Adaptation of a cubic smoothing spline algortihm for multi-channel data stitching at the National Ignition Facility

C. Brown, A. Adcock, S. Azevedo, J. Liebman, E. Bond

January 7, 2011

SPIE Photonics West
San Francisco, CA, United States
January 22, 2011 through January 27, 2011

Disclaimer

This document was prepared as an account of work sponsored by an agency of the United States government. Neither the United States government nor Lawrence Livermore National Security, LLC, nor any of their employees makes any warranty, expressed or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States government or Lawrence Livermore National Security, LLC. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States government or Lawrence Livermore National Security, LLC, and shall not be used for advertising or product endorsement purposes.

Adaptation of a cubic smoothing spline algortihm for multi-channel data stitching at the National Ignition Facility

Charles G. Brown Jr.*, [a], Aaron B. Adcock[a,b], Stephen G. Azevedo[a], Judith A. Liebman[a], and Essex J. Bond[a]

[a]Lawrence Livermore National Laboratory, 7000 East Ave., Livermore, CA, USA 94550-9234;
[b]Stanford University, 450 Serra Mall, Stanford, CA, USA 94305

ABSTRACT

Some diagnostics at the National Ignition Facility (NIF), including the Gamma Reaction History (GRH) diagnostic, require multiple channels of data to achieve the required dynamic range. These channels need to be stitched together into a single time series, and they may have non-uniform and redundant time samples. We chose to apply the popular cubic smoothing spline technique to our stitching problem because we needed a general non-parametric method. We adapted one of the algorithms in the literature, by Hutchinson and deHoog, to our needs. The modified algorithm and the resulting code perform a cubic smoothing spline fit to multiple data channels with redundant time samples and missing data points. The data channels can have different, time-varying, zero-mean white noise characteristics. The method we employ automatically determines an optimal smoothing level by minimizing the Generalized Cross Validation (GCV) score. In order to automatically validate the smoothing level selection, the Weighted Sum-Squared Residual (WSSR) and zero-mean tests are performed on the residuals. Further, confidence intervals, both analytical and Monte Carlo, are also calculated. In this paper, we describe the derivation of our cubic smoothing spline algorithm. We outline the algorithm and test it with simulated and experimental data.

Keywords: Cubic spline, smoothing

1. INTRODUCTION

The National Ignition Facility (NIF) is a 192-beam pulsed laser system completed in May 2009 at the Lawrence Livermore National Laboratory (LLNL) and now producing experimental results for the study of inertial confinement fusion and the physics of extreme energy densities and pressures [1]. The initial goals of NIF include demonstration of thermonuclear burn (ignition) of deuterium and tritium fuel (D-T) in a laboratory setting. One of the ways to measure the energy yield over very short time-scales (20 ns) and large dynamic range is with carefully-timed measurements of gamma-rays emitted by the imploding D-T target [2]. The Gamma Reaction History (GRH) diagnostic has been successfully deployed for this purpose at the OMEGA laser [3] and is now operational at NIF, producing many time waveforms for each imploded target shot. To obtain the required orders-of-magnitude in dynamic range for GRH, experimenters are recording multiplexed channels of the same event, each with different voltage offsets, attenuation, and even time scales. The absolute noise levels or uncertainties for each channel are different, and some of the channels may have saturated regions that cannot be used. The goal of this paper is to describe a method that accurately estimates the "true" signal, with estimated error bars, that combines all data from the multiplexed noisy waveforms into a single composite, or "stitched", signal.

A "stitching" method for NIF data must be automatic, robust, and yield high-quality results. One approach to spline smoothing with repeated time samples is addressed in [4], which models the true signal as a stochastic process and solves for its estimate using a Kalman filter approach. However, based on past work by other researchers on OMEGA GRH data and ease of implementation, we chose to modify the cubic smoothing spline technique in [5] to our needs. The modified algorithm and the resulting code perform a cubic smoothing spline fit to multiple data channels with redundant time samples and missing data points. The data channels can have different, time-varying, zero-mean white noise characteristics. The method we employ automatically

*brown207@llnl.gov

determines an optimal smoothing level by minimizing the Generalized Cross Validation (GCV) score [6]. In order to automatically detect failure of the smoothing level selection or excessive bias in the fit, the Weighted Sum-Squared Residual (WSSR) and zero-mean tests are performed on the residuals. Furthermore, confidence intervals, both analytical and Monte Carlo, are also calculated. In the first section of this paper, we describe the derivation of our cubic smoothing spline algorithm and provide error estimates. Then we apply it to simulated and experimental data. The evaluation of the results motivates the need for the model accuracy checks.

