

In *Bayesian Inference and Maximum Entropy Methods in Science and Engineering*, pp. 604-614, A. Mohammad-Djafari, ed., AIP Conf. Proc. **568**, AIP, Melville, NY, 2001.

INFERENCE FROM ROSSI TRACES

KENNETH M. HANSON AND JANE M. BOOKER
Los Alamos National Laboratory, MS P940
Los Alamos, New Mexico 87545, USA[†]

Abstract.

We present an uncertainty analysis of data taken using the Rossi technique in which the horizontal oscilloscope sweep is driven sinusoidally in time while the vertical axis follows the signal amplitude. The analysis is aimed at determining the logarithmic derivative of the amplitude as a function of time. Within the Bayesian framework used, inferences are obtained with the Markov Chain Monte Carlo technique, which produces random samples from the posterior probability distribution expressed in terms of the parameters.

Key words: Rossi technique, Bayesian inference, uncertainty estimation, Markov Chain Monte Carlo (MCMC), Rossi alpha, smoothing splines, hyperparameter estimation

1. Introduction

We present a Bayesian analysis of data acquired using the Rossi technique. This analysis problem is interesting because the inferred time-dependent signal is not linearly related to the basic measurements. Rather than seeking the maximum posterior estimate, we will emphasize the probabilistic character of Bayesian analysis by using Markov Chain Monte Carlo (MCMC) to make inferences. The MCMC samples from the posterior can be displayed in terms of the inferred signal to visualize its overall uncertainties. The posterior mean estimate of the time-dependent curve and its uncertainties are obtained.

2. Rossi technique

It often happens that one wants to record a signal that is monotonically increasing with time. When the signal is supraexponential, most of the amplitude increase may occur near the end of the time interval being recorded. If the signal is being recorded on an analog oscilloscope, the trace may fall mostly outside the oscilloscope's central region where its linearity is usually best.

[†]Email: kmh@lanl.gov, jmb@lanl.gov; WWW: <http://www.lanl.gov/home/kmh/>

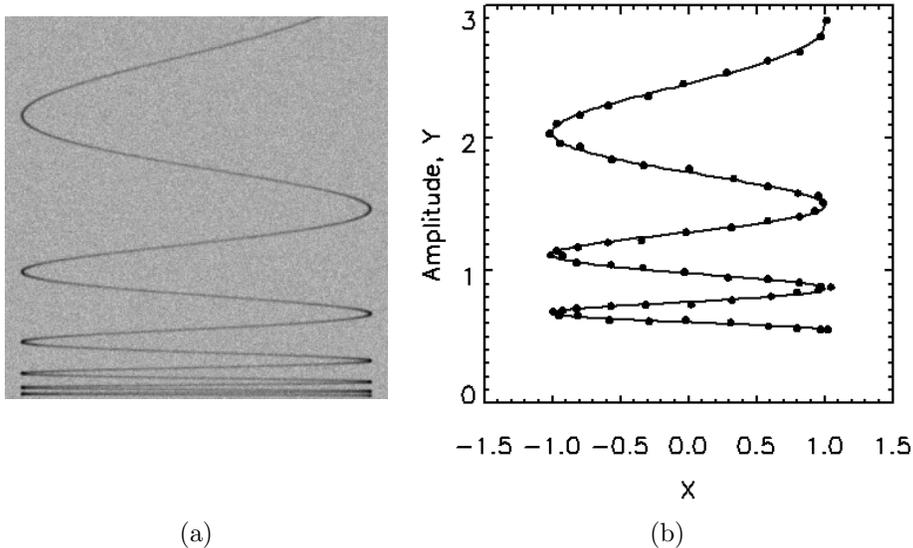


Figure 1. (a) A simulated photograph of a Rossi trace on an analog oscilloscope and (b) the points that might be obtained by manually reading a portion of such a film, compared to the underlying true Rossi trace.

