Blind Identification and Equalization of CDMA Signals Using the Levenvberg-Marquardt Algorithm

Mohammed Boutalline, Imad Badi, Belaid Bouikhalene, Said Safi

Abstract—In this paper we describe the Levenvberg-Marquardt (LM) algorithm for identification and equalization of CDMA signals received by an antenna array in communication channels. The synthesis explains the digital separation and equalization of signals after propagation through multipath generating intersymbol interference (ISI). Exploiting discrete data transmitted and three diversities induced at the reception, the problem can be composed by the Block Component Decomposition (BCD) of a tensor of order 3 which is a new tensor decomposition generalizing the PARAFAC decomposition. We optimize the BCD decomposition by Levenvberg-Marquardt method gives encouraging results compared to classical alternating least squares algorithm (ALS). In the equalization part, we use the Minimum Mean Square Error (MMSE) to perform the presented method. The simulation results using the LM algorithm are important.

Keywords—Identification and equalization, communication channel, Levenvberg-Marquardt, tensor decomposition

I. INTRODUCTION

IN this paper, we consider a multi-user access based on spread spectrum code DS-CDMA (Direct Sequence Code Division Multiple Access). In the current systems, Knowledge of CDMA codes at the receiver is operated to effect separation of user signals. In each communication system an identification of the transmission channel between the transmitter and the receiver is necessary in order to correct its distortions. Several methods exist, the most commonly used methods are learning by sending occasionally a known sequence between the transmitter and receiver. This operation is called equalization. However, if the channel varies rapidly over time, it is necessary to periodically send the learning sequence, which limits the useful rate (about 25% the total flow rate is spent learning in GSM, UMTS up to 50%) and entails the use of an important part of bandwidth. in order to save the resource of bandwidth, other techniques are called blind, not using training sequence to make the channel identification but they use structural information on sequences transmitted (such as higher order statistics. Other algorithms called semi-blind combine the last two approaches.

Conventionally, the problem of channel identification is based on the following matrix algebraic formulation:

$$Y = H \bullet S, \tag{1}$$

Which the objective is to find the parameters of the transmission channel H and / or data transmitted S from the received data Y . The transmission symbols by CDMA signals to an antenna array involve three diversities (spatial, spectral and temporal) which give the received signal a multilinear algebraic structure. Our approach is to collect samples of the signal in a tensor of order three of which each dimension characterizes a diversity. The decomposition of this tensor allows to extract the contribution of each user. The system structure differs depending on the case of spreading by direct paths only. In the literature, the blind estimation of the signal from each user [1] can be obtained by decomposition in terms of rank 1 tensor observations or PARAFAC decomposition (for PARAllel FACtor) [2], [3]. In the article [4], the authors showed that for a multipath channel generating the ISI, the equalization and blind separation of CDMA signals can be carried out jointly by a more general multilinear decomposition that PARAFAC, using the block components decomposition (BCD) tensor comments [1], [5], [6].

The presented algorithm in [6], [4], to calculate the BCD is an Alternating Least Squares (ALS). However, this algorithm converges slowly and sometimes it remains sensitive to bad conditioning extracted data. In this article, we show that it is possible to get rid of these disadvantages by calculating the BCD by the optimization method of Levenberg-Marquardt (LM) [7], which is a technique of Gauss-Newton.

II. PROBLEM STATEMENT

A. Analytical Model

In a system which uses the MIMO technique, we consider R users transmit simultaneously in the same bandwidth to a network with K antennas. Each signal is spread by the CDMA codes of length I, such as $T_s = IT_c$, where T_s is the symbol period and T_c the chip period.

Let the sequences $\left\{ S_j^{(r)} \right\}_{j=1}^J$ and $\left\{ C_i^{(r)} \right\}_{i=1}^I$, representing respectively J successive symbols transmitted by the user r and the I chips its spreading sequence. We note $h_r(t)$ the spreading waveform of this user :

$$h_r(t) = \sum_{i=1}^I C_i^{(r)} g(t - iT_c) \tag{2}$$

Where g(t) represents the filter shaping (raised cosinus).

