SYMBOL-RATE SECOND-ORDER BLIND CHANNEL ESTIMATION IN CODED TRANSMISSIONS

Javier Villares

Technical University of Catalonia (UPC)
Signal Theory and Communications Department
Jordi Girona, 1-3 08034 Barcelona (Spain)
javi@gps.tsc.upc.edu

ABSTRACT

This paper proves that second-order symbol-rate blind channel identification is feasible if the transmitted bit stream is encoded and the channel estimator conveniently exploits the code redundancy. This result is surprising because most error correcting codes do not alter the symbols correlation. Despite this, it can be shown that second-order methods can exploit the symbols fourth-order statistical information, which contains part of the code redundant structure. Analytical and numerical results are presented for the BPSK (Binary Phase Shift Keying) modulation and simple error-correcting block codes. The extension to more sophisticated coding and modulation schemes seems straightforward although computationally intensive.

***Index Terms*—** Parameter estimation, channel coding

1. INTRODUCTION

In single antenna communication systems, second-order blind channel estimation requires to oversample the received cyclostationary signal to have multiple symbol-rate versions of the channel impulse response (CIR) [1]. In that case, classical subspace methods [1][2], (cyclic) correlation matching methods [3], or a mixing of both approaches [4], can be used to obtain consistent CIR estimates under mild identifiability constraints. Otherwise, if the signal is processed at one sample per symbol, the general belief is that higher-order techniques are needed to identify the channel. In this paper, it is proved that second-order channel identification is still possible at one sample per symbol provided that the transmitted bit stream is coded using standard error control codes.

Although in most cases the encoder does not modify the correlation of the transmitted symbols, it introduces statistical dependency in the coded sequence that is manifested in the higher-order statistics of the coded symbols. Apparently, second-order techniques could not use this higher-order information. However, it was proved by the authors in [5][6][7] that optimal second-order estimators can exploit the symbols fourth-order statistical information.

Thus far, this information was found to be relevant in some multivariate estimation problems if the transmitted symbols had constant modulus [8][9]. In this contribution, it is shown that the fourth-order cumulants (kurtosis) of the transmitted symbols, which appear naturally in the formulation of the optimal second-order estimator, provide sufficient information about the code redundancy to yield accurate CIR estimates without the need of oversampling the received signal.

Finally, notice that the code redundancy is also exploited by some iterative schemes in the literature based on the turbo principle [10][11]. However, they are inherently higher-order methods performing joint decoding and channel estimation.

This work has been financed by the Spanish/Catalan Science and Technology Commissions and FEDER funds from the European Commission: TEC2006-06481/TCM, TEC2004-04526, TIC2003-05482, and 2005SGR-00639.

2. SIGNAL MODEL

In this section, the proposed framework for channel estimation is presented in detail. The symbol-rate samples at the matched filter output are given by

$$y(n) = \sum_{l=0}^{L-1} h_l x(n-l) + w(n) \tag{1}$$

where h_l is l-th element of the complex-valued CIR, $x(n) \in \{-1, 1\}$ the sequence of transmitted uncorrelated *coded* BPSK symbols and $w(n)$ the zero-mean white Gaussian noise term of known variance σ_w^2. For convenience, we consider that $\sum_{l=0}^{L-1} \|h_l\|^2 = 1$ and, therefore, σ_w^2 is the inverse of the signal-to-noise ratio (SNR) at the matched filter output.

Let us consider that $x(n)$ is the output of a given (N, K) block code. In that case, $x(n)$ is the concatenation of BPSK-modulated codewords of length N. The binary version of these codewords is computed using bitwise operations as follows:

$$\mathbf{x}' = \mathbf{s}' \cdot \mathbf{G} \tag{2}$$

where $\mathbf{G}$ is the code generating matrix and $\mathbf{s}'$ is the block of the K independent bits entering into the encoder. Notice that the prime sign $(\,)'$ is used to denote the binary representation of the associated BPSK symbol, i.e.,

$$\begin{aligned} \mathbf{x} &= 1 - 2\mathbf{x}' \\ \mathbf{s} &= 1 - 2\mathbf{s}'. \end{aligned} \tag{3}$$

