Performance of Semi-Blind Channel Estimation Approximations in LTE Downlink

Mrs. Poornima*, Prof. G. Laxminarayana** and Prof. D. Srinivas Rao***
*Associate Professor, Dept. of ECE, Sphoorthy Engineering College, Hyderabad
padaraju_poorni@yahoo.com
** Professor, Dept. of ECE, Anurag College of Engineering, Hyderabad
gln1955@gmail.com
*** Professor, Dept. of ECE, JNTUH CEH, Kukatpally, Hyderabad
dsraoece@gmail.com

Abstract: In this paper, we analyzed the Semi-Blind approach channel estimation in MIMO-OFDM systems, and in particular for LTE downlink. Proposed Semi-Blind approaches lead to significant improvements in the estimation accuracy, both from an MSE and BER perspective, compared to the typical pilot based technique. However, exploiting the true discrete distribution of the unknown symbols is computationally demanding, therefore we propose the use of two approximations on the unknown symbols: the Gaussian and the Constant Modulus assumption to significantly improve the accuracy with respect to the pilot based approach, while reducing the computational overhead incurred when using true discrete distribution of the unknown symbols.

Keywords: Long Term Evolution; MIMO-OFDM; Semi-Blind Channel Estimation; Expectation Maximization.

Introduction

Long Term Evolution (LTE) is a new communication technology based on Orthogonal Frequency Division Multiple Access (OFDMA) in the downlink (DL) and Single Carrier Frequency Division Multiple Access (SCFDMA) in the uplink (UL). Additionally, LTE downlink transmission model is based on multiple antenna architecture on the transmitter and receiver side [1]. Orthogonal Frequency Division Multiplexing (OFDM) has been widely applied in wireless communication systems due to its high data rate transmission and its robustness to multipath channel delay [2], [3].

Although combination of MIMO with OFDM presents a solution to increase the capacity and the reliability of wireless channels, due to time varying channel characteristics of the radio signals it is particularly challenging from a channel estimation perspective.

Channel estimation plays a crucial role in the performance of wireless communication systems, since its knowledge is utilized to detect the data symbols [1]. With enhancing the channel estimation performance, the performance of the entire system can also be improved. The channel estimation can be classified into three categories: training-based methods, blind methods and semi-blind methods. For pure training-based schemes, a long training is necessary in order to obtain a reliable MIMO-OFDM channel estimate which reduces the system bandwidth efficiency considerably. Blind methods which do not require any training symbols achieve high system throughput at the expense of high computational complexity. The blind channel identification methods can be classified into higher-order statistics based techniques [6]-[8] and second order statistics based techniques [9], [10].

Joint blind channel estimation and data estimation detection has been proposed based on the iterative least squares with projection [11]-[13]. This scheme estimated the channel and data iteratively but the convergence of the scheme depends on the initialization of the channel model. Semi-blind schemes on the other hand require less computational complexity than blind methods and fewer training symbols than training-based methods, to provide the initial MIMO-OFDM channel estimation and exchange the information between the channel estimator and the data detector iteratively so as to provide Channel state information [CSI] to the Transmitter, so that it can adjust antenna weights, transmitting power, modulation and coding instantaneously.

This paper is organized as: In section II, we introduce ML semi-blind channel estimation of MIMO-OFDM FIR channels from a general perspective, without any prior assumption on the distribution of the transmitted signal. A framework on Expectation-Maximization algorithm to solve the maximization problem is presented in Section-III.

System Model

Let's consider a MIMO-OFDM system with N sub-carriers, T transmitting and R receiving antennas ($T\times R$ MIMO). Let $X_n(k)$ be the MIMO signal transmitted on sub-carrier n at time k (this is a $T \times 1$ vector). With OFDM, the time domain signal is obtained with the Inverse DFT transformation, through the relation

$$x^{(k)}(p) = \frac{1}{\sqrt{N}} \sum_n X_n(k) e^{i2\pi \frac{pn}{N}} \quad , \quad where \quad p = -CP...N-1 \tag{1}$$

Here $x^{(k)}(p)$ is the pth sample of the kth MIMO-OFDM symbol, where this latter term refers to the ordered set of the symbols transmitted on all the sub-carriers, that is $\{X_n(k), n=0...N-1\}$. These samples are then transmitted in sequence through the channel across the antennas array. Since the channel is FIR of length L, the output of the model at time k depends only on the transmitted symbols at times k-L+1...k. The symbol transmitted on the kth symbol is given as:

$$\begin{aligned} y^{(k)}(p) &= \sum_{l=0}^{L-1} h_l x^{(p)}(p-l) + \tilde{\eta}^{(k)}(p) \\ &= \frac{1}{\sqrt{N}} \sum_n \sum_{l=0}^{L-1} h_l e^{i2\pi \frac{(p-l)n}{N}} X_n(k) + \tilde{\eta}^{(k)}(p), \qquad where \quad p = 0...N-1 \end{aligned} \tag{2}$$

If H_n be the frequency domain channel, defined as $\sqrt{N}$ times the DFT of the time domain channel h_l, we obtain

$$y^{(k)}(p) = \frac{1}{\sqrt{N}} \sum_n H_n X_n(k) e^{i2\pi \frac{pn}{N}} + \tilde{\eta}^{(k)}(p), \qquad where \quad p = 0...N-1 \tag{3}$$

At the receiver, the time-domain signal is processed using the N-points DFT. On subcarrier m we have

$$\begin{aligned} Y_m(k) &= \frac{1}{\sqrt{N}} \sum_{p=0}^{N-1} y^{(k)}(p) e^{-i2\pi \frac{pm}{N}} \\ &= \frac{1}{N} \sum_n H_n X_n(k) \sum_{p=0}^{N-1} e^{i2\pi \frac{p(n-m)}{N}} + \frac{1}{N} \sum_{p=0}^{N-1} \tilde{\eta}^{(k)}(p) e^{i2\pi \frac{-pm}{N}} \\ &= \sum_n H_n X_n(k) \delta_{nm} + \eta_m(k) \;\; = H_m X_m(k) + \eta_m(k) \end{aligned} \tag{4}$$

where η_m is the noise vector on sub-carrier m at time k.

