This document is a an assignment report for the Advance Digital Communication Course, Fall 2007
Isfahan University of Technology, Isfahan, Iran - Instructor Prof. Omidi

Blind Channel Estimation using Subspace and LMMSE for SIMO OFDM Systems: A Review

Norooz Motamedi

***Abstract*—A review of two blind channel estimation algorithms is presented. The subspace method and the LMMSE approach is reviewed in this paper. We outline basic ideas behind several developments, the assumptions and identifiably conditions required by these approaches, and the algorithm characteristics.**

***Index Terms*—Blind Channel Estimation, Equalization**

I. Introduction

INTERSYMBOL interference (ISI) is a limiting factor in many communication systems. ISI can arise from time-varying multi-path fading, which can be severe in, for example, a mobile communication system. Other channel impairments that contribute to ISI include symbol clock jitter, carrier phase jitter, etc. To achieve high-speed reliable communication, channel estiamtion and equalization are necessary to overcome the effects of ISI [1].

Designs of receivers that remove channel distortions require either the knowledge of the channel or the access to the "training" signals. The latter is the choice in many communication systems design. The transmission of training signals decreases communications throughput although, for time invariant channels, the loss is insignificant because only one training sequence is necessary. For time varying channels, however, the loss of throughput becomes a problem [2].

The blind channel estimation means that the channel is estimated without any training sequence; instead, the identification is achieved by using only the channel output with certain a priori statistical information on the input. Such methods can increase the transmission capability due to the elimination of training signals [3].

Earlier approaches to blind identification use the higher-order statistics of the output. These methods, although reliable and robust in some applications, require a large number of data samples and a large amount of computation. In fast changing environments, such as in cellular communications, their applications may be limited. The method proposed by Tong et al. [4] solved this problem. This method explored the cyclostational properties of an over-sampled communication signal and cause the blind channel estimation to be accomplished based on second-order statistics (SOS) of the channel output. Since the SOS of scalar system output do not contain enough information to identify a possibly non-minimum phase system, and since the temporal over-sampling technique converts a stationary communication sequence into a cyclostationary process, it was, for a while, believed that cyclostationarity was the only reason to the success of the algorithm [5].

Classical solutions of blind identification in digital communication systems are based upon data sampled at the baud rate, although it has been known that fractionally spaced equalizers are more robust under timing uncertainties. Since communication channels are non-minimum phase generally, the SOS of baud-rate-sampled stationary signals is inadequate for channel identification. The phase information is available in the cyclostationary sequence [5].

Some of the early statistics-based methods suffer from the performance degradation caused by the model mismatch when only a limited number of observations are available. The desire for more data-efficient algorithm led to the development of a class of subspace based [6] blind identification algorithms. These techniques significantly outperform several previously statistics-based methods, especially for short data sequences [5].

Despite many good features, the performance of these subspace-based algorithms may be basically limited by the nature of the channel. For example, singularity of the channel matrix can cause divergence of the subspace, and result in failure of the subspace approaches [5].

By reviewing recent surveys [2], [5], the purpose of this paper is to review some blind channel estimation approaches. We provide a systematic summary of some algorithms in the area of blind channel estimation. Various existing algorithms are classified into the moment-based and the maximum likelihood (ML) methods. If input is assumed to be random with prescribed statistics, the corresponding blind channel estimation schemes are considered to be statistical. On the other hand, if the source does not have a statistical description, or although the source is random but the statistical properties of the source are not used, the corresponding estimation algorithms are deterministic [2]. Fig. 1 shows a map for different classes of algorithms.

