Performance Analysis of ML-Based Feedback Carrier Phase Synchronizers for Coded Signals

Nele Noels[(a)], Heidi Steendam[(a)] and Marc Moeneclaey[(a)]

ABSTRACT

This paper considers carrier phase recovery in transmission systems with an iteratively decodable error-control code (turbo codes, low density parity check (LDPC) codes), whose large coding gains enable reliable communication at very low signal-to-noise ratio (SNR). We compare three types of feedback phase synchronizers which are all based upon the maximum likelihood (ML) estimation principle: a data-aided (DA) synchronizer, a non-code-aided (NCA) synchronizer and an iterative code-aided (CA) synchronizer.

We introduce a blockwise forward-backward recursive phase estimator and we show that the mean square phase error (MSPE) of the NCA synchronizer equals that of the DA synchronizer when the carrier phase is constant and the loop filter gain is the same for both synchronizers. When the signal is affected by phase noise, the NCA synchronizer (as compared to the DA synchronizer) yields a larger MSPE due to phase fluctuations. We also show that, at the normal operating SNR of the considered code, the performance of the CA synchronizer is very close to that of a DA synchronizer that knows all transmitted symbols in advance.

1. INTRODUCTION

The last decade has seen the development of powerful error correcting codes such as turbo codes and LDPC codes. The impressive bit error rate (BER) performance of the associated iterative decoding processes implicitly assumes coherent detection, meaning that the carrier phase must be recovered accurately before the data is decoded. However, since the decoder usually operates at extremely low SNR values, accurate carrier recovery is a challenging task. Numerous efforts to tackle this problem have resulted in a myriad of different receivers [1]-[10].

In [1],[2] the phase estimator ignores error-control coding and assumes that the transmitted symbols are mutually independent (NCA operation), whereas in [3]-[10] the code properties are exploited in the phase estimation process (CA operation). In [11] it was shown that the second approach is potentially more accurate.

The iterative scheme in [5], which is based on the expectation-maximization algorithm, is optimal in the sense that it converges to the true ML carrier phase estimate [12],[13]. The algorithm does not require modification of the decoder operation, and the resulting receiver is only marginally more complex than the conventional receiver that a priori knows the exact value of the phase. Unfortunately, its performance rapidly degrades in the presence of a time-varying carrier phase.

In [2], [6], [7], [8] and [10], feedback phase estimation has been adopted to cope with carrier phase variations. The ML-based receiver proposed in [10] combines the low complexity from the approach in [5] with the ability to automatically track a slowly varying carrier phase. Simulation results in [10] show the interesting potential of this approach. As opposed to the algorithms in [2], [6] and [7], the derivation of the phase estimation algorithm stems directly from the ML criterion and can therefore be seen as the feedback counterpart of the receiver presented in [5]. Moreover, its computational complexity is lower than that of the algorithms in [8] and [9], which modify the decoder operation by either taking into account the phase statistics or using per-survivor phase estimates inside the decoder.

This contribution zooms in on the approach that was adopted in [10]. By means of theoretical analysis and computer simulations we compare the tracking performances resulting from the iterative CA synchronizer from [10], the DA synchronizer which knows all transmitted symbols in advance, and the NCA synchronizer which neglects the underlying encoding rule. It is shown that CA feedback phase estimation outperforms NCA feedback estimation when the phase to be estimated varies with time; when the carrier phase is constant over the observation interval, both synchronizers yield essentially the same MSPE. We also show that, at the normal operating SNR of the considered code, the performance of the CA synchronizer is very close to that of a DA synchronizer that knows all transmitted symbols in

[(a)] The authors are with the Telecommunications and Information Processing Dept., Ghent University, Sint-Pietersnieuwstraat 41, B-9000 Ghent, Belgium.
E-mail : {nnoels, hs, mm}@telin.ugent.be Fax : +32 9 264 4295

advance. This illustrates the optimality of the CA synchronizer.

2. MAXIMUM LIKELIHOOD CRITERION

We consider the transmission of an arbitrary sequence of complex-valued symbols $\mathbf{a} = (a_0, a_1, ..., a_{K-1})$ over an additive white Gaussian noise (AWGN) channel. The joint probability mass function of the symbols $\{a_k\}$ is denoted as $p(\mathbf{a})$. Assuming linear modulation using square-root Nyquist transmit pulses, and matched filtering at the correct decision instants, the discrete-time baseband observation is given by

$$r_k = a_k e^{j\theta} + w_k, \qquad k = 0, ..., K-1 \qquad (1)$$

where θ denotes the unknown carrier phase[1], and the sequence $\{w_k\}$ consists of independent zero-mean complex-valued Gaussian noise terms; $\mathrm{Re}[w_k]$ and $\mathrm{Im}[w_k]$ are statistically independent, and have a variance equal to $N_0/2$.

Let us denote by $\tilde{\theta}$ a trial value of the true carrier phase θ that has to be estimated by the synchronizer. Then, the ML estimate of the carrier phase is the value of $\tilde{\theta}$ that makes zero the derivative of the *log-likelihood function* $L(\tilde{\theta}) = \ln(p(\mathbf{r};\tilde{\theta}))$ with respect to $\tilde{\theta}$ [14]. The probability density $p(\mathbf{r}\,|\,\mathbf{a},\tilde{\theta})$ of $\mathbf{r}$ resulting from (1), given the data sequence $\mathbf{a}$ and a trial value $\tilde{\theta}$ of the carrier phase, is (within a factor not depending on $(\mathbf{a};\tilde{\theta})$) given by

$$p(\mathbf{r}\,|\,\mathbf{a};\tilde{\theta}) = \exp\left(\frac{-1}{N_0}\sum_{k=0}^{K-1} |\, r_k - a_k e^{j\tilde{\theta}}\,|^2\right) \qquad (2)$$

