# A Bayesian Analysis of Extrasolar Planet Data for HD 208487

P. C. Gregory

*Physics and Astronomy Department, University of British Columbia,
Vancouver, BC V6T 1Z1*

**Abstract.** Precision radial velocity data for HD 208487 has been re-analyzed using a new Bayesian multi-planet Kepler periodogram. The periodgram employs a parallel tempering Markov chain Monte Carlo algorithm with a novel statistical control system. We confirm the previously reported orbit (Tinney et al. 2005) of 130 days. In addition, we conclude there is strong evidence for a second planet with a period of $998^{+57}_{-62}$ days, an eccentricity of $0.19^{+0.05}_{-0.18}$, and an $M\sin i = 0.46^{+0.05}_{-0.13}M_J$.

**Keywords:** Bayesian methods, time series analysis, extrasolar planets, HD 208487

## INTRODUCTION

High-precision radial-velocity measurements of the reflex radial velocity of host stars have lead to the discovery of over 150 planetary candidate companions to solar-type stars both in single and multiple star systems (e.g., Eggenberger et al. 2004). Most of these planets have $M\sin i$ comparable to or greater than Jupiter, where $M$ is the mass of the planet and $i$ is the angle between the orbital plane and a plane perpendicular to the line of sight to the planetary system. However, as Doppler precision increases, lower mass planets are also being discovered and there are now five planets with an $M\sin i$ less than the mass of Neptune (Extrasolar Planets Encyclopaedia, http://vo.obspm.fr/exoplanetes/encyclo/index.php).

Progress is also being made in the development of improved algorithms for the analysis of these data. In Gregory (2005a) we described a Bayesian Markov chain Monte Carlo (MCMC) approach and demonstrated its effectiveness in the analysis of data from HD 73526 (Tinney et al. 2003). This approach has the following advantages:

a) Good initial guesses of the parameter values are not required. A MCMC is capable of efficiently exploring all regions of joint prior parameter space having significant probability. There is thus no need to carry out a separate search for the orbital period(s).

b) The method is also capable of fitting a portion of an orbital period, so search periods longer than the duration of the data are possible.

c) The built-in Occam's razor in the Bayesian analysis can save considerable time in deciding whether a detection is believable. More complicated models contain larger numbers of parameters and thus incur a larger Occam penalty, which is

automatically incorporated in a Bayesian model selection analysis in a quantitative fashion. The analysis yields the relative probability of each of the models explored.

d) The analysis yields the full marginal posterior probability density function (PDF) for each model parameter, not just the maximum *a posterior* (MAP) values and a Gaussian approximation of their uncertainties.

e) The inclusion of a novel statistical control system automates the otherwise time consuming process of selecting an efficient set of Gaussian parameter proposal distribution $\sigma$'s for use in the MCMC algorithm.

In this paper we will describe a significant improvement to this algorithm and demonstrates its use in the re-analysis of data from HD 208487 obtained by Tinney et al. (2005). See Ford (2004) for an interesting application of MCMC methods to the interacting planetary system, GJ 876, in which mutual perturbations are detectable that can be modeled with two precessing Keplerian orbits.

## ALGORITHM IMPROVEMENT

Basically the algorithm was designed as a Bayesian tool for nonlinear model fitting. It is well known that nonlinear models can yield multiple solutions (peaks in probability) in the prior parameter space of interest. For the Kepler model with sparse data the target probability distribution can be very spiky. This is particularly a problem for the orbital period parameter, $P$, which spans the greatest number of decades. In general, the sharpness of the peak depends in part on how many periods fit within the duration of the data. The previous implementation of the parallel tempering algorithm employed a proposal distribution for $P$ which was a Gaussian in the logarithmic of $P$. This resulted in a constant fractional period resolution instead of a fixed absolute resolution, increasing the probability of detecting narrow spectral peaks at smaller values of $P$. However, this proved not to be entirely satisfactory because for the HD 73526 data set (Tinney et al. 2005) one of the three probability peaks (the highest) was not detected in two out of five trials (Gregory 2005a).

The latest implementation employs a proposal distribution in discrete frequency space with a frequency interval $\Delta v$ given by $\Delta v \times$ data duration $=$ resolution (typically resolution $= 0.007$). The resolution corresponds to the fraction of the trial period that the last data point will be shifted by a change in search frequency of $\Delta v$. The width of any probability peak is directly related to this quantity so that a proposal distribution of this form corresponds more closely to achieving a uniform number of samples per peak. Using this search strategy on the HD 73526 data set all three peaks were detected in five out of five trials.

