

Signal Parameter Estimation in the Presence of Timing Noise

Julius Kusuma and Vivek K Goyal

Signal Transformation and Information Representation Group

Massachusetts Institute of Technology

Cambridge, MA 02139

Email: {kusuma, vgoyal}@mit.edu

Abstract— We consider the problem of estimating the parameters of a signal when the sampling instances are perturbed by signal-independent timing noise. The classical techniques consider timing noise to induce a signal-independent additive white Gaussian noise term on the sample values. We reject this simplification of the problem and give alternative methodologies. For the problem of delay estimation when the pulse shape and amplitude of the signal are known, we give an iterative algorithm that shows superior performance compared to the traditional method which relies on maximizing the cross-correlation.

I. INTRODUCTION

Both uniform and non-uniform sampling of analog signals have been studied extensively, including approximation and noise robustness properties [1]. This theory builds upon the celebrated sampling theorem of Shannon, Nyquist and Kotelnikov, which applies to the class of *bandlimited signals*. The recent work of Vetterli *et al.* [2] provided examples of nonbandlimited signals that can be recovered perfectly from uniform samples. Though not bandlimited, the signals they could recover are specified by a finite number of degrees of freedom per unit time, a property termed *finite rate of innovation* (FRI) in [2]. For example, a signal that is composed of K scaled and delayed copies of a known pulse is completely determined by K pairs of amplitudes and delays $\{a_k, t_k\}_{k=1}^K$. Many classes of signals and corresponding sampling kernels have been considered and algorithms for their perfect reconstruction in the exact and noiseless cases have been proposed [3], [4].

In almost any real application the actual sampling instances are not the same as the ideal, desired sampling instances. We call this difference *timing noise*. In many current and emerging applications such as in wideband communication systems and mechanically-actuated probes this timing noise is not an insignificant source of uncertainty. An analog-to-digital converter (ADC) is triggered by a clock signal which in turn is often generated from a lower-frequency clock signal from a dedicated clock circuitry. Hence the timing noise may have a strong structure which we may be able to estimate. In electronic system design, timing noise is often modeled as additional additive white Gaussian noise (AWGN) [5]. For example, one may say, “*The 74-ps jitter means that we have a 2 dB loss.*” However, the effect of timing noise in general is neither additive, white nor Gaussian. Moreover, knowledge

of the input signal may be used to estimate the timing noise. This is the key insight of this work. We focus on signals with a known pulse shape but unknown delay, and will demonstrate that we can improve the performance of time delay estimation by modeling timing noise explicitly.

II. PROBLEM STATEMENT

Let the sequence z_n represent the *timing noise* in the clock signal: it affects the instances at which the input signal $x(t)$ is evaluated. Let $s_n = nT + z_n$ be these instances. Then the observed samples from the ADC are given by:

$$y_n = x(s_n) = x(nT + z_n), \quad n = 0, 1, \dots, N - 1. \quad (1)$$

We say that the timing noise z_n is additive in the *timing domain*, which is the argument of the signal $x(\cdot)$. Suppose that we obtain N samples of (1). In this paper we will focus on the canonical problem of signal delay estimation when the pulse shape is perfectly known to the receiver. That is, we restrict our input signal to be $x(t) = g(t - \tau)$, where τ is the unknown delay and $g(t)$ is the known pulse shape. This signal is completely described by one parameter, namely τ , and hence its rate of innovation in the aperiodic case is zero.

In this document we describe a few models for timing noise and describe estimation algorithms for signal parameter estimation. We focus on time delay estimation of the input signal $x(t)$, but first we illustrate how knowledge of the structure of the input signal $x(t)$ can be used for estimating the jitter. The most interesting case will be when AWGN is present in addition to timing noise. In this case, let w_n be the AWGN:

$$y_n = x(nT + z_n) + w_n, \quad n = 0, 1, \dots, N - 1. \quad (2)$$

In contrast to the timing noise term, we say that the w_n is additive in the *observation domain*.

The problem of sampling in the presence of timing noise is not new. Balakrishnan [6] considered bandlimited signals and signals consisting of deterministic line spectral terms, and gave an interpolation technique which minimizes the reconstruction error both for white and correlated jitter. Butzer studied the effect of timing noise on bandlimited signals in [7], [8]. In his work he considered the \mathcal{L}_2 error between the original signal and its sinc interpolation after sampling in the presence of bounded timing noise. He derived an upper bound for the

reconstruction error that is dependent on the supremum of the first order derivative of the input signal: the steeper the slopes of the signal then the stronger the effect of timing noise on the sampled values. We will see later that in the case of parameterized signals and in the presence of additive noise, this is not the case. The most similar previous work is that of Narasimhamurti and Awad [9]. The authors considered the problem of estimating the phase of a sinusoid sampled using a noisy clock with accumulated jitter when the frequency and amplitude are known. They proposed a state-space approach via writing the relationship between successive sample phase using trigonometric identities.

