Approximations to Joint-ML and ML Symbol-Channel Estimators in MUD CDMA

Thomas Fabricius
Informatics and Mathematical Modelling
Technical University of Denmark
Copenhagen, Denmark

Ole Nørklit
Algorithm Design
Nokia Mobile Phones
Copenhagen, Denmark

Abstract—**In this contribution will we conceptually derive two symbol-channel estimators, the joint-ML and the ML, both having exponential complexity. Pragmatically we derive three approximations with polynomial complexity, one to the joint-ML: the pseudo-joint-ML; two to the ML: the naive-ML and the linear-response-ML. We asses the resulting average bit error rates empirically. Performance gains of several dBs are observed from using the ML based approximations compared to the joint-ML.**

Keywords—**CDMA, Multiuser Detection, Blind channel estimation**

I. Introduction

Traditional code division multiple access (CDMA) systems are interference limited due to receivers designed using a code matched filter or correlator followed by detection. However, communication networks with high data rates, force the use of short spreading codes. This will generate additional multiple access interference (MAI) and in such situations interference will limit the receiver performance. As shown by Verdú [1] CDMA systems need not to be interference limited; when using an optimal multiuser detector (MUD) instead of the traditional receiver. These optimal receivers have exponential complexity, so sub-optimal solutions to the multiuser detection problem have been proposed [1], [2]. In traditional settings the channel is estimated from pilot sequences, but blind estimation of the channel has received focus, aiming to increase the spectral efficiency. A blind linear detector that minimises the output energy (MOE) is reported in [1], asymptotically the detector approaches the linear MMSE detector. In [3] the received signal is projected into a subspace, where a linear decorrelating or linear MMSE detector is formed, by only knowing the signature waveform from the user of interest. In [4] the likelihood for the unknown channel is established by using the statistical properties of the symbols and the signature waveform of the user of interest. Besides taking the a priori statistics of the symbols into account, the next natural step is to include the a priori statistics of the fading. The last approach is employed in [5], resulting in a Space Alternating Generalized EM (SAGE) algorithm.

Our contribution is twofold, first we derive the joint-ML symbol-channel estimator and an approximation to this, denoted pseudo-joint-ML detector. Secondly we derive the ML detector for the channel i.e. marginalising out the symbols, and two approximations to this: naive-ML and Linear-Response-ML. The reason for proposing these new estimators, is to reveal the better performance achieved from using both the apriori statistics of the symbols and the properties of the signature waveform in addition to a correct ML estimator of the channel. The approximations applied are inspired from [6] on Independent Component Analysis (ICA), being similar to Blind MUD CDMA [7]. The derivation is based on a relatively simple channel model: K Synchronous Users in Flat Fading CDMA; but the methods also applies to more general models. The paper is organised as follows: First we introduce the channel model; then we review optimal symbol detection according to both minimum probability of error and minimum BER. Next we derive the joint-ML detector and its approximation: the pseudo-joint-ML. Lastly we derive the ML detector and its approximations: the naive-ML and the linear-response-ML. The performance of the detectors are assessed empirically and discussed, followed by a conclusion.

II. K Users CDMA in Flat fading

Assuming a flat slow fading channel[1], and BPSK modulation the received base band CDMA signal can be modelled as

$$y(t) = \sum_{i=1}^{N} \sum_{k=1}^{K} A_k b_{k,i} s_k(t - iT) + \sigma n(t) \qquad (1)$$

where t is time and $s_k(t)$ is the k'th users signature waveform with unit energy and T is the bit duration, $b_{k,i} \in [-1, 1]$ is user k's i'th bit in the block of duration N, A_k is the k'th users fading coefficient, and $n(t)$ is a stationary white complex Gaussian process with unit spectral density.