Although the considered encoder is linear in the Galois field, it is performing a nonlinear transformation of the input data symbols. For example, if we consider the following $(6, 3)$ encoder

$$\mathbf{G} = \begin{bmatrix} 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 & 1 \end{bmatrix}, \tag{4}$$

the codewords will be constructed as follows:

$$\mathbf{x} = \begin{bmatrix} s_1 & s_2 & s_3 & s_1 s_3 & s_1 s_2 & s_2 s_3 \end{bmatrix}, \tag{5}$$

where s_n stands for the n-th element of vector $\mathbf{s}$. Therefore, the modulus-2 binary sum corresponds to the product of the associated BPSK symbols. In the proposed example, the encoder is implicitly introducing a quadratic transformation but, in general, the order of the nonlinearity will be given by the maximum number of ones in the columns of $\mathbf{G}$.

The channel estimator will process consecutive codewords of length N samples, having the following expression:

$$\mathbf{y} = \mathbf{H}\tilde{\mathbf{x}} + \mathbf{w} \tag{6}$$

where $\mathbf{H}$ is the matrix carrying out the CIR convolution and $\tilde{\mathbf{x}}$ is the column vector that contains the $N+L-1$ BPSK symbols of the current codeword ($\mathbf{x}$) and the previous interfering codewords. For instance, if the CIR has length $L = 3$ and the encoding in (4) is considered, the content of matrix $\mathbf{H}$ would be

$$\mathbf{H} = \begin{bmatrix} h_2 & h_1 & h_0 & 0 & 0 & 0 & 0 & 0 \\ 0 & h_2 & h_1 & h_0 & 1 & 1 & 0 & 0 \\ 0 & 0 & h_2 & h_1 & h_0 & 0 & 0 & 0 \\ 0 & 0 & 0 & h_2 & h_1 & h_0 & 0 & 0 \\ 0 & 0 & 0 & 0 & h_2 & h_1 & h_0 & 0 \\ 0 & 0 & 0 & 0 & 0 & h_2 & h_1 & h_0 \end{bmatrix}, \tag{7}$$

involving $L-1 = 2$ interfering symbols from the previous codeword (first two columns of $\mathbf{H}$) and the next $N = 6$ symbols from the current codeword.

The proposed estimator assumes word synchronization, meaning that the beginning of the codewords is known. In addition, symbol synchronization is assumed although it is actually optional because the sampling errors could be included into the CIR.

3. SECOND-ORDER CHANNEL ESTIMATION

The real and imaginary parts of the channel coefficients is stacked in the parameter vector $\boldsymbol{\Theta}$ as follows

$$\boldsymbol{\Theta} = \begin{bmatrix} \Re(h_0) & \dots & \Re(h_{L-1}) & \Im(h_1) & \dots & \Im(h_{L-1}) \end{bmatrix}^T \tag{8}$$

where the superscript T denotes transposition. Notice that the first channel tap is assumed real-valued to overcome the inherent phase ambiguity of second-order methods. The phase ambiguity could be solved exploiting the non-circularity of the BPSK modulation [12]. However, this approach is out of the scope of this work.

The vector of parameters $\boldsymbol{\Theta}$ can be estimated from the sample covariance matrix

$$\hat{\mathbf{R}} = \mathbf{y}\mathbf{y}^H \tag{9}$$

where the superscript H denotes transpose conjugate. Again, the non-circular property is not studied in the paper and, thus, the improper sample covariance matrix $\mathbf{y}\mathbf{y}^T$ is not considered for channel estimation.

A closed-loop implementation is proposed for the channel estimator allowing channel tracking in time-varying scenarios [13]. A closed-loop scheme is composed of a discriminator (error detector) and a loop filter. The discriminator is aimed to estimate the error between the actual CIR and the CIR estimate in the last iterate, say $\hat{\boldsymbol{\Theta}}$. After acquisition, in the steady-state, the estimation error $\mathbf{e} = \hat{\boldsymbol{\Theta}} - \boldsymbol{\Theta}$ has zero mean and its variance is determined by the noise equivalent loop bandwidth [13].