III. PARAMETER ESTIMATION IN THE PRESENCE OF TIMING NOISE: DELAY ESTIMATION

Our main contribution in this section, and in its generalizations, is the use of the notion of *consistency*. We define consistency in parameter estimation in the presence of timing noise as follows.

Definition 1 (Consistency with timing noise only). Let parameter vector θ parameterize the signal $x_\theta(t)$, and let $\{y_n\}_{n=0}^{N-1}$ be samples taken at times $\{nT + z_n\}_{n=0}^{N-1}$. Estimates of the unknown parameter vector θ and timing noise vector $\{z_n\}_{n=0}^{N-1}$ are said to be *consistent with the observations* when $y_n = x_\theta(nT + z_n)$ for $n = 0, 1, \dots, N - 1$.

Throughout this paper, when some of the variables are given we also use \mathcal{C} to denote the values of the free variables that are consistent with the given ones. For example, $(\mathbf{z}, \theta) \in \mathcal{C}$ means values of (\mathbf{z}, θ) that are consistent with a given \mathbf{y} .¹ It will be made clear by the context of the statement.

In this section we introduce a model in which every clock tick that drives the sampling signal of the ADC is subjected to jitter. The signal model is given by:

$$y_n = g(nT + z_n - \tau), \quad n = 0, 1, \dots, N - 1.$$

In the above, the sequence z_n is white Gaussian. See [10, Sec. 3.5.4] for an overview of the frequency domain characterization of oscillator noise. Following Definition 1 for delay estimation, the signal is given by $x(t) = g(t - \tau)$, basically a delayed version of a known pulse shape $g(\cdot)$ at delay τ . Given the observations $\{y_n\}, n = 0, \dots, N - 1$, we wish to estimate τ . The only noise in the system is the jitter timing noise. An estimate of delay and timing noise of the signal $\{\tau; z_0, \dots, z_{N-1}\}$ is consistent with observations $\{y_n\}$ for a given pulse shape $g(t)$ when $y_n = g(nT + z_n - \tau)$ for $n = 0, 1, \dots, N - 1$.

Instead of modeling the effect of sampling clock jitter as an AWGN term, we propose to find the estimate $\hat{\tau}$ that minimizes the *jitter noise* and gives a *consistent* solution following Definition 1. Note that for any given estimate of the delay τ there may be more than one sequence of jitter estimate that produce a consistent reconstruction. We represent all the consistent sequences via the ambiguity set.

¹Bold letters are vectors of samples of the same name, e.g., $\mathbf{z} = (z_0, z_1, \dots, z_{N-1})$.

Definition 2. Consider the problem of estimating the delay of a signal consisting of a known pulse shape $g(t)$. Given a sequence of samples y_n we define the ambiguity set for each n as follows:

$$\mathcal{A}_g[n] = \{x | g(x) = y_n\}. \quad (3)$$

The use of this set is to allow treatment of pulse shapes $g(\cdot)$ which are not invertible maps. When the reconstruction algorithm makes a mistake in choosing the correct member of the ambiguity set, we refer to this error as the *ambiguity error*.

Consider the maximum-likelihood (ML) estimate of τ :

$$\hat{\tau}_{ML} = \arg \max_{\tau} \ln Pr(\mathbf{y}; \tau). \quad (4)$$

The ML estimation problem given in (4) is difficult to evaluate for general pulse shapes. We follow the Expectation-Maximization (EM) approach [11], and simplify the computation using the idea of *consistency* introduced above instead. The EM algorithm has been used to solve missing data problems, and estimation problems which are difficult to express or compute explicitly. It can be shown to give the ML solution under mild conditions. Following EM terminology we consider the observations y_n to be the *incomplete data*, and the unobserved timing noise term z_n to be a hidden variable which together with y_n form the *complete data* set. Therefore instead of the ML problem of (4) we use the complete data set and consider maximizing $Pr(\mathbf{y}, \mathbf{z}; \tau)$ over τ instead. Of course, the hidden variables z_n are not available. However, both the timing noise $\{z_n\}$ and τ are in the timing domain. If we restrict ourselves only to values of \mathbf{z} and τ which are consistent, then the values of \mathbf{z} are completely determined by τ unless the signal pulse shape contains constant terms. Hence given the ML estimate of one we can derive the ML estimate of the other term. This observation means that the solution to the above can also be obtained by optimizing over \mathbf{z} . We obtain:

$$\hat{\tau} = \arg \max_{\tau, \mathbf{z}} \ln Pr(\mathbf{y} | \mathbf{z}; \tau) Pr(\mathbf{z}; \tau) \quad (5)$$