In the Rossi technique for displaying a time-dependent signal [1], the horizontal sweep of an analog oscilloscope is driven sinusoidally in time and the vertical sweep is driven by the signal to be recorded. The major advantage of this unusual recording technique is that the trace dwells in the central region of the oscilloscope face, where the linearity of the oscilloscope response is best. The sinusoid signal also provides a built-in time marker. Figure 1a represents a simulated photograph of a Rossi oscilloscope trace with these features.

The first step in interpreting the Rossi traces is to read the photograph taken of the oscilloscope signal. In the reading process a technician manually places cross hairs over the Rossi trace at a succession of locations. A computer records the cross-hair positions for subsequent analysis. The points in Fig. 1b show the kind of data that one might get from such a reading process. Uncertainties in the positions of the points are included in these simulated data by displacing the true readings by perturbations in x and y randomly drawn from Gaussian distributions.

We assume that the signal being recorded is band limited and has a known time-resolution function. To be specific, the frequency response of the input circuitry is assumed to drop to 50% at the Rossi frequency f_R . We do not attempt to recover frequencies in $\alpha(t)$ higher than $2f_R$, which would result in the ill-posed problem of deblurring or signal recovery.

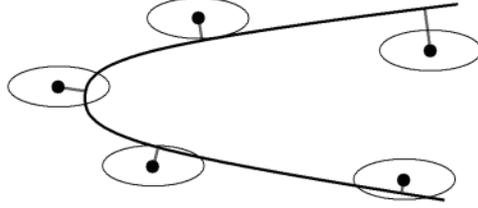


Figure 2. The contribution to the likelihood for each data point is based on the components of the shortest vector between the data point and the modeled Rossi curve.

3. Model-based analysis

To establish our notation, the vertical axis in Fig. 1b is proportional to the signal amplitude, whose time dependence is $y(t)$. The horizontal axis is sinusoidally driven in time as

$$x(t) = x_R \cos [2\pi f_R(t - t_0)] \quad , \quad (1)$$

where f_R is the Rossi frequency and t_0 is the time at which the Rossi sweep starts. We assume that t_0 and f_R are precisely known. We describe the measured points along the trace as the data set, $\{x_i, y_i\}$.

The aim of the present analysis is to determine from the measurements not merely $y(t)$, but the relative time rate of change of the signal amplitude as a function of time:

$$\alpha(t) = \frac{1}{y} \frac{dy}{dt} \quad . \quad (2)$$

This quantity, often referred to as the Rossi alpha [2,3], is a measure of the criticality of an assembly of fissile material in which the amplitude signal y is the measured flux from the assembly. The units of alpha are $[t^{-1}]$.

Our approach is to model the alpha curve in terms of cubic splines. For a specified $\alpha(t)$, Eq. (2) is integrated to get

$$y(t) = y_0 \exp\left(\int \alpha(t) dt\right) \quad , \quad (3)$$

where y_0 is the y amplitude at $t = 0$, which we assume is known. This $y(t)$, together with (1), predicts the Rossi curve. Inferences about model parameters are based on the comparison of the predicted Rossi curve to the $\{x_i, y_i\}$ data set.

3.1. UNCERTAINTY MODEL FOR THE MEASUREMENTS

Figure 2 shows our approach to assigning the likelihood, which quantifies the probability of the measurements \mathbf{d} for any specified Rossi curve. We propose using for the minus-log-likelihood

$$-\log[p(\mathbf{d}|\alpha(t))] = \frac{1}{2}\chi^2 = \frac{1}{2} \sum_i \left[\frac{(x_i - x_i^*)^2}{\sigma_x^2} + \frac{(y_i - y_i^*)^2}{\sigma_y^2} \right] \quad , \quad (4)$$

where (x_i, y_i) is the measured position of the i th data point and (x_i^*, y_i^*) is the position of the nearest point on the Rossi curve. The uncertainties in x and y , given by σ_x and σ_y , are assumed to be independent of i . These two uncertainties may be chosen to be different to allow for larger uncertainties in the peak regions, where σ_x dominates, than elsewhere. In any given application, one should confirm that the likelihood model properly captures the probability distribution of the actual uncertainties in the data. The quadratic form of this expression comes from the assumption that the uncertainties follow Gaussian distributions. The sum over contributions from individual data points is valid only if the uncertainties in measuring a point are independent of other measurements. Gull [4] used a similar model for the likelihood to tackle the complex problem of fitting a straight line to data points that have uncertainties in both x and y .

