Lower Bounds for the Performance of Iterative Timing Recovery at low SNR

Aravind R. Nayak, John R. Barry and Steven W. McLaughlin
School of Electrical and Computer Engineering
Georgia Institute of Technology
Atlanta , GA 30332 USA

Abstract

We consider the problem of timing recovery at low signal-to-noise ratio. We first derive a lower bound for the timing estimation error variance in the presence of a random walk timing jitter, for the PR-IV channel. Then, we look at the trained phase-locked loop, which gives a heuristic lower bound for the performance of iterative timing recovery schemes involving phase locked loops.

1 Introduction

The push for higher recording densities has motivated the development of iterative error-control codes of unprecedented power, whose large coding gains enable low error rates at very low SNR [1] [2]. In addition, the iterative decoding technique has been extended to turbo equalization, where the equalizer and the decoder iterate [3]. Consequently, timing recovery, which typically derives no benefit from coding, must be performed at an SNR lower than ever before.

At high SNR, the timing recovery process can be separated from the decoding process with little penalty; timing recovery can use an instantaneous decision device to provide tentative decisions that are adequately reliable, which can then be used to estimate the timing error. In essence, the timing recovery process is able to ignore the presence of the code, and assume instead that neighboring symbols are independent. At low SNR, however, timing recovery and decoding are intertwined. The timing recovery process must exploit the presence of the code to get reliable decisions, and the decoder must be fed well-timed samples to function properly.

In principle one could formulate the problem of jointly determining the maximum-likelihood (ML) estimates of the timing offsets and message bits, but the complexity would be prohibitive. A solution based on the expectation-maximization (EM) algorithm would also be complex [4]. A method for jointly performing the tasks of timing recovery and turbo equalization was proposed in [5], with complexity comparable to a conventional turbo equalizer.

In this paper, we look at fundamental limits to the performance of timing recovery systems. We present a lower bound on the timing estimation error variance based on the Cramér-Rao bound. Due to the nature of the system model chosen, this bound is not achievable, and the closest we can get is by ML estimation, which is prohibitively complex. Therefore, to get a tighter practical bound, we look at the performance of a trained phase-locked loop (PLL) which gives a heuristic lower bound for the performance of iterative PLL-based receivers.

2 Lower Bound on Timing Error Variance

2.1 System Model

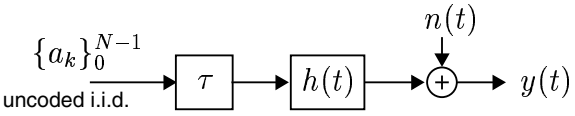

Figure 1: Block diagram of the system model.

Consider the system shown in Figure 1, where the channel output waveform $y(t)$ is given by

$$y(t) = \sum_{k=0}^{N-1} a_k h(t - kT - \tau_k) + n(t), \tag{1}$$

where T is the bit period, $a_k \in \{\pm 1\}$ are the N *i.i.d.* data symbols, $h(t)$ is the channel impulse response, $n(t)$ is additive white Gaussian noise, and τ_k is the unknown timing offset for the k^{th} symbol. This is the generic model of bandlimited communication channels with timing error. Our motivation is in magnetic recording where $h(t)$ is some so-called partial response (PR) target. We consider $h(t) = p(t) - p(t-2T)$, which is the perfect PR-IV pulse, where $p(t) = \sin(\pi t/T)/(\pi t/T)$ is a 0% excess bandwidth pulse. We model the timing offset as a random walk, according to

$$\tau_{k+1} = \tau_k + w_{k+1} = \tau_{-1} + \sum_{j=0}^{k+1} w_j, \tag{2}$$

where $w_k \sim \mathcal{N}(0, \sigma_w^2)$ are *i.i.d.* and σ_w^2 determines the severity of the timing jitter. The random walk model is chosen because of its simplicity and because of its ability to model a wide range of channels by varying a single parameter. We assume perfect acquisition by setting $\tau_{-1} = 0$.