Note that $Pr(\mathbf{y} | \mathbf{z}; \tau)$ restricts us to the consistent set, or in other words,²

$$Pr(\mathbf{y} | \mathbf{z}; \tau) = \begin{cases} 1, & (\mathbf{y}, \mathbf{z}, \tau) \in \mathcal{C}; \\ 0, & (\mathbf{y}, \mathbf{z}, \tau) \notin \mathcal{C}. \end{cases}$$

Since the jitter z_n is assumed to be Gaussian, we wish to find the consistent estimate that minimizes $\|\mathbf{z}\|^2$:

$$\min_{\tau, \mathbf{z}} \frac{\|\mathbf{z}\|^2}{2\sigma_z^2} \quad s.t. \quad y_n = g(nT + z_n - \tau), \forall n. \quad (6)$$

When the pulse shape does not contain any constant sections, then we only have to perform the above optimization over τ . When g is invertible, let $\hat{a}_n = g^{-1}(y_n)$. In this case the term $a_n = nT + z_n - \tau$ is completely determined by y_n . Given any estimate of \mathbf{a} , the optimal choice of τ is always given by

² $Pr(a|b; c)$ means the conditional probability of a conditioned on a random variable b and a non-random parameter c .

one that is the mean of the values of $\{a_n - nT\}$. Therefore we need to only optimize over τ , or over \mathbf{z} . The constraints given above are the *consistency constraints*. From here it is straightforward to obtain $\hat{\tau}$ that minimizes $\|\mathbf{z}\|^2$:

$$\hat{\tau}_{ML} = \frac{1}{N} \sum_n (a_n - nT). \quad (7)$$

However, in general the pulse shapes of interest are not invertible maps. For example, the Gaussian pulse is not an invertible map. Therefore we need to resolve the ambiguity in estimating the noisy sampling instances. For each n we have to consider all possible ambiguous values of a_n which yield the same y_n . Then we have to find the minimization of $\|\mathbf{z}\|^2$ over all possible ambiguous sets given by the cardinal product $\mathbf{a} \in \vec{\mathcal{A}} = \otimes_n \mathcal{A}_g[n]$.

In summary, a brute-force method for solving the optimization problem of 6 is given by:

- 1) For each n , compute $\mathcal{A}_g[n]$ from y_n , and obtain $\vec{\mathcal{A}} = \otimes_n \mathcal{A}_g[n]$.
- 2) For each $\mathbf{a} \in \vec{\mathcal{A}}$, compute

$$\hat{\tau}(\mathbf{a}) = \frac{1}{N} \sum_n (a_n - nT), z_n(\mathbf{a}) = a_n - (nT - \hat{\tau}).$$

- 3) Choose $\hat{\tau}(\mathbf{a})$ which minimizes $\|\mathbf{z}\|_2^2$.

In most cases, the ambiguity sets \mathcal{A}_n will consist only of discrete points, and a simplified solution is possible. In that case, the solution can be obtained by simply choosing which member of each ambiguity set gives the best overall solution. Recall that any admissible a_n is of form $a_n = nT + z_n - \tau$. When z_n is small, this is simply a line of slope T in the variable n , and its intercept is the delay term $-\tau$. A simpler algorithm can be used that is an iteration between *line regression* – fixing the slope to always be nT – and nearest-neighbor association or quantization, that solves the ambiguity problem. The standard line regression is optimal in the mean-square sense, and therefore optimal for white Gaussian timing noise z_n . A sample result is given in Fig. 1.

The proposed system is compared to the pulse fitting procedure which seek to minimize the \mathcal{L}_2 norm by maximizing the cross-correlation between the estimate and the measurement. This latter system is a gradient-search algorithm, and we allow multiple initial conditions and pick the best result. We emulate the UWB system of the MIT MTL lab, in which a 2 ns Gaussian pulse located near the center of a 100 ns frame is subjected to a Gaussian jitter of standard deviation of 75 ps and 200 ps. There are 200 samples taken within the frame. The results are given in the table below.

	$\sigma_z = 0.2$ -ns no AWGN	$\sigma_z = 0.75$ -ps no AWGN	$\sigma_z = 0.75$ -ps 10-dB AWGN
consistency	0.0109	0.0031	0.9407
grad-search LS	0.0625	0.0557	0.1064

While the new algorithm consistently outperforms the conventional method, it is not robust to AWGN. A small amount of AWGN may have an adverse effect on the performance depending on the pulse shape. We will address this by extending our model of consistency.

