AN ITERATIVE SOFT DECISION DIRECTED LINEAR TIMING ESTIMATOR

C. Herzet, V. Ramon, L. Vandendorpe∗

Université catholique de Louvain
Communications and Remote Sensing Laboratory
2 place du Levant, 1348 Louvain-la-Neuve, Belgium
{herzet,ramon,vdd}@tele.ucl.ac.be

M. Moeneclaey

Ghent University
DIGCOM Research Group, TELIN Department
41 St. Pietersnieuwstraat, 9000 Gent, Belgium
mm@telin.rug.ac.be

ABSTRACT

This paper addresses the issue of estimating the symbol timing in turbo receivers. We propose a linear timing estimator which takes benefit from the soft information delivered by the turbo system at each iteration to compute the timing estimate. This estimator is then compared to a synchronizer based on the expectation-maximization (EM) algorithm [1]. Performance of the proposed synchronizer is illustrated by simulation results and proves better than that of the EM-based counterpart regarding both the convergence speed and range.

1. INTRODUCTION

Since the discovery of turbo codes by Berrou and Glavieux [2], the so-called turbo principle has been applied to various operations performed by a receiver : joint demodulation and decoding, joint equalization and decoding, ... Such receivers have become famous due to the surprisingly low bit-error rate (BER) they can reach at very low signal-to-noise ratios (SNR). However, in order to benefit from this performance turbo receivers have to know some parameters like the carrier phase, the timing offset, ... However, in such iterative systems the low SNR operating point may make the synchronization hard to perform. Moreover, as turbo receivers are designed to reach very low BER a small synchronization error may lead to important degradation and quite accurate estimates are therefore required.

In this context one may be interested in exploiting the soft information on bits or symbols delivered by such iterative receivers in order to help the synchronizer. This approach, which is often referred to as turbo-synchronization, has already been studied in several contributions. In particular, [3] presents a unifying framework for ML synchronization in turbo systems by means of the expectation-maximization (EM) algorithm [4].

∗This work has been supported by the Interuniversity Attraction Poles Programme - Belgium State - Federal Office for Scientific, Technical and Cultural Affairs. V. Ramon would like to thank the belgian NSF for its financial support

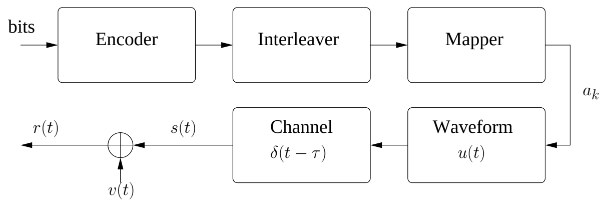

Fig. 1. Transmitter.

The present paper, we will focus on the particular issue of timing synchronization in turbo demodulation receiver. In particular we will derive a linear timing estimator which takes benefit from the soft information delivered by the turbo system at each iteration. The proposed estimator will then be discussed and compared to an EM-based synchronizer proposed in [1].

2. SYSTEM MODEL

We will focus on a bit-interleaved coded modulation (BICM) scheme. The transmitter (Fig. 1) is then made up of a binary convolutional encoder and a constellation mapper separated by a bit interleaver. In the baseband formalism, the signal at the transmitter output may then be written as

$$s(t) = \sum_k a_k \, u(t - kT), \tag{1}$$

where a_k's are complex symbols belonging to constellation alphabet $\mathcal{A}$, T is the symbol period and $u(t)$ is a unit-energy square-root raised-cosine pulse with roll-off α. Assuming that $s(t)$ is sent over an AWGN channel introducing a time delay τ, the received signal is

$$r(t) = \sum_k a_k \, u(t - kT - \tau) + v(t), \tag{2}$$

where $v(t)$ is the complex envelope of an additive white gaussian noise with passband two-sided power spectral density $N_0/2$. At the receiver (Fig. 2), after anti-aliasing filter-

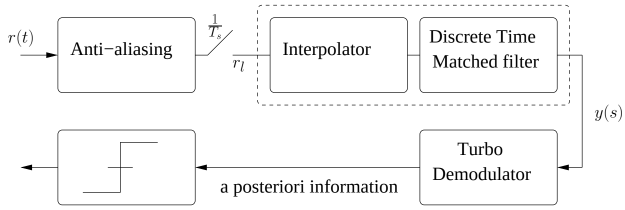

Fig. 2. Receiver.