3.2. SPLINE EXPANSION

We model the function of interest, alpha as a function of time, in terms of cubic B splines, which are chosen for their smoothness properties. For uniformly spaced basis functions, we write the continuous alpha curve as

$$\alpha(t) = \sum_{k=1}^K a_k \phi\left(\frac{t-t_k}{\Delta t}\right) , \quad (5)$$

where $\phi\left(\frac{t-t_k}{\Delta t}\right)$ is a basis function centered on the knot time t_k and Δt is the spacing between knots. To respect the assumed band limitation of the system mentioned above, we choose $\Delta t = 0.25 f_R^{-1}$. The corresponding Nyquist frequency is $2f_R$, high enough to accommodate signals with the assumed 50% attenuation at f_R .

The cubic B-spline basis function is defined as:

$$\phi(x) = \begin{cases} 1 - \frac{3}{2}|x|^2 + \frac{3}{4}|x|^3 & , |x| \leq 1 \\ \frac{1}{4}|2 - |x||^3 & , 1 < |x| < 2 \\ 0 & , |x| \geq 2 \end{cases} . \quad (6)$$

Figure 3 shows the spline curve with the knot positions, the t_k in Eq. (5), for a linear alpha dependence over three Rossi cycles. There are 15 knots in all; 13 lie within the data interval ($0 \leq t \leq 3$). Two additional spline knots are present but not shown, one beyond either end of the interval covered by the data. These are included to provide the same functional dependence at the ends of the interval as elsewhere. This approach differs from the usual assumption in spline theory that either the first or second derivative of the function is zero at the end of the interval [5].

3.3. BAYESIAN INFERENCE

Our goal is to make inferences about the spline model for alpha from the x_i, y_i data points. In the Bayesian approach, the uncertainty in the value of a model parameter is represented by a probability density function (pdf). Bayes law gives

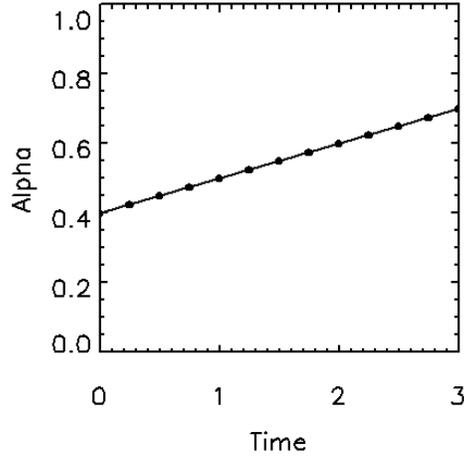


Figure 3. The model used to interpret the Rossi data. We represent alpha as a function of time in terms of a spline expansion. The spline knot positions are shown as dots.

the pdf for the vector of model parameters \mathbf{a} , predicated on measurements \mathbf{d} :

$$p(\mathbf{a}|\mathbf{d}, \mathcal{I}) = \frac{p(\mathbf{d}|\mathbf{a}, \mathcal{I}) p(\mathbf{a}, \mathcal{I})}{p(\mathbf{d}|\mathcal{I})} , \quad (7)$$

where $p(\mathbf{d}|\mathbf{a}, \mathcal{I})$ is the likelihood and $p(\mathbf{a}, \mathcal{I})$ is the prior on the parameters. The symbol \mathcal{I} represents all background information about the situation at hand, including information about the experiment, the apparatus, and so forth. It is meant to remind us that analysis should not be done in a vacuum; prior knowledge always plays a role. From here on, we will drop \mathcal{I} from the probability expressions. The quantity $p(\mathbf{a}|\mathbf{d}, \mathcal{I})$ is called the posterior and summarizes our knowledge about the parameters after we combine the measurements with what we knew beforehand.