We process the signal $y(t)$ with the conventional matched filter bank, each filter indexed by $k' \in [1; K]$, followed by integration and sampling at each bit instant i to get the sufficient statistics[2] (for A, b, and σ^2)

$$z_{k',i} = \int_0^T y(t) s_{k'}(t - iT) dt = \sum_{k=1}^{K} A_k b_{k,i} r_{k,k'} + n_{k',i} \qquad (2)$$

[1] We assume nothing about the distribution of the fading.

[2] We can understand this change to sufficient statistics as a projection onto a subspace spanned by the signature waveforms.

where $r_{k,k'}$ is the correlation between signal $s_k(t)$ and $s_{k'}(t)$, and $n_{k',i}$ is a Gaussian random variable with both real and imaginary part having zero mean and variance σ^2. Using the above we have the joint likelihood for A, b, and σ^2

$$p\left(Z\left|B,A,R,\sigma^2\right.\right)=\frac{1}{|2\pi\sigma^2 R|^N}\cdot \exp-\left[\mathrm{Tr}\left(\frac{1}{2\sigma^2}(Z-RAB)^{\mathrm{T}}R^{-1}(Z-RAB)\right)\right] \tag{3}$$

where we have introduced matrix notation using $(Z)_{k',i}=z_{k',i}$, $(B)_{k',i}=b_{k',i}$, $(A)_{k,k'}=\delta(k-k')A_k$, and $(R)_{k,k'}=r_{k,k'}$, $(\cdot)^{\mathrm{T}}$ means transpose and complex conjugate, and Tr is the usual matrix trace.

III. Optimal Detectors

Given the noisy data and the channel one goal is to minimise the expected BER, another to minimise the probability of error. The expected BER is defined as

$$\mathrm{BER}=\frac{1}{2}\langle 1-\frac{1}{NK}\,\mathrm{Tr}\,B^{\mathrm{T}}\widehat{B}\rangle_{p(Z,B|A,\sigma^2)} \tag{4}$$

where the average $\langle\cdot\rangle_{p(Z,B|A,\sigma^2)}$, as indicated, is taken with respect to all bit realizations B and all noise realizations Z. The joint distribution $p(Z,B\left|A,\sigma^2\right.)$ is found by the likelihood (3) multiplied by the a priori distribution of the transmitted symbols $p(B)$

$$p(Z,B\left|A,\sigma^2\right.)=p\left(Z\left|B,A,\sigma^2\right.\right)p(B) \tag{5}$$

where the dependence on R is omitted due to the fact that it is known.

It can be shown, under some regularity conditions, that the expected BER subject to the constraint that $\left|(\widehat{B})_{k',i}\right|=1$ is minimised by

$$\widehat{B}=\mathrm{sgn}\langle B\rangle_{p(B|Z,A,\sigma^2)} \tag{6}$$

for all given received data Z, sgn working element wise.

The probability of error is defined as the probability that at least one of the NK bits in the block are detected wrongly. We can define $p(e=1\left|B\right.)=1-\delta(NK-\mathrm{Tr}\,B^{\mathrm{T}}\widehat{B})$ and so the probability of error becomes

$$p(e=1\left|A,\sigma^2\right.)=1-\langle\delta(NK-\mathrm{Tr}\,B^{\mathrm{T}}\widehat{B})\rangle_{p(Z,B|A,\sigma^2)} \tag{7}$$

which is shown to be minimised for any received data Z by

$$\widehat{B}=\underset{B\in[-1;1]^{NK}}{\arg\max}\,p(B\left|Z,A,\sigma^2\right.). \tag{8}$$

Equation (8) is the well known Maximum A Posterior (MAP) solution, which reduces to maximum likelihood for uniform prior distribution $p(B)$.

The complexity per bit for both methods are $\mathcal{O}(2^K)$ due to the summation and exhaustive search, which is intractable due to the exponential law in K.

