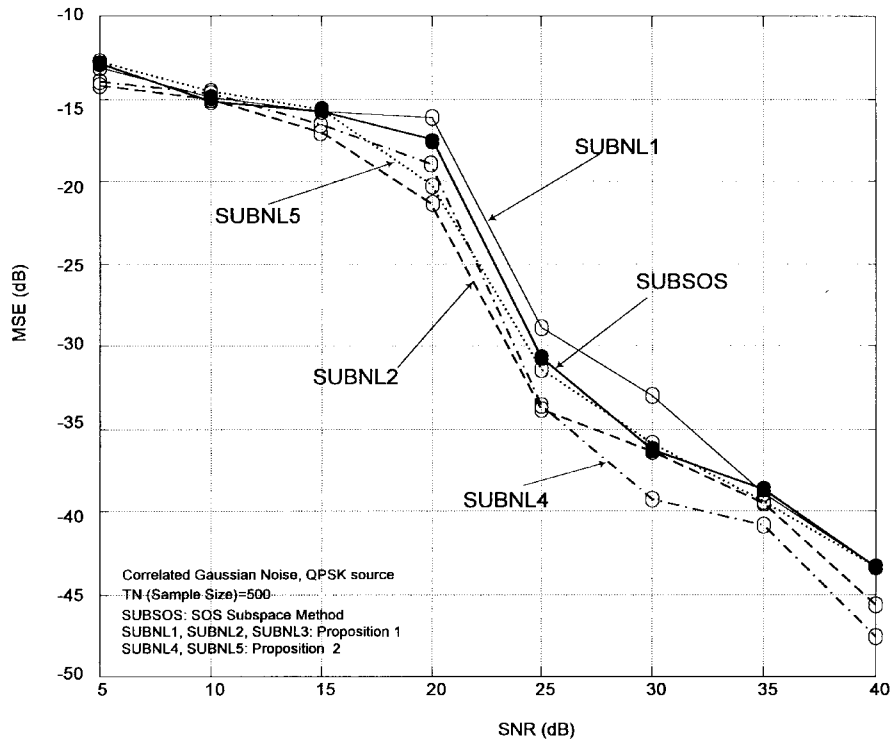


Fig. 3. Mean squared error versus SNR of the channel estimates for correlated Gaussian noise. Black dots indicate performance of the subspace method based on the correlation matrix of [3]. The estimation of the statistics is made using 500 samples.

case. In fact, we introduced two matrices Γ_y or $\Gamma_y^{(1,2)}$ that can be used to identify the column space of the filtering matrix \mathbf{H} exactly as in [3], where the authors used the autocorrelation matrix \mathbf{R}_y . Some important assumptions must be satisfied, and we showed that simple polynomial approximations of the ZMNL function simplify the conditions to inequalities that are satisfied by a wide class of nonlinearities, hence resulting in a very attractive practical approach. The results of computer simulations were shown to emphasize the improvement of performance with respect to the subspace method of [3], especially in a colored and/or high noise environment. Bit error rate results in a more realistic cellular communication system simulation show the effectiveness of the method.

APPENDIX

Proof of (4) and (5): Using (3), the generic k, i element of the matrix $\mathbf{G}_y^{(l,m)} = E_Y \{ \mathbf{y}_n^{(l)} \mathbf{y}_n^{(m)H} \}$ is given by

$$[\mathbf{G}_y^{(l,m)}]_{k,i} = \sum_{n=0}^{\infty} \frac{1}{(n+1)!n!} \bar{E}_Y \cdot \{ g(y_{t-i+1}^{(m)}, \bar{y}_{t-i+1}^{(m)})^{(n+1,n)} \} \times \text{cum}[y_{t-k+1}^{(l)}, y_{t-i+1}^{(m)}; n, \bar{y}_{t-i+1}^{(m)}; n+1]. \quad (15)$$

Due to the linearity property of cumulants and the independence of the source symbols, we can write for $n > 0$

$$\begin{aligned} & \text{cum}[y_{t-k+1}^{(l)}, y_{t-i+1}^{(m)}; n, \bar{y}_{t-i+1}^{(m)}; n+1] \\ &= \sum_{k_1} h_l(k_1 - k + 1) [h_m(k_1 - i + 1)]^n \\ & \quad \times [\bar{h}_m(k_1 - i + 1)]^{n+1} \\ & \quad \cdot \text{cum}[x_t; n+1, \bar{x}_t; n+1] \end{aligned} \quad (16)$$