For the case where the timing offset remains constant over the duration of the packet, the Cramér-Rao bound on the timing estimation error variance for generic channels is well known [6]. For the magnetic recording channel, a variety of symbol rate timing recovery techniques are available [7]. Our goal is to arrive at a lower bound on timing estimation error variance in face of a time-varying timing offset.

At the receiver, a front-end lowpass filter (with impulse response $p(t)/T$) is used to eliminate out-of-band noise from the readback waveform $y(t)$, producing the bandlimited waveform $r(t)$. We then sample $r(t)$ at instants kT to get $r_k = \sum_l a_l h(kT - lT - \tau_l) + n_k$, where $\{n_k\}$ are *i.i.d.* $\mathcal{N}(0, \sigma^2)$. Define the observation vector $\boldsymbol{r} = [r_{-M}\ r_{-M+1} \ldots r_{N+M-1}]^T$, where we have collected $N + 2M$ samples, and where we will eventually let $M \to \infty$.

We aim to estimate the data vector $\boldsymbol{a} = [a_0\ a_1 \ldots a_{N-1}]^T$. One way is to estimate the timing vector $\boldsymbol{\tau} = [\tau_0\ \tau_1 \ldots \tau_{N-1}]^T$ and use this information in estimating $\boldsymbol{a}$. So, our problem can be reformulated as estimation of the parameter $\boldsymbol{\theta_1} = [\boldsymbol{\tau}^T\ \boldsymbol{a}^T]^T$, which has $2N$

unknown elements. The analysis that follows requires $\boldsymbol{a}$ to come from an open subspace of R^N. Therefore, we add an independent zero-mean Gaussian random variable of variance σ_a^2 to each of the symbols a_k. As σ_a becomes smaller, we approximate the discrete pdf better.

The Cramér-Rao bound gives a lower bound on the estimation error variance of unbiased estimators of deterministic parameters. This has been extended to get a lower bound on the mean-squared estimation error for any unbiased estimator $\hat{\boldsymbol{\theta}}_{\mathbf{1}}(\boldsymbol{r})$ which is given by [8]

$$E[(\hat{\theta}_{1i}(\boldsymbol{r}) - \theta_{1i})^2] \geq \boldsymbol{J}_{\theta_1}^{-1}(i,i), \tag{3}$$

where $\hat{\theta}_{1i}(\boldsymbol{r})$ and θ_{1i} denote the i^{th} element of $\hat{\boldsymbol{\theta}}_{\mathbf{1}}(\boldsymbol{r})$ and $\boldsymbol{\theta_1}$ respectively, the expectation is taken over $\boldsymbol{r}$ and $\boldsymbol{\theta_1}$, and $\boldsymbol{J}_{\theta_1}^{-1}(i,i)$ is the i^{th} diagonal element of the inverse of the matrix $\boldsymbol{J}_{\theta_1}$ defined by

$$\boldsymbol{J}_{\theta_1} = E\left\{\left[\frac{\partial}{\partial \boldsymbol{\theta_1}} \ln f_{\boldsymbol{r},\boldsymbol{\theta_1}}(\boldsymbol{r},\boldsymbol{\theta_1})\right]\left[\frac{\partial}{\partial \boldsymbol{\theta_1}} \ln f_{\boldsymbol{r},\boldsymbol{\theta_1}}(\boldsymbol{r},\boldsymbol{\theta_1})\right]^T\right\}, \tag{4}$$

where the expectation is taken over $\boldsymbol{r}$ and $\boldsymbol{\theta_1}$, and $f_{\boldsymbol{r},\boldsymbol{\theta_1}}(\boldsymbol{r},\boldsymbol{\theta_1})$ is the joint probability density of $\boldsymbol{r}$ and $\boldsymbol{\theta_1}$. More generally,

$$E[(\hat{\boldsymbol{\theta}}_{\mathbf{1}}(\boldsymbol{r}) - \boldsymbol{\theta_1})(\hat{\boldsymbol{\theta}}_{\mathbf{1}}(\boldsymbol{r}) - \boldsymbol{\theta_1})^T] \geq \boldsymbol{J}_{\theta_1}^{-1}. \tag{5}$$

