

ROBUST BLIND CHANNEL IDENTIFICATION AND EQUALIZATION BASED ON MULTI-STEP PREDICTORS

David Gesbert
PAB/RGF/RCN
France Telecom-CNET
92131 Issy-Les-Mûr, France
gesbert@issy.cnet.fr

Pierre Duhamel
Dept. Signal
Telecom Paris
75013 Paris, France
duhamel@sig.enst.fr

ABSTRACT

This contribution deals with the problem of blind channel identification and equalization based on the (temporally or spatially) oversampled channel output. A novel algorithm is presented which builds on a *multistep prediction* (MSP) approach and can be viewed as a certain generalization of the initial Linear Prediction Algorithm (LPA) proposed in 1994. Our algorithm improves on most recent related works in that it is theoretically and practically unsensitive to the critical and expected problem of channel length mismatch. Moreover, the MSP scheme improves on the conventional LPA by increasing the robustness of this earlier algorithm. In contrast with the LPA, the proposed prediction scheme exploits the full channel structure, thus providing more statistical efficiency in channel identification. A direct symbol recovery algorithm (requiring no channel estimate) is also straightforwardly drawn from our approach.

1. INTRODUCTION

Blind multichannel identification and equalization exploiting the channel diversity induced by sensor arrays and/or fractional sampling (single input/multiple output) has attracted a lot of attention in the last four years. Numerous methods can now be found in the literature, based on various second-order criteria, which offer promising alternatives to the previously reported higher-order based techniques. Basically, the channel diversity framework (in which the same discrete source signal is observed through several linear filters sampled at source's rate) induces a useful "signal overdetermination" which can be successfully exploited in terms of signal/noise subspaces decompositions. Reminiscent of classical narrow band array processing techniques, most second-order methods developed so far do rely on such decompositions, in either a stochastic (see for instance [8, 12]) or deterministic (see [5, 6]) way. Other subspace methods also incorporate alphabet related knowledge to further (i) improve the performances and/or (ii) address the equalization problem with more than one source (multiple input/multiple output) [7].

It is known however that subspace approaches generally do not tolerate mismatched or even ill-defined channel lengths since they require explicit knowledge of the "source" (or "noise") subspace dimension to work [2]. This is a major limitation to their practical applicability. In contrast, the alternative Linear Prediction Algorithm, proposed and

then refined by Slock *et al.* [1, 2, 3] proved to be consistent in the presence of channel order errors. Note that other authors also reported similar robustness properties of their algorithms, but on experimental basis only [10, 11]).

Though this important property makes the LPA one of the potentially most attractive solutions to the blind multichannel equalization problem, the initial algorithm of [1, 2] lacks robustness because it fails in exploiting the channel structure completely. In particular, the algorithm overall performances rely on the particular realization of the multichannel precursor coefficient. The reason for this is that only a one-step predictor is used in [1, 2], yielding a prediction error signal with uncontrollable symbol-to-noise ratio.

Here we investigate a more general framework based on a multistep prediction approach which, importantly, is still consistent in the presence of channel length mismatch. The main result here is that several multistep prediction error signals can be used to triangularize the multichannel system and then combined to exhibit the full channel structure. In contrast, the initial LPA only involves a one-step predictor, leading to causal channel equalization. A channel identification algorithm is derived which is statistically more efficient than that of the LPA, at the expense of a higher computational cost. A direct linear non-causal symbol estimation procedure is also easily drawn from the MSP approach.

Notations: $E()$ statistical expectation; $()^*$ complex conjugation; $()^t$ transposition; $()^+$ transpose-conjugation; $|\cdot|$ L_2 -norm of a vector or matrix; $\mathbf{0}_{i \times k}$: $i \times k$ all-zero matrix. I_k identity matrix of size k .