ing, $r(t)$ is sampled at a rate of $1/T_s$ (with $T_s < T/(1+\alpha)$) leading to samples

$$r_l \triangleq r(lT_s) = \sum_k a_k\, u(lT_s - kT - \tau) + v_l, \quad (3)$$

where v_l is a white gaussian noise with variance $2N_0/T_s$. Samples r_l are passed through a discrete-time matched filter i.e.

$$\begin{aligned} y(s) &= \sum_l r_l\, u^*(lT_s - s) \\ &= \sum_k a_k\, x(s - kT - \tau) + \sum_l v_l\, u^*(lT_s - s), \quad (4) \end{aligned}$$

where s denotes the time at which the matched filter output is computed and $x(t)$ is a raised-cosine filter with roll-off α. Finally, we assume that statistics $y(s)$ are processed in a turbo demodulator. Such a device introduced in [5] performs iterative joint demodulation and decoding through the exchange of extrinsic information between a soft-input soft-ouput (SISO) demodulator and a SISO decoder.

3. EM ALGORITHM AND SYNCHRONIZATION

The expectation-maximization (EM) algorithm [4] is a method which enables to compute iteratively the maximum-likelihood (ML) estimate of a parameter. In [1], we applied this algorithm in a turbo receiver for timing synchronization of a linearly-modulated signal. The proposed implementation of the maximization step was based on a steepest-descent method. At iteration n, the correction performed on the current timing estimate $\hat{\tau}^{(n)}$ was then equal to

$$\begin{aligned} \hat{\epsilon}^{(n)} &\triangleq \hat{\tau}^{(n+1)} - \hat{\tau}^{(n)} \\ &= \beta \sum_k |\eta_k^{(n)}| \\ &\times \mathrm{Re}\Big\{ e^{-j\arg(\eta_k^{(n)})} \Big(\dot{y}(kT + \hat{\tau}^{(n)}) - \sum_{k'} \eta_{k'}^{(n)} \dot{x}_{k-k'} \Big) \Big\}, \end{aligned} \quad (5)$$

where $\dot{y}(kT + \hat{\tau}^{(n)})$ is the time derivative of the matched filter output evaluated at $kT + \hat{\tau}^{(n)}$, $\dot{x}_{k-k'}$ represents the time derivative of a raised-cosine filter evaluated at $kT - k'T$, β is a weighting factor and $\eta_k^{(n)}$ denotes symbol a_k a posteriori average values computed from the soft information delivered by the turbo demodulator at iteration n. Note that the terms subtracted from the matched filter output derivative in (5) cancel out two by two and one may rewrite this equation as

$$\hat{\epsilon}^{(n)} = \beta \sum_k |\eta_k^{(n)}|\, \mathrm{Re}\Big\{ e^{-j\arg(\eta_k^{(n)})}\, \dot{y}(kT + \hat{\tau}^{(n)}) \Big\}. \quad (6)$$

The interest of the full expression (5) will however appear in section 6. One may show that $\mathrm{Re}\{ e^{-j\arg(\eta_k^{(n)})}\, \dot{y}(kT + \hat{\tau}^{(n)})\}$ actually represents the component of $\dot{y}(kT + \hat{\tau}^{(n)})$ which is in phase with $\eta_k^{(n)}$. Correction $\hat{\epsilon}^{(n)}$ may then be viewed as a weighted sum of the projections of the matched filter output derivatives on the a posteriori average values $\eta_k^{(n)}$. In other words, correction $\hat{\epsilon}^{(n)}$ is simply the result of a linear operation on a subset of the available observations. Now, the components in quadrature also contain information about the timing error and one may wonder why not to benefit from them in the computation of $\hat{\epsilon}^{(n)}$. In the same way, since (6) reduces to a linear operation, why not to implement an efficient and unbiased linear estimator to compute the new timing estimate ? These questions will be addressed in the sequel of this paper.