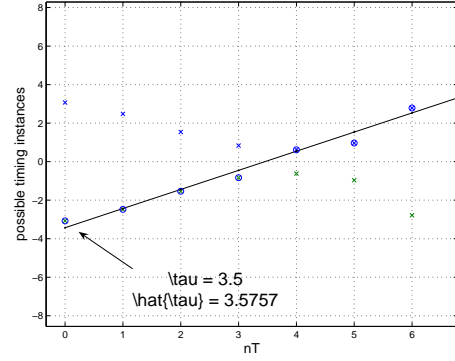


Fig. 1. Sampling at noisy instances, with only white timing noise. In this set, $N = 16$, Gaussian pulse with width 2, $T = 2$, $\tau = 3.4$. The X's are the ambiguity set members, and the O's are those that are estimated to be the true sampling points. The line is the regression, and τ is given by the intersection of the line with the y -axis.

A. Example signal: constant slope

Consider $g(t) = a \cdot t$. We have that $x(t) = g(t - \tau) = a(t - \tau)$. This pulse shape is invertible, and intuitively we can see that it simply converts the problem of timing noise into the well-known problem of additive noise. Let $\mathcal{G}(a, b^2)$ be the evaluation of a Gaussian PDF of mean 0 and variance b^2 at a . It can be shown that the distribution is given by:

$$Pr(y_n; \tau) = \mathcal{G}(x(nT), (a\sigma_z)^2), \quad (8)$$

$$\ln Pr(\mathbf{y}; \tau) = \text{constant} - \sum_{n=0}^{N-1} \frac{(y_n - a(nT - \tau))^2}{2a^2\sigma_z^2}. \quad (9)$$

From here it is then easy to see that the ML estimator and the Cramér-Rao bound are given by:

$$\hat{\tau}_{ML} = \arg \min_{\tau} \sum_{n=0}^{N-1} (y_n - (anT - a\tau))^2 = \frac{\sum_{n=0}^{N-1} (y_n - anT)}{aN},$$

$$\text{var}(\hat{\tau}) \geq \frac{\sigma_z^2}{N}.$$

Note that in this case the value of the slope a does not affect the performance of the ML estimator; in the case where there is AWGN we will see that the slope in fact determines the performance of the estimator.

IV. PULSE DELAY ESTIMATION IN THE PRESENCE OF DRIFT TIMING NOISE

In this section we first consider the delay estimation problem in which the observations are given by

$$y_n = x(nT + \Delta n) = g(nT + \Delta n - \tau).$$

In this case, the timing noise consists of a pure drift term determined by Δ . Drift occurs when the clock frequency between two references are not exactly the same. Suppose that the drift term Δ is relatively small; in most electronic applications the drift can be guaranteed to be below 100 ppm. However, many crystal oscillators are cut adaptively, *i.e.*, not

cut to dimension but to the desired frequency, and hence their variance in frequency as they age may become large.

From the set of measurements $\{y_n\}$, we can again find the ambiguity sets \mathcal{A}_n as defined above. The problem now becomes a clustering and line regression problem. We wish to find a set of points from the ambiguity set such that a line regression of these points will yield a line with slope close to T , because we know that the drift term is small relative to T . We implement a simple iterative routine which repeats the following two steps, except that the slope is no longer T but is instead $(1 + \Delta)T$:

- 1) From each ambiguity set indexed in n , find the point closest to the current estimate of the line slope and offset.
- 2) From the set of nearest points, update the estimate of the slope and offset.

In the presence of drift and white Gaussian jitter, the algorithm above can be used with no modification. After all, line regression is a least-squares fit, which is optimal for additive Gaussian noise. We obtain results similar to that shown in Fig. 1.

V. PARAMETER ESTIMATION IN THE PRESENCE OF TIMING NOISE AND AWGN: DELAY ESTIMATION

In the previous part we have seen that one problem with the consistency-based algorithm in the presence of AWGN is that a small amount of additive noise may lead to a very different answer to the ambiguity set \mathcal{A}_n , and hence give a very poor estimate $\hat{\tau}$. Extending our model to include AWGN, the observations are given by

$$y_n = x_\theta(nT + z_n) + w_n, \quad n = 0, 1, \dots, N - 1. \quad (10)$$

We include the additive noise term w_n in our definition of consistency.

Definition 3 (Consistency with timing noise and AWGN).

