# MAXIMUM LIKELIHOOD MULTICHANNEL ESTIMATION UNDER REDUCED RANK CONSTRAINT

*P. Forster, T. Asté*

Laboratoire Electronique et Communication, CNAM, 292 rue St Martin,75141 Paris Cedex 03, France
e-mail:forster,aste@cnam.fr

**ABSTRACT**

This paper deals with the maximum likelihood estimation of the multichannel impulse response in a mobile communication system whose base stations are equipped with antennas arrays. The following problem is solved: using the training sequence, find the maximum likelihood multichannel impulse response from one mobile to the base station under a reduced rank constraint in the presence of gaussian noise and jammers with unknown covariance matrix. Our results find applications in equalization (the reduced rank channel estimate can be used in a Viterbi algorithm), and in the estimation of the directions of arrival (DOA) of the paths from the mobile to the base station. In this last application, a MUSIC like algorithm is developped using the estimated channel subspace.

## 1. INTRODUCTION

There has been recently an increasing interest for using antennas arrays at base stations in mobile communication systems like GSM. In these systems, estimation of the mobile-to-base multichannel impulse response from the training sequence is a preliminary step before equalization. However, when propagation occurs through a limited number of path, this multichannel impulse response exhibits some simple structure: its rank is equal to the number of paths which can be much less than the number of antennas. We develop in this context a maximum likelihood estimation of this reduced rank multichannel impulse response in presence of gaussian noise *with unknown covariance matrix*. As shown in this paper, a closed-form solution to that problem exists which does not require optimization techniques: this is in contrast with the full modelling problem that is addressed in [1]. Our result finds applications in equalization (the reduced rank channel estimate can be used in a Viterbi algorithm), and in the estimation of the directions of arrival of the paths from the mobile to the base station. In this second application, a MUSIC like algorithm can be developped using the estimated channel subspace: this algorithm is a simple alternative to the complex iterative procedure developped in [1].

This paper is organised as follows. Section 2 formulates the problem, whose solution is given in section 3. In section 5, applications in equalization are presented using the reduced rank channel estimate in a Viterbi equalizer. Finally, an application to the estimation of the paths DOA's is developped in section 6 together with some simulations results.

Due to limited space, the estimation of the number of paths is not treated here.

## 2. NOTATIONS AND PROBLEM FORMULATION

Consider an array of $m$ sensors that receives the signals emitted by a mobile in a Time Division Multiple Access (TDMA) communication system (e.g. GSM) over a frequency selective channel. These signals are corrupted by noise due to jammers and receivers so that the sampled array output can be written during the $l$-th communication burst :

$$\mathbf{y}_l(t) = \mathbf{H}_l\,\mathbf{x}_l(t) + \mathbf{e}_l(t) \tag{1}$$

where:

- $\mathbf{H}_l$ is the $m \times q$ channel impulse response over the $l$-th burst ;
- $\mathbf{x}_l(t) = [x_l(t) \cdots x_l(t-q+1)]^T$ is a vector built from the information bearing signals $x_l(t)$ ;
- $\mathbf{e}_l(t)$ is the vector of noise at array inputs ;
- $t$ varies from 1 to the burst length $N_b$ ;
- $l$ varies from 1 to the number $L$ of processed bursts.

The channel maximum length $q$ is assumed to be known: for instance, it corresponds to the duration of 5 symbols in GSM. The channel impulse response $\mathbf{H}_l$ varies from one burst to another, but physical considerations and experimental results [2] allow us to assume that propagation from the mobile to the array occurs through a limited number of paths whose directions remain *unchanged* over the processed bursts

This work was supported by France Télécom, CNET, 38-40 rue du G$^a$ $l$ Leclerc, 92131 Issy Les Moulieaux Cedex, France

(as long as their number is not to large). So , by extending the model used in [1] over successive bursts, we can write:

$$\mathbf{H}_l = \mathbf{A}\,\mathbf{B}_l \tag{2}$$

where $\mathbf{A}$ is the $m \times n$ paths steering vectors matrix, $n$ is the number of paths from the mobile to the array, and $\mathbf{B}_l$ is an unknown $n \times q$ complex matrix that changes from one burst to another due to the mobile displacement. Therefore, all the columns of $\mathbf{H}_l$ for $l = 1, \cdots, L$ lie in the same $n$-dimensional subspace spanned by the columns of $\mathbf{A}$ and referred to as the *signal-subspace*. Hence, we make the following assumption.