Maximum Likelihood Channel Estimation In MIMO-OFDM FIR Channel

As the order of the MIMO system increases to achieve an acceptable estimation accuracy more pilots have to be collected at the receiver. This in turn is achieved either enlarging the observation time, or allocating more pilots on the OFDM grid. However, the first approach (larger observation time) compromises the ability of the receiver to track fast-varying channels, which is not acceptable in a Mobile Communication System. On the other hand, the second approach (more pilots on the OFDM grid) compromises the bandwidth efficiency of the system, since the pilots represent a waste of bandwidth. Therefore, it becomes important to exploit other information at the receiver than relying solely on pilots. Semi-blind channel estimators mitigates the above two problems by exploiting the known symbols for the estimate but also blind information in order to enhance the estimation accuracy.

Let's assume K OFDM symbols are transmitted. The input-output relation of this system is given by

$$Y_n = H_n X_n + \eta_n \qquad \forall \quad n = 0...N-1 \tag{5}$$

where Y_n is the R×K observation matrix, H_n is the channel matrix, X_n is the T × K matrix of the transmitted symbols, and η_n is the noise matrix at the receiver on subcarrier n. We don'take any assumption on the distribution of the transmitted symbols, X_n may carry either pilots, unknown symbols, or both.

The Maximum Likelihood solution is obtained by maximizing the likelihood, or equivalently minimizing the negative log-likelihood of the observations with respect to the parameters of the model. As the channel is FIR of length L, in order to enforce the functional constraint of the frequency domain channel taps, the ML solution is determined with respect to the

channel coefficients in the time domain, that is $\left\{h_l\left(r,t\right),\ \ \forall\, l,r,t\right\}$.Then, stacking the time domain channel entries on the column vector *h*, with entries $h\left(RTl+Tr+t\right)=h_l\left(r,t\right)$, the likelihood of the observations conditioned on *h* is given by:

$$p\left(Y\middle|h\right)=E_X\left[p\left(Y\middle|X,h\right)\right] \tag{6}$$

where the notation $E_\alpha\left[f\left(\alpha,\beta\right)\right]$, represents the expectation of $f\left(\alpha,\beta\right)$ with respect to the prior distribution of α, whereas the notation $E_\alpha\left[f\left(\alpha,\beta\right)|\beta\right]$ represents the expectation of $f(\alpha,\beta)$ with respect to the distribution of α conditioned on β (this expectation is a function of β).Under regularity conditions (differentiability of the likelihood function with respect to its argument *h*), a necessary condition for the ML solution is that it is solution to the likelihood equation, which is obtained by calculating the gradient of $-\ln p\left(Y\middle|h\right)$ with respect to the parameter vector *h* (the gradient operator is indicated with the notation Δ_h), and equaling it to zero. We obtain

$$-\Delta_h \ln p\left(Y\middle|h\right)=-\frac{1}{p(Y|h)}\Delta_h p\left(Y\middle|h\right)=-\frac{1}{p(Y|h)}\Delta_h E_X\left[p(Y|X,h)\right] \tag{7}$$

Where, we used the fact that $\dfrac{\partial \ln f\left(\alpha\right)}{\partial\left(\alpha\right)}=\dfrac{1}{f\left(\alpha\right)}\dfrac{\partial f\left(\alpha\right)}{\partial\alpha}$.

Then, since the prior distribution of the transmitted symbols does not depend on the channel entries, we can move the gradient within the expectation term, obtaining

$$-\Delta_h \ln p\left(Y\middle|h\right)=-\frac{1}{p\left(Y\middle|h\right)}E_X\left[\Delta_h p\left(Y\middle|X,h\right)\right]=-\frac{1}{p\left(Y\middle|h\right)}E_X\left[p\left(Y\middle|X,h\right)\Delta_h \ln p\left(Y\middle|X,h\right)\right] \tag{8}$$

Finally, using the fact that $E_X\left[\dfrac{p\left(Y\middle|X,h\right)}{p\left(Y\middle|h\right)}f\left(X\right)\right]=E_X\left[f\left(X\right)\middle|Y,h\right]$, and equaling the gradient to zero, the likelihood can be given as

$$-\Delta_h \ln p\left(Y\middle|h\right)=-E_X\left[\Delta_h \ln p(Y|X,h)|Y,h\right]=0 \tag{9}$$