(since cumulants of a Gaussian process vanish for order greater than two), whereas for $n = 0$

$$\begin{aligned} & \text{cum}[y_{t-k+1}^{(l)}, y_{t-i+1}^{(m)}; n, \bar{y}_{t-i+1}^{(m)}; n+1] \\ &= \sum_{k_1} h_l(k_1 - k + 1) \bar{h}_m(k_1 - i + 1) \sigma_x^2 \\ & \quad + [\mathbf{R}_n^{(l,m)}]_{k,i} \end{aligned} \quad (17)$$

where $\sigma_x^2 = E_x \{ x_n \bar{x}_n \}$, and

$$[\mathbf{R}_n^{(l,m)}]_{k,i} = E_n \{ \mathbf{n}_n^{(l)} \mathbf{n}_n^{(m)H} \}.$$

Substituting (16) and (17) into (15), we obtain the component-wise version of (4) and (5)

$$\begin{aligned} [\mathbf{G}_y^{(l,m)}]_{k,i} &= \sum_{n=0}^{\infty} \frac{1}{(n+1)!n!} \text{cum}[x_t; n+1, \bar{x}_t; n+1] \\ & \quad \times \{ \mathbf{H}_l(\mathbf{H}_m^T)^{[n]} \bar{E}_Y \{ g(y_{t-i+1}^{(m)}, \bar{y}_{t-i+1}^{(m)})^{(n+1,n)} \} \}_{k,i} \\ & \quad + K_g^{(m)} [\mathbf{R}_n^{(l,m)}]_{k,i} \end{aligned}$$

where $K_g^{(m)} = \bar{E}_Y \{ g(y_t^{(m)}, \bar{y}_t^{(m)})^{(1,0)} \}$.

Proof of Proposition 1: First, observe that the $\hat{\mathbf{a}}^{y^{(m)}}$ minimizing $MS E_2(\mathbf{a}^{y^{(m)}})$ is

$$\begin{aligned} \hat{\mathbf{a}}^{y^{(m)}} &= E_Y \{ \tilde{\mathbf{z}}_y^{(m)} \tilde{\mathbf{z}}_y^{(m)H} \}^\dagger E_Y \{ \tilde{\mathbf{z}}_y^{(m)*} g(y_t^{(m)}, \bar{y}_t^{(m)}) \} \\ &= \begin{bmatrix} m_{y^{(m)},1,1}^{(2)} & m_{y^{(m)},1,2}^{(2)} \\ m_{y^{(m)},2,1}^{(2)} & m_{y^{(m)},2,2}^{(2)} \end{bmatrix} E_Y \{ \tilde{\mathbf{z}}_y^{(m)*} g(y_t^{(m)}, \bar{y}_t^{(m)}) \} \end{aligned} \quad (18)$$

where

$$\begin{aligned} \hat{\mathbf{a}}^{y^{(m)}} &= [\hat{a}_1^{y^{(m)}}, \hat{a}_2^{y^{(m)}}]^T \\ \tilde{\mathbf{z}}_y^{(m)} &= [y_t^{(m)}, y_t^{(m)2} \bar{y}_t^{(m)}]^T. \end{aligned}$$

Now, define $\mathbf{K}_n^{(f)}$ as in (5) and $\tilde{\mathbf{B}}_{x,y}^{(f)}$ as in (4) for $f_2(u, \bar{u})$. From (6), $\Gamma_y = \mathbf{H} \tilde{\mathbf{B}}_{x,y}$ and since $\mathbf{K}_n^{(f)} \simeq \mathbf{K}_n$, we have that

TABLE IV
MA MODEL USED TO CORRELATE THE DISCRETE-TIME NOISE VECTOR SAMPLES ($\mathbf{B}(1) = \mathbf{B}(3) = \mathbf{0}$)