*Assumption 1:* The $m \times L\,q$ matrix formed with channel impulse responses from L successive bursts

$$\mathbf{H} = [\mathbf{H}_1 \cdots \mathbf{H}_L] \tag{3}$$

has rank $n$, where the number $n$ of paths satisfies $n < min(m, L\,q)$.

The next assumption is about noise and is essentially the same as the corresponding one in [1]: again, we implicitly assume that the jammers geometry remains unchanged over the processed bursts.

*Assumption 2:* Noise $\mathbf{e}_l(t)$ is zero-mean, Gaussian, circular, uncorrelated from one time instant to another, with unknown covariance matrix $\mathbf{Q}$.

*Assumption 3:* The emitted signals $x_l(t)$ are known (e.g. training sequence).

As shown in [1], the Maximum Likelihood Estimation (MLE) of the steering vector matrix $\mathbf{A}$ is a highly non-linear problem that has no explicit solution : one must resort to complex optimization techniques. It turns out, as demonstrated in the next section, that a much more simple solution can be obtained if we look only for a MLE of the signal subspace without taking into account the steering-vectors structure: in other words, we perform an *unstructured* estimation of the signal subspace. This subspace can then be utilized for estimating the paths DOA's by a MUSIC like scheme. When obtained from the training sequence, the corresponding reduced rank channel impulse responses could also be used for demodulation with a Viterbi algorithm. Now, the addressed problem can be stated as follows.

*Problem Formulation:* Given L bursts of data, calculate the rank n maximum likelihood estimate of $\mathbf{H} = [\mathbf{H}_1 \cdots \mathbf{H}_L]$ and the noise covariance matrix $\mathbf{Q}$.

## 3. PROBLEM SOLUTION

Let us define the following covariance matrices:

$$\widehat{\mathbf{R}}_{xx_l} = N_b^{-1} \sum_{t=1}^{N_b} \mathbf{x}_l(t)\mathbf{x}_l^*(t)\,, \tag{4}$$

$$\widehat{\mathbf{R}}_{yy_l} = N_b^{-1} \sum_{t=1}^{N_b} \mathbf{y}_l(t)\mathbf{y}_l^*(t)\,, \tag{5}$$

$$\widehat{\mathbf{R}}_{yx_l} = N_b^{-1} \sum_{t=1}^{N_b} \mathbf{y}_l(t)\mathbf{x}_l^*(t)\,, \tag{6}$$

$$\begin{aligned}\widehat{\mathbf{W}} &= L^{-1} \sum_{l=1}^{L} (\widehat{\mathbf{R}}_{yy_l} - \widehat{\mathbf{R}}_{yx_l}\,\widehat{\mathbf{R}}_{xx_l}^{-1}\,\widehat{\mathbf{R}}_{yx_l}^*)\,, \\ &= \widehat{\mathbf{W}}^{1/2}\,\widehat{\mathbf{W}}^{*/2}\end{aligned} \tag{7}$$

$$\widetilde{\mathbf{R}} = L^{-1}\,\widehat{\mathbf{W}}^{-1/2} \left( \sum_{l=1}^{L} \widehat{\mathbf{R}}_{yx_l} \widehat{\mathbf{R}}_{xx_l}^{-1} \widehat{\mathbf{R}}_{yx_l}^* \right) \widehat{\mathbf{W}}^{-*/2}\,. \tag{8}$$

