BLIND FIR CHANNEL ESTIMATION IN MULTICHANNEL CYCLIC PREFIX SYSTEMS

Dirk T.M. Slock[*]

Eurecom Institute
2229 route des Crêtes, B.P. 193, 06904 Sophia Antipolis Cedex, FRANCE
Email: dirk.slock@eurecom.fr

ABSTRACT

In this paper, we revisit a number of classical blind estimation techniques for FIR multichannels when applied to communication systems that are based on the introduction of a cyclic prefix. These techniques include techniques based on deterministic modeling of the unknown symbols such as (signal and noise) subspace fitting methods, subchannel response matching (SRM), deterministic maximum likelihood (DML), and techniques based on a Gaussian white noise model for the unknown symbols such as Gaussian ML (GML) methods and covariance matching. The presence of a cyclic prefix transforms spatiotemporal channels into a set of parallel spatial channels, coupled by the discrete Fourier transform (DFT) of the FIR channel impulse response. The associated blind channel estimation methods become computationally much more attractive and also become more straightforward to analyze and to compare in terms of performance. Working in the DFT domain reveals immdiately that temporal whiteness of the additive noise is unessential, only spatial whiteness matters. Furthermore, the blind channel identifiability conditions become extremely weak when Zero Padded (ZP) systems are considered.

1. INTRODUCTION

A wealth of blind channel estimation techniques have been introduced for spatio-temporal channels over the past decade, based on the singularity of the received signal power spectral density matrix [11]. This singularity can be exploited to separate the white noise contribution. The main problem characteristic in fact that allows channel identifiability is the minimum phase characteristic of the Single-Input Multiple-Output (SIMO) or MIMO matrix channel transfer function of the spatio-temporal channel. Spatio-temporal channels arise in mobile communications when multiple antennas or polarizations or beams are used at the receiver. Physical multi-channels can also arise in xDSL systems when the receiver has access to a complete cable bundle. Other problem formulations that lead to multi-channel models are the use of oversampling at the receiver or the decoupling of in-phase and in-quadrature components when real symbols get modulated. Or the reception of multiple signal copies in ARQ protocols.

A variety of blind symbol/channel estimation strategies can be developed depending on the amount of a priori information that gets formulated on the unknown symbols. In general, the less structure that gets exploited about the symbol alphabet, the less problems are encountered with local minima. Of course, more estimation accuracy is obtained by exploiting more information. A reasonable strategy is hence to exploit a progressive range of algorithms exploiting increasing a priori information levels. The algorithm at the next level can be initialized with the estimate obtained at the previous level of a priori information.

The memory introduced by a convolutive channel leads to the requirement of having to treat all available data in a contiguous observation interval in one shot if no suboptimality is allowed. This leads to problem formulations with large convolution matrices, large covariance matrices and high complexity. Attempts have been made by our own group to introduce asymptotic approximations, by approximating large Toeplitz convolution matrices by circulant matrices, to allow transformation to the frequency domain, or by others by introducing approximate DFT operations.

Cyclic prefixes have been introduced in a number of existing systems such as OFDM systems for ADSL and wireless LANs. To combine the benefits of exploiting diversity and simplifying equalization, multicarrier CDMA systems have been proposed which combine a spreading over tones with the OFDM approach. If the DFT matrix is used as spreading matrix (corresponding to full spreading over all tones), then the spreading-OFDM combination leads in fact to direct transmission in the time domain but with a cyclic prefix [5]. Such a system has been proposed for one of the

[*]Institut Eurécom's research is partially supported by its industrial members: Bouygues Télécom, Fondation d'entreprise Groupe Cegetel, Fondation Hasler, France Télécom , Hitachi, ST Microelectronics, Swisscom, Texas Instruments, Thales. This paper was also presented at the Asilomar Conf. on Signals, Systems and Computers, Pacific Grove, CA, Nov. 2002, but without trace in the proceedings.

high data rate WLAN standards. Cyclic prefixes have also been introduced in a number of communication scheme proposals.

