NOISE VARIANCE ESTIMATION IN DS-CDMA AND ITS EFFECTS ON THE INDIVIDUALLY OPTIMUM RECEIVER

R. Gaudel, F. Bonnet, J.B. Domelevo-Entfellner

ENS Cachan
Campus de Ker Lann
35170 Bruz, France

A. Roumy

IRISA-INRIA
Campus de Beaulieu
35042 Rennes Cedex, France

ABSTRACT

In the context of synchronous random DS-CDMA (Direct Sequence Code Division Multiple Access) communications over a mobile network, the receiver that minimizes the per-user bit error rate (BER) is the symbol Maximum a posteriori (MAP) detector. This receiver is derived under the hypothesis of perfect channel state information at the receiver. In this paper we consider the case where the channel noise variance is estimated and analyze the effect of this mismatch. We show that the Bit Error Rate (BER) is piecewise monotonic wrt. the estimated noise variance, reaching its minimum for the true channel variance. We also provide an upper bound of the individually optimum receiver performance under noise variance mismatch. Thus we give a theoretical justification for the usual bias towards noise variance underestimation adopted by the community.

1. INTRODUCTION

Code Division Multiple Access (CDMA) is still the industry standard for today's mobile networks and is likely to remain at the core of some next generation technologies. We can think of 3GPP2, i.e. cdma2000, HDSPA, adopted by the 3GPP community and based on a wideband CDMA, or China's standard based on a Time-Division CDMA. All these prospects make it still highly beneficial to study the CDMA model.

The optimal multiuser receiver [7], in the sense of minimum per user bit error rate (BER) is the symbol Maximum a posteriori (MAP) detector and is also referenced as individually optimum receiver [8]. The derivation and analysis of this receiver [7] assume that the channel characteristics (and in particular the channel noise variance) are perfectly known at the receiver. However, the receiver does not know perfectly the noise variance and has to estimate it.

Various methods have been proposed to estimate the channel noise variance or equivalently the signal to noise ratio (SNR). Most methods [3, 4, 5] compute a variance estimate based on the moments of the received observation. The interest in noise variance estimators has grown with the introduction of powerful turbo-codes [2], decoded by means of the symbol MAP decoder that needs to know the channel noise variance. Then, in the context of turbo-codes, [6, 9] study the effect of SNR mismatch and conclude that overestimation of SNR is less detrimental than underestimation. In this paper we give a theoretical justification for this result in the context of an instantaneous mixture that is synchronous random DS-CDMA.

We compute here the performance behavior of the individually optimum receiver wrt. the noise variance mismatch and prove that the function BER (σ_e) decreases monotonically from $\sigma_e = 0$ to $\sigma_e = \sigma$, and then increases from $\sigma_e = \sigma$ to $\sigma \to \infty$, where the individually optimum receiver behaves like a simple bank of Matched Filters (MF).

The rest of the present article is organized as follows: in Section 2 we provide the reader with the theoretical background concerning the communication model used, and in Section 3 we give the results obtained. We finally draw some conclusions in Section 4.

Throughout this paper we will use the notation $y_{i:j} = (y_i, \ldots, y_j)$ for any sequence $\{y_n\}$.

2. THEORETICAL BACKGROUND

In this section, the transmitter model (subsection 2.1) and the individually optimum receiver will be presented. For the sake of the following analysis, we will also present two other receivers: the conventional detector (subsection 2.2) and the jointly optimum receiver (subsection 2.4).