4. UNBIASED AND MINIMUM-VARIANCE LINEAR ESTIMATOR

Consider the following linear model

$$\mathbf{z} = \mathbf{H}\,\boldsymbol{\Theta} + \mathbf{n}, \quad (7)$$

where $\mathbf{H}$ is a random matrix, $\boldsymbol{\Theta}$ is a deterministic but unknown vector of parameters to be estimated and $\mathbf{n}$ is a zero-mean noise vector with covariance matrix $\mathbf{R_n}$. From the observation vector $\mathbf{z}$ we would like to construct a linear estimator of $\boldsymbol{\Theta}$ which is both unbiased and efficient i.e.

$$\hat{\boldsymbol{\Theta}} = \mathbf{F}\,\mathbf{z}, \quad (8)$$

with matrix $\mathbf{F}$ such that

$$\mathrm{E}[\hat{\boldsymbol{\Theta}}] = \boldsymbol{\Theta}, \quad (9)$$

and

$$\mathrm{E}[(\hat{\Theta}_i - \Theta_i)^2] \text{ is minimum } \forall\, i, \quad (10)$$

where Θ_i and $\hat{\Theta}_i$ respectively denote the i^{th} element of $\boldsymbol{\Theta}$ and $\hat{\boldsymbol{\Theta}}$. By analogy with what is done in [6] for a non-random matrix $\mathbf{H}$, one can show that the matrix $\mathbf{F}$ which solves this constrained minimization problem is given by

$$\mathbf{F} = \Big(\mathrm{E}[\mathbf{H}]^T\, \mathbf{R_n}^{-1} \mathrm{E}[\mathbf{H}] \Big)^{-1} \mathrm{E}[\mathbf{H}]^T\, \mathbf{R_n}^{-1}, \quad (11)$$

where $\mathrm{E}[\cdot]$ represents the mathematical expectation operator and $\mathbf{R_n}$ is the covariance matrix of the noise vector $\mathbf{n}$.

5. SOFT LINEAR TIMING ESTIMATOR

Consider the time derivative of the matched filter output

$$\dot{y}(kT+\hat{\tau}^{(n)}) = \sum_{k'} a_{k'}\,\dot{x}(kT-k'T-\epsilon^{(n)}) + e(k), \quad (12)$$

where $\epsilon^{(n)} \triangleq \tau - \hat{\tau}^{(n)}$ represents the timing estimation error at iteration n and $e(k)$ is the contribution of the background noise to the derivative of the matched filter output at time $kT+\hat{\tau}^{(n)}$. Let us now expand $\dot{y}(kT+\hat{\tau}^{(n)})$ into a sum of its components in phase and in quadrature with $\eta_k^{(n)}$ i.e.

$$\begin{aligned}\dot{y}(kT+\hat{\tau}^{(n)}) &= \dot{y}_I(kT+\hat{\tau}^{(n)})\, e^{j\arg(\eta_k^{(n)})} \\ &\quad + \dot{y}_Q(kT+\hat{\tau}^{(n)})\, e^{j\frac{\pi}{2}+j\arg(\eta_k^{(n)})}, \quad (13)\end{aligned}$$

with

$$\dot{y}_I(kT+\hat{\tau}^{(n)}) = \mathrm{Re}\Big\{\, e^{-j\arg(\eta_k^{(n)})}\,\dot{y}(kT+\hat{\tau}^{(n)})\Big\} \quad (14)$$

$$\dot{y}_Q(kT+\hat{\tau}^{(n)}) = \mathrm{Im}\Big\{\, e^{-j\arg(\eta_k^{(n)})}\,\dot{y}(kT+\hat{\tau}^{(n)})\Big\}. \quad (15)$$

Together components $\dot{y}_I(kT+\hat{\tau}^{(n)})$ and $\dot{y}_Q(kT+\hat{\tau}^{(n)})$ actually contain the same information as $\dot{y}(kT+\hat{\tau}^{(n)})$ and we may use them to estimate the timing error. Let us then linearize (14) around $\epsilon = 0$, we have

$$\dot{y}_I(kT+\hat{\tau}^{(n)}) = \ h_I(k)\,\epsilon + w_I(k) + e_I(k), \quad (16)$$

where

$$h_I(k) \triangleq -\mathrm{Re}\Big\{e^{-j\arg(\eta_k^{(n)})}\sum_{k'} a_{k'}\,\ddot{x}(kT-k'T)\Big\} \quad (17)$$

$$w_I(k) \triangleq \mathrm{Re}\Big\{e^{-j\arg(\eta_k^{(n)})}\sum_{k'} a_{k'}\,\dot{x}(kT-k'T)\Big\} \quad (18)$$

$$e_I(k) \triangleq \mathrm{Re}\Big\{e^{-j\arg(\eta_k^{(n)})}e(k)\Big\}. \quad (19)$$