Now, let's assume that the noise at the receiver is a zero mean Gaussian process, independent across the sub-carriers and across time, with covariance matrix Cov(η_n) (or equivalently precision matrix $B_{\eta_n}=Cov\left(\eta_n\right)^{-1}$ on sub-carrier *n*. Under this assumption, when conditioned on the transmitted symbols and on the channel, the observations are independent across sub-carriers and across time, therefore we can split the probability density function (PDF) $p(Y|X,H)$ into the product of the PDFs of the observations on each sub-carrier, and equivalently we can express $\ln p(Y|H,X)$ as the sum of those densities. Then, making explicit the probability density function on each sub-carrier we obtain

$$-\ln p\left(Y\middle|X,h\right)=-\sum_{n=0}^{N-1}K_n \ln\left(\frac{\left|B_{\eta_n}\right|}{\pi^R}\right)+\sum_{n=0}^{N-1}trace\left[B_{\eta_n}\left(Y_n-H_nX_n\right)\left(Y_n-H_nX_n\right)^H\right] \tag{10}$$

Where, K_n is the number of observations used for the estimate on sub-carrier *n*, and H_n is the frequency domain channel tap on sub-carrier *n*, whose entries are linear functions of the parameter vector *h* through the DFT transform. The derivative of this term with respect to the time-domain channel matrix entry $h_l(r,t)$ is given by

$$-\frac{\partial \ln p\left(Y\middle|X,h\right)}{\partial h_l\left(r,t\right)}=\sum_{n=0}^{N-1}trace\left[B_{\eta_n}\partial\left(r,t\right)X_n\left(Y_n-H_nX_n\right)^H\right]e^{-i2\pi\frac{\ln}{N}}=\sum_{n=0}^{N-1}\left[X_n\left(Y_n-H_nX_n\right)^H B_{\eta_n}\right]_{tr}e^{-i2\pi\frac{\ln}{N}} \tag{11}$$

Finally, equaling the derivative to zero, we obtain the entries of the likelihood equation (9).

$$-\frac{\partial \ln p\left(Y\middle|h\right)}{\partial h_l\left(r,t\right)}=-\sum_{n=0}^{N-1}E_X\left[X_n\left(Y_n-H_nX_n\right)^H B_{\eta_n}|Y,h\right]_{tr}e^{-i2\pi\frac{\ln}{N}} \tag{12}$$

Since the above equation has to be satisfied for all the transmitting-receiving antennas pairs (*r*,*t*) and for all channel taps *l*, we can rewrite it in matrix form with respect to the indexes *t* and *r*, obtaining the following set of equations

$$-\frac{\partial \ln p\left(Y\middle|h\right)}{\partial h_l^*}=-\sum_{n=0}^{N-1}B_{\eta_n}E_{X_n}\left[Y_nX_n^H-H_nX_nX_n^H|Y,h\right]e^{i2\pi\frac{\ln}{N}}=0,\qquad \forall\quad l=0\ldots L-1 \tag{13}$$

The solutions to equation (13) are stationary points of the negative log-likelihood function, however they are not guaranteed to be absolute minima of the function. Furthermore, observe that the solution depends on the posterior expectation and the posterior correlation of the transmitted symbols after observing *Y*. In the case of training sequence estimation, X_n is the matrix containing solely the pilot symbols, which is a deterministic quantity independent of the channel realization and of the observations, therefore for this case the above equation reduces to:

$$\sum_{n=0}^{N-1} B_{\eta_n} H_n X_n X_n^H e^{i2\pi \frac{\ln}{N}} = \sum_{n=0}^{N-1} B_{\eta_n} Y_n X_n^H e^{i2\pi \frac{\ln}{N}} \tag{14}$$

When both pilot symbols and blind information are used for the estimation, we can split equation (13) into the sum of the contribution coming from the pilot symbols and the contribution from the blind information, that is, using the superscripts (*tr*) and (*bl*) to distinguish pilot from blind observations, symbols and noise, we can rewrite equation (13) as:

$$-\frac{\partial \ln p\left(Y \middle| h\right)}{\partial h_l^*} = -\sum_{n=0}^{N-1} B_{\eta_n} \left(Y_n^{(tr)} - H_n X_n^{(tr)}\right) X_n^{(tr)H} e^{i2\pi \frac{\ln}{N}} - \sum_{n=0}^{N-1} B_{\eta_n} E_{V_n}(bl)\left[\left(Y_n^{(bl)} - H_n C V_n^{(bl)}\right) V_n^{(bl)H} \middle| Y_n^{(bl)}, h\right] C^H e^{i2\pi \frac{\ln}{N}} \tag{15}$$

$$= 0. \qquad \forall \quad l = 0 \ldots L-1$$

Since this equation involves the calculation of the posterior expectation of the transmitted symbols and their correlation conditioned on the observations *Y*, the solution to this equation then depends on the assumptions we use on the prior distribution of the unknown symbols. From the point of view of the estimation accuracy, the optimal solution consists in using the true discrete distribution of the symbols. However this solution is computationally very demanding, since it requires the computation of the posterior probabilities for any possible combination of transmitted symbols. Moreover, it is not scalable to MIMO systems since the number of symbol combinations grows exponentially with the transmission rank.

Time Discrete Distribution of Unknown Symbols in ML Estimation

With reference to the system model and the set of assumptions the unknown symbols are drawn uniformly from a finite discrete constellation $\square^{S\times 1}$, where *S* is the transmission rank, independently across the sub-carriers and across time.

Therefore, for all the unknown symbols $V_n^{(bl)}(k)$, we have

$$p\left(V_n^{(bl)}(k)\right) = \frac{1}{\left|\square\right|^S} \qquad \forall \quad V_n^{(bl)}(k) \in \square^{S\times 1} \tag{16}$$

Now, based on a set of observations, collected on the observation matrix *Y*, and a set of pilot symbols $X^{(tr)}$, the goal is to determine the ML estimate of the channel, which is solution to the likelihood equation, given by

$$\Delta_h \ln p\left(Y \middle| X^{(tr)}, h\right) = 0 \tag{17}$$

where the gradient is calculated with respect to the time-domain channel matrix *h*, in order to enforce the channel length constraint.