2.1. Transmitter model

Consider a K-user synchronous DS-CDMA system. User k is assigned a signature $s_k(t), t \in [0, T]$ and a data symbol b_k to be transmitted over the channel with a signal amplitude A_k. The information concerning user k is therefore a signal

A. Roumy is supported in part by the Network of Excellence in Wireless Communications (NEWCOM), E. C. Contract no. 507325.

sig(k,t) during a time period T:

$$\text{sig}(k,t) = A_k b_k s_k(t)$$

It follows that the received continuous-time real baseband signal is

$$y(t) = \sum_{k=1}^{K} A_k b_k s_k(t) + n(t)$$

where $n(t)$ is a zero-mean random Gaussian noise with variance σ^2. In this paper we consider BPSK data modulation, the b_k taking their values in the alphabet $\{-1,+1\}$. Moreover the transmitted symbols are assumed to be equally probable.

We consider here also random DS-CDMA such that the assigned signatures are correlated. We define as ρ_{ij}, the correlation between the signatures of user i and j:

$$\rho_{ij} \triangleq \int_0^T s_i(t) s_j(t) dt \tag{1}$$

2.2. Conventional detector

The Conventional detector consists of a bank of matched filters followed by decision devices.
The bank of matched filters is a bench of K correlators, one for each user: if we consider the k-th correlator to implement the simple function

$$y_k = \int_0^T y(t) s_k(t) dt \tag{2}$$

the receiver obtains from the signal $y(t)$ a series of K values $y_k, k \in \{1,...,K\}$. It will then decide about b_k being ± 1 considering its estimate $\hat{b}_k = \text{sgn}(y_k)$.

2.3. Individually Optimum Detector

The Individually Optimum Detector minimizes the individual probability of error (or BER of each user). It computes the most probable symbol given the signal received during the period T:

$$\hat{b}_k = \arg\max_{b_k} \quad p\left(b_k \mid \{y(t)\}_{0 \le t < T}\right) \tag{3}$$

This detector is therefore also called symbol MAP detector. Under the hypothesis of white Gaussian noise, the rule in the 2-user case for this receiver to decide whether $b_1 = \pm 1$ is as follows [8]:

$$\hat{b}_1 = \text{sgn}\left(y_1 - \frac{\sigma^2}{2A_1} \log \left(\frac{\cosh\left[\frac{A_2 y_2 + A_1 A_2 \rho_{12}}{\sigma^2}\right]}{\cosh\left[\frac{A_2 y_2 - A_1 A_2 \rho_{12}}{\sigma^2}\right]} \right) \right)$$

The symmetric equation holds for $\hat{b}_2$. More generally the set of K scalars $y_{1:K})$ is a sufficient statistic for b_k, where y_k is the output of the k-th matched filter (2).

2.4. Jointly optimum Detector

The Jointly optimum Detector minimizes the joint probability of error (i.e. averaged BER of all the users). The decision rules are such to maximize the joint probability of the K-uple $b_{1:K}$ given the signal received during the period T:

$$\hat{b}_{1:K} = \arg\max_{b_{1:K}} \quad p\left(b_{1:K} \mid \{y(t)\}_{0 \le t < T}\right) \tag{4}$$

This receiver can also be called sequence MAP detector. Under the white Gaussian noise assumption, the decision rules for b_1 and b_2 in the 2-user case read [8]

$$\begin{aligned} \hat{b}_1 = \text{sgn}\big(A_1 y_1 \quad &+ \quad \tfrac{1}{2}\left|A_2 y_2 - A_1 A_2 \rho_{12}\right| \\ &- \quad \tfrac{1}{2}\left|A_2 y_2 + A_1 A_2 \rho_{12}\right|\big) \\ \hat{b}_2 = \text{sgn}\big(A_2 y_2 \quad &+ \quad \tfrac{1}{2}\left|A_1 y_1 - A_1 A_2 \rho_{12}\right| \\ &- \quad \tfrac{1}{2}\left|A_1 y_1 + A_1 A_2 \rho_{12}\right|\big) \end{aligned}$$

Here again the set of K scalars $y_{1:K}$ is a sufficient statistic for $b_{1:K}$.

Remark. We notice that the decision rules of the conventional and jointly optimum receiver do not depend on the noise variance but not those of the individually optimum do.

