Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY

Mon. Not. R. Astron. Soc. **403,** 731–747 (2010) doi:10.1111/j.1365-2966.2009.16233.x

A Bayesian periodogram finds evidence for three planets in 47 Ursae Majoris

Philip C. Gregory[1★] and Debra A. Fischer[2,3★]

[1]Physics and Astronomy Department, University of British Columbia, 6224 Agricultural Rd., Vancouver, BC V6T 1Z1, Canada
[2]Department of Physics and Astronomy, San Francisco State University, San Francisco, CA 9413, USA
[3]Department of Astronomy, Yale University, New Haven, CT 06520, USA

Accepted 2009 December 7. Received 2009 December 3; in original form 2009 August 27

ABSTRACT

A Bayesian analysis of 47 Ursae Majoris radial velocity data confirms and refines the properties of two previously reported planets with periods of 1079 and 2325 d. The analysis also provides orbital constraints on an additional long-period planet with a period of ∼10 000 d. The three-planet model is found to be 10^5 times more probable than the next most probable model which is a two-planet model. The non-linear model fitting is accomplished with a new hybrid Markov chain Monte Carlo (HMCMC) algorithm which incorporates parallel tempering, simulated annealing and genetic crossover operations. Each of these features facilitate the detection of a global minimum in χ^2. By combining all three, the HMCMC greatly increases the probability of realizing this goal. When applied to the Kepler problem, it acts as a powerful multiplanet Kepler periodogram.

The measured periods are 1078 ± 2 d, 2391^{+100}_{-87} d and 14002^{+4018}_{-5095} d, and the corresponding eccentricities are 0.032 ± 0.014, $0.098^{+.047}_{-.096}$ and $0.16^{+.09}_{-.16}$. The results favour low-eccentricity orbits for all three. Assuming the three signals (each one consistent with a Keplerian orbit) are caused by planets, the corresponding limits on planetary mass ($M \sin i$) and semimajor axis are ($2.53^{+.07}_{-.06} M_J$, 2.10 ± 0.02 au), ($0.54 \pm 0.07\ M_J$, 3.6 ± 0.1 au) and ($1.6^{+0.3}_{-0.5} M_J$, $11.6^{+2.1}_{-2.9}$ au), respectively. Based on a three-planet model, the remaining unaccounted for noise (stellar jitter) is $5.7\ \mathrm{m\,s^{-1}}$.

The velocities of model fit residuals were randomized in multiple trials and processed using a one-planet version of the HMCMC Kepler periodogram. In this situation, periodogram peaks are purely the result of the effective noise. The orbits corresponding to these noise-induced periodogram peaks exhibit a well-defined strong statistical bias towards high eccentricity. We have characterized this eccentricity bias and designed a correction filter that can be used as an alternate prior for eccentricity to enhance the detection of planetary orbits of low or moderate eccentricity.

Key words: methods: numerical – methods: statistical – techniques: radial velocities – stars: individual: 47 Ursae Majoris – planetary systems.

1 INTRODUCTION

Improvements in precision radial velocity (RV) measurements and continued monitoring are permitting the detection of lower amplitude planetary signatures. One example of the fruits of this work is the detection of a superearth in the habitable zone surrounding Gliese 581 (Udry et al. 2007). This and other remarkable successes on the part of the observers is motivating a significant effort to improve the statistical tools for analysing RV data (e.g. Loredo & Chernoff 2003; Loredo 2004; Cumming 2004; Ford 2005, 2006; Gregory 2005a,b; Ford & Gregory 2007; Cumming & Dragomir 2010). Much of the recent work has highlighted a Bayesian Markov chain Monte Carlo (MCMC) approach as a way to better understand parameter uncertainties and degeneracies and to compute model probabilities.

Gregory (2005a,b,c and 2007a,b,c) presented a Bayesian MCMC algorithm that makes use of parallel tempering (PT) to efficiently explore a large model parameter space starting from a random location. It is able to identify any significant periodic signal component in the data that satisfies Kepler's laws and thus functions

★E-mail: gregory@phas.ubc.ca (PCG); debra.fischer@yale.edu (DAF)

as a Kepler periodogram.[1] This eliminates the need for a separate periodogram search for trial orbital periods which typically assume a sinusoidal model for the signal that is only correct for a circular orbit. In addition, the Bayesian MCMC algorithm provides full marginal parameter distributions for all the orbital elements that can be determined from RV data. The algorithm includes an innovative two-stage adaptive control system (CS) that automates the selection of efficient Gaussian parameter proposal distributions.

The latest version of the algorithm, Gregory (2009), incorporates a genetic crossover operation into the MCMC algorithm. The new adaptive hybrid MCMC (HMCMC) algorithm incorporates PT, simulated annealing and genetic crossover operations. Each of these techniques was designed to facilitate the detection of a global minimum in χ^2. Combining all three in an adaptive HMCMC greatly increases the probability of realizing this goal.