In other words, $h_I(k)$ represents the sensitivity of $\dot{y}_I(kT+\hat{\tau}^{(n)})$ to a timing error, $w_I(k)$ is the perturbation introduced by the presence of adjacent symbols (self-noise) and $e_I(k)$ is the contribution of the background noise. Expressions similar to (16), (17), (18) and (19) may be found for $\dot{y}_Q(kT+\hat{\tau}^{(n)})$. If we then define by $\dot{\mathbf{y}}$ the vector of components $\dot{y}_I(kT+\hat{\tau}^{(n)})$ and $\dot{y}_Q(kT+\hat{\tau}^{(n)})$, by $\mathbf{h}$ the vector of sensitivities $h_I(k)$ and $h_Q(k)$, by $\mathbf{w}$ the vector of self-noise components $w_I(k)$ and $w_Q(k)$, and by $\mathbf{e}$ the vector of background noise contributions $e_I(k)$ and $e_Q(k)$, we have

$$\dot{\mathbf{y}} = \mathbf{h}\,\epsilon + \mathbf{w} + \mathbf{e}. \quad (20)$$

Assume we would like to construct an unbiased and minimum variance estimate of ϵ from the observation vector $\dot{\mathbf{y}}$. Using model (20) we have from section 4

$$\hat{\epsilon}^{(n)} = \Big(\mathrm{E}[\mathbf{h}]^T\mathbf{R}^{-1}\mathrm{E}[\mathbf{h}]\Big)^{-1}\mathrm{E}[\mathbf{h}]^T\mathbf{R}^{-1}(\dot{\mathbf{y}} - \mathrm{E}[\mathbf{w}] - \mathrm{E}[\mathbf{e}]) \quad (21)$$

where

$$\mathbf{R} = \mathbf{R_w} + \mathbf{R_e},$$

since the contribution of the self-noise and of the background noise are independent. Note that the expectation operators which appear in (21) should be over the a priori probability distribution of transmitted symbols a_k. In this paper however, we would like to benefit from the symbol a posteriori information delivered by the turbo receiver in order to perform the synchronization. Consequently, a priori expectations $\mathrm{E}[\cdot]$ will be approximated by expectations using the soft a posteriori information provided by the SISO decoder. Note also that (21) has a high computation cost since it requires the inversion of covariance matrix $\mathbf{R}$. Therefore, in order to overcome this problem, we will assume that matrix $\mathbf{R}$ is diagonal. Estimator (21) then reduces to

$$\begin{aligned}\hat{\epsilon}^{(n)} &= \beta'\sum_k \frac{\mathrm{E}[h_I(k)]}{\sigma^2_{w_I(k)}+\sigma^2_{e_I(k)}} \\ &\quad\times \mathrm{Re}\Big\{\, e^{-j\arg(\eta_k^{(n)})}\,\Big(\dot{y}(kT+\hat{\tau}^{(n)}) - \sum_{k'}\eta_{k'}^{(n)}\dot{x}_{k-k'}\Big)\Big\} \\ &\quad + \beta'\sum_k \frac{\mathrm{E}[h_Q(k)]}{\sigma^2_{w_Q(k)}+\sigma^2_{e_Q(k)}} \\ &\quad\times \mathrm{Im}\Big\{\, e^{-j\arg(\eta_k^{(n)})}\,\Big(\dot{y}(kT+\hat{\tau}^{(n)}) - \sum_{k'}\eta_{k'}^{(n)}\dot{x}_{k-k'}\Big)\Big\},\end{aligned} \quad (22)$$

with

$$\beta' = \Big(\sum_k \frac{\mathrm{E}[h_I(k)]}{\sigma^2_{w_I(k)}+\sigma^2_{e_I(k)}} + \sum_k \frac{\mathrm{E}[h_Q(k)]}{\sigma^2_{w_Q(k)}+\sigma^2_{e_Q(k)}}\Big)^{-1},$$

and where

$$\mathrm{E}[h_I(k)] = -\,\mathrm{Re}\Big\{e^{-j\arg(\eta_k^{(n)})}\sum_{k'}\eta_{k'}^{(n)}\,\ddot{x}(kT-k'T)\Big\} \quad (23)$$