The optimal sample covariance based discriminator, which minimizes its variance in the steady-state, was formulated in [5] and it is particularized now for the CIR estimation problem:

$$\hat{\mathbf{e}} = (\mathbf{D}^H\mathbf{Q}^{-1}\mathbf{D})^{-1}\mathbf{D}^H\mathbf{Q}^{-1}\text{vec}(\hat{\mathbf{R}} - \mathbf{R}) \tag{10}$$

with

$$\begin{aligned} \mathbf{R} &= E\left\{\hat{\mathbf{R}}\right\} = \mathbf{H}\mathbf{H}^H + \sigma_w^2\mathbf{I}_N \\ \mathbf{D} &= \left[\text{vec}(\tfrac{\partial \mathbf{R}}{\partial \theta_1}) \ \dots \ \text{vec}(\tfrac{\partial \mathbf{R}}{\partial \theta_{2L-1}})\right] \\ \mathbf{Q} &= E\left\{\text{vec}(\hat{\mathbf{R}} - \mathbf{R})\text{vec}^H(\hat{\mathbf{R}} - \mathbf{R})\right\} \\ &= \mathbf{R}^* \otimes \mathbf{R} + (\mathbf{H}^* \otimes \mathbf{H})\mathbf{K}(\mathbf{H}^* \otimes \mathbf{H})^H \end{aligned}$$

where θ_n is the n-th entry in $\boldsymbol{\Theta}$, vec( ) denotes column-wise vectorization, $\otimes$ stands for the Kronecker product, $\mathbf{I}_N$ is the N-by-N identity matrix and $\mathbf{K}$ is the kurtosis matrix containing the fourth-order cumulants of vector $\tilde{\mathbf{x}}$:

$$\mathbf{K} = E\left\{\text{vec}(\tilde{\mathbf{x}}\tilde{\mathbf{x}}^H - \mathbf{I}_{N+L-1})\text{vec}^H(\tilde{\mathbf{x}}\tilde{\mathbf{x}}^H - \mathbf{I}_{N+L-1})\right\} \tag{11}$$

This matrix plays a prominent role in this contribution because it supplies the whole fourth-order statistical information about the received symbols that second-order estimators are able to exploit. It was shown in [6][7] that second-order methods can be improved for medium-to-high SNR by means of matrix $\mathbf{K}$ if and only if this matrix has some eigenvalues equal to -1. This property was verified in case of constant modulus alphabets in [5][6]. In addition, the main conclusion of this paper is that the rank of the subspace of $\mathbf{K}$ associated to the eigenvalue -1 increases in case of coded transmissions. For example, focusing on the $(6,3)$ code in (4) and the channel matrix in (7), $\mathbf{K}$ has 36 eigenvalues equal to -1 in the uncoded case and 44 in the coded case. This expansion allows reducing the estimator variance at medium-to-high SNR.

Matrix $\mathbf{K}$ could be computed analytically from the higher-order moments of the BPSK constellation if the encoder modular operations are expressed as products of the corresponding BPSK symbols, as done in (5). However, it is recommended to evaluate $\mathbf{K}$ by means of a short computer simulation taking into account that the elements of $\mathbf{K}$ take values in the set $\{-1, 0, 1\}$.

It can be shown that the channel is identifiable from the sample covariance matrix if and only if the Jacobian matrix $\mathbf{D}$ has full column rank. Actually, the estimator variance is strongly related to the condition number of matrix $\mathbf{D}$. The condition number is defined as the ratio between the largest and smallest singular value of $\mathbf{D}$ and it states how difficult is to separate the different CIR components.

Using the results in [5] and considering a first-order loop, the channel estimator variance is given by

$$VAR = \sum_{l=0}^{L-1} E\|\hat{h}_l - h_l\|^2 \simeq 2NB_n\text{Tr}\left((\mathbf{D}^H\mathbf{Q}^{-1}\mathbf{D})^{-1}\right) \tag{12}$$

where $\hat{h}_l$ is the estimator of h_l, Tr( ) is the trace operator and B_n is the noise equivalent loop bandwidth normalized to the symbol period. Simulations in next section have validated the correctness of the last expression.