The *likelihood function* $p(\mathbf{r};\tilde{\theta})$ of the carrier phase is obtained by averaging $p(\mathbf{r}\,|\,\mathbf{a};\tilde{\theta})$ over the symbol vector $\mathbf{a}$, i.e. $p(\mathbf{r};\tilde{\theta}) = \sum_{\mathbf{a}} p(\mathbf{r}\,|\,\mathbf{a};\tilde{\theta})\,p(\mathbf{a})$. From a similar reasoning as in [5] and [11], the derivative $L'(\tilde{\theta})$ of the log-likelihood function with respect to $\tilde{\theta}$ can be manipulated into the following form:

$$L'(\tilde{\theta}) = \frac{d\ln(p(\mathbf{r};\tilde{\theta}))}{d\tilde{\theta}} = \frac{2}{N_0}\sum_{k=0}^{K-1} \mathrm{Im}\left[A_k^*(\mathbf{r},\tilde{\theta}) r_k e^{-j\tilde{\theta}}\right] \qquad (3)$$

where

$$A_k(\mathbf{r},\tilde{\theta}) = E\left[a_k\,|\,\mathbf{r};\tilde{\theta}\right] = \sum_{m=0}^{M-1} \Pr[a_k = s_m\,|\,\mathbf{r};\tilde{\theta}]\, s_m \qquad (4)$$

is the a posteriori expectation of the symbol a_k conditioned on $\mathbf{r}$ and $\tilde{\theta}$, with $\Pr[a_k = s_m\,|\,\mathbf{r};\tilde{\theta}]$ denoting the marginal a posteriori probability (APP) of the symbol a_k, and $\{s_0, s_1, ..., s_{M-1}\}$ the set of constellation points with symbol energy $E_s = \frac{1}{M}\sum_{m=0}^{M-1} |\, s_m\,|^2$.

When the data symbol vector $\mathbf{a}$ consists of known pilot symbols $(a_{p,0}, a_{p,1}, ..., a_{p,K-1})$, we obtain $\Pr[a_k = s_m\,|\,\mathbf{r};\tilde{\theta}]$ equal to 1 for $s_m = a_{p,k}$ and zero otherwise, yielding $A_k(\mathbf{r},\tilde{\theta}) = a_{p,k}$ in (4). The log-likelihood function that correspond to the transmission of pilot symbols is denoted $L_{ps}(\tilde{\theta})$.

In the case of uncoded transmission, the symbols $\{a_k\}$ are statistically independent, so the APPs of a_k reduce to:

$$\Pr[a_k = s_m\,|\,\mathbf{r};\tilde{\theta}] = \Pr[a_k = s_m\,|\,r_k;\tilde{\theta}] = \frac{F(s_m e^{j\tilde{\theta}}, r_k)}{\sum_{i=0}^{M-1} F(s_i e^{j\tilde{\theta}}, r_k)} \qquad (5)$$

where

$$F(s_m e^{j\tilde{\theta}}, r_k) = \exp\left(\frac{-1}{N_0}\,|\, r_k - s_m e^{j\tilde{\theta}}\,|^2\right) \qquad (6)$$

As (5) depends only on r_k, we will denote the corresponding a posteriori average of the symbol a_k as $A_k(r_k;\tilde{\theta})$. The log-likelihood function that corresponds to the transmission of statistically independent symbols $\{a_k\}$ is denoted $L_{ind}(\tilde{\theta})$.

This paper considers systems with an iteratively decodable error-control code (turbo-, LDPC-codes). The data symbol vector $\mathbf{a} = (a_0, a_1, ..., a_{K-1})$ is obtained from the encoding of a sequence of information bits and a proper mapping of the coded bits on the signal constellation. In this case, the APPs in (4) are a function of *all* components of the vector $\mathbf{r}$. To avoid the computational complexity associated with their exact evaluation[2], the marginal APPs are approximately computed by means of the iterative application of the sum-product (SP) algorithm on a factor graph with cycles [15]. If the cycles in the graph are large (which is reasonable for well-designed turbo

---

[1] The carrier phase is initially assumed to be constant over the observation interval. Later the observation model will be extended to allow a time varying carrier phase.

[2] In principle, the exact marginal APPs $\Pr[a_k\,|\,\mathbf{r};\tilde{\theta}]$ can be obtained as a summation of joint APPs $\Pr[\mathbf{a}\,|\,\mathbf{r};\tilde{\theta}]$, which in turn can be computed from (2) and Bayes' rule. However, the computational complexity of this procedure increases exponentially with the sequence length K.

and LDPC codes), this iterative procedure (after convergence) yields marginal APPs that are very close to the correct marginal APPs. The corresponding log-likelihood function is denoted $L_{coded}(\tilde{\theta})$.

3. ML-BASED PHASE TRACKING FOR CODED SIGNALS

The general structure of a first order discrete-time *feedback* carrier synchronizer or *phase-locked loop* (PLL) is shown in Figure 1 [16]. The phase estimate is updated once per symbol interval, according to the following forward[3] recursion

$$\hat{\theta}_{k+1} = \hat{\theta}_k + \lambda x_k \tag{7}$$

In (7), λ is the loop filter gain, and x_k denotes the phase error detector (PED) output. The recursion starts with an initial phase estimate $\hat{\theta}_0$, that can be obtained from a feedforward synchronizer operating on a short pilot sequence [16].

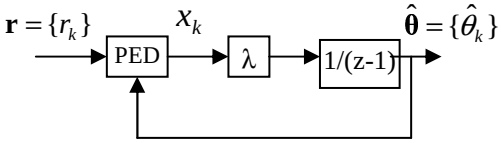

Figure 1: General structure of a discrete-time feedback carrier synchronizer

In the following, we consider three types of ML-based PEDs. The DA PED (based on $L'_{ps}(\tilde{\theta})$) assumes that all data symbols are known. The NCA PED (based on $L'_{ind}(\tilde{\theta})$) assumes that the data symbols are independent, whereas the CA PED (based on $L'_{coded}(\tilde{\theta})$) takes the code properties into account. We obtain from (3) that

$$x_k = \frac{1}{E_s}\begin{cases} \mathrm{Im}[a_k^* r_k e^{-j\hat{\theta}_k}], & DA\,operation \\ \mathrm{Im}[A_k^*(r_k,\hat{\theta}_k) r_k e^{-j\hat{\theta}_k}], & NCA\,operation \\ \mathrm{Im}[A_k^*(\mathbf{r},\hat{\theta}_k) r_k e^{-j\hat{\theta}_k}], & CA\,operation \end{cases} \tag{8}$$

Comparison of the PED outputs for NCA and CA operation with that for DA operation indicates that the a posteriori mean A_k can be considered as a soft decision (SD) regarding a_k, based upon the received sample r_k or the received sample sequence $\mathbf{r}$ and the phase estimate $\hat{\theta}_k$. Note from (8) that the DA and the NCA PED output depend only on $(r_k,\hat{\theta}_k)$; this is in contrast with the CA PED output whose computation depends on the entire vector $\mathbf{r}$: all K samples $(r_0, r_1, ..., r_{K-1})$ have to be rotated over an angle $-\hat{\theta}_k$, and fed to the SP algorithm for producing the SD $A_k(\mathbf{r},\hat{\theta}_k)$. Hence, in the case of CA operation, the entire received block must be processed K times, whereas the received block is processed only once in the case of DA or NCA operation.