Let parameter vector θ parameterize the signal $x_\theta(t)$, and let $\{y_n\}_{n=0}^{N-1}$ be samples taken at times $\{nT + z_n\}_{n=0}^{N-1}$. Estimates of the unknown parameter vector θ , timing noise vector $\{z_n\}_{n=0}^{N-1}$ and AWGN $\{w_n\}_{n=0}^{N-1}$ are said to be *consistent with the observations* when $y_n = x_\theta(nT + z_n) + w_n$ for $n = 0, 1, \dots, N - 1$.

Let \mathcal{C} denote the set of consistent estimates.

Directly expressing $Pr(\mathbf{y}; \theta)$ is difficult in general. We again follow the Expectation-Maximization algorithm (EM) [11] approach by working with the complete data set:

$$\arg \max_{\theta, \mathbf{z}} \ln Pr(\mathbf{y}, \mathbf{z}; \theta) = \arg \max_{\theta, \mathbf{z}} \ln Pr(\mathbf{w}|\mathbf{z}; \theta) Pr(\mathbf{z}; \theta). \quad (11)$$

The computation of (11) can be done either in a nested manner or iteratively between the two variables τ and \mathbf{z} . For the problem of delay estimation, we are concerned only with the delay τ as the unknown parameter. The model is then given by

$$y_n = g(nT + z_n - \tau) + w_n, \quad n = 0, 1, \dots, N - 1. \quad (12)$$

Similarly to the case where only timing noise is present, we can obtain an ML estimate of τ from $\{z_n\}$ and vice-versa. An estimate of delay, timing noise and additive noise of the signal $\{\tau; z_0, \dots, z_{N-1}; w_0, \dots, w_{N-1}\}$ is consistent with observations $\{y_n\}$ for a given pulse shape $g(t)$ if and only if $y_n = g(nT + z_n - \tau) + w_n$ for $n = 0, 1, \dots, N - 1$.

Conditioned on τ the estimates of $\{w_n, z_n\}$ can be computed separately for each n , based on the observations y_n . Moreover, for each n a consistent w_n is completely determined by fixing τ and z_n . Indeed, using the model we can set:

$$w_n = y_n - g(nT + z_n - \tau), \quad n = 0, 1, \dots, N - 1. \quad (13)$$

We will take advantage of this in maximizing the joint probability $Pr(w_n, z_n|y_n; \tau)$ by writing:

$$\begin{aligned} Pr(w_n, z_n|y_n; \tau) &= Pr(w_n|z_n, y_n; \tau) Pr(z_n|y_n; \tau), \quad (14) \\ Pr(w_n|z_n, y_n; \tau) &= Pr(w_n = y_n - g(nT + z_n - \tau)|z_n, y_n; \tau). \end{aligned}$$

Consider computing the estimate of the hidden variables $\{w_n, z_n\}$ and the unknown parameter τ iteratively. In essence we iterate between:

- **Estimating hidden variables:**

$$(\hat{\mathbf{w}}_{k+1}, \hat{\mathbf{z}}_{k+1}) = \arg \max_{\mathbf{w}, \mathbf{z}} Pr(\mathbf{w}, \mathbf{z}|\mathbf{y}, \hat{\tau}_k).$$

- **Estimating unknown parameter:**

$$\hat{\tau}_{k+1} = \arg \max_{\theta} Pr(\mathbf{y}|\hat{\mathbf{w}}_k, \hat{\mathbf{z}}_k; \tau).$$

In the estimation of the hidden variables we can use the insights of the previous part, namely that the estimation can be distributed in n , and that w_n is uniquely determined by τ, z_n following (13). We obtain (15).

$$\begin{aligned} \arg \max_{(w_n, z_n)} Pr(w_n, z_n|y_n, \hat{\tau}_k) & \quad (15) \\ &= \arg \max_{(w_n, z_n)} \ln Pr(z_n|y_n; \tau) Pr(w_n|z_n, y_n; \tau) \\ &= \arg \max_{z_n} \frac{\|z_n\|^2}{\sigma_z^2} + \frac{\|y_n - g(nT + z_n - \tau)\|^2}{\sigma_w^2}. \end{aligned}$$

Estimation of the unknown delay is simpler: given \mathbf{z} the only uncertainty in \mathbf{y} is the AWGN term. Hence,

$$\arg \max_{\tau} Pr(\mathbf{y}|\hat{\mathbf{w}}_k, \hat{\mathbf{z}}_k; \tau) = \arg \min_{\tau} \|\mathbf{y} - g(\mathbf{n}T + \mathbf{z} - \tau)\|^2, \quad (16)$$

where \mathbf{n} denotes the vector $[0, 1, \dots, N - 1]^T$ and g denotes the appropriate N -fold product of the pulse shape. The optimization in (16) has an intuitive interpretation: given our knowledge of the timing noise we can construct a template pulse shape which we can then use for maximizing the cross-correlation with the sampled signal.