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

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

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

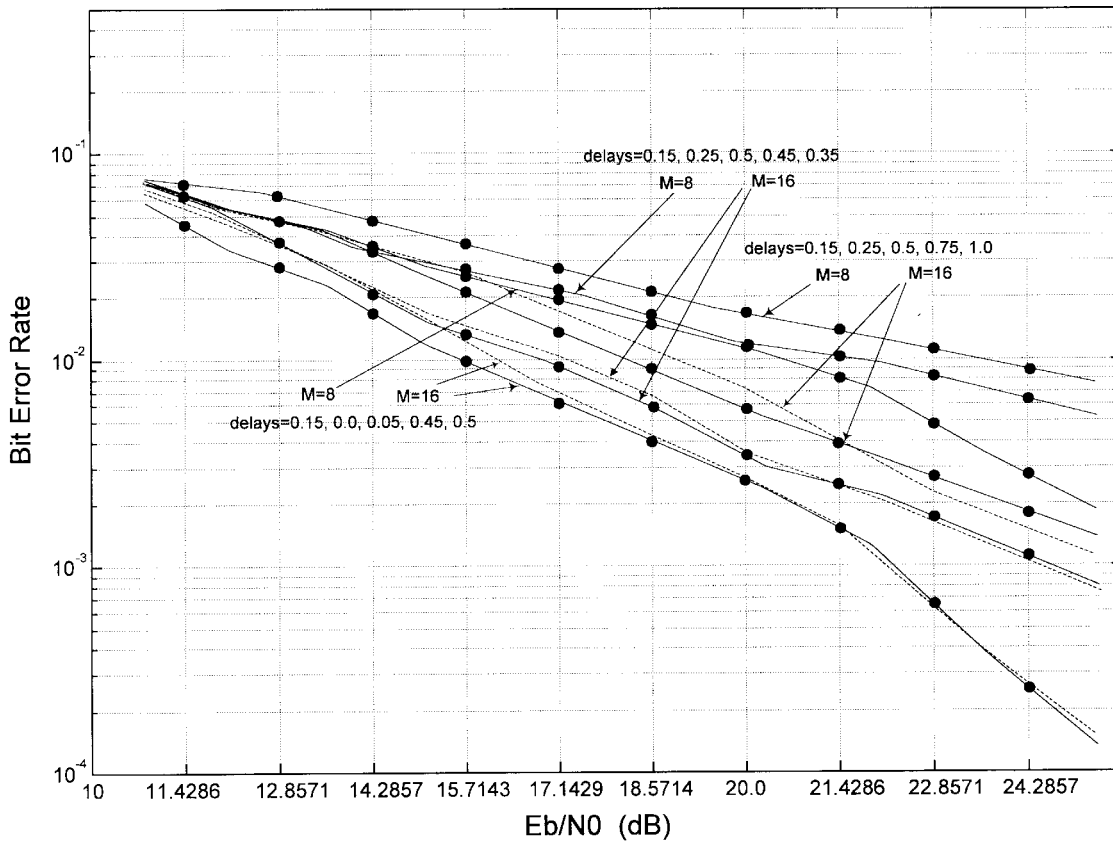


Fig. 4. Results of simulations in terms of bit error rate for different multipath power delay profiles. The number of samples used to estimate the channels is 5000. The channel estimation method is SUBNL4 for the solid curves. The subspace method (SUBSOS) results are shown by means of dashed curves. The number of paths is the value of M of the M algorithm. Indicated delays for the double exponential power delay profile are normalized to symbol period.

$\tilde{\mathbf{B}}_{x,y} \simeq \tilde{\mathbf{B}}_{x,y}^{(f)} = \sum_{n=1}^{\infty} (1/n!(n+1)!) (\mathbf{H}^T)^{[n]} \mathbf{K}_n^{(f)} \gamma_n$ must be full rank. The terms $n \geq 2$ in $\tilde{\mathbf{B}}_{x,y}^{(f)}$ are 0; in fact, $f_2(y_t^{(m)}; \bar{y}_t^{(m)})^{(1,0)} = 2\hat{a}_2^{y^{(m)}} |y_t^{(m)}|^2 + \hat{a}_1^{y^{(m)}}$, $f_2(y_t^{(m)}; \bar{y}_t^{(m)})^{(2,1)} = 2\hat{a}_2^{y^{(m)}}$, and of course, $f_2(y_t^{(m)}; \bar{y}_t^{(m)})^{(n+1,n)} = 0$ for $n \geq 2$ so that we have $\tilde{\mathbf{B}}_{x,y} \simeq (\mathbf{H}^T)^{[1]} \mathbf{K}_1^{(f)} (\gamma_1/2)$. If $(\mathbf{H}^T)^{[n]}|_{n=1}$ is full rank (as assumed), the rank condition required by $\tilde{\mathbf{B}}_{x,y}$ reduces to a full-rank condition on the matrix $\mathbf{K}_1^{(f)}$. Particularly, since $\mathbf{K}_1^{(f)}$ is diagonal, it is required that all its diagonal elements are different from 0. This condition is

expressed in (12) because the generic diagonal element of $\mathbf{K}_1^{(f)}$ is $[\mathbf{K}_1^{(f)}]_{j,j} = \overline{E}_Y \{ f_2(y_t^{(m)}; \bar{y}_t^{(m)})^{(2,1)} \}$, and $f_2(y_t^{(m)}; \bar{y}_t^{(m)})^{(2,1)}$ is given by $f_2(y_t^{(m)}; \bar{y}_t^{(m)})^{(2,1)} = 2\hat{a}_2^{y^{(m)}}$, with $\hat{a}_2^{y^{(m)}}$ given in (18).