$$\sigma^2_{e_I(k)} = -\,N_0\ \ \ddot{x}(0) \quad (24)$$

$$\sigma^2_{w_I(k)} = \sum_l\ \dot{x}^2(kT-lT)\ \sigma^2_{a_I(l)}, \quad (25)$$

and $\sigma^2_{a_I(l)}$ is the a posteriori variance of the component of symbol a_l which is in phase with η_k. Likewise, $\sigma^2_{e_I(k)}$ and $\sigma^2_{w_I(k)}$ respectively represent the variance of the background noise and the self noise components which are in phase with $\eta_k^{(n)}$. Expressions similar to (23), (24) and (25) may be found for $\mathrm{E}[h_Q(k)]$, $\sigma^2_{e_Q(k)}$ and $\sigma^2_{w_Q(k)}$. In the sequel, the estimator (22) will be referred to as SLITE (Soft LInear Timing Estimator).

6. SLITE AND EM ESTIMATOR COMPARISON

In this section we will compare the SLITE with the EM estimator. We see from (5) and (22) that both estimators use the components of a modified version of the matched filter output derivative which are in phase with $\eta_k^{(n)}$, i.e.

$$\mathrm{Re}\Big\{ e^{-j\arg(\eta_k^{(n)})} \Big(\dot{y}(kT+\hat{\tau}^{(n)}) - \sum_{k'} \eta_{k'}^{(n)} \dot{x}_{k-k'}\Big)\Big\}.$$

However, the SLITE estimator also uses the in quadrature components. These components mainly contain information about the timing error coming from adjacent symbols. Their contribution will therefore become more important as the actual timing error ϵ increases. Note that the term which is subtracted from $\dot{y}(kT+\hat{\tau}^{(n)})$ actually corresponds to a soft cancellation of the self noise. Although this soft operation appears in both the EM and the SLITE approaches, the subtracted soft terms will cancel out two by two in the EM case when we perform the summation over all the observations. This is actually due to the particular coefficients used to weight the observations, namely $|\eta_k^{(n)}|$. Notice also that in the EM estimator the weighting of an observation is only a function of the confidence we have in the corresponding symbol. On the contrary, the SLITE proposes to weight each observation by a coefficient proportional to the mean of the sensitivity and inversely proportional to variances of the noises (self and background) which affect the matched filter output derivative. Consequently, as we see from (23) and (25), the coefficients in the SLITE are weighted sums of the confidence we have in all the symbols.

Let us now consider the performance of the estimators after a large number of iterations. If the turbo system converges i.e. manages to decrease the BER a posteriori expectations $\eta_k^{(n)}$ will become closer and closer to the actual transmitted symbol. Consequently, (5) reduces to the maximization of a data-aided likelihood function via a steepest-descent method. The EM estimator will then be asymptotically unbiased and efficient. Concerning the SLITE, we see from (25) that the self-noise variances which affect the coefficient reduce to zero. Therefore, (22) becomes

$$\begin{aligned}\hat{\epsilon}^{(n)} = &\sum_k \frac{h_I(k)}{h_I(k)+h_Q(k)} \\ &\times \mathrm{Re}\Big\{ e^{-j\arg(a_k)} \Big(\dot{y}(kT+\hat{\tau}^{(n)}) - \sum_{k'} a_{k'}^{(n)} \dot{x}_{k-k'}\Big)\Big\} \\ &+ \sum_k \frac{h_Q(k)}{h_I(k)+h_Q(k)} \\ &\times \mathrm{Im}\Big\{ e^{-j\arg(a_k^{(n)})} \Big(\dot{y}(kT+\hat{\tau}^{(n)}) - \sum_{k'} a_{k'}^{(n)} \dot{x}_{k-k'}\Big)\Big\},\end{aligned} \tag{26}$$

which is the equation of a least-square (LS) estimator. This is a consequence of our diagonal approximation of covariance matrix $\mathbf{R}$. One may show [6] that if the noise which affects the observations is not white the LS estimator is not efficient. We should therefore observe a degradation with respect to the Cramer-Rao bound in the mean squared error of the SLITE.