In order to avoid the high computational complexity resulting from the CA PED, the following *iterative* CA PLL has been proposed in [10]. During the i-th iteration, the FB synchronizer generates estimates $(\hat{\theta}_0^{(i)}, \hat{\theta}_1^{(i)}, ..., \hat{\theta}_{K-1}^{(i)})$ essentially according to (7), but with the PED output given by

$$x_k = \frac{1}{E_s}\mathrm{Im}[A_k^*(\mathbf{r},\hat{\boldsymbol{\theta}}^{(i-1)}) e^{-j\hat{\theta}_k^{(i)}}], \quad iterative\,CA\,operation \tag{9}$$

where $\hat{\boldsymbol{\theta}}^{(i-1)} = (\hat{\theta}_0^{(i-1)}, \hat{\theta}_1^{(i-1)}, ..., \hat{\theta}_{K-1}^{(i-1)})$, and $A_k(\mathbf{r},\hat{\boldsymbol{\theta}}^{(i-1)})$ is the a posteriori expectation of the symbol a_k conditioned on $\mathbf{r}$ and $\hat{\boldsymbol{\theta}}^{(i-1)}$. Hence, from the phase vector $\hat{\boldsymbol{\theta}}^{(i-1)}$, the received vector $\mathbf{r}$ is processed to compute $A_k(\mathbf{r},\hat{\boldsymbol{\theta}}^{(i-1)})$ for $k = 0,1,...,K-1$, after which the PLL generates the phase vector $\hat{\boldsymbol{\theta}}^{(i)}$. The iterative process is initialized by means of a phase vector $\hat{\boldsymbol{\theta}}^{(0)}$, that can be obtained from a PLL with NCA operation. When convergence is achieved after n iterations, the vector $\mathbf{r} = (r_0, r_1, ..., r_{K-1})$ has been processed n times. When $n << K$, considerable savings in computation time have been obtained as compared to the non-iterative PLL that uses the CA PED output from (8). Moreover, when applied to a turbo or LDPC receiver with iterative MAP detection/decoding, the proposed phase estimation/compensation scheme yields very low additional complexity when the synchronizer iterations are merged with the decoder iterations [4],[5],[10], i.e., after each synchronizer iteration only one decoder iteration is performed without resetting extrinsic probabilities.

4. Tracking Performance Analysis

4.1 Analytical Results

Computing the exact tracking performance of the iterative CA feedback phase estimator is much more difficult than for the NCA and DA synchronizers, because of the iterations involved and the dependence of the soft decisions on the entire phase vector. Instead we will proceed assuming that, at the normal operating SNR of the considered error-correcting code, the MSPE resulting from the iterative CA phase estimator converges to the MSPE resulting from a fictitious DA

[3] We speak of a *forward* recursion when the phase updating is performed from the first symbol interval to the last.

phase estimator that knows all data symbols in advance.

A motivation for this assumption reads as follows. Note that in (8) the CA PED output reduces to the DA PED output when the APP $\Pr[s_m \mid \mathbf{r}; \hat{\theta}_k]$ is one for $s_m = a_k$ and zero otherwise. This indicates that the CA PLL essentially behaves like the DA PLL, provided that the ratios $R(s_m, a_k \mid \mathbf{r}; \hat{\theta}_k) = \Pr[s_m \mid \mathbf{r}; \hat{\theta}_k] / \Pr[a_k \mid \mathbf{r}; \hat{\theta}_k]$ are likely to be much smaller than 1 for all $s_m \neq a_k$ and all $k = 0,1,...,K-1$. Let us introduce the indicator function $I_y(k)$, which equals one when $R(s_m, a_k \mid \mathbf{r}; \hat{\theta}_k) \geq y$ for at least one $s_m \neq a_k$, and equals zero otherwise. Then we obtain

$$E\left[\frac{1}{K}\sum_{k=0}^{K-1} I_y(k)\right] = \frac{1}{M^{\rho K}} \sum_{\mathbf{a}\in\xi} \left(\frac{1}{K}\sum_{k=0}^{K-1} \Pr\left[\bigcup_{s_m \neq a_k} R(s_m, a_k \mid \mathbf{r}; \hat{\theta}_k) \geq y \right] \right) \quad (10)$$

where ξ denotes the set of legitimate coded symbol sequences of length K. We assume that $p(\mathbf{a}) = M^{-\rho K}$ for $\mathbf{a} \in \xi$ and $p(\mathbf{a}) = 0$ otherwise, where the quantities ρ and M denote the rate of the code and the number of constellation points, respectively. With $y = 1$ and $\hat{\theta}_k = \theta$ for all k, equation (10) is nothing but the (very small) symbol error rate resulting from an optimal maximum a posteriori probability symbol decoder [17]. Hence, for small phase errors, the fraction of symbol intervals for which $I_1(k) = 1$ is very small, so that we can safely assume that the CA PLL operation closely resembles the DA PLL operation, at the normal operating SNR of the code.

We will now compute the performance of the DA and the NCA phase estimator. Assuming that at the low SNR supported by capacity-approaching codes it is not possible to compute reliable data decisions without taking into account the code structure, we expect the NCA PLL to perform significantly worse than a DA PLL with perfect knowledge on the data symbols.