EM-algorithm for the determination update of the channel estimate solution depends only on the posterior first and second order statistics of the unknown symbols. To compute the posterior first and second order statistics when current channel estimate *h*(*j*) is given as:

$$\begin{cases} \Lambda_{vv}^{(n,j)} = E_{V_n^{(bl)}}\left[V_n^{(bl)} V_n^{(bl)H} \middle| Y_n^{(bl)}, h^{(j)}\right] \\ \tilde{V}_n^{(j)} = E_{V_n^{(bl)}}^{\left(q^{(i)}\right)}\left[V_n^{(bl)} \middle| Y_n^{(bl)}, h^{(j)}\right] \end{cases}, \quad \forall \quad n = 0 \ldots N-1 \tag{18}$$

The unknown symbol on sub-carrier *n* at time *k*, using Bayes' rule we have

$$p\left(V_n^{(bl)}(k) \middle| Y_n^{(bl)}(k), h^{(j)}\right) = \rho\, p\left(Y_n^{(bl)}(k) \middle| V_n^{(bl)}(k), h^{(j)}\right) p\left(V_n^{(bl)}(k)\right) \tag{19}$$

Where, ρ is the normalization factor, independent of the value of the unknown symbol.

From the posterior distribution on the unknown symbols $q_{nk}^{(j)}(\beta)$, we can calculate the two matrices $\tilde{V}_n^{(j)}$ and $\Lambda_{vv}^{(n,j)}$ defined as:

$$\begin{cases} \tilde{V}_n^{(j)}(k) = \sum_{\beta \in \square^{S\times 1}} \beta . q_{nk}^{(j)}(\beta) \\ \Lambda_{VV}^{(n,j)} = \sum_{\beta \in \square^{S\times 1}} \beta\beta^H \Sigma_k q_{nk}^{(j)}(\beta) \end{cases} \tag{20}$$

It is observed that, the complexity of this algorithm depends on the computations required to calculate the posterior distribution for each point of the constellation $\square^{S\times 1}$. Eg. if M be the constellation order, M^S posterior probabilities have to be calculated for each unknown symbols. Hence, this is not scalable to higher order MIMO systems, since the number of posterior probabilities which need to be computed grows exponentially with the transmission rank.

Gaussian Approximation for the Unknown Symbols in ML Estimation

Assuming that the distribution of the unknown symbols is circular Gaussian, implies that the distribution of the observations conditioned on the channel matrix is a multivariate Gaussian, therefore the marginalization over the discrete distribution of the unknown symbols leads to a mixture of Gaussians. However, we can approximate this distribution with a single multivariate Gaussian.

If the best multivariate Gaussian is given as $q(X)$ which can be used as an approximation of the true distribution $p(X)$. Then, measure of closeness of a distribution to another is given by Kullback–Leibler divergence, which for continuous distributions is defined as:

$$KL(p\|q) = \int_D p(X) \ln\left(\frac{p(X)}{q(X)}\right) dX \tag{21}$$

Where, $p(X)$ is the true PDF and $q(X)$ is the PDF we want to use to approximate $p(X)$. To approximate $p(X)$ with a multivariate Gaussian $q(X)$ with mean m and covariance matrix Σ. Then, the best m and Σ are obtained by minimizing the Kullback–Leibler divergence with respect to m and Σ. It can be easily shown, by calculating the derivative and equaling it to zero, that the solution is given by

$$\begin{cases} m = E[X] \\ \Sigma = E\left[(X-m)(X-m)^H\right] \end{cases}. \tag{22}$$

Where, the expectation is taken with respect to the true distribution $p(X)$. To approximate the distribution of the observations corresponding to the unknown symbols with a multivariate Gaussian $q(Y)$ with mean m_Y and covariance matrix Σ_Y. On sub-carrier n at time k the mean value of the observations is given by

$$mY_n = E\left[Y_n(k)\right] = E\left[H_n X_n(k) + W_n(k)\right] = 0 \tag{23}$$

Where, we used the fact that the noise and the unknown symbols are zero mean.
Similarly, for the covariance matrix we obtain

$$\Sigma Y_n = E\left[Y_n(k) Y_n(k)^H\right] = H_n E\left[X_n(k) X_n(k)^H\right] H_n^H + Cov(\eta_n) \tag{24}$$

The higher is the noise variance at the receiver with respect to the power of the symbols, the larger is the lobe of each multivariate Gaussian, the more overlap there is between pairs of multivariate Gaussians. Therefore, we expect this approximation to perform well especially in the low-SNR regime and higher constellation order. Since the posterior expectation of the unknown symbols is a function of the channel matrix, Expectation-Maximization algorithm is used to determine a local maximum to the likelihood function.

For the given channel estimation $h^{(j)}$, the posterior distribution and second order statistics of the unknown symbols is calculated as:
Using Bayes' rule, the posterior distribution of the unknown symbol on sub-carrier n at time k is given by

$$p\left(V_n^{(bl)}(k) \middle| Y_n^{(bl)}(k), h\right) = \mu p\left(Y_n^{(bl)}(k) \middle| V_n^{(bl)}(k), h\right) p\left(V_n^{(bl)}(k)\right) \tag{25}$$

Where, μ is the normalization factor, which does not depend on the unknown symbols.