7. SIMULATION RESULTS

In this section the performance of the SLITE will be studied and compared to that of the EM estimator through simulation results. At the transmitter, we consider a rate-$\frac{1}{2}$ non-systematic convolutional encoder with polynomial generators $(g_1, g_2) = (5, 7)_8$ and use 16-QAM modulation. A mapping proposed by ten Brink in [5] and referred to as medium unconditioned bit-wise mutual information mapping is used. The pulse waveform is a square-root raised cosine with roll-off 0.2. The interleaver is totally random and a different permutation is used at each frame. At the receiver, the matched filter outputs are computed from the samples thanks to an interpolator and a discrete-time matched filter. The considered interpolator [7] is designed in order to minimize, on the bandwidth of the useful signal $s(t)$, the quadratic error between the ideal interpolator frequency response and the interpolator frequency response. The number of taps of the interpolator is set to 21. The timing estimate is initialized to 0 at the first iteration. At each turbo iteration, a new timing estimation is computed using the posterior soft information delivered by the turbo demodulator. The matched filter output is then corrected according to this new estimate and a new turbo iteration is performed. Note that the matched filter output derivatives which appear in (5) and (22) are simply computed via an early-late approximation with time separation $T/16$.

Fig. 3 represents the SLITE (in red) and EM estimator (in blue) performance (mean and mean squared error) as well as the corresponding bit error rate (BER) and frame error rate (FER). The simulation has been run for frames of 500 16-QAM symbols and 12 turbo iterations have been performed. An E_b/N_0-ratio of 4dB and a normalized timing offset τ/T ranging from 0 to 0.5 have been considered. The dashed curves represent the data-aided (DA) Cramer-Rao bound in Fig. 3.b and the BER reached by a perfectly synchronized system in Fig. 3.c.

In Fig. 3.a we can see that the mean of both estimators is the same after a large number of iterations. Note however that the SLITE exhibits a faster convergence towards the actual timing offset τ than the EM method. As expected this difference between the two estimators increases when the timing error increases since then the SLITE fully benefits from its "extended" observation set (namely both in-phase and in-quadrature components). We observe the same behavior for the mean squared error in Fig. 3.b. Note however that, this time, the final performance is no longer the same. Indeed, whereas the EM estimator reaches the Cramer-Rao bound the SLITE remains slightly above. As mentioned in section 6, this behavior is actually due to the diagonal ap-

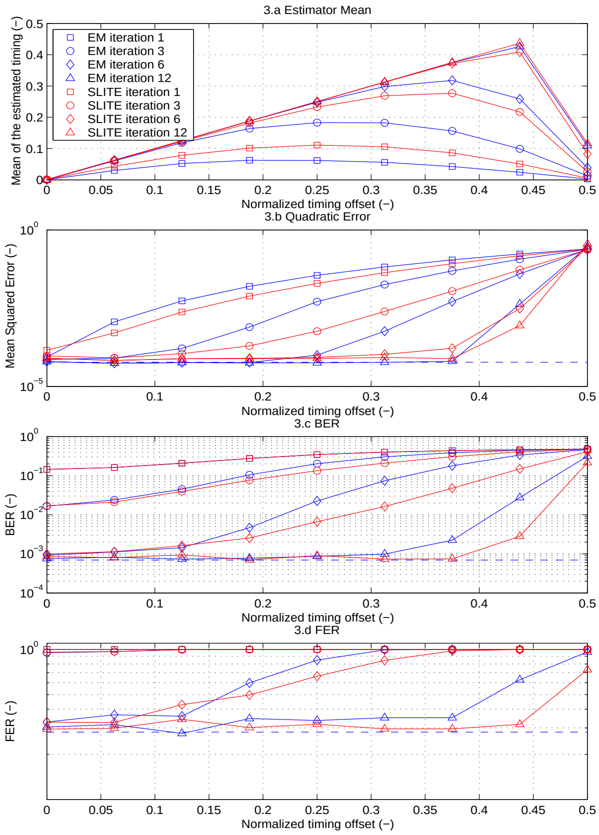

Fig. 3. Estimator mean, mean squared error, BER and FER for ten Brink's mapping at $E_b/N_0 = 4$dB versus normalized timing offset

proximation for matrix $\mathbf{R}$. Notice however that the gap between the two methods is very small. Therefore, as see at Fig. 3.c and 3.d the estimate provided by both the estimator is accurate enough and the reached BER and FER do not suffer from any degradation with respect to the performance of a perfectly synchronized system.

Fig. 4 represents the SLITE (in red) and EM estimator (in blue) mean as well as the frame error rate (FER) reached by the system for an anti-Gray mapping. The simulation has been run for frames of 300 16-QAM symbols at an $E_b/N_0 = 4$dB. Note that anti-Gray mapping enables to reach lower BER but is also more sensitive to timing errors. Therefore, for a high timing offset the soft information delivered by the turbo demodulator will be of lower quality. We may notice from Fig. 4.a and 4.b that SLITE makes a better use of the available information and is therefore less sensitive than the EM approach to the degradation of the information delivered by the turbo system.