The introduction of a cyclic prefix renders the transformation to the frequency domain clean and exact even for a finite data length. The resulting algorithmic simplifications will be detailed for a number of classical blind channel estimation methods mentioned above. Furthermore, the same framework can be used to analyze the performance of the algorithms and the algorithmic simplifications also translate into much simplified performance expressions, which allow a direct and insightful analytical performance comparison between a number of algorithms.

Another significant consequence of the presence of a cyclic prefix is its impact on or insight in identifiability. Since the spatiotemporal channel gets transformed into a bank of parallel spatial channels at a number of tones, the requirement on the noise for blind identifiability is to be spatially white at each tone. However, it is not at all required that the variance be the same at each tone. This implies also that for non cyclic prefix systems, channel identifiability should be possible whenever the noise spectral density matrix is a scalar multiple of the identity matrix with the scalar multiple being a fairly arbitrary scalar power spectral density. This constitutes a new finding.

2. MIMO CYCLIC PREFIX BLOCK TX SYSTEMS

Consider a MIMO system with q inputs x_l, $p > q$ outputs y_i per (symbol/sample) period

$$\begin{aligned}\underbrace{\mathbf{y}[m]}_{p\times 1} &= \sum_{l=1}^{q}\sum_{j=0}^{L_l}\underbrace{\mathbf{h}^l[j]}_{p\times 1}\,\underbrace{x_l[m-j]}_{1\times 1} \\ &= \sum_{j=0}^{L}\underbrace{\mathbf{h}[j]}_{p\times q}\,\underbrace{\mathbf{x}[m-j]}_{q\times 1} = \underbrace{H(q)}_{p\times q}\,\underbrace{\mathbf{x}[m]}_{q\times 1}\end{aligned} \quad (1)$$

where $H(q) = \sum_{j=0}^{L} \mathbf{h}[j]\, q^{-j}$ is the MIMO system transfer function corresponding to the z transform of the impulse response $\mathbf{h}[.]$. Equation (1) mixes time domain and z transform domain notations to obtain a compact representation. In $H(q)$, z is replaced by q to emphasize its function as an elementary time advance operator over one sample period. Its inverse corresponds to a delay over one sample period: $q^{-1}\mathbf{x}[n] = \mathbf{x}[n-1]$.

Consider a (OFDM or single-carrier) CP block transmission system with N samples per block. The introduction of a cyclic prefix of K samples means that the last K samples of the current block (corresponding to N samples) are repeated before the actual block. If we assume w.l.o.g. that the current block starts at time 0, then samples $\mathbf{x}[N-K]\cdots\mathbf{x}[N-1]$ are repeated at time instants $-K, \ldots, -1$. This means that the output at sample periods $0, \ldots, N-1$ can be written in matrix form as

$$\begin{bmatrix}\mathbf{y}[0]\\ \cdots \\ \mathbf{y}[N-1]\end{bmatrix} = \mathrm{Y}[0] = \mathrm{H}\,\mathrm{X}[0] + \mathrm{V}[0] \quad (2)$$

where the matrix H is not only (block) Toeplitz but even (block) circulant: each row is obtained by a cyclic shift to the right of the previous row. Consider now applying an N-point FFT to both sides of (2) at block m:

$$F_{N,p}\mathrm{Y}[m] = F_{N,p}H\,F_N^{-1} \quad F_N\mathrm{X}[m] + F_{N,p}\mathrm{V}[m] \quad (3)$$

or with new notations:

$$\mathrm{U}[m] \quad = \quad \mathcal{H} \qquad \mathrm{A}[m] \quad + \quad \mathrm{W}[m] \quad (4)$$

where $F_{N,p} = F_N \otimes I_p$ (Kronecker product: $A \otimes B = [a_{ij}B]$), F_N is the N-point $N \times N$ DFT matrix, $\mathcal{H} = \mathrm{diag}\,\{\mathbf{h}_0, \ldots, \mathbf{h}_{N-1}\}$ is a block diagonal matrix with diagonal blocks $\mathbf{h}_k = \sum_{l=0}^{L}\mathbf{h}[l]\; e^{-j2\pi\frac{1}{N}kl}$, the $p \times 1$ channel transfer function at tone k (frequency $= k/N$ times the sample frequency). In OFDM, the transmitted symbols are in $\mathrm{A}[m]$ and hence are in the frequency domain. The corresponding time domain samples are in $\mathrm{X}[m]$. The OFDM symbol period index is m. In Single-Carrier (SC) CP systems, the transmitted symbols are in $\mathrm{X}[m]$ and hence are in the time domain. The corresponding frequency domain data are in $\mathrm{A}[m]$. The components of V are considered white noise, hence the components of W are white also. At tone (subcarrier) $n \in \{0, \ldots, N-1\}$ we get the following input-output relation