2. ALGORITHM DERIVATION

In this section we derive our algorithm for fitting a cubic smoothing spline to data with repeated time samples. The introductory portion of our derivation and its notation are taken from a combination of [4], [7], [8].

Following [4], we assume we have measurements from multiple data channels with samples $y_{n,p}$ at unique times t_n, with $n = 1, 2, ..., N$ and $p = 1, 2, ..., P_n$. N is the number of unique time samples, and $t_1 \geq t_2 \geq ... \geq t_N$. P_n is the number of redundant measurements for each unique time sample t_n. We form $y_{n,p}$ into a vector $\mathbf{y} = (y_{1,1}, y_{1,2}, ..., y_{1,P_1}, y_{2,1}, y_{2,2}, ..., y_{2,P_2}, ..., y_{N,1}, y_{N,2}, ..., y_{N,P_N})^T$ [4], where superscipt "T" denotes the transpose. Let the true function values at the unique times $t_n, n = 1, 2, ..., N$ be $\mathbf{a} = (a_1, a_2, ..., a_N)^T$. Let the zero-mean, independent, white noise random variables added to the true values be formed into the vector $\eta = (\eta_{1,1}, \eta_{1,2}, ..., \eta_{1,P_1}, \eta_{2,1}, \eta_{2,2}, ..., \eta_{2,P_2}, ..., \eta_{N,1}, \eta_{N,2}, ..., \eta_{N,P_N})^T$. Our measurement model in matrix form then is

$$\underbrace{\mathbf{y}}_{M\times 1} = \underbrace{F}_{M\times N}\underbrace{\mathbf{a}}_{N\times 1} + \underbrace{\eta}_{M\times 1} , \tag{1}$$

where η is a column vector representing the zero-mean, independent, white noise added to each of the measurements. The matrix F is the $M \times N$ matrix of zeros and ones, with ones in rows $m = 1, 2, ..., P_1$ of column $n = 1$; ones in rows $m = P_1 + 1, P_1 + 2, ..., P_1 + P_2$ of column $n = 2$; and so on up to column $n = N$, where there are ones in rows $m = \left(\sum_{i=1}^{N-1} P_i\right) + 1, \left(\sum_{i=1}^{N-1} P_i\right) + 2, ..., \left(\sum_{i=1}^{N-1} P_i\right) + P_N$ [4]. M is the total number of all of the measurements: $M = \sum_{n=1}^{N} P_n$.

The function to be minimized is obtained by altering the one in [7] to account for redundant time samples:

$$\hat{\mathbf{a}} = \arg\min_{a} \left\{ (\mathbf{y} - F\mathbf{a})^T W^{-2} (\mathbf{y} - F\mathbf{a}) + \alpha \mathbf{a}^T Q T^{-1} Q^T \mathbf{a} \right\} , \tag{2}$$

where W is the $M \times M$ diagonal matrix with the standard deviations of the elements of η along it main diagonal and zeros elsewhere [7], Q is the $N \times (N-2)$ matrix with rows $j = 1, 2, ..., N$ and columns $i = 1, 2, ..., N-2$ as follows [7]:

$$Q_{j,i} = \begin{cases} h_i^{-1}, & j = i \\ -(h_i^{-1} + h_{i+1}^{-1}), & j = i+1 \\ h_{i+1}^{-1}, & j = i+2 \\ 0, & \text{otherwise} \, , \end{cases} \tag{3}$$

with $h_i = t_{i+1} - t_i$, and T is the $(N-2) \times (N-2)$ matrix with rows $i = 1, 2, ..., N-2$ and columns $j = 1, 2, ..., N-2$ as follows [7]:

$$T_{i,j} = \begin{cases} \frac{1}{6} h_i, & j = i-1 \\ \frac{1}{3}(h_i + h_{i+1}), & j = i \\ \frac{1}{6} h_{i+1}, & j = i+1 \\ 0, & \text{otherwise} \, . \end{cases} \tag{4}$$