2.5. Decision regions

Since $y_{1:K}$, is a sufficient statistic for b_k and for $b_{1:K}$ [7], the three receivers presented above can be compared by plotting the decision regions for each receiver in a K-dimensional space. This space is the projection of the infinite-dimensional space in which $\{y(t)\}_{0 \le t < T}$ lives onto the space spanned by the vector of the K correlated signatures $s_{1:K}$ where s_k stands for $\{s_k(t)\}_{0 \le t < T}$. In this finite space, it is possi-

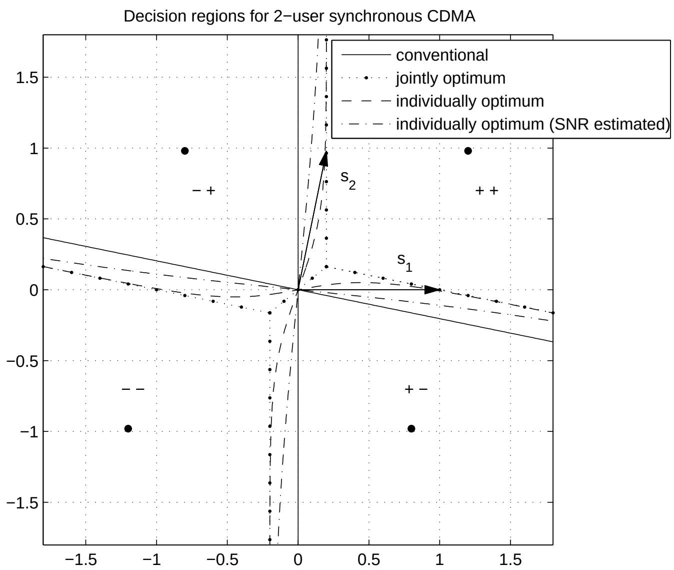

Fig. 1. Decision regions for the 2-user case.

$$f_\sigma(y_{2:K}) \triangleq \frac{\sigma^2}{2A_1}\log\frac{\displaystyle\sum_{e_2\in\{\pm1\}}\cdots\sum_{e_K\in\{\pm1\}}\exp\left[\frac{1}{\sigma^2}\sum_{k=2}^{K}\left(A_k e_k y_k - \sum_{\substack{j=1:\\ j\neq k, e_1=+1}}^{K} A_k A_j e_k e_j \rho_{jk}\right)\right]}{\displaystyle\sum_{e_2\in\{\pm1\}}\cdots\sum_{e_K\in\{\pm1\}}\exp\left[\frac{1}{\sigma^2}\sum_{k=2}^{K}\left(A_k e_k y_k - \sum_{\substack{j=1:\\ j\neq k, e_1=-1}}^{K} A_k A_j e_k e_j \rho_{jk}\right)\right]} \quad (5)$$

ble to draw the region decisions, i.e. the set of $y_{1:K}$ points (each point being the projection of one specific received signal $\{y(t)\}_{0\leq t<T}$) where a particular decision $\hat{b}_{1:K}$ is taken [8].

For the sake of clarity, we plot in figure 1 the decision regions for the 2-user case. The signal amplitudes are set to $1 = A_1 = A_2$ and the correlation ρ_{12} to 0.2. σ^2 equals .5. The vectors $(s_1(t), s_2(t))$ are represented: they span the 2-dimensional space but do not form an orthogonal basis since they are correlated. The four dots in the figure correspond to the four possible hypotheses for $(b_1, b_2) \in \{+1, -1\}^2$. The three first curves represent the boundaries of the decision regions for the three receivers seen above, $(+,+)$ meaning that in that region the decision $\widehat{(b_1, b_2)} = (+1, +1)$ is taken.

In the rest of the paper all the figures are for the 2-user case, with the simulation settings given above. However all the proofs are given for the general K-user case.