IV. Joint-ML Symbol-Channel Estimation

Constructing the joint-ML estimator, we divide the maximisation into two parts; one for the symbols given the channel and one for the channel given the symbols. Using $-\log$ of the joint likelihood (3) the symbol maximisation becomes

$$\widehat{B}=\underset{B\in[-1;1]^{NK}}{\arg\min}\left(Z-RAB\right)^{\mathrm{T}}R^{-1}(Z-RAB)) \tag{9}$$

which is identical to (8) for known A. The update for the channel given the symbols is a convex optimisation problem, hence we can take the partial derivative with respect to the channel A and σ^2, equate to zero and solve

$$\mathrm{diag}(A)=\mathrm{diag}(Z\widehat{B}^{\mathrm{T}})^{\mathrm{T}}Q^{-1} \tag{10}$$

$$\sigma^2=\frac{\mathrm{Tr}\left(R^{-1}ZZ^{\mathrm{T}}+A^{\mathrm{T}}R^{\mathrm{T}}A\widehat{B}\widehat{B}^{\mathrm{T}}-2A\widehat{B}Z^{\mathrm{T}}\right)}{NK} \tag{11}$$

where $(Q_{JML})_{k',k}=(\widehat{B}\widehat{B}^{\mathrm{T}})_{k',k}(R)_{k',k}+NI_{k,k'}$ and $\mathrm{diag}(\cdot)$ of a matrix means the diagonal of that matrix arranged as a column vector. Since the symbols are discrete the optimisation in equation (9) has to be carried out by enumeration[3], i.e. exhaustive search, to surely obtain the global joint-ML solution. We see that the evaluation of σ^2_{JML} can be omitted if only the symbol estimates are of interest. Asymptotically $(N\to\infty)$ the joint-ML detector will minimise the probability of error, since the channel estimate will approach the actual channel and equation (9) becomes equivalent to equation (8). This results in the complexity being equivalent to the complexity of (8).

A. Pseudo-joint-ML Estimation

Since the complexity of the joint-ML estimator is exponential in K we derive a local search algorithm for the symbols. Taking the partial derivative of $-\log$ of the joint likelihood (3) with respect to B, subject to the constraint $|(B)_{k',i}|=1$, equating to zero, and solving yields

$$B=\mathrm{sgn}\left(\mathrm{Re}\,A^{\mathrm{T}}\left(Z-(R-I)AB\right)\right) \tag{12}$$

where sgn acts on each element. Taking partial derivatives with respect to A and σ^2, equating to zero and solving yields the equations (10) and (11) for A and σ^2.

The equations holds in minima since the gradient of the joint likelihood equals zero. Solving the equations can be carried out in many ways leading to e.g. hard decision successive interference cancellation [2] (HSIC) or hard decision parallel interference cancellation (HPIC). Updating equation (12) suffers from local minima and hence depends on the initial values. Thus, it does not in general achieve the joint-ML solution, this is the reason for the name pseudo-joint-ML. In this context we can see the update of A and σ^2 as channel estimation using hard decision feedback from equation (12). We can see the pseudo-joint-ML as trying to minimise the probability of error.

[3]Equation (10) can eventually be substituted back into (9) before the enumeration.

V. ML CHANNEL ESTIMATOR

Since the joint likelihood is the likelihood for the symbols jointly with the channel, we now derive the correct likelihood for the channel alone.

In ML settings it is common to distinguish between hidden variables (incomplete data), visible variables (complete data besides the incomplete data) and parameters[4]. The likelihood of the channel parameters A and σ^2 is found as the joint likelihood times the prior of the hidden variables B marginalized with respect to the hidden variables B. The likelihood for A and σ^2 then becomes

$$p(Z\,|A,R,\sigma^2\,) = \sum_{B\in[-1,1]^{NK}} \frac{p(Z\,|B,A,R,\sigma^2\,)}{2^{NK}} \qquad (13)$$

where the prior is $p(B) = 2^{-NK}$. Taking partial derivatives of $-\log$ of the likelihood (13) with respect to A, σ^2 respectively, equating to zero and solving we get

