Least-squares estimation of time-base distortion of sampling oscilloscopes

— **Source link** ☑

Chih-Ming Wang, Paul D. Hale, Kevin J. Coakley

Institutions: National Institute of Standards and Technology

Related papers:

- Time-base nonlinearity determination using iterated sine-fit analysis

- Estimating the Magnitude and Phase Response of a 50 GHz Sampling Oscilloscope Using the "Nose-to-Nose" Method

- Accurate spectral estimation based on measurements with a distorted-timebase digitizer

- An improved sine-wave fitting procedure for characterizing data acquisition channels

- Compensation of Random and Systematic Timing Errors in Sampling Oscilloscopes

Least-Squares Estimation of Time-Base Distortion of Sampling Oscilloscopes

C. M. Wang, Paul D. Hale, and Kevin J. Coakley

Abstract—**We present an efficient least-squares algorithm for estimating the time-base distortion of sampling oscilloscopes. The method requires measurements of signals at *multiple* phases and frequencies. The method can accurately estimate the order of the harmonic model that is used to account for the amplitude nonlinearity of the sampling channel. We study several practical problems related to the time-base distortion estimation, such as the effect of averaging and sample size requirements. We also compare the relative performance of various methods for estimating time-base distortion using simulated and measured data.**

Index Terms—**Curve fitting, harmonic distortion, least-squares methods, mean-square error methods, timing jitter.**

I. INTRODUCTION

HIGH-SPEED sampling oscilloscopes have been used for many years as qualitative tools for measuring temporal waveforms. For these oscilloscopes to serve as accurate metrological instruments over their entire bandwidth, several effects must be compensated for [1], [2]. These effects include time-base distortion (TBD), timing jitter, timing drift [3], [4], additive noise, and impedance mismatch [5]. Our work focuses on characterizing TBD in high speed digital sampling oscilloscopes. If left uncorrected, TBD can cause significant errors in pulse width, step transition, and time interval measurements. Discontinuities in the TBD can severely distort a short pulse waveform. Such discontinuities must be detected and avoided in even the crudest measurements. After estimation of the TBD, the measured waveform must be adjusted accordingly [6] to insure high accuracy results.

The model of a discrete time signal is given by

$$y_i = f(t_i) + \epsilon_i,$$

where the ith sample y_i is a function of actual time of sampling t_i plus the additive noise ϵ_i. The actual time t_i can be written as

$$t_i = (i-1)T_s + g_i + \tau_i,$$

where $(i-1)T_s$ is the ideal sample time and T_s is the sampling interval. Deviations between the ideal and actual times have two components: a deterministic part g_i, called TBD and a random component τ_i, called jitter.

A number of methods have been developed to estimate g_i. Earlier methods, such as the "sinefit" [7], [8] and "analytic signal" [9], use sine-wave data of a known frequency with single or multiple starting phases to estimate the TBD. Stenbakken and Deyst [10] compare the performance of these two methods by simulations. Their results show that the sinefit method performs better than the analytic-signal method when the TBD has discontinuities, whereas the analytic-signal method performs better when the TBD is slowly changing. In addition, the analytic-signal method performs better when the TBD is large. For implementation, the sinefit method requires multiple waveforms of different starting phases to carry out the analysis, while the analytic-signal method requires only one waveform. The analytic-signal method, however, needs to drop data near the ends of the record, so the TBD at those samples can not be obtained.

Recent work [11], [12] on TBD estimation has been concentrated on least-squares methods that use waveforms of *multiple* phases and frequencies. The advantage of such an approach is that there are enough data to accurately estimate the TBD (including discontinuities). It also allows the harmonic distortions to be estimated and separated from the TBD. In this paper, we present an efficient least-squares algorithm for estimating the time-base and other distortions *simultaneously*. We study several practical problems related to TBD estimation, such as the effect of averaging and sample size requirements. We also compare the relative performance of various methods for estimating TBD using simulated and measured data.

II. LEAST-SQUARES METHOD

The model of the waveforms of multiple phases and frequencies is given by

$$y_{ij} = \alpha_j + \sum_{k=1}^{h} [\beta_{jk} \, \cos(2\pi k f_j t_{ij}) + \gamma_{jk} \, \sin(2\pi k f_j t_{ij})] + \epsilon_{ij}$$

where y_{ij} is the measured signal at time t_{ij} (the ith actual sample time of the jth experiment), f_j is the frequency used in the jth experiment, and β_{jk} and γ_{jk} are the amplitudes of the kth harmonic of the jth experiment. The number of harmonics h is assumed to be finite. The additive noises ϵ_{ij} are assumed to be independently and identically distributed (iid) with zero means and standard deviations $\sigma_\epsilon(j)$. The model

allows different additive noise standard deviations for different experiments. The model also assumes that t_{ij} is given by

$$t_{ij} = (i-1)T_s + g_i + \tau_{ij}$$

where T_s and g_i are as defined before and τ_{ij} are the random jitters and assumed to be iid (and independent of ϵ_{ij}) with zero means and standard deviations $\sigma_\tau(j)$. There will be m experiments with n samples for each experiment; that is, $i = 1, 2, \cdots, n$, and $j = 1, 2, \cdots, m$.

Let

$$\boldsymbol{\theta} = (g_1, g_2, \cdots, g_n, \alpha_1, \beta_{11}, \gamma_{11}, \cdots,$$
$$\beta_{1h}, \gamma_{1h}, \cdots, \alpha_m, \beta_{m1}, \cdots, \gamma_{mh})^t$$

be the column vector of the unknown parameters of the model. The number of unknowns is $n + m(2h + 1)$. Define

$$z_{ij}(\boldsymbol{\theta}) = \alpha_j + \sum_{k=1}^{h} [\beta_{jk} \cos(2\pi k f_j((i-1)T_s + g_i))$$
$$+ \gamma_{jk} \sin(2\pi k f_j((i-1)T_s + g_i))] \quad (1)$$

and

$$SS(\boldsymbol{\theta}) = \sum_{i,j} (y_{ij} - z_{ij}(\boldsymbol{\theta}))^2.$$

Then the LS estimate of $\boldsymbol{\theta}$, denoted by $\hat{\boldsymbol{\theta}}$, is the solution of

$$\min_{\hat{\boldsymbol{\theta}}} SS(\hat{\boldsymbol{\theta}}). \quad (2)$$

A Gauss-Newton type of iterative procedure can be used to solve (2). If the procedure is implemented directly, it would require $O(n^3)$ operations at each iterative step. This is not acceptable for a large n (in our problems, $n = 4096$). The model in (1) has a special structure, however, that can be exploited to obtain an algorithm which requires only $O(n)$ operations at each step. The detailed derivation of the algorithm is given in Appendix A.