## RE-ANALYSIS OF HD 208487

With the exception of the change mentioned in the previous section, the details of the algorithm are as described in Gregory (2005a). Four different models $(M_{2j}, M_2, M_{1j}, \& M_{1s})$ were compared. $M_{1j} \& M_{2j}$ are one and two planet models which

**TABLE 1.** Parameter prior probability distributions.

| Parameter | Prior | Lower bound | Upper bound |
|---|---|---|---|
| $P_1$ & $P_2$ (d) | Jeffreys | 0.5 | 4422 |
| $K_1$ & $K_2$ (m s$^{-1}$) | Modified Jeffreys* | 0 ($K_a = 1$) | 400 |
| V (m s$^{-1}$) | Uniform improper prior | all reals | |
| $e_1$ & $e_2$ | Uniform | 0 | 1 |
| $\chi_1$ & $\chi_2$ | Uniform | 0 | 1 |
| $\omega_1$ & $\omega_2$ | Uniform | 0 | $2\pi$ |
| $s$ (m s$^{-1}$) | Modified Jeffreys | 0 ($s_a = 5$, typical measurement error) | 100 |

* Since the prior lower limits for $K$ and $s$ include zero, we used a modified Jeffrey's prior of the form

$$p(K|Model,I) = \frac{1}{K + K_a} \frac{1}{\ln\left(\frac{K_a + K_{\max}}{K_a}\right)} \tag{1}$$

For $K \ll K_a$, $p(K|Model,I)$ behaves like a uniform prior and for $K \gg K_a$ it behaves like a Jeffreys prior. The $\ln\left(\frac{K_a + K_{\max}}{K_a}\right)$ term in the denominator ensures that the prior is normalized in the interval 0 to $K_{\max}$.

in addition to the known measurement errors include an empirical estimate of stellar jitter which is due in part to flows and inhomogeneities on the stellar surface (e.g., Wright 2005). For HD 208487 this is estimated to be somewhere in the range $2 - 4$ m s$^{-1}$. We used a value of 3 m s$^{-1}$. In model $M_{1s}$ we neglect the stellar noise jitter but incorporates an additional noise parameter, $s$, that can allow for any additional noise beyond the known measurement uncertainties [1]. For example, suppose that the star actually has two planets, and the model assumes only one is present. In regard to the single planet model, the velocity variations induced by the unknown second planet acts like an additional unknown noise term. Stellar jitter, if present, will also contribute to this extra noise term. In general, nature is more complicated than our model and known noise terms. Marginalizing $s$ has the desirable effect of treating anything in the data that can't be explained by the model and known measurement errors as noise, leading to the most conservative estimates of orbital parameters (see Sections 9.2.3 and 9.2.4 of Gregory (2005b) for a tutorial demonstration of this point). If there is no extra noise then the posterior probability distribution for $s$ will peak at $s = 0$.

The parameter prior probability distributions used in this analysis are given in Table 1. For the $V$ parameter, we use a uniform prior which spans all real values and set $p(V|Model,I) = 1$. While this is an 'improper' (non-normalized) prior, the normalization cancels when comparing different models, since it appears exactly once in each

---

[1] In the absence of detailed knowledge of the sampling distribution for the extra noise, other than that it has a finite variance, the maximum entropy principle tells us that a Gaussian distribution would be the most conservative choice (i.e., maximally non committal about the information we don't have). We will assume the noise variance is finite and adopt a Gaussian distribution with a variance $s^2$. Thus, the combination of the known errors and extra noise has a Gaussian distribution with variance $= \sigma_i^2 + s^2$, where $\sigma_i$ is the standard deviation of the known noise for i$^{th}$ data point.

**TABLE 2.** Bayes factors for model comparison.

| Model | Mean $s$ $(m\ s^{-1})$ | $p(D\|\text{Model}, I)$ | Bayes factor |
|---|---|---|---|
| $M_{1j}$ | | $2.4 \pm 0.1 \times 10^{-55}$ | 1.0 |
| $M_{1s}$ | 8 | $7.0 \pm 0.4 \times 10^{-56}$ | 0.3 |
| $M_2$ | | $6.5 \pm 0.7 \times 10^{-54}$ | 27.0 |
| $M_{2j}$ | | $2.2 \pm 0.2 \times 10^{-54}$ | 9.1 |