2. MULTICHANNEL REPRESENTATION

Consider the sample output of a 1-input/ L -output baseband communication channel with finite-order (M) impulse response $\mathbf{h}(0), \mathbf{h}(1), \dots, \mathbf{h}(M)$:

$$\mathbf{x}(n) = \sum_{k=0}^M s(n-k)\mathbf{h}(k) + \mathbf{b}(n). \quad (1)$$

Here $s(n)$ and $\mathbf{b}(n)$ denote respectively the mutually and individually white symbol (unit-variance) and noise (with variance σ_b^2) sequences. In our L -channel framework, $\mathbf{x}(n), \mathbf{h}(k), \mathbf{b}(n)$ are all $L \times 1$ vector quantities. Stacking $N \geq M$ successive samples in

$$\mathbf{X}_N(n) = [\mathbf{x}(n)^t, \mathbf{x}(n-1)^t, \dots, \mathbf{x}(n-N+1)^t]^t,$$

$$\mathbf{B}_N(n) = [\mathbf{b}(n)^t, \mathbf{b}(n-1)^t, \dots, \mathbf{b}(n-N+1)^t]^t$$

yields the well-known linear model, where both channel and symbols are unknown desired quantities:

$$X_N(n) = \mathcal{H}_N S_K(n) + B_N(n) \quad (2)$$

where $K = M + N$ is the number of symbols involved between time n and $n - N + 1$. We have in addition:

$$\mathcal{H}_N = \begin{pmatrix} h(0) & \cdots & h(M) & 0 & \cdots & 0 \\ & & & \vdots & & \\ 0 & \cdots & 0 & h(0) & \cdots & h(M) \end{pmatrix}$$

$$S_K(n) = [s(n), \dots, s(n - K + 1)]^t.$$

The fundamental channel "length and zero" condition for second-order deconvolution can be restated as follows:

- (H1) At least one channel has exact degree M (the other degrees being possibly lower).
- (H2) The z -transforms of the channel impulse responses do not share any common zero.

Under (H1), (H2), $\mathcal{H}_N$ is a left-invertible matrix [4].

3. PREDICTION-ERROR METHODS

Generally speaking, prediction error methods aim at recovering the channel inverse through the whitening of the observations, under the fundamental i.i.d input condition. For usual (mixed-phase) monochannel systems, the whiteness condition is too weak to allow phase and amplitude equalization. In the monochannel case, the prediction criterion (if one wishes to use it) then must be augmented with a non-linear higher-order based criterion (CMA or other) as was proposed for instance in [9].

As an important difference, there exist exact FIR channel inverses in the *multichannel context*. Relying on this property, the main idea of prediction-error methods for multichannel equalization and identification consists in noting that the *noise free* received signal $\mathbf{x}(n)$, while MA by construction, is at the same time a finite-order (N) AR signal by nature under (H3) [2]. Consequently, the input sequence $s(n)$ is also found to be the finite-horizon innovation process of $\mathbf{x}(n)$, hence can be recovered by prediction-error filtering. The basis for the conventional LPA [1, 2] is as follows:

Lemma 3..1 *Let P be a $LN \times L$ complex-valued matrix of prediction coefficients. Let $\tilde{\mathbf{x}}(n)$ be the $L \times 1$ prediction error defined by: $\tilde{\mathbf{x}}(n) = \mathbf{x}(n) - P^+ X_N(n-1)$. The minimum prediction error variance $J(P) = E \|\tilde{\mathbf{x}}(n)\|^2$ is attained if and only if:*

$$\tilde{\mathbf{x}}(n) = (I_L, -P^+) X_{N+1}(n) = h(0)s(n) \quad (3)$$

Note that the channel order M need not be known to use the prediction algorithm (in fact, one should choose N so as to at least overestimate M) which makes the method robust towards the likely order estimation errors, in contrast with most subspace approaches. Based on (3), a possible approach to blind channel equalization consists in 1) solving the noise free prediction problem corresponding to the minimization of $J(P)$, and 2) using the prediction-error filter $(I, -P^+)$ as a particular channel FIR causal inverse as proposed in [1]. Since however the performances of this

equalization method critically depends on the SNR (uncontrollable) on $h(0)s(n)$, it is preferable to identify the channel first. To this end, the prediction error $\tilde{\mathbf{x}}(n)$ can be used as a training sequence against which one may correlate the measured signals to estimate the channel characteristics as in a input/output method.