In the subsequent discussion, all the TBD estimation is performed using the LS method.

III. EFFECTS OF AVERAGING

The choice of input frequencies is important in TBD estimation. Guidelines are given in [12] for selecting good sets of frequencies. In practice, a pair of appropriate frequencies is used. At each of the two frequencies, signals are sampled at different starting phases. In general, the phases measured and the number of phases used need not be the same at each frequency. Since the estimation is usually carried out offline, a large number of waveforms may be available. Due to computational and other constraints, we may not be able to use all the data at once to estimate the TBD; we need to do "averaging." To illustrate, suppose two frequencies are chosen. For each frequency, six waveforms using different starting phases are sampled. These 12 waveforms are labeled 1 to 12. Four possible ways can then be used to estimate the TBD. We could use all the data at once, or use waveforms, say, 1, 2, 3, 7, 8, and 9 to estimate the TBD and average it with another TBD estimate obtained from the remaining waveforms. Similarly,

we can average TBD estimates obtained from three sets of four waveforms, or six sets of two waveforms. We denote these methods by "1/12," "2/6," "3/4," and "6/2." The notation "M/N" means that the TBD estimate is the average of M individual TBD estimates, each of which was obtained using N waveforms with half from one frequency and another half from another frequency. To get a unique solution to (2), we must impose a constraint on the parameters. For example, we might require that $g_1 = 0$, or $\Sigma_{i=1}^n g_i = 0$. That is, the LS TBD estimates are unique only up to an arbitrary translation. This condition also exists for other methods. For example, the TBD estimates obtained using the analytic–signal method would differ by an arbitrary translation due to different starting phases. Thus, TBD estimates must be shifted to a common level before averaging. A commonly used criterion is to minimize the "distance" among the shifted TBD estimates. Let $\hat{g}_{ij}, i = 1, 2, \cdots, n$, be the TBD estimate obtained from the jth set of N waveforms, $j = 1, 2, \cdots, M$. We want to find constants ν_j that minimize

$$\sum_{i=1}^{n} \mathcal{D}(\hat{g}_{i1} - \nu_1, \hat{g}_{i2} - \nu_2, \cdots, \hat{g}_{iM} - \nu_M),$$

where $\mathcal{D}(\cdot, \cdots, \cdot)$ is an appropriate distance metric. Two possible distance metrics are considered in Appendix B to obtain ν_j. Once these "offset adjustments" are found, each $\hat{g}_{ij}$ is then shifted accordingly, and the final TBD estimate $\hat{g}_i$ can be obtained as the mean of the shifted $\hat{g}_{ij}$, that is,

$$\hat{g}_i = \frac{1}{M} \sum_{j=1}^{M} (\hat{g}_{ij} - \nu_j).$$

In simulation experiments, we study the effects of averaging. TBD can take on different forms for different oscilloscopes [11]. The simulation parameters used here are closely related to those we observe in our laboratory. We used an 8 ns time window with 4096 samples ($n = 4096$). The nominal TBD g_i is shown in Fig. 1. It has discontinuities at 4 ns intervals and includes some quadratic and sinusoidal modulation. Its exact form is given in Appendix C. We assumed that there was no harmonic distortion ($h = 1$). The two frequencies used were 9.75 GHz and 10.25 GHz. At each frequency, 12 phases, $(i-1)2\pi/12, i = 1, 2, \cdots, 12$, were used to generate a total of 24 waveforms. The additive and jitter error standard deviations used were 1% of the amplitude and 80% of the sample period, respectively. Based on these 24 waveforms, six averaging methods, 12/2, 6/4, 4/6, 3/8, 2/12, and 1/24, were employed to estimate g_i. The root–mean–square (RMS) error of the estimate $\hat{g}_i$ for each method is calculated as (adjusted for the arbitrary translation)

$$S_g = \sqrt{\frac{1}{4096} \sum_{i=1}^{4096} (\hat{g}_i - g_i)^2}. \quad (3)$$

The process was repeated 100 times. Fig. 2 plots the 100 RMS errors of TBD estimates for each method.

Fig. 2 indicates that the largest gain in performance improvement was obtained when we increased the number of waveforms in each frequency from 1 to 2 signals nearly in
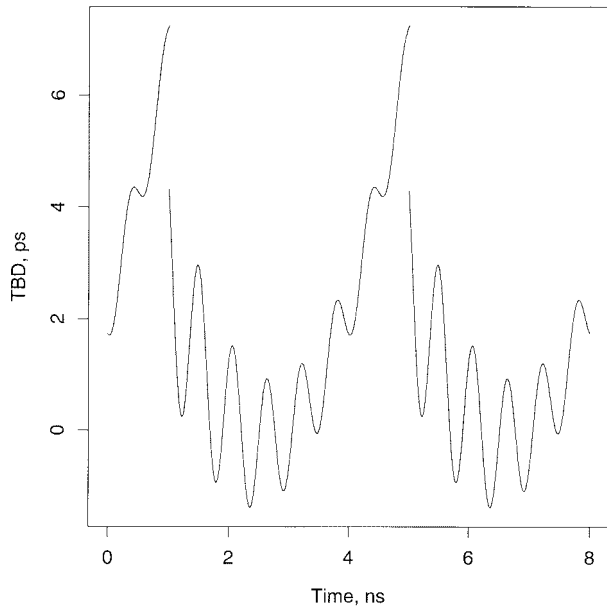
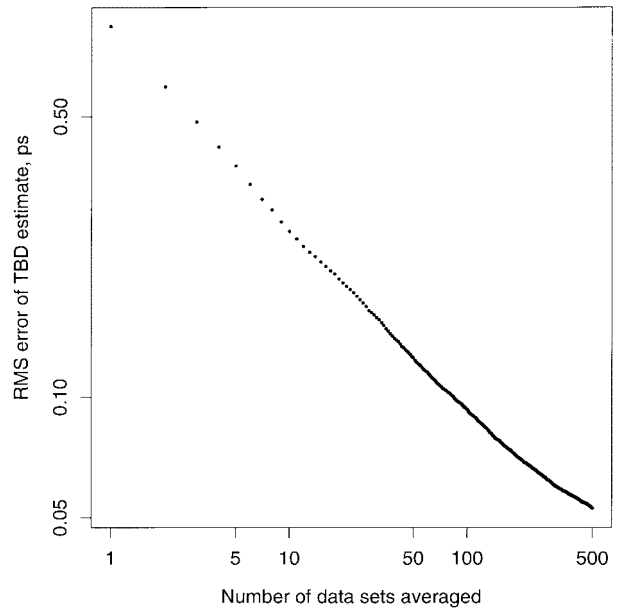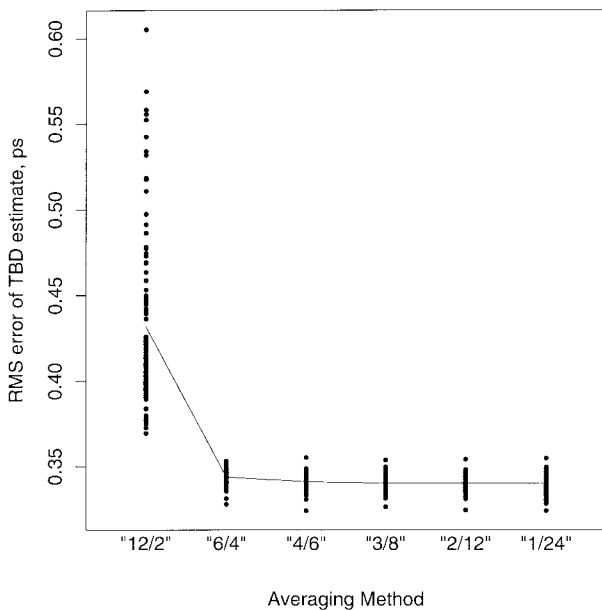Fig. 1. The time-base distortion used in the simulation.