**TABLE 3.** Model parameter estimates

| Parameter | Tinney et al. (2005) | New analysis | Parameter | New analysis |
|---|---|---|---|---|
| $P_1$ (d) | $130 \pm 1$ | $129.5^{+0.8}_{-0.3}$ | $P_2$ (d) | $998^{+57}_{-62}$ |
| $K_1$ (m s$^{-1}$) | $20 \pm 2$ | $16.0^{+1.5}_{-1.8}$ | $K_2$ (m s$^{-1}$) | $9.8^{+1.0}_{-2.7}$ |
| $e_1$ | $0.32 \pm 0.10$ | $0.22^{+.14}_{-.07}$ | $e_2$ | $0.19^{+.05}_{-.18}$ |
| $\omega_1$ (deg) | $126 \pm 40$ | $126^{+12}_{-39}$ | $\omega_2$ (deg) | $157^{+112}_{-60}$ |
| $a_1$ (AU) | $0.49 \pm 0.04$ | $0.492^{+.002}_{-.001}$ | $a_2$ (AU) | $1.92^{+.07}_{-.08}$ |
| $M_1 \sin i$ ($M_J$) | $0.45 \pm 0.05$ | $0.37^{+.03}_{-.04}$ | $M_2 \sin i$ ($M_J$) | $0.46^{+.05}_{-.13}$ |
| Periastron passage 1 (JD - 2,450,000) | $1002.8 \pm 10$ | $1010^{+8}_{-20}$ | Periastron passage 2 (JD - 2,450,000) | $446^{+186}_{-120}$ |
| RMS residuals (m s$^{-1}$) | 7.2 | 4.2 | | |
| Reduced $\chi^2$ | 1.27 | 0.83 | | |

model. Note, that during the MCMC iterations there is nothing to constrain $P_1$ to be less than $P_2$. After the MCMC has completed the parameters are redefined so that $P_1 < P_2$.

Table 2 shows the Bayes factors for the 4 models. The Bayes factor is just the ratio of the global likelihood of a particular model to model $M_{1j}$, taken as a reference. Section 5.3 of Gregory (2005a) describes how the Bayes factors are calculated. Assuming equal prior probability for the one and two planet models then Table 2 indicates that a two planet model is favored over a one planet model.