We assume that for each user given r, the p th path is characterized by its delay τ_{rp}, the angle of arrival on the antenna array θ_{rp} and attenuation β_{rp}. Let L the length of the channel impulse response at the symbol rate, which means that the ISI contains L consecutive symbols. At the reception,

Mohammed Boutalline and Imad Badi are with the Department of Computer Science, Faculty of Sciences and Techniques, Sultan Moulay Slimane University, Beni Mellal, Morocco (e-mail: boutalline@gmail.com).

Said Safi and Belaid Bouikhalene are with the Department of Mathematics and Informatics, Polydisciplinary Faculty, Sultan Moulay Slimane University, Beni Mellal, Morocco.

the overall signal is sampled the chip rate and the observation interval of duration $JTs = IJTc$, for which we assume the channel is stationary.

In order to simplify notation, we consider the same number of paths P and interfering symbols L for each user. The ith sample associated with the jth symbol of global signal received by the kth antenna can be written :

$$y_{ijk} = \sum_{r=1}^{R} \sum_{p=1}^{P} a_k(\theta_{rp}) \sum_{l=1}^{L} h_{rp}(i + (l-1)I) s_{j-l=1}^{(r)} \quad , \quad (3)$$

Where $a_k(\theta_{rp})$ is the amplification of the kth antenna coefficient and θ_{rp} represent the incidence angle, in which the attenuation was incorporated θ_{rp} and where $h_{rp}(i + (l-1)I)$ is the sample of $h_{rp}(t - \tau_{rp})$ at the instant $t = (i + (l-1)I)T_c$. In the analytical model of (1), indices i, j and k corresponding to three differences induced by the system spectral diversity provided by the CDMA spreading $(i = 1 \ldots I)$, temporal diversity $(j = 1 \ldots J)$ and spatial diversity $(k = 1 \ldots K)$. Collect the IJK samples y_{ijk} into a tensor $Y \in \mathbb{C}^{I \times J \times K}$ along with order 3, ie a cube.

B. Algebraic Model

The algebraic multilinear model is strictly equivalent to the analytical model of (3) :

$$Y = \sum_{r=1}^{R} H_r \bullet S_r \bullet A_r \quad , \quad (4)$$

Equation (4) represents the block components decomposition (BCD) of the tensor of observations Y, shown in Fig 1. This decomposition is interpreted as follows : Y is a combination of all users, where each contribution is characterized by three components blocks :

- Tensor $H_r \in \mathbb{C}^{I \times L \times P}$
- Matrix $A_r \in \mathbb{C}^{K \times P}$
- Toeplitz matrix $S_r \in \mathbb{C}^{J \times L}$
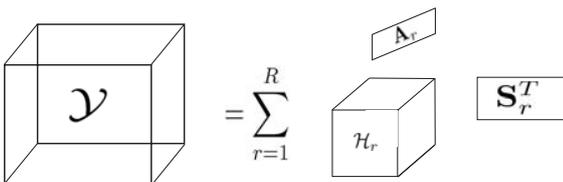
Fig. 1 Representation of the BCD

For a given path (p fixed index), each end portion of H_r is a matrix of size $I \times L$ which contains the IL successive samples $h_{rp}(i + (l-1)I)$ de $h_r(t - \tau_{rp})$. The first column of the matrix S_r contains the J successive symbols of the user r. This matrix has a Toeplitz structure caused by the presence of ISI. Finally, the matrix A_r contains the answer of K antennas in P different angles of arrival, i.e., $[A_r]_{k,p} = a_k(\theta_{rp})$. The BCD can be seen as a generalization of the PARAFAC decomposition, which corresponds to the particular case where $L = 1$ et $P = 1$.

Given the observation of Y only, the problem of blind equalization and separation of CDMA signals received involves the BCD of Y, to estimate unknown components H_r, S_r and A_r. The separation of R contributions based on the uniqueness of the decomposition, which has been demonstrated in [5], and implies a maximum number of users allowable simultaneously in the system. In particular, the uniqueness can still be guaranteed if the number of users exceeds the number of antennas $(R > K)$, providing that I and J are large enough. we impose a structure on Toeplitz matrix S_r in the equalization, i.e., the vectors $S_r \in \mathbb{C}^{(J+L-1) \times 1}$ generators of these matrices will be updated at each step, instead of matrices them selves. This approach is deterministic and is based on the algebraic structure of the tensor observations. Therefore, the sources are not necessarily statistically independent, the CDMA codes not necessarily orthogonal, and the geometry of the antenna array is unknown. In addition, This method can be used for relatively short frames, which is more flexible considerably the constraint of stationarity, compared to purely statistical methods.