In order to allow a time varying carrier phase, the observation model (1) is modified into

$$r_k = a_k e^{j\theta_k} + w_k, \qquad k = 0, ..., K-1 \quad (11)$$

where θ_k is the phase during the k-th symbol interval. An often used phase noise (PN) model is based on a discrete Wiener process (random walk)

$$\theta_k = \theta_{k-1} + \Delta_k \quad (12)$$

characterized by independent and identically distributed (i.i.d.) Gaussian increments Δ_k with zero mean and standard deviation σ_Δ, descriptive of the phase noise intensity. It is assumed that $\{w_k\}$ and $\{\Delta_k\}$ are statistically independent, and that θ_0 is uniformly distributed in [-π,π]. We define the phase estimation error during the k-the symbol period as $\phi_k = \theta_k - \hat{\theta}_k$.[4]

The DA and NCA PED outputs x_k from (8) that depend only on $(r_k, \hat{\theta}_k)$ can be decomposed as the sum of their average $g(\phi_k)$ and their zero-mean statistical fluctuation $N_k(\phi_k)$:

$$x_k = g(\phi_k) + N_k(\phi_k) \quad (13)$$

with

$$g(\phi_k) = E[x_k] \qquad N_k(\phi_k) = x_k - E[x_k] \quad (14)$$

denoting the PED characteristic and the loop noise of the synchronizer, respectively. We show in the Appendix that $g(0) = 0$. Assuming small phase errors, the following linearization applies [16]:

$$x_k = g'(0)\phi_k + N_k(0) \quad (15)$$

where $g'(0)$ is the slope of the PED characteristic and $N_k(0)$ is the loop noise at $\phi_k = 0$. Substituting (15) into (7) we obtain:

$$\phi_{k+1} = (1 - \lambda g'(0))\phi_k + \Delta_{k+1} - \lambda N_k(0) \quad (16)$$

where $E[\Delta_k \Delta_{k'}] = \sigma_\Delta^2 \delta_{k-k'}$. We show in the Appendix that

$$E[N_k(0) N_{k'}(0)] = \left(\frac{N_0}{2E_s} g'(0)\right) \delta_{k-k'} \quad (17)$$

where

$$g'(0) = \begin{cases} 1, & \text{DA operation} \\ S_{NCA}\left(\frac{E_s}{N_0}\right) \leq 1, & \text{NCA operation} \end{cases} \quad (18)$$

and $S_{NCA}(\infty) = 1$. Solving (16) yields,

$$\phi_k = \left(\sum_{m=0}^{k-1} (1 - \lambda g'(0))^m \left(\Delta_{k-m-1} + \lambda N_{k-m-1}(0) \right) \right) + (1 - \lambda g'(0))^k \phi_0 \quad (19)$$

The phase error (19) at the output of the PLL consists of two contributions, that are caused by the noise (AWGN, PN) affecting the observation, and by the initial phase error ϕ_0, respectively. In all practical

---

[4] This definition of the estimation error agrees with the framework in [16]. It should be noted, however, that the estimation error is usually defined as the inverse of ϕ_k, i.e. according to $\hat{\theta}_k - \theta_k$.

cases the quantity $|1-\lambda g'(0)|$ is smaller than 1, so that the phase error (19) exhibits a decaying acquisition transient near the start of the observation interval. Assuming a uniformly distributed initial phase error, the mean acquisition time (T_{acq}) in the absence of noise is well approximated by [16]

$$\frac{T_{acq}}{T} = \frac{1}{2B_L T} \cong \frac{2}{\lambda g'(0)} \tag{20}$$

where $B_L T$ is the one-side bandwidth (normalized to the symbol rate) of the closed-*loop filter* with impulse response $h_k = \lambda g'(0)(1-\lambda g'(0))^k$ and z-transform $H(z) = \lambda g'(0)/(z-1+\lambda g'(0))$. We have

$$\begin{aligned} B_L T &= T \int_0^{1/2T} \left|H(e^{j2\pi fT})\right|^2 df = \frac{1}{2}\sum_{k=0}^{\infty}(h_k)^2 \\ &= \frac{\lambda g'(0)}{2(2-\lambda g'(0))} \cong \frac{\lambda g'(0)}{4} \end{aligned} \tag{21}$$

The approximation in (20) and (21) is valid for small λ. It follows from (20) that a larger λ results in a faster acquisition. The same goes for a larger $g'(0)$. At the end of the acquisition period the phase error enters the tracking mode, during which (19) can be safely approximated by

$$\phi_k = (\lambda g'(0))^{-1}\sum_{m=0}^{\infty} h_m\left(\Delta_{k-m-1} + \lambda N_{k-m-1}(0)\right) \tag{22}$$

It is easily seen from (22) that the steady state phase error has zero mean and that the steady state MSPE is given by

$$E[\phi^2] = MSPE_{AWGN} + MSPE_{PN} \tag{23}$$

with

$$\begin{aligned} MSPE_{AWGN} &= E\left[(N_k(0))^2\right]\frac{2B_L T}{g'(0)^2} \cong \frac{N_0}{E_s}\frac{\lambda}{4} \\ MSPE_{PN} &= \sigma_\Delta^2 \frac{2B_L T}{(\lambda g'(0))^2} \cong \sigma_\Delta^2 \frac{1}{2\lambda g'(0)} \end{aligned} \tag{24}$$

The linearized steady state MSPE from (24) consist of two contributions: an additive noise contribution and a phase noise contribution. The PN contribution is inversely proportional to the slope of the PED characteristic, whereas the AWGN contribution does not depend on the slope of PED characteristic. This is because, for given values of λ and E_s/N_0, the reduction of the PED slope $g'(0)$ of the NCA PLL (as compared with the DA PLL) is precisely compensated by the reduction of the loop bandwidth $2B_L T$ and of the phase noise variance $E\left[(N_k(0))^2\right]$. A larger value of λ yields a larger AWGN contribution but a smaller PN contribution, and vice versa. When the carrier phase is time-invariant ($\sigma_\Delta = 0$), the MSPE can be made arbitrarily small by reducing the value of λ. A small λ, however, implies a large acquisition time. When the carrier phase is time-varying ($\sigma_\Delta \neq 0$), there exists an optimal value for λ that minimizes the steady state MSPE. Solving this optimization problem yields

$$\begin{aligned} \lambda_{opt} &= \sqrt{\frac{2E_s}{N_0}\sigma_\Delta^2 \frac{1}{g'(0)}} \\ MSPE_{\min} &= \sqrt{\frac{N_0}{2E_s}\sigma_\Delta^2 \frac{1}{g'(0)}} \end{aligned} \tag{25}$$

where λ_{opt} and $MSPE_{\min}$ denote the optimal value for λ and the corresponding minimum value for the linearized steady state MSPE, respectively.

