LA-UR- 00-4256

Approved for public release; distribution is unlimited.

Title:	INFERENCES FROM ROSSI TRACES
Author(s):	KENNETH M. HANSON, DX-3 JANE M. BOOKER, TSA-1
Submitted to:	PROCEEDINGS BAYESIAN INFERENCE AND MAXIMUM ENTROPY METHODS GIF SUR YVETTE, FRANCE JULY 9-13, 2000

Los Alamos
NATIONAL LABORATORY

Form 836 (8/00)

DISCLAIMER

INFERENCES 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 done within a Bayesian framework. Complete inferences are obtained by using 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, Markov Chain Monte Carlo (MCMC), Rossi alpha, smoothing splines

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 going for a maximum posterior estimate, we will emphasize the probabilistic character of Bayesian analysis by using 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 is obtained, along with uncertainty estimates.

2. Rossi technique

It often happens that one wants to record a signal that is monotonically increasing with time. If that signal is supra-exponential, most of the amplitude increase may occur at the end of the time interval being recorded. If the signal is being recorded on an oscilloscope, the trace may fall mostly outside the oscilloscope's central sweet spot, the area in which the linearity is best.

In the Rossi technique for displaying a time-dependent signal, the horizontal sweep of an analog oscilloscope is driven sinusoidally in time and the vertical sweep

[†]Email: kmh@lanl.gov, jmb@lanl.gov

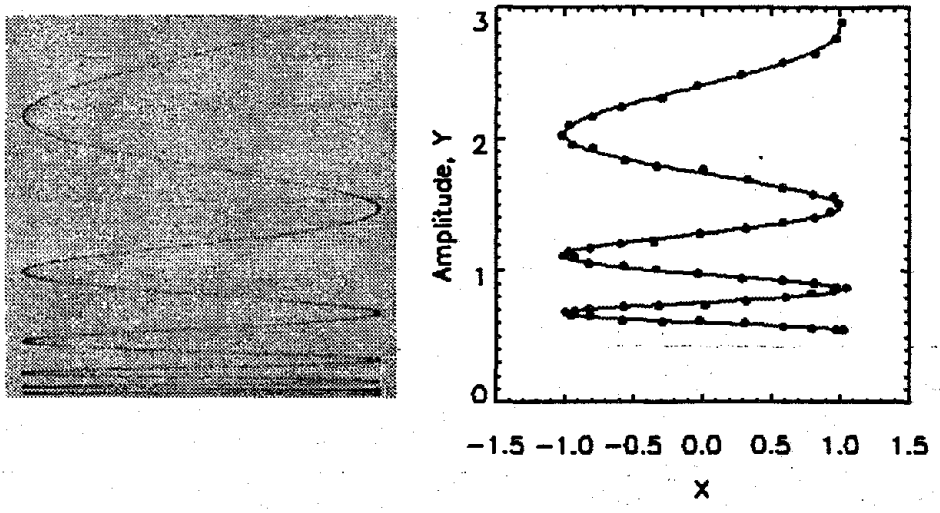

Figure 1. A simulated photograph of a Rossi trace on an analog oscilloscope (left) and the points that might be obtained by manually reading a portion of such a film, compared to the underlying true Rossi trace.

is driven by the signal to be recorded. The major advantage of this unusual recording technique is that the trace is confined to the central region of the oscilloscope face, the region in which the linearity of the oscilloscope response is best. The sinusoid signal also provides a built-in time marker. Figure 1 represents a simulated photograph of a Rossi oscilloscope trace.

The first step in interpreting these Rossi traces is to read the photograph taken of the oscilloscope signal. The reading process amounts to tabulating the position of a succession of points by manually placing cross hairs on the Rossi trace by a technician. A computer records the cross-hair positions for subsequent analysis. Figure 1 shows the kind of data 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 an amount chosen by randomly drawing from Gaussian distributions for the x and y displacements.