Butler & Marcy (1996) first reported a 1090 d companion to 47 UMa using data from Lick Observatory. With additional velocity measurements over 13 yr, Fischer et al. (2002) announced a long-period second planet, 47 UMa c, with a period of 2594 ± 90 d and a mass of 0.76 M_J. Naef et al. (2004) reported ELODIE (the fibre fed echelle spectrograph of Observatoire de Haute-Provence) observations of 47 UMa and noted that the second planet was not evident in their data. Wittenmyer, Endl & Cochran (2007) reported that there is still substantial ambiguity as to the orbital parameters of the proposed planetary companion 47 UMa c. They gave a period of 7586 d for one orbital solution and 2594 d for two others. In their latest work, Wittenmyer et al. (2009), their best-fitting two-planet model now calls for $P_2 = 9660$ d. In this paper, we present a Bayesian analysis of the latest Lick Observatory measurements and a combined Lick plus McDonald Observatory (Wittenmyer et al. 2009) data set.

We also report on an investigation of the behavior of the Bayesian HMCMC Kepler periodogram to noise. The noise data sets were formed by randomly interchanging velocity measurements.

2 THE ADAPTIVE HYBRID MCMC

The adaptive HMCMC is a very general Bayesian non-linear model fitting program. After specifying the non-linear model, data and priors, Bayes theorem dictates the target joint probability distribution for the model parameters which can be very complex. To compute the marginals for any subset of the parameters, it is necessary to integrate the joint probability distribution over the remaining parameters. In high dimensions, the principal tool for carrying out the integrals is MCMC based on the Metropolis algorithm. The greater efficiency of an MCMC stems from its ability, after an initial burn-in period, to generate samples in parameter space in direct proportion to the joint target probability distribution. In contrast, straight Monte Carlo (MC) integration randomly samples the parameter space and wastes most of its time sampling regions of very low probability.

An important feature that prevents the HMCMC from becoming stuck in a local probability maximum is PT. Multiple MCMC chains are run in parallel. The joint probability density distribution for the parameters ($\boldsymbol{X}$) of model M_i, for a particular chain, is given by

$$p(\boldsymbol{X}|D, M_i, I, \beta) \propto p(\boldsymbol{X}|M_i, I) \times p(D|\boldsymbol{X}, M_i, I)^{\beta}. \quad (1)$$

Each MCMC chain corresponding to a different β, with the value of β ranging from 0 to 1. When the exponent $\beta = 1$, the term on the left-hand side of the equation is the target joint probability distribution for the model parameters, $p(\boldsymbol{X}|D, M_i, I)$. It is the posterior probability of a particular choice of parameter vector, $\boldsymbol{X}$, given the data represented by D, the model choice M_i and the prior information I. In general, the model parameter space of interest is a continuum so $p(\boldsymbol{X}|D, M_i, I)$ is a probability density distribution. The first term on the right-hand side of the equation, $p(\boldsymbol{X}|M_i, I)$, is the prior probability density distribution of $\boldsymbol{X}$, prior to the consideration of the current data D. The second term, $p(D|\boldsymbol{X}\ M_i, I)$, is called the likelihood and it is the probability that we would have obtained the measured data D for this particular choice of parameter vector $\boldsymbol{X}$, model M_i and prior information I. At the very least, the prior information, I, must specify the class of alternative models (hypotheses) being considered (hypothesis space of interest) and the relationship between the models and the data (how to compute the likelihood). For further details of the likelihood function for this problem see Gregory (2005b). In many situations, the log of the likelihood is simply proportional to the familiar χ^2 statistic. If we later acquire another data set D' then the new prior, $p(\boldsymbol{X}|M_i, I')$, is equal to our previous posterior, $p(\boldsymbol{X}|D, M_i, I)$, i.e. $I' = I, D$. An exponent $\beta = 0$, yields a broader joint probability density equal to the prior. The reciprocal of β is analogous to a temperature, the higher the temperature the broader the distribution.

For parameter estimation purposes eight chains ($\beta = \{0.09, 0.13, 0.20, 0.29, 0.39, 0.52, 0.72, 1.0\}$) were employed. At an interval of 10 iterations, a pair of adjacent chains on the tempering ladder is chosen at random and a proposal made to swap their parameter states. A MC acceptance rule determines the probability for the proposed swap to occur (e.g. Gregory 2005a, equation 12.12). This swap allows for an exchange of information across the population of parallel simulations. In low β (higher temperature) simulations, radically different configurations can arise, whereas in higher β (lower temperature) states, a configuration is given the chance to refine itself. The lower β chains can be likened to a series of scouts that explore the parameter terrain on different scales. The final samples are drawn from the $\beta = 1$ chain, which corresponds to the desired target probability distribution. For $\beta \ll 1$, the distribution is much flatter. The choice of β values can be checked by computing the swap acceptance rate. When they are too far apart, the swap rate drops to very low values.