2.2 Computing $\boldsymbol{J}_{\theta_1}$

Instead of directly computing $\boldsymbol{J}_{\theta_1}$, we define $\boldsymbol{\theta_2} = [\boldsymbol{w}^T\ \boldsymbol{a}^T]^T$, where $\boldsymbol{w} = [w_0\ w_1\ \ldots\ w_{N-1}]^T$, compute $\boldsymbol{J}_{\theta_2}$, and then compute $\boldsymbol{J}_{\theta_1}$ using $\boldsymbol{J}_{\theta_2}$.

The vector $\boldsymbol{\theta_2}$ consists of independent variables. After some manipulations, where we use Bayes' rule, Parseval's theorem and let $M \to \infty$, we get

$$\boldsymbol{J}_{\theta_2} = \begin{bmatrix} \frac{1}{\sigma^2}\boldsymbol{H_1} + \frac{1}{\sigma_w^2}\boldsymbol{I} & \mathbf{0} \\ \mathbf{0} & \frac{1}{\sigma^2}\boldsymbol{H_2} + \frac{1}{\sigma_a^2}\boldsymbol{I} \end{bmatrix}, \quad \text{where} \tag{6}$$

$$\boldsymbol{H_1} = (1+\sigma_a^2)E_{h'}\begin{bmatrix} N & N-1 & N-2 & \ldots & 1 \\ N-1 & N-1 & N-2 & \ldots & 1 \\ N-2 & N-2 & N-2 & \ldots & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & 1 & \ldots & 1 \end{bmatrix}, \tag{7}$$

$$\boldsymbol{H_2} = \begin{bmatrix} 2 & 0 & -1 & 0 & \ldots & \ldots & 0 \\ 0 & 2 & 0 & -1 & \ddots & & \vdots \\ -1 & 0 & 2 & 0 & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & 0 & 2 & 0 & -1 \\ \vdots & & \ddots & -1 & 0 & 2 & 0 \\ 0 & \ldots & \ldots & 0 & -1 & 0 & 2 \end{bmatrix}, \tag{8}$$

and $E_{h'}$ is the energy in the derivative of $h(t)$, which evaluates to

$$E_{h'} = \frac{1}{T^2}\left(\frac{2\pi^2}{3} - 1\right). \tag{9}$$

The vectors $\boldsymbol{\theta_1}$ and $\boldsymbol{\theta_2}$ are related by the linear transformation

$$\boldsymbol{\theta_1} = \begin{bmatrix} \boldsymbol{A}_{N\times N} & \mathbf{0}_{N\times N} \\ \mathbf{0}_{N\times N} & \boldsymbol{I}_{N\times N} \end{bmatrix} \boldsymbol{\theta_2} = \boldsymbol{T}\boldsymbol{\theta_2}, \tag{10}$$

where

$$\boldsymbol{A} = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ 1 & 1 & 0 & \dots & 0 \\ 1 & 1 & 1 & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & 0 \\ 1 & 1 & 1 & \dots & 1 \end{bmatrix} \text{ is a lower triangular matrix.} \tag{11}$$

Therefore the corresponding Fisher information matrices are related by [9]

$$\boldsymbol{J}_{\theta_1} = \boldsymbol{T}^{-T}\boldsymbol{J}_{\theta_2}\boldsymbol{T}^{-1}, \quad \text{where} \quad \boldsymbol{T}^{-1} = \begin{bmatrix} \boldsymbol{A}^{-1} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{I} \end{bmatrix}. \tag{12}$$