Let $\widehat{\mathbf{\Pi}}_s$ be the projector onto the $n$ largest eigenvalues of $\widetilde{\mathbf{R}}$, and $\widehat{\mathbf{\Pi}}_n = \mathbf{I} - \widehat{\mathbf{\Pi}}_s$ the projector onto the $m-n$ smallest eigenvalues. Then, the rank n maximum likelihood estimate $\widehat{\mathbf{H}} = \left[\widehat{\mathbf{H}}_1 \cdots \widehat{\mathbf{H}}_L\right]$ of $\mathbf{H} = [\mathbf{H}_1 \cdots \mathbf{H}_L]$ and the maximum likelihood estimate $\widehat{\mathbf{Q}}$ of $\mathbf{Q}$ are given by:

$$\widehat{\mathbf{H}}_l = \widehat{\mathbf{W}}^{1/2}\,\widehat{\mathbf{\Pi}}_s\,\widehat{\mathbf{W}}^{-1/2}\,\widehat{\mathbf{R}}_{yx_l}\,\widehat{\mathbf{R}}_{xx_l}^{-1}\,, \tag{9}$$

$$\widehat{\mathbf{Q}} = \widehat{\mathbf{W}} + \widehat{\mathbf{W}}^{1/2}\,\widehat{\mathbf{\Pi}}_n\,\widetilde{\mathbf{R}}\,\widehat{\mathbf{\Pi}}_n\,\widehat{\mathbf{W}}^{*/2}\,. \tag{10}$$

*Proof:* The solution amounts to maximize the log-likelihood function $\mathcal{L}(\mathbf{A}, \mathbf{B}_1, \cdots, \mathbf{B}_L, \mathbf{Q})$ where $\mathbf{A}$ $(m \times n)$, $\mathbf{B}_l$ $(n \times q)$ for $l = 1, \cdots, L$ and $\mathbf{Q}$ (hermitian $m \times m$) are globally unknown:

$$\begin{aligned}\mathcal{L}(\mathbf{A}, \mathbf{B}_1, \cdots, \mathbf{B}_L, \mathbf{Q}) = -m\,L\,N_b\,log\pi - L\,N_b\,log|\mathbf{Q}| - \\ \sum_{l=1}^{L}\sum_{t=1}^{N_b} (\mathbf{y}_l(t) - \mathbf{A}\,\mathbf{B}_l\,\mathbf{x}_l(t))^*\,\mathbf{Q}^{-1}\,(\mathbf{y}_l(t) - \mathbf{A}\,\mathbf{B}_l\,\mathbf{x}_l(t))\,.\end{aligned} \tag{11}$$

Due to limited space, the proof cannot be presented here.

*Remark:* When *no constraint* is set on the $\mathbf{H}_l$'s, the maximum likelihood estimates of $\mathbf{H}_l$ and $\mathbf{Q}$ are:

$$\widehat{\mathbf{H}}_l = \widehat{\mathbf{R}}_{yx_l}\,\widehat{\mathbf{R}}_{xx_l}^{-1} \tag{12}$$

$$\widehat{\mathbf{Q}} = \widehat{\mathbf{W}} \tag{13}$$

Note that $\widehat{\mathbf{R}}_{yx_l}\,\widehat{\mathbf{R}}_{xx_l}^{-1}$ in (12) is also the MMSE estimate of $\mathbf{H}_l$ obtained by minimizing

$$MSE = \sum_{t=1}^{N_b} \|\mathbf{y}_l(t) - \mathbf{H}_l\,\mathbf{x}_l(t)\|^2$$

with respect to $\mathbf{H}_l$.

## 4. INTERPRETATION

We derive in this section some basic statistical properties of the two matrices $\widehat{\mathbf{W}}$ (7) and $\widetilde{\mathbf{R}}$ (8) involved in the previous section. These results are used next to interpret the solution (9).

First, it can be checked that $LN_b\,\widehat{\mathbf{W}}$ has a complex Wishart distribution with $L\,(N_b-q)$ degrees of freedom and parameter matrix $\mathbf{Q}$. So,

$$E[\widehat{\mathbf{W}}] = \frac{N_b-q}{N_b}\mathbf{Q} \tag{14}$$

which shows that $\widehat{\mathbf{W}}$ is an estimate of the noise covariance matrix (up to a scale factor).