$$\underbrace{\mathbf{u}_n[m]}_{p\times 1} = \underbrace{\mathbf{h}_n}_{p\times q}\,\underbrace{a_n[m]}_{q\times 1} + \underbrace{\mathbf{w}_n[m]}_{p\times 1} \quad (5)$$

where the symbol $a_n[m]$ belongs to some finite alphabet (constellation) in the case of OFDM.

3. SOME GENERALITIES FOR CP SYSTEM METHODS

In what follows, we shall see that for methods and performance analysis, we get a cost function or information at each tone for the channel response at that tone, and to get the cost function or information for the temporal channel response, it suffices to sum up the cost functions or informations over the tones after transforming back to the time domain. To be a bit more explicit, let $\overline{\mathbf{h}}_k = \mathrm{vec}(\mathbf{h}_k)$ and let $\mathbf{h}$ be the vectorized channel impulse response. Then there exists transformation matrices G_k (containing DFT portions) such that

$$\overline{\mathbf{h}}_k = G_k\,\mathbf{h}\,. \quad (6)$$

Now, if at tone k we have a cost function of the form

$$\overline{\mathbf{h}}_k^H \; Q_k \; \overline{\mathbf{h}}_k \tag{7}$$

then this induces a cost function for the overall channel impulse response of the form

$$\mathbf{h}^H \left[\sum_{k=0}^{N-1} G_k^H \; Q_k \; G_k \right] \; \mathbf{h} \tag{8}$$

and similarly for Fisher information matrices. So in what follows, we shall concentrate on the cost function for a given tone.

4. DETERMINISTIC SYMBOLS CASE

Algorithms that fall under this category are

- subspace fitting (MIMO)
- Subchannel Response Matching (SRM)/ Cross Relation (CR) method (SIMO)
- DML, IQML, DIQML, PQML (SIMO)
- singular prediction parameters (MIMO): $P(z)\,H(z) = \mathbf{h}[0] \;\Rightarrow\; (\mathbf{h}^{\perp}[0]\,P(z))\;H(z) = 0$
- deterministic approach by itself of limited use in MIMO case unless channels of different sources of same length: the case of spatial multiplexing MIMO systems

5. SIGNAL SUBSPACE FITTING

Let us focus in particular on the signal subspace fitting method (noise subspace fitting can be similarly formulated for the SIMO case using the linear noise subspace parameterization in terms of the channel, considered in the next section. For the (spatiotemporally) white noise case (and assuming spatiotemporally white symbols for simplicity), the eigendecomposition of the covariance matrix of a block of signal in the time domain can in fact easily be computed from the eigendecompositions at each tone! Indeed

$$\begin{aligned} R_{\mathrm{YY}} &= \sigma_x^2 \, \mathrm{H}\,\mathrm{H}^H + \sigma_v^2 \, I_{Np} \\ \Rightarrow \;\; & F_{N,p} R_{\mathrm{YY}} F_{N,p} \\ &= \sigma_x^2 \, F_{N,p} \mathrm{H} F_{N,p}^{-1} F_{N,p} \mathrm{H}^H F_{N,p}^{-1} + \sigma_v^2 \, F_{N,p} F_{N,p}^{-1} \\ &= F_{N,p} \left[\sigma_x^2 \, \mathcal{H}\,\mathcal{H}^H + \sigma_v^2 \, I_{Np} \right] F_{N,p}^{-1} \end{aligned} \tag{9}$$

where the matrix in square brackets is block diagonal. Hence the eigenvectors in the time domain are the DFTs of the eigenvectors at each tone, and the eigenvalues are the same in time or frequency domain. This exact relationship no longer holds for the eigenvectors based on sample covariances in time and frequency domain due to the noise (it remains true in the absence of noise). Nevertheless this relationship encourages us to develop subspace fitting problems in the frequency domain, involving eigendecompositions of N $p \times p$ matrices instead of the eigendecomposition of one $Np \times Np$ matrix. Furthermore, it is clear that the eigenvectors do not change in the frequency domain if the noise variance starts to differ from tone to tone (this would correspond to cyclostationary colored noise in the time domain, which can model stationary colored noise with some approximation if N is large). Let $\widehat{\mathrm{E}}$ denote a sample average, then the details of the signal subspace fitting method are