Clearly, the fact that the prediction error signal only depends on very partial channel information (namely, only the first coefficient of the impulse response) may result in unreliable (symbol or channel) estimates for some particular realization of the channel coefficients. A partial answer to this problem was investigated in [3] in terms of moving the dependence onto $h(i), i > 0$, instead of $h(0)$. In this paper, we investigate a more general and efficient solution which builds on a generalization of the initial prediction framework.

4. MULTI-STEP PREDICTION

Let $\hat{M}$ be an arbitrary estimated channel order such that $\hat{M} \geq M$. Consider the *i-step prediction* problem at the multichannel output, corresponding to the minimum variance estimation of $\mathbf{x}(n)$, given the past samples: $\mathbf{x}(n-i)$, $\mathbf{x}(n-i-1), \dots, \mathbf{x}(n-N)$, where N is chosen such that $N > \hat{M}$.

An *i-step predictor*, denoted by P_i , is an $L(N-i+1) \times L$ complex-valued matrix minimising

$$J_i(P_i) = E \|\mathbf{x}(n) - P_i^+ X_{N-i+1}(n-i)\|^2$$

It is found with previous notations that

- $J() = J_1()$.
- $P = P_1$.

Lemma 4.2 (System Triangularisation) *Let the *i-step prediction error filter*, for $i = 1, \hat{M} + 1$, be defined as*

$$\Omega_i = (I_L, \mathbf{0}_{L \times L(i-1)}, -P_i^+)^t.$$

*We temporarily assume a noise-free model ($\sigma_b^2 = 0$). Then, assuming $(N+1)(L-1) \geq LM$, the *i-step prediction error signal* is given by, for all i of $1, \hat{M} + 1$:*

$$\tilde{\mathbf{x}}_i(n) \stackrel{\text{def}}{=} \Omega_i^+ X_{N+1}(n) = \sum_{k=0}^{i-1} h(k)s(n-k) \quad (4)$$

Proof For $i \geq M + 1$, i then being greater than the channel memory, the prediction has clearly no effect: $\tilde{\mathbf{x}}_i(n) = \mathbf{x}(n)$, since there is no correlation between predicted and prediction variables. The proof of (4) for $i \leq M$ is as follows. Using the well-known orthogonality principle at optimality, it comes from (2):

$$\begin{aligned} E(X_{N-i+1}(n-i)\tilde{\mathbf{x}}_i(n)^+) &= \mathbf{0} \\ E(\mathcal{H}_{N-i+1} S_{K-i+1}(n-i) X_{N+1}(n)^+) \Omega_i &= \mathbf{0} \\ \mathcal{H}_{N-i+1} E(S_{K-i+1}(n-i) S_{K+1}(n)^+) \mathcal{H}_{N+1}^+ \Omega_i &= \mathbf{0} \\ \mathcal{H}_{N-i+1} (\mathbf{0}_{K-i+1 \times i}, I_{K-i+1}) \mathcal{H}_{N+1}^+ \Omega_i &= \mathbf{0} \quad (5) \end{aligned}$$

Under (H1), (H2), $\mathcal{H}_{N-i+1}$ has full column rank. Moreover, the prediction error filter is strongly structured, so that (5)

yields

$$\begin{aligned} (0_{K-i+1 \times i}, I_{K-i+1}) \mathcal{H}_{N+1}^+ \Omega_i &= \mathbf{0} \\ \Omega_i^+ \mathcal{H}_{N+1} &= (\mathbf{h}(0), \dots, \mathbf{h}(i-1), \mathbf{0}, \dots, \mathbf{0}) \\ \text{then } \Omega_i^+ X_{N+1}(n) &= \sum_{k=0}^{i-1} \mathbf{h}(k) s(n-k) \quad \square \end{aligned}$$

Remark: In a batch formulation, the predictors are obtained through resolution of several different Yule-Walker equations, instead of one in the LPA. For the i -step problem, Yule-Walker equations are classically given by

$$R_{N-i+1} P_i = E(X_{N-i+1}(n-i) \mathbf{x}(n)^+), \quad (6)$$

where R_{N-i+1} is the covariance of $X_{N-i+1}(n)$, and from which the noise contribution has to be removed in some way to estimate unbiased predictors in a noisy situation. To this end, the noise level can be estimated from the smallest eigenvalues of R_N in a batch implementation. Non trivial adaptive procedures for computing unbiased solutions to (6) can also be found in [13].