4. NUMERICAL RESULTS

In this section, the steady-state variance of the studied closed-loop estimators is evaluated numerically averaging 500 independent realizations of the data symbols (BPSK), the noise and the channel impulse response. The channel is modeled as a complex-valued zero-mean Gaussian vector of exponentially decreasing variance, i.e.,

$$E\left\{\|h_l\|^2\right\} = \exp(-l/DS) \tag{13}$$

where DS is the channel delay spread normalized to the symbol interval. Recall that the phase of the first coefficient is set to 0 to

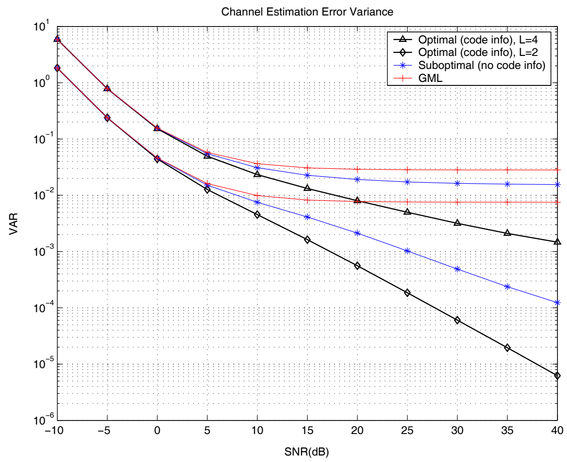

Fig. 1. Channel estimator variance as a function of the SNR for the optimal (code info), suboptimal (no code info) and GML second-order estimators. A constant power delay profile ($DS \gg 1$) with $L = 4$ taps is simulated. The loop bandwidth is set to 1.2×10^{-3} and the $(6, 3)$ coding in (4) is considered.

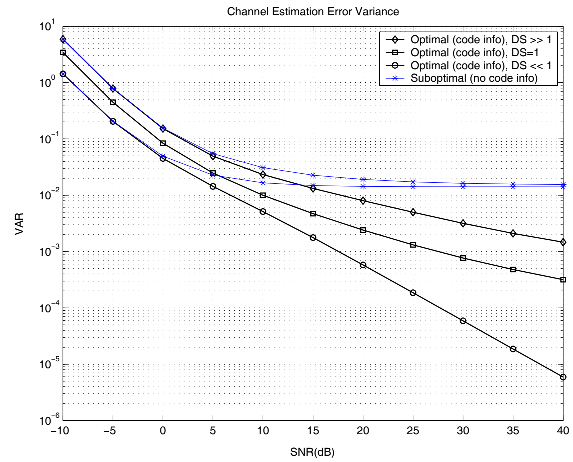

Fig. 2. Channel estimator variance as a function of the SNR for different values of the delay spread (DS). The CIR length is set to $L = 4$ and the loop bandwidth to 1.2×10^{-3}. The $(6, 3)$ coding in (4) is considered.

overcome the phase ambiguity of the sample covariance matrix (9). Notice also that the 500 channel realizations used in the variance calculus correspond to "well-behaved" channel responses having a condition number lower than 10.

Three channel estimators are simulated:

- Optimal estimator (code info) : the optimal sample covariance based estimator, which exploits the coding fourth-order information by means of matrix $\mathbf{K}$ (11).

- Suboptimal estimator (no code info) : the quadratic estimator proposed in [5] that only exploits the constant modulus property of the symbols. In this case, $\mathbf{K}$ is obtained from (11) assuming independent BPSK symbols.

- GML estimator : the quadratic Gaussian maximum likelihood estimator derived assuming Gaussian symbols [5], i.e., $\mathbf{K} = \mathbf{0}$.