The smoothing parameter, α, controls the tradeoff between the two terms of Equation 2: if α is very small, the estimate minimizes the sum of the squared residuals, while if α is very large, the estimate is a least-squares fit of a straight line to the data [7]. In order to solve Equation 2, we can determine the $\mathbf{a}$ that satisfies

$$\frac{d}{d\mathbf{a}} [(\mathbf{y} - F\mathbf{a})^T W^{-2} (\mathbf{y} - F\mathbf{a}) + \alpha \mathbf{a}^T Q T^{-1} Q^T \mathbf{a}] = 0 \tag{5}$$

using matrix calculus [9], [10], which gives us the estimate $\hat{\mathbf{a}}_\alpha$:

$$\hat{\mathbf{a}}_\alpha = (\Lambda + \alpha Q T^{-1} Q^T)^{-1} F^T W^{-2} \mathbf{y} \,, \tag{6}$$

where $\Lambda = F^T W^{-2} F$. The straightforward solution in Equation 6 is not desirable from a computational standpoint, since the matrix that is inverted is full, so we seek a form that is more computationally tractable. Starting with Equation 6 and using $Q^T \hat{\mathbf{a}}_\alpha = T \hat{\mathbf{c}}_\alpha$, which is a property of cubic splines [7], we arrive at the following set of equations:

$$\begin{aligned} \hat{\mathbf{c}}_\alpha &= (T + \alpha Q^T \Lambda^{-1} Q)^{-1} Q^T \Lambda^{-1} F^T W^{-2} \mathbf{y} && (7) \\ \hat{\mathbf{a}}_\alpha &= \Lambda^{-1} F^T W^{-2} \mathbf{y} - \alpha \Lambda^{-1} Q \hat{\mathbf{c}}_\alpha \,. && (8) \end{aligned}$$

Equations 7 and 8 are similar to Equations 2.5 and 2.6 in [5] and form a tractable solution for $\hat{\mathbf{a}}_\alpha$, given α. We modify the algorithm in [5] to solve for $\hat{\mathbf{a}}_\alpha$ with an α given by GCV score minimization.

As in [5], we define the matrix B_α as

$$B_\alpha = T + \alpha Q^T \Lambda^{-1} Q \,, \tag{9}$$

so that

$$\hat{\mathbf{c}}_\alpha = B_\alpha^{-1} Q^T \Lambda^{-1} F^T W^{-2} \mathbf{y} \,. \tag{10}$$

Setting $\hat{\mathbf{u}}_\alpha = \hat{\mathbf{c}}_\alpha$, we find

$$B_\alpha \hat{\mathbf{u}}_\alpha = Q^T \Lambda^{-1} F^T W^{-2} \mathbf{y} \,, \tag{11}$$

which we can solve efficiently using banded solvers or Cholesky decomposition [5].

Next, with foresight to the GCV score computation, we set $\hat{\mathbf{v}}_\alpha = W^{-1}(I - F A_\alpha)\mathbf{y}$, where A_α is defined by substituting Equation 7 into Equation 8 [5]:

$$\hat{\mathbf{a}}_\alpha = \underbrace{[\Lambda^{-1} F^T W^{-2} - \Lambda^{-1} Q B_\alpha^{-1} Q^T \Lambda^{-1} F^T W^{-2}]}_{A_\alpha} \mathbf{y} \,. \tag{12}$$

After substitution of A_α into the equation $\hat{\mathbf{v}}_\alpha = W^{-1}(I - F A_\alpha)\mathbf{y}$, linear algebra leads to

$$\hat{\mathbf{v}}_\alpha = W^{-1}[(I - F \Lambda^{-1} F^T W^{-2})\mathbf{y} - \alpha F \Lambda^{-1} Q \hat{\mathbf{u}}_\alpha] \,, \tag{13}$$

and

$$\begin{aligned} \hat{\mathbf{a}}_\alpha &= (F^T F)^{-1} F^T (\mathbf{y} - W \hat{\mathbf{v}}_\alpha) && (14) \\ \hat{\mathbf{c}}_\alpha &= \hat{\mathbf{u}}_\alpha \,. && (15) \end{aligned}$$

Now we outline the computation of the GCV score. First, we express the GCV score V_α for a given α as [4], [5], [11]

$$V_\alpha = \frac{\frac{1}{M} ||W^{-1}(I - F A_\alpha)\mathbf{y}||^2}{[\frac{1}{M} Tr(I - F A_\alpha)]^2} \,. \tag{16}$$