$$\mathrm{diag}(A) = \mathrm{diag}(Z\langle B\rangle^{\mathrm{T}})Q^{-1} \qquad (14)$$

$$\sigma^2 = \frac{\mathrm{Tr}\left(R^{-1}ZZ^{\mathrm{T}} + A^{\mathrm{T}}R^{\mathrm{T}}A\langle BB^{\mathrm{T}}\rangle - 2A\langle B\rangle Z^{\mathrm{T}}\right)}{NK} \qquad (15)$$

with

$$(Q)_{k',k} = (\langle BB^{\mathrm{T}}\rangle)_{k',k}(R-I)_{k',k} + NI_{k',k}. \qquad (16)$$

where we define $\langle\cdot\rangle = \langle\cdot\rangle_{p(B|Z,A,R,\sigma)}$ i.e. the posterior expectation. We see that the ML estimator, contrary to pseudo-joint-ML estimator, needs knowledge of the noise variance. The ML estimator of A and σ^2 is a more efficient estimator than the joint-ML estimator. This can be explained by the fact that the ML estimate of the channel depends on $\langle B\rangle$ and $\langle BB^{\mathrm{T}}\rangle$ which fluctuates less than the B that maximises the joint likelihood and BB^{T} used in the joint-ML channel estimate (10). Since we have to calculate $\langle B\rangle$, we get the optimal BER symbol estimator, equation (6): $\widehat{B} = \mathrm{sgn}\langle B\rangle$, as a side product. Asymptotically the channel estimate approaches the actual channel, implying the symbol estimate asymptotically minimises the expected BER. Unfortunately this also implies the same complexity as (6).

A. Approximate ML Channel Estimator

Because of the high complexity to obtain $\langle B\rangle$ and $\langle BB^{\mathrm{T}}\rangle$, the aim now is to approximate these in a less costly way than doing the exact likelihood estimates.

With $-\log$ likelihood defined by

$$-\log p(Z\,|A,R,\sigma^2\,) = -\log\sum_{B\in[-1,1]^{NK}} \frac{p(Z\,|B,A,R,\sigma^2\,)}{2^{NK}} \qquad (17)$$

[4]This is contrary to the Bayesian approach where we only have hidden variables and visible variables, the first including the parameters.

we now propose an arbitrary distribution $q(B)$ which has support where the posterior distribution $p(B\,|Z,A,R,\sigma^2\,)$ has support. Using Jensens inequality we obtain

$$-\log p(Z\,|A,R,\sigma^2\,) \le -\sum_{B\in[-1,1]^{NK}} q(B)\log\frac{p(Z\,|B,A,R,\sigma^2\,)}{q(B)2^{NK}} \qquad (18)$$

Equality is obtained if $q(B) = p(B\,|Z,A,R,\sigma^2\,)$ which can be proved by Bayes rule. Making this choice will obviously have the same complexity as calculating the likelihood exactly. Now, the idea is to come up with a distribution $q(B)$ that makes the calculation of the right side of (18) tractable and makes the bound tight. The resulting distribution being an approximation to the posterior distribution of B. If $q(B)$ is parameterised we can use (18) as a cost function to minimise with respect to the parameters under the constraint that $q(B)$ sums to one.

Here we take

$$q(B) = \prod_{k=1}^{K}\prod_{i=1}^{N} q_{ki}(B_{ki}) \qquad (19)$$

i.e. to factorise completely, where

$$q_{ki}(B_{ki}) = \left[\frac{1+m_{ki}}{2}\right]^{\left(\frac{1+B_{ki}}{2}\right)}\left[\frac{1-m_{ki}}{2}\right]^{\left(\frac{1-B_{ki}}{2}\right)}, \qquad (20)$$

is chosen as a Bernoulli distribution with parameter $\frac{1+m_{ki}}{2}$. This parameterisation is chosen to make $\langle B_{ki}\rangle_{q_{ki}(B_{ki})} = m_{ki}$. This approximation is in the physics literature [8] referred to as the naive approximation.