We apply the optimization procedures of (15) and (16) iteratively. For both steps we will use a steepest-descent approach. The performance is strictly better than that of using only correlation, because it also allows for compensation in the timing domain. It can be shown the the gradients are:

$$\frac{\partial}{\partial \tau} (\cdot)_n = 2(y_n - g(nT + z_n - \tau))g'(nT + z_n - \tau).$$

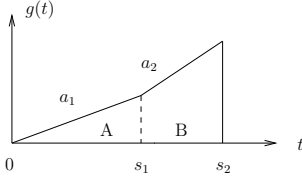


Fig. 2. A simple pulse shape with two different slopes.

$$\frac{\partial}{\partial z_n}(\cdot)_n = \frac{2z_n}{\sigma_z^2} + \frac{2}{\sigma_w^2}(y_n - g(nT + z_n - \tau))g'(nT + z - \tau).$$

The performance evaluation of this method is promising: in all cases, the performance of the FRI-based technique is superior to that of standard correlation-based method as shown below:

$N = 8, T = 1, \sigma_g = 1$				
$\hat{\tau}$	$\sigma_z = 0.1$	$\sigma_z = 0.01$	$\sigma_z = 0.1$	$\sigma_z = 0.01$
RMS	$\sigma_w = 0.1$	$\sigma_w = 0.1$	$\sigma_w = 0.01$	$\sigma_w = 0.01$
FRI	0.257	0.303	0.282	0.192
Xcorr	0.345	0.394	0.307	0.315
$N = 16, T = 0.5, \sigma_g = 1$				
$\hat{\tau}$	$\sigma_z = 0.1$	$\sigma_z = 0.01$	$\sigma_z = 0.1$	$\sigma_z = 0.01$
RMS	$\sigma_w = 0.1$	$\sigma_w = 0.1$	$\sigma_w = 0.01$	$\sigma_w = 0.01$
FRI	0.170	0.177	0.211	0.065
Xcorr	0.298	0.339	0.281	0.281

A. Analytical result: constant slope

Recall the basic example in which the pulse shape is given by $g(t) = a \cdot t$. Clearly, this function is invertible. It can easily be derived that

$$Pr(y_n; \tau) = \mathcal{G}(y_n - a(nT - \tau), \sigma_w^2 + a^2\sigma_z)$$

Hence the ML estimate and Cramér-Rao bound are simply given by:

$$\hat{\tau}_{ML} = \frac{1}{N} \sum_{n=0}^{N-1} \left(\frac{x_n}{a} - nT \right), \quad \text{var}(\hat{\tau}) \geq \frac{\sigma_z^2 + \frac{\sigma_w^2}{a^2}}{N}. \quad (17)$$

In contrast to the case where only white timing noise is present, in this case it is clear that the slope a determines the relative effect of the white timing noise versus the additive white noise: the steeper the slope, the better the performance. In estimating the delay, the slope does not change the performance in the case of pure white timing noise since the desired parameter and the nuisance are both in the timing domain. However, when AWGN is present the effect of the AWGN in the observation domain is independent of the slope by definition, whereas the delay term τ from the timing domain is amplified by a . This result should be compared with the non-parametric analytical result of Butzer [7].

B. Optimizing the pulse shape for delay estimation

In order to gain an insight into the best design of pulse shapes, consider a pulse consisting of two piecewise-linear regimes, denoted A and B , as shown in Fig. 2. The slopes at the different regimes are a_1 and a_2 respectively. Without loss of generality, let $a_2 \geq a_1$. Let n_A, n_B be the number of samples that fall within each regime.

The ML estimator can be derived from the log-likelihood function and is a linear combination of the samples $\{y_n\}_{n=0}^{N-1}$. The Cramér-Rao bound is given by

$$\text{var}(\hat{\tau}) \geq \left(\frac{n_A}{\sigma_z^2 + \frac{\sigma_w^2}{a_1^2}} + \frac{n_B}{\sigma_z^2 + \frac{\sigma_w^2}{a_2^2}} \right)^{-1}. \quad (18)$$

The result of (18) can be used to give us a very important insight into the design of pulse shapes. Suppose we wish to compare the performance of the pulse given in Fig. 2 and the triangular pulse of the previous part. For fairness, we will fix the same N, T , the signal energy, and the domain ambiguity of both pulses. Let the signal energy of a pulse $g(t)$ be denoted by $\mathcal{E} = \int |g(t)|^2 dt$. Then it can be shown that for all values of $N, T, \mathcal{E}, \sigma_z, \sigma_w$ the minimum value of (18) is achieved by $s_1 = 0$, because it is strictly decreasing in s_1 by the assumption that $a_2 \geq a_1$, and that for any fixed s_1 it is strictly decreasing in a_2 and a_1 . This means that the triangle pulse is optimal in the sense that it gives the lowest Cramér-Rao bound for the estimate of pulse delay, which we know can be achieved with the estimator derived from the log-likelihood function.