The denominator in (7) is called the evidence. It can be thought of as the probability of the data (given the model)

$$p(\mathbf{d}) = \int p(\mathbf{d}|\mathbf{a}) p(\mathbf{a}) d\mathbf{a} . \quad (8)$$

This denominator ensures the proper normalization of posterior, $\int p(\mathbf{a}|\mathbf{d}) d\mathbf{a} = 1$. Because it does not depend on the parameters, it can be ignored when one is concerned only with the parameters \mathbf{a} . However, as we shall see later, the evidence becomes the focus of attention when we are concerned about hyperparameters or selecting the best model to describe the data [6].

3.4. SMOOTHNESS PRIOR

We know that the alpha curve must possess a certain degree of smoothness because $y(t)$ is band limited. While the spline representation is supposed to provide smooth curves, it tends to produce oscillations in some applications [7]. To promote the

smoothness of a function, one often chooses [5,8] to minimize the integral over the interval T of the square of the second derivative of the function:

$$S(\alpha) = T^3 \int_T \left(\frac{d^2\alpha}{dt^2} \right)^2 dt . \quad (9)$$

The T^3 factor makes S depend only on the shape of $\alpha(t)$ and independent of T . The units of S are those of α^2 or $[t^{-2}]$. Using the expansion for $\alpha(t)$, Eq. (5), this functional can be expressed in terms of the coefficients \mathbf{a} . Oscillations in the splines can be controlled through the prior in Bayes law, Eq. (7). The minus-log-prior on \mathbf{a} is taken to be $\lambda S(\mathbf{a})$, where λ determines the strength of this prior. To make this quantity dimensionless, λ must have units of $[t^2]$. The parameter λ is called a hyperparameter because it directly affects a pdf, instead of the model describing the signal.

An alternative approach to controlling spline oscillations [7] is to allow the number and positions of the knots to vary, both of which are chosen by maximizing the evidence using a process similar to that described in Sect. 5.2. That approach involves much more difficult calculations but has the advantage that the smoothness is variable and adapts to the requirements of the data.

3.5. SYSTEMATIC EFFECTS

In describing the measurements in Sect. 2, we did not mention several important aspects of the measurement process. For example, it is essential to determine the location of the baseline for the amplitude measurements, y_B , since all values of the y position of the data points are relative to this baseline. If y_B is measured in a manner similar to that for the data points, we expect the uncertainty in y_B to be comparable to that in the y position of the data points. To include this uncertainty in our analysis, we add to the minus-log-likelihood for the data points, Eq. (4),

$$\frac{1}{2} \left[\frac{(y_B - y_B^*)^2}{\sigma_{y_B}^2} \right] , \quad (10)$$

where y_B is the measured value of the baseline and y_B^* is the value in our model. The standard deviation in measuring the baseline, σ_{y_B} , is assumed to be 0.02.

To take y_B into account in our model, y in Eq. (2) has to be replaced by $y - y_B^*$. This baseline represents a systematic uncertainty because its value impacts many other model parameters.

Another systematic effect is the amplitude of the Rossi sweep, x_R , in Eq. (1). This is a parameter in the model that we are using to predict the x, y data. It is not directly measured in our scenario. Since the minus-log-likelihood is a sum over contributions from measurements, there is no term in the minus-log-likelihood for x_R , as there is for y_B above. However, since x_R is a parameter in our model, which influences predictions for the Rossi curve, it can be inferred in our analysis.

There are several more aspects of the experiment that might contribute significant systematic effects, e.g., y_0 , t_0 , f_R , and geometrical distortions in the x, y measurements. We ignore them to simplify the present analysis.

