Preprint, Astrophysical Journal, accepted June 10, 1996

BAYESIAN PERIODIC SIGNAL DETECTION: ANALYSIS OF ROSAT OBSERVATIONS OF PSR 0540-693

P. C. Gregory
Department of Physics, University of British Columbia, Vancouver, British Columbia, V6T 1Z1

T.J. Loredo
Department of Astronomy, Space Sciences Building, Cornell University, Ithaca, New York, 14853

ABSTRACT

We briefly summarize the main features of our Bayesian method for detecting and measuring periodic signals in event arrival time data, and compare it to the more common FFT and epoch folding (EF) methods in a case study using ROSAT observations of the X-ray pulsar PSR 0540-693. Although originally detected using an FFT analysis of Einstein Observatory data (Seward et al.,1984), the same approach failed to detect the pulsar signal in ROSAT Observatory data. We show that both EF and our Bayesian method can reliably detect the pulsar in the ROSAT data, and we demonstrate several important advantages of the Bayesian method. These include the ease of interpretation of the results, and more accurate period determinations. Further, the Bayesian approach accounts for our prior uncertainty in the signal period and phase in a manner that does not require specification of the number and location of "independent" frequencies or phases, an aspect of more traditional approaches that lends a troubling subjectivity to conclusions based on them. Finally we present the shape of the pulsar lightcurve derived from the ROSAT data by using our Bayesian method.

Subject headings: pulsars, Bayesian methods, period detection, FFT, epoch folding, Gregory-Loredo method, x-ray astronomy

1. INTRODUCTION

In 1979 X-ray emission was detected (Long and Helfand 1979) from a previously known radio supernova remnant, SNR 0540-693 (le Marne 1968; Mathewson and Clarke 1973), located in the Large Magellanic Cloud. X-ray spectral observations (Clarke et al. 1982) showed that the X-ray emission was nonthermal and hence probably from a pulsar-driven synchrotron nebula. Dramatic confirmation of this hypothesis was provided by the discovery of a 50 ms X-ray pulsar at the center of the nebula in 1984 (Seward et al. 1984). The X-ray observations showed the pulsar was spinning down rapidly, with a characteristic age, $P/(2\,\dot{P})$, of only 1660 yr. The pulsar was subsequently detected at optical wavelengths (Middleditch and Pennypacker 1985) and at radio wavelengths (Manchester et al. 1993). The evidence at radio, optical and X-ray wavelengths show that SNR 0540-693 is a very close analog of the Crab Nebula.

Although the X-ray pulsations were originally detected using an FFT analysis of Einstein Observatory data (Seward et al. 1984), the same approach failed to detect the pulsar signal in ROSAT Observatory data. Further details of the Einstein detection and ROSAT observations are provided in section 4.

In many fields the FFT (Fast Fourier Transform) is the standard tool for for the detection of a periodic signal. It was introduced largely for intuitive reasons, but Jaynes has recently provided a formal justification for using the FFT by applying the principles of Bayesian inference to the following (Jaynes 1987). Suppose we are analyzing data consisting of samples of a continuous function contaminated with additive noise, and that we assign Gaussian probability distributions with zero means to the noise components. Jaynes showed that, presuming the possible periodic signal is sinusoidal (but with unknown amplitude, frequency, and phase), the Schuster periodogram exhausts all the information in the data relevant to assessing the possibility that a signal is present, and to estimating the frequency and amplitude of such a signal. If the data are evenly spaced, the Schuster periodogram is determined by the FFT, thus justifying the use of the FFT for such detection and measurement problems.

A corollary of Jaynes's analysis is that for any other problem (e.g., nonsinusoidal lightcurve, or nongaussian probabilities) use of the FFT is not optimal; more information can be extracted from the data if we use more sophisticated statistics. Jaynes made this point himself, and it has been amply demonstrated in the work of Bretthorst (1988), who has applied similar methods to signal detection and estimation problems with nonsinusoidal models with Gaussian noise probabilities (e.g., models with multiple sinusoids, or with "chirp" or exponentially decaying amplitude).

To illustrate this point in an intuitive fashion, consider application of FFT methods to data produced in the following two situations:

a) Lightcurve consisting of a single sinusoid plus Gaussian white noise.

b) Lightcurve consisting of a narrow pulse with the same signal energy as in (a) and same Gaussian white noise.

In (b) the same signal energy is spread out over many harmonics. If the signal to noise in (a) is just sufficient for a reliable detection of a peak above the noise at the correct frequency then case (b) is expected to result in a non detection. Intuitively we would expect a method which did not disperse the pulse signal energy would be more optimal for the detection problem. A well known "frequentist" statistical method that overcomes this particular problem for Poisson data is "epoch folding", which makes use of the χ^2 statistic (Leahy et al. 1983), (Collura et al. 1987). However, because of it's great speed advantage the FFT continues to be widely used in situations in which it is not optimal. In some cases the price paid for speed is very large as we shall illustrate below.

A great strength of Bayesian analysis is that it enables us to find optimal solutions to well posed statistical problems. We became interested in finding a method which would be optimal for detecting a periodic signal in data consisting of the arrival times of events, when we had no specific prior information about the shape of the signal. Given some data, our goal was to develop Bayesian answers to the questions, "Is there evidence for a periodic signal? If so, what are the best estimates of its period and shape?" Our initial motivation was the detection of X-ray pulsars in data consisting of a set of arrival times for individual X-ray photons, some of which are background events. The lightcurves of X-ray pulsars exhibit a very wide range of shapes, and arrival time data is best modeled with the Poisson distribution rather than with the Gaussian distribution. These properties of the problem indicated to us that a Bayesian analysis would lead to valuable new procedures for analyzing such data.

In secion 2 we will summarize briefly the main features of the method we developed to address these questions. We have also generalized the method to order to treat the problem of detecting nonperiodic variability. In section 3 we compare our method to the epoch folding method. Finally in section 4 we apply our method to arrival time data for the X-ray pulsar PSR 0540-693.

2. BRIEF SUMMARY OF THE GREGORY-LOREDO METHOD

Full details of the theoretical development of the method are given in our earlier papers (Gregory and Loredo 1992a, b; henceforth referred to as GL1 and GL2). In this section we present a nonmathematical summary of the method. The problems we are addressing are the detection and measurement of a periodic signal in an arrival time series when we have no specific prior knowledge of the existence of such a signal or of its characteristics, including its shape. The first question is a detection problem, "Is there evidence for a periodic signal regardless of the exact shape, period and phase of the signal?"

To address the detection problem, we used Bayes's theorem to compare a constant rate model for the signal to models with periodic structure. The periodic models describe the signal plus background rate as a stepwise distribution in m bins per period, for various values of m. Clearly, such a model is capable of approximating a lightcurve of arbitrary shape. For each choice of m the model contains $m + 2$ parameters: the trial period and phase, the average rate, plus $m - 1$ parameters describing the normalized shape of the lightcurve (for the normalized shape, specifying the fraction in any $m - 1$ bins fixes the fraction in the remaining bin). The constant model has only one parameter, the average rate; formally, it is equivalent to the $m = 1$ case of the periodic model. For each choice of m, we compute the odds (short for odds ratio), $O_{per}(m)$ which is the ratio of the probability for the m bin periodic model to the probability of the constant model. For the detection problem we need to marginalize out all the parameters, including treating m as a parameter as well, leading to a global odds, O_{per}. Marginalization involves integrating the joint probability for all parameters of a model over the prior range in each parameter. A remarkable feature of the stepwise model is that it enables marginalization of the m shape parameters to be performed analytically, leaving only the period and phase to be integrated numerically.