In order to illustrate the main points of this work, four simulations have been carried out:

- Figure 1 : the performance of the GML estimator is dramatically degraded at high SNR due to the so-called *self-noise* caused by the random BPSK symbols. The reason is that the *symbol-rate* channel matrix $\mathbf{H}$ is not invertible since it has more inputs than outputs. Consequently, oversampling –or another form of diversity– is required to cancel out the self-noise floor at high SNR.

On the other hand, in case of coded transmissions, the redundancy of the error-correcting code is partially reflected in the kurtosis matrix $\mathbf{K}$ (11). Effectively, if this information is considered, the optimal estimator outperforms the subtpimal solution derived assuming independent BPSK symbols.

- Figure 2 : the optimal and suboptimal second-order estimator is evaluated for $L = 4$ and different values of the delay spread. Two asymptotic scenarios are considered: almost multiplicative channel ($DS \ll 1$) and almost constant power delay profile ($DS \gg 1$). It is shown that the estimator performance is lower and upper bounded by these two asymptotic cases. When DS goes to zero, the estimator variance is inversely proportional to the SNR (Fig. 1). For intermediate values of DS, the estimator variance decays with a lower slope.

Notice that, if the CIR length is $L = 4$, the suboptimal estimator (no code info) exhibits a variance floor at high SNR that is independent of the value of DS.

- Figure 3 : the optimal and suboptimal second-order estimator is evaluated for different values of L, assuming an almost constant power delay profile ($DS \gg 1$). In the figure, it is observed how the redundancy of the selected code (4) becomes insufficient to mitigate the self-noise as the number of taps increases. In this case, the improvement with respect to the suboptimal second-order estimator vanishes. It seems that self-noise is avoided if and only if the number of independent data symbols included in the observation is not greater than N (observation size). For the simulated $(6, 3)$ code, $L = 4$ becomes this limit because, in that case, only $2K = 6 \leq N$ out of the $N + L - 1 = 9$ involved symbols are independent.

- Figure 4 : the optimal and suboptimal second-order estimator is evaluated for different values of L (CIR duration) and the well-known $(7, 4)$ Hamming code:

$$\mathbf{G} = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 \end{bmatrix}. \tag{14}$$

Using this coding it is possible to reduce the variance achievable with the $(6, 3)$ code at high SNR. Although both codings introduce the same number of parity bits and the memory of the closed-loop estimator is frozen to the same value, the second encoding reduces the self-noise variance at high SNR, mainly for the longest channel response ($L = 6$).

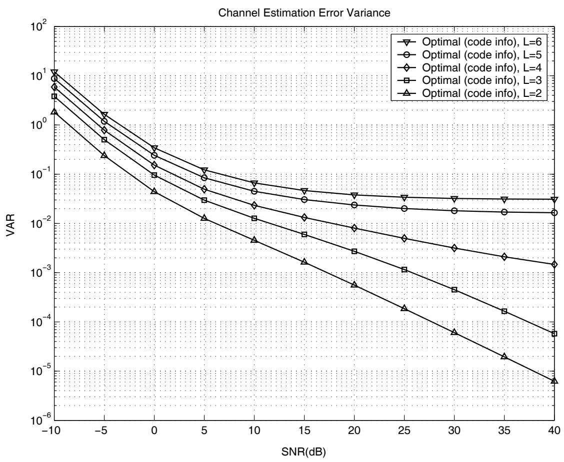

Fig. 3. Channel estimator variance as a function of the SNR for different values of the CIR duration (L). A constant power delay profile is simulated ($DS \gg 1$). The loop bandwidth is set to 1.2×10^{-3} and the code is the $(6,3)$ block code in (4).

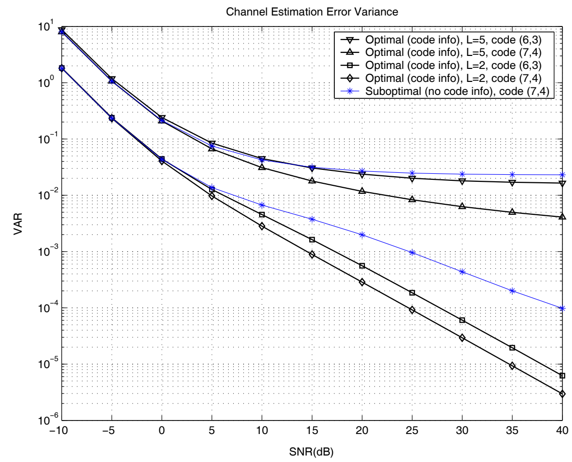

Fig. 4. Channel estimator variance as a function of the SNR for two different block encoders. The loop bandwidth is 1.2×10^{-3} and the CIR length is set to $L = 2$ or $L = 5$.