4.1. Utilization of the predictors

(4) offers a simple generalization of (3): The one-step prediction filter completely eliminates the ISI while, more generally, the i -step prediction error filter reduces the IIS order from M to $i-1$. Then, the action of the multistep prediction error filters can be seen as a triangularization of the multichannel system. The different filters can be straightforwardly exploited using pair-wise combinations as follows.

Let $Z(n)$ be the $L(\tilde{M}+1) \times 1$ vector signal defined according to:

$$\begin{aligned} Z(n) &= (X_{N+1}(n)^t \Omega_1^*, X_{N+1}(n+1)^t (\Omega_2 - \Omega_1)^*, \dots, \\ &\quad X_{N+1}(n + \tilde{M})^t (\Omega_{\tilde{M}+1} - \Omega_{\tilde{M}})^*)^t \end{aligned} \quad (7)$$

involving delayed versions of all prediction error signals. In the absence of noise, it comes:

$$Z(n) = h s(n); \quad (8)$$

where $h = (\mathbf{h}(0)^t, \mathbf{h}(1)^t, \dots, \mathbf{h}(M)^t, \mathbf{0}_{1 \times L(\tilde{M}-M)})^t$ is the column-wise global channel impulse response. Note that, in contrast with (3), (8) involves *all* channel coefficients, thus removing the problematic exclusive dependency on $\mathbf{h}(0)$.

In a practical situation though, we deal with a noisy $Z(n)$ signal:

$$Z(n) = h s(n) + \tilde{B}(n), \quad (9)$$

where $\tilde{B}(n)$ is a filtered noise sequence obtained from: (please recall (7))

$$\begin{aligned} \tilde{B}(n) &= (B_{N+1}(n)^t \Omega_1^*, B_{N+1}(n+1)^t (\Omega_2 - \Omega_1)^*, \dots, \\ &\quad B_{N+1}(n + \tilde{M})^t (\Omega_{\tilde{M}+1} - \Omega_{\tilde{M}})^*)^t \end{aligned}$$

4.2. Channel Identification

Equation (9) suggests to simply extract h (up to a scalar constant) from $Z(n)$, for channel identification. It is however statistically more efficient to deal with the covariance matrix of $Z(n)$:

$$R_Z \stackrel{\text{def}}{=} E(Z(n)Z(n)^+) = h h^+ + \sigma_b^2 R_B, \quad (10)$$

where R_B is the normalized covariance of $\tilde{B}(n)$. Note that R_Z (resp. R_B) is easily expressed in terms of the prediction coefficients and of the received signal (resp. noise) correlation coefficients. A straightforward procedure then consists in extracting the dominant eigenvector of the denoised covariance $R_Z - \sigma_b^2 R_B$ in order to obtain a channel estimate.

4.3. Symbol recovery

A channel estimate can be used for non-linear (maximum likelihood) symbol recovery typically. However, a direct linear estimation of the transmitted symbols is also possible through a proper combination of the entries of $Z(n)$. Let U be a $L(\tilde{M}+1) \times 1$ complex-valued combiner:

$$U^+ Z(n) = \rho \hat{s}(n), \quad (11)$$

where ρ is a scalar undetermination. Note that the estimation in (11) is strictly non causal, since $\hat{s}(n)$ carries both past and future received information.