4.2 Numerical Results and Discussion

By means of example we consider a BPSK signal constellation, an observation interval of K=999 symbol periods and a rate 1/3 turbo code. The turbo encoder consists of the parallel concatenation of two identical non-recursive systematic convolutional encoders with generator polynomials $(21)_8$ and $(37)_8$ in octal notation, separated by a pseudo-random interleaver of size 333 bits. Figure 2 shows the BER after 1, 2 and 10 iterations of the coherent turbo decoder/detector.

In Figure 3, the slope $S_{NCA}\left(\frac{E_s}{N_0}\right)$ of the NCA PED is plotted against the signal-to-noise ratio E_s/N_0. Monte Carlo simulation techniques were used to evaluate the statistical expectation involved in the expression of $S_{NCA}\left(\frac{E_s}{N_0}\right)$ (see equation (A10) of the Appendix). We observe that, at the normal operating SNR of the turbo code (say, BER<10^{-3}), the slope of the NCA PED characteristic is strictly smaller than 1, i.e. than is the slope of the DA PED.

The tradeoff on λ in case $\sigma_\Delta \neq 0$ is illustrated in Figure 4 showing the numerical evaluation of (24) for E_s/N_0 = -2.77 dB, which is in the operating range of the turbo code from Figure 2, and σ_Δ = 2 degrees. As the slope of the DA PED is larger than that of the NCA PED, the optimized DA PLL yields a smaller acquisition time and a smaller MSPE than the optimized NCA PLL. The optimal loop filter gain λ is around 4.75 x 10^{-2} with the NCA PED, and reduces to 3.6 x 10^{-2} with the DA PED. The minimum MSPE for the NCA PLL is a factor of 0.57 larger than the minimum MSPE for the DA PLL. This is consistent with (25) and with the behavior of $S_{NCA}\left(\frac{E_s}{N_0}\right)$ depicted in Figure 3

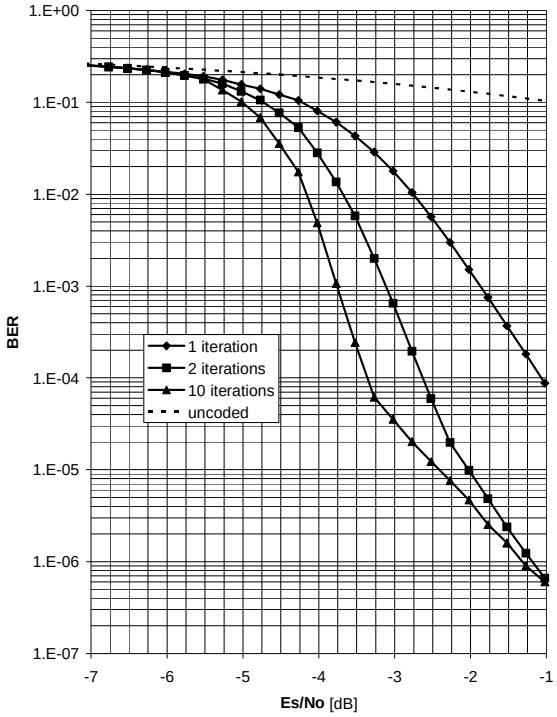

Figure 2: BER of coherent turbo receiver for a rate 1/3 turbo coded BPSK signal in AWGN

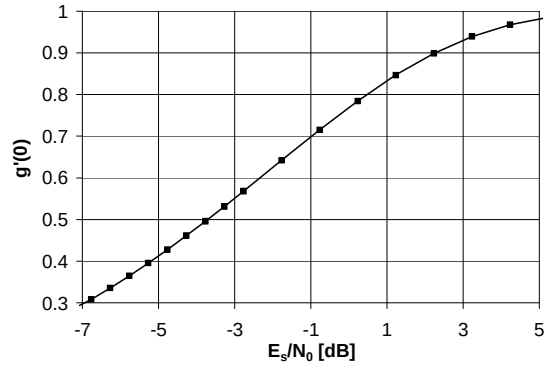

Figure 3: slope of the NCA PED characteristic at $\phi_k = 0$, in the case of BPSK

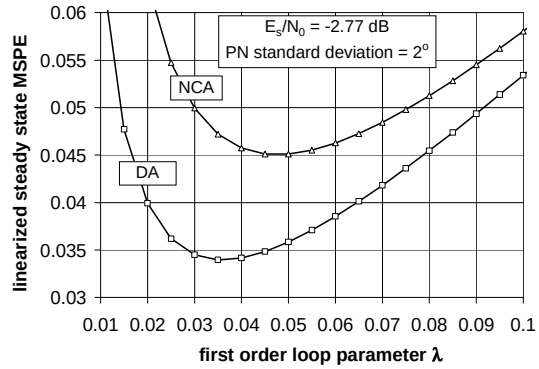

Figure 4: linearized steady state MSPE as a function of the loop parameter λ

In Figure 5 we have plot the simulated MSPE at the output of the DA, the NCA and the iterative CA PLL as a function of the symbol index k. We have taken E_s/N_0 = -2.77 dB and λ = 0.04. The carrier phase is assumed to be either constant over the observation interval (CCP), or to perform a random walk with σ_Δ = 2 degrees (WPN). The iteration i=0 of the CA PLL is an NCA recursion. In both cases, we find that the MSPE of the iterative CA PLL becomes essentially equal to the MSPE of the DA PLL after only 2 iterations (i.e. for iterations i>1). This confirms the validity of the assumption made at the beginning of Section 4.1. As the slope of the NCA PED is smaller than that of the DA PED, the acquisition time and the PN contribution to the tracking MSPE are larger for NCA operation than for DA operation (CA operation, i>1). Conversely, the AWGN contribution to the tracking MSPE, is the same for NCA operation and for DA operation (CA operation, i>1). This implies that, when the carrier phase is time-invariant (CCP), the tracking MSPE can not be reduced by performing, after the initial NCA recursion, iterations in the CA mode.