The vertical y axis is proportional to the signal amplitude, whose time dependence is $y(t)$. The horizontal x axis is sinusoidally driven in time as

$$x(t) = x_R \cos(2\pi f_R(t - t_0)) \ , \tag{1}$$

where f_R is the Rossi frequency and t_0 is the time at which the Rossi sweep starts. 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, called alpha, as a function of time:

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

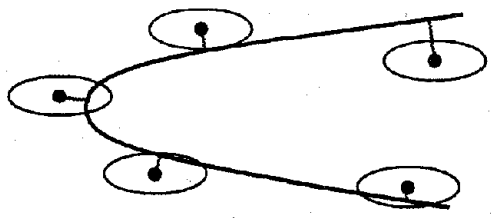

Figure 2. The contribution to the likelihood for each data point is based on the components of the shortest vector between the data point and the modeled Rossi curve.

This quantity, often referred to as the Rossi alpha [1], 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. We assume that the signal being recorded is band limited and has a known time-resolution function. In this analysis, we will assume that the frequency response of the input circuitry drops to 50% at the Rossi frequency f_R. We will not attempt to recover higher frequencies in $\alpha(t)$ than beyond the pass band of the input signal, which constitutes the ill-posed problem of deblurring or signal recovery.

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

$$-\log[p(\mathbf{d}|\mathbf{a})] = \tfrac{1}{2}\chi^2 = \tfrac{1}{2}\sum_i \left[\frac{(x_i - x_i^*)^2}{\sigma_x^2} + \frac{(y_i - y_i^*)^2}{\sigma_y^2}\right] , \qquad (3)$$

where (x_i, y_i) is the measured position of the ith 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 the σ_x and σ_y, are assumed to independent of i. Of course, in any given application, it is best to confirm that the likelihood model properly adheres to 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 individual data points is valid only if the uncertainties in measuring a point is independent of other measurements. Gull [2] used a similar model for the likelihood to tackle the complex problem of fitting a straight to data points that have uncertainties in both x and y.

3. Model-based analysis

3.1. SPLINE EXPANSION

We directly model the function of interest, alpha vs. time, in terms of a cubic B spline. The spline is chosen for its smoothness properties. For uniformly spaced

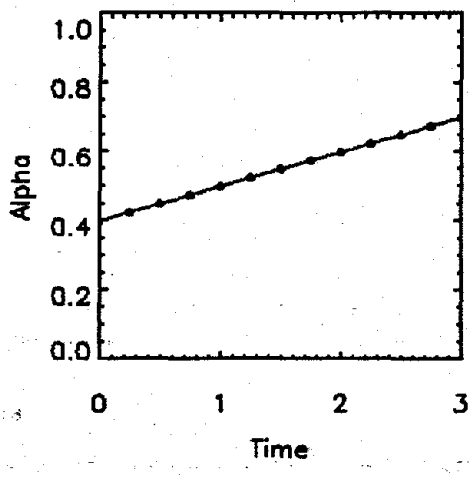

Figure 3. The model used to interpret the Rossi data. In the alpha domain, we use a spline to model alpha as function of time. The spline knot positions are shown as dots.

basis functions, we write the continuous alpha curve as

$$\alpha(t) = \sum_{k=1}^{K} a_k \, \phi\left(\frac{t - t_k}{\Delta t}\right) \ , \tag{4}$$

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

The cubic B-spline basis function is defined as:

$$\phi(x) = \begin{cases} 1 - \frac{3}{2}x^2 + \frac{3}{4}|x|^3 & , \ |x| \le 1 \\ \frac{1}{4}|2 - x|^3 & , \ 1 < |x| < 2 \\ 0 & , \ |x| \ge 2 \ . \end{cases} \tag{5}$$

Figure 3 shows the spline curve with the knot positions, the t_k in Eq. (4), for a linear alpha dependence over three Rossi cycles. 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 for the same functional dependence in the end intervals as elsewhere. This approach differs from the usual assumption that either the first or second derivative of the function is zero at the end of the interval [3].