The Bayesian posterior probability for a periodic model contains a term which quantifies Ockham's razor, penalizing successively more complicated periodic models for their greater complexity even though they are assigned equal prior probabilities. The calculation thus balances model simplicity with goodness-of-fit, allowing us to determine both whether there is evidence for a periodic signal and the optimum number of bins for describing the structure in the data. Unlike traditional "frequentist" calculations which seek to draw conclusions by *maximizing* a statistic measuring discrepancy from a constant signal, the Bayesian calculations *average* the relevant statistics over uncertain parameters needed to specify the signal; the Ockham's razor terms come from this averaging procedure. An important consequence of using averages rather than maxima is that the outcome of the Bayesian calculation does not depend on the number of periods examined in a search for periodicity, but only on the range examined. Also, prior information about the range of periodicity (e.g., from earlier measurements) is easily and objectively handled in the Bayesian formalism.

Once a signal is detected, we also use Bayes's theorem to address measurement questions, such as estimating the frequency and shape of the signal.

In GL3 (Gregory and Loredo 1993) we extended the method to arbitrarily binned data. We showed that the same method is directly applicable to binned data even if the binning is irregular in size. Secondly, we demonstrated the use of the same method to reliably detect nonperiodic variability as first proposed in Appendix C of GL2. The stepwise periodic models referred to above transform to nonperiodic models when the parameter representing the trial period is set equal to the data duration. In this case a signal means any significant variation in the Poisson time series compared to a constant rate model. Again we assume no specific knowledge about the shape of this variation.

3. COMPARISON WITH EPOCH FOLDING METHOD

In some respects the Gregory-Loredo (GL) method is similar to the popular epoch folding (EF) method. Both involve dividing the trial period into m bins and binning the data modulo the trial period for a given trial phase. In EF the χ^2 statistic is used to decide at some significance level, whether a constant model can be rejected and, thus indirectly infer the presence of a periodic signal. The significance associated with rejection of the constant model depends on the number of independent trial periods, $N_{(per)}$, searched. Unfortunately the value of $N_{(per)}$ can vary by a factor of 2-3 depending on what underlying "true" lightcurve is assumed. It can also depend on the locations of the periods examined, which should be specified a priori (i.e., one must not locate the maximum and then choose periods spaced so that one of them falls precisely on the maximum).

An additional problem concerns the choice of signal phase, ϕ (defined so that the bin which an event at time t falls in is determined by $(\omega t + \phi)$, with $\omega = 2\pi/P$ for period P). Typically, EF analysts have ignored the signal phase, implicitly setting $\phi = 0$. In our own simulations, we have found that χ^2 can easily vary by as much as a factor of 2 as ϕ varies over $[0, 2\pi/m]$. Since the significance is roughly an exponential function of χ^2, accounting for possible values of ϕ is crucial. Rather than examining a grid of ϕ values and facing the resulting problem of determining the number of "independent" ϕ values to search, Collura et al. (1987) average χ^2 over phase, and use simulations to show that, at least asymptotically, the phase-averaged χ^2, which we denote by $\langle\chi^2\rangle_\phi$, still follows the χ^2 distribution. In the Bayesian treatment, signal phase is an uninteresting *nuisance parameter* which is removed rigorously by marginalization, that is, by integrating the joint distribution for ϕ and other parameters over ϕ. Since the joint distribution is roughly proportional to $e^{-\chi^2/2}$, and not to χ^2 itself, the result of marginalization can be quite different from averaging χ^2.

Returning now to the GL method, one useful quantity is the frequency (or trial period) dependent odds, $O_{per}(f, m)$, specifying how the data favor a periodic model with an a priori specified frequency over the constant model. It turns out that this is a maximum for the choice of period which gives rise to a distribution of the events in the m bins which is least likely to occur by chance. Formally the probability for any particular choice of signal parameters is inversely proportional to the multiplicity of the distribution of the resulting distribution of events in bins. If the number of events is large then the multiplicity is simply related to the configurational entropy of the distribution. Thus $O_{per}(f, m)$ is a maximum for a trial period leading to a distribution with the lowest entropy (most organized).

In the GL method the reciprocal of the multiplicity is the sufficient statistic that parallels the χ^2 statistic in EF. The two methods were compared using simulated data for a low count rate pulsar in GL2. While the GL method resulted in a global odds, O_{per} of 13:1 in favor of the presence of a periodic signal, the EF method failed to achieve a significant detection.

Finally the GL method automatically includes a term which quantifies Ockham's razor, allowing us to determine the optimum number of bins for describing the structure in the lightcurve *from the data.* In EF, a particular value of m must be specified a priori, with no reference to the evidence for structure in the data.

The above comparison has dealt mainly with the signal detection problem. Once a periodic signal is detected we want to estimate the frequency and shape of the periodic signal together with their credible regions. (A *credible* region, defined by integrating a probability distribution over the parameter space, is the Bayesian counterpart to the familiar *confidence* regions of frequentist statistics, that are defined in terms of integrals over the sample space; the two are identical for simple Gaussian estimation problems, but not in general.) These parameter estimation problems can be rigorously treated by applying Bayes's theorem as shown in GL2. However, we are not aware of any astronomical work that uses the EF method to find rigorous estimates and confidence regions for the period of a detected signal. Most studies seem to consider the "half power" width of the χ^2 peak to give a rough estimate of the uncertainty in the frequency estimate provided by the χ^2 minimum. In comparison, the uncertainty associated with the Bayesian period estimate can be orders of magnitude smaller than that found from the "half power" width of the χ^2 peak, as we will demonstrate in section 4.2.2.

4. APPLICATION TO X-RAY PULSAR DATA

In this section we present some background about the 50 ms pulsar in the Large Magellanic Cloud (LMC) regarding its initial detection and explore what role the GL method might play.

In 1979 X-ray emission was detected (Long and Helfand 1979) from a previously known radio supernova remnant, SNR 0540-693 (le Marne 1968; Mathewson and Clarke 1973), located in the Large Magellanic Cloud. X-ray spectral observations (Clarke et al. 1982) showed that the X-ray emission was nonthermal and hence probably from a pulsar-driven synchrotron nebula. Dramatic confirmation of this hypothesis was provided by the discovery of a 50 ms X-ray pulsar at the center of the nebula in 1984 (Seward et al. 1984). The X-ray observations showed the pulsar was spinning down rapidly, with a characteristic age, $\mathrm{P}/(2\,\dot{\mathrm{P}})$, of only 1660 yr. The pulsar was subsequently detected at optical wavelengths (Middleditch and Pennypacker 1985) and at radio wavelengths (Manchester et al. 1993). The evidence at radio, optical and X-ray wavelengths show that SNR 0540-693 is a very close analog of the Crab Nebula.