Fig. 2. RMS errors of the TBD estimate for the six averaging methods. The line connects the means of the 100 RMS errors.

Fig. 3. RMS errors of the TBD estimate for $M/4$ method. The horizontal axis is the value of M.

quadrature. Moreover, increasing the number of waveforms further in averaging does not significantly improve the performance. For the computational speed, the "1/24" method was, on average, 4.8 times slower than the "6/4" method. Similar results were obtained when the actual and assumed harmonic order was 3 ($h = 3$). Based on these results, we conclude that when multiple sets of waveforms are used to estimate the TBD by averaging, it is sufficient to have only 4 waveforms. Each set contains two signals in quadrature from each of the two frequencies.

Having decided to use the method $M/4$, the next question that comes immediately to mind is what is the proper value of M to use in order to obtain an adequate TBD estimate. Both simulated and real data were used to answer this question. We first generated 500 sets of four waveforms using the same simulation parameters stated above. Fig. 3 plots the RMS error of the TBD estimate as a function of M on the log scale. That is, Fig. 3 plots the value of S_g against the value of M used in the $M/4$ method to calculate S_g. It shows that the RMS error drops precipitously as M increases from one to ten and levels off when M is greater than 200. We also measured 500 sets of four waveforms. Each data set of four waveform measurements contained a 9.75 GHz signal and nearly quadrature signal, and a 10.25 GHz signal and nearly quadrature signal. The signals were generated using an inexpensive 100 kHz—3.2 GHz synthesized signal generator multiplied by a $5\times$ multiplier. The resulting signal was filtered and amplified to give spurious harmonics of the input signal less than -60 dB (re: carrier) and spurious harmonics of the output signal < -50 dB (re: carrier). The oscilloscope was triggered using the fundamental signal generated by the signal generator and the relative phase of the measured waveform was set by changing the trigger level of the oscilloscope. The additive noise standard deviation was estimated to be about 1% of the amplitude, and jitter noise standard deviation was between 740 and 900 fs depending on the trigger level (phase) and frequency. For the measured data, since the TBD is not known and hence S_g cannot be computed, a different criterion must be used. The criterion is the standard deviation of the TBD estimate, which is defined as

$$\sqrt{\frac{1}{M-1} \sum_{j=1}^{M} (\hat{g}_{ij} - \hat{g}_i)^2}.$$

Fig. 4 plots the standard deviation of the TBD estimate for the measured data. The number on the top of each subplot is the value of M used in TBD estimation. Fig. 4 shows that
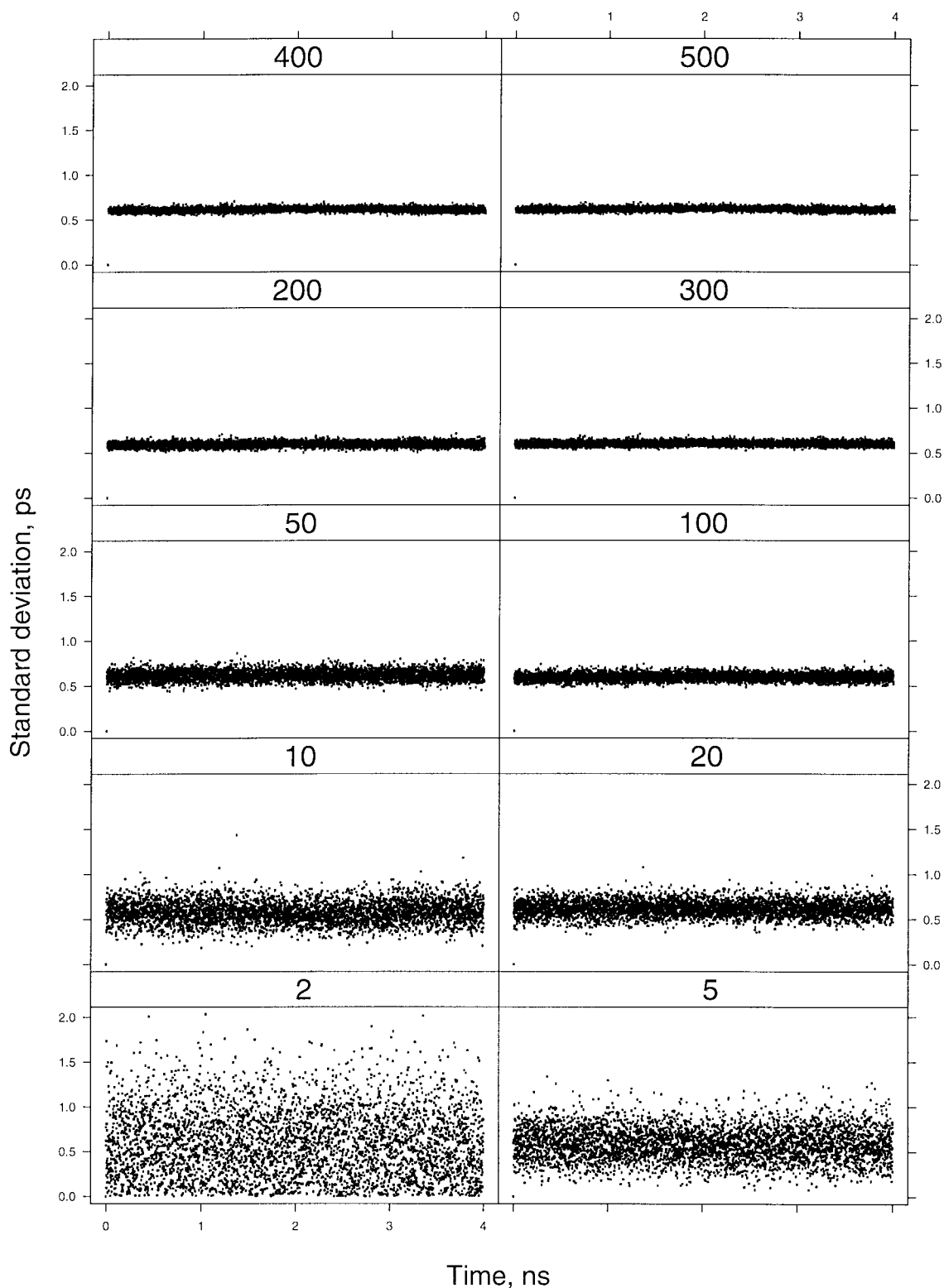
Fig. 4. Standard deviations of TBD estimates for $M/4$ method. The number of the top of each subplot is the value of M.