Now, $p\left(Y_n^{(bl)}(k)\middle|V_n^{(bl)}(k),h\right)$ is a Gaussian PDF with mean $H_n C V_n^{(bl)}(k)$ and covariance $Cov(\eta_n)$ (precision B_{η_n}), and the unknown symbols $V_n^{(bl)}(k)$ are Gaussian distributed with zero mean and covariance $\sigma^2{}_S I_S$. Keeping only the terms depending on the symbol $V_n^{(bl)}(k)$ and including others in normalization factor μ we have

$$p\left(V_n^{(bl)}(k)\middle|Y_n^{(bl)}(k),h\right)=\mu\exp\left\{-V_n^{(bl)}(k)^H\left(C^H H_n^H \beta_{\eta_n} H_n C+\frac{1}{\sigma_S^2}I_S\right)V_n^{(bl)}(k)\right\}.\exp\left\{2real\left(V_n^{(bl)}(k)^H C^H H_n^H B_{\eta_n} Y_n^{(bl)}(k)\right)\right\} \tag{26}$$

However, when conditioned on $Y_n^{(bl)}(k)$ and h, $V_n^{(bl)}(k)$ is Gaussian distributed with mean $mV_n(k)$ and covariance matrix $\Sigma V_n(k)$). Therefore we have also:

$$p\left(V_n^{(bl)}(k)\middle|Y_n^{(bl)}(k),h\right)=\lambda\exp\left\{-V_n^{(bl)}(k)^H \Sigma V_n(k)^{-1}\Sigma V_n(k)^{-1}V_n^{(bl)}(k)\right\}.\exp\left\{2real\left(V_n^{(bl)}(k)^H \Sigma V_n(k)^{-1} mV_m(k)\right)\right\} \tag{27}$$

where λ is the normalization factor.

Comparing the above expression with Eq.(24) we have the following two equalities for the posterior covariance matrix $\Sigma_{V_n}(k)$ and for the posterior mean $m_{V_n}(k)$ of the unknown symbols at time k on sub-carrier n, given the current update of the channel matrix $h^{(j)}$:

$$\begin{cases} \Sigma_{V_n}^{(j)} = \left(C^H H_n^{(j)H} B_{\eta_n} H_n^{(j)} C+\frac{1}{\sigma_n^2}I_S\right)^{-1} \\ mV_n(k)^{(j)} = \Sigma_{V_n} C^H H_n^{(j)} B_{\eta_n} Y_n^{(bl)}(k) \end{cases} \tag{28}$$

Where, for the covariance term we dropped the time index k since it is independent from it.
Then, stacking the posterior mean of the unknown symbols on a matrix using the time k as column index, and we have

$$m_{V_n}^{(j)} = \Sigma_{V_n}^{(j)} C^H H_n^{(j)H} B_{\eta_n} Y_n^{(bl)} \tag{29}$$

From the posterior mean and covariance we can calculate the posterior first and second order moments of the unknown symbols as

$$\begin{cases} \tilde{V}_n^{(j)} = m_{V_n}^{(j)} \\ \Lambda_{nn}^{(n,j)} = m_{V_n}^{(j)} m_{V_n}^{(j)H} + K_n^{(bl)} \Sigma_{V_n}^{(j)} \end{cases} \tag{30}$$

These matrices are then used to update the channel matrix.

Constant Modulus Approximation for the Unknown Symbols

We propose a Semi-Blind MIMO-OFDM FIR channel estimation technique based on the assumption that the unknown symbols are drawn from a constant modulus alphabet. By constant modulus, it is meant a modulation technique with the property that all the points in the constellation have the same amplitude. the Gaussian assumption means that we have two degrees of uncertainty on the transmitted symbols: amplitude and phase. Conversely, the points in a constant modulus constellation have only one degree of freedom, the phase, since the amplitude is fixed. While in the Gaussian assumption the phase of the symbols is uniformly distributed in the range [0,2π) and the amplitude is Rayleigh distributed. In the Constant Modulus, the phase of the symbols is assumed to be uniformly distributed in the range [0,2π). Therefore, given the less degree of freedom on the unknown symbols, we expect to achieve a more accurate estimate than the Gaussian assumption.
In this work, we propose an alternative algorithm, based on a Taylor series expansion of the posterior probabilities of the unknown symbols, for the limit case of the constellation order M going to infinity. This algorithm performs well also with a short sequence of blind observations, as we will show in the simulation results. Assuming rank-one transmission (S = 1), and

assuming that the unknown symbols $V_n(k)$ are drawn from a constant modulus alphabet, the term $V_n(k)V_n(k)^H$ is deterministically equal to the symbol power σ_s^2, independently of the observations and of the channel realization. Therefore:

$$E_{V_n^{(bl)}}\left[V_n^{(bl)}V_n^{(bl)H}\,\middle|\,Y_n^{(bl)},h\right]=K_n^{(bl)}\sigma_s^2 \tag{31}$$

Assuming for now $S > 1$, and letting $V_{nk} \in \square^{S\times 1}$ be the unknown symbol transmitted on sub-carrier n at time k, and Y_{nk} the corresponding observation, the posterior mean of the unknown symbol is given by

$$E_{nk}\left[V_{nk}\,\middle|\,Y_{nk},h\right]=\sum_{\alpha\in\square^{S\times 1}}\alpha\, p\left(\alpha\,\middle|\,Y_{nk},h\right) \tag{32}$$