- $\mathbf{r}_k = \mathrm{E}\,\mathbf{u}_k[n]\mathbf{u}_k^H[n] = \sigma_a^2\,\mathbf{h}_k\mathbf{h}_k^H + \sigma_{w_k}^2 I_p$ $= V_{\mathcal{S},k}\Lambda_{\mathcal{S},k}V_{\mathcal{S},k}^H + \sigma_{w_k}^2 V_{\mathcal{N},k}V_{\mathcal{N},k}^H$
- $\widehat{\mathbf{r}}_k = \widehat{\mathrm{E}}\,\mathbf{u}_k\mathbf{u}_k^H = \widehat{V}_{\mathcal{S},k}\widehat{\Lambda}_{\mathcal{S},k}\widehat{V}_{\mathcal{S},k}^H + \widehat{V}_{\mathcal{N},k}\widehat{\Lambda}_{\mathcal{N},k}\widehat{V}_{\mathcal{N},k}^H$
- signal subspace fitting cost function

  $$\min_h \sum_{k=0}^{N-1} \|\mathbf{h}_k^H \widehat{V}_{\mathcal{N},k}\|_F^2$$

- cost function per tone $\Rightarrow$ not very costly to introduce optimal weighting

6. MORE DETERMINISTIC APPROACHES

- Deterministic (symbols) ML (DML):

  $$\max_h \sum_{k=0}^{N-1} \mathrm{tr}\,\{P_{\mathbf{h}_k}\widehat{\mathbf{r}}_k\} \;\Leftrightarrow\; \min_h \sum_{k=0}^{N-1} \mathrm{tr}\,\{P_{\mathbf{h}_k}^{\perp}\widehat{\mathbf{r}}_k\}$$

- IQML: in the SIMO case, we can introduce a linear (in the channel parameters) parameterization of the noise subspace, $\mathbf{h}_k^{\perp}$ so that $P_{\mathbf{h}}^{\perp} = P_{\mathbf{h}^{\perp}} \;\Rightarrow$

  $$\min_h \sum_{k=0}^{N-1} \mathrm{tr}\,\{(\mathbf{h}_k^{\perp H}\mathbf{h}_k^{\perp})^{-1}\mathbf{h}_k^{\perp H}\widehat{\mathbf{r}}_k\mathbf{h}_k^{\perp}\}$$

- Subchannel Response Matching (SRM)/Cross Relation method (CR):

  $$\min_h \sum_{k=0}^{N-1} \mathrm{tr}\,\{\mathbf{h}_k^{\perp H}\widehat{\mathbf{r}}_k\mathbf{h}_k^{\perp}\}$$

- Denoised IQML (DIQML):

  $$\min_h \sum_{k=0}^{N-1} \mathrm{tr}\,\{(\mathbf{h}_k^{\perp H}\mathbf{h}_k^{\perp})^{-1}\mathbf{h}_k^{\perp H}(\widehat{\mathbf{r}}_k - \widehat{\sigma}_{w_k}^2 I_p)\mathbf{h}_k^{\perp}\}$$

  Of course, one can now go further in denoising and replace $\widehat{\mathbf{r}}_k - \widehat{\sigma}_{w_k}^2 I_p$ by its pure signal subspace part.