the range of standard deviations decreases drastically as M increases from 2 to 10 but the standard deviation remains fixed at about 0.7 ps up to $M = 500$. Both studies suggest that a reasonable starting value of M to use in TBD estimation is about 20. Incremental and sequential improvement on TBD estimation can be made as more data become available. The $M/4$ method with a moderate value of M enables us to obtain the uncertainty of the TBD estimate.

IV. SELECTION OF HARMONIC ORDER

A measured signal is usually contaminated with harmonics from the signal generator and distortion in the oscilloscope.

Fig. 5. Residual errors in TBD and in amplitude versus the harmonic order used in LS fits. The model with $h = 4$ was used in simulation.

Consequently, knowledge of the correct harmonic order is important for accurate TBD estimation. The proper harmonic order can be easily determined by examining the residuals from the LS fit. To illustrate, we first use simulated data. Twenty sets of four wa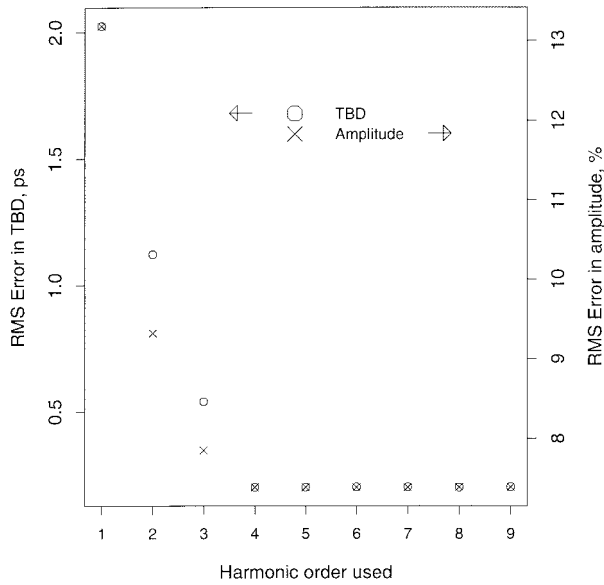veforms (20/4) with $h = 4$ were generated. The amplitudes corresponding to the second, third, and fourth harmonics were approximately 14, 7, and 3.5% of the fundamental amplitude. The rest of the simulation parameters were the same as those used before. Different harmonic orders, from one to nine, were used in TBD estimation. Two residual errors from the LS fit corresponding to different harmonic orders were calculated. The first is the residual error in TBD, S_g of (3). The second is the residual error in amplitude. The residual error in amplitude from the LS fit based on 4 waveforms and a harmonic model of order h^* is given by

$$S_A = \sqrt{\frac{\sum_{i=1}^{4096} \sum_{j=1}^{4} (y_{ij} - \hat{y}_{ij})^2}{4 \times 4096 - 4096 - 4(2h^* + 1)}}$$

where y_{ij} is the measured signal and $\hat{y}_{ij}$ is the LS predicted signal. The denominator is the number of degrees of freedom, which is the difference between the number of measurements (4×4096) and the number of parameters fitted in the model. For 20/4 method, we calculated the average residual error in amplitude as

$$\overline{S}_A = \sqrt{\tfrac{1}{20} \sum S_A^2}$$

where the summation is carried out for the 20 individual S_A^2.

Fig. 5 plots values of S_g and $\overline{S}_A$ as functions of h^*, the harmonic order used in LS fits. It shows that the correct harmonic order is the value of h^* where its corresponding S_g or $\overline{S}_A$ starts leveling off; that is, S_g or $\overline{S}_A$ does not change significantly. In this case, $h^* = 4$. We applied this simple technique on the measured data mentioned in the previous
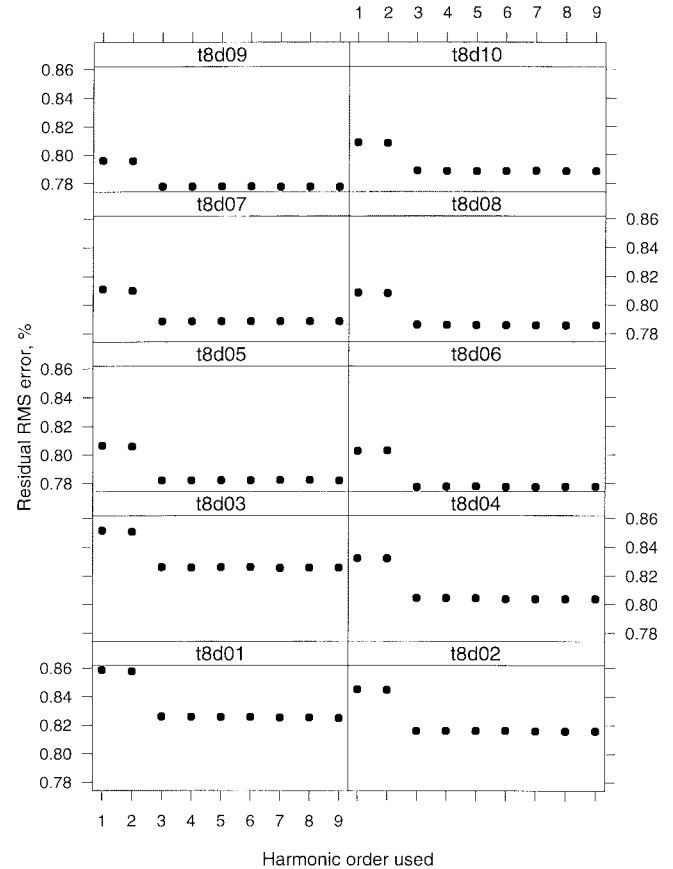

Fig. 6. Residual error in amplitude versus the harmonic order used in LS fits for ten data sets.

section. Fig. 6 plots the value of S_A (S_g is not available for real data) as a function of h^* for the first 10 (out of 500) data sets. It suggests that the proper order to use is 3.

