

A Reliable method for Blind Channel Identification using Burst Data

Dan Raphaeli, *Senior Member, IEEE* and Udi Suissa

Department of Electrical Engineering - Systems, Faculty of Engineering, Tel-Aviv University, Tel-Aviv 66978

E-mail: danr@eng.tau.ac.il - udint@inter.net.il

Abstract—In this paper we present a new approach for blind identification of Single Input Single Output (SISO) and Multiple Inputs Single Output (MISO) FIR channels with nonminimum phase. The approach is based on minimizing a cost function built of the problem unknown parameters and a vector of measurements achieved by passing the received data through a known parallel set of FIR filters followed by samples averaging. Averaging is done according to certain functions that have Higher Order Statistics (HOS) content and that their asymptotical mean can be expressed in closed form. The main advantages of this approach are its high probability of identification success when considering statistical channels, its ability to obtain reliable channel estimates in low SNR using short records of samples and its unsensitivity to overestimation of the channel order.

Index Terms—Blind Channel Identification, HOS, SISO, MISO, nonminimum phase channels, cumulants

I. INTRODUCTION

BLIND Identification of nonminimum phase linear systems has attracted a great attention in recent years. One of its most important applications is for the design of optimal receivers in digital communication where data bursts with independent channel between them (for example due to mobility, frequency hopping, unknown source) are used and where knowing the current channel impulse response is required without wasting bandwidth by transmission of training sequences. The only available information at the receiver for the sake of channel identification is the received data and some statistical information about the transmitted input data where the most common assumption in digital communication systems is that data is i.i.d with known distribution and noise is AWGN.

In general, blind channel identification methods can be partitioned to those based on *Second Order Statistics* (SOS) relative to others based on *Higher Order Statistics* (HOS). SOS methods are applicable when we have time or antenna diversity (assuming no common zeros among the diversity channels [1] [2]) because then the received data is no longer stationary but cyclostationary and channel phase can be identified. In this paper we will concentrate on the SISO channel case for which only HOS methods can be exploited (assuming no cyclostationarity is induced in the transmitter [3]).

Most methods for channel identification based on HOS can be divided into three main categories [12]: Closed form solutions [5] [4], Linear Algebra solutions [5] [8] and Non-linear optimization solutions [11] [7]. Closed form solutions are theoretically important but are not practical. Linear Algebra solutions are usually simpler than the Non-linear

ones but are less accurate. Non-linear optimization solutions give much accurate channel estimation but they require complex non-linear optimization processes and they can converge to a local minimum. In all those methods the main disadvantage is the requirement for relatively large records of samples (e.g thousands for FIR channels of five taps) to enable cumulants estimation with sufficiently low variance. The method that have the best performance, at least to our knowledge, when using relatively short records of samples (hundreds when other methods require thousands) is the EVI method [9] [10] to which we will compare our new method.

The method to be presented doesn't use cumulants in an explicit way. It can incorporate implicitly cumulants of any order and as we will see later it can use the Cumulants Generating Function itself. Its main advantage is its high probability of identification success when considering *Maximum Likelihood Sequence Estimation* (MLSE) performance over large number of random channels. In digital communication, MLSE success or failure is a more suitable criteria for channel estimation success than the common known criteria of Normalized Mean Square Error (NMSE) between the true channel and the estimated one. MLSE success, i.e having low enough BER, suggests that further improvement is possible for example when continuing channel estimation improvement and data detection with a probabilistic approach like [13] or a *Per-Survivor-Processing* (PSP) based method [14]. MLSE failure indicates that estimation is too bad to continue with.

Other advantages of our method are its ability to give very good channel estimates using relatively small records of samples and for very low SNR, its robustness to overestimation of the channel order (other methods require knowledge of the exact order) its relative simplicity and flexibility to adapt to various types of HOS functions where usually in explicit cumulants fitting methods cumbersome managing and measuring of all cumulants is required.