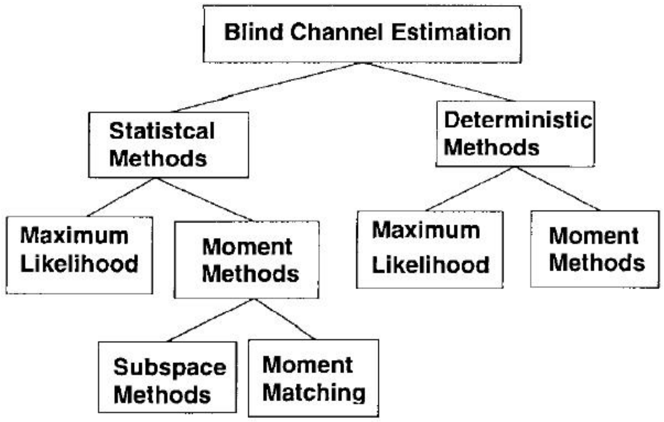

Fig.1. Classification of blind channel estimator [2]

II. CHANNEL ESTIMATION USING SUBSAPCE METHOD

Many recent blind channel estimation techniques exploit subspace structures of observation. The key idea is that the channel vector is in a one-dimensional subspace of the observation statistics. These methods, which are often referred to as subspace algorithms, have the attractive property that the channel estimates can often be obtained in a closed form from optimizing a quadratic cost function. Subspace methods can sometimes be considered part of the moment methods [2].

A. Problem Formulation

Let x_n, denote the symbol emitted by the digital source at time *nT, (T* is the symbol duration). The base-band signal at receiver is given [3] by:

$$y(t)=\sum_{m=-\infty}^{+\infty} x_m h(t-mT)+w(t) \tag{1}$$

Where *w(t)* is a band-limited complex stationary process, assumed to be independent from the emitted symbols, and *h(t)* is the overall response of the transmission filter, receiver filter, channel response, and modulation/demodulation.

Taking into account that the channel has finite support, the complex envelope of the signal received on the i[th] sensor after sampling is:

$$y_n^{(i)}=\sum_{m=0}^{M} x_{n-m} h_m^{(i)}+w_n^{(i)},\ \ 0\le i<L \tag{2}$$

$$y_n^{(i)}=y(t_0+nT); w_n^{(i)}=w(t_0+nT); h_m^{(i)}=h(t_0+mT)$$

Each sequence $y_n^{(i)}$ depends on a discrete-time impulse response $H^{(i)}$ characterizing the i[th] channel:

$$H^{(i)}\stackrel{def}{=}[h_0^{(i)},h_1^{(i)},\ldots,h_M^{(i)}]^T \tag{3}$$

Stacking N successive samples of the received signal sequence, we obtain:

$$Y_n^{(i)}=H_N^{(i)}X_n+W_n^{(i)}$$

$$Y_n^{(i)}=[y_n^{(i)},y_{n-1}^{(i)},\ldots,y_{n-N+1}^{(i)}]^T$$

$$H_N^{(i)}=\begin{bmatrix} h_0^{(i)} & \ldots & h_M^{(i)} & 0 & \ldots & \ldots & 0 \\ 0 & h_0^{(i)} & \cdots & h_M^{(i)} & 0 & \ldots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 & h_0^{(i)} & \cdots & h_M^{(i)} \end{bmatrix}$$

$$Dim.: N\times(N+M)$$

$$X_n=[x_n,x_{n-1},\ldots,x_{n-N-M+1}]^T;\quad Dim.:(N+M)\times 1$$

$$W_n^{(i)}=[w_n^{(i)},w_{n-1}^{(i)},\ldots,w_{n-N+1}^{(i)}]^T$$

Hence, the set of measurements depending on the same set of input symbols is given by:

$$\begin{pmatrix} Y_n^{(0)} \\ \vdots \\ Y_n^{(L-1)} \end{pmatrix}=\begin{pmatrix} H_N^{(0)} \\ \vdots \\ H_N^{(L-1)} \end{pmatrix}X_n+\begin{pmatrix} W_n^{(0)} \\ \vdots \\ W_n^{(L-1)} \end{pmatrix} \tag{5}$$

This linear system has dimension of LN×(M+N) [6].