V. COMPARISON OF METHODS

In this section, we compare the performance of the LS method with other known methods. We first compare the LS method with another least-squares based method proposed recently by Stenbakken and Deyst [12]. Their method, also requiring waveforms of multiple phases and frequencies, uses a two-stage approach. Specifically, the method first estimates amplitude parameters α_j, β_{jk}, and γ_{jk} of (1) for fixed TBD g_i using the ordinary least squares and then estimates g_i for fixed amplitude parameters using the weighted nonlinear least squares. These two steps are repeated until results converge. An advantage of this approach is that it requires fewer operations at each iterative step than the LS method which estimates all the parameters simultaneously. The comparison was done by simulation. We employed the same simulation parameters used before for g_i (Fig. 1), frequencies (9.75 and 10.25 GHz), starting phases (0 and 90°), and elapsed time (8 ns with $n = 4096$). We used $h = 3$, three additive error standard deviations of the 0.5, 1, and 2% of the amplitude, and three jitter error standard deviations of the 20, 50, and 80% of the sample period. Optimal weighting schemes were used in both methods. Fig. 7 plots the ratio of the RMS error in TBD
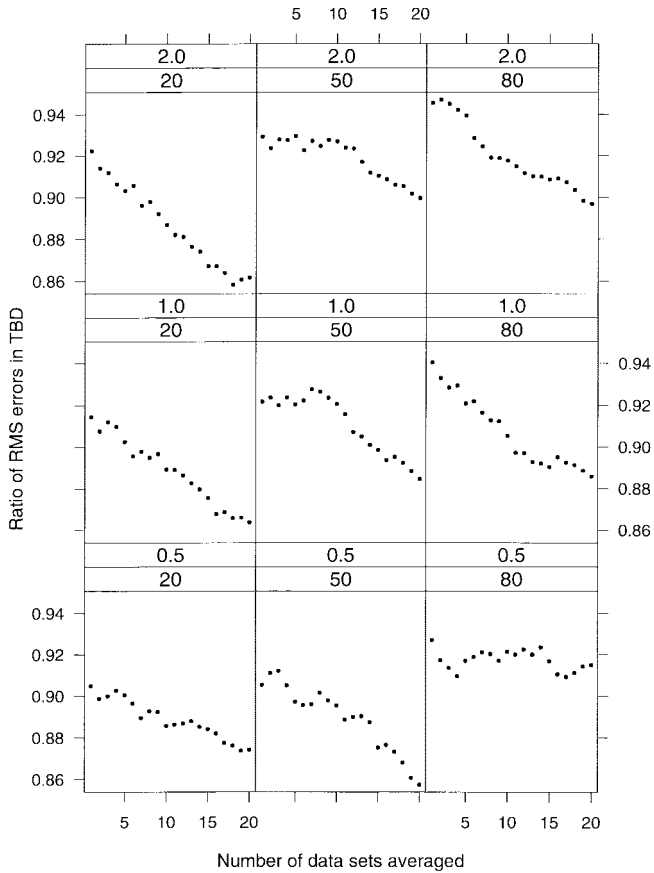
Fig. 7. Ratios of residual error in TBD of the LS method to the Sten-bakken–Deyst method. The horizontal axis is the value of M used in $M/4$. The top and bottom numbers in each subplot are the value of additive and jitter error standard deviations.

Fig. 8. Ratios of residual error in TBD of the LS method to the ana-lytic-signal method. The horizontal axis is the value of M used in $M/4$. The top and bottom numbers in each subplot are the value of additive and jitter error standard deviations.

(S_g) of the LS method to the Stenbakken–Deyst method as a function of M, the number of data sets averaged in $M/4$, for the 9 combinations of additive and jitter error standard deviations. The top and bottom numbers in each subplot are, respectively, the value (in percentage) of additive and jitter error standard deviations.

Fig. 7 indicates that, for simulation parameters considered, the LS method *uniformly* produces smaller RMS errors than does the Stenbakken–Deyst method. The difference can be as much as 14% for some cases. There seems to be no discernible correlation between relative performance and additive/jitter error standard deviation. For the execution speed, the Sten-bakken–Deyst method runs about 45% faster than the LS method at each iterative step. In our implementation, however, the Stenbakken–Deyst method generally requires more itera-tions than the LS method to converge with the same stopping criteria, the total execution time for both methods is very close.

Next, we compare the LS method with the analytic-signal method, a nonleast-squares method. Since the analytic-signal method does not perform well with the presence of discon-tinuities in the TBD or harmonic distortion, we used $h = 1$ and a 2 ns window ($n = 1024$ between times 3 ns and 5 ns in Fig. 1) where g_i is smoothly varying. In implementing the analytic-signal method, 8% of the samples from either end were dropped. Fig. 8 plots the ratio of the RMS error

in TBD of the LS method to the analytic-signal method as a function of M for the 9 combinations of additive and jitter error standard deviations. It shows that the analytic-signal method can be competitive when the jitter is large and M is small (the break-even point for $\sigma_\tau = 80\%$ is $M = 36$; that is, the LS method would outperform the analytic-signal method when $M \geq 36$). However, when the jitter is small, or the TBD has discontinuities, or there is harmonic distortion, or there are many waveforms available, the LS method is preferable.

VI. CONCLUSIONS

We described an efficient least-squares algorithm for esti-mating the TBD of sampling oscilloscopes based on wave-forms of multiple phases and frequencies. The method can accurately estimate the TBD even when it has discontinuities. It can also determine the correct order of the harmonic model. We showed that the TBD estimate can be updated and its performance improved sequentially as more measurements become available. The method compares favorably with other procedures. We applied the method to simulated and real data.