If Y is the tensor of observations and $\hat{Y}$ an estimation of this tensor, the calculation of BCD of Y is to minimize the following cost function:

$$\phi = \frac{1}{2}\left\|Y - \hat{Y}\right\|_F^2 = \frac{1}{2}\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}|y_{ijk} - \hat{y}_{ijk}|^2, \quad (5)$$

Where $\|\Delta\|_F$ is the Frobenius norm defined between two vectors A and B by $\|A.B\|_F = \sqrt{tr(A^T B)}$, where $tr(.)$ denotes the trace and the superscript T denotes the matrix transpose [8], [9]. In other terms, given Y, the calculation of BCD is to find components blocks $\hat{H}_r$, $\hat{S}_r$ and $\hat{A}_r$ of tenseur $\hat{Y}$ which minimizes ϕ. In the next section, we propose to calculate the BCD by a Levenberg-Marquardt algorithm, which offers better performance than the proposed algorithm (ALS) in [4].

III. LEVENVBERG-MARQUARDT ALGORITHM

The Levenberg-Marquardt (LM) is an improvement of the conventional Gauss-Newton method, to solve the nonlinear least squares regression. The method is presented in detail in[10]. This is the recommended method for non-linear least squares problems, because it is more efficient compared to more general optimization algorithms (such as Quasi-Newton methods or simplex; see also nonlinear estimation for a presentation of other methods in the case of regression / non-linear estimation).

The Levenberg-Marquardt algorithm is based on iterative procedure. To start a minimization, the user has to provide an initial guess for the parameter vector p. In many cases, an uninformed standard guess like $p^T = (1, 1, \ldots, 1)$ will work fine; in other cases, the algorithm converges only if the initial guess is already somewhat close to the final solution [11], [12], [13].

In each iteration step, the parameter vector p is replaced by a new estimate $p + q$. To determine q, the functions $f_i(p + q)$

are approximated by their linearizations :

$$\mathbf{f(p+q) \approx f(p) + J_q} \qquad (6)$$

where J is the Jacobian of f at p. At a minimum of the sum of squares S, we have $\nabla_q S = 0$. Differentiating the square of the following expression :

$$S(a) = \sum_{i=1}^{m} [y_i - f(t_i|a)]^2$$

give us the following equation :

$$\mathbf{(J^T J)q = J^T[y - f(p)]} \qquad (7)$$

from which q can be obtained by inverting $J^T J$. The key to the LM Algorithm is to replace this equation by a 'damped version'

$$(J^T J + \lambda . I)q = -J^T f \qquad (8)$$

The damping factor λ is adjusted at each iteration. If reduction of S is rapid a smaller value can be used bringing the algorithm closer to the GNA, whereas if an iteration gives insufficient reduction in the residual λ can be increased giving a step closer to the gradient descent direction. A similar damping factor appears in *Tikhonov regularization*, which is used to solve linear *ill-posed problems*.

If a retrieved step length or the reduction of sum of squares to the latest parameter vector p fall short to predefined limits, the iteration is aborted and the last parameter vector p is considered to be the solution.

Choice of Damping Parameter

Various more or less heuristic arguments have been put forward for the best choice for the damping parameter λ. Theoretical arguments exist showing why some of these choices guaranteed local convergence of the algorithm; however these choices can make the global convergence of the algorithm suffer from the undesirable properties of steepest-descent, in particular very slow convergence close to the optimum.

The absolute values of any choice depends on how well-scaled the initial problem is. Marquardt recommended starting with a value λ_0 and a factor $\nu_0 > 1$. Initially setting $\lambda = \lambda_0$ and computing the residual sum of squares $S(p)$ after one step from the starting point with the damping factor of $\lambda = \lambda_0$ and secondly with λ/ν. If both of these are worse than the initial point then the damping is increased by successive multiplication by ν until a better point is found with a new damping factor of $\lambda\nu^k$ for some k.

If use of the damping factor λ/ν results in a reduction in squared residual then this is taken as the new value of λ (and the new optimum location is taken as that obtained with this damping factor) and the process continues; if using λ/ν resulted in a worse residual, but using λ resulted in a better residual then λ is left unchanged and the new optimum is taken as the value obtained with λ as damping factor.