3. PERFORMANCE DEGRADATION WHEN THE INDIVIDUALLY OPTIMUM RECEIVER DOES NOT RECEIVE THE TRUE σ

In this section, we calculate the performance degradation when the individually optimum receiver is given a noise variance different from the true one. Without any loss of generality, we assume now that the user of interest is $k = 1$.

3.1. Preliminary properties of the individually optimum receiver

The noise being Gaussian, the symbols being equiprobable and $y_{1:K}$ being a sufficient statistic for b_k, the performance of the individually optimum receiver are completely determined by the likelihood function

$$p(y_{1:K} \mid b_{1:K}) = \alpha \exp\left[\frac{1}{\sigma^2}\sum_{k=1}^{K}\left(A_k b_k y_k - \sum_{j=1, j\neq k}^{K} A_k A_j b_k b_j \rho_{jk}\right)\right]$$

where α is a multiplicative coefficient independent of $y_{1:K}$ and $b_{1:K}$. It follows that the individually optimum receiver takes the decision

$$\begin{cases} \hat{b}_1 = +1 & \text{iff} \quad p_+(y_{1:K}, \sigma) > p_-(y_{1:K}, \sigma) \\ \hat{b}_1 = -1 & \text{otherwise} \end{cases}$$

where p_+ and p_- are defined as:

$$\begin{aligned} p_+(y_{1:K}, \sigma) &\triangleq \alpha\, e^{\frac{A_1 y_1}{\sigma^2}} \sum_{e_2\in\{\pm1\}}\cdots\sum_{e_K\in\{\pm1\}} \\ &\exp\left[\frac{1}{\sigma^2}\sum_{k=2}^{K}\left(A_k e_k y_k - \sum_{\substack{j=1:\\ j\neq k, e_1=+1}}^{K} A_k A_j e_k e_j \rho_{jk}\right)\right] \\ p_-(y_{1:K}, \sigma) &\triangleq \alpha\, e^{-\frac{A_1 y_1}{\sigma^2}} \sum_{e_2\in\{\pm1\}}\cdots\sum_{e_K\in\{\pm1\}} \\ &\exp\left[\frac{1}{\sigma^2}\sum_{k=2}^{K}\left(A_k e_k y_k - \sum_{\substack{j=1:\\ j\neq k, e_1=-1}}^{K} A_k A_j e_k e_j \rho_{jk}\right)\right] \end{aligned} \quad (6)$$

The decision rules can be rewritten as:

$$\begin{cases} \hat{b}_1 = +1 & \text{iff} \quad y_1 > f_\sigma(y_{2:K}) \\ \hat{b}_1 = -1 & \text{otherwise} \end{cases}$$

where $f_\sigma(y_{2:K})$ is defined in (5).

Property 1. [8] studies the limit behavior of the decision boundaries for the individually optimum receiver. First it shows that the minimum bit error rate decisions converge as $\sigma \to +\infty$ to those of the conventional detector:

$$\lim_{\sigma\to+\infty} \operatorname{sgn}\left(y_1 - f_\sigma(y_{2:K})\right) = \operatorname{sgn} y_1$$

Then it shows that the individually optimum decisions converge as $\sigma \to 0$ to the jointly optimum decisions.

Property 2. The performance of the individually optimum receiver are evaluated through the probability of error P_e for user 1. Conditioned on all possible realizations of the random variable $b_{2:K}$, this probability reads

$$\begin{aligned} P_e = &\sum_{b_2\in\{\pm1\}}\cdots\sum_{b_K\in\{\pm1\}} \frac{1}{2^K}\mathbb{P}(\hat{b}_1 = +1 \mid b_1 = -1, b_{2:K}) \\ &+ \sum_{b_2\in\{\pm1\}}\cdots\sum_{b_K\in\{\pm1\}} \frac{1}{2^K}\mathbb{P}(\hat{b}_1 = -1 \mid b_1 = +1, b_{2:K}) \end{aligned}$$