Possible solutions for U include:

- The optimal solution in terms of output SNR: $\text{argmax}(U^+ R_Z U) / (U^+ R_B U)$. (this solution is however numerically ill-posed due to the rank deficiency of R_B , which itself results from the redundancy in $\tilde{B}(n)$).
- The (sub-optimal) maximum output power (MOP) solution: $\text{argmax}(U^+ R_Z U) / (U^+ U)$. This solution is more appropriate due to better conditioning. It is found that $U \rightarrow h$ when $\sigma_b^2 \rightarrow 0$.

5. SIMULATIONS

We illustrate the behaviour of the (batch based) MSP scheme above and provide a comparison with the one step LPA of [2]. Monte-Carlo simulations are conducted in the following context: $L = 2$ channels of degree $M = 4$, driven by a white QPSK sequence. Channel one: $(-0.05 + 0.27j) + (-0.37 - 0.01j)z^{-1} + (0.02 - 0.07j)z^{-2} + (-0.21 - 0.03j)z^{-3} + (0.5 - 0.6j)z^{-4}$. Channel two: $(0.25 + 0.27j) + (-0.1 + 0.38j)z^{-1} + (0.22 - 0.05j)z^{-2} + (0.26 + 0.14j)z^{-3} + (0.17 - 0.72j)z^{-4}$. The output SNR is defined as in [12].

• Fig. 1 shows the statistical performance in terms of channel estimate variance for the scheme presented in 4.2.. An overestimated channel order is fed into the algorithm, with $\tilde{M} = 5$. Results are averaged over 100 independent realizations and the normalized variance (NV) is plotted for different sample sizes:

$$NV = \frac{<|\hat{\mathbf{h}} - \mathbf{h}|^2>_{100}}{|\mathbf{h}|^2} \quad (12)$$

where $< \cdot >_{100}$ denotes the averaging operation. A severe 5 dB SNR is simulated. We plot the result achieved by the Subspace Method of [12] under similar conditions.

The MSP algorithm consistently provides better results than the LPA, for any sample size. The advantage of a prediction approach over a subspace approach is also confirmed: the Subspace Method does not work in the presence of channel length mismatch.

• Fig. 2 compares the results of direct linear symbol estimations based on (a) combining the prediction error signals

in the MSP scheme, as described in 4.3., and (b) combining the one-step prediction error signals in a conventional LPA. The SNR is set to 25 dB. In each experiment, we use the MOP combiner and plot the equalized constellations. The MSP algorithm linearly combines $L(\hat{M} + 1)$ error signals and outperforms the LPA which only uses L causal error signals. The one-step predictor eliminates the ISI but dramatically degrades the symbol to noise ratio. Note however that in both cases, the SNR in the linear symbol estimates is directly related to the predictors gains, which cannot be fully controlled, hence limiting the applicability of these direct equalization techniques.

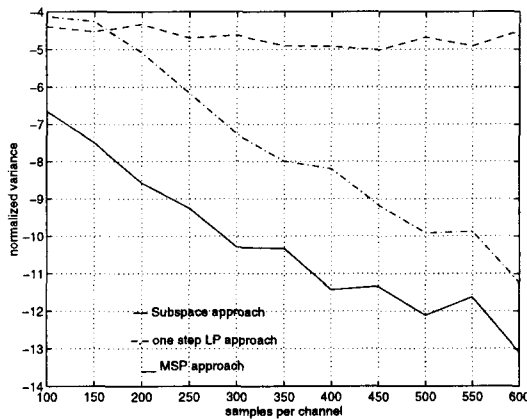


Figure 1. Channel identification with the MSP method. Comparison with LPA and Subspace Algorithm. RSB= 5dB, $N = 10$.