APPENDIX A
DERIVATION OF THE LS ALGORITHM

The iterative procedure starts with an initial guess θ_0 and produces a sequence $\theta_1, \theta_2, \cdots$ which, we hope, converges to

$\hat{\boldsymbol{\theta}}$. The $(l+1)$th iterate $\boldsymbol{\theta}_{l+1}$ is given by

$$\boldsymbol{\theta}_{l+1} = \boldsymbol{\theta}_l + \boldsymbol{b}_l,$$

where $\boldsymbol{b}_l$ is the so-called Gauss-Newton step and is obtained from the solution of the LS problem

$$\min_{\boldsymbol{b}} \|\boldsymbol{J}_l \boldsymbol{b} - (\boldsymbol{y} - \boldsymbol{z}_l)\| \tag{4}$$

where $\| \cdot \|$ is the 2-norm,

$$\boldsymbol{y} = (y_{11}, y_{12}, \cdots, y_{1m}, y_{21}, y_{22}, \cdots, y_{2m},$$
$$\cdots, y_{n1}, y_{n2}, \cdots, y_{nm})^t$$
$$\boldsymbol{z}_l = (z_{11}(\boldsymbol{\theta}_l), \cdots, z_{1m}(\boldsymbol{\theta}_l),$$
$$\cdots, z_{n1}(\boldsymbol{\theta}_l), \cdots, z_{nm}(\boldsymbol{\theta}_l))^t$$

and $\boldsymbol{J}_l$ is the $mn \times (n + m(2h+1))$ Jacobian matrix of the vector-valued function $\boldsymbol{z}(\boldsymbol{\theta})$ evaluated at $\boldsymbol{\theta}_l$. That is shown in the first equation at the bottom of the page.

The solution of (4) involves a QR decomposition of matrix $\boldsymbol{J}_l$. For a $u \times v$ matrix the QR decomposition requires $O(uv^2)$ operations [13]. Since $\boldsymbol{J}_l$ is $mn \times (n + m(2h+1))$ (and $n \gg m$, $n \gg h$) it requires $O(n^3)$ operations to compute its QR decomposition. The matrix $\boldsymbol{J}_l$ has a special structure, however, that can be exploited to solve (4) more efficiently. This is due to the fact that

$$\frac{\partial z_{ij}}{\partial g_k} = 0 \quad \text{if } i \neq k.$$

In particular, let U be the matrix containing the first n columns of $\boldsymbol{J}_l$ and V the remaining columns, that is,

$$\boldsymbol{J}_l = (\boldsymbol{U} \quad \boldsymbol{V}).$$

Then

$$\boldsymbol{U} = \begin{pmatrix} \boldsymbol{u}_1 & 0 & \cdots & 0 \\ 0 & \boldsymbol{u}_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \boldsymbol{u}_n \end{pmatrix}$$

where

$$\boldsymbol{u}_i = \begin{pmatrix} \dfrac{\partial z_{i1}}{\partial g_i} & \dfrac{\partial z_{i2}}{\partial g_i} & \cdots & \dfrac{\partial z_{im}}{\partial g_i} \end{pmatrix}^t$$

and (see the second equation shown at the bottom of the page) where

$$\boldsymbol{v}_{ij} = \begin{pmatrix} \dfrac{\partial z_{ij}}{\partial \alpha_j} & \dfrac{\partial z_{ij}}{\partial \beta_{j1}} & \dfrac{\partial z_{ij}}{\partial \gamma_{j1}} & \cdots & \dfrac{\partial z_{ij}}{\partial \beta_{jh}} & \dfrac{\partial z_{ij}}{\partial \gamma_{jh}} \end{pmatrix}^t.$$

To save storages, U and V can be stored in compact forms as

$$(\boldsymbol{u}_1 \quad \boldsymbol{u}_2 \quad \cdots \quad \boldsymbol{u}_n)$$

and

$$\begin{pmatrix} \boldsymbol{v}_{11}^t & \boldsymbol{v}_{12}^t & \cdots & \boldsymbol{v}_{1m}^t \\ \boldsymbol{v}_{21}^t & \boldsymbol{v}_{22}^t & \cdots & \boldsymbol{v}_{2m}^t \\ \vdots & \vdots & \vdots & \vdots \\ \boldsymbol{v}_{n1}^t & \boldsymbol{v}_{n2}^t & \cdots & \boldsymbol{v}_{nm}^t \end{pmatrix}$$

The normal equations associated with (4) are given by

$$\boldsymbol{J}_l^t \boldsymbol{J}_l \boldsymbol{b} = \boldsymbol{J}_l^t (\boldsymbol{y} - \boldsymbol{z}_l) \tag{5}$$

Let $\boldsymbol{b}^t = (\boldsymbol{b}_1^t \boldsymbol{b}_2^t)$ with $\boldsymbol{b}_1^t$ is $1 \times n$ and $\boldsymbol{b}_2^t$ is $1 \times m(2h+1)$. Then (5) becomes

$$\begin{pmatrix} \boldsymbol{U}^t \boldsymbol{U} & \boldsymbol{U}^t \boldsymbol{V} \\ \boldsymbol{V}^t \boldsymbol{U} & \boldsymbol{V}^t \boldsymbol{V} \end{pmatrix} \begin{pmatrix} \boldsymbol{b}_1 \\ \boldsymbol{b}_2 \end{pmatrix} = \begin{pmatrix} \boldsymbol{U}^t(\boldsymbol{y} - \boldsymbol{z}_l) \\ \boldsymbol{V}^t(\boldsymbol{y} - \boldsymbol{z}_l) \end{pmatrix}. \tag{6}$$

Solving the top equations of (6) for $\boldsymbol{b}_1$ in terms of $\boldsymbol{b}_2$ yields

$$\boldsymbol{b}_1 = (\boldsymbol{U}^t \boldsymbol{U})^{-1}[\boldsymbol{U}^t(\boldsymbol{y} - \boldsymbol{z}_l) - \boldsymbol{U}^t \boldsymbol{V} \boldsymbol{b}_2] \tag{7}$$

Since $\boldsymbol{U}^t \boldsymbol{U}$ is a diagonal matrix, $\boldsymbol{b}_1$ can be easily obtained. We need, however, to find a solution for $\boldsymbol{b}_2$ first. Substituting (7) into the bottom equations of (6) gives

$$\boldsymbol{V}^t(\boldsymbol{I} - \boldsymbol{U}(\boldsymbol{U}^t \boldsymbol{U})^{-1} \boldsymbol{U}^t) \boldsymbol{V} \boldsymbol{b}_2$$
$$= \boldsymbol{V}^t(\boldsymbol{I} - \boldsymbol{U}(\boldsymbol{U}^t \boldsymbol{U})^{-1} \boldsymbol{U}^t)(\boldsymbol{y} - \boldsymbol{z}_l). \tag{8}$$