- WSSF/large sample Gaussian (symbols) ML (GML):

$$\max_{h} \sum_{k=0}^{N-1} \text{tr} \left\{ P_{\mathbf{h}_k} \widehat{V}_{\mathcal{S},k} \widetilde{\widehat{\Lambda}}^2_{\mathcal{S},k} \widehat{\Lambda}^{-1}_{\mathcal{S},k} \widehat{V}^H_{\mathcal{S},k} \right\}$$

7. GAUSSIAN SYMBOLS APPROACHES

- Tone-wise covariance analysis

$$\mathbf{r}_k = \mathrm{E}\, \mathbf{u}_k[n]\mathbf{u}_k^H[n] = \sigma_a^2\, \mathbf{h}_k\mathbf{h}_k^H + \sigma_{w_k}^2 I_p$$

  $\Rightarrow$ separate noise variance identifiable at every tone, this corresponds to a circulant noise covariance matrix in the time domain. This also suggests for non-CP systems that $S_{\mathbf{vv}}(z) = S(z)\, I_p$ with some scalar $S(z)$ should be identifiable. In fact, since spatially colored noise can be handled blindly to some extent, it should also be possible to handle $S_{\mathbf{vv}}(z) = S(z)\, R_{\mathbf{vv}}$ for some spatial covariance matrix $R_{\mathbf{vv}}$.

- GML: has the same gradient as WCM below.
- Weighted Covariance Matching (WCM):

$$\min_{h,\sigma^2} \sum_{k=0}^{N-1} \text{tr} \left\{ \mathbf{r}_k^{-1}(\mathbf{r}_k - \widehat{\mathbf{r}}_k)\mathbf{r}_k^{-1}(\mathbf{r}_k - \widehat{\mathbf{r}}_k) \right\}$$

- Linear prediction based methods:

$$P(z)\, H(z) = \mathbf{h}[0]$$

  becomes

$$P_k\ \mathbf{h}_k = \mathbf{h}[0]$$

  tonewise.

- For MIMO, the proper exploitation of the Gaussian case is quite advantageous over the deterministic symbols approach.

8. ZERO PADDED (ZP) SYSTEMS

These are a special case of CP systems in which the cyclic prefix signal is zero. The result can be considered as zero padding or guard intervals. In the ZP case, the description of the signal as in (2) changes. Due to the fact that the last K (block) entries of X are zero, the non-zero part of X multiplies a reduced form of H in which the last K (block) columns have been removed. The result is a tall banded block triangular block Toeplitz matrix $\overline{\text{H}}$. As a result of this reduction in signal, the signal subspace dimension has increased whereas the noise subspace dimension has correspondingly increased. Nevertheless a fast computation of the eigenvectors of the theoretical covariance matrix R_{YY} remains possible. The noise subspace vectors comprise those of the CP case, which can be computed by DFT from the noise subspace vectors at the various tones. The extra noise suspace vectors only have non-zero portions in the first and the last K (block) rows, since those are the (two triangular) portions in the last K columns of H that are non-zero. Hence, in the time domain, it suffices to compute a covariance martix of (block) size $2K$, project it on the signal subspace from the CP case (partial DFTs of signal subspaces in the frequency domain) and then compute the noise subspace of the resulting matrix of actual size $2Nq$.

9. SOME CHANNEL IDENTIFIABILITY CONSIDERATIONS

Consider a genuine CP system first. Apart from the fact that convolution becomes circulant instead of linear, not much changes from the blind multichannel estimation problem in non-CP systems. This means that in the deterministic symbols case, only a (non-unique) irreducible factor of the channel can be identified (in fact only $P_{H(z)}$ can be identified (projection onto the column space of $H(z)$)), and the estimation is sensitive to channel length overdetermination. In the Gaussian symbols case on the other hand, robustness to channel length overdetermination and channel zeros arises, and the channel unidentifiablity gets reduced to a block diagonal unitary mixture matrix in general, depending on the relation between the channel lengths seen from the different channel inputs. The remaining ambiguity on possible channel zeros becomes discrete in the SIMO case. In fact, what can be identified more in the Gaussian case w.r.t. the determistic symbols case is the quantity $H^{\dagger}(z)\, H(z)$. The number of tones in the CP symbol should equal or exceed the filter length: $N \geq L{+}1$.

In the ZP case, as the channel convolution matrix $\overline{\text{H}}$ becomes block triangular, the only requirement for it to be full column rank is for $\mathbf{h}[0]$ to be full column rank. In that case, the channel is identifiable up to a static mixture (unitary in the Gaussian symbols case), regardless of channel zeros or channel length overdetermination.