To reduce the problem of slow convergence of LMA, the several improvements to the Levenberg-Marquardt algorithm existed [14] in order to improve both its convergence speed and robustness to initial parameter guesses. the update is done on the usual step to include a geodesic acceleration correction term, explore a systematic way of accepting uphill steps that may increase the residual sum of squares due to Umrigar and Nightingale, and employ the Broyden method to update the Jacobian matrix.

IV. SIMULATION RESULTS

We present in this section, the simulation results to solve the problem of blind equalization and separation by simulating the performance of the receiver based on multilinear BCD. First we illustrate the impact of data conditioning on the performance of LM and ALS algorithms for signal without noise. The identification process is initialized by random spreading codes with the following parameters:

- Length I=16 ;
- Sequences j = 30 QPSK symbols ;
- K=4 antennas ;
- L=3 interfering symbols per user;
- P=2 main paths per user;
- R= 5 users.

Fig. 2 shows the evolution of ϕ depending on the number of iterations, for several values of conditioning A, denoted $\kappa(A)$. The speed of convergence of ALS decreases drastically when $\kappa(A)$ increases, because ϕ encounters a bearing whose length depends on $\kappa(A)$. The LM algorithm provides a quadratic speed of convergence for the final iterations and abode very sensitive to the value of $\kappa(A)$. The Fig. 3 illustrates the effect for near-far of the non-noisy data. The tensor of observations is generated as follows :

$$\mathbf{Y} = \sum_{\mathbf{r=1}}^{\mathbf{R}} \alpha_{\mathbf{r}} \frac{\mathbf{y_r}}{\|\mathbf{y_r}\|_{\mathbf{F}}} \qquad (9)$$

Where α_r is the coefficient used to weigh the power of the contribution of each user.
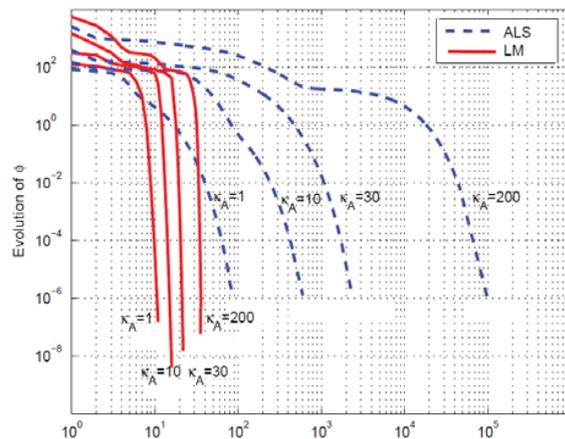
Fig. 2 Evolution of ϕ in terms of $k(A)$

In order to see the impact of the near-far effect on the observations at the reception, we perform several simulations where the data are obtained from a random (1000 simulations).
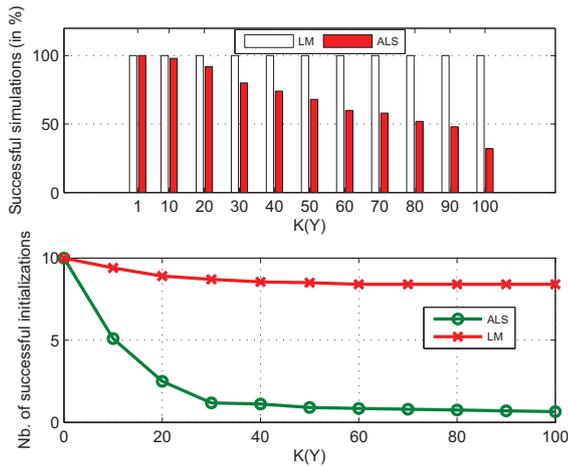
Fig. 3 Number of simulations and successful initializations

For each simulation performed, 10 initializations are tested. The process stops when the calculated value of the difference between the last two terms is less than or equal to 10^{-7} ($\|\hat{y}(n) - \hat{y}(n)\| \leq 10^{-7}$)for the case of two algorithms. It is seen that the ratio $K(y)$ ($\max(\alpha_r) / \min(\alpha_r)$) has an important value while the near-far effect has an influence on the observations.