Now, using Bayes' rule, we can write the posterior distribution as

$$p\left(\alpha\,\middle|\,Y_{nk},h\right)=\mu\, p\left(Y_{nk}|\alpha,h\right)p\left(\alpha\right) \tag{33}$$

Where, μ is the normalization factor, independent from μ, and the prior distribution $p(\alpha)$ is a constant with respect to α, since the symbols are drawn uniformly from the alphabet, therefore $p(k)=\frac{1}{|\square|^S}$

Then, under the assumption that the noise is Gaussian with zero mean and precision matrix $B_{\eta_n}=Con(\eta_n)^{-1}$.

$$E_{V_{nk}}\left[V_{nk}\,\middle|\,Y_{nk},h\right]=\frac{\sum_{\alpha\in C^{S\times 1}}\alpha\exp\left\{-trace\left[B_{\eta_n}\left(Y_{nk}-H_nC\alpha\right)\left(Y_{nk}-H_nC\alpha\right)^H\right]\right\}}{\sum_{\alpha\in C^{S\times 1}}\exp\left\{-trace\left[B_{\eta_n}\left(Y_{nk}-H_nC\alpha\right)\left(Y_{nk}-H_nC\alpha\right)^H\right]\right\}} \tag{34}$$

For the exponential term in the above expression we have

$$\exp\left\{-trace\left[B_{\eta_n}\left(Y_{nk}-H_nC\alpha\right)\left(Y_{nk}-H_nC\alpha\right)^H\right]\right\}=\mu\exp\left\{-\alpha^HC^HH_n^HB_{\eta_n}H_nC\alpha\right\}\exp\left\{2real\left(Y_{nk}^HB_{\eta_n}H_nC\alpha\right)\right\} \tag{35}$$

Where, μ is a constant which does not depend on α.

Let $\Upsilon_{S_1S_2}=\left(C^HH_n^HB_{\eta_n}H_nC\right)_{s_1s_2}$ and $\xi_s=\left(Y_{nk}^HB_{\eta_n}H_nC\right)_s$, then we can rewrite the above exponential term as

$$\exp\left\{-trace\left[B_{\eta_n}\left(Y_{nk}-H_nC\alpha\right)\left(Y_{nk}-H_nC\alpha\right)^H\right]\right\}=\mu\exp\left\{-\sum_{s_1}\Upsilon_{s_1s_2}\left|\alpha_{s_1}\right|^2\right\}\exp\left\{-\sum_{s_1s_2\neq s_1}\left(\Upsilon_{s_1s_2}\alpha_{s_2}\alpha_{s_1}^*\right)\right\}\exp\left\{2real\sum_s\alpha_s\xi_s\right\} \tag{36}$$

In the case $S>1$ there is one more term in the expression for the posterior expectation, given by $\exp\left\{-\sum_{s_1s_2\neq s_1}\left(\Upsilon_{s_1s_2}\alpha_{s_2}\alpha_{s_1}^*\right)\right\}$, which keeps into account the correlation between the symbols across the transmission streams. We defined the scalar function:

$$g\left(x\right)=\frac{\sum_{n=0}^{+\infty}\frac{1}{n!\left(n+1\right)!}x^{2n+1}}{\sum_{n=0}^{+\infty}\frac{1}{\left(n!\right)^2}x^{2n}}\qquad\forall\ x\geq 0 \tag{37}$$

Since it is not possible to solve analytically the above sum, we seek for an approximation. Let $\tilde{g}_N\left(x\right)$ be the function obtained by taking the first N terms of the numerator and denominator

$$\tilde{g}_N(x) = \frac{\sum_{n=0}^{N} \frac{1}{n!(n+1)!}(x)^{2n+1}}{\sum_{n=0}^{N} \frac{1}{(n!)^2}(x)^{2n}} \qquad \forall \ x \geq 0 \tag{38}$$

for different values of N.

$$\lim_{N \to +\infty} \tilde{g}_N(x) = g(x) \tag{39}$$

It is observed that, the series of functions $\lim_{N \to +\infty} \tilde{g}_N(x) = g(x)$ approaches the black curve $\{\tilde{g}_N\}$ for growing values of N, which is equal to zero for $x = 0$ and converges to one for growing values of x. Therefore we expect $\{\tilde{g}_N\}$ to be close to the approximation $g(x)$.

With the statistical properties of the term $\sigma_s \rho_{nk}$ in the low and high-SNR ranges and white Gaussian noise at the receiver with variance σ_w^2 and considering the term $\sigma_s \rho_{nk}$ as a random variable, its mean and variance are given by:

$$\begin{cases} E\left[\sigma_s \rho_{nk}\right] = 0 \\ E\left[\sigma_s^2 \left|\rho_{nk}\right|^2\right] = \sigma_s^2 C^H H_n^H B_{\eta_n} E\left[Y_n^{(bl)}(k) Y_n^{(bl)}(k)^H\right] B_{\eta_n} H_n C \end{cases} \tag{40}$$