10. CONCLUDING REMARKS

CP systems allow for the development of highly structured blind multichannel estimation algorithms. The resulting limited complexity of these algorithms is such that it becomes affordable to consider implementing the best performing algorithms, for instance based on a combination of subspace decompositions and maximum likelihood techniques. In particular, ML techniques based on Gaussian symbol models become quite accessible in CP systems.

In the CP context, it should be fairly straightforward to answer questions such as to determine the optimal block size N from a subspace based estimation point of view: this is the issue of trading block size for temporal averaging and hence suspace estimation quality.

CP systems simplify blind channel estimation based on deterministic or Gaussian symbol models. However, for OFDM systems, it is clear that furthermore it should remain straightforward to incorporate more precise marginal distribution information for the symbols at the various tones, such as Higher-Order statistics, Constant Modulus properties, Finite Alphabet constraints or other information-theoretic criteria. Along the same lines, the development of semiblind techniques is quite straightforward in OFDM systems, with pilot tones, due to the orthogonality between pilot and data tones.

11. REFERENCES

[1] A. Medles and D.T.M. Slock "Semiblind Channel Estimation for MIMO Spatial Multiplexing Systems". In Proc. *Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, CA, USA, Nov. 2001.

[2] B.M. Sadler, R.J. Kozick, and T. Moore, "Bounds on Bearing and Symbol Estimation with Side Information," *IEEE Transactions on Signal Processing*, vol. 49, no. 4, pp. 822–834, April 2001.

[3] A. Medles, D.T.M. Slock and E. De Carvalho "Linear prediction based semi-blind estimation of MIMO FIR channels". In Proc. *Third Workshop on Signal Processing Advances in Wireless Communications (SPAWC'01)*, pp. 58-61, Taipei, Taiwan, March 2001.

[4] M. Guillaud and D.T.M. Slock "Channel Modeling and Associated Inter-Carrier Interference Equalization for OFDM Systems with High Doppler Spread". In Proc. *IEEE Int'l Conf. on Acoustics, Speech, and Sig. Proc. (ICASSP)*, Hong Kong, 2003.

[5] D. Falconer, S.L. Ariyavisitakul, A. Benyamin-Seeyar and B. Eidson "Frequency-Domain Equalization for Single-Carrier Broadband Wireless Systems". *IEEE Comm. Mag.*, pp. 58-66, Apr. 2002.

[6] P.P. Vaidyanathan and B. Vrcelj "Fast and robust blind-equalization based on cyclic prefix". In Proc. *IEEE Int'l Conf. on Communications (ICC)*, pp. 1-5, May 2002.

[7] P.P. Vaidyanathan and B. Vrcelj "Theory of fractionally spaced cyclic-prefix equalizers". In Proc. *IEEE Int'l Conf. on Acoustics, Speech, and Sig. Proc. (ICASSP)*, pp. 1277-1280, Orlando, FL, USA, 2002.

[8] E. de Carvalho and D.T.M. Slock "Semi-Blind Methods for FIR Multichannel Estimation". Chapter 7 in *Signal Processing Advances in Communications*, Vol. 1: Trends in Channel Estimation and Equalization, G.B. Giannakis, P. Stoica, Y. Hua and L. Tong, editors, Prentice Hall, 2000.

[9] M. Kristensson and B. Ottersten "Optimum Subspace Methods". Chapter 6 in *Signal Processing Advances in Communications*, Vol. 1: Trends in Channel Estimation and Equalization, G.B. Giannakis, P. Stoica, Y. Hua and L. Tong, editors, Prentice Hall, 2000.

[10] H. Wang, Y. Lin and B. Chen "Data Efficient Blind OFDM Channel Estimation using Receiver Diversity". *IEEE Trans. Sig. Proc.*, 2003.

[11] D.T.M. Slock "From Sinusoids in Noise to Blind Deconvolution in Communications". In *Communications, Computation, Control and Signal Processing: a Tribute to Thomas Kailath*, A. Paulraj, V. Roychowdhury and C.D. Schaper, editors, Kluwer Academic Publishers, Boston, 1997.