In order to compute the numerator of Equation 16, we calculate [5]

$$F_\alpha = \hat{\mathbf{v}}_\alpha^T \hat{\mathbf{v}}_\alpha \,. \tag{17}$$

The computation of the denominator of Equation 16 is more involved. First, we compute the five central bands of B_α^{-1} using the rational Cholesky decomposition and Theorem 3.1 of [5]. Then we use the resulting five central bands of B_α^{-1} to compute $Tr(I - F A_\alpha)$ as follows:

$$Tr(I - F A_\alpha) = Tr(I - F \Lambda^{-1} F^T W^{-2}) + \alpha Tr(Q^T \Lambda^{-1} F^T W^{-2} F \Lambda^{-1} Q B_\alpha^{-1}) \,. \tag{18}$$

V_α is then computed as [5]

$$V_\alpha = \frac{\frac{1}{M} F_\alpha}{[\frac{1}{M} Tr(I - F\Lambda^{-1}F^T W^{-2} + Q^T \Lambda^{-1} F^T W^{-2} F \Lambda^{-1} Q B_\alpha^{-1})]^2} \,. \tag{19}$$

Equations 11, 13, 14, and 19 are computed repeatedly to find the minimum α. The $\hat{\mathbf{a}}_\alpha$ corresponding to the minimum α is taken to be the estimate of the unknown true function at times $t_n, n = 1, 2, ..., N$. Similarly, $\hat{\mathbf{c}}$ is the $\hat{\mathbf{u}}_\alpha$ corresponding to the minimum α. The resulting cubic spline can then be computed by standard techniques.

3. ERROR ESTIMATES

Further linear algebra manipulations yield error estimates for $\hat{\mathbf{a}}$. The covariance matrix for $\hat{\mathbf{a}}$ is found to be

$$C = (I - \alpha\Lambda^{-1} Q B_\alpha^{-1} Q^T)\Lambda^{-1} F^T W^{-2} F \Lambda^{-1} (I - \alpha \Lambda^{-1} Q B_\alpha^{-1} Q^T)^T \,. \tag{20}$$

The standard deviations are computed by taking the square root of the main diagonal elements of C. Unfortunately, computation of C involves inversion of the full matrix B_α, which might be prohibitively expensive. Another option is Monte Carlo estimates of C, which can be tractable, depending on the number of realizations required. Note that these error estimates are simply composed of the errors caused by imperfect measurements. They are the “R-errors” in [12], [13]. They do not include the “F-errors”, or errors due to smoothing [12], [13]. Part of the reason for the model accuracy checks in Section 5 is to address possible F-errors.

4. APPLICATION TO SIMULATED AND REAL DATA

In Figure 1, we show the results of applying the algorithm to real GRH data from OMEGA (July 30, 2009). We use standard deviations of 0.05, 0.1, 0.5, and 1.0 for the noise on channels one through four, respectively. Also, we scale the time values from seconds to nanoseconds before calculating the various matrices that involve time. The upper left plot in Figure 1 shows the actual data for separate channels in each subplot. The upper right plot shows the actual data with the cubic smoothing spline estimate superimposed. The bottom left plot shows the V_α curve for various α values. The estimate used corresponds to the α with the minimum value of V_α. The lower right plot in Figure 1 shows the corresponding standard deviation estimates. Note that the cubic smoothing spline appears to fit rather well over most of the signal, except at very sharp peaks and troughs.

In order to more clearly demonstrate the performance of the cubic smoothing spline fitting method, we craft simulated data from the OMEGA GRH data used in Figures 1 by fitting a sum of gaussians to the data and adding zero-mean gaussian white noise with standard deviations of 0.05, 0.1, 0.5, and 1.0 for the noise on channels one through four, respectively. Figure 2 shows plots for the results for one realization of the simulated data. Since we have the “true” values for the simulation, we plot these in the upper right subplot of Figure 2. In the same subplot we also plot the residual between the cubic smoothing spline and the true value. We see good results everywhere except at the very sharp peaks. In an effort to quantitatively evaluate the performance of the algorithm, we ran 100 realizations and found the mean value at 233.7760 ns (near the apex of one very sharp peak) to be 29.7, compared with the true value of 31.9.