The X-ray pulsations were initially detected from an FFT analysis of the data obtained from the Einstein Observatory. Data was obtained in seven intervals in the period 1979 April 1 to 1980 December 17. The initial FFT search was carried out with 2500 photons spanning a duration of approximately 10^4 s from a data set commencing 1979 September 8. By good fortune the period search range extended down to a low period limit of 49 ms (F. Seward, private communication 1993). The true pulsar signal (50 ms) turned out to be the second highest peak in the initial FFT. Confidence in the reality of the signal was established from FFT runs on other data sets. Epoch folding was then used to obtain the lightcurve and a best period (epoch November 9.9174) of $P = 0.050208679$ s ± 3 ns from observations during 1979 November 9-11, containing 21,000 photons. The period uncertainty was estimated from the "half power" width of the χ^2 peak. In addition a period derivative of $\dot{P} = 4.79 \times 10^{-13}$ s s$^{-1} \pm 0.5\%$ was determined from two observations $\sim$ 600 days apart. Later radio observations (Manchester and Peterson 1989) also provided a measurement of the braking index.

The pulsar was recently re-observed with the ROSAT Observatory by F. Seward and colleagues, but this time an FFT search failed to detect the pulsar. We decided to try out our method on a sample of ROSAT data, using as prior information the results from the earlier detection and to compare with epoch folding (EF) results. A sample data set of 3305 photons was kindly provided by F. Seward. The data spanned an interval of 116,341 s and contained many gaps.

4.1. Epoch Folding Results

We first show the results of an EF analysis of the data over a trial frequency range of 6×10^{-4} Hz. This range was determined by the earlier measurements of the period and period derivatives and their uncertainties. For both methods the binning of the individual photons was accomplished with an algorithm that incorporated the known period derivative and second derivative (Gouiffes et al. 1992). For this short data set the second derivative was found to be unimportant. We also allowed for the effects of data gaps by correcting for differences in the effective integration times in the different bins.

To account for the unknown phase, we averaged χ^2 over phase for each examined period, following Collura et al. (1987). Phase averaging involves integrating χ^2 over a phase range corresponding to one bin. This integration was calculated exactly as follows. The value of χ^2 is a stepwise function of ϕ; its value changes only when ϕ changes enough to move an event across a bin boundary. Thus, starting at $\phi = 0$, as we bin the individual events to compute χ^2, we note the distance in ϕ required to move each event to the next bin boundary from which a set of phase increments is computed. The desired integral is then the sum of the products of the constant χ^2 within a given increment times the width of the increment. The resulting plot of $\langle\chi^2\rangle_\phi$ versus frequency for $m = 5$ bins is shown in Figure 1. Out of curiosity we computed χ^2 versus frequency for a single phase. The result was qualitatively similar but the resulting curve (not shown) was much noisier in appearance.

The significance associated with rejection of the constant model depends on the number of independent trial periods, N_{per}, searched. Rigorously, both N_{per} and the location of the frequencies must be determined a priori. The regularly spaced vertical bars correspond to our a priori choice of independent frequencies. Thus for the given period search range $N_{per} = 144$ and the maximum $\langle\chi^2\rangle_\phi$ of 77 corresponds to a χ^2 significance of 1.0×10^{-13}. For this particular data set and period search range, the EF method decisively rejects the constant signal hypothesis, indirectly supporting the hypothesis that a periodic signal is present.

The "half power" width of the χ^2 peak for this data corresponds to a frequency uncertainty of $\pm 3.6 \times 10^{-6}$ Hz.

4.2. Gregory-Loredo Method Results

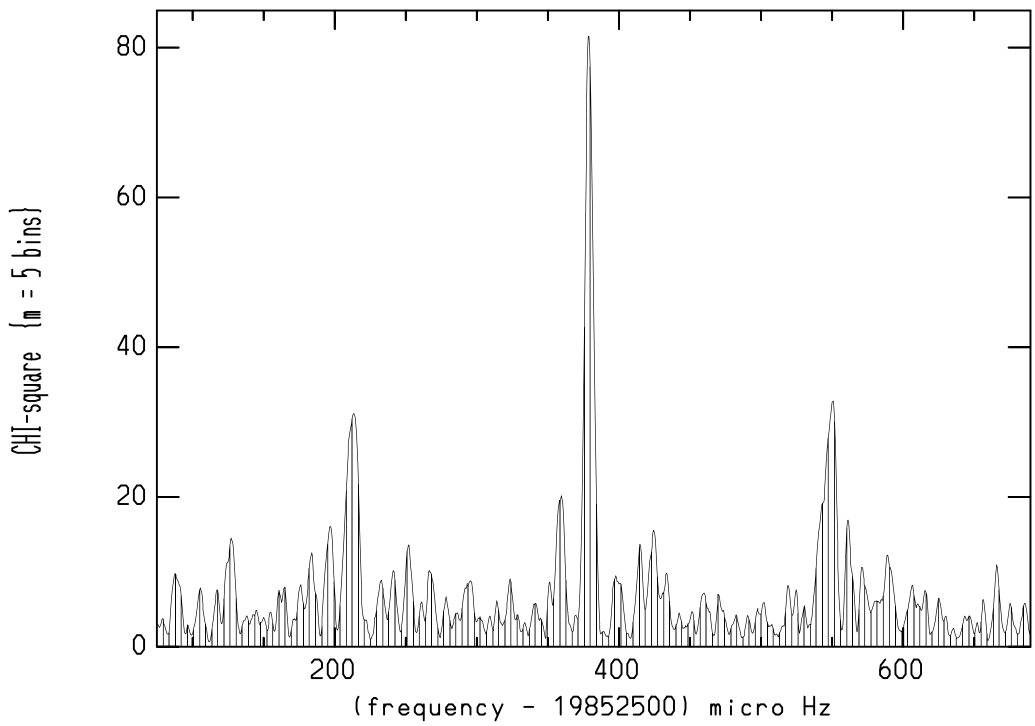

Fig. 1.— The $\langle\chi^2\rangle_\phi$ statistic versus trial frequency from epoch folding analysis using $m = 5$ bins. The vertical bars indicate the locations of independent frequencies.

4.2.1. Detection

The GL analysis consists of computing the odds in favor of the periodic class of models over the constant model, using the same period range. This requires marginalizing the joint distribution for the model parameters and the data over all of the periodic model parameters. As already noted, the required integrations over the shape parameters can be done analytically; the remaining integrals over the frequency (or period) and phase, and the sum over the number of bins, m, must be done numerically. The integrals extend over the prior range for each parameter. The phase integration can be carried out exactly, using the same procedure as outlined for the phase averaging of χ^2 above, as described in Appendix A of GL2.

We performed these calculations with the ROSAT data as follows. After first averaging over phase (after the analytic integration over the shape parameters) we obtained a useful intermediate function, the frequency dependent odds, $O_{per}(f, m)$. This is the odds in favor of the *particular* periodic model that has m bins and frequency f. That is, it is the odds in favor of such a model if we knew a priori that any possible periodic signal would have frequency f and be described by a stepwise function with m bins. Figure 2 is a plot of the log_{10} of $O_{per}(f, m)$ as a function of f for the ROSAT data, for $m = 5$ bins. Comparison of Figures 1 and 2 indicates that the arrangements of peaks and their relative heights are very similar but the peak structure differs in detail; we will say more about this structure presently. It is also apparent that for most of the frequency range the Ockham penalty arising from integration over the unknown shape and phase parameters results in a $O_{per}(f, m) \ll 1$. Three peaks exceed an odds of 1 ($log_{10} = 0$). The central one is $\approx 10^{11}$ larger than the next largest.