II. PRESENTATION OF THE METHOD

The received samples at the output of a SISO linear discrete-time channel with finite impulse response $\mathbf{h} = [h_0, \dots, h_{L-1}]$ are expressed as

$$y(k) = \sum_{m=0}^{L-1} h(m)I(k-m) + n(k) \quad (1)$$

where $\{I(k)\}$ is the transmitted data sequence and $\{n(k)\}$ is an additive noise. Our problem is to estimate the channel

coefficients $\{h(k)\}$ from a record of samples. The only information that we have is that $\{I(k)\}$ is an i.i.d sequence with zero mean and known nongaussian distribution. We also know that the noise is AWGN with zero mean. The channel order is unknown but can be bounded.

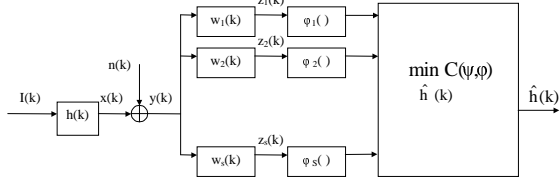


Fig. 1. General system description

The steps to be taken in our new method, are as described in Figure-1 where:

1. Pass the received data $y(k)$ through a parallel set of S FIR filters $\mathbf{w}_1, \dots, \mathbf{w}_S$ that fulfill some requirements (in time domain and frequency domain) to be detailed later.
2. At each of the FIR filters outputs measure the mean of a pre-defined function $f(\cdot)$ that has HOS content. The measurement will be done simply by

$$\varphi_i = \varphi(z_i) = g \left(\left[\frac{1}{N_s} \sum_{k=1}^{N_s} f(z_i(k)) \right] \right) \quad (2)$$

where N_s is the number of available samples, $z_i(k)$ is the i^{th} filter output and $g(\cdot)$ is any deterministic function including the trivial case of $g(c) = c$. In order to have more accurate samples averaging, especially when using small records of samples, ignoring the first $q-1$ samples of the received block is preferable. q is an upper bound of the channel length.

The function $f(\cdot)$ to be used must be a one that its expectation can be expressed explicitly in terms of channel coefficients, AWGN noise variance $No/2$ and the known coefficients of the FIR filters set. So that

$$\psi_i = \psi(\mathbf{w}_i, \mathbf{h}, No) = \lim_{N_s \rightarrow \infty} \varphi_i = g(E[f(z_i(k))]) \quad (3)$$

Averaging can also be done on functions that use both input and outputs samples (like cross-cumulants). Using multiple functions by applying different functions to different sections of the FIR filters outputs is also an option.

3. Solve the following problem to get the vector of unknowns

$$\min_{\hat{\mathbf{h}}(k), N_o} C(\varphi, \psi) \quad (4)$$

where $C(\cdot)$ is a cost function. In this paper we will give simulation results for solving the following simple Least Squares problem

$$\min_{\hat{\mathbf{h}}(k), N_o} \sum_{i=1}^S (\varphi_i - \psi_i)^2 \quad (5)$$

Solving (5) will be based on using the known *Gauss-Newton-Levenberg-Marquardt* method in what he have termed “Span Mode” and “Scan Mode” (see later on), where the following definitions will be used: \mathbf{W} is a $S \times M$ matrix which its rows are \mathbf{w}_i , $i = 1..S$, $\theta = [\hat{h}(0) \dots \hat{h}(\tilde{L}-1), N_o]$ is the vector of unknown parameters where $\tilde{L} = 2q-1$ for the “Span Mode” and $\tilde{L} = q$ for the “Scan Mode” and finally \mathbf{A} is the *Jacobian* matrix of $\psi(\hat{\mathbf{h}}, \mathbf{W}, N_o)$. The method works by

$$\theta_{p+1} = \theta_p - (A^T A + \lambda I)^{-1} A^T (\varphi - \psi) \quad (6)$$

where p is the *iteration* index of the algorithm and λ is the *Levenberg-Marquardt*(LM) parameter which insures convergence.