If we fit the cubic smoothing spline to only the data around the peak, the bias of the fit decreases (Figure 3). Again, we ran 100 realizations and found the mean value at 233.7760 ns to be 31.5, compared with the true value of 31.9. Such behavior indicates that robust application of the cubic smoothing spline fitting method described in this report must include model accuracy checks, which is the subject of Section 5.

5. MODEL ACCURACY CHECKS

In an algorithm that automatically chooses the model parameters, it is desirable to have a check on the accuracy of the model. If the model is inaccurate, then neither the prediction nor the error bounds can be trusted. The model accuracy can be checked using a few facts about the noise and using the difference between the measurement and the model, called either the error or innovation sequence. The model accuracy methods in this section are summarized from [14].

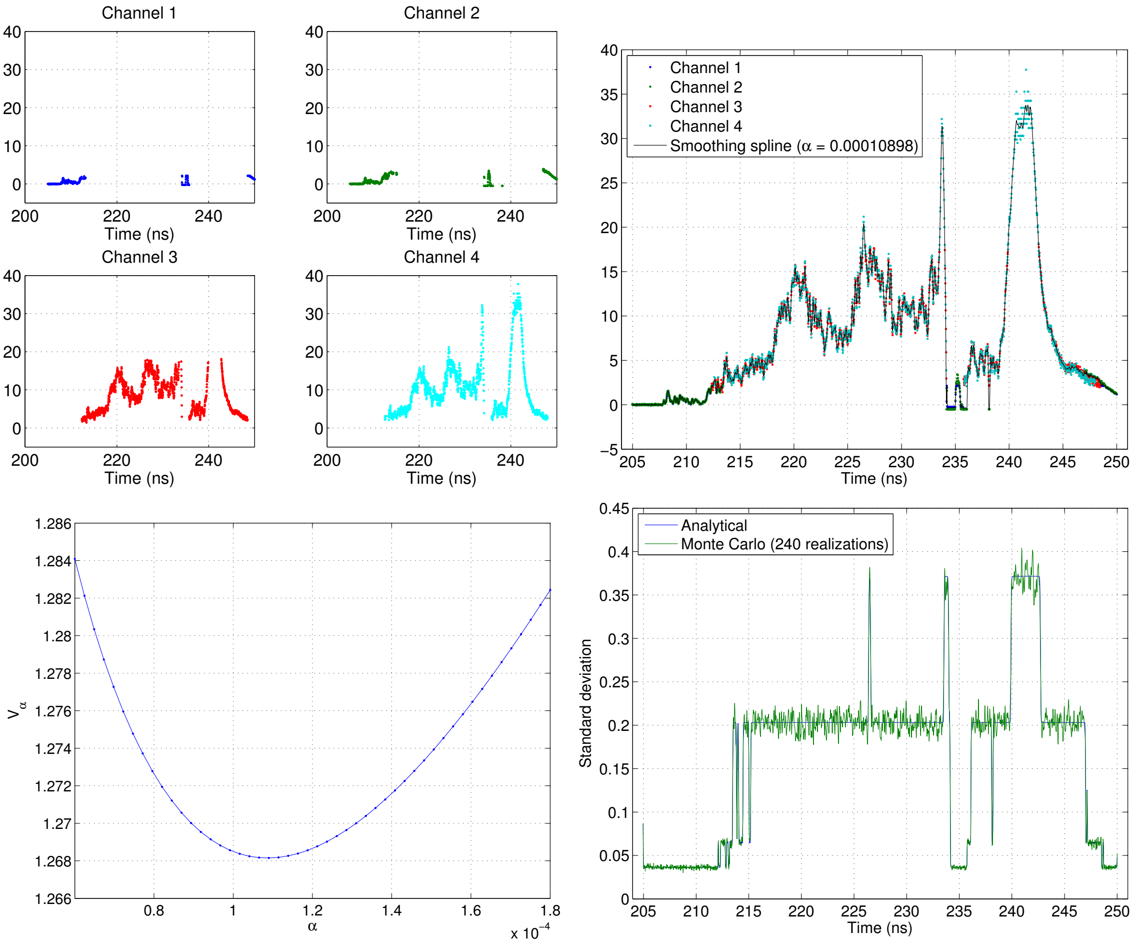

Figure 1. Results of applying the cubic smoothing spline algorithm to real GRH data from OMEGA. The upper left plot shows the actual data for separate channels as subplots. The upper right plot shows the actual data with the cubic smoothing spline estimate superimposed. The bottom left plot displays the V_α curve for various α values. The lower right plot shows the corresponding standard deviation estimates.