Since we dont know the exact frequency a priori we marginalize (average) over the prior frequency range to obtain $O_{per}(m)$. The frequency marginalization introduces an additional Ockham's penalty. This is apparent in Figure 3, which plots $O_{per}(m)$ versus the number of bins, m. Although the largest value of $O_{per}(f, m)$ in Figure 2 exceeded 10^{14}, the odds in favor of the $m = 5$ model is about 2×10^{11}. The reduction is the price paid for not knowing the frequency a priori.

If we wish to choose a single stepwise model to represent the signal, the GL analysis allows us to determine the optimum number of bins. This is apparent from Figure 3, where $O_{per}(m)$ has an obvious maximum for $m = 3$. This maximum arises from a competition between the evidence in the data arguing for more structure and Ockham's razor, which penalizes successively more complicated periodic models for their greater complexity.

The answer to the detection problem comes from the global odds, O_{per}, which requires a further marginalization of the joint probability for the data and the number of bins over m. The resulting odds of 2.6×10^{11} indicates near certainty in the presence of a periodic signal.

The high value of O_{per} is due in part to our prior knowledge of the frequency of PSR0540-693: our relatively small prior uncertainty in the period resulted in a relatively small Ockham penalty.

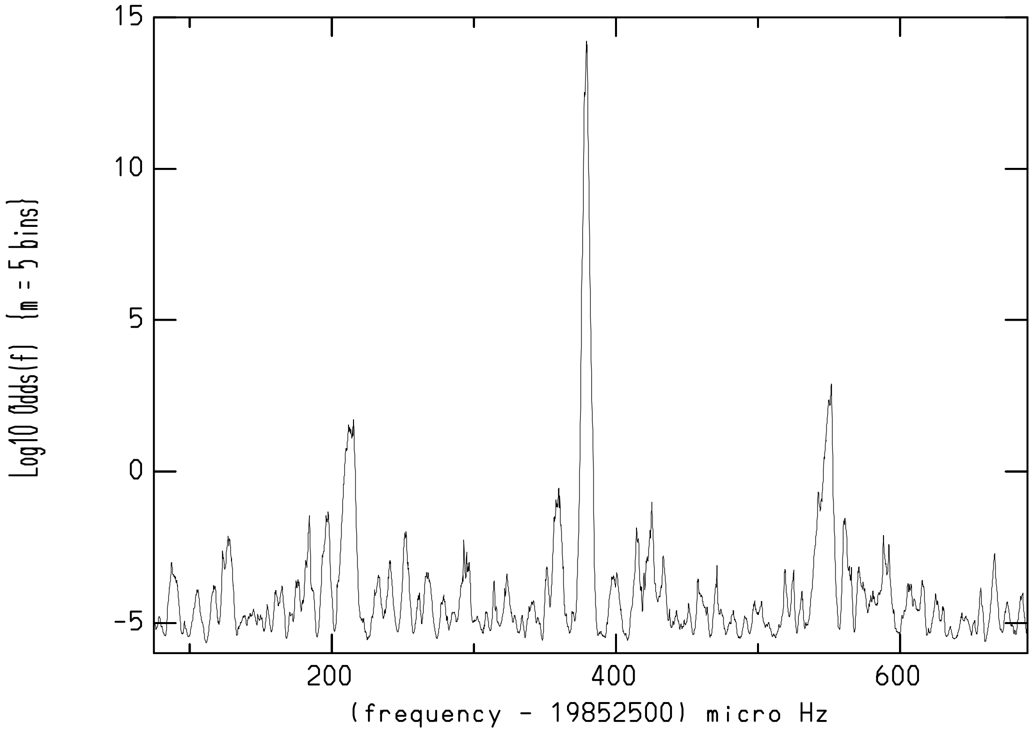

Fig. 2.— Log_{10} of the frequency dependent Odds versus trial frequency for $m = 5$ bins.

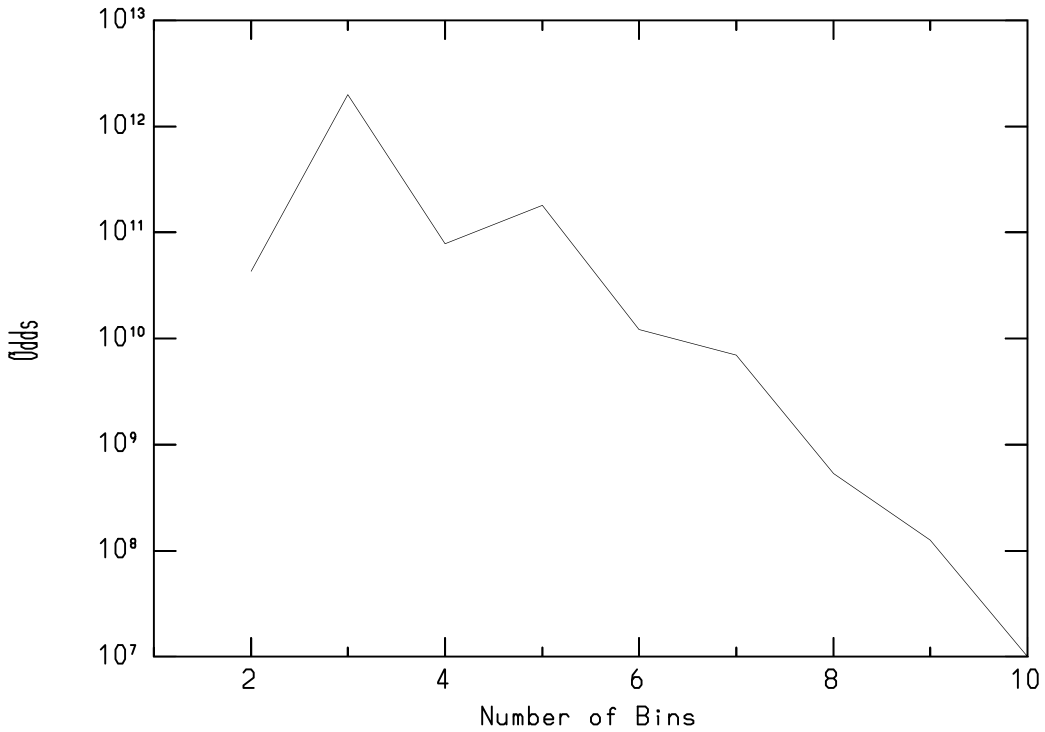

Fig. 3.— The Odds versus m, the number of bins.

However, the periodic signal in these data is so strong that it would have been detected even if no prior knowledge of the pulsar period had been available. To see this, suppose our a priori period search range extended from the rotational breakup period of a neutron star (≈ 1.5 ms), to half the duration of the data. In the Bayesian calculation we would have to marginalize over the larger period range. However, we can find a lower limit on the resulting odds in favor of the periodic model class without doing any further calculations, as follows.