Since channel length is bounded to q and the solution of (5) depends on the initial channel guess then we have used two approaches. The “Span Mode” approach uses, as an initial channel guess, a vector of $\tilde{L} = 2q-1$ taps where only the middle tap gets an arbitrary value (not zero) while the rest of the taps equal zero. \tilde{L} enables covering of all the possible channel shifts. To further suppress the energy out of the maximum q main taps we have added an additional equation to the set of equations as followed

$$\sum_{|m|=q}^{2q-2} \alpha(|m|-q) R_{\hat{h}^2 \hat{h}^2}(m) = 0 \quad (7)$$

where $\alpha(i)$, $i = 0..q-2$ is a weighting vector that we increase dynamically as the equations set solution error decreases and $R_{\hat{h}^2 \hat{h}^2}(m)$ is the autocorrelation function of $\{\hat{h}^2(k)\}$.

The “Scan Mode” approach uses $\tilde{L} = q$ initial channel estimations where in each of these initial points one different tap is active while the others equal zero. Activating the LM method for a few number of iterations per a starting point and then selecting the one with the minimum intermediate error gave us a very good initial guess to continue with.

III. DEMONSTRATION OF THE METHOD

In this paper we will demonstrate the method for two specific functions assuming that transmitted data is BPSK modulated with $\{I(K)\}$ having equal probability of being $\{+1, -1\}$. We will use each of the functions separately to solve the problem of identifying the unknown channel (we can also use a mixed set of the equations). The channel coefficients will be real and for avoiding scaling ambiguity we will assume that channel norm equals one (in many papers scaling ambiguity is solved by assuming that $h(0) = 1$. We think that it is not a good idea in cases that $h(0)$ is very close to zero).

The first statistical function that we will use is

$$\psi(\mathbf{w}_i, \mathbf{h}, No) = \log E \left[e^{jz_i(k)} \right] \quad (8)$$

and the second one is

$$\psi(\mathbf{w}_i, \mathbf{h}, No) = E [z_i^4(k)] \quad (9)$$

The first function is actually the Cumulants Generating Function (CGF) of the received samples, $\log E \left[e^{j\mathbf{w}_i^T \mathbf{y}} \right]$ see [12]. The second function as we will see later is composed of 4th and 2nd order statistics of the received data.

Defining $g_i(k) = h(k) * w_i(k)$, $s_i(k) = x(k) * w_i(k)$ (signal part at the filters outputs), $v_i(k) = n(k) * w_i(k)$ (noise part at the filters outputs) and assuming that each of the FIRs $w_i(k)$ is of order $M-1$ we can express $\psi(\mathbf{w}_i, \mathbf{h}, N_0) = \log E \left[e^{jz_i(k)} \right]$ in closed-form as followed

$$\begin{aligned} \log E \left[e^{jz_i(k)} \right] &= \\ &= \log E \left[e^{j \sum_{l=k}^{k-M-L+2} g_i(k-l) I(l)} \right] + \log E \left[e^{j \sum_{p=0}^{M-1} w_i(p) n(k-p)} \right] \\ &= \log \prod_l E \left[e^{j I(l) g_i(k-l)} \right] + \log \prod_p E \left[e^{j w_i(p) n(k-p)} \right] \\ &= \log \prod_l \left(\frac{1}{2} e^{j g_i(k-l)} + \frac{1}{2} e^{-j g_i(k-l)} \right) + \log \prod_p e^{-\frac{1}{2} w_i^2(p) \frac{N_0}{2}} \\ &= \sum_{m=0}^{M+L-2} \log(\cos(g_i(m))) - \sum_{p=0}^{M-1} \frac{1}{2} w_i^2(p) \frac{N_0}{2} \quad (10) \end{aligned}$$

where in (10) we have calculated, explicitly, the expectation of $e^{j I(l) g_i(k-l)}$ knowing that the probability of $I(k)$ being +1 or -1 equals $\frac{1}{2}$. Furthermore we have used the fact that the *Characteristic Function* of a Gaussian random variable $n(k)$, $E \left[e^{jwn} \right]$, equals $e^{-\frac{1}{2} w^2 \sigma_n^2}$.