Combining (6) and (12), we get

$$\begin{aligned} \boldsymbol{J}_{\theta_1} &= \begin{bmatrix} \boldsymbol{J}_1^{11} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{J}_1^{22} \end{bmatrix}, \qquad \text{where} \\ \boldsymbol{J}_1^{11} &= \frac{1}{\sigma^2}\boldsymbol{A}^{-T}\boldsymbol{H_1}\boldsymbol{A}^{-1} + \frac{1}{\sigma_w^2}\boldsymbol{A}^{-T}\boldsymbol{A}^{-1} \qquad \text{and} \\ \boldsymbol{J}_1^{22} &= \frac{1}{\sigma^2}\boldsymbol{H_2} + \frac{1}{\sigma_a^2}\boldsymbol{I}. \end{aligned} \tag{13}$$

To compute $\boldsymbol{A}^{-1}$, recognize the fact that we need now the inverse mapping of $\tau_k = \sum_{j=0}^{k} w_j$, which is $w_k = \tau_k - \tau_{k-1}$. Therefore,

$$\boldsymbol{A}^{-1} = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ -1 & 1 & 0 & \dots & 0 \\ 0 & -1 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & -1 & 1 \end{bmatrix}. \tag{14}$$

Combining (7) and (13), we get

$$\boldsymbol{J}_1^{11} = \frac{1}{\sigma_w^2}\begin{bmatrix} \lambda & -1 & 0 & \dots & 0 \\ -1 & \lambda & -1 & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & -1 & \lambda & -1 \\ 0 & \dots & 0 & -1 & \lambda - 1 \end{bmatrix}, \tag{15}$$

where

$$\lambda = 2 + (1 + \sigma_a^2)\left(\frac{2\pi^2}{3} - 1\right)\frac{\sigma_w^2}{\sigma^2 T^2}. \tag{16}$$

2.3 Inverting $\boldsymbol{J}_{\theta_1}$

From (13), we see that

$$\boldsymbol{J}_{\theta_1}^{-1} = \begin{bmatrix} [\boldsymbol{J}_1^{11}]^{-1} & \mathbf{0} \\ \mathbf{0} & [\boldsymbol{J}_1^{22}]^{-1} \end{bmatrix}. \tag{17}$$

Letting $\sigma_a \to 0$, we see from (13) that $[\boldsymbol{J}_1^{22}]^{-1} \to \mathbf{0}_{N\times N}$. To invert $\boldsymbol{J}_1^{11}$, we follow a three step procedure. First we do a Cholesky factorization of $\sigma_w^2\boldsymbol{J}_1^{11}$ to get $\sigma_w^2\boldsymbol{J}_1^{11} = \boldsymbol{L}\boldsymbol{L}^T$, where $\boldsymbol{L}$ is a lower triangular matrix. Then, we solve $\boldsymbol{L}\boldsymbol{C} = \boldsymbol{I}$. Next, solve $\boldsymbol{L}^T\boldsymbol{B} = \boldsymbol{C}$, and finally, $[\boldsymbol{J}_1^{11}]^{-1} = \sigma_w^2\boldsymbol{B}$. Recognizing that $\boldsymbol{B}$, which is the inverse of a symmetric matrix, is symmetric, we get

$$[\boldsymbol{J}_1^{11}]_{ij}^{-1} = \sigma_w^2[\boldsymbol{B}]_{ij} = \sigma_w^2 \frac{a_{\max(i,j)} n_{\min(i,j)}}{n_N - n_{N-1}}, \tag{18}$$

where

$$a_j = \frac{\eta^{N-j} + \eta^{-N+1+j}}{\eta+1}, \quad n_j = \frac{\eta^{j+2} - \eta^{-j}}{\eta^2-1}, \quad \text{and} \quad \eta = \frac{\lambda + \sqrt{\lambda^2-4}}{2}. \tag{19}$$