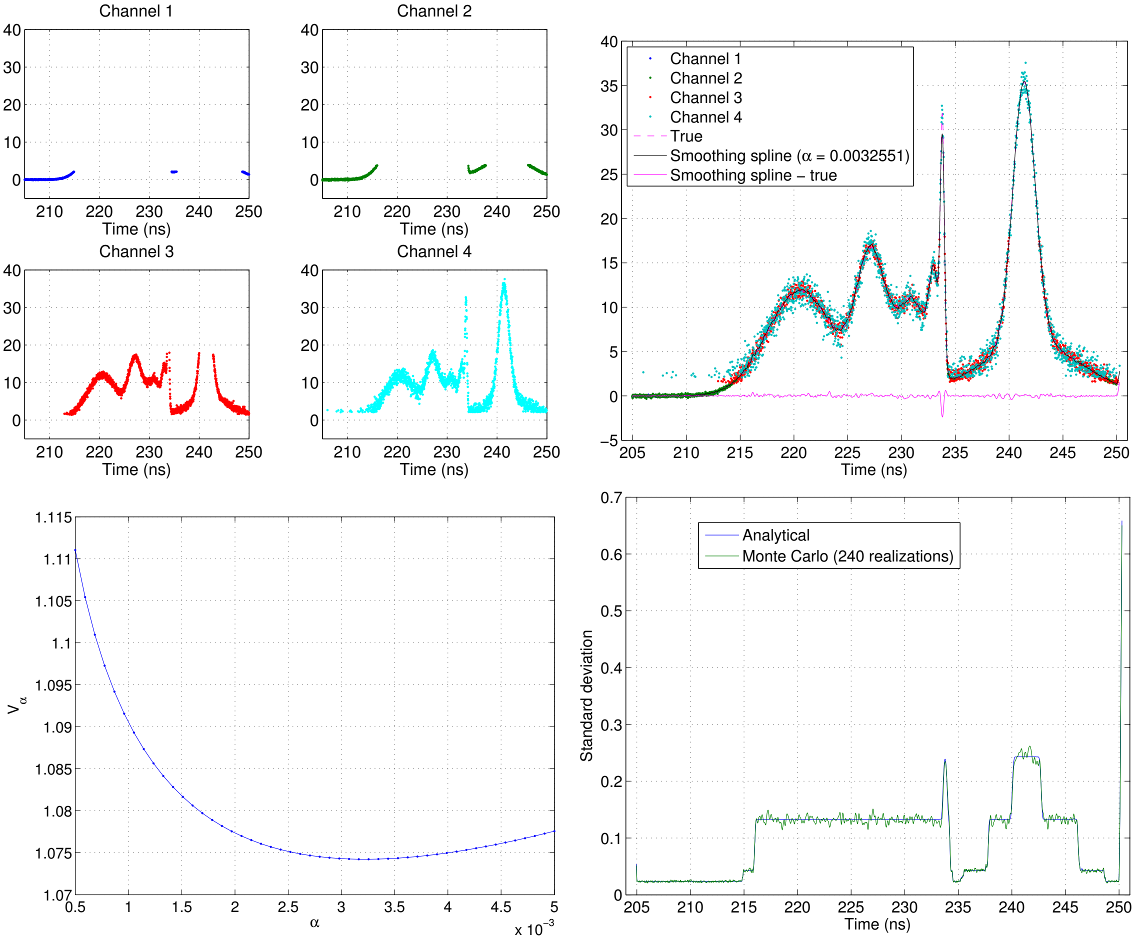

Figure 2. Results of applying the algorithm to simulated data. See Figure 1 for an explanation of the plots. The "true" values and the actual errors (smoothing spline minus true) are added to the upper right subplot of this figure.