B. Subspace-based Identification

A blind identification procedure consists in estimating the L(M+1)×1 vector H of channel coefficients:

$$H=[H^{(0)T},\ldots,H^{(L-1)T}]^T \tag{6}$$

From (5), we can see that:

$$Y_n=H_N X_n+W_n \tag{7}$$

So, the identification is based on the LN×LN autocorrelation matrix of the measurement vector:

$$R_{yy}=E\{Y_n Y_n^H\} \tag{8}$$

Since the additive noise is assumed independent of the emitted sequence, the autocorrelation matrix is expressed:

$$R_{yy}=H_N R_{xx} H_N^H+R_{ww} \tag{9}$$

The source covariance matrix R_{xx} has dimension *(N+M)×(N+M), a*nd is assumed to be full rank but otherwise unknown. The noise covariance matrix R_{ww} is of size *LN×LN*. To ease the derivations, the noise is assumed to be white [6].

C. Subspace Decomposition

Let $\lambda_0\ge\lambda_1\ge\cdots\ge\lambda_{LN-1}$ denote the eigenvalues of R_{yy}. Since R_{xx} is full rank, the signal part of autocorrelation matrix $H_N R_{xx} H_N^H$ has rank M+N, since $R_{ww}=\sigma^2 I$, hence:

$$\begin{aligned} \lambda_i &> \sigma^2 \quad for \quad i=0,\cdots,M+N-1 \\ \lambda_i &= \sigma^2 \quad for \quad i=M+N,\cdots,LN-1 \end{aligned} \tag{10}$$

Denote the unit norm eigen-vectors related with the eigenvalues $\lambda_0,\ldots,\lambda_{M+N-1}$ by $S_0,\ldots,S_{M+N-1}$ and those corresponding to $\lambda_{M+N},\ldots,\lambda_{LN-1}$ by $G_0,\ldots,G_{LN-M-N-1}$

The autocorrelation matrix is thus also expressed as:

$$R_{yy} = S\,diag(\lambda_0,\cdots,\lambda_{M+N-1})S^H + \sigma^2 GG^H$$
$$S = [S_0,\ldots,S_{M+N-1}] \qquad LN\times(M+N) \tag{11}$$
$$G = [G_0,\ldots,G_{LN-M-N-1}] \quad LN\times(LN-M-N)$$

The columns of matrix *S* span the signal subspace (dimension M+N), while the columns of *G* span its orthogonal complement, the noise subspace. Also, the signal subspace is the spanned by the columns of the filtering matrix H_N. By orthogonality between the noise and the signal subspace, the columns of H_N are orthogonal to any vector in the noise subspace. Hence, we have:

$$G_i^H H_N = 0 \qquad 0 \le i < LN - M - N \tag{12}$$

Under some conditions detailed in the theorem below, the noise subspace (matrix *G*) uniquely determines the channel coefficients up to a multiplicative constant.

Theorem 1: assume that N≥M and matrix H_{N-1} is full rank. Let $H_N^{'}$ be a new nonzero matrix with same dimension as H_N. The range of new matrix is included in the range of H_N iff the corresponding *H* and $H^{'}$ are proportional. Hence, both of matrices share the same column space iff they are proportional.

We show in the next section how by using this theorem the channel coefficients can be estimated even in the cases where R_d is full rank and unknown [6].

D. Subspace-Based Parameter Estimation Scheme

In practice, sample estimates $\hat{G}_i$, of the noise eigenvectors are available and (12) is solved in the least squares sense and leads to minimize the following quadratic form:

$$q(H) = \sum_{i=0}^{LN-M-N-1} \left|\hat{G}_i H_N\right|^2 \tag{13}$$

As you can see, *q(H)* depend on vector *H* rather than on the filtering matrix H_N. This is conveniently done by application of following Lemma, which requires the following notations.