Conclusion

In this paper we have presented a new soft decision directed iterative linear estimator of symbol timing. At each iteration

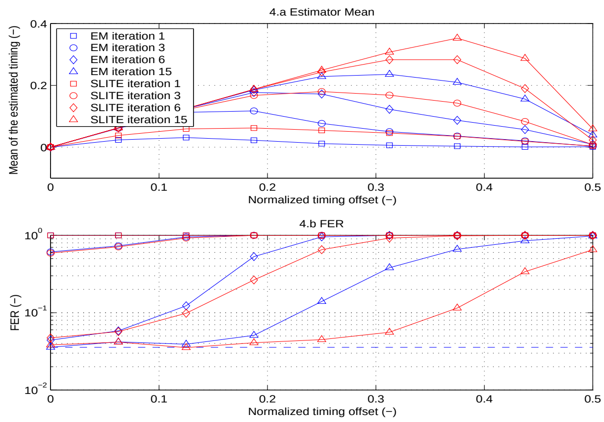

Fig. 4. Estimator mean and FER for anti-Gray mapping at E_b/N_0=4.5dB versus normalized timing offset

of a turbo process, this estimator computes a new timing estimate by taking into account the soft information delivered by the turbo system. The proposed estimator have been compared to an EM-based method and proves to have both a higher range and a faster speed of convergence.

A. BEST LINEAR UNBIASED ESTIMATOR

A.1. Problem Statement

Consider the following linear model

$$\mathbf{z} = \mathbf{H}\,\mathbf{\Theta} + \mathbf{n}, \tag{27}$$

where $\mathbf{H}$, $\mathbf{\Theta}$ and $\mathbf{n}$ are mutually independent real random elements. From the observation vector $\mathbf{z}$ we would like to construct a linear estimator of $\mathbf{\Theta}$ which is both unbiased and efficient by design i.e.

$$\hat{\mathbf{\Theta}} = \mathbf{F}\,\mathbf{z}, \tag{28}$$

with matrix $\mathbf{F}$ such that

$$\mathrm{E}[\hat{\mathbf{\Theta}}] = \mathbf{\Theta}, \tag{29}$$

and

$$\mathrm{E}[(\hat{\Theta}_i - \Theta_i)^2] \text{ is minimum } \forall\, i, \tag{30}$$

where Θ_i and $\hat{\Theta}_i$ respectively denote the i^{th} element of $\mathbf{\Theta}$ and $\hat{\mathbf{\Theta}}$.

A.2. Unbiasedness Constraint

Replacing $\hat{\mathbf{\Theta}}$ by its expression in (29) we find that the unbiasedness constraint implies for matrix $\mathbf{F}$ to be such that

$$\mathbf{F}\ \mathrm{E}[\mathbf{H}] = \mathbf{I}, \tag{31}$$

where $\mathbf{I}$ represents a unitary matrix. Rewriting $\mathbf{F}$ in a way more convenient for the subsequent derivations

$$\mathbf{F} = \begin{pmatrix} \mathbf{f}_1^{'} \\ \mathbf{f}_2^{'} \\ \vdots \\ \mathbf{f}_n^{'} \end{pmatrix}, \tag{32}$$

the constraint (31) may also be written as

$$\mathrm{E}[\mathbf{H}'] \begin{pmatrix} \mathbf{f}_1 & \mathbf{f}_2 & \dots & \mathbf{f}_n \end{pmatrix} = \begin{pmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \dots & \mathbf{e}_n \end{pmatrix}, \tag{33}$$

where $\mathbf{e}_i$ is a unit vector,

$$\mathbf{e}_i = \mathrm{col}\begin{pmatrix} 0, & \dots & 0, & 1, & 0, & \dots & 0, \end{pmatrix}, \tag{34}$$

in which the nonzero element occurs in the ith position.