3.2. BAYESIAN INFERENCE

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

the vector of model parameters **a**, predicated on some data **d**:

$$p(\mathbf{a}|\mathbf{d},\mathcal{I}) = \frac{p(\mathbf{d}|\mathbf{a},\mathcal{I})\,p(\mathbf{a},\mathcal{I})}{p(\mathbf{d}|\mathcal{I})} \quad , \tag{6}$$

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

The denominator in (6) is called the evidence and can be thought of as the probability of the data (given the model)

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

This quantity is required to ensure the proper normalization of posterior, $\int p(\mathbf{a}|\mathbf{d},\mathcal{I})d\mathbf{a} = 1$. It can be ignored when one is concerned only with the parameters **a**. However, as we shall see later, it becomes the focus of attention when we are concerned about hyperparameters or selecting the best model to describe the data [4].

In our situation 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 smoothness to the curves, it tends to produce oscillations, as we shall see later. These oscillations can be controlled through the prior in Eq. (6). To promote the smoothness of a function, one often chooses[3,5] to minimize the integral over the interval T of the square of the second derivative of the function:

$$S(\alpha) = T^3 \int_T \left(\frac{d^2\alpha}{dt^2}\right)^2 dt \quad . \tag{8}$$

The T^3 factor makes S dimensionless. Using the expansion for $\alpha(t)$, Eq. (4), this functional can be expressed in terms of the coefficients **a**. Following common practice, the minus-log-prior on $\alpha(t)$ is taken to be $\lambda S(\mathbf{a})$, where λ determines the strength of this prior. The parameter λ is called a hyperparameter because it directly affects a pdf, instead of the model describing the signal.

3.3. SYSTEMATIC EFFECTS

In describing the measurements, we glossed over several important aspects of the measurement process. For example, it is essential to determine the location of the baseline for the amplitude measurements, y_0, since all values of the y position of the data points must be referred to this baseline. If y_0 is measured in a manner similar to that for the data points, we expect the uncertainty in y_0 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. (3),

$$\frac{1}{2}\left[\frac{(y_0 - y_0^*)^2}{\sigma_{y_0}^2}\right] \quad , \tag{9}$$

where y_0 is the measured value of the baseline and y_0^* is the value in our model.

To take y_0 into account in our model, y in Eq. (2) has to be replaced by $y - y_0$ to calculate the predicted $\alpha(t)$ curve. This baseline represents a systematic uncertainty because its value impacts many other the 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. Since it is not directly measured in our scenario, there is no contribution to the likelihood. It must be inferred from the data points.

There are potentially several more aspects of the measurement process that might be important, e.g., geometrical distortions. We will ignore them in this paper to simplify the present analysis.

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

$$-\log[p(\mathbf{a}|\mathbf{d})] = \tfrac{1}{2}\chi^2 + \tfrac{1}{2}\left[\frac{(y_0 - y_0^*)^2}{\sigma_{y_0}^2}\right] + \lambda S(\mathbf{a}) \ , \tag{10}$$

where $\frac{1}{2}\chi^2$ is given by Eq. (3) and $S(\mathbf{a})$ by Eq. (8). The first two terms represent likelihood contributions and the last term comes from the prior on smoothness.

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 [6–8].

The simplest MCMC approach is to use the Metropolis algorithm [9] 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 of 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.

5. Results

Figure 4 shows five samples drawn from the posterior for our model for two values of λ. 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 [10]. By showing a representative set of plausible solutions, 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 a characterization of the posterior from which quantitative estimates of the uncertainty in the inferred models may be derived. The uncertainty in any aspect