With the above $q(B)$ we write out the right side of equation (18) and take the partial derivative with respect to the m_{ki}'s under the constraint $\sum q(B) = 1$ and equating to zero. This yields the equation for the matrix $(M)_{ki} = m_{ki}$

$$M = \tanh\left(\frac{1}{\sigma^2}\,\mathrm{Re}\left(A^{\mathrm{T}}Z - A^{\mathrm{T}}(R-I)AM\right)\right) \qquad (21)$$

We see that (21) only differs from pseudo-joint-ML updates in (12) by $\tanh(\frac{\cdot}{\sigma^2})$ instead of $\mathrm{sgn}(\cdot)$.

To update A and σ^2 we use the ML estimate (14) and (15) described in the previous section with the approximation that $\langle B\rangle \simeq \langle B\rangle_{q(B)} = M$ and $\langle BB^{\mathrm{T}}\rangle \simeq \langle BB^{\mathrm{T}}\rangle_{q(B)} = \langle B\rangle_{q(B)}\langle B^{\mathrm{T}}\rangle_{q(B)} = MM^{\mathrm{T}}$, where we have used that $q(B)$ factorises.

We see that the algorithm obtained is very similar to that of the pseudo-joint-ML, except we get soft decisions instead of hard decisions. The way we solve equation (21) gives either soft SIC (SSIC) or soft PIC (SPIC). If we use successive updates i.e. update one M_{ki} at a time we can prove that the bound (18) is made tighter in each update of M_{ki}. Since the updates of A and σ^2 are convex given M this also tightens the bound. The result being a Generalized EM algorithm[5].

[5]For this to be true both steps have to do hill-climbing towards the true likelihood, but not necessarily going to a minimum in each step. The proof is left out due to limited space.

B. Linear Response Correction to Channel Estimate

The ML estimate in (14) and (15) requires knowledge of both $\langle B\rangle$ and $\langle BB^{\mathrm{T}}\rangle$. In the previous section these were approximated by taking averages with respect to $q(B)$. This gave a naive approximation to $\langle BB^{\mathrm{T}}\rangle$ due to the fact that the correlations are neglected when $q(B)$ is chosen to factorise. A correction to the estimate of $\langle BB^{\mathrm{T}}\rangle$ can be found by using the following identity for the second cumulant

$$\begin{aligned}\chi_{kik'i'} &= \langle B_{ki}B^{\mathrm{T}}_{k'i'}\rangle - \langle B_{ki}\rangle\langle B^{\mathrm{T}}_{k'i'}\rangle \\ &= \sigma^2 \frac{\partial\langle B_{ki}\rangle}{\partial\,\mathrm{Re}(Z^{\mathrm{T}}A)_{i'k'}}\end{aligned} \tag{22}$$

Proof:

$$\begin{aligned}\sigma^2 \frac{\partial\langle B_{ki}\rangle}{\partial\,\mathrm{Re}(Z^{\mathrm{T}}A)_{i'k'}} &= \sigma^2 \frac{\partial \sum B_{ki} p(B\,|Z,A,R,\sigma^2)}{\partial\,\mathrm{Re}(Z^{\mathrm{T}}A)_{i'k'}} \\ &= \sigma^2 \frac{\partial \frac{1}{p(Z|A,R,\sigma^2)} \sum B_{ki} p(B,Z\,|A,R,\sigma^2)}{\partial\,\mathrm{Re}(Z^{\mathrm{T}}A)_{i'k'}} \\ = \frac{\sum B_{ki}B_{k'i'}p(B,Z\,|A,R,\sigma^2)}{p(Z\,|A,R,\sigma^2)} \\ - \frac{\sum B_{ki}p(B,Z\,|A,R,\sigma^2)\sum B_{k'i'}p(B,Z\,|A,R,\sigma^2)}{p(Z\,|A,R,\sigma^2)^2} \\ &= \langle B_{ki}B^{\mathrm{T}}_{k'i'}\rangle - \langle B_{ki}\rangle\langle B^{\mathrm{T}}_{k'i'}\rangle \qquad \square\end{aligned}$$