The derived two planet model parameter values are given in Table 3 and compared with the one planet parameters given in Tinney et al. (2005). Figures 1 and 2 show results for the most favored model $M_2$. The upper panel of Figure 1 shows the raw data, the middle panel shows the best two planet fit and the lower panel shows the residuals which have an RMS = 4.2 m s$^{-1}$. Panel (a) of Figure 2 shows the data, minus the best fitting $P_2$ orbit, for two cycles of $P_1$ phase. The best fitting $P_1$ orbit is overlaid. Panel (b) shows the data plotted versus $P_2$ phase with the best fitting $P_1$ orbit removed.
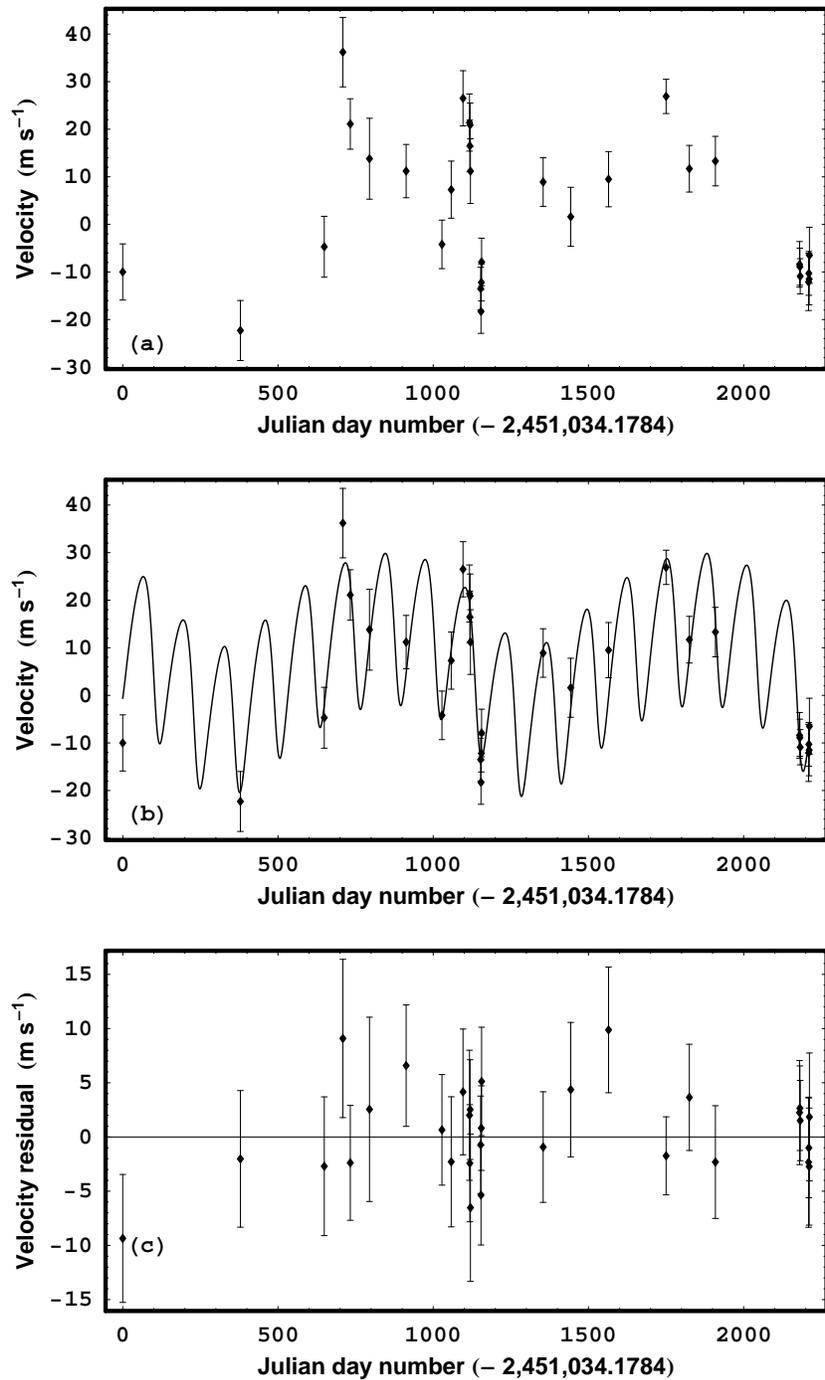
**FIGURE 1.** The raw data is shown in panel (a) and the best fitting two planet ($P_1 = 130$ day, $P_2 = 998$ day) model versus time is shown in (b). Panel (c) shows the residuals.
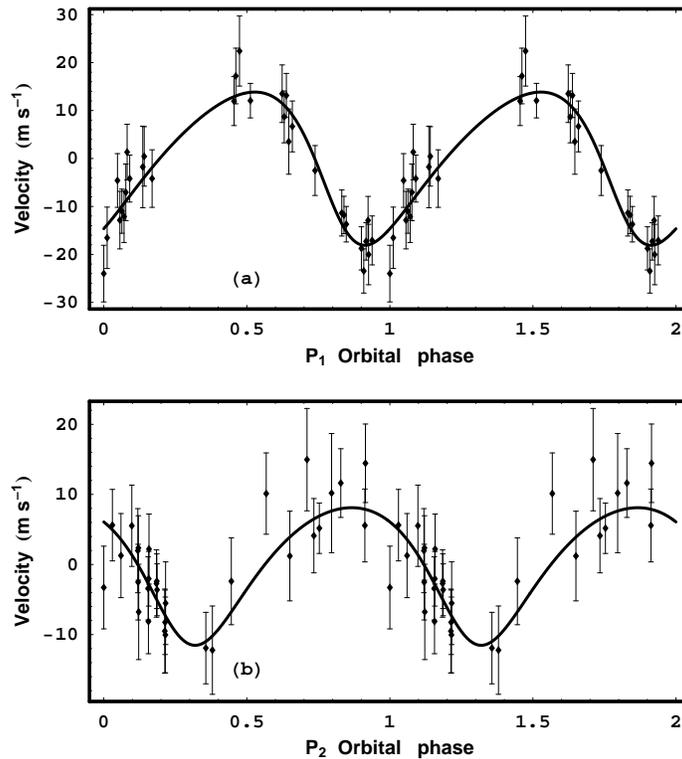
**FIGURE 2.** Panel (a) shows the data, with the best fitting $P_2$ orbit subtracted, for two cycles of $P_1$ phase with the best fitting $P_1$ orbit overlaid. Panel (b) shows the data plotted versus $P_2$ phase with the best fitting $P_1$ orbit removed.