The symmetries of the channel (the channel is output symmetric) and of the receiver (the decision regions are central symmetrics) imply that:

$$\mathbb{P}(\hat{b}_1 = -1 \mid b_1 = 1, b_{2:K}) = \mathbb{P}(\hat{b}_1 = +1 \mid b_1 = -1, -b_{2:K})$$

Thus P_e may be written considering only the case where $b_1 = -1$:

$$P_e = \sum_{b_2 \in \{\pm 1\}} \cdots \sum_{b_K \in \{\pm 1\}} \frac{1}{2^{K-1}} \mathbb{P}(\hat{b}_1 = +1 \mid b_1 = -1, b_{2:K})$$

Using the notation p_- introduced in (6), this probability becomes:

$$P_e = \int_{\mathcal{A}} p_-(y_{1:K}, \sigma)\, dy_{1:K} \tag{7}$$

where $\mathcal{A}$ corresponds to the region where the decision $\hat{b}_1 = +1$ is taken i.e.

$$\mathcal{A} = \{y_{1:K} : y_1 > f_\sigma(y_{2:K})\}$$

3.2. Performance at the limits

We now assume that the receiver is given a noise variance σ_e different from the true one σ. The fourth curve in Figure 1 shows the evolution of the decisions boundary for the individually optimum receiver when the noise variance is estimated ($\sigma_e^2 = 2$).

Proposition 1. The probability of error of the individually optimum receiver under noise variance mismatch converges to the one of the conventional detector as the estimated noise variance σ_e tends to $+\infty$ and to the one of the jointly optimum receiver as σ_e tends to 0.

Proof. It follows directly from the convergence of the decision regions (see Property 1 in Section 3.1) and from the definition of the probability of error (7). □

3.3. Monotonic increase and decrease of the probability of error

Having determined the limit behavior of the probability of error under noise variance mismatch, we would like to further investigate the behavior of the receiver in the range of all possible estimated noise variances.

Proposition 2. The probability of error of the individually optimum receiver under noise variance mismatch is a piecewise monotonic function of the estimated noise variance σ_e. It decreases monotonically from $\sigma_e = 0$ to $\sigma_e = \sigma$, and then increases from $\sigma_e = \sigma$ to $\sigma_e \to +\infty$, where the individually optimum receiver behaves as the conventional detector.

Proof. Consider now two receivers that estimate the noise variance to σ_e and to σ'_e. Without loss of generality we assume that $\sigma_e > \sigma'_e$. We define two sets $\mathcal{A}$ and $\mathcal{A}'$ which correspond to the integration domains:

$$\begin{aligned} \mathcal{A} &= \{y_{1:K} : y_1 > f_{\sigma_e}(y_{2:K})\} \\ \mathcal{A}' &= \{y_{1:K} : y_1 > f_{\sigma'_e}(y_{2:K})\} \end{aligned}$$

Then we introduce the difference between both probabilities of error:

$$\begin{aligned} \Delta P_e &= P_e - P'_e \\ &= \int_{\mathcal{A}} p_-(y_{1:K}, \sigma)\, dy_{1:K} - \int_{\mathcal{A}'} p_-(y_{1:K}, \sigma)\, dy_{1:K} \end{aligned}$$

It is important to note that the densities depend on the observation and thus on the true variance whereas the decision regions depend on the receiver and therefore on the estimated variance. We now use the short-hand notation for ΔP_e:

$$\Delta P_e = \int_{\mathcal{A}} p_- - \int_{\mathcal{A}'} p_-$$

Proof part 1: partitioning the space. The domains $\mathcal{A}$ and $\mathcal{A}'$ overlap. To determine the non-overlapping areas, we introduce:

$$\mathcal{B} = \{y_{1:k} | f_{\sigma_e}(y_{2:k}) < f_{\sigma'_e}(y_{2:k})\}.$$

It can be easily checked that on $\mathcal{B}$, $\mathcal{A}'$ is included in $\mathcal{A}$ ($\mathcal{A}' \cap \mathcal{B} \subseteq \mathcal{A} \cap \mathcal{B}$) whereas $\mathcal{A}$ is included in $\mathcal{A}'$ on $\mathcal{B}^c$. It follows that

$$\begin{aligned} \Delta P_e &= \int_{\mathcal{A} \cap \mathcal{B}} p_- - \int_{\mathcal{A}' \cap \mathcal{B}} p_- + \int_{\mathcal{A} \cap \mathcal{B}^c} p_- - \int_{\mathcal{A}' \cap \mathcal{B}^c} p_- \\ &= \int_{(\mathcal{A} \setminus \mathcal{A}') \cap \mathcal{B}} p_- - \int_{(\mathcal{A}' \setminus \mathcal{A}) \cap \mathcal{B}^c} p_- \end{aligned}$$

Proof part 2: reducing the number of integrals. This quantity can be further simplified noticing that:

$$\begin{aligned} \int_{(\mathcal{A}' \setminus \mathcal{A}) \cap \mathcal{B}^c} p_-(y_{1:k}) dy_{1:k} &= \int_{-(\mathcal{A}' \setminus \mathcal{A}) \cap \mathcal{B}^c} p_-(-y_{1:k}) dy_{1:k} \\ &= \int_{(\mathcal{A} \setminus \mathcal{A}') \cap \mathcal{B}} p_-(-y_{1:k}) dy_{1:k} \\ &= \int_{(\mathcal{A} \setminus \mathcal{A}') \cap \mathcal{B}} p_+(y_{1:k}) dy_{1:k} \end{aligned}$$

The first equality is obtained by the change of variable $y_{1:k} \to z_{1:k} = -y_{1:k}$. The second equality is due to the fact that, for a given σ_e, $f_{\sigma_e}(y_{2:k})$ is an odd function of $y_{2:k}$, $f_{\sigma_e}(-y_{2:k}) = -f_{\sigma_e}(y_{2:k})$ (and this is immediate from the definition of the function (5)). The last equality follows from the symmetry of the channel. We get

$$\Delta P_e = \int_{(\mathcal{A} \setminus \mathcal{A}') \cap \mathcal{B}} p_- - p_+$$

Proof part 3. Finally we show that the function $\sigma_e \to P_e(\sigma_e)$ decreases on $[0, \sigma]$ and increases on $[\sigma, \infty[$.

First consider the interval: $(\sigma_e, \sigma'_e) \in [0, \sigma]$.
Since $\sigma'_e < \sigma$, for all $y_{1:k}$ belonging to $(\mathcal{A} \backslash \mathcal{A}') \cap \mathcal{B}$, we have $f_\sigma(y_{2:k}) < f_{\sigma'_e}(y_{2:k})$. And by definition of $\mathcal{A}$, $f_{\sigma'_e}(y_{2:k}) < y_1$. It follows that on $(\mathcal{A} \setminus \mathcal{A}') \cap \mathcal{B}$, $f_\sigma(y_{2:k}) < y_1$ which is equivalent to $p_-(y_{1:k}) - p_+(y_{1:k}) < 0$ by definition of $f_\sigma(y_{2:k})$ (5). Thus ΔP_e is the integral of a negative fonction, so ΔP_e is negative.

A similar argument holds for the interval $[\sigma_e, \infty[$. □

Proposition 2 allows us to derive upper and lower bounds of the BER for the individually optimum receiver under noise variance mismatch. In fact, as a corollary we have that

Corollary 2. The error probability under noise variance mismatch is lower bounded by the one of individually optimum receiver under perfect variance knowledge (this is clear since this receiver achieves minimum error probability) and upper bounded by the maximum between the BER of the conventional detector and the BER of the jointly optimum receiver.