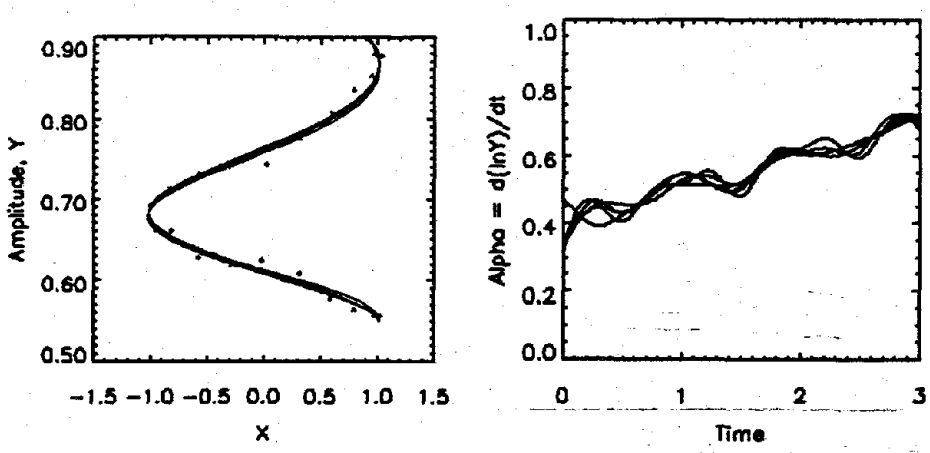

Figure 4. Five widely separated samples from an MCMC sequence, shown in both the data domain (left) and the alpha domain (right). The hyperparameter λ is 0.04 for this case. Viewing these samples provides one with an understanding of the type of curves that are admissible within the framework of the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

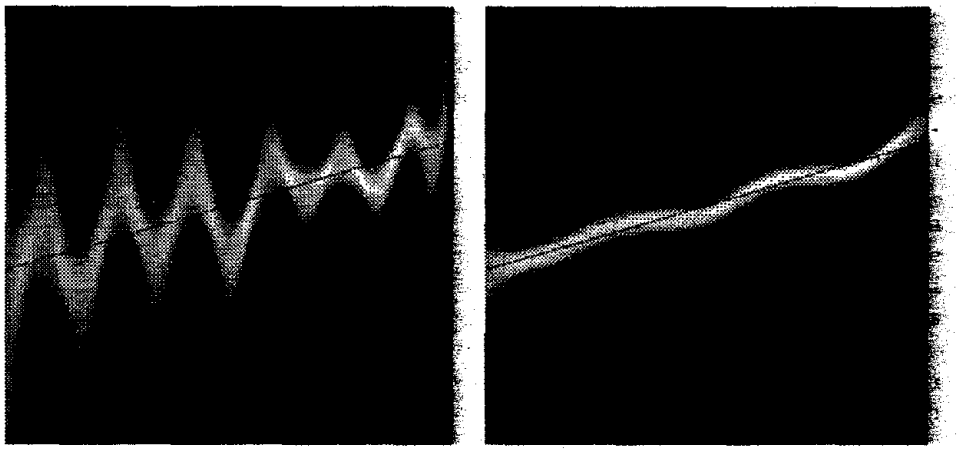

Figure 5. The posterior distributions for the alpha curves for two values of the hyperparameter that controls the strength of the smoothness prior: $\lambda = 0.0004$ on the left and $\lambda = 0.4$ on the right. The true alpha curve is shown as a straight line.

of the model may be estimated with respect to any type of uncertainty measure desired, for example, in terms of variance. 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_0 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).