2.4 The lower bound

From (3), we see that we are interested in the diagonal entries of $[\boldsymbol{J}_1^{11}]^{-1}$. Combining (16), (18), and (19), and simplifying, the lower bound can be expressed as

$$\frac{E[(\hat{\tau}_i(\boldsymbol{r}) - \tau_i)^2]}{T^2} \geq \frac{[\boldsymbol{J}_1^{11}]_{ii}^{-1}}{T^2} = h \cdot f(i), \tag{20}$$

where

$$h = \frac{\sigma_w^2}{T^2}\frac{\eta}{\eta^2-1} \quad \text{is the steady state value,} \tag{21}$$

$$f(i) = \tanh\left((N+\frac{1}{2})\ln\eta\right)\left[1 - \frac{\sinh\left((N-2i-\frac{3}{2})\ln\eta\right)}{\sinh\left((N+\frac{1}{2})\ln\eta\right)}\right], \tag{22}$$

$$\eta = \frac{\lambda+\sqrt{\lambda^2-4}}{2} \quad \text{and} \quad \lambda = 2 + \left(\frac{2\pi^2}{3} - 1\right)\frac{\sigma_w^2}{\sigma^2T^2}, \tag{23}$$

where we have taken the limit $\sigma_a \to 0$.

Figure 2 plots the bound for SNR = 5 dB (SNR defined as $10\log_{10}(1/\sigma^2)$), and for $\sigma_w/T \in \{0.005\%,\ 0.05\%,\ 0.5\%,\ 5\%\}$. Along with the bounds, we have plotted a horizontal line, whose height is h. As σ_w/T increases, we see that the shaping function $f(i)$ tends towards a constant, independent of i. The steady state bound can be rewritten as

$$h = \frac{\sigma\frac{\sigma_w}{T}}{\sqrt{K_1(4 + K_1\frac{\sigma_w^2}{\sigma^2T^2})}}, \tag{24}$$

where $K_1 = (2\pi^2/3) - 1$. For $\sigma_w^2/T^2 \ll \sigma^2$, which would be the case, for example, with $SNR \sim 5dB$ and $\sigma_w/T < 10\%$, we can approximate

$$h \approx K_2\frac{\sigma_w}{T}\sigma, \tag{25}$$

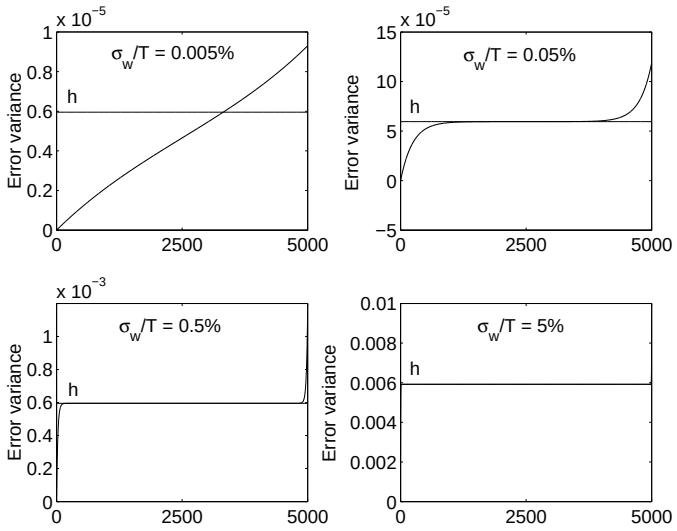

Figure 2: The lower bound on timing estimation error variance at SNR = 5.0 dB.

where $K_2 = 1/(2\sqrt{K_1})$. So, the steady state lower bound is linear in both σ_w/T and σ. If we look at the other extreme $\sigma_w^2/T^2 \gg \sigma^2$, we get

$$h \approx \frac{\sigma^2}{K_1}, \tag{26}$$

independent of σ_w/T. This behavior is illustrated in Figure 3.

Figure 3-(a) plots h *vs.* σ for σ_w/T going from 0.1 to 1 in steps of 0.1. The arrow shows the direction of increasing σ_w/T. Figure 3-(b) plots h *vs.* σ_w/T for σ from 0.5 to 5 in steps of 0.5. This is the case where we have $\sigma_w^2/T^2 \ll \sigma^2$, and h is linear in both σ and σ_w/T as given by (25). Figures 3-(c) and 3-(d) deal with the other extreme. In Figure 3-(c), we have h *vs.* σ for σ_w/T going from 0.03 to 0.3 in steps of 0.03, and in Figure 3-(d), we have h *vs.* σ_w/T for σ from 0.01 to 0.1 in steps of 0.01. We see that whenever $\sigma_w^2/T^2 \ll \sigma^2$, h is quadratic in σ and independent of σ_w/T, as given by (26).