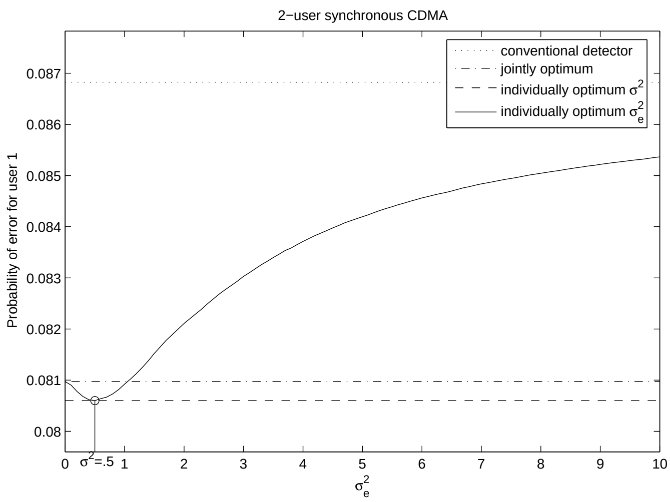

Fig. 2. Effect of noise variance mismatch mismatch on the individually optimum receiver

Figure 2 illustrates proposition 2 and its corollary for the 2-user case. As it is analytically proven above, the error probability of the individually optimum receiver is a piecewise monotonic function of the estimated noise variance. In the case of positive signal to noise ratio for user 1 ($A_1 > \sigma^2$), the BER of the conventional detector is greater than the BER of the jointly optimum receiver, which is known to be close to the BER of the individually optimum receiver [1, page 814]. This justifies a well known result in the community that underestimation of the noise variance is less detrimental than overestimation.

4. CONCLUSION

In this paper we have studied the behavior of the individually optimum receiver when it has partial knowledge of the noise variance. We have shown that the error probability is piecewise monotonic.

It follows that the error probability under noise variance mismatch is lower bounded by the one of individually optimum receiver under perfect variance knowledge (this is clear since this receiver achieves minimum error probability) and upper bounded by the maximum between the BER of the conventional detector and of the BER of the jointly optimum receiver. This shows that underestimation of the noise variance is less detrimental than overestimation.

5. REFERENCES

[1] S. Benedetto and E. Biglieri, *Principles of Digital Transmission with Wireless Applications,* NewYork: Kluwer/Plenum, 1999.

[2] C. Berrou, A. Glavieux, P. Thitimajshima, “Near Shannon limit error-correcting and decoding: Turbo codes,” *Proc. ICC,* Geneva, pp. 1064-1070, May 1993.

[3] R. Matzner, “An SNR estimation algorithm for complex baseband signals using higher order statistics,” *Facta Universitatis, Series: Electronics and Energetics,* Vol. 6, No. 1, pp. 41–52, 1993.

[4] D.R. Pauluzzi and N.C. Beaulieu. “A comparison of SNR estimation techniques for the AWGN channel.” *IEEE Transactions on Communications,* Vol. 48, (10), pp. 1681–1691, Oct. 2000.

[5] H. Shin and J. H. Lee, “A channel reliability estimation for turbo decoding in rayleigh fading channels with imperfect channel estimates,” *IEEE Communications Letters*, pp. 503- 505, Nov. 2002.

[6] T.A. Summers and S.G. Wilson, “SNR mismatch and online estimation in turbo decoding,” *IEEE Transactions on Communications,* Vol. 46, (4), pp. 421-423, April. 1998.

[7] S. Verdú, “Minimum Probability of error for asynchronous Gaussian multiple-access channels,” *IEEE Trans. on Inf. Th.,* Vol. 32, pp. 85-96, Jan. 1986.

[8] S. Verdú, *Multiuser detection,* Cambridge Univ. Press, 1998.

[9] A. Worm, P. Hoeher and N. Wehn. “Turbo-decoding without SNR estimation,” *IEEE Communications Letters*, pp. 193-195, June 2000.