For $\psi(\mathbf{w}_i, \mathbf{h}, N_0) = E \left[z_i^4(k) \right]$ we will get:

$$E \left[z_i^4(k) \right] = E \left[s_i^4(k) \right] + 6E \left[s_i^2(k) \right] E \left[v_i^2(k) \right] + E \left[v_i^4(k) \right] \quad (11)$$

where we have used the fact that odd moments of $s_i(k)$ and $v_i(k)$ equal zero, since their probability distribution is even. The first part of (11) is

$$E \left[s_i^4(k) \right] = E \left[\left(\sum_m I(k-m) g_i(m) \right)^4 \right] =$$

$$E \left[I^4(k) \sum_k g_i^4(k) + 3E^2 \left[I^2(k) \right] \left[\left(\sum_k g_i^2(k) \right)^2 - \sum_k g_i^4(k) \right] \right] \quad (12)$$

using the fact that only the terms with indexes $m = n \neq p = l$, $m = l \neq p = n$, $m = p \neq l = n$ and $m = n = l = p$ give contribution in the multiple sums.

The second part of (11) is simply

$$6E \left[s_i^2(k) \right] E \left[v_i^2(k) \right] = 6 \frac{N_0}{2} \sum_q w_i^2(q) \sum_p g_i^2(p) \quad (13)$$

and the third part of (11) is

$$E \left[v_i^4(k) \right] = 3 \left(\frac{N_0}{2} \right)^2 \left(\sum_p w_i^2(p) \right)^2 \quad (14)$$

where we have used the 4th order cumulant definition (see [12]) to extract $E \left[n(k-m)n(k-t)n(k-l)n(k-p) \right]$ and the known fact that all cumulants of order higher than two of a gaussian process equals zero. So finally we get

$$\begin{aligned} E \left[z_i^4(k) \right] &= 3 \left(\sum_k g_i^2(k) \right)^2 - 2 \sum_k g_i^4(k) + 6 \frac{N_0}{2} \sum_p w_i^2(p) \sum_k g_i^2(k) \\ &\quad + 3 \left(\frac{N_0}{2} \right)^2 \left(\sum_p w_i^2(p) \right)^2 \quad (15) \end{aligned}$$

Using the closed-form expressions of (15) and (10) it is now simple to express the *Jacobian* matrix \mathbf{A} where for the case of $\log E \left[e^{jz_i(k)} \right]$ the gradients are

$$\frac{\partial \psi_i}{\partial \hat{h}(j)} = - \sum_{m=0}^{M+L-2} w_i(m-j) \tan \sum_{k=0}^{L-1} h(k) w_i(m-k) \quad (16)$$

$$\frac{\partial \psi_i}{\partial N_0} = - \frac{1}{4} \sum_{p=0}^{M-1} w_i^2 \quad (17)$$

and for the case of $E \left[z_i^4(k) \right]$ the gradients are

$$\begin{aligned} \frac{\partial \psi_i}{\partial \hat{h}(j)} &= \\ &= (12 \sum_{k=0}^{M+L-2} g_i^2(k) + 6N_0 \sum_{p=0}^{M-1} w_i^2(p)) \sum_k w_i(k-j) g_i(k) \\ &\quad - 8 \sum_{k=0}^{M+L-2} w_i(k-j) g_i^3(k) \quad (18) \\ \frac{\partial \psi_i}{\partial N_0} &= 3 \sum_{p=0}^{M-1} w_i^2(p) \sum_{k=0}^{M+L-2} g_i^2(k) + 6 \frac{N_0}{4} \left(\sum_{p=0}^{M-1} w_i^2(p) \right)^2 \quad (19) \end{aligned}$$

Now that we have ψ_i and \mathbf{A} we just have to measure φ_i to complete the equations set. The measurement will be done by simple empirical averaging on each of the FIR filters outputs where $i = 1, \dots, S$ according to (2).