3.6. THE FULL POSTERIOR

To summarize, the full minus-log-posterior for our analysis is

$$-\log[p(\mathbf{a}|\mathbf{d})] = \frac{1}{2}\chi^2 + \frac{1}{2} \left[\frac{(y_B - y_B^*)^2}{\sigma_{y_B}^2} \right] + \lambda S(\mathbf{a}) , \quad (11)$$

where $\frac{1}{2}\chi^2$ is given by Eq. (4) and $S(\mathbf{a})$ by Eq. (9). The first two terms represent likelihood contributions and the last term comes from the prior on smoothness. There are a total of 17 variable parameters in our model about which we may draw inferences on the basis of Eq. (11), namely the 15 coefficients in the spline expansion (5), x_R , and y_B .

4. Markov Chain Monte Carlo

The Markov Chain Monte Carlo (MCMC) technique provides a means to generate a random sequence of model realizations that sample the posterior probability distribution of a Bayesian analysis. The sequence may be used to make inferences about the model uncertainties that derive from measurement uncertainties. The usefulness of MCMC in Bayesian inference is well established [9–11].

The simplest MCMC approach is to use the Metropolis algorithm [12] to construct the sequence. In the Metropolis algorithm, one tries to move from the current position in parameter space by randomly selecting a trial step from a symmetric probability distribution. The trial step is either accepted or rejected on the basis of the probability at the new position relative to the previous one. This algorithm is widely employed because of its simplicity. We use the Metropolis algorithm and omit the details for lack of space. Many issues need to be paid attention, including choice of the trial pdf, assurance of convergence of the sequence to the posterior, and burn in. By using a suitable trial function, we achieve an estimated efficiency for the MCMC sampling of 0.17% for our example (at $\lambda = 0.4$).

5. Results

We demonstrate our analysis technique using the simulated data shown in Fig. 1b. The 61 data points are based on an alpha curve that rises linearly from 0.4 to 0.7 at the times of 0 and 3 (Rossi cycles), respectively. The x, y values of each data point are generated by adding random perturbations drawn from Gaussian distributions with $\sigma_x = 0.02$ and $\sigma_y = 0.01$.

5.1. POSTERIOR SAMPLING USING MCMC

Figure 4 shows five samples drawn from the posterior, Eq. (11), for our model for $\lambda = 0.04$, somewhat lower than its optimal value. Because successive samples in an MCMC sequence are highly correlated, these five samples are separated by 2000 steps to minimize correlations between them. This kind of display of model realizations is a good way to visualize the characteristics of an inferred model [13]

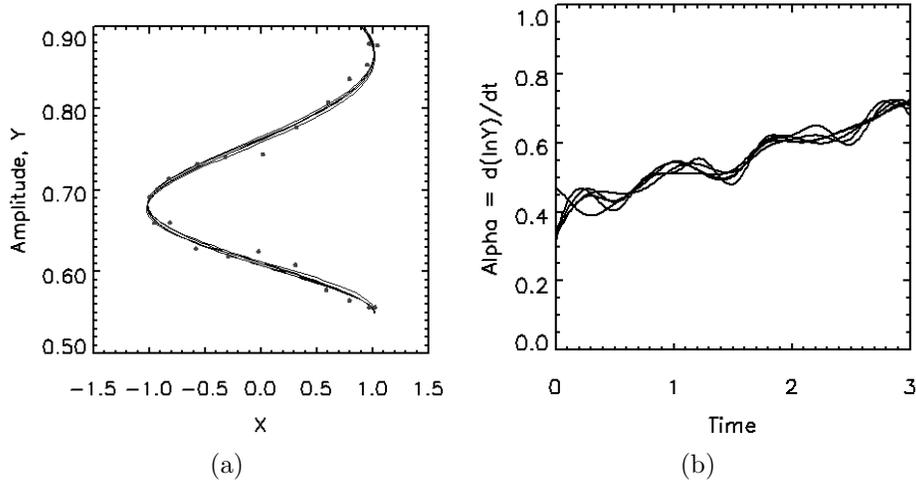


Figure 4. Five widely separated samples from an MCMC sequence, shown in both (a) the data domain and (b) the alpha domain. The hyperparameter λ is 0.04 for this case. Viewing these samples provides one with an understanding of the type of curves that are admissible within the framework of the model used to interpret the data.

as well as its uncertainties. By showing a representative set of plausible solutions consistent with the data, the degree of variability of this presentation provides the viewer with a visual impression of the degree of uncertainty in the inferred model.