Theorem 1. Optimal pulse shape in the presence of white Gaussian timing and additive noise: *For pulse delay estimation in the presence of white Gaussian timing noise and white Gaussian additive noise, the triangle pulse is optimal for a given N, T, \mathcal{E} . The Cramér-Rao bound is given in (18).*

VI. PULSE DELAY ESTIMATION IN THE PRESENCE OF DRIFT, JITTER AND AWGN

We now consider the problem where there is a drift term and AWGN. The model is given by:

$$y_n = g(nT + n\Delta) + w_n, \quad n = 0, 1, \dots, N-1. \quad (19)$$

We solve the estimation problem by considering both τ and Δ as the desirable, unknown parameters.

$$\arg \max_{(\tau, \Delta)} \ln Pr(\mathbf{y}; \tau, \Delta). \quad (20)$$

Following the idea developed previously, we consider the additive white noise term $\{w_n\}$ as a hidden variable. From (19) we have that when τ and Δ are fixed, then w_n is completely determined by $w_n - x_\theta(nT + n\Delta)$. Moreover, for fixed Δ the problem of finding τ is simply the classical delay estimation in the presence of AWGN using a new template pulse.

Since the only source of noise in this problem is the AWGN term w_n , the problem is simply a minimum mean-square error. The optimization problem is given by

$$\begin{aligned} \arg \max_{\Delta, \tau} P(\mathbf{y}; \Delta, \tau) &= \arg \max_{\Delta, \tau \in \mathcal{C}} \ln P(\mathbf{y}; \Delta, \tau) \\ &\approx \arg \min_{\Delta, \tau} \|\mathbf{w}\|_2^2 \quad \text{s.t.} \quad y_n = g(nT + \Delta \cdot n - \tau) + w_n \\ &= \arg \min_{\Delta, \tau} \sum_n \|y_n - g(nT + \Delta \cdot n - \tau)\|^2. \end{aligned}$$

In other words, using the approximation via consistency above, the estimator is a line regression in the space of Δ and τ which

minimizes the \mathcal{L}_2 -distance between the reconstruction signal and the observation.

One method of implementing the optimization is via a gradient search iterating between updating $\hat{\tau}$ and updating $\hat{\Delta}$. It can be shown that the gradients are:

$$\frac{\partial}{\partial \tau}(\cdot)_n = 2(y_n - g(nT + \Delta \cdot n - \tau))g'(nT + \Delta \cdot n - \tau).$$

$$\frac{\partial}{\partial \Delta}(\cdot)_n = 2n(g(nT + \Delta \cdot n - \tau) - y_n)g'(nT + \Delta \cdot n - \tau).$$

The new proposed method is far superior to the classical method which considers only cross-correlation with a template pulse shape without compensating for timing noise. The performance is shown in below. The drift is uniformly distributed between $[-\Delta_r, +\Delta_r]$. The new system significantly outperforms the conventional method which uses only correlation. This is simply because the drift induces a warping of the template signal.

$N = 8, T = 1, \sigma_g = 1$				
	$\Delta_r = 0.1$ $\sigma_w = 0.1$	$\Delta_r = 0.01$ $\sigma_w = 0.1$	$\Delta_r = 0.1$ $\sigma_w = 0.01$	$\Delta_r = 0.01$ $\sigma_w = 0.01$
Δ RMS FRI	0.0057	0.0056	0.0046	0.0431
τ RMS FRI	0.0224	0.0218	0.0209	0.0160
τ RMS Xcorr	0.7007	0.5613	0.1956	0.1698

Further extending our model, we now consider pulse delay estimation in the presence of drift, jitter and AWGN. The maximum-likelihood optimization is written as:

$$\arg \max_{\tau, \Delta} Pr(\mathbf{y}; \tau, \Delta).$$

Following our previous approach, we can attempt instead to estimate the hidden variables explicitly:

$$\begin{aligned} & \arg \max_{\tau, \Delta} Pr(\mathbf{y}, \mathbf{w}, \mathbf{z}; \tau, \Delta) \\ &= \arg \max_{\tau, \Delta} \ln Pr(\mathbf{y}; \tau, \Delta | \mathbf{w}, \mathbf{z}) Pr(\mathbf{w}) Pr(\mathbf{z}) \\ &= \arg \min_{\tau, \Delta} \frac{\|\mathbf{z}\|^2}{\sigma_z^2} + \frac{\|\mathbf{w}\|^2}{\sigma_w^2}, \quad (\tau, \Delta) \in \mathcal{C}. \end{aligned}$$