If the model is accurate, the innovation sequence should consist of only noise, since the signal will have been mostly extracted from the measurement. If the noise is zero-mean and white, this implies that the innovation sequence will be zero-mean and white. Checking if the innovation sequence is zero-mean is straightforward and is done by computing the mean of the innovation sequence in a window centered at each sample point. If the noise is gaussian, then the WSSR test is a tractable test for whiteness. The formula,

$$\mathrm{WSSR}(p,K) = \sum_{k=p-K+1}^{p} \mathbf{e}^T(k) R_{ee}^{-1}(k) \mathbf{e}(k) \tag{21}$$

where p is the time sample being tested, $\mathbf{e}(k)$ represents the innovation sequence at sample k, K is the window size, and $R_{ee}(k)$ is the covariance of the innovation at sample k. Since the covariance of the innovation sequence is known on each channel, and each sample is assumed to be independent, the covariance of each sample is available and the WSSR statistic can be calculated. For a white, zero-mean innovation sequence, these statistics should be near zero, within a selected threshold.

For both the zero-mean test and the WSSR test, two parameters can be chosen. The first is the window length K. A longer window will smooth the resulting curve while too short of a window will not contain enough samples for the measurement. We choose this parameter by experiment. The second is the threshold at which a sample is considered to fail the test, which is chosen by giving a level of significance. For example, if a significance

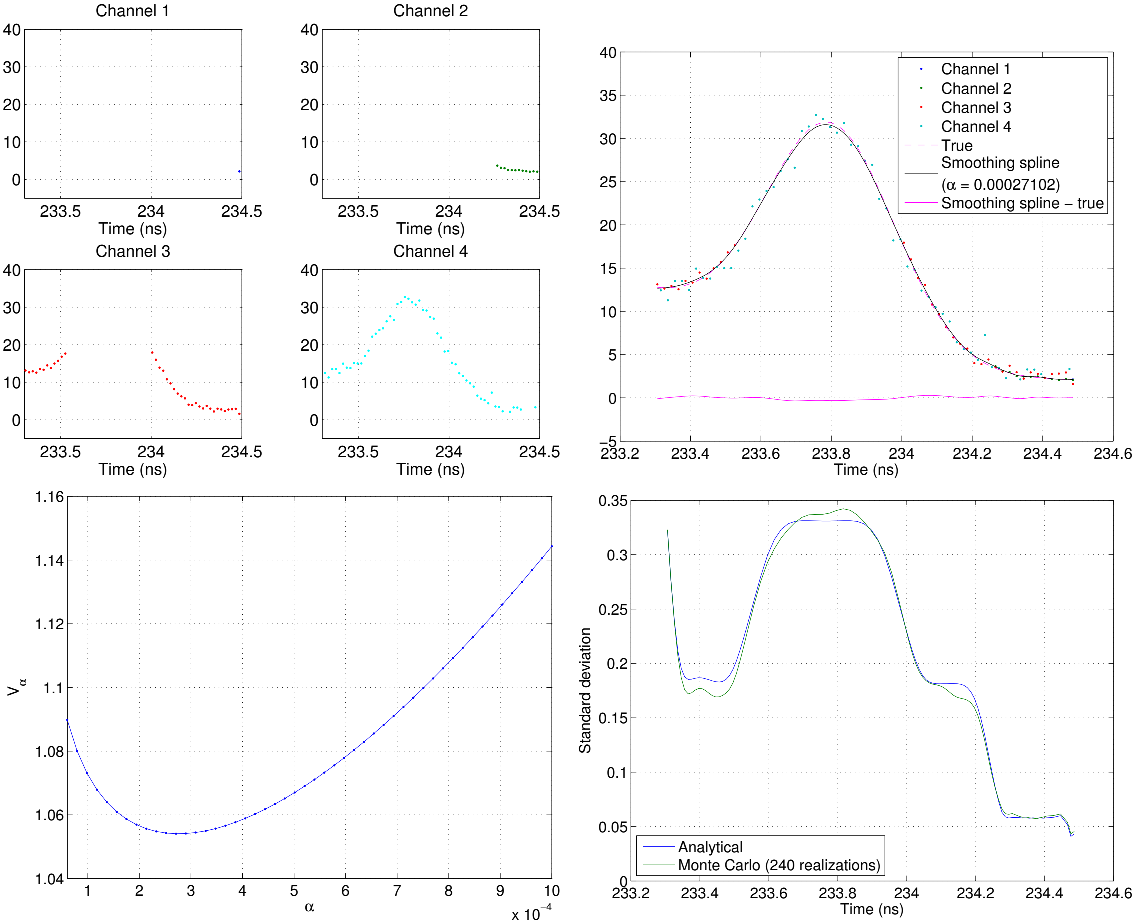

Figure 3. Results of applying the algorithm to a subset of the simulated data in Figure 2 which contains only the peak at 233.7760 ns.