Of course, MCMC is more than a tool for visualizing uncertainties; it provides samples from the posterior from which quantitative estimates of the uncertainty in the inferred models may be derived. The uncertainty in any aspect of the model may be estimated with respect to any type of uncertainty measure desired, for example, in terms of variance or confidence interval. A notable advantage of MCMC is that the results are obtained with marginalization with respect to any nuisance parameters. In our problem, we are not interested in the two systematic parameters, y_B and x_R . The uncertainties in these parameters are integrated out by the MCMC process. An MCMC sequence can also be used to estimate the posterior mean (as an alternative to the posterior mode).

To average over our MCMC sequence, we lay the curves down on a pixelated image that spans the region of interest. Each curve adds a value to the pixels it covers. At the end of the process, the value of each pixel in the image is proportional to the number of curves that fell on top of it. Figure 5 shows two such images. Figure 5a is obtained with a minimal prior. It demonstrates the tendency of splines to oscillate. The peaks and valleys of the envelope occur at the spline knots where the uncertainties are largest. Figure 5b is for $\lambda = 0.4$, which is approximately the value favored by maximizing the evidence in Bayes law, Eq. (7), as discussed next. The severe oscillations seen in Fig. 5a are well controlled by the stronger smoothness prior in Fig. 5b. The remaining wiggles in Fig. 5b are caused by random errors in the particular x, y data set that we are analyzing. The choice of the hyperparameter λ is clearly very important to making the correct inferences.

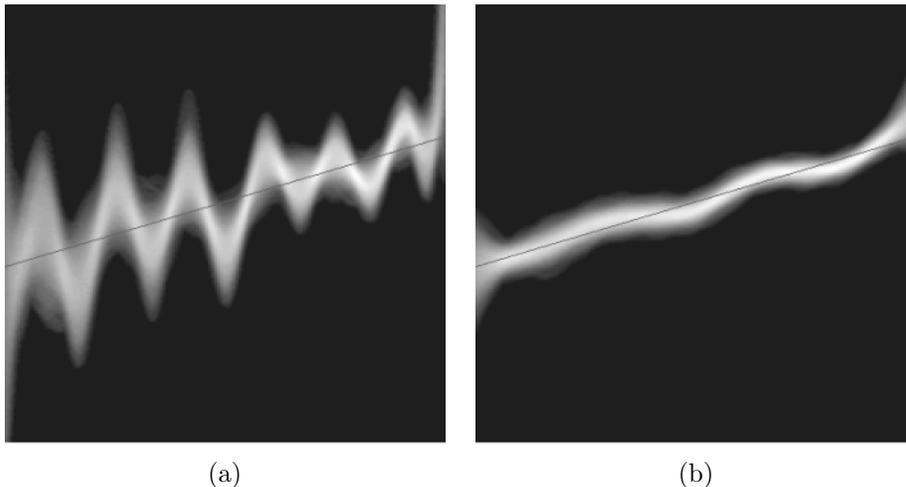


Figure 5. The posterior distributions for the alpha curves for two values of the hyperparameter that controls the strength of the smoothness prior: (a) $\lambda = 0.0004$ and (b) $\lambda = 0.4$. The true alpha curve is shown as a straight line. These images have the same axes as Fig. 4b.