In the case of noisy data by additive white Gaussian noise with zero mean , we consider a simulation is successful if the latter leads to $\phi <= 10^{-5}$ after the convergence of the algorithm. The top of Fig. 3 shows the percentage of successful simulations based on the value of $K(y)$.

This leads us to say that the ALS algorithm is very influenced by the near-far effect and convergence takes a little time because it is relative to the high value of $K(y)$. On the contrary the LM algorithm is a little sensitive to this effect (see Fig. 2).

The function cost is stagnant so the shutdown process is satisfied. The bottom of Fig. 3 shows the successful simulations where we calculated the number of successful initializations. In fact we can conclude that the ALS requires a number of iterations and resets all larger than $K(y)$ or $K(y)$ is great it shows us that the LM algorithm is certainly favorable to solve the problem of equalization.

Fig. 2, illustrates the near-far effect on data daubed by additive white Gaussian noise (AWGN) which has the Gaussian distribution with mean zero and whose variance is given by the tensor noise in dB :

$$SNR = 10 \, log_{10} \left(\frac{\|Y\|_F^2}{\|N\|_F^2} \right) \qquad (10)$$

In the following, we take this values : the spreading code length $I = 16$, frames symbols QPSK $J = 30$, $K = 4$ antennas, ISI on $L = 2$ symbols, $P = 2$ paths and $R = 3$ users. Then the equation tensor the observations is :

$$\hat{Y} = Y + N \qquad (11)$$

Where

$$Y = \sum_{r=1}^{3} \alpha_r \frac{y_r}{\|y_r\|_F}$$

Fig. 4, 5 and 6 illustrate the performance of the presented algorithms. These algorithms are compared using the Minimum Mean Squares Error (MMSE), which is recognized by a perfect identification of the channel parameters and the response of the antennas (unblinded approach). The results presented show that the evolution of BER by 3 users with coefficients $\alpha = 1$, $\alpha = 5,5$ and $\alpha = 10$ representing users 1,2 and 3 respectively . We conclude that the LM algorithm effectively approach the performance of MMSE (about 2 dB difference between two curves) against the ALS algorithm.
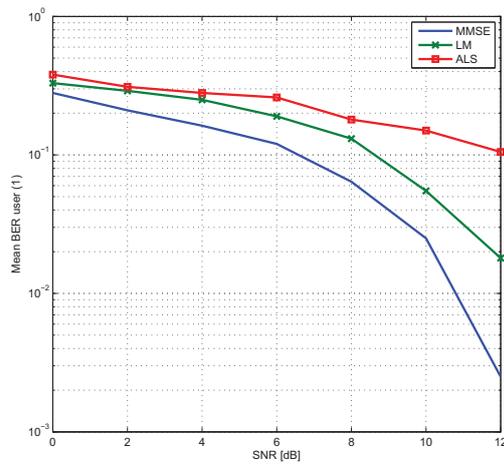
Fig. 4 BER of user 1
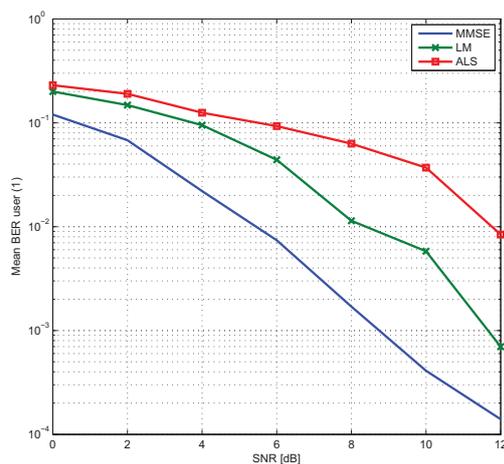
Fig. 5 BER of user 2