Next, using basic properties of Wishart matrices [4], it can be shown that the mean of $\widetilde{\mathbf{R}}$ (8) is given by:

$$E[\widetilde{\mathbf{R}}] = \alpha\,\mathbf{Q}^{-1/2}\mathbf{A}\,\Sigma\,\mathbf{A}^*\mathbf{Q}^{-*/2} + \beta\,\mathbf{I} \tag{15}$$

where $\alpha$ and $\beta$ are given by:

$$\begin{aligned}\alpha &= \frac{L\,N_b}{L\,(N_b-q)-m}\\ \beta &= \frac{L\,q}{L\,(N_b-q)-m}\\ \Sigma &= \frac{1}{L\,N_b}\sum_{l=1}^{L}\sum_{t=1}^{N_b}\mathbf{B}_l\,\mathbf{x}_l(t)\,\mathbf{x}_l(t)^*\mathbf{B}_l^*.\end{aligned}$$

Thus, from (15), the signal subspace of $E[\widetilde{\mathbf{R}}]$ is spanned by the columns of the whitened steering vectors matrix $\mathbf{Q}^{-1/2}\mathbf{A}$. Therefore, the projector $\widehat{\mathbf{\Pi}}_s$ onto the $n$ largest eigenvalues of $\widetilde{\mathbf{R}}$ is an estimate of the projector onto the columns of $\mathbf{Q}^{-1/2}\mathbf{A}$.

Turning our attention to the channel impulse response estimate $\widehat{\mathbf{H}}_l$, we observe that the computation of $\widehat{\mathbf{H}}_l$ in (9) can be interpreted as follows:

- whiten the MMSE estimate (12) (multiplication by $\widehat{\mathbf{W}}^{-1/2}$ in (9)) ;
- project it onto the estimated signal subspace (multiplication by $\widehat{\mathbf{\Pi}}_s$);
- cancel the whitening (multiplication by $\widehat{\mathbf{W}}^{1/2}$).

## 5. APPLICATION TO MLSE

Maximum Likelihood Sequence Estimation (MLSE) requires both the channel impulse response and the noise covariance: our maximum likelihood estimates of channel and noise can be used for that purpose. In most mobile communication systems, a training sequence is available: we assume that this is the case in this paragraph. For simplicity and coherence with preceding sections, the *last bust* of received data is numbered $L$, and the $L-1$ previous ones are numbered from 1 to $L-1$. Using the training sequences of the $L-1$ previous bursts, and the one of the present burst, one can compute the MLE of the current burst impulse response $\widehat{\mathbf{H}}_L$ (9) and the noise covariance matrix $\widehat{\mathbf{Q}}$ (10).

In the sequel, primes will be used to distinguish information bearing data from training sequence data. Let $x'_L(1)$, $\cdots$, $x'_L(N'_b)$ be the $N'_b$ unknown symbols of the last burst, and set $\mathbf{x}'_L(t) = [x'_L(t)\,\cdots\,x'_L(t-q+1)]^T$. Denote by $\mathbf{y}'_L(t)$ the array output data. Then, using our MLE $\widehat{\mathbf{H}}_L$ and $\widehat{\mathbf{Q}}$, the MLSE is obtained by minimizing the following function with respect to the unknown symbols:

$$f_{\widehat{\mathbf{Q}}}(x'_L(1)\,\cdots\,x'_L(N'_b)) = \sum_{t=1}^{N'_b}(\mathbf{y}'_L(t)-\widehat{\mathbf{H}}_L\,\mathbf{x}'_L(t))^*\,\widehat{\mathbf{Q}}^{-1}\,(\mathbf{y}'_L(t)-\widehat{\mathbf{H}}_L\,\mathbf{x}'_L(t))\ . \tag{16}$$

An equivalent criterion $f_{\widehat{\mathbf{W}}}(x'_L(1)\,\cdots\,x'_L(N'_b))$ is obtained if one replaces in expression (16) above $\widehat{\mathbf{Q}}$ (10) by $\widehat{\mathbf{W}}$ (7). Indeed, the inverse of $\widehat{\mathbf{Q}}$ (10) can be written:

$$\widehat{\mathbf{Q}}^{-1} = \widehat{\mathbf{W}}^{-1} - \widehat{\mathbf{W}}^{-*/2}\,\widehat{\mathbf{\Pi}}_n\,(\mathbf{W}+\widetilde{\mathbf{R}}^{-1})^{-1}\,\widehat{\mathbf{\Pi}}_n\,\widehat{\mathbf{W}}^{-1/2}\ ,$$

which yields , noting that $\widehat{\mathbf{\Pi}}_n\,\widehat{\mathbf{W}}^{-1/2}\,\widehat{\mathbf{H}}_L = 0$ :