Notations: Let $V^{(0)},\ldots,V^{(L-1)}$ be L arbitrary N×1 vectors and let V be the LN×1 vector defined as $V=[V^{(0)T},\ldots,V^{(L-1)T}]^T$. Denote

$$V_{M+1}^{l} = \begin{pmatrix} v_0^{(l)} & \cdots & v_{N-1}^{(l)} & 0 & \cdots & \cdots & 0 \\ 0 & v_0^{(l)} & \cdots & v_{N-1}^{(l)} & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 & v_0^{(l)} & \cdots & v_{N-1}^{(l)} \end{pmatrix}$$

$$V_{M+1} = [V_{M+1}^{(0)T},\cdots,V_{M+1}^{(L-1)T}]^T \quad Dim.: L(M+1)\times(M+N)$$

Lemma 1: $V^T H_N = H^T V_{M+1}$

By using Lemma, (13) can be written as follows:

$$q(H) = \sum_{i=0}^{LN-M-N-1} \hat{G}_i^H H_N H_N^H \hat{G}_i = \sum_{i=0}^{LN-M-N-1} H^H \hat{g}_i \hat{g}_i^H H$$
$$q(H) = H^H Q H;\; Q = \sum_{i=0}^{LN-M-N-1} \hat{g}_i \hat{g}_i^H \tag{15}$$

Where $\hat{g}_i$ is the L(M+1)×(M+N) filtering matrix associated with the vector $\hat{G}_i$, defined according to (14).

By theorem 1, if true autocorrelation matrix was available, the true channel coefficients are the unique (up to a scalar factor) vector *H* such that *q(H)*=0. In contrast, when only an *estimate* of the autocorrelation matrix is available, the quadratic form has not exactly rank L(M+1). Hence, estimation of *H* can be obtained by minimizing *q(H)* subject to a properly chosen constraint avoiding the trivial solution *H*=0. Different constraints on *H* provide different solutions. We have classically considered minimization subject to linear and quadratic constraints:

- *Quadratic constraint:* Minimize *q(H)* subject *to* |H|=1. The solution is the unit-norm eigenvector related to the smallest eigenvalue of matrix *Q*.
- *Linear constraint:* Minimize *q(H)* subject to $c^H H = 1$. Where *c* is a L(M+1)×1 vector. The solution is proportional to $Q^{-1}C$.

The first choice is more natural but involves the computation of an additional eigenvector. The second solution depends on the choice of an arbitrary constraint vector c. The computational cost of the second solution is lower since it amounts to solving a linear system rather than extracting an eigenvector [6].

E. Signal Subspace

It is shown before that minimizing a constrained quadratic form involving the noise eigenvectors give the channel coefficients. This quadratic form is equivalently rewritten in terms of the signal eigenvectors as:

$$\begin{aligned} q(H) &= N|H|^2 - \sum_{i=0}^{M+N-1} \left|\hat{S}_i^H H_N\right|^2 \\ &= N|H|^2 - H^H\left(\sum_{i=0}^{M+N-1} \hat{s}_i \hat{s}_i^H\right)H \\ &= N|H|^2 - H^H \tilde{Q} H \end{aligned} \tag{16}$$

Where $\hat{s}_i$ denotes the filtering matrix of size L(M+1)×(N+M) associated to eigenvector $\hat{S}_i$. The minimization of (13) under the constraint |H|=1 is thus equivalent to the maximization of $\tilde{q}(H) = H^H \tilde{Q} H$ under the same constraint. This maximization is easily implemented by looking for the maximum eigenvalue of $\tilde{Q}$.

Under the unit norm constraint, both the noise and signal subspace give identical solutions. However, computing the coefficients of the quadratic form involves LN-N-M terms in the former case and M+N in the latter [6].

F. Deterministic Subspace Approach

Subspace method based on the property that the channel is in a unique direction. It may not be robust against modeling errors, especially when the channel matrix is close to being singular. The second disadvantage is that they are often more computationally expensive [2].

Deterministic subspace methods do not use a priori statistical information. A more useful property of deterministic subspace methods is the finite sample convergence property. Without presence of noise, the estimator produces the exact channel using only a finite number of samples if the identifiability condition is satisfied. Therefore, these methods are most effective at high SNR and for small data sample applications. On one hand, deterministic methods can be applied to a much wider range of source signals; on the other hand, not using the source statistics affects its asymptotic performance [2].