But since we don't want to calculate $\langle B_{ki}\rangle$, we employ the following approximation

$$\frac{\partial\langle B_{ki}\rangle_{p(B|Z,A,R,\sigma)}}{\partial\,\mathrm{Re}(Z^{\mathrm{T}}A)_{i'k'}} \simeq \frac{\partial M_{ki}}{\partial\,\mathrm{Re}(Z^{\mathrm{T}}A)_{i'k'}} \tag{23}$$

which is called the linear response correction [8].

We now do the partial derivative on

$$\begin{aligned}M_{ki} = \tanh(&\frac{1}{\sigma^2}[\mathrm{Re}\left((A^{\mathrm{T}}Z)_{ki}\right) \\ &- \sum_{l=1}^{K} \mathrm{Re}\left((A_{kk})^*(R-I)_{kl}A_{ll}M_{li}\right)])\end{aligned} \tag{24}$$

from (21) to yield the approximation (23) for the second cumulant

$$\begin{aligned}\chi_{kik'i'} &\simeq \sigma^2 \frac{\partial M_{ki}}{\partial\,\mathrm{Re}(Z^{\mathrm{T}}A)_{i'k'}} \\ &= \delta(i-i')(1-M^2_{ki'})[\delta(k-k') \\ &\quad - \frac{1}{\sigma^2}\sum_{l=1}^{K}\mathrm{Re}\left((A)^*_{kk}(R-I)_{kl}A_{ll}\right)\chi_{ki'k'i'}].\end{aligned} \tag{25}$$

Isolating $\chi_{kik'i'}$ we get

$$\chi_{kik'i'} = \delta(i-i')\left((C_{i'}^{-1} + \frac{1}{\sigma^2}A^{\mathrm{T}}(R-I)A)^{-1}\right)_{kk'} \tag{26}$$

where the diagonal matrix $C_{i'}$ is defined as $(C_{i'})_{kk} = 1 - M^2_{ki'}$.

Now we can estimate

$$(\langle BB^{\mathrm{T}}\rangle_{p(B|Z,A,R,\sigma)})_{kk'} \simeq \sum_{i=1}^{N}\chi_{kik'i} + (MM^{\mathrm{T}})_{kk'} \tag{27}$$

which can be used directly in (14) and (15) to improve the estimate of A and σ^2.

C. Complexity of the Approximate Algorithms

The updates in equation (12) and (21) have the same complexity of order $\mathcal{O}(K)$ per bit. The updates for A in equation (14) has a complexity of $\mathcal{O}(K)$ for $N > K$ and $\mathcal{O}(\frac{K^2}{N})$ per bit for $N < K$. The linear response corrected estimate requires $\mathcal{O}(K^2)$ due to the fact that at each bit instant we have to invert a $K \times K$ matrix.

VI. Simulations and Discussion

Data are generated according to the distribution (3), R being the correlations of the signature waveforms, here calculated on the basis of random PN-sequences, with N_c chips. The channel coeffiecients A are all constructed to have unit length and random phase, the block length is fixed to $N = 100$. For each block simulated new random sequences are generated. We define the system load as the ratio of the number of users to the spreading gain $\alpha = \frac{K}{N_c}$.

We use the three proposed detectors as follows: We first make an initial guess on A by using 4 preamble bits, then we use this to get an initial guess on $B = \mathrm{sgn}((RA)^{-1}Z)$ or $B = \tanh((RA)^{-1}Z)$ corresponding to either joint-ML or ML algorithms. We calculate the initial noise variance by equation (11). Then we process two stages of each algorithm, with successive updates in decending power order.