Since

$$\boldsymbol{P} = \boldsymbol{I} - \boldsymbol{U}(\boldsymbol{U}^t \boldsymbol{U})^{-1} \boldsymbol{U}^t$$

is an idempotent matrix, that is, $\boldsymbol{P}^2 = \boldsymbol{P}$, (8) is the normal equations associated with the following LS problem

$$\min_{\boldsymbol{b}_2} \|\boldsymbol{P} \boldsymbol{V} \boldsymbol{b}_2 - (\boldsymbol{y} - \boldsymbol{z}_l)\| \tag{9}$$

$$\boldsymbol{J}_l = \begin{pmatrix} \dfrac{\partial z_{11}}{\partial g_1} & \dfrac{\partial z_{11}}{\partial g_2} & \cdots & \dfrac{\partial z_{11}}{\partial g_n} & \dfrac{\partial z_{11}}{\partial \alpha_1} & \dfrac{\partial z_{11}}{\partial \beta_{11}} & \cdots & \dfrac{\partial z_{11}}{\partial \gamma_{mh}} \\ \dfrac{\partial z_{12}}{\partial g_1} & \dfrac{\partial z_{12}}{\partial g_2} & \cdots & \dfrac{\partial z_{12}}{\partial g_n} & \dfrac{\partial z_{12}}{\partial \alpha_1} & \dfrac{\partial z_{12}}{\partial \beta_{11}} & \cdots & \dfrac{\partial z_{12}}{\partial \gamma_{mh}} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \dfrac{\partial z_{nm}}{\partial g_1} & \dfrac{\partial z_{nm}}{\partial g_2} & \cdots & \dfrac{\partial z_{nm}}{\partial g_n} & \dfrac{\partial z_{nm}}{\partial \alpha_1} & \dfrac{\partial z_{nm}}{\partial \beta_{11}} & \cdots & \dfrac{\partial z_{nm}}{\partial \gamma_{mh}} \end{pmatrix} \Bigg|_{\boldsymbol{\theta} = \boldsymbol{\theta}_l}$$

$$\boldsymbol{V}^t = \begin{pmatrix} \boldsymbol{v}_{11} & 0 & \cdots & 0 & \boldsymbol{v}_{21} & 0 & \cdots & 0 & \cdots & \boldsymbol{v}_{n1} & 0 & \cdots & 0 \\ 0 & \boldsymbol{v}_{12} & \cdots & 0 & 0 & \boldsymbol{v}_{22} & \cdots & 0 & \cdots & 0 & \boldsymbol{v}_{n2} & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & \boldsymbol{v}_{1m} & 0 & 0 & \cdots & \boldsymbol{v}_{2m} & \cdots & 0 & 0 & \cdots & \boldsymbol{v}_{nm} \end{pmatrix}$$

The solution of $\boldsymbol{b}_2$ then involves a QR decomposition of $\boldsymbol{PV}$. This is an acceptable solution since it requires only $O(nm^3)$ operations.

The iterative procedure above may converge very slowly or may oscillate widely. The following algorithm is used to speed up the convergence. At the lth iteration, let interval $\mathcal{I}$ be

$$\mathcal{I} = \begin{cases} (0,2) & \text{if } SS(\boldsymbol{\theta}_{l+1}) < SS(\boldsymbol{\theta}_l) \\ (-0.5,0.5) & \text{if } SS(\boldsymbol{\theta}_{l+1}) > SS(\boldsymbol{\theta}_l) \end{cases}$$

then a combination of the golden-section search and successive parabolic interpolation [14] is used to find $\delta \in \mathcal{I}$ such that $SS(\boldsymbol{\theta}_l + \delta \boldsymbol{b}_l)$ is minimum. Set $\boldsymbol{\theta}_{l+1} = \boldsymbol{\theta}_l + \delta \boldsymbol{b}_l$ and the iterative cycle begins again.

Weights can be used in this procedure. Let $\boldsymbol{W}$ be an $mn \times mn$ diagonal matrix with diagonal elements as the weights for y_{ij}. To incorporate the weights, simply premultiply $\boldsymbol{U}$, $\boldsymbol{V}$, and $(\boldsymbol{y} - \boldsymbol{z})$ by $\boldsymbol{W}^{1/2}$ and use them in (7) and (9) to obtain $\boldsymbol{b}$ for each iteration.

The most commonly used weight for y_{ij} is $1/\text{var}(y_{ij})$. Since, in practice, the variance of y_{ij} is unknown, it must be estimated. We can obtain the estimate of $\text{var}(y_{ij})$ either from independent, repeated experiments or (if we have prior information on the additive and jitter noises) from the model

$$\text{var}(y_{ij}) \approx \sigma_\epsilon^2(j) + \left(\frac{\partial z_{ij}}{\partial t_{ij}}\right)^2 \sigma_\tau^2(j).$$

APPENDIX B
METHODS FOR OBTAINING OFFSET ADJUSTMENTS

Two possible metrics are considered here. The first is absolute deviation from the mean. Let

$$\hat{g}_{i\bullet} = \frac{1}{M} \sum_{j=1}^{M} \hat{g}_{ij}$$

and define the distance metric as the sum of the distance between $\hat{g}_{ij} - \nu_j$ and $\hat{g}_{i\bullet}$; that is

$$\mathcal{D}(\hat{g}_{i1} - \nu_1, \hat{g}_{i2} - \nu_2, \cdots, \hat{g}_{iM} - \nu_M) = \sum_{j=1}^{M} |\hat{g}_{ij} - \nu_j - \hat{g}_{i\bullet}|.$$

The expression we need to minimize is

$$\sum_{j=1}^{M} \sum_{i=1}^{n} |\hat{g}_{ij} - \hat{g}_{i\bullet} - \nu_j|.$$

Given $y_1, y_2, \cdots, y_k$, the value of ν that minimizes

$$\sum_{i=1}^{k} |y_i - \nu|$$

is $\nu = \text{median}(y_1, y_2, \cdots, y_k)$. Thus the solution for ν_j is

$$\nu_j = \text{median}(\hat{g}_{1j} - \hat{g}_{1\bullet}, \hat{g}_{2j} - \hat{g}_{2\bullet}, \cdots, \hat{g}_{nj} - \hat{g}_{n\bullet}). \quad (10)$$

The second metric is the sum of the squared distance between $\hat{g}_{ij} - \nu_j$ and $\hat{g}_{i\bullet}$. The expression needs to be minimized is then

$$\sum_{j=1}^{M} \sum_{i=1}^{n} (\hat{g}_{ij} - \hat{g}_{i\bullet} - \nu_j)^2$$

Given $y_1, y_2, \cdots, y_k$, the value of ν that minimizes

$$\sum_{i=1}^{k} (y_i - \nu)^2$$

is $\nu = \Sigma_{i=1}^{k} y_i / k$. Thus the solution of ν_j is

$$\nu_j = \frac{1}{n} \sum_{i=1}^{n} (\hat{g}_{ij} - \hat{g}_{i\bullet}) = \hat{g}_{\bullet j} - \hat{g}_{\bullet \bullet}$$