Following the algorithm steps described in section-II, using each of the measured values and asymptotical expressions of the two functions presented, will give us an estimation of the unknown channel coefficients and the unknown AWGN variance.

Referring to the function $\log E \left[e^{jz_i(k)} \right]$, we can see that it is very close to the $E \left[z_i^4(k) \right]$ case if we replace $\log(\cos(\cdot))$ with its dominant *Taylor* series expansion terms.

IV. SIMULATIONS AND RESULTS

In this section we will present some simulation results. Our simulations were obtained on 2000 random real channels for small records of samples and on 10000 random real channels for large blocks of samples, each channel is of 5 taps. Each channel was normalized to unit energy. The modulation we have used is BPSK. For our method we have selected $\psi = E \left[z_i^4(k) \right]$ using 55 FIR filters in "Span

Mode“ and comparison was made with the EVI [9],[10] method.

We also did some comparisons with the GM method but the number of samples that were required for the GM method was of the order of thousands and even tens of thousands and it also had very low probability of success (about 40% – 50%) as we will define it later.

Most of the papers dealing with Blind Channel Identification methods use the NMSE as the main parameter for performance comparison between methods

$$NMSE = \frac{\sum_{k=1}^L (h_k - \hat{h}_k)^2}{\sum_{k=1}^L h_k^2} \quad (20)$$

and this is usually done with regard to a very few specific channels.

A more suitable criteria which we will use in this paper, is the probability of MLSE failure when using random set of channels. From statistical empirical results on thousands of channels estimations we have observed that the MLSE BER in case the NMSE is larger than 0.15 is very high (e.g > 0.05) relative to the expected BER when using the true channel. This suggests that if $NMSE > 0.15$ then no further improvement of the estimation can be achieved by using additional steps of data aided iterations and it is considered as a failure.

From figure (2) we can see that for all block lengths starting from as low as 144 samples per block and up to 1000 samples per block and for all SNR, our method has about twice lower probability of failure relative to EVI. In both methods the probability of success is above 90% for SNR of 15dB and up and for all block lengths. For block lengths higher the 500 samples we reach more than 99% of success. Regarding results where $NMSE < 0.15$ we saw by histograms that the new method and EVI give about the same values.

Another result that we can report is that the new method is not sensitive to *Critical Channels* or to cases where we have few identical taps. Cases to which other methods are sensitive [10].

Regarding co-channel interferer presence, we have used random channels of 3 taps each and random interferer channels of also 3 taps each. The C/I that we have used was of 10dB and modulation type was BPSK. Using our new method with the modified expression of (3) according to the MISO problem formulation we have identified the channels with NMSE as good as in the SISO cases.

V. CONCLUSION

We present a new blind channel identification method for the SISO and MISO cases where no oversampling or antenna diversity is present. The new method gives a very high probability of channel identification when looking at a large number of random channels. The method is proved by simulations to have advantage relative to other methods when considering small records of samples, low SNR and unknown channel order. There is still work to be done in investigating how many equations are required, how to

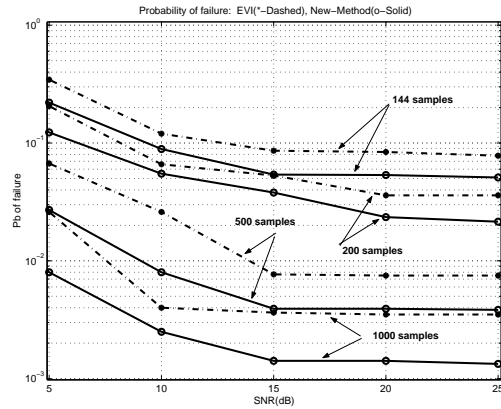


Fig. 2. Probability of failure, EVI and New method

select better the FIR filters coefficients and in analysing its performance and its identifiability criteria.