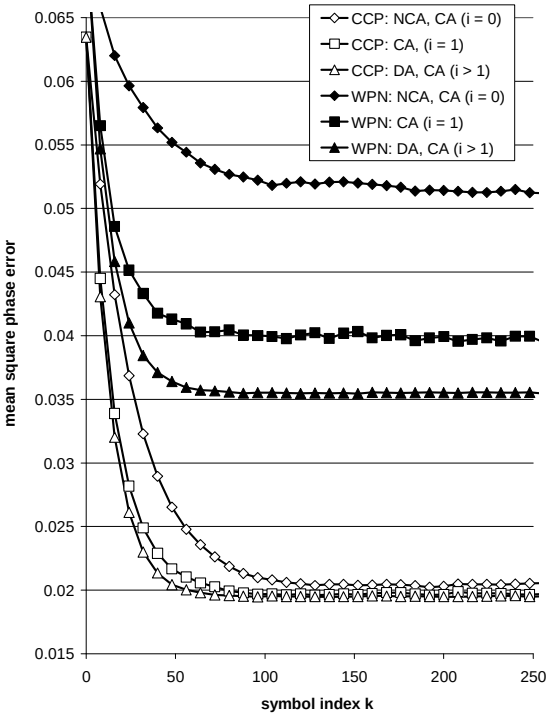

Figure 5: MSPE of a first order PLL with λ = 0.04, tracking a constant carrier phase (CCP) or Wiener phase noise with σ_Δ = 2 degrees (WPN)

5. FORWARD-BACKWARD PHASE UPDATING

During the acquisition period at the beginning of the observation interval (see Figure 5) the MSPE may

assume large values. Assuming that the acquisition transient is no longer present at the end of the observation interval, accurate phase estimates at the beginning of the observation interval can be obtained by carrying out an additional recursion, using as initial phase estimate the estimate obtained at the end of the first recursion and updating the phase estimates from the last symbol to the first according to the following *backward* recursion

$$\hat{\theta}_{k-1} = \hat{\theta}_k + \lambda x_k \qquad \text{(backward recursion)} \qquad (26)$$

The result of this procedure is shown in Figure 6 for a first order DA PLL with λ = 0.04, and for E_s/N_0 = -2.77 dB and σ_Δ = 0. For a DA PLL or a NCA PLL, the edge effect that arises near the end of the observation interval can be explained as follows. Substituting (15) into (26) we obtain:

$$\phi_{k-1} = (1-\lambda g'(0))\phi_k - \Delta_k - \lambda N_k(0) \qquad (27)$$

with $E[N_k(0)N_{k'}(0)] = g'(0)\delta_{k-k'}$, $E[\Delta_k \Delta_{k'}] = \sigma_\Delta^2 \delta_{k-k'}$ and $g'(0)$ given by (18). Taking into account that ϕ_{K-1} is given by (22) with $k = K$, solving (27) yields

$$\phi_k = \left(\sum_{m=0}^{k_\Delta - 1} (1-\lambda g'(0))^m \left(-\Delta_{k+m} + \lambda N_{k+m+1}(0)\right) \right) + (1-\lambda g'(0))^{k_\Delta} \left(\sum_{m=0}^{\infty} (1-\lambda g'(0))^m \left(\Delta_{K-m-1} + \lambda E_s N_{K-m-1}(0)\right) \right) \qquad (28)$$

where $k_\Delta = (K-1) - k$. Approximating $(1-\lambda g'(0))^n$ by 0 for $n \geq K-1$ and taking into account that the noise samples are mutually independent, it can be verified that

$$E[(\phi_k)^2] = MSPE_{AWGN}\left(1 + f_{k_\Delta}\right) + MSPE_{PN}\left(1 - \lambda^2 (1-\lambda g'(0)) f_{k_\Delta}\right) \qquad (29)$$

where

$$f_n = 2\lambda\beta g'(0)\, n(1-\lambda g'(0))^{2n-1} \qquad (30)$$

and $MSPE_{AWGN}$ and $MSPE_{PN}$ are given by (24). The terms proportional to f_{k_Δ} do not occur in (23) and are a result of the reusing of the samples. The function f_n is plot in Figure 7, for $\lambda g'(0) << 1$. For $n = 0$, it equals zero. As the product of a linearly increasing function and an exponentially decreasing function, f_n first increases and then decreases with increasing n. For infinite n, f_n converges to zero. This explains the shape of the edge effect in the MSPE near the end of the observation interval (where k_Δ is small). Near the start of the observation interval (where k_Δ is large), f_{k_Δ} is negligibly small and (29) reduces to (23). Assuming that $n(1-\lambda g'(0))^{2n-1}$ is essentially zero for $n \geq K/2$, the edge effect can be circumvented by taking as final phase estimates the estimates at instants $k = 0,1,...,\frac{K}{2}-1$ from the forward recursion and those with indices $k = \frac{K}{2}, \frac{K}{2}+1, ..., K-1$ from the backward recursion.

In the above we have considered only two successive recursions, one in the forward and one in the backward direction, but the forward-backward phase updating principle is easily extended to a higher number of recursions. In order to reduce the effect of acquisition transients, the iterative CA PLL is modified as follows. Instead of performing only forward recursions, an alternation of forward (for iterations i = 0, 2, 4, ...) and backward (for iterations i = 1, 3, 5, ...) recursions is carried out, with each recursion using as initial phase estimate the estimate obtained at the end of the previous iteration.

Let us reconsider the turbo-coded BPSK signal from Section 4. Figure 8 shows the MSPE performance of the NCA and CA feedback phase estimators with forward-backward phase updating and λ = 0.04, for several values of the phase noise variance σ_Δ^2 as a function of E_s/N_0. The simulated MSPE (continuous lines) is compared with the linearized analytical result from Section 4 (dashed lines). The discrepancy (which is quite small for E_s/N_0 larger than about -4 dB) between the simulations and the computations is the result of the random occurrence of non-linear phenomena such as cycle slips and hang-ups [16]. In the absence of phase noise, the CA algorithm yields essentially the same MSPE as the conventional NCA algorithm since the latter is given 'more time' by the forward-backward recursion to let its transient fade out. In the presence of phase noise, the MSPE of the CA synchronizer is lower than that of the conventional NCA algorithm. The relative advantage of the CA synchronizer over the NCA synchronizer increases with the value of the phase noise variance. It speaks for itself that the estimation accuracy increases with E_s/N_0.