III. LMMSE CHANNEL ESTIMATION

LMMSE is widely used in the OFDM channel estimation since it is optimum in minimizing the MSE of the channel estimates in the presence of AWGN. LMMSE uses additional information like the SNR. LMMSE is a smoother/ interpolator /extrapolator, and hence is very attractive for the channel estimation of OFDM based systems with pilot sub-carriers. However, the computational complexity of LMMSE is very high due to extra information incorporated in the estimation technique [8].

A. System Description

We assume that the use of a cyclic prefix (CP) both preserves the orthogonality of the tones and eliminates ISI between consecutive OFDM symbols. Further, the channel is assumed to be slowly fading, so it is considered to be constant during one OFDM symbol. The number of tones in the system is N.

Under these assumptions we can describe the system as a set of parallel Gaussian channels, with correlated attenuations. The attenuations on each tone are given by h_k.

In matrix notation we describe the OFDM system as

$$y = Xh + w \tag{17}$$

Where y is the received vector, X is a diagonal matrix containing the transmitted signaling points, *h* is a channel attenuation vector, and *w* is a vector of IID complex zero-mean Gaussian noise. LMMSE of the variable h is given [9] by:

$$\hat{h} = R_{yh} R_{yy}^{-1} y \tag{18}$$

Where R_{yh} is the cross-correlation between variables y and h.

The LMMSE estimate of the channel given the received data and the transmitted symbols, is [10]:

$$\begin{aligned} \hat{h}_{LMMSE} &= R_{hh}(R_{hh} + \sigma^2 (XX^H)^{-1})^{-1} \hat{h}_{LS} \\ \hat{h}_{LS} &= X^{-1} y \end{aligned} \tag{19}$$

Where $\hat{h}_{LS}$, is the Least-Square (LS) estimate of h and R_{hh} is the channel autocorrelation matrix. Without loss of generality, we assume that the variance of the channel attenuation is normalized to unity.

The LMMSE estimator (19) is very complex since a matrix inversion is needed every time the data in changes. We reduce the complexity by averaging over the transmitted data. We replace the term $(XX^H)^{-1}$ in with its expectation $E\{(XX^H)^{-1}\}$. Simulations indicate that the performance degradation is negligible [10]. Assuming the same signal constellation on all tones and equal probability on all constellation points, we have $E\{(XX^H)^{-1}\}=E\{|1/x_k|^2\}I$. Defining the average SNR as $E\{|x_k|^2\}/\sigma^2$, we obtain the simplified estimator:

$$\begin{aligned} \hat{h}_{LMMSE} &= R_{hh}(R_{hh} + \frac{\beta}{SNR} I)^{-1} \hat{h}_{LS} \\ \beta &= E\{|x_k|^2\} E\{|1/x_k|^2\} \end{aligned} \tag{20}$$

Where β is a constant depending on the signal constellation.

Because X is no longer a factor in the matrix calculation, the matrix inversion does not need to be calculated with data changing. Furthermore, if channel autocorrelation and SNR are known the channel estimated matrix needs to be calculated only once. Under these conditions the estimation requires N multiplications per tone. To further reduce the complexity, we proceed with the low-rank approximations below [10].

B. Optimal Low-Rank Approximations

Optimal rank reduction is achieved by using the singular value decomposition (SVD). The SVD of the channel autocorrelation matrix is:

$$R_{hh} = U \Lambda U^H \tag{21}$$

Where U is a unitary matrix containing the singular vectors and Λ is a diagonal matrix containing the singular values $\lambda_1 \geq \cdots \geq \lambda_N$ on its diagonal [10].