The odds is proportional to the ratio of the probability for the data presuming the periodic models (requiring integrals over all model parameters for those models) to the probability for the data presuming a constant signal (requiring marginalization only over the unknown average rate). We can schematically write the former as

$$p(D \mid M_{per}) = \int df\, p(f, D \mid M_{per}) = \int df\, p(f \mid M_{per})\, p(D \mid f, M_{per}).$$

The first factor in the final equation is just the prior distribution for f, which is proportional to $1/f$ (i.e., it is flat in $log(f)$). For the calculation above, we integrated over a narrow frequency interval $[f_l, f_u]$, so that

$$p(D \mid M_{per}) = \frac{1}{\ln(f_u/f_l)} \int_{f_l}^{f_u} df\, \frac{p(D \mid f, M_{per})}{f}.$$

Had previous measurements of f not been available, we would have to use the larger range, $[f_l', f_u']$, with $f_l' = 2/T$ (with T the duration of the data), and $f_u' \approx 667$ Hz. But assume for the moment that there is no other peak in the larger period range. Then the value of the integral over the larger range will be the same as that over the range already examined; only the normalization factor out front will differ. The resulting odds, O_{per}', will then be related to that just calculated according to

$$O_{per}' \approx O_{per} \times \frac{\ln(f_u/f_l)}{\ln(f_u'/f_l')}.$$

This gives $O_{per}' = 4.5 \times 10^5$. This is greatly reduced compared to O_{per}, due to the much larger Ockham penalty we must pay for not knowing the frequency. But this still provides very strong evidence for the presence of a periodic signal, despite the fact that it was undetected by FFT techniques.

If other significant peaks were present in the larger period range, O_{per}' would be larger; the previous result is a lower limit. Such peaks are only likely to occur if strong signals are present at periods other than that already known for the pulsar

4.2.2. Parameter Estimation

Now that we have established the existence of a periodic signal we are ready to proceed with estimating the frequency and shape of the periodic signal together with their credible regions.

Figure 4 shows a plot of the posterior probability for the frequency of the periodic signal showing a sharp peak centered on 19.85287838 Hz ($P = 50.3705297$ ms) with 68.3% of the probability distribution lying in the range $+0.8$ and -0.2×10^{-6} Hz of the peak. For comparison the $\langle\chi^2\rangle_\phi$ peak is also shown (diamonds). The GL results indicate that the χ^2 "half power" width greatly overestimates the uncertainty of the period.

A feature of the GL analysis is that all parameter estimates improve with increasing evidence, in this case more photons. In contrast, the "half power" width of the χ^2 peak depends mainly on the duration of the data and is not sensitive to the number of photons provided there are sufficient number to produce a well defined peak. To illustrate these two points we created three x-ray data sets spanning the same elapsed time. The first was the original data set of photon arrival times, the second consisted of every second arrival time and the third consisted of every forth arrival time. Figure 5 shows the marginal posterior probability for the frequency of the periodic signal for the three data sets. The posterior probability density function becomes progressively narrower as the number of photons increases.

In contrast, the "half power" width of the χ^2 peak, which is sometimes used as a rough estimate of frequency resolution, is mainly determined by the duration of the data and to a lesser extent on the number of events detected. Figure 6 shows the χ^2 peak corresponding to the same three data sets. The half power width does not exhibit a significant narrowing.

We can understand this and other differences between the GL and EF results as follows. In GL2 we showed that the Bayesian joint posterior for the frequency and phase, $p(f, \phi)$ of the periodic signal could be roughly approximated by

$$p(f, \phi) \propto e^{\chi^2(f,\phi)/2}.$$

Thus a change in χ^2 that is small relative to the peak value can strongly affect parameter estimates, since the relevant probabilities are approximated by the *exponential* of χ^2.

This approximate connection between GL and EF indicates that if we are interested in the frequency alone, marginalizing the phase corresponds to averaging the exponential of $\chi^2/2$ over phase. Figure 7 shows a closeup of the largest frequency peak, comparing the marginal probability densities for f computed exactly and with the approximation obtained by averaging $e^{\chi^2(f,\phi)/2}$ over phase. Both calculations assume $m = 3$ bins. The resulting distributions are quite similar. Formally, to get the approximate expression, we must assume that that the binned data are nearly uniform, as is the case for frequencies that do not correspond to that of the true signal. Nevertheless, this calculation indicates that the approximation may be adequate even for frequencies where the signal is strong. However, little is gained by the approximation, since calculation of the exact expression is a negligible fraction of the computational burdern of the GL method. The main computation burden is binning the events, just as it is in the EF method.

The importance of accounting for uncertainty in the phase is illustrated in Figure 8, which shows the marginal posterior probability density for bin phase for $m = 3$ bins. If only a single bin

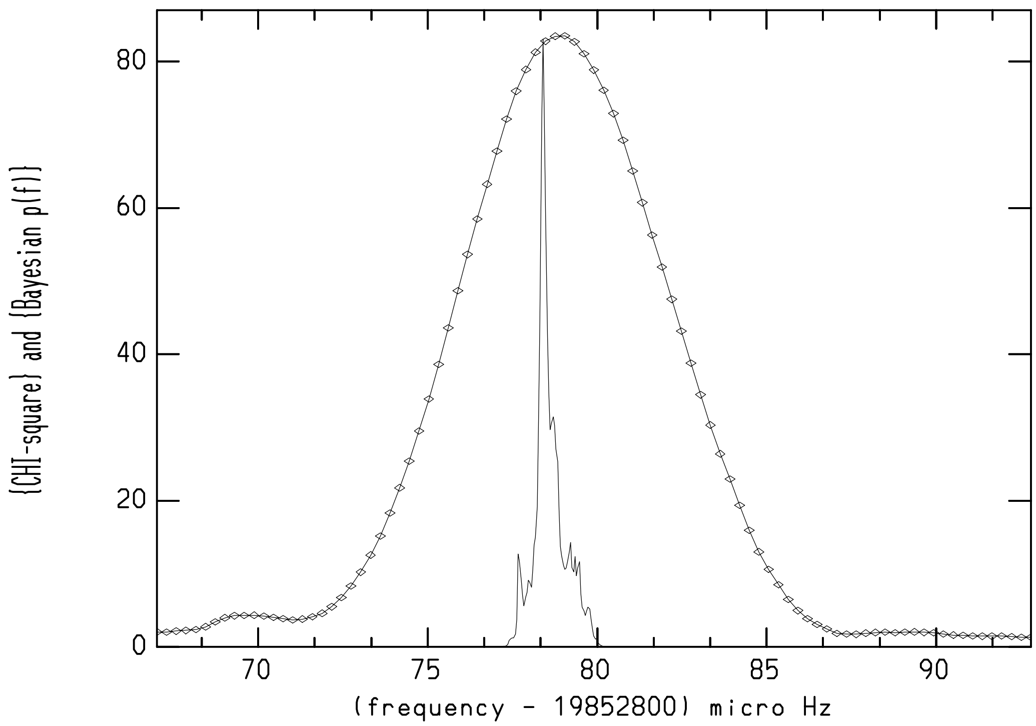

Fig. 4.— Close up of largest frequency peak comparing the GL marginal probability density for f to the Epoch Folding $\langle\chi^2\rangle_\phi$ statistic (diamonds). The $\langle\chi^2\rangle_\phi$ statistic versus trial frequency results from epoch folding analysis using $m = 5$ bins.