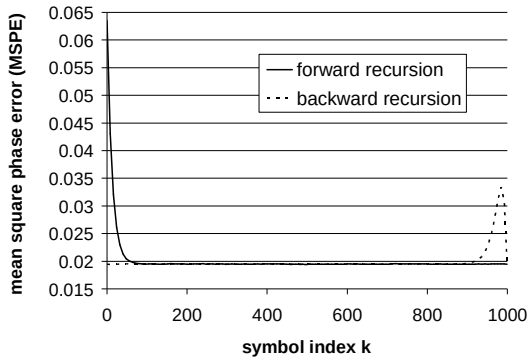

Figure 6: MSPE of a first order DA PLL with λ = 0.04, tracking a constant carrier phase and using forward-backward phase updating

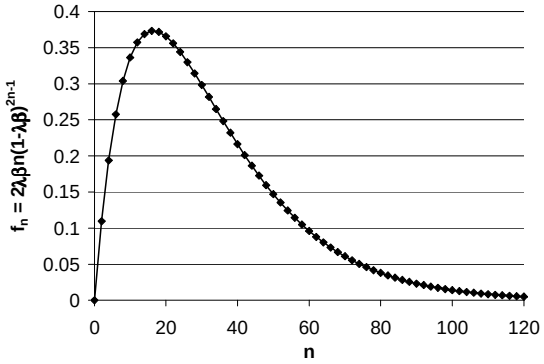

Figure 7: $f_n = 2\lambda g'(0)n(1-\lambda g'(0))^{2n-1}$ as a function of n

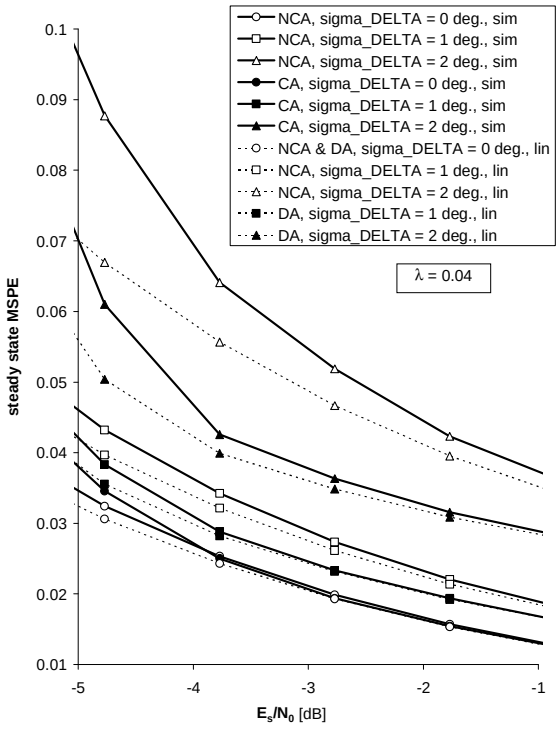

Figure 8: MSPE of a first order ML-based PLL with λ = 0.04. CA operation versus NCA operation, Simulation versus analytical computation.

6. CONCLUSIONS

This contribution has studied the effectiveness of CA and NCA ML-based feedback phase synchronizers at the low SNRs supported by powerful iteratively decodable codes such as turbo codes or LDPC codes. Under normal operating conditions, the MSPE resulting from the iterative CA synchronizer converges to the MSPE resulting from a DA synchronizer that knows all data symbols in advance. This illustrates the optimality of the CA ML-based feedback phase synchronizer. By virtue of a forward-backward multiple-recursion estimator, the linearized MSPE of the NCA feedback synchronizer equals that of the CA feedback synchronizer, when the carrier phase is essentially constant over the observation interval and the loop filter gain is the same for both synchronizers. Conversely, the presence of Wiener phase noise results in a NCA feedback synchronizer MSPE that is larger than the CA feedback synchronizer MSPE. The NCA synchronizer also yields a larger acquisition time than the CA synchronizer. However, assuming that the acquisition transient is shorter than the observation interval, the effect of the acquisition transient on the phase error can be circumvented by carrying out an alternation of forward and backward phase updating recursions, with each recursion using as initial phase estimate the estimate obtained at the end of the previous recursion.

APPENDIX

We first consider NCA feedback phase synchronization. Taking in equation (3) $\mathbf{r}$ = r_k and $\tilde{\theta} = \hat{\theta}_k$, we find that the NCA PED output from (8) can be rewritten as follows

$$x_k = \frac{1}{E_s}\mathrm{Im}[A_k^*(r_k,\hat{\theta}_k)r_k e^{-j\hat{\theta}_k}] = \frac{N_0}{2E_s}\frac{d\ln\left(p(r_k;\hat{\theta}_k)\right)}{d\hat{\theta}_k} \quad \text{(A1)}$$

where the probability density $p(r_k;\hat{\theta}_k)$ is (within an irrelevant factor) given by $\frac{1}{M}\sum_{i=0}^{M-1}F(s_i e^{j\hat{\theta}_k}, r_k)$ with $F(s_i e^{j\hat{\theta}_k}, r_k)$ as in (6).

The decomposition (13)-(14) of (A1) as the sum of the PED characteristic $g(\phi_k)$ and the loop noise $N_k(\phi_k)$ yields

$$g(\phi_k) = \frac{N_0}{2E_s}E\left[\frac{d\ln\left(p(r_k;\hat{\theta}_k)\right)}{d\hat{\theta}_k}\right] \quad \text{(A2)}$$

$$N_k(\phi_k) = \frac{N_0}{2E_s}\frac{d\ln\left(p(r_k;\hat{\theta}_k)\right)}{d\hat{\theta}_k} - g(\phi_k) \quad \text{(A3)}$$

Taking the first and the second derivative (with respect to the true carrier phase θ_k) of both sides of the

normalization constraint $\int p(r_k;\theta_k)dr_k = 1$, and using $df(x)/dx = f(x)\{d\ln(f(x))/dx\}$ it can be verified that

$$E\left[\frac{d\ln(p(r_k;\theta_k))}{d\theta_k}\right] = 0 \quad \text{(A4)}$$

and

$$E\left[\left(\frac{d\ln(p(r_k;\theta_k))}{d\theta_k}\right)^2\right] = -E\left[\frac{d^2\ln(p(r_k;\theta_k))}{d\theta_k{}^2}\right] \quad \text{(A5)}$$

It follows directly from $\phi_k = \theta_k - \hat{\theta}_k$ and (A4) that $g(0) = 0$. Taking into account that $\phi_k = \theta_k - \hat{\theta}_k$ and (11), the PED slope $g'(0)$ and the loop noise $N_k(0)$ at $\phi_k = 0$ are given by

$$g'(0) = -\frac{N_0}{2E_s}E\left[\frac{d^2\ln(p(r_k;\theta_k))}{d\theta_k{}^2}\right] \quad \text{(A6)}$$

$$N_k(0) = \frac{N_0}{2E_s}\frac{d\ln(p(r_k;\theta_k))}{d\theta_k} \quad \text{(A7)}$$