In the above, we have taken advantage of the independence across n and the fact that in the set of consistent estimates w_n is completely determined by τ, Δ, z_n . As a first attempt, we simply use the algorithm used in the previous section - which takes into account only drift and AWGN - and examine the performance when there is jitter below. Recall that the update equation for Δ is in form of a line regression, which usually works well in the presence of additive white Gaussian noise in that (timing) domain. Indeed, the performance degradation is very small when compared to the case where white jitter is not present, as shown below. We note that since we are able to estimate the drift term very well in this set of simulations, the performance of our algorithm is much less dependent on the distribution of the drift term than the performance of the classical technique is.

$N = 8, T = 1, \sigma_g = 1$				
	$\Delta_r = 0.1$ $\sigma_z = 0.1$ $\sigma_w = 0.1$	$\Delta_r = 0.01$ $\sigma_z = 0.1$ $\sigma_w = 0.1$	$\Delta_r = 0.1$ $\sigma_z = 0.1$ $\sigma_w = 0.01$	$\Delta_r = 0.01$ $\sigma_z = 0.1$ $\sigma_w = 0.01$
Δ RMS FRI	0.0063	0.0057	0.0346	0.0022
τ RMS FRI	0.0241	0.0250	0.0246	0.0147
τ RMS Xcorr	3.9270	4.2117	1.3047	1.1856

VII. CONCLUSION

In this paper we have proposed a new approach to signal parameter estimation when the samples are subjected to timing noise. We extended this to include AWGN. Throughout we have focused on the problem of delay estimation when the pulse shape is known exactly. Taking advantage of the special structure in the problem, we developed algorithms which are shown to be superior to the previous method which ignores timing noise and considers it instead as an additional AWGN term. The approach proposed in this paper can be extended to other signal parameter estimation problems, and can include prior information of the parameters of the signal which we wish to estimate.

Interestingly, the effect of the pulse shape on the performance of parameter estimation in the presence of timing noise, and in the presence of both timing and additive noise, depend on the nature of the parameter of interest. We have seen in Theorem 1 that in the case of delay estimation, a pulse consisting of one linear piece than a pulse of many linear pieces of different slopes.

Finally, note that our algorithm not only gives an estimate of the delay $\hat{\tau}$, but also gives an estimate of the jitter \hat{z}_n , and hence of the sampling instances \hat{s}_n . This estimate can be used for other purposes, for example for improving future estimates in the case that the variables of interest are correlated.

REFERENCES

- [1] M. Unser, "Sampling-50 years after Shannon," *Proceedings of the IEEE*, vol. 88, no. 4, pp. 569–587, April 2000.
- [2] M. Vetterli, P. Marziliano, and T. Blu, "Sampling signals with finite rate of innovation," *IEEE Transactions on Signal Processing*, vol. 50, no. 6, pp. 1417–1428, June 2002.
- [3] P. L. Dragotti, M. Vetterli, and T. Blu, "Exact sampling results for signals with finite rate of innovation using Strang-Fix conditions and local kernels," in *Proc. IEEE Conf. on Acoustics, Speech, and Signal Proc.*, Philadelphia, PA, March 2005.
- [4] J. Kusuma and V. K. Goyal, "Multichannel sampling for parametric signals with a successive approximation property," in *IEEE Conf. on Image Processing*, August 2006, accepted for publication.
- [5] B. Drakhlis, "Calculate oscillator jitter by using phase-noise analysis," *Microwaves & RF*, pp. 82–90, 157, January 2001.
- [6] A. V. Balakrishnan, "On the problem of time jitter in sampling," *IRE Trans. Inf. Theory*, vol. 3, pp. 226–236, April 1962.
- [7] P. L. Butzer, "A survey of the Whittaker-Shannon sampling theorem and some of its extensions," *J. Math. Res. Expos.*, no. 3, pp. 183–212, 1983.
- [8] J. R. Higgins, *Sampling Theory in Fourier and Signal Analysis*. Oxford, UK: Oxford Science Publications, 1996.
- [9] N. Narasimhamurti and S. Awad, "Estimating the parameters of a sinusoid sampled with a clock with accumulated jitter," in *Proc. IEEE Instr. and Meas. Tech. Conf.*, May 1997, pp. 1132–1135.
- [10] H. Meyr and G. Ascheid, *Synchronization in Digital Communication*. New York, NY: Wiley, 1990.
- [11] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *J. Roy. Stat. Soc., Ser. B*, vol. 39, no. 1, pp. 1–38, 1977.