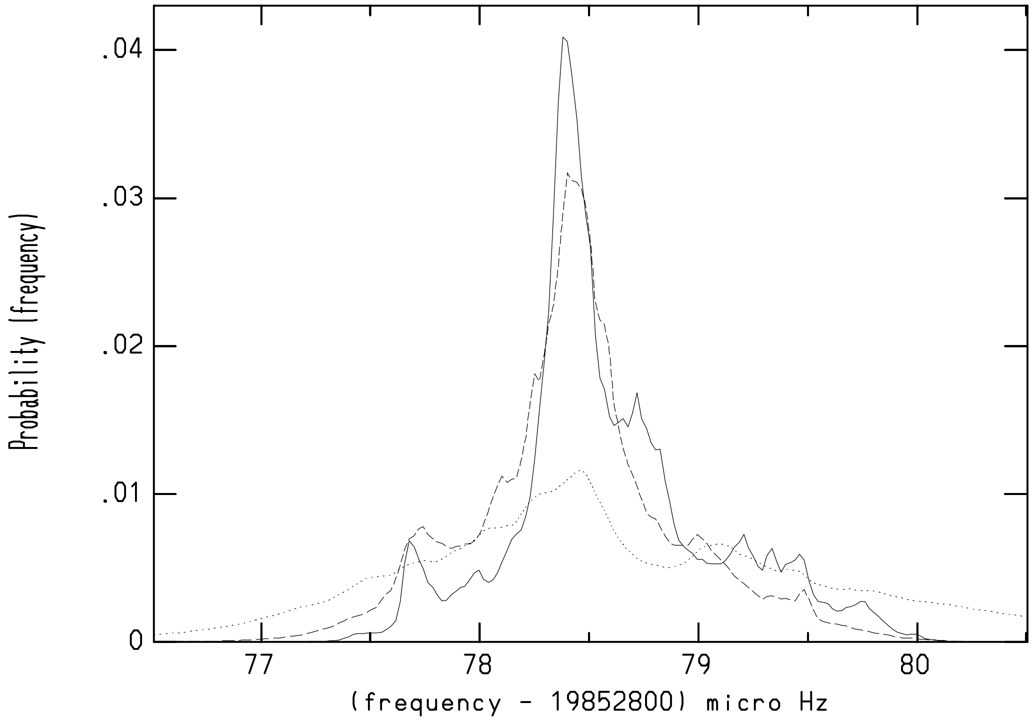

Fig. 5.— Dependence of the marginal probability density for frequency on the number of events detected within a fixed time interval. The solid, dashed and dotted curves correspond to 3304, 1652, and 826 events respectively.

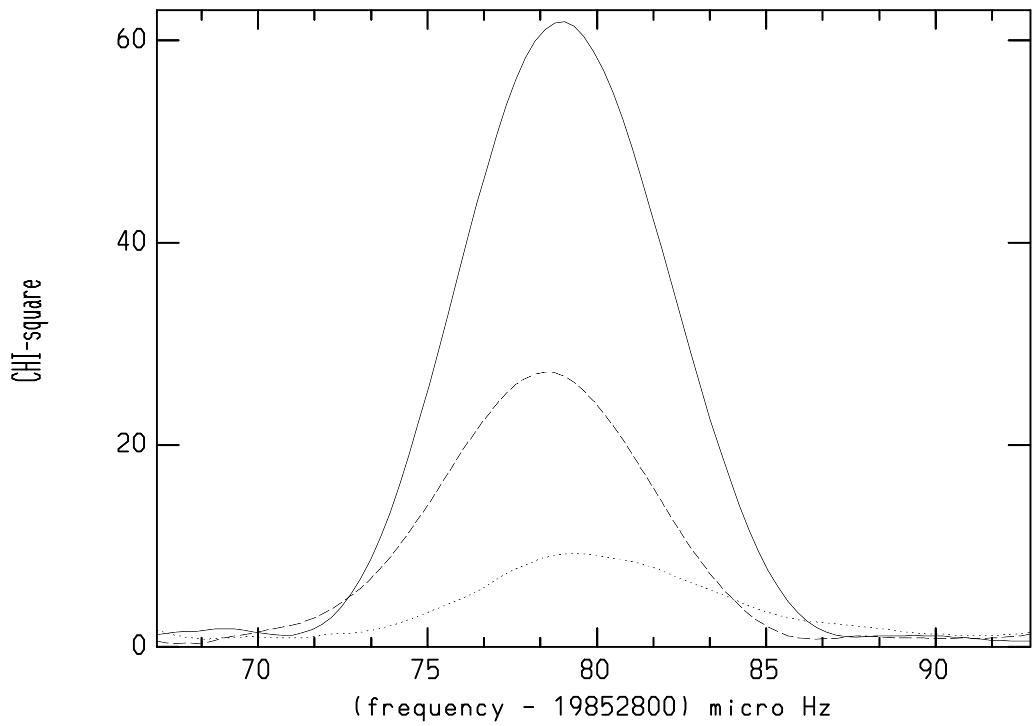

Fig. 6.— Dependence of $\langle\chi^2\rangle_\phi$ frequency behavior on the number of events detected within a fixed time interval. The solid, dashed and dotted curves correspond to 3304, 1652, and 826 events respectively.

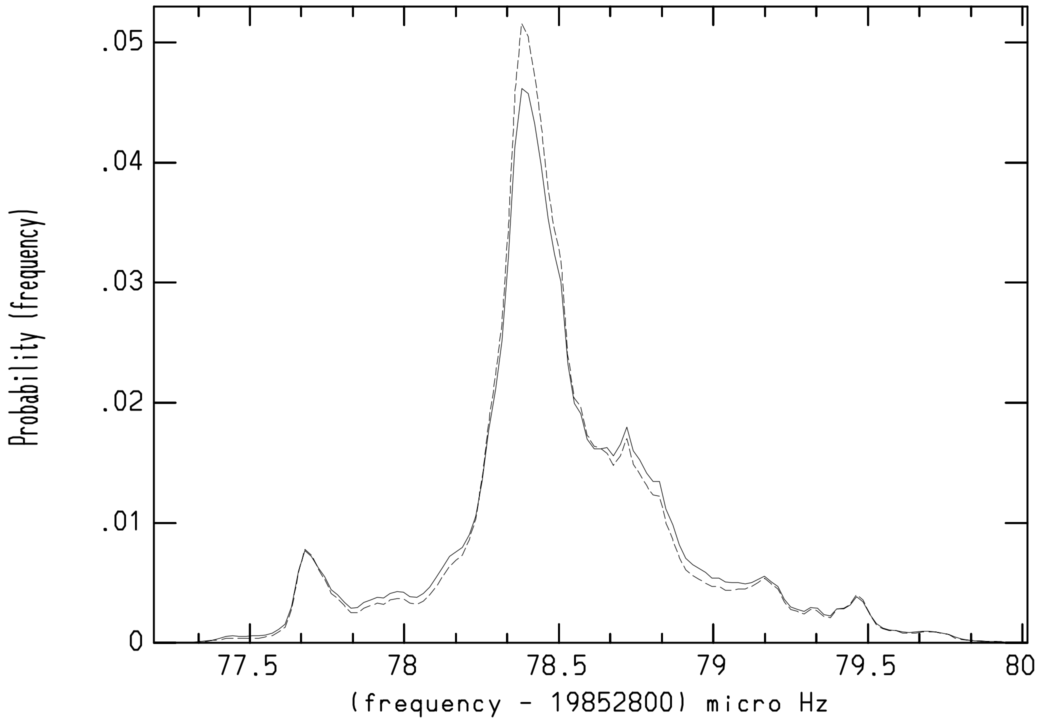

Fig. 7.— Close up of largest frequency peak comparing the marginal probability density for f computed exactly (solid) and using the $e^{\chi^2/2}$ approximation (dashed), for $m = 3$ bins.

phase were considered, the resulting odds could vary by a factor of $> 10^6$ depending on the value selected.