To verify the performance of the Levenvberg-Marquardt algorithm, we test it on a simulation example. In the case, the excitation input e (k) is random and i.i.d (identically and
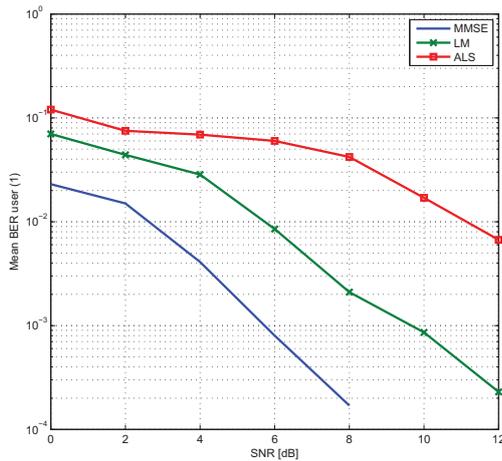
Fig. 6 BER of user 3

independent distributed). The signals y (k), s (k) and n (k) are related by (12) and (13). In order to measure the influence of noise, we define the signal to noise ratio.

$$y(n) = h(n) * e(n) = \sum_{i=0}^{q} h(i)e(n-i) \quad (12)$$

If the output $y(n)$ of the system is noisy by a white or colored Gaussian noise $w(n)$, the available output $s(n)$, practically, is given by the following equation:

$$s(n) = y(n) + w(n) \quad (13)$$

In order to measure the influence of noise, we define the signal to noise ratio (signal-to-noise ratio (SNR)) as :

$$SNR = 10log\left(\frac{\sigma_y^2(k)}{\sigma_w^2(k)}\right) \quad (14)$$

To measure the precision of the estimated parameters compared with real ones, we define the normalized mean square error (MSE) for each iteration :

$$EQM = \sum_{i=0}^{q} \left(\frac{h(i) - \hat{h}(i)}{h(i)}\right)^2 \quad (15)$$

V. APPLICATION : BRAN A CHANNEL

We test the performance of the equalization of MC-CDMA systems are evaluated using the LM algorithm. This evaluation is conducted by calculating the Binary error rate (BER) for the MMSE equalizer, using measured and estimated parameters of BRAN A channel.

The model representing the BRAN A channel , consisting of 18 paths, is given its impulse response by (16)

$$h_A(n) = \sum_{i=0}^{N_T} h_i \delta(n - \tau_i) \quad (16)$$

Where $\delta(n)$ is the Dirac function, h_i is the amplitude of the path i, $N_T = 18$ is the number of paths and τ_i temporal delay of path. Table I shows the values of the impulse response of the channel A BRAN.

TABLE I
DELAY AND MAGNITUDES OF 18 TARGETS OF BRAN A
RADIO CHANNEL

Delay τ_i [ns]	Mag.C_i [dB]	Delay τ_i [ns]	Mag.C_i [dB]
0	0	90	-7.8
10	-0.9	110	-4.7
20	-1.7	140	-7.3
30	-2.6	170	-9.9
40	-3.5	200	-12.5
50	-4.3	240	-13.7
60	-5.3	290	-18
70	-6.1	340	-22.4
80	-6.9	390	-26.7

We estimate the parameters of the BRAN A radio channel using the LM algorithm, This procedure allows us to have a good estimation of the parameters of the impulse response of the channel. For different signal to noise ratio, we represent in Fig.7 the estimated amplitude and phase of the BRAN A channel. From Fig. 7, we observe that the estimated values are very closed to measured one.
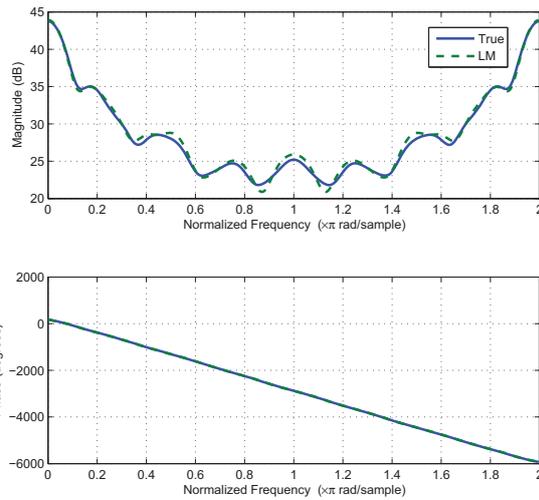
Fig. 7 Amplitude and phase of BRAN A channel response

We represent in Fig. 8, the BER simulation results obtained, using the measured and estimated data for BRAN A channel. Equalization is done by the MMSE technique. From Fig. 8, we observe that the blind equalization using the LM algorithm, give us approximately the same results using measured data. This is due to a good estimate of the amplitude and phase.