In figure (1) we present Monte Carlo studies of the algorithms. Every point is simulated until 100 bit errors were reached, giving an accuracy on BER of 10% [6].

In the left plot of figure (1) we have simulated the Bit Error Rate of user 1 where SNR_1 is the corresponding signal to noise ratio for $K = 20$ users and the number of chips $N_c = 24$. The conventional detector fails to estimate the transmitted bits due to the high SIR produced by the non-orthogonal signature waveforms. A rutine calculation shows that $SIR_1 = \frac{1}{\sum_{k=2}^{K} r^2_{1k}}$, where we can use $|r_{1k}| \simeq \frac{1}{\sqrt{N_c}}$ on average when using random PN-sequences. This implies $SIR_1 = \frac{N_c}{K-1}$ which for the given case yields a bit error rate of $Q(\sqrt{SIR_1}) = 0.13$ in the absence of noise. At a BER of $P_b = 10^{-3}$, linear-response-ML detector gains $1dB$ compared to naive-ML detector and $2dB$ compared to pseudo-joint-ML detector. The loss down to the single user bound is around 7 dB.

[6] Strictly this is an optimistic error bar, since the bit errors inside a block probably are correlated due to the estimate of A and σ^2 but at the desired BER=10^{-3} we mainly observed single errors.

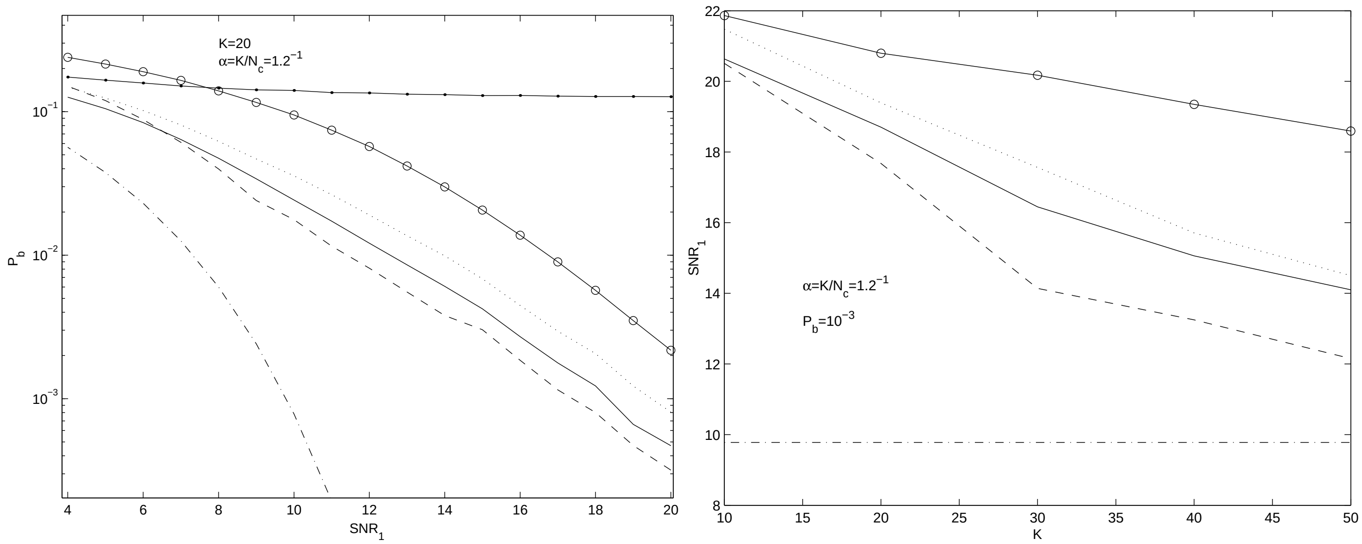

Fig. 1. Left plot: BER for user 1 for the three proposed methods, the single user bound, conventional, and the initial detector, at $K = 20$ and $N_c = 24$. Right plot: SNR of user 1 to obtain a BER$= 10^{-3}$ for different number of users at a constant load $\alpha = \frac{K}{N_c} = 1.2^{-1}$. Conventional detector: solid-dot, Initial detector: solid with round markers, pseudo-joint-ML: dotted, naive-ML: solid, linear-response-ML: dashed, and Single user bound: dash-dot.