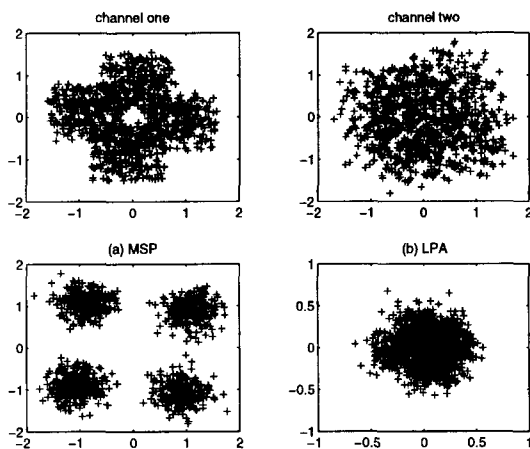


Figure 2. Output of the equalizer combining channel one and channel two. fig.(a): MSP method. fig.(b): LPA method. RSB= 25 dB, $N = 10$. 1000 samples are used.

6. CONCLUSION

The Multi Step Prediction approach for blind channel identification/equalization offers a simple generalization of the initial Linear Prediction Algorithm. Like the LPA, it is consistent in the presence of channel length mismatch, thus improving on the vast majority of existant (typically subspace-

based) solutions. It is also statistically more efficient than the LPA because it exploits more completely the channel structure and exploits typically more second-order information, at the expense of a higher computational cost. The MSP approach can be used for channel identification as well as for direct symbol recovery. In the latter case however, the equalization performances are limited by the (hardly controllable) gains of the prediction error filters.

REFERENCES

- [1] D. Slock, "Blind fractionally-spaced equalization, perfect reconstruction filter-banks and multichannel linear prediction," in *Proc. ICASSP*, vol. 4, pp. 585-588, 1994.
- [2] K. Abed Meraim, P. Duhamel, D. Gesbert, P. Loubaton, S. Mayrargue, E. Moulines, and D. Slock, "Prediction error methods for time-domain blind identification of multichannel FIR filters," in *Proc. ICASSP*, 1995.
- [3] D. T. M. Slock and C. B. Papadias, "Further results on blind identification and equalization of multiple FIR channels," in *Proc. ICASSP*, 1995.
- [4] Y. Li and Z. Ding, "Blind channel identification based on second-order cyclostationary statistics," in *Proc. ICASSP*, vol. 4, pp. 81-84, 1993.
- [5] H. Liu, G. Xu, and L. Tong, "A deterministic approach to blind identification of multichannel FIR systems," in *Proc. ICASSP*, vol. 4, pp.581-584,1994.
- [6] H. Liu and G. Xu, "A deterministic approach to blind symbol estimation," in *IEEE SP Letters*, vol 1, n 12, dec. 1994
- [7] A. J. Van der Veen, S. Talwar, and A. Paulraj, "Blind estimation of multiple digital signals transmitted over FIR channels," *IEEE S.P. Letters*, vol. 2, pp. 99-102, May 1995.
- [8] L.Tong, G.Xu, and T.Kailath, "Blind identification and equalization based on second-order statistics: a time-domain approach," *IEEE Tr. on Information Theory*, vol. 40, no. 2 , pp. 340-349, March 1994.
- [9] O. Macchi, C. Faria Da Rocha, and J. Travaros-Romano, "Egalisation adaptive autodidacte par retro-prédiction et prédiction," in *Proc. Colloque GRETSI*, pp. 491-494, 1993.
- [10] D. Gesbert, P. Duhamel, and S. Mayrargue, "On-line blind multichannel equalization based on mutually referenced filters," *IEEE Trans. on Signal Processing*, to appear. Shorter version in *Proc. Eusipco'94*.
- [11] Z. Ding, "An outer-product decomposition algorithm for multichannel blind deconvolution," in *proc. 8th IEEE Signal Processing Workshop on SSPA*, 1996.
- [12] E. Moulines, P. Duhamel, J.F. Cardoso, and S. Mayrargue, "Subspace methods for the blind identification of multichannel FIR filters," *IEEE Trans. SP*, vol. 43, no. 2, pp. 516-526, February 1995.
- [13] D. Gesbert and P. Duhamel, "Unbiased adaptive blind channel identification and equalization," in *IEEE Tr. on Signal Processing*, submitted Dec. 1996. Shorter version in *Proc. 29th Asilomar Conf. on Signals, Systems, and Computers*, 1995.