$$f_{\widehat{\mathbf{Q}}}(x'_L(1)\,\cdots\,x'_L(N'_b)) = f_{\widehat{\mathbf{W}}}(x'_L(1)\,\cdots\,x'_L(N'_b)) + \sum_{t=1}^{N'_b}\mathbf{y}'_L(t)^*\,\widehat{\mathbf{W}}^{-*/2}\,\widehat{\mathbf{\Pi}}_n\,(\mathbf{W}+\widetilde{\mathbf{R}}^{-1})^{-1}\,\widehat{\mathbf{\Pi}}_n\,\widehat{\mathbf{W}}^{-1/2}\,\mathbf{y}'_L(t)\ .$$

Since in the above equation the second term of the right hand side does not depend on the emitted symbols, minimizing $f_{\widehat{\mathbf{W}}}$ and $f_{\widehat{\mathbf{Q}}}$ are clearly equivalent.

The next simulation illustrates this approach. The array is circular, with radius $R = 0.7\,\lambda$ where $\lambda$ is the wave-length, and it consists of $m = 10$ equi-spaced sensors. The channel impulse response has length $q = 5$. There are $n = 2$ signal paths in (2) with DOA's $0°$ and $30°$. The signal to noise ratio (excluding jammers) is $E_b/N_0 = -2$ dB. Three jammers are present with DOA's $-50°$, $-30°$ and $60°$ and equal power: the jammer to signal ratio is 7 dB. Each burst consists of $N'_b = 150$ information symbols $\{$+1,-1$\}$ and a training sequence of $N_b = 26$ symbols. The maximum likelihood estimates are obtained by processing the training sequences of $L = 100$ bursts. Three demodulation schemes are compared: the MLSE using exact channel $\mathbf{H}_L$ and noise covariance $\mathbf{Q}$, the MLSE using the unconstrained MLE channel estimate (12) and noise covariance $\widehat{\mathbf{W}}$ (7), and the MLSE using the rank constrained MLE channel estimate (9) and noise covariance $\widehat{\mathbf{W}}$ (7). Bit Error Rates (BER) are displayed in the next table. The BER using the rank constrained MLE is *three times* smaller than the BER using the unconstrained MLE.

| | known channel | unconstrained MLE channel | rank constrained MLE channel |
|---|---|---|---|
| BER | 1.28 % | 7.44 % | 2.65 % |

## 6. APPLICATION TO PATHS DOA ESTIMATION

In some applications for mobile communication systems, the paths DOA have to be determined. This is the case for instance when one tries to design a beamformer for down-link communication based on up-link data in Frequency-Division Duplex systems such as GSM [3]. The results of section 3 can be used to that purpose as discussed now.

We have shown in section 3 that

$$\widehat{\mathbf{H}}_l = \widehat{\mathbf{W}}^{1/2}\,\widehat{\mathbf{\Pi}}_s\,\widehat{\mathbf{W}}^{-1/2}\,\widehat{\mathbf{R}}_{yx_l}\,\widehat{\mathbf{R}}_{xx_l}^{-1} \qquad (17)$$

is the maximum likelihood estimate of $\mathbf{H}_l = \mathbf{A}\,\mathbf{B}_l$. By writing

$$\mathbf{H}_l = \widehat{\mathbf{W}}^{1/2}\,\widehat{\mathbf{W}}^{-/2}\,\mathbf{A}\,\mathbf{B}_l \qquad (18)$$

and comparing equations (17) and (18), we deduce that projector $\widehat{\mathbf{\Pi}}_s$ is an estimate of the projector onto the columns of $\widehat{\mathbf{W}}^{-/2}\,\mathbf{A}$. Consequently, the steering vector $\mathbf{a}(\theta)$ of a path with DOA $\theta$ satisfies $\widehat{\mathbf{\Pi}}_n\,\widehat{\mathbf{W}}^{-1/2}\mathbf{a}(\theta) \approx 0$ where $\widehat{\mathbf{\Pi}}_n = \mathbf{I} - \widehat{\mathbf{\Pi}}_s$ . Thus, following the MUSIC scheme, the paths DOA can be estimated by looking for the maxima of the function $h(\theta)$ defined by :

$$h(\theta) = \frac{\mathbf{a}(\theta)^*\,\widehat{\mathbf{W}}^{-1}\,\mathbf{a}(\theta)}{\mathbf{a}(\theta)^*\,\widehat{\mathbf{W}}^{-*/2}\,\widehat{\mathbf{\Pi}}_n\,\widehat{\mathbf{W}}^{-1/2}\mathbf{a}(\theta)}\,. \qquad (19)$$