5.2. ESTIMATION OF HYPERPARAMETER λ

In the Bayesian approach to determining hyperparameters [6], the pdf for λ can be obtained by integrating the joint distribution for λ and \mathbf{a} over \mathbf{a} :

$$p(\lambda|\mathbf{d}) = \int p(\mathbf{a}, \lambda|\mathbf{d}) d\mathbf{a} \propto \int p(\mathbf{d}|\mathbf{a}, \lambda) p(\mathbf{a}, \lambda) = p(\lambda) \int p(\mathbf{d}|\mathbf{a}, \lambda) p(\mathbf{a}) d\mathbf{a} , \quad (12)$$

where we have assumed that the prior on λ is independent of that on \mathbf{a} . For a fairly flat prior on λ , the dependence of this expression on λ is dominated by the integral on the right-hand side of the equation, in other words, the evidence defined in Eq. (8). Within a scaling factor, the integral can be approximated as the peak value of the product $p(\mathbf{d}|\mathbf{a}, \lambda) p(\mathbf{a})$ times its volume in the 17-dimensional parameter space. The volume is proportional to the determinant of the covariance matrix, which we estimate using the MCMC technique by the second moments, $\langle (a_i - \bar{a}_i)(a_j - \bar{a}_j) \rangle$, where the brackets indicate an average over the posterior and \bar{a}_i is the posterior mean value of the i th parameter a_i , that is, $\bar{a}_i = \langle a_i \rangle$.

For this approach to work, it is necessary for the $\alpha(t)$ curve to possess some structure that doesn't minimize the prior. The peak in the evidence comes about when a stronger prior draws the solution away from the actual signal and thus causes the minus-log-likelihood to increase. For the linear $\alpha(t)$ signal shown in Fig. 3, the evidence shows no peak because the linear signal that minimizes the smoothness prior also comes close to minimizing χ^2 . Thus, for this auxiliary calculation, we use a band-limited step function for $\alpha(t)$, which steps from 0.4 to 0.7 at $t = 1.5$. As shown in Fig. 6, the evidence exhibits a maximum around $\lambda = 0.4$, indicating that this is the preferred value.

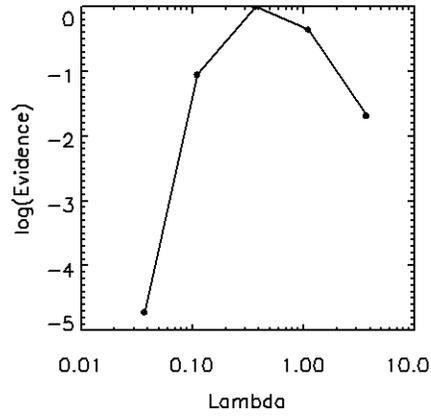


Figure 6. Plot of the normalized evidence versus the hyperparameter λ for a steplike alpha curve.

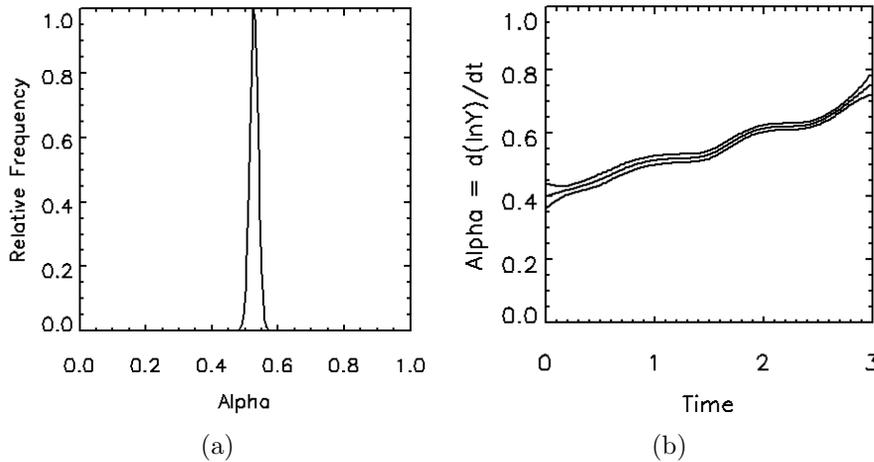


Figure 7. (a) A vertical slice through Fig. 5b ($\lambda = 0.4$) at the $t = 1.5$, which estimates the posterior distribution for alpha at that time. (b) Analysis of Fig. 5b yields the posterior mean estimate for $\alpha(t)$ and its one-standard deviation uncertainty band.