5. CONCLUSION

In this paper, it is proved that second-order blind channel estimators are able to use the redundancy of error-control codes to mitigate the self-noise variance. The code redundant structure is manifested in the fourth-order cumulants (kurtosis) of the transmitted symbols. This matrix appears naturally in the formulation of optimal second-order estimators and also provides information on the discrete nature of the transmitted symbols. It is shown that, using this fourth-order information, it is possible to obtain self-noise free channel estimates at one sample for symbol. This result is relevant because all the second-order techniques in the literature work with more than one sample per symbol.

6. REFERENCES

[1] L. Tong and G. Xu and T. Kailath, *A New Approach to Blind Identification and Equalization of Multipath Channels"*, Proc. of the 25th Asilomar Conference, Pacific Grove (USA), pp. 856-860, Nov. 1991.

[2] E. Moulines and P. Duhamel and J. Cardoso and S. Mayrargue, *"Subspace Methods for Blind Identification of Multichannel FIR Filters"*, IEEE Trans. on Signal Processing, vol. 43, no., pp. 516-526, Feb. 1995.

[3] G. Giannakis and S.D. Halford, *"Asymptotically Optimal Blind Fractionally Spaced Channel Estimation and Performance Analysis"*, IEEE Trans. on Signal Processing, vol. 45, no. 7, pp. 1815-1830, Jul. 1997.

[4] H.H. Zeng and L. Tong, *"Blind Channel Estimation Using the Second-Order Statistics: Algorithms"*, IEEE Trans. on Signal Processing, vol. 48, no. 8, pp. 1919-19300, Aug. 1997.

[5] J. Villares and G. Vazquez, *"Second-Order Parameter Estimation"*, IEEE Trans. on Signal Processing, vol. 53, no. 7, pp. 2408-2420, Jul. 2005.

[6] J. Villares and G. Vazquez, *"The Gaussian Assumption in Second-Order Estimation in Digital Communications"*, Submitted to IEEE Trans. on Signal Processing, Jul. 2006.

[7] J. Villares, *Sample covariance based parameter estimation for digital communications*, Ph.D. dissertation, Technical University of Catalonia (UPC), Barcelona, Spain, Sep. 2005. Available at http:\\gps-tsc.upc.es\comm.

[8] J.R. Treichler, B.G. Agee, *"A new approach to multipath correction of constant modulus signals,"* IEEE Trans. on Accoustics, Speech and Signal Processing, vol. 31, pp. 459-471, Apr. 1983.

[9] A.J. van der Veen, A. Paulraj, *"An analytical constant modulus algorithm,"* IEEE Trans. on Signal Processing, vol. 44, pp. 1136-1155, May 1996.

[10] R.R. Lopes, J.R. Barry, *"Exploiting error-control coding in blind channel estimation,"* Proc. of the IEEE Global Telecommunications Conference, San Antonio (USA), pp. 1317-1321, Nov. 2001

[11] J. Garcia-Frias, J. Villasenor, *Combined blind equalization and turbo decoding,* Proc. of IEEE Commun. Theory Mini-Conf., Vancouver (Canada), pp. 52-57, Jun. 1999.

[12] T. Fusco, M. Tanda *"ML-based symbol timing and frequency offset estimation for OFDM systems with noncircular transmissions,"* IEEE Trans. on Signal Processing, vol.54, no.9, pp. 3527- 3541, Sep. 2006.

[13] M. Meyr, G. Ascheid *Synchronization in Digital Communications. Phase-, Frequency-Locked Loops, and Amplitude Control,* John Wiley and Sons, 1990.