level of 5% is desired, using the gaussian distribution the zero-mean threshold is

$$\tau_{zero\,mean} = 1.96\sqrt{\frac{R_{ee}}{K}}\,, \tag{22}$$

and the WSSR threshold is

$$\tau_{WSSR} = K + 1.96\sqrt{2K}\,. \tag{23}$$

Thus, if less than 5% of the innovation sequence fails the test, the sequence is statistically zero-mean and white, implying that the model captures all of the information present in measurement.

6. SUMMARY AND CONCLUSIONS

We have modified an existing cubic smoothing spline algorithm by Hutchinson and deHoog [5] to accept redundant time samples. The resulting algorithm performs well on both simulated and real data. However, it exhibits smoothing errors for sharp peaks when their time span is small compared with the total time range of the data. We have applied model accuracy checks in an effort to detect when such smoothing errors, and other failures of the code, occur. Future work in making the routine more robust should include time-varying smoothing parameters or variable knot location methods, such as in [12], [13]. One approach could be to use model error checks to localize the smoothing errors and then use the model error checks to guide adaptation of the smoothing parameters on a more local level.

ACKNOWLEDGMENTS

We thank Kirk Miller (NSTec Special Technologies Laboratory), Hans Herrmann (Los Alamos National Laboratory), and the other GRH personnel for the data used in this paper. This work performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344.

REFERENCES

[1] Moses, E., "Ignition on the National Ignition Facility: a path towards inertial fusion energy," *Nuclear Fusion* **49**, 104022 (9 pp.) (October 2009).

[2] Herrmann, H., Mack, J., Young, C., Malone, R., Stoeffl, W., and Horsfield, C., "Cherenkov radiation conversion and collection considerations for a gamma bang time/reaction history diagnostic for the NIF," *Review of Scientific Instruments* **79**, 10E531 (3 pp.) (October 2008).

[3] McEvoy, A. M., Herrmann, H. W., Horsfield, C. J., Young, C. S., Miller, E. K., Mack, J. M., Kim, Y., Stoeffl, W., Rubery, M., Evans, S., Sedillo, T., and Ali, Z. A., "Gamma bang time analysis at omega," *Review of Scientific Instruments* **81**(10), 10D322 (2010).

[4] Ansley, C. F. and Kohn, R., "Spline smoothing with repeated values," *Journal of Statistical Computation and Simulation* **25**(3), 251–258 (1986).

[5] Hutchinson, M. and de Hoog, F., "Smoothing noisy data with spline functions," *Numerische Mathematik* **47**(1), 99–106 (1985).

[6] Craven, P. and Wahba, G., "Smoothing noisy data with spline functions-estimating the correct degree of smoothing by the method of generalized cross-validation," *Numerische Mathematik* **31**, 377–403 (1979 1979).

[7] Fessler, J., "Nonparametric fixed-interval smoothing with vector splines," *IEEE Transactions on Signal Processing* **39**, 852–9 (April 1991).

[8] Koenker, R., "Smoothing Things Over: Some Notes on Cross-Validated Smoothing Splines." Online: www.econ.uiuc.edu/∼roger/courses/574/lectures/L17.pdf (2009).

[9] Rodgers, C., [*Inverse Methods for Atmospheric Sounding: Theory and Practice*], World Scientific (2000).

[10] Ruf, C., "AOSS 585 - Introduction to Remote Sensing and Inversion Theory," Class Notes (2001).

[11] Woltring, H., "A FORTRAN package for generalized, cross-validatory spline smoothing and differentiation," *Advances in Engineering Software* **8**, 104–13 (April 1986).

[12] Blair, J., "Error estimates derived from the data for least-squares spline fitting," in [*2007 IEEE Instrumentation and Measurement Technology Conference Proceedings*], 964–9 (2007).

[13] Blair, J., "Optimal knot selection for least-squares fitting of noisy data with spline functions," in [*2008 IEEE Instrumentation and Measurement Technology Conference (IMTC '08)*], 27–32 (2008).

[14] Candy, J. V., [*Model-Based Signal Processing*], Wiley - IEEE Press (2006).