REFERENCES

- [1] L. Tong, G. Xu and T. Kailath. “Blind Identification and Equalization Based on Second-Order Statistics: A Time Domain Approach”. *IEEE Transactions on Information Theory*, Vol. 40, No 2, pp. 340-349, March 1994.
- [2] J.K. Tungait. “On Blind Identifiability of Multipath Channels Using Fractional Sampling and Second-Order Cyclostationary Statistics”. *IEEE Transactions on Information Theory*, IT-41(1), pp. 308-311, January 1995.
- [3] A. Chevreuril, E. Serpedin, P Loubaton, G.B. Giannakis. “Blind Channel Identification and Equalization Using Periodic Modulation Precoders: Performance Analysis”. *IEEE Transactions on Signal Processing*, Vol. 48, No. 6, pp. 1570-1586, June 2000.
- [4] A. Swami, J.M. Mendel. “Closed-Form Recursive Estimation of MA Coefficients using Autocorrelations and Third Order Cumulants”. *IEEE Transactions on Acoustics, Speech and Signal Processing*, ASSP-37(11), pp. 1794-1795, November 1989.
- [5] G.B Giannakis J.M. Mendel “Identification of Nonminimum Phase System Using Higher Order Statistics”. *IEEE Transactions on Acoustics, Speech and Signal Processing*, Vol. 37, No. 3, pp. 360-77, March 1989.
- [6] J. Tungait. “Identification of linear stochastic system via second-and-fourth-order Cumulant Matching”. *IEEE Transactions on Information Theory*, Vol. 33, pp. 393-407, MayJ 1987.
- [7] B. Friedlander, B. Porat. “Performance analysis of parameter estimation algorithm based on high-order moments”. *Int J. Adaptive Control and Signal Processing*, Vol. 3, pp. 191-229, 1989.
- [8] Jose A.R Fonollosa, J. Vidal “System Identification Using a Linear Combination of Cumulant Slices”. *IEEE Transactions on Signal Processing*, Vol. 41, No. 7, pp. 2405-2412, July 1993.
- [9] B. Jelonnek K.D Kammeyer. “Eigenvector Algorithm for Blind Equalization”. In *Proc. IEEE Signal Proc. Workshop on Higher-Order-Statistics*, pp. 19-23, South Lake Tahoe, California. June 1993.
- [10] D. Boss, B. Jelonnek, K.D Kammeyer. “Eigenvector Algorithm for Blind MA System Identification”. *ELSEVIER Signal Processing*, Vol. 66, No. 1, pp. 1-28, April 1998.
- [11] K.S III, M. Rosenblatt. “Deconvolution and estimation of transfer function phase and coefficients for non-Gaussian linear processes”. In *the Annals of Statistics*, Vol. 10, pp. 1195-1208, 1982.
- [12] Jerry M. Mendel “Tutorial on Higher-Order Statistics (Spectra) in Signal Processing and System Theory: Theoretical Results and Some Applications”. *Proceedings of the IEEE*, Vol. 79, No. 3, pp. 278-305, March 1991.
- [13] C. Anton-Haro, J.A.R Fonollosa, J.R Fonollosa “Blind Channel Estimation and Data Detection Using Hidden Markov Models”. *IEEE Tran. on Signal Processing*, Vol. 45, No. 1, pp. 241-247, January 1997.
- [14] R. Raheli, A. Polydoros, C. Tzou “Per-Survivor Processing: A general approach to MLSE in Uncertain Environments”. *IEEE Trans. On Comm*, Vol. 43, No. 2/3/4, pp. 354-362, Feb/March/April 1995.