The plot to the right shows the performance for a fixed system load $\alpha = 1.2^{-1}$ i.e. fixed spectral efficiency and fixed BER $P_b = 10^{-3}$. At first sight it seems a bit contra intuitive that the performance is increased for increasing K, the SIR$\simeq \alpha^{-1}$ so we could expect the performance to be equal for all K. But in the derivation of the algorithms we have implicitly assumed that the MAI can be approximated by a gaussian variable, i.e. using the Central Limit theorem. The conclusion is that MAI looks more and more gaussian for increasing number of users K and hence performance is increased for fixed load α. As expected, due to the soft tentative decission, naive-ML perform better than pseudo-joint-ML for all K, but it is hard to quantize whether this performance gain is due to the better channel estimate or the soft vs. hard tentative decission. For the linear-response-ML vs. the naive-ML we see that we gain up to 2 dB, this because of the improved channel estimate, since the tentative decisions in both cases are soft. In future work will we examine the dependence on N and the near-far resistance.

VII. Conclusion

In this contribution three approximative detectors for joint symbol-channel estimation are derived. One approximating the joint-ML detector, and two approximating the ML detector for the channel. The empirical results shows a bennefit of using approximations to ML for the channel instead of the approximating joint-ML detector. Furthermore can it be concluded that the cross-correlations between individual bits are important for the ML channel estimate.

Acknowledgements

The authors would like to thank Professor Dr. Bernard Fleury, Aalborg University Denmark and Dr. Niels Mørch, Nokia Denmark for their carefull reading of the manuscript.

References

[1] Sergio Verdú, *Multiuser Detection*, Cambridge University Press, The Pitt Building, Trumpington Street, Cambridge, CB2 1RP United Kingdom, 1998.

[2] M. Varanasi, B Aazhang, "Near-optimum detection in synchronous code-division multiple-access systems," *IEEE Trans. Communication*, vol. 39, no. 5, pp. 725–735, May 1991.

[3] Wang Xiaodong, H.V. Poor, "Blind multiuser detection: a subspace approach," *IEEE Trans. on Information Theory*, vol. 44, pp. 677–690, Mar. 1998.

[4] M. F. Bugallo, J. Mí guez, L. Castedo, "A Maximum Likelihood Approach to Blind Multiuser Interference Cancellation," *IEEE Trans. On Signal Processing*, vol. 49, no. 6, pp. 1228–1239, June 2001.

[5] A. Kocian, B.H. Fleury, "Iterative joint symbol detection and channel estimation for DS/CDMA via the SAGE algorithm," in *Proc. 11th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, 2000 (PIMRC'2000).*, 2000, vol. 2, pp. 1410–1414.

[6] P. A.d.F.R. Højen-Sørensen, O. Winther, L. K. Hansen, "Ensemble Learning and Linear Response Theory for ICA," in *In Advances in Neural Information Processing Systems 13 (NIPS'2000).*, T. K. Leen, T. G. Dietterich, V. Tresp, Ed. 2001, pp. 542–548, MIT Press.

[7] J. Joutsensalo, T. Ristaniemi, "Learning algorithms for blind multiuser detection in CDMA downlink ," in *Proc. 9th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, 1998 (PIMRC'1998).*, 1998, vol. 3, pp. 1040–1044.

[8] Giorgi Parisi, *Statistical Field Theory*, Addison-Wesley Publishing Company, New York, 1994.