The following simulation illustrates this approach. The array is circular, with radius $R = 0.7\,\lambda$ where $\lambda$ is the wavelength, and consists of $m = 10$ equi-spaced sensors: the 3-dB beamwidth is $2\,\theta_3 = 30°$.

The channel impulse response has length $q = 5$. The training sequence consists of 26 symbols (+1 or -1), and $N_b = 26 - q + 1 = 22$ data are processed at each burst. There are $n = 2$ signal paths in (2) with DOA's $\theta_1 = 10°$ and $\theta_2 = 20°$. The number of processed burst is $L = 100$ which corresponds to about 0.5s for GSM. Matrices $\mathbf{B}_l$, $l = 1, \cdots, L$ in (2) are independent from one burst to another, and their elements are complex gaussian distributed with zero mean and variance one. Their are three gaussian jammers with DOA's $30°$, $340°$ and $350°$ and power 2.5 each. Uncoherent backgound noise power is also 2.5. Note that there are three sources ($10°$, $20°$ and $30°$) within a beamwidth and one of them is a jammer ($30°$). Figure 1 displays the paths DOA's estimates obtained from 50 independent trials: the 2 paths are detected, and jammers are rejected.

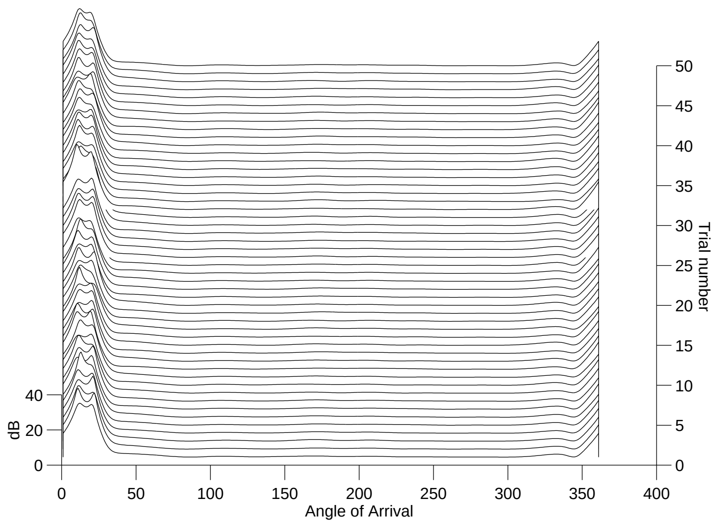


Figure 1: Paths DOA's estimates.

## 7. CONCLUSION

We have developped a maximum likelihood estimator of the multichannel impulse response under a reduced rank constraint. This estimator finds applications in equalization and channel analysis (paths DOA estimation). When used in a Viterbi equalizer, the new channel estimate yields better performances compared to the MMSE channel estimate.

## 8. REFERENCES


[1] M. Viberg, P. Stoica, and B. Ottersten. Maximum Likelihood Array Processing in Spatially Correlated Noise Fields Using Parameterized Signals. *IEEE Trans. Signal Processing*, 45(4):996–1004, April 1997.

[2] J.P. Rossi, J.P. Barbot, and A.J. Levy. Theory and Measurement of the Angle of Arrival and Time Delay of UHF Radiowaves Using a Ring Array. *IEEE Trans. Signal Processing*, 45(5):876–884, May 1997.

[3] P. Zetterberg and B. Ottersten. The Spectrum Efficiency of a Base-Station Antenna Array System for Spatially Selective Transmission. *IEEE Trans. on Vehicular Technology*, 44(3):651–660, 1995.

[4] D. Maiwald and D. Kraus. On Moments of Complex Wishart and Complex Inverse Wishart Distributed Matrices. in *Proc. ICASSP 97*, Munich, Germany, 1997, pp. 3817–3820.