# CONCLUSIONS

In this paper we presented an improved version of the Bayesian parallel tempering MCMC algorithm of Gregory (2005a), which is a powerful Kepler periodogram for detecting multiple planets in precision radial velocity data. It can easily be modified to incorporate precessing Keperian orbits. It is also capable of fitting a portion of an orbital period, so early detection of periods longer than the duration of the data are possible. From a re-analysis of the HD 208487 data of Tinney et al (2005), we find strong evidence for a second planet with a period of $998^{+57}_{-62}$ days, an eccentricity of $0.19^{+0.05}_{-0.18}$, and an $M \sin i = 0.46^{0.05}_{0.13} M_J$.

# ACKNOWLEDGMENTS

# REFERENCES

1. Eggenberger, A., Udry, S., & Mayor, M. 2004, *Astronomy & Astrophysics* 417, 353.
2. Ford, E. B. 2005, *Astronomical Journal*, 129, 1706.
3. Gregory, P. C., *Astrophysical Journal*, in press 2005a.
4. Gregory, P. C. 2005b, 'Bayesian Logical Data Analysis for the Physical Sciences: A Comparative approach with *Mathematica* Support', Cambridge University Press.
5. Tinney, C. G., Butler, R. P., Marcy, G. W., Jones, H. R. A., Penny, A. J., McCarthy, C., Carter, B. D., & Bond, J. 2003, *Astrophysical Journal*, 587, 423.
6. Tinney, C. G., Butler, R. P., Marcy, G. W., Jones, H. R. A., Penny, A. J., McCarthy, C., Carter, B. D., & Fischer, D. A. 2005, *Astrophysical Journal*, 623, 1171.
7. Wright, J. T. 2005, *Publication of the Astronomical Soc. of the Pacific*, 117, 657.

HD 208487. Quite the same Wikipedia. Just better. Ä  The discovery was made using Bayesian analysis of the radial velocity dataset to determine the planetary parameters.[2] However, further analysis revealed that an alternative 2-planet solution for the HD 208487 system was possible, with a planet in a 28-day orbit instead of the 908-day orbit postulated, and it was concluded that activity on the star is more likely to. be responsible for the residuals to the one-planet solution than the presence of a second planet.[3]. A Bayesian multi-planet Kepler periodogram has been developed based on a fusion Markov chain Monte Carlo algorithm (FMCMC). FMCMC is a new general purpose tool for nonlinear model fitting. It incorporates parallel tempering, simulated annealing and genetic crossover operations. Each of these features facilitate the detection of a global minimum in chi-squared in a highly multi-modal environment. By combining all three, the algorithm greatly increases the probability of realizing this goal. View via Publisher. Â  A Bayesian Analysis of Extrasolar Planet Data for HD 208487. P. Gregory. Computer Science, Physics. The planetary system HD 208487 hosts at least one planet. System parameters. Primary system name. HD 208487. Alternative system names. N/A. Â  If you spot an error or if you can contribute additional data to this entry, please send an e-mail to exoplanet@hanno-rein.de. Please include the corrected xml file and a reference to where the new data is coming from, ideally a scientific paper. If you are fluent with git and github, you can also create a pull request or open an issue on the Open Exoplanet Catalogue repository. Please include the reference to the relevant scientific paper in your commit message. 3) FMCMC analysis of HD 208487 and Gliese 581 RV data equations and priors. 4) Demonstration of FMCMC in Mathematica. 5) Bayesian model selection for exoplanets. Â  The Bayesian nonlinear processing of C(fn) also yields, when the data give evidence for them, arbitrarily sharp spectral peaks. Probability density. Power density. Â  Application to extra-solar planet data. Sparse radial velocity measurements for HD 73526. (ref. Tinney, G. C. 2003, Astrophsical Journal, 587, p. 423). An automatic Bayesian Kepler periodogram has been developed for identifying and characterizing multiple planetary orbits in precision radial velocity data. The periodogram is powered by a parallel tempering Markov chain Monte Carlo (MCMC) algorithm which is capable of efficiently exploring a multiplanet model parameter space. Â  Improvements in the periodogram and further tests using data from HD208487 have resulted in the detection of a second planet with a period of 90982-92d, an eccentricity of 0.370.26-0.20, a semimajor axis of 1.870.13-0.14au and an M sini = 0.45+0.11-0.13 MJ. The revised parameters of the first planet are period = 129.8 +/- 0.4d, eccentricity = 0.20 +/- 0.09, semimajor axis = 0.51 +/- 0.02au and M sini = 0.41 +/- 0.05MJ.