Here we define $\hat{g}_{\bullet j}$ with a subscript replaced by a dot to be the mean when averaged over the subscript that has been replaced by that dot. Since $\hat{g}_{\bullet \bullet}$, the overall mean, is a fixed constant (independent of subscript j), ν_j is simply

$$\nu_j = \hat{g}_{\bullet j} \quad (11)$$

The offset adjustment in (10) is more robust against the presence of outliers in $\hat{g}_{ij}$. If there is no outlier, both offset adjustments produce almost identical results. In this paper, we use the ν_j of (11) because the ease of computing. Furthermore, if $\nu_j = \hat{g}_{\bullet j}$, then

$$\hat{g}_i = \frac{1}{M} \sum_{j=1}^{M} (\hat{g}_{ij} - \hat{g}_{\bullet j}) = \frac{1}{M} \sum_{j=1}^{M} \hat{g}_{ij} - \hat{g}_{\bullet \bullet}$$

That is, no shift for individual $\hat{g}_{ij}$ is needed.

APPENDIX C
NOMINAL TIME-BASE DISTORTION

TBD shown in Fig. 1 is given by

$$\begin{aligned} g_t = &\Phi(1-t)[(a_1(t+1) + (0.35t + 0.35)^2 \\ &- a_2 e^{-0.35(t+3)} \sin(3.5\pi(t+3.5))] \\ &+ \Phi(t-1)\Phi(5-t)[a_1(t-3) + (0.35t - 1.05)^2 \\ &- a_2 e^{-0.35(t-1)} \sin(3.5\pi(t-0.5))] \\ &+ \Phi(t-5)[a_1(t-7) + (0.35t - 2.45)^2 \\ &- a_2 e^{-0.35(t-5)} \sin(3.5\pi(t-4.5))] \end{aligned}$$

where $a_1 = 0.001$, $a_2 = 0.002$, and

$$\Phi(t) = \begin{cases} 0, & \text{if } t < 0 \\ 1, & \text{otherwise} \end{cases}$$

REFERENCES

[1] W. R. Scott Jr., "Error corrections for an automated time-domain network analyzer," *IEEE Trans. Instrum. Meas.*, vol. IM-35, pp. 300–303, 1986.
[2] J. P. Deyst, N. G. Paulter, T. Daboczi, G. N. Stenbakken, and M. Souders, "Fast-pulse oscilloscope calibration system," *IEEE Trans. Instrum. Meas.*, vol. 47, pp. 1037–1041, 1998.
[3] K. J. Coakley and P. D. Hale, "Alignment of noisy signals," submitted for publication.
[4] J. Verspecht, "Calibration of a measurement system for high frequency nonlinear devices," Ph.D. dissertation, Vrije Univ., Brussels, Belgium, pp. 70–72, 1995.

[5] T. Dhuane, L. Martens, and D. D. Zutter, "Calibration and normalization of time-domain network analyzer measurements," *IEEE Trans. Microwave Theory Tech.*, vol. 42, pp. 580–589, 1994.

[6] Y. Rolan, J. Schoukens, and G. Vandersteen, "Signal reconstruction for nonequidistant finite length sample sets: A KIS approach," *IEEE Trans. Instrum. Meas.*, vol. 47, pp. 1046–1052, 1998.

[7] "IEEE standard for digitizing waveform recorders," IEEE Std. 1057-1994, 1994.

[8] R. Pintelon and J. Schoukens, "An improved sine wave fitting procedure for characterizing data acquisition channels," *IEEE Trans. Instrum. Meas.*, vol. 45, pp. 588–593, 1996.

[9] J. Verspecht, "Accurate spectral estimation based on measurements with a distored-timebase digitizer," *IEEE Trans. Instrum. Meas.*, vol. 43, pp. 210–215, 1994.

[10] G. N. Stenbakken and J. P. Deyst, "Comparison of time base nonlinearity measurement techniques," *IEEE Trans. Instrum. Meas.*, vol. 47, pp. 34–39, 1998.

[11] G. Vandersteen, Y. Rolain, and J. Schoukens, "System Identification for data acquisition characterization," in *Proc. Instrum. Meas. Tech. Conf.*, IMTC'98, May 18–21, 1998.

[12] G. N. Stenbakken and J. P. Deyst, "Time-base nonlinearity determination using iterated sine-fit analysis," *IEEE Trans. Instrum. Meas.*, vol. 47, pp. 1056–1061, 1998.

[13] G. H. Golub and C. F. Van Loan, *Matrix Computations, 2nd ed.* Baltimore MD: John Hopkins Univ. Press, p. 211, 1989.

[14] G. E. Forsythe, M. A. Malcolm, and C. B. Moler, *Computer Methods for Mathematical Computations*. Englewood Cliffs, NJ: Prentice-Hall, p. 182, 1977.

Paul D. Hale received the Ph.D. degree in applied physics from the Colorado School of Mines, Golden, in 1989.

He has been with the Optoelectronics Division, National Institute of Standards and Technology (NIST), Boulder, CO, since 1984. He has conducted research in birefringent devices, mode-locked fiber lasers, fiber chromatic dispersion, broadband lasers, interferometry, polarization standards, and photodiode frequency response. He is presently leader of the High Speed Measurements Project in the Sources and Detectors Group. His current interests are in precision optoelectronic frequency response measurement and related areas.

Dr. Hale, along with a team of four scientists, received the Department of Commerce Gold Medal in 1994 for measuring fiber cladding diameter with an uncertainty of 30 nm. In 1998, he received a Department of Commerce Bronze Medal, along with C. M. Wang and three other scientists, for developing measurement techniques and standards to determine optical polarization parameters.

Kevin J. Coakley received the Ph.D. degree in statistics from Stanford University, Stanford, CA, in 1989.

He is a Mathematical Statistician with the National Institute of Standards and Technology (NIST), Boulder, CO. His research interests include computer intensive statistical methods, signal processing, imaging and planning and analysis of experiments in physical science and engineering.

C. M. Wang received the Ph.D. degree in statistics from Colorado State University, Fort Collins, in 1978.

He is a Mathematical Statistician with the Statistical Engineering Division, National Institute of Standards and Technology (NIST), Boulder, CO. His research interests include interval estimation on variance components, statistical graphics and computing, and the application of statistical methods to physical sciences.

Dr. Wang is a Fellow of the American Statistical Association.