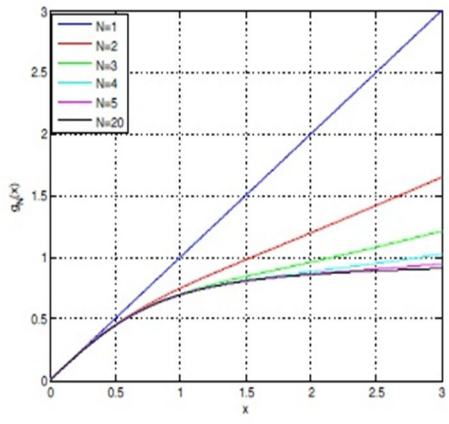


Fig 1. $\tilde{g}_N(x)$ for different values of N

In the low-SNR regime we have $\frac{\sigma_s^2}{\sigma_w^2} \ll 1$, therefore for the variance of $\sigma_s \rho_{nk}$ we have:

$$E\left[\sigma_s^2 \left|\rho_{nk}\right|^2\right] \ll \frac{\sigma_s^2}{\sigma_w^2} C^H H_n^H H_n C \ll 1 \tag{41}$$

which means that $\sigma_s \rho_{nk}$ is statistically small, and accordingly $g\left(\sigma_s \left|\rho_{nk}\right|\right)$, that is the amplitude of the posterior expectation, is small since in the low-SNR regime the observations carry mostly noise, and very few information about the transmitted symbols, therefore the posterior mean is close to the prior mean, which is zero.

In the high-SNR regime we have $\frac{\sigma_s^2}{\sigma_w^2} \gg 1$, therefore for the variance of $\sigma_s \rho_{nk}$ we have:

$$E\left[\sigma_s^2 \left|\rho_{nk}\right|^2\right] = \frac{\sigma_s^4}{\sigma_w^4}\left(C^H H_n^H C\right)^2 \gg 1 \tag{42}$$

which means that $\sigma_s \rho_{nk}$ is statistically large, and accordingly $g\left(\sigma_s \left|\rho_{nk}\right|\right)$ is close to 1. Similarly, this high-SNR regime behavior is the one expected, since the observations carry mostly information about the transmitted symbols, therefore the posterior mean is close to the true transmitted symbol, or equivalently it is close to the circle of amplitude σ_s.

Therefore, we can statistically associate large values of $\sigma_s \left|\rho_{nk}\right|$ to the high-SNR regime, and small values to the low-SNR regime.Since it is not practical to use the truncated series expansion, we want to approximate the curve *g(x) g20(x))* with another simpler function. We verified that one close approximation is of the form $\hat{g}(x,\alpha) = 1 - e^{-\alpha x}$, for some positive real α. In fact this function is also equal to zero for $x = 0$, is strictly lower than one for $x>0$ and converges to 1 for $x \to +\infty$. The coefficient was determined by minimizing the Mean Square Error between the approximation and $\tilde{g}_{20}(x) \square g(x)$.

Using this approach, we determined the optimum coefficient to be α =1 .0639. Therefore, the approximation to the posterior expectation of the unknown symbols can be written as

$$E_{V_n^{(bl)}(k)}\left[V_n^{(bl)}(k) \middle| Y_n^{(bl)}(k), h\right] \square \sigma_s e^{i\theta_{nk}} \left(1 - e^{-1.0639.\sigma_s \left|\rho_{nk}\right|}\right) \tag{43}$$

In figure 2 we show curve $g_{20}(x)$ and the approximation $\hat{g}(x, 1.0639)$, as well as the error on the amplitude.

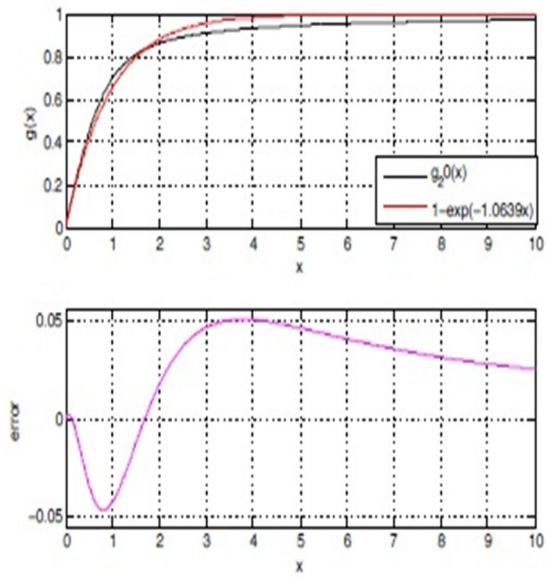
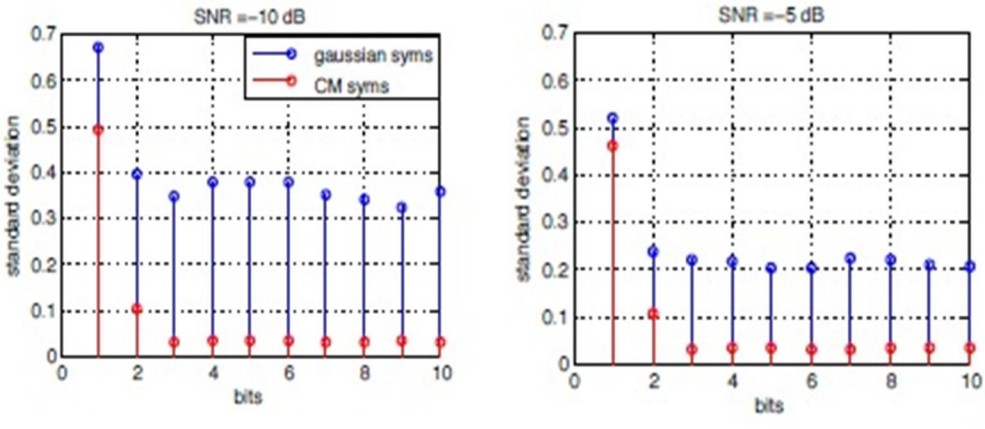