The remaining parameter estimation problem is the shape of the pulsar lightcurve. In EF this is accomplished by plotting the binned data for some choice of the number of bins and using the best estimate for the frequency, resulting in a stewise lightcurve. The number of bins is usually chosen using an informal criterion for significant bin-to-bin variation, based on the standard deviation of the counts in each bin (equal to the square root of the number of events in the bin). A smoother version can be obtained by computing a set of points for a number of different choices of bin phase ranging over one bin, with the drawback that the points are no longer independent.

From the Bayesian point of view, estimates of the lightcurve shape should be based on the posterior distribution for the normalized signal strength at each instant. In GL2 we described a useful summary of this complicated distribution, simply by plotting the mean and its standard deviation as a function of time. The calculation involves a sum over the number of bins, and integrals over frequency and phase, and completely accounts for our uncertainty in these parameters through these integrals. The normalized shape estimate based on the ROSAT data is plotted in Figure 9 as a pair of curves one standard deviation above and below the mean normalized intensity.

The left lightcurve estimate in Figure 9 exhibits a dominant 3 bin shape because of the pronounced maximum in odds for $m = 3$ which is evident in Figure 3. This result highlights the major weakness of the GL method, namely that it does not incorporate any prior information about the expected smoothness of the lightcurve. At present we have not found a computationally efficient way of incorporating such information. One way to achieve a smoother shape estimate is to fix m equal to a large value as shown in the right plot of figure 9, but this results in a larger standard deviation estimate. A better approach would be to use the GL method for detection and period estimation, and then use a more complicated Bayesian method for the shape estimation. Although the additional complication of a method that could incorporate smoothness would likely prohibit its application to the detection problem (because of the large number of frequencies and phases one must examine), once the frequency is known, such methods could be used to provide better shape estimates.

5. DISCUSSION

Analysts who are unfamiliar with Bayesian analysis often want to interpret the odds in terms of a frequentist or classical statistical concept, such as the significance with which one can reject the constant model. But this frequentist quantity is not a direct measure of the probability for the hypothesis that a periodic signal is present. The Bayesian result is based on this probability, so desired by our intuitions. The odds is simply a ratio of this probability to that for the constant

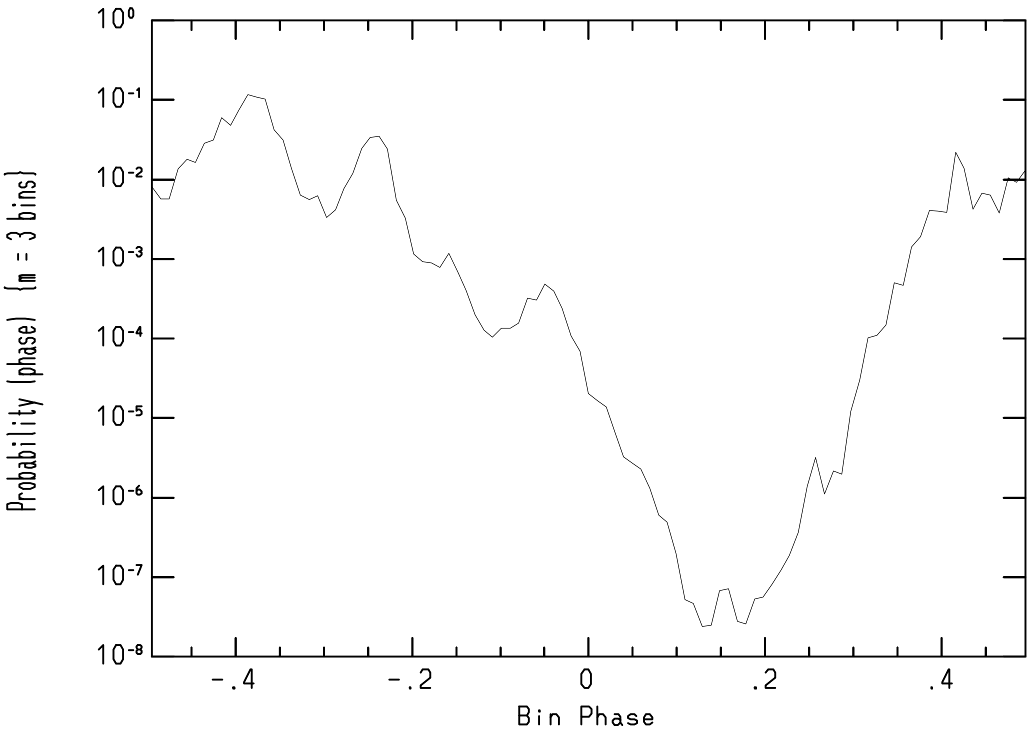

Fig. 8.— Marginal posterior density for bin phase.

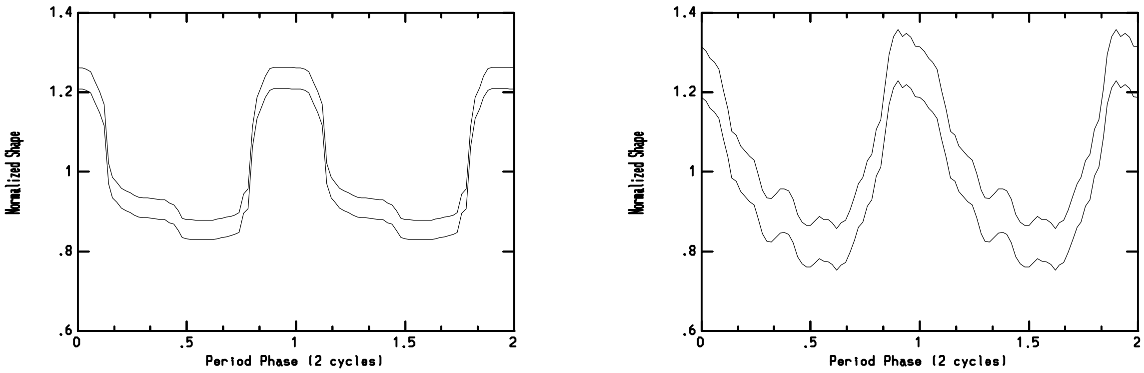

Fig. 9.— Normalized lightcurve shape estimates for the pulsar PSR 0540-69. The two curves in each plot are ± 1 standard deviation estimates. The left plot corresponds to marginalizing over the bin range $m = 2$ to $m = 10$. The right plot is for $m = 12$ bins.

model:

$$O_{per} = p(M_{per} \mid D, I)/p(M_c \mid D, I)$$

where $p(M_{per} \mid D, I)$ is the probability of the periodic class of models given the data (D) and prior information (I), and $p(M_c \mid D, I)$ is the same for the constant model, M_c. If our hypothesis space is complete then,

$$p(M_{per} \mid D, I) + p(M_c \mid D, I) = 1$$

and thus,

$$p(M_{per} \mid D, I) = \frac{O_{per}}{(1 + O_{per})}.$$

Accordingly an $O_{per} = 7.5 \times 10^5$ translates to a probability that a periodic signal is present of 0.9999987.