Proof of Proposition 2: From (8), we have $\mathbf{I}_y^{(1,2)} = \mathbf{H} (\tilde{\mathbf{B}}_{x,y}^{(1,2)} - \tilde{\mathbf{B}}_{x,y}^{(f)}) = \mathbf{H} \tilde{\mathbf{B}}_{x,y}^{(1,2)}$, which satisfies the rank condition if matrix $\tilde{\mathbf{B}}_{x,y}^{(1,2)}$ is full rank. Define $f_{12}(u, \bar{u})$ and $f_{22}(u, \bar{u})$ as quadratic polynomial approximations of $g_1(u, \bar{u})$ and $g_2(u, \bar{u})$, respectively, $\mathbf{K}_1^{(f)}$, $\mathbf{K}_2^{(f)}$ for $f_{12}(u, \bar{u})$, $f_{22}(u, \bar{u})$, as in (5)

and $\widetilde{\mathbf{B}}\mathbf{1}_{x,y}^{(f)}$, $\widetilde{\mathbf{B}}\mathbf{2}_{x,y}^{(f)}$ as in (4) for $f_{12}(u, \bar{u})$, $f_{22}(u, \bar{u})$, respectively. In addition, $\widetilde{\mathbf{B}}\mathbf{1}_{x,y}^{(f)} - \widetilde{\mathbf{B}}\mathbf{2}_{x,y}^{(f)} = \mathbf{B}\mathbf{1}\mathbf{2}_{x,y}^{(f)}$. Since the expansion $\sum_{n=1}^{\infty} (1/n!(n+1)!) (\mathbf{H}^T)^{[n]} (\mathbf{K}\mathbf{1}_n^{(f)} \mathbf{K}\mathbf{2}_0^{(f)} - \mathbf{K}\mathbf{2}_n^{(f)} \mathbf{K}\mathbf{1}_0^{(f)}) \gamma_n$ is 0 for $n \geq 2$ (see proof of Proposition 1), and $\mathbf{K}\mathbf{1}_n^{(f)} \simeq \mathbf{K}\mathbf{1}_n$, $\mathbf{K}\mathbf{2}_n^{(f)} \simeq \mathbf{K}\mathbf{2}_n$, we can write

$$\begin{aligned} \widetilde{\mathbf{B}}\mathbf{1}\mathbf{2}_{x,y} &\simeq \widetilde{\mathbf{B}}\mathbf{1}\mathbf{2}_{x,y}^{(f)} \\ &= \frac{1}{2} (\mathbf{H}^T)^{[1]} (\mathbf{K}\mathbf{1}_1^{(f)} \mathbf{K}\mathbf{2}_0^{(f)} - \mathbf{K}\mathbf{2}_1^{(f)} \mathbf{K}\mathbf{1}_0^{(f)}) \gamma_1 \end{aligned} \quad (19)$$

and express the rank condition $[(\mathbf{H}^T)^{[n]}]_{n=1}$ is full rank by assumption] as the generic diagonal element of $\mathbf{K}\mathbf{1}_1^{(f)} \mathbf{K}\mathbf{2}_0^{(f)} - \mathbf{K}\mathbf{2}_1^{(f)} \mathbf{K}\mathbf{1}_0^{(f)}$ being different from 0, that is

$$\begin{aligned} \overline{E}_Y \{f_{12}(y_t^{(m)}, \bar{y}_t^{(m)})^{(2,1)}\} \overline{E}_Y \{f_{22}(y_t^{(m)}, \bar{y}_t^{(m)})^{(1,0)}\} \\ \neq \overline{E}_Y \{f_{22}(y_t^{(m)}, \bar{y}_t^{(m)})^{(2,1)}\} \overline{E}_Y \{f_{12}(y_t^{(m)}, \bar{y}_t^{(m)})^{(1,0)}\} \end{aligned} \quad (20)$$

for $m = 1, 2, \dots, K$. Substituting into (20) the explicit expressions for $f_{12}(y_t^{(m)}; \bar{y}_t^{(m)})^{(1,0)}$, $f_{22}(y_t^{(m)}; \bar{y}_t^{(m)})^{(1,0)}$, $f_{12}(y_t^{(m)}; \bar{y}_t^{(m)})^{(2,1)}$, $f_{22}(y_t^{(m)}; \bar{y}_t^{(m)})^{(2,1)}$ (as in the proof of Proposition 1), and then using (18), with $g_1(u, \bar{u})$, $g_2(u, \bar{u})$, we obtain conditions (13) and (14).