Consider next the shaping function $f(i)$. We can expect this to be a non-decreasing function of the symbol index i, due to the random walk model being used for the timing jitter. Since we start off with perfect acquisition, we expect the variance to rise from 0, as is indeed the case. As i increases, we reach the steady state bound. But as we approach the end of the packet, we see a curious exponential deviation from the steady state value, governed by

$$\frac{[\boldsymbol{J}_1^{11}]^{-1}_{(N-j)(N-j)}}{T^2} - h \approx \frac{\sigma_w^2}{T^2}\frac{\eta^{2-2j}}{\eta^2-1}, \tag{27}$$

the shape of which is independent of N as shown in Figure 4, where SNR = 5 dB and $\sigma_w/T = 0.05\%$. As N increases, this end effect affects a smaller and smaller fraction of symbols. In other words, the steady state bound becomes more representative.

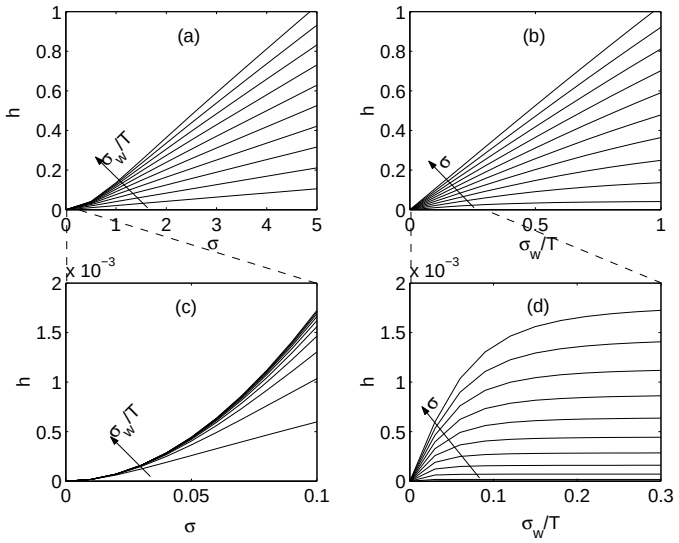

Figure 3: Steady state bound at the two extremes.

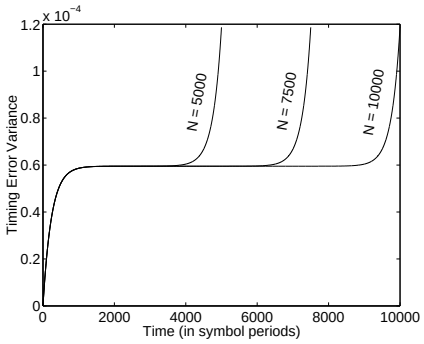

Figure 4: Shape of the end effect is independent of packet length.

Let r be the ratio of the variance of the last symbol to the steady state value. From (20) and (27), $r \approx 1 + \eta^{-1}$. From (23), $\eta \geq 1$. When $\sigma_w^2/T^2 \ll \sigma^2$, $\eta \approx 1$ and therefore, $r \approx 2$. When $\sigma_w^2/T^2 \gg \sigma^2$, $\eta \gg 1$ and therefore, $r \approx 1$. So, for any particular σ_w/T, as SNR ranges from $-\infty$ to ∞, r goes from 2 to 1. In other words, the end effect becomes less and less significant as SNR increases.

3 PLL-based timing recovery

The lower bound derived in the previous section applies to timing recovery systems in general. In practice, timing recovery systems usually involve phase locked loops, as described below.