Because of the statistical properties of $\{r_k\}$, the loop noise at $\phi_k = 0$ is white[5] and its power spectral density is given by

$$E[(N_k(0))^2] = \left(\frac{N_0}{2E_s}\right)^2 E\left[\left(\frac{d\ln(p(r_k;\theta_k))}{d\theta_k}\right)^2\right] \quad \text{(A8)}$$

Taking into account (A5) and (3), we obtain

$$g'(0) = S_{NCA}\left(\frac{E_s}{N_0}\right)$$
$$E\left[(N_k(0))^2\right] = \frac{N_0}{2E_s}S_{NCA}\left(\frac{E_s}{N_0}\right) \quad \text{(A9)}$$

with

$$S_{NCA}\left(\frac{E_s}{N_0}\right) = \frac{2}{E_sN_0}E\left[\mathrm{Im}^2\left[A_k^*(r_k;\theta_k)r_ke^{-j\theta_k}\right]\right] \quad \text{(A10)}$$

It can be verified that S_{NCA} from (A10) equals the ratio of the modified Cramer-Rao bound (MCRB) to the true Cramer-Rao bound (CRB) related to the estimation of an unknown but deterministic carrier phase from the noisy observation of K uncoded data symbols [14]. As CRB ≥ MCRB, and CRB converges to MCRB for large SNR, it is found that $0 \le S_{NCA}\left(\frac{E_s}{N_0}\right) \le 1$, and $S_{NCA}(\infty) = 1$.

[5] In the case of uncoded transmission this follows directly from the fact that r_k and $r_{k'}$ are independent for $k \neq k'$, but it can be shown analytically (outside the scope of this paper) that this property holds independently of the code properties.

The above analysis remains valid for DA operation provided that we replace $A_k(r_k;\hat{\theta}_k)$ with a_k, and $p\left(r_k;\hat{\theta}_k\right)$ with $p\left(r_k;a_k,\hat{\theta}_k\right)$, which is (within an irrelevant factor) given by $F(a_ke^{j\hat{\theta}_k}, r_k)$. In this case, we obtain

$$g'(0) = 1$$
$$E\left[(N_k(0))^2\right] = \frac{N_0}{2E_s} \quad \text{(A11)}$$

References

[1] A. D'Amico, A. D'Andrea and R.Reggiannini, "Efficient Non-Data-Aided Carrier and Clock Recovery for Satellite DVB at Very Low Signal-to-Noise Ratio," IEEE J. on Sel. Areas in Commun., vol. 19, NO. 12, Dec. 2001, pp. 2320-2330

[2] L. Lu and G. Wilson, "Synchronization of Turbo Coded Modulation Systems at Low SNR," in Proc. Int. Conf. on Commun., Atlanta, GA, 1998, pp. 428-432

[3] L. Zhang and A. Burr, "Iterative Carrier Phase Recovery suited for Turbo-Coded systems," IEEE Trans. On Wireless Commun., vol. 3, NO. 6, Nov. 2004, pp. 2267-2276

[4] V. Lottici and M. Luise, "Embedding carrier phase recovery into Iterative Decoding of Turbo-Coded Linear Modulations," IEEE Trans. Commun., vol. 52, NO. 4, Apr. 2004, pp. 661-669

[5] N. Noels, V. Lottici, A. Dejonghe, H. Steendam, M. Moeneclaey, M. Luise and L. Vandendorpe, "A Theoretical Framework for Soft Information Based Synchronization in Iterative (Turbo) Receivers," *EURASIP Journal on Wireless Communications and Networking*, vol. 2005, NO. 2,, April 2005, pp. 117-129

[6] W. Oh and K. Cheun, "Joint decoding and carrier recovery algorithms for turbo codes," *IEEE Commun. Lett.*, vol. 6, Sept. 2001, pp. 375-377

[7] C. Langlais and M. Helard, "Phase carrier recovery for turbo codes over a satellite link with the help of tentative decisions," in *Proc. Intern. Symp. on Turbo Codes & Relat. Topics*, Brest, France, Sept. 2000, pp. 439-442

[8] A. Anastasopoulos and K.M.Chugg, " Adaptive iterative detection for phase tracking in turbo-coded systems," *IEEE Trans. Commun.*, vol. 49, Dec. 2001, pp 2135-2144

[9] G. Colavolpe, A. Barbieri and G. Caire, "Iterative decoding in the presence of strong phase noise," submitted to *IEEE J. on Sel. Areas in Commun.*, available at www. eurecom.fr/~caire

[10] N. Noels, H. Steendam and M. Moeneclaey, "A Maximum-Likelihood Based Feedback Carrier Synchronizer for Turbo-Coded Systems", in Proc. *IEEE Vehicular Technology Conference (VTC) spring 2005*, Stockholm, Sweden, May 2005

[11] N. Noels, H. Steendam, M. Moeneclaey, "The Cramer-Rao Bound for Phase Estimation from Coded

Linearly Modulated Signals," *IEEE Commun. Lett.*, vol. 7, No. 5, May 2003, pp. 207-209

[12] C. Georghiades and D. Snyder, "The Expectation-Maximization Algorithm for Symbol Unsynchronized Sequence Detection," IEEE Trans. On Commun., vol. 39, NO. 1, Jan. 1991, pp. 54-61

[13] R.A. Boyles, "On the convergence of the EM algorithm," J. Roy. Statist. Soc. B 45, No 1, pp. 47-50

[14] H. L Van Trees, "*Detection, Estimation, and Modulation Theory*," New York, Wiley, 1990

[15] F. Kschinschang, B. Frey and H.-A. Loeliger, "Factor Graphs and the sum-product algorithm," IEEE Trans. Inform. Theory, vol. 47, NO. 2, Feb. 2001, pp. 498-519

[16] H. Meyr, M. Moeneclaey and S. A. Fechtel, "*Digital Communication Receivers, Synchronization, Channel Estimation and Signal Processing*," New York, Wiley, 1998

[17] L.R. Bahl, J. Cocke, F. Jelinek and J. Raviv, "Optimal decoding of linear codes for minimizing symbol error rate," *IEEE Trans. Inform. Theory*, vol IT-20, March 1974, pp. 248-287