A.3. Mean Squared Error Expression

Knowing that the estimation of Θ_i is given by

$$\hat{\Theta}_i = \mathbf{f}_i^{'}\,\mathbf{z}, \tag{35}$$

the mean squared error associated with the estimation of this parameter may then be written as

$$\mathrm{E}\big[(\Theta_i - \hat{\Theta}_i)^2\big] = \mathrm{E}\big[(\Theta_i - \mathbf{f}_i^{'}\,\mathbf{z})^2\big] \tag{36}$$

Developing (36), we get

$$\mathrm{E}\big[(\Theta_i - \hat{\Theta}_i)^2\big] = \mathrm{E}\big[\Theta_i^2 - \mathbf{\Theta}^{'}\mathbf{H}^{'}\mathbf{f}_i\,\Theta_i - \Theta_i\mathbf{f}_i^{'}\,\mathbf{H\Theta} + \mathbf{f}_i^{'}\,(\mathbf{H\Theta\Theta}^{'}\mathbf{H}^{'} + \mathbf{nn}^{'})\,\mathbf{f}_i\big] \tag{37}$$

Using then the unbiasedness constraint (33), we finally come up with

$$\mathrm{E}\big[(\Theta_i - \hat{\Theta}_i)^2\big] = \mathbf{f}_i^{'}\,(\mathrm{E}[\mathbf{HR_\Theta H}^{'}] - \mathrm{E}[\mathbf{H}]\mathbf{R_\Theta}\mathrm{E}[\mathbf{H}^{'}] + \mathbf{R_n})\,\mathbf{f}_i \tag{38}$$

where $\mathbf{R_\Theta}$ and $\mathbf{R_n}$ respectively denote the autocorrelation of the parameter vector $\mathbf{\Theta}$ and of the noise vector $\mathbf{n}$.

A.4. Derivation of the Estimator

Using the Lagrange's method for handling the minimization problem (**??**) with constraint (**??**), we find that the expression of matrix $\mathbf{F}$ is given by

$$\mathbf{F} = \Big(\mathrm{E}[\mathbf{H}^{'}]\mathbf{R}^{-1}\mathrm{E}[\mathbf{H}]\Big)^{-1}\mathrm{E}[\mathbf{H}^{'}]\mathbf{R}^{-1} \tag{39}$$

where $\mathbf{R}$ has been defined as

$$\mathbf{R} = \mathrm{E}[\mathbf{HR_\Theta H}^{'}] - \mathrm{E}[\mathbf{H}]\mathbf{R_\Theta}\mathrm{E}[\mathbf{H}^{'}] + \mathbf{R_n} \tag{40}$$

Note that if $\mathbf{H}$ is deterministic, (39) reduces to

$$\mathbf{F} = \Big(\mathbf{H}^{'}\mathbf{R_n}^{-1}\mathbf{H}\Big)^{-1}\mathbf{H}^{'}\mathbf{R_n}^{-1} \tag{41}$$

which is also the expression of the maximum-likelihood estimator if the noise $\mathbf{n}$ which affects the observation vector $\mathbf{z}$ is gaussian.

B. REFERENCES

[1] C. Herzet, V. Ramon, L. Vandendorpe, and M. Moeneclaey, "Em algorithm-based timing synchronization in turbo receivers," submitted to IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP', 2003.

[2] C. Berrou and A. Glavieux, "Near optimum error correcting coding and decoding: Turbo codes," in *IEEE Transactions on Communications*, Oct. 1996, vol. 44, pp. 1261–1271.

[3] N. Noels, C. Herzet, A. Dejonghe, V. Lottici, H. Steendam, M. Moeneclaey, M. Luise, and L. Vandendorpe, "Turbo-synchronization: an em algorithm interpretation," submitted to IEEE International Conference on Communications, ICC', 2003.

[4] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum-likelihood from incomplete data via the em algorithm," *J. Roy. Stat. Soc.*, vol. 39, no. 1, pp. 1–38, Jan. 1977.

[5] S. ten Brink, J. Speidel, and J. C. Yan, "Iterative demapping and decoding for multilevel modulation," in *IEEE Global Communications Conference, GLOBECOM'*, Sidney, Australia, Nov. 1998, pp. 579–584.

[6] J.M. Mendel, *Lessons in Estimation Theory for Signal Processing Communications and Control*, Prentice Hall Signal Processing Series, Englewood Cliffs, NJ, 1995.

[7] H. Meyr, M. Moeneclaey, and S. Fetchel, *Digital Communication Receivers : Synchronization, Channel Estimation and Signal Processing*, Wiley Series in Telecommunications and Signal Processing, USA, 1998.