5.3. INFERENCES ABOUT THE ALPHA CURVE

As argued above, each column of Fig. 5b represents the posterior distribution for alpha at a given time, as shown in Fig. 7a. The mean value of alpha for this distribution is 0.53. Its rms deviation of 0.013 or 2.4% is approximately consistent with the assumed $\sigma_y = 0.01$ and $\sigma_{y_B} = 0.02$, which are, respectively, 0.9% and 1.8% of the amplitude $y = 1.15$ at that time. By analyzing each column of Fig. 5b, one can determine the mean of the posterior in alpha and its rms deviation as a function of time. Figure 7b shows the posterior mean and the mean plus and minus one standard deviation. These latter curves represent an uncertainty envelope for

the alpha curves.

One should be aware that the uncertainty envelope in Fig. 7b only refers to the pdf at any particular time, without regard to any other time. Details contained in the posterior distribution have been marginalized out. Specifically, the correlations in uncertainties from one time to another can not be inferred from this envelope. To get an idea of these correlations, one has to go back to the MCMC samples to visualize the correlations, as in Fig. 4. One can quantify the correlations. For example, the correlation between the uncertainties at two different times can be estimated by computing the cross correlation for the MCMC sequence.

Acknowledgements

This work has been supported by the United States Department of Energy under contract number W-7405-ENG-36. For useful discussions and insights, we thank Kent Croasdell, Rick Collinsworth, Layle Zongker, Jamie Langenbrunner, Tom Gorman, Shane Reese, Greg Cunningham, Paul Goggans, Martial Honnell, and Jack Jacobson.

References

1. C. W. Franklin, "Recovery of time-varying exponential signals and other rapidly developing functions," Tech. Rep. 1183-2233, EGG, Santa Barbara, 1970.
2. J. D. Orndoff, "Time scale measurements by the Rossi method," Tech. Rep. LA-744, Los Alamos Scientific Laboratory, Los Alamos, 1949.
3. G. A. Linenberger, J. D. Orndoff, and H. C. Paxton, "Enriched-uranium-hydride critical assemblies," *Nucl. Sci. Eng.*, **7**, pp. 44–57, 1960.
4. S. F. Gull, "Bayesian data analysis: straight-line fitting," in *Maximum Entropy and Bayesian Methods*, J. Skilling, ed., pp. 511–518, Kluwer Academic, Dordrecht, 1989.
5. C. deBoer, *A Practical Guide to Splines*, vol. 27 of *Applied Mathematical Sciences*, Springer, New York, 1978.
6. D. J. C. MacKay, "Bayesian interpolation," *Neural Comput.*, **4**, pp. 415–447, 1992.
7. R. Fischer, K. M. Hanson, V. Dose, and W. von der Linden, "Background estimation in experimental spectra," *Phys. Rev. E*, **61**, pp. 1152–1161, 2000.
8. C. H. Reinsch, "Smoothing by spline functions," *Numer. Math.*, **10**, pp. 237–248, 1967.
9. W. R. Gilks, S. Richardson, and D. J. Spiegelhalter, *Markov Chain Monte Carlo in Practice*, Chapman and Hall, London, 1996.
10. A. Gelman, J. B. Carlin, H. S. Stern, and D. B. Rubin, *Bayesian Data Analysis*, Chapman & Hall, London, 1995.
11. J. Besag, P. Green, D. Higdon, and K. Mengersen, "Bayesian computation and stochastic systems," *Stat. Sci.*, **10**, pp. 3–66, 1995.
12. N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, "Equations of state calculations by fast computing machine," *J. Chem. Phys.*, **21**, pp. 1087–1091, 1953.
13. J. Skilling, D. R. T. Robinson, and S. F. Gull, "Probabilistic displays," in *Maximum Entropy and Bayesian Methods*, W. T. Grandy, Jr. and L. H. Shick, eds., pp. 365–368, Kluwer Academic, Dordrecht, 1991.