For the system described by (1), at the receiver, we first have a front-end low-pass filter (with impulse response $p(t)/T$) to eliminate out-of-band noise from the received waveform

$y(t)$, to get the bandlimited waveform $r(t)$. Based on the estimates $\hat{\tau}_k$ of τ_k produced by a timing recovery system, the waveform $r(t)$ is then sampled at the timing instants $kT + \hat{\tau}_k$, producing samples $r_k = \sum_l a_l h(kT + \hat{\tau}_k - lT - \tau_l) + n_k$, where $\{n_k\}$ are *i.i.d.* $\mathcal{N}(0, \sigma^2)$.

For simplicity and because our model has no frequency offset component, we restrict ourselves to a first-order PLL, which updates its estimate of τ_k according to

$$\hat{\tau}_{k+1} = \hat{\tau}_k + \alpha \hat{\epsilon}_k, \tag{28}$$

where α is the PLL gain, and where $\hat{\epsilon}_k$ is the detector's estimate of the estimation error $\epsilon_k = \tau_k - \hat{\tau}_k$. The widely used Mueller and Müller timing-error detector (TED) generates this estimate according to [10]:

$$\hat{\epsilon}_k = \frac{3T}{16}(r_k \hat{d}_{k-1} - r_{k-1}\hat{d}_k), \tag{29}$$

where $\hat{d}_k$ is an estimate of $d_k = a_k - a_{k-2} \in \{0, \pm 2\}$, typically obtained by a memoryless three-level quantization of r_k. The constant $3T/16$ ensures that there is no bias at high SNR, so that $E[\hat{\epsilon}_k] = \epsilon_k$.

Performance of the Mueller and Müller TED can be improved by using soft estimates $\tilde{d}_k$ in place of hard estimates $\hat{d}_k$ in (29) [11]. Choosing $\tilde{d}_k = E[d_k | r_k]$ leads to a memoryless soft slicer of the form

$$\tilde{d}_k = \frac{2\sinh(2r_k/\sigma^2)}{\cosh(2r_k/\sigma^2) + e^{2/\sigma^2}}. \tag{30}$$

In the preceding analysis of Section 2, we assumed a system with no coding and let the symbols $\{a_k\}$ be independent. In [5], an error-control coded system was considered, where turbo equalization was implemented at the receiver. In this setting, better soft decisions about the symbols, coming from the turbo equalizer, were fed to the TED at the end of each turbo iteration, and the new timing estimates were used to refine the observations. This process continued, and as iterations progressed, the SNR requirement reduced significantly.

4 Comparing PLL-based structures and the CRB

Instead of feeding the TED decisions about the received symbols, if we allow it to have access to the actual transmitted symbols $\{d_k\}$, then we have a trained PLL, where the TED generates $\hat{\epsilon}_k$ according to

$$\hat{\epsilon}_k = \frac{3T}{16}(r_k d_{k-1} - r_{k-1} d_k). \tag{31}$$

The performance of trained PLL gives a heuristic lower bound for the performance of receiver structures that use the PLL for timing recovery. In Figure 5, we plot the steady state CRB and the performance of the trained PLL for the following system parameters: $\sigma_w/T = 0.5\%$, block length $N = 500$, and the PLL performance being averaged over 1000 trials. As seen in the figure, the performance of the PLL is a strong function of the gain

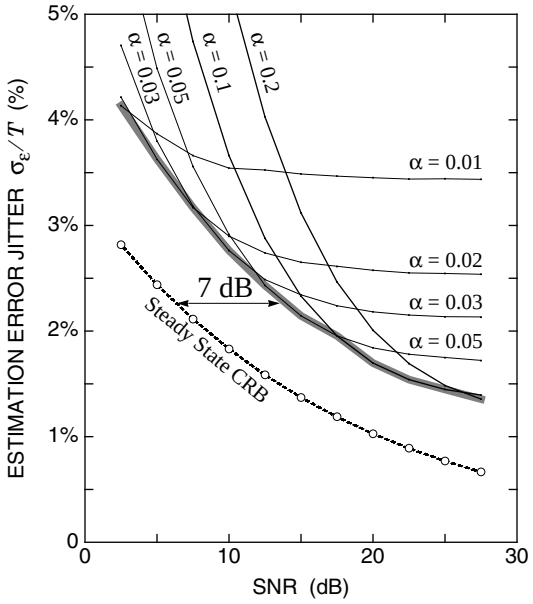

Figure 5: Trained PLL and the Cramér-Rao Lower Bound.