REFERENCES

- [1] G. Jacovitti and G. Scarano, "Hybrid nonlinear moments in array processing and spectrum analysis," *IEEE Trans. Signal Processing*, vol. 42, July 1994.
- [2] G. Scarano, "Cumulant series expansion of hybrid nonlinear moments of complex random variables," *IEEE Trans. Signal Processing*, vol. 39, Apr. 1991.
- [3] 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, Feb. 1995.
- [4] S. Anderson, M. Millnert, M. Viberg, and B. Wahlberg, "An adaptive array for mobile communication systems," *IEEE Trans. Veh. Technol.*, vol. 40, Feb. 1991.
- [5] J. G. Proakis, *Digital Communications*. New York: McGraw-Hill, 1989.
- [6] J. S. Bendat, *Nonlinear System Analysis and Identification from Random Data*. New York: Wiley, 1990.
- [7] P. Jung, "Performance evaluation of a novel M-detector for coherent receiver antenna diversity in a GSM-type mobile radio system," *IEEE J. Select. Areas Commun.*, vol. 13, pp. 80–88, Jan. 1995.

A Wavelet Time-Scale Deconvolution Filter Design for Nonstationary Signal Transmission Systems through a Multipath Fading Channel

Bor-Sen Chen, Yue-Chiech Chung, and Der-Feng Huang

Abstract—This study attempts to develop a time-scale deconvolution filter for optimal signal reconstruction of nonstationary processes with a stationary increment transmitted through a multipath fading and colored noisy channel with stochastic tap coefficients. A deconvolution filter based on wavelet analysis/synthesis filter bank is proposed to solve this problem via a three-stage filter bank. A fractal signal transmitted through a multipath fading and noisy channel is provided to demonstrate the design procedure's effectiveness and to exhibit the signal reconstruction performance of the proposed optimal time-scale deconvolution filter.

Index Terms—Fading channels, fractal signal, wavelet.

I. INTRODUCTION

There has been considerable research on the signal reconstruction (deconvolution) of a stationary signal in a time-invariant channel [1]–[6], [10], [16], [17]. Signal transmission is seriously affected by the nonstationary environment, i.e., the so-called multipath fading effect [18], [25], [28]. Currently, efficiently treating the reconstruction problem in the multipath fading channel case is a relatively difficult task. Adaptive filtering algorithms have been developed to estimate the coefficients of a channel to update the reconstruction filter [9], [21], [27]. However, a large amount of computation is necessary to update the adaptive algorithm in every sampling time. Furthermore, if the transmission channel is time-varying fading and signals are nonstationary, constructing a real-time adaptive deconvolution filter is not easy.

Chen and Lin [3] addressed the restoration problem of fractal signal [7], [8], [12], [14], [23] on a linear time-invariant channel by the wavelet filter bank method. However, multipath fading in the transmission channel with stochastic tap coefficients and colored noise is considered in this work. To the best of our knowledge, there is still not an efficient way of treating the important deconvolution problem of a nonstationary signal transmitted through a multipath time-varying fading channel.

In this framework, we prove that the output of a fading channel with a stationary increment process as the input is still a stationary increment property. Then, a wavelet analysis filter bank is first developed to transform the transmitted signal from the time domain to the time-scale domain. Then, in each subband of the wavelet filter bank, a Wiener filter is designed to optimally reconstruct the transmitted signal in the time-scale domain. Finally, a wavelet synthesis filter bank is developed to transform the estimate of the transmitted sequence in the time-scale domain back to the estimate of transmitted sequence $\hat{s}(t)$ in time domain (see Fig. 1).

Manuscript received November 6, 1996; revised August 23, 1998. This work was supported by the National Science Council under Contract NSC 87-2218-E-007-032. The associate editor coordinating the review of this paper and approving it for publication was Dr. Xiang-Gen Xia.

The authors are with the Department of Electrical Engineering, National Tsing Hua University, Hsin-Chu, Taiwan, R.O.C. (e-mail: bschen@moti.ee.nthu.edu.tw).

Publisher Item Identifier S 1053-587X(99)03243-2.