So if the SNR is equal to $16dB$, we will have only one wrong bit when 10^2 bits are received, but if the SNR is equal to $20dB$ we have one wrong bit for 10^3 bits received.
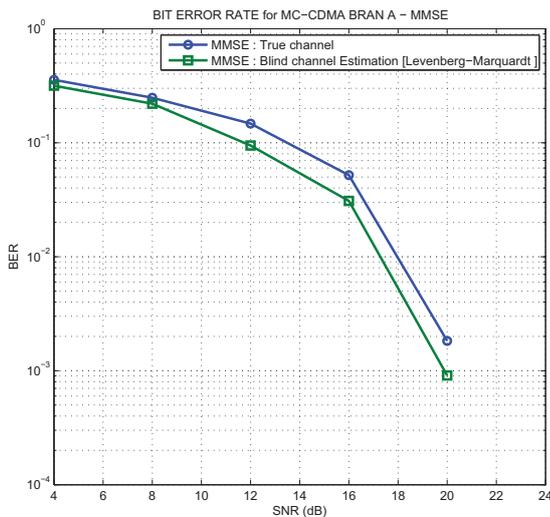
Fig. 8 BER of the estimated and measured of BRAN A channel

VI. Conclusion

According to the obtained results, the blind equalization problem of CDMA signals received by an antenna array can be formulated in terms of multilinear algebra. The problem is solved using the Levenberg-Marquardt algorithm based on block components decomposition of a tensor with three order. The levemberg-Marquardt optimisation method provides better performance than the classical alternate least square algorithm, especially for noisy data. According to the encouraging results the multilinear formulation can be used in other problems where three diversities are usable(the spatial diversity, spectral and temporal).

References

[1] L. D. Lathauwer, "Decompositions of a higher-order tensor in block terms - part 1 : Lemmas for partitioned matrices," *SIAM J. Matrix Anal. Appl, USA*, 2006.
[2] R. A. Harshman, "Foundations of the parafac procedure : Model and conditions for an 'explanatory' multi-mode factor analysis," *UCLA Working*, vol. Papers in Phonetics, 16 :, pp. 1–84, 1970.
[3] N. D. Sidiropoulos, G. B. Giannakis, and R. Bro, "Blind parafac receivers for ds-cdma systems," *Signal Processing, IEEE Transactions on*, vol. 48, no. 3, pp. 810–823, 2000.
[4] D. Nion and L. D. Lathauwer, "A block factor analysis based receiver for blind multi-user access in wireless communications," *In Proc. IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 825–828, 2006.
[5] L. D. Lathauwer, "Decompositions of a higher-order tensor in block terms - part 2 : Definitions and uniqueness," *SIAM J. Matrix Anal. Appl, USA*, 2006.
[6] L. D. Lathauwer and D. Nion, "Decompositions of a higher-order tensor in block terms - part 3 : Alternating least squares algorithms," *SIAM J. Matrix Anal. Appl, USA*, 2006.
[7] D. Marquardt, "An algorithm for least-squares estimation of non-linear parameters," *Technical University of Denmark*, pp. 431–441, 1963.
[8] C. F. V. L. Gene H. Golub, *Matrix Computationsn*. Johns Hopkins University Press, 3rd edition, October 15, 1996.
[9] N. J. Higham, *Accuracy and Stability of Numerical Algorithms*. Soc for Industrial & Applied Math, 1 edition, 1996.
[10] J. J. More, "The levenberg-marquardt algorithm: implementation and theory," *Springer, Berlin*, vol. 630, pp. 105–116, 1977.
[11] K. Levenberg, "A method for the solution of certain problems in least squares," *Quart. Appl. Math.*, vol. 2, pp. 164–168, 1944.
[12] P. E. Gill and W. Murray, "Algorithms for the solution of the nonlinear least-squares problem," *SIAM J. Numer. Anal*, vol. vol. 15, no 5, pp. 977–992, 1978.
[13] H. Gavin, "The levenberg-marquardt method for nonlinear least squares curve-fitting problems," *Department of Civil and Environmental Engineering, Duke University*, 2011.
[14] M. K. Transtrum and J. P. Sethna, "Improvements to the levenberg-marquardt algorithm for nonlinear least-squares minimization," *Journal of Computational Physics*, January 30, 2012.