parameter α, and therefore, α has to be optimized for each SNR. The PLL error variance is plotted for various values of α. Taking the minimum of the error variance over all α gives us the best performance we can expect using the trained PLL. We see that the trained PLL is about 7 dB away from the CRB.

This gap of 7 dB has to be put in perspective by the fact that the CRB is not attainable in this case. For the CRB to be attainable, the *a posteriori* density $f_{\boldsymbol{r}|\boldsymbol{\theta_2}}(\boldsymbol{r}|\boldsymbol{\theta_2})$ needs to be Gaussian, which is not the case here.

It would be of interest to get the actual error variance of the trained ML timing estimator and the joint ML estimator that estimates both the timing offsets and the transmitted symbols. The first one would give us the penalty with respect to the CRB due to the fact that the required *a posteriori* density is not Gaussian, while the second would give us the penalty due to the fact that we do not know the transmitted symbols at the receiver. A better indicator of the effectiveness of PLL-based receivers would then be the gap from the joint ML estimator. These are open problems, and future research should deal with these.

Acknowledgement: Research supported by the National Storage Industries Consortium. Aravind Nayak would like to thank Renato Lopes and Andrew Thangaraj for the discussion on inverting the Fisher information matrix.

References

[1] C. Berrou, A. Glavieux, and P. Thitimajshima, "Near Shannon Limit Error-correcting Coding and Decoding: Turbo-codes," *Proc. of the IEEE International Conference on*

Communications 1993, vol. 2, pp. 1064-1070, May 1993.

[2] R. Gallager, "Low-Density Parity-Check Codes," *IRE Transactions on Information Theory*, vol. IT-8, pp. 21-28, January 1962.

[3] D. Raphaeli and Y. Zarai, "Combined Turbo Equalization and Turbo Decoding," *Proc. of the IEEE Global Telecommunications Conference 1997*, vol. 2, 639-643, November 1997.

[4] C. Georghiades and D. Snyder, "The Expectation Maximization Algorithm for Symbol Unsynchronized Sequence Detection," *IEEE Transactions on Communication*, vol. 39, pp. 54-61, January 1991.

[5] A. Nayak, J. Barry, and S. McLaughlin, "Joint Timing Recovery and Turbo-Equalization for Partial Response Channels," *Proc. of the 2002 IEEE International Magnetics Conference*, BR03, May 2002.

[6] H. Meyr, M. Moeneclaey, and S. Fechtel, *Digital Communication Receivers: Synchronization, Channel Estimation, and Signal Processing*, pp. 54-58, John Wiley and sons, 1998.

[7] P. Aziz and S. Surendran, "Symbol Rate Timing Recovery for Higher Order Partial Response Channels," *IEEE Journal on Selected Areas on Communications*, vol. 19, no. 4, pp. 635-648, April 2001.

[8] H. Van Trees, *Detection, Estimation, and Modulation Theory*, vol. 1, pp. 72-85, John Wiley and sons, 1968.

[9] L. Scharf, *Statistical Signal Processing: Detection, Estimation, and Time Series Analysis*, pp. 230, Addison-Wesley, 1991.

[10] K. Mueller and M. Müller, "Timing Recovery in Digital Synchronous Data Receivers," *IEEE Transactions on Communication*, vol. com-24, no. 5, pp. 516-31, May 1976.

[11] A. Nayak, J. Barry, and S. McLaughlin, "Timing Recovery at Low SNR," *NSIC Quarterly Report*, March 2000.