Fig 2. Plot of function $g(x)$ and its approximation $1-e^{-1.0639x}$



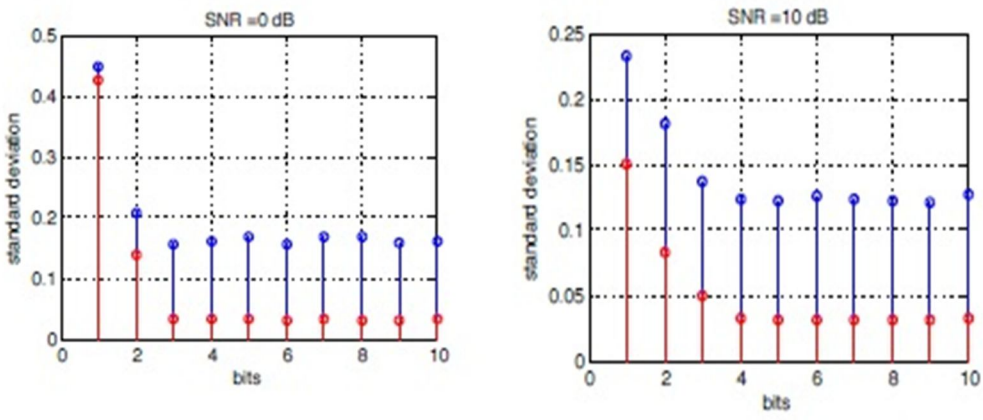


Fig 3. Gaussian approximation versus CM with uniform phase approximation, standard deviation on the posterior expectation; $N = L = 1$, $R = T = 1$.

Conclusion

It is interesting to compare the closeness of the posterior expectation using the Gaussian approximation (MMSE detector) and using the Constant Modulus approximation for the transmitted symbols to the true posterior expectation calculated averaging over the true discrete distribution of the symbols. Figure 3 shows the standard deviation of the error between the true posterior expectation and the approximated posterior expectation for different SNR and different number of bits per symbol, for the two cases where the symbols are assumed to be Gaussian distributed and where they are assumed to be Constant Modulus with phase uniformly distributed in $[0,2\pi)$. In this latter case the posterior expectation is calculated using the approximation to the posterior mean given by 3.69. It is worth noticing that the Constant Modulus approximation proposed leads to a significant improvement compared to the Gaussian assumptions, even for a small number of bits (the 2 bits case is particularly interesting, since this corresponds to the 4-QAM constellation used in the LTE system). Moreover, the standard deviation decreases over the number of bits, since the more bits there are, the more evenly the symbols are distributed on the unit circle, and the better their phase can be approximated as being uniform in $[0,2\pi)$.

References

[1] A.K.Jagannatham and B.D.Rao. Whitening-rotation-based semi-blind MIMO channel estimation. IEEE Transactions on Signal Processing, 54(3):861–869, March 2006.

[2] Pramod Viswanath David Tse. Fundamentals of Wireless Communication. Cambridge University Press, 2005.

[3] Nikolas P. Galatsanos Dimitris G. Tzikas, Aristidis C. Likas. The variational approximation for bayesian inference. IEEE Signal Processing Magazine, November 2008.

[4] J Ketonen, M Juntti, J Ylioinas and J R. Cavallaro, “Decision-Directed Channel Estimation Implementation for Spectral Efficiency Improvement in Mobile MIMOOFDM,” Springer Science, DOI 10.1007/s11265-013-0833-4, 2013.

[5] A Zaier and R Bouallègue, "Blind Channel Estimation Enhancement for MIMO- OFDM Systems under High Mobility Conditions," International Journal of Wireless & Mobile Networks (IJWMN) Vol. 4, No. 1, pp. 207214, February 2012.

[6] “LTE Release 12 and Beyond,” Ericsson, January 2013

[7] Q. Li, G. Li, W. Lee, M. il Lee, D. Mazzarese, B. Clerckx, and Z. Li, “MIMO techniques in WiMAX and LTE: a feature overview,” IEEE Commun. Magazine, vol. 48, no. 5, pp. 86–92, May. 2010.

[8] M.Sushanth Babu and Prof.K.Kishan Rao, “Performance Analysis of Trained and Semi-Blind Based Channel Estimation in Coded MIMO HSDPA Cellular System”,International Conference on Electronics & Communication Engineering (ICECE), PP. 99-103,25th Nov.2012, Chennai, INDIA. ISBN: 978-93-82208-38-9

[9] M.Sushanth Babu and Prof.K.Kishan Rao, “Fast Converging Semi-Blind SNR Estimation for Wireless MIMO-OFDM Systems,” IEEE International Conference on Signal Processing, Communications and Computing. ICSPCC 2011, Xi’an, Chinna.

[10] M.Sushanth Babu, Prof.K.Kishan Rao,P.Krishna, V.Ugendhar , “BER PerformanceAnalysis for HSDPA with MIMO – SBCE Predictors”, IEEE International Conference –ICICI-BME, Indonesia, Nov-2009,pp.196-200.

[11] P. Venkateswarlu and R. Nagendra “Channel Estimation Techniques in MIMO-OFDM LTE Systems”, Int. Journal of Engineering Research and Applications,ISSN : 2248-9622, Vol. 4, Issue 7(Version 1), July 2014, pp.157-161.