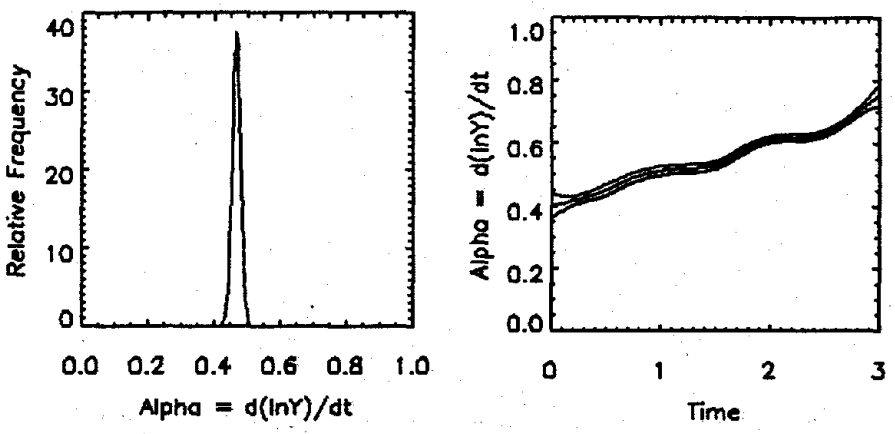

Figure 6. On the left, a vertical slice through Fig. 5 ($\lambda = 0.4$) at the time 1.5, which gives the posterior distribution for alpha at that time. Analysis of Fig. 5 (right) yields the posterior mean and one standard deviation uncertainty band for alpha.

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. The one on the left is obtained with a minimal prior. It demonstrates the tendency of the splines to oscillate. The peaks and troughs of the envelope occur at the spline knots. The image on the right is for $\lambda = 0.4$, which is approximately the value favored by maximizing the evidence in Bayes law, Eq. (6). The severe oscillations seen on the left are well controlled by the smoothness prior. The remaining wiggles are caused by fluctuations away from the true linear alpha curve in the particular set data we are analyzing.

The pdf for λ can be obtained by integrating the joint distribution for λ and $\mathbf{a}$,

$$p(\mathbf{a}, \lambda | \mathbf{d}) = \frac{p(\mathbf{d}|\mathbf{a}, \lambda)\, p(\mathbf{a}, \lambda)}{p(\mathbf{d})} \ , \tag{11}$$

over $\mathbf{a}$:

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

For a fairly flat prior on λ and $\mathbf{a}$, the most probable λ occurs when the evidence $p(\mathbf{d}|\mathbf{a}, \lambda)$ is largest [4].

This approach is followed here. However, for this to work, it is necessary for the $\alpha(t)$ curve to possess some structure that doesn't minimize the prior, For this calculation we used a step function for $\alpha(t)$. The evidence exhibits a broad maximum around $\lambda = 1$.

As argued above, each column of Fig. 5 represents the posterior distribution for alpha at a given time, as seen in Fig. 6 (left). One can thus determine the posterior mean alpha and the rms deviation of the posterior as a function of time,

as shown in Fig. 6 (right). The plot on the right shows the posterior mean and the mean plus or minus one standard deviation. These curves represent an uncertainty envelope for the alpha curves. However, it has to be realized that they really only refer to the pdf at any particular time without regard to any other time. Details contained in the posterior distirbution have been marginalized out. Specifically, the correlations in uncertainties from one time to another can not be inferred from this envelope. Rather, to get an idea of these correlations, one has to go back to the MCMC samples and either visualize the correlations, as in Fig. 4. One can quantify the correlations, for example, by computing the cross correlation over the MCMC sequence between two different times, to obtain an estimate of the covariance between the two uncertainties.

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, Shane Reese, Greg Cunningham, Paul Coggans, and Jack Jacobson.

References

1. G. A. Linenberger, J. D. Orndoff, and H. C. Paxton, "Enriched-uranium hydride critical assemblies," *Nucl. Sci. Eng.*, **7**, pp. 44–57, 1960.
2. 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.
3. C. deBoer, *A practical guide to splines*, vol. 27 of *Applied Mathematical Sciences*, Springer, New York, 1978.
4. D. J. C. MacKay, "Bayesian interpolation," *Neural Comput.*, **4**, pp. 432–447, 1992.
5. C. H. Reinsch, "Smoothing by spline functions," *Numer. Math*, **10**, pp. 237–248, 1967.
6. W. R. Gilks, S. Richardson, and D. J. Spiegelhalter, *Markov Chain Monte Carlo in Practice*, Chapman and Hall, London, 1996.
7. A. Gelman, J. B. Carlin, H. S. Stern, and D. B. Rubin, *Bayesian Data Analysis*, Chapman & Hall, London, 1995.
8. J. Besag, P. Green, D. Higdon, and K. Mengersen, "Bayesian computation and stochastic systems," *Stat. Sci.*, **10**, pp. 3–66, 1995.
9. 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.
10. 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.