Since the delay spread in OFDM is usually much less than the symbol duration to remove ISI, the channel frequency response at different frequencies are highly correlated. So, a few singular values have value significantly larger than zero [11]. It is shown [10] that the optimal rank-p estimator with p largest singular values is:

$$\hat{h}_p = U \Delta_p U^H \hat{h}_{LS} \tag{22}$$

Where Δ_p is a diagonal matrix with entries:

$$\delta_k = \begin{cases} \dfrac{\lambda_k}{\lambda_k + \dfrac{\beta}{SNR}}, k = 1, \cdots, p \\ 0, \qquad k = p+1, \cdots N \end{cases} \tag{23}$$

The dimension of the space (p) of time and band-limited signals is needed in the low-rank estimator. It is shown that this dimension is about 2BT+1, where B is the bandwidth and T is the time interval of the signal. Accordingly, the magnitude of the singular values should become small after about L+1 values, where L is the length of the CP ($2B=1/T_s$, $T=LT_s$ and $2BT+1=L+1$) [10].

The low-rank estimator can be interpreted as first projecting the LS estimates onto a subspace and then performing the estimation. If the subspace has a small dimension and can describe the channel well, the complexity of the estimator will be low while showing a good performance. A block diagram of the rank-p estimator in (22) is shown in Fig.2.

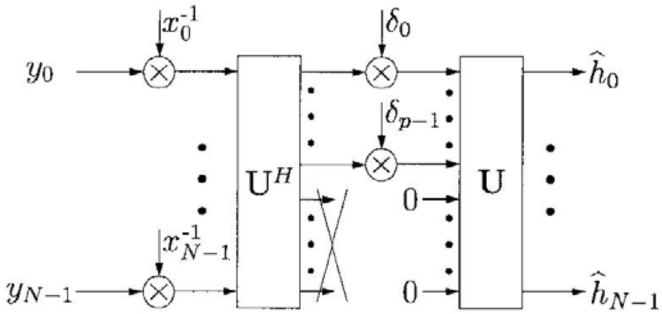

Fig.2 block diagram of the rank-p estimator [10]

IV. CONCLUSION

Channel estimation is a standard linear system identification problem with the training sequence as the pilot input signal. In many applications, the pilot signals may not be easy to use or they may present an extra problem, for example requiring more bandwidth in communication systems. Blind channel estimation and equalization eliminates the need for a pilot signal and simplifies the requirements for channel estimation and equalization. In particular, recent developments in blind estimation research have led to a class of rapidly converging and data efficient algorithms that can effectively estimate the channel with a small number of data points. In this paper, we reviewed some of the basic approaches in blind estimation.

REFERENCES

[1] Tong L., et al, "Blind Identification and Equalization based on Second-Order Statistics: A Time Domain Approach," IEEE Transaction on Information Theory, Vol.40, No.2, pp.340-349, March 1994.

[2] Tong L., Perreau S., " Multichannel Blind Identification: From Subspace to Maximum Likelihood Methods", Proceeding of the IEEE, Vol.86, No. 10, pp.1951-1968, October 1998.

[3] Tong L., et al, "Blind Channel Identification based on Second-Order Statistics: A Frequency Domain Approach," IEEE Transaction on Information Theory, Vol.41, No.1, pp.329-334, January 1995.

[4] Gardner W.A., *Cyclostationarity in Communications and Signal Processing*, IEEE press, 1994.

[5] Liu Hui, et al, " Recent development in blind channel equalization: from cyclostationarity to subspace", Signal Processing, Vol.50, pp.83-99, 1996.

[6] Moulines E., et al, "Subspace Methods for Blind Identification of Multichannel FIR Filters," IEEE Transaction on Signal Processing, Vol.43, No.2, pp.516-525, February 1995.

[7] Proakis J., D*igital Communication,*. 4th Ed, McGraw Hill, 2001.

[8] Ozdemir M.K., Arslan H., " Channel Estimation for Wireless OFDM Systems", IEEE Communications Surveys and Tutorials, pp.18-48, 2nd Quarter 2007.

[9] Kailath T., *Lectures on Wiener and Kalman Filtering,* Springer, 1985.

[10] Edfords O., et al, "OFDM Channel Estimation by Singular Value Decomposition", IEEE Transaction on Communications, Vol.46, No.7, pp. 931-939, July 1998.

[11] Li Ye (Geoffrey), Stuber G., *Orthogonal Frequency Division Multiplexing for Wireless Communications,* Springer, 2006.