In EF the χ^2 statistic is used to decide at some confidence level whether a constant model can be rejected and thus indirectly infer the presence of a periodic signal. In this approach the alternative hypothesis of interest (i.e. periodic signal present) is not directly included in the analysis. It is our opinion that the Bayesian approach of directly comparing the probability of two or more competing models is the better way of conducting inference in part because the Bayesian calculation automatically includes a quantitative Ockham' razor (see GL2) penalizing the periodic class of models for their extra complexity.

Two interesting foot notes to this discussion follow. (a) One could argue that our hypothesis space is not complete. For example it is possible that neither hypothesis is true and that, instead, the data might contain a nonperiodicly varying signal. It turns out that we can readily allow for this hypothesis in our Bayesian analysis (see GL3). The nonperiodic models are just a special case of the periodic class when the trial period is equal to the duration of the data. (b) The optimum way to search for a signal whose shape is truely unkown is to include smooth models in the periodic class in addition to the stepwise model so that both spiky and smooth signals can be optimally detected. In principle a smooth parametric model can more sensitively detect such signals because they can model the signal with fewer parameters, and so pay a smaller Ockham penalty. (A forthcoming publication will describe Bayesian analyses with such models.) However, our simulations demonstrate (GL2) that the GL method can usefully detect smooth signals and accurately estimate their period and shape. In this regard we draw attention to an independent comparison (Rots et al. 1994) of the Gregory-Loredo method to 3 other frequently used period searching algorithms (FFT, Scargle Periodogram, and Epoch Folding). In this comparison the GL method provided the best overall performance with respect to accuracy and the ability to find periods and amplitudes in low signal-to-noise data.

It is apparent from the above discussion and GL2 that the GL method has many advantages over EF. What about the computational speeds of the two methods? For the same number of data points and trial periods the two methods are comparable in speed (provided one is accounting for phase uncertainty in the EF method). Because the Bayesian approach permits more information to be extracted from the data (e.g. better weak signal detection, more accurate period and optimum number of bins) this requires in principle examining more trial periods (but only in the vicinity of a peak) and exploring a wider range in the number of bins, which increases the computational burden relative to EF by a constant factor. The need to consider different numbers of bins is the price we pay for allowing the data to determine the complexity of the lightcurve (rather than setting it a priori, as required in frequentist EF analyses). On the other hand the GL method incorporates a quantified Ockham's razor so much less time is spent exploring insignificant noise peaks.

6. SUMMARY

In this paper we have attempted to illustrate in an informal fashion some of the advantages of the Bayesian period signal detection method developed by Gregory and Loredo (1992a, 1992b, 1993) when applied to x-ray pulsar data, by comparing it with two commonly used methods, Epoch Folding (EF) and the FFT in a case study using ROSAT data from observations of PSR 0540-693. For the data set considered the FFT failed to detect a significant signal, which was readily detected by both the GL and EF methods. Both the GL and EF methods involve dividing a trial period into m bins and binning the data modulo the trial period. In EF the χ^2 statistic is used to decide at some confidence level whether a constant model can be rejected and thus indirectly infer the presence of a periodic signal. In EF the value of m must be decided a priori, as must the number and locations of the periods or frequencies searched. Uncertainty in the bin phases is usually not accounted for; when it has been considered, a phase-averaged χ^2 was used.

The GL method employs Bayes's theorem to address both the signal detection problem and the estimation problem of measuring the characteristics of the detected signal. The Bayesian posterior probability for a periodic model contains a term which quantifies Ockham's razor, penalizing successively more complicated models for their greater complexity, allowing us to determine both whether there is evidence for a periodic signal and the optimum number of bins for describing the structure in the data. The marginalization procedure involves integrating functions of the data over frequency, rather than maximizing, and thus avoids all of the problems associated with choosing the number and location of periods to be searched. It also involves a similar integral over the binning phase, which can be approximated by averaging the *exponential* of χ^2, not χ^2 itself, as is done in some applications of EF.

Once a periodic signal is detected the GL method provides rigorous estimates (with uncertainties) of the signal parameters, in contrast to EF. In particular, we showed that the common practice of using the EF "half power" width of the χ^2 to estimate the uncertainty in the measured frequency can give orders of magnitude less precise results than the Bayesian calculation.

This research was supported in part by a grant from the Canadian Natural Sciences and Engineering Research Council at the University of British Columbia, and by a NASA GRO Fellowship (NAG5–1758), NASA grant NAGW–666, and NSF grants AST–87–14475 and AST–89–13112 at Cornell University.

REFERENCES

Bretthorst, G. L. 1988, Bayesian Spectrum Analysis and Parameter Estimation, (Springer-Verlag, New York)

Clark, D. H., Tuohy, I.R., Long, K. S., Szymkowiak, A. E., Dopita, M. A., Mathewson, D. S., and Culhane, J. L. 1982, ApJ, 255,440

Collura, A., Maggio, A., Sciortino, S., Serio, S., Vaiana, G. S., and Rosner, R. 1987, ApJ, 315, 340

Gouiffes, C., Finley, J. P., and Ögelman, H. 1992, ApJ, 394, 581

Gregory, P. C., and Loredo, T. J. 1992a, in Maximum Entropy and Bayesian Methods, ed. C. R. Smith, G. J. Erickson, and P. O. Neudorfer (Kluwer, Dordrecht), 79

Gregory, P. C., and Loredo, T. J. 1992b, ApJ, 398, 146

Gregory, P. C., and Loredo, T. J. 1993, in Maximum Entropy and Bayesian Methods, Paris, 1992, ed. Mohammad-Djafari and G. Demoment (Kluwer, Dordrecht), 225

Jaynes, E. T. 1987, in Maximum Entropy and Bayesian Spectral Analysis and Estimation Problems, ed. C. R. Smith and G. J. Erickson (D. Reidel, Dordrecht-Holland), 1

Leahy, D. A., Darbro, W., Elsner, R. F., Weisskopf, M. C., Sutherland, P. G., Kahn, S., and Grindley, J. E. 1983, ApJ, 266, 160

Le Marne, A. E. 1968, MNRAS, 139, 461

Long, K. S., and Helfand, D. J. 1979, ApJ, 234, L77

Manchester, R. N., Mar, D. P., Lyne, A. G., Kaspi, V. M., and Johnston, S. 1993, ApJ, 403, L29

Manchester, R. N., and Peterson, B. A. 1989, ApJ, 342, L23

Mathewson, D. S., and Clark, J. N. 1973, ApJ, 180, 725

Middleditch, J., and Pennypacker, C. 1985, Nature, 313, 659

Rots, A. H., Lochner, J. C., Schlegel, E. M., and Smale, A. P., HEAD meeting Napa Valley, November 1994

Seward, F. D., Harnden, F. R., and Helfand, D. J. Ap. J. (Letters), 287, L19