Each parallel chain employs the Metropolis algorithm. At each iteration a proposal to jump to a new location in parameter space is generated from independent Gaussian proposal distributions (centred on the current parameter location), one for each parameter. In general, the σ's of these Gaussian proposal distributions are different because the parameters can be very different entities. Also if the σ's are chosen too small, successive samples will be highly correlated and will require many iterations to obtain an equilibrium set of samples. If the σ's are too large, then proposed samples will very rarely be accepted. The process of choosing a set of useful proposal σ's when dealing with a large number of different parameters can be very time consuming. In PT MCMC, this problem is compounded because of the need for a separate set of Gaussian proposal σ's for each chain (different tempering levels). This process is automated by an innovative two-stage statistical CS (Gregory 2007b; Gregory 2009) in which the error signal is proportional to the difference between the current joint parameter acceptance rate and a target acceptance rate, typically 25 per cent (Roberts, Gelman & Gilks 1997). A schematic of the full adaptive CS is shown in Fig. 1.

The first-stage CS, which involves annealing the set of Gaussian proposal distribution σ's, was described in Gregory (2005a). An initial set of proposal σ's ($\approx$10 per cent of the prior range for each

[1]Following on from the pioneering work on Bayesian periodograms by Jaynes (1987) and Bretthorst (1988)

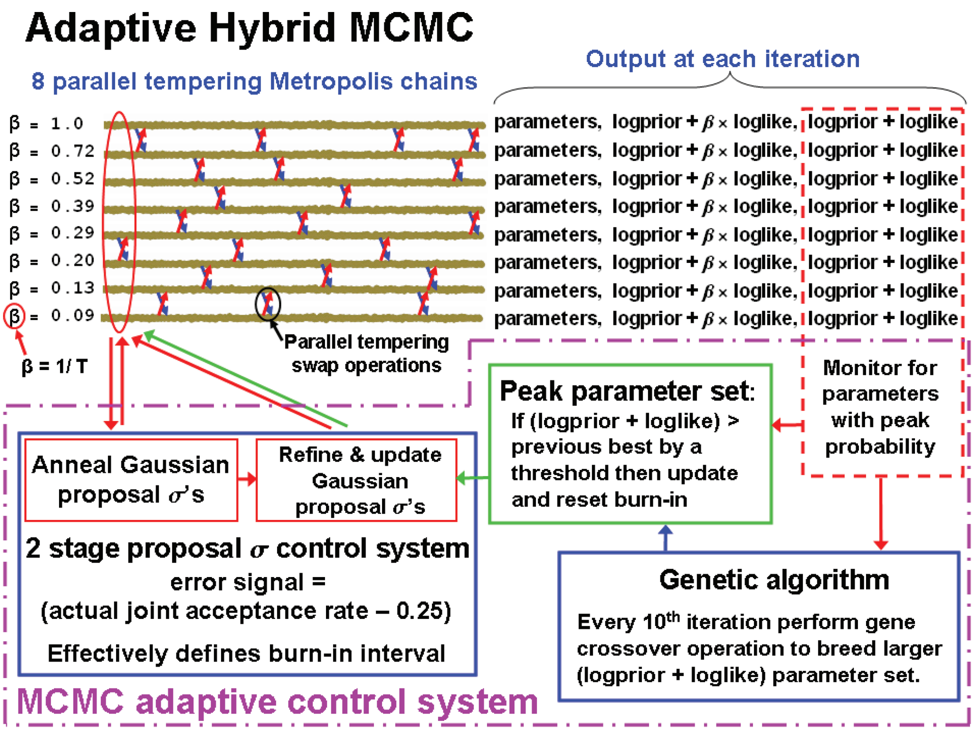

Figure 1. Schematic of the operation of the adaptive HMCMC algorithm.

parameter) is used for each chain. During the major cycles, the joint acceptance rate is measured based on the current proposal σ's and compared to a target acceptance rate. During the minor cycles, each proposal σ is separately perturbed to determine an approximate gradient in the acceptance rate for that parameter. The σ's are then jointly modified by a small increment in the direction of this gradient. This is done for each of the parallel chains. Proposals to swap parameter values between chains are allowed during major cycles but not within minor cycles.

The annealing of the proposal σ's occurs while the MCMC is homing in on any significant peaks in the target probability distribution. Concurrent with this, another aspect of the annealing operation takes place whenever the Markov chain is started from a location in parameter space that is far from the best-fitting values. This automatically arises because all the models considered incorporate an extra additive noise (Gregory 2005b), for reasons discussed in Section 3, whose probability distribution is Gaussian with zero mean and with an unknown standard deviation s. When the χ^2 of the fit is very large, the Bayesian Markov chain automatically inflates s to include anything in the data that cannot be accounted for by the model with the current set of parameters and the known measurement errors. This results in a smoothing out of the detailed structure in the χ^2 surface and, as pointed out by Ford (2006), allows the Markov chain to explore the large-scale structure in parameter space more quickly. The chain begins to decrease the value of s as it settles in near the best-fitting parameters. An example of this is shown in Fig. 2. In the early stages, s is inflated to around 38 m s^{-1} and then decays to a value of $\approx$4 m s^{-1} over the first 9000 iterations. This is similar to simulated annealing, but does not require choosing a cooling scheme.

Although the first-stage CS achieves the desired joint acceptance rate, it often happens that a subset of the proposal σ's are too small leading to an excessive autocorrelation in the MCMC iterations for these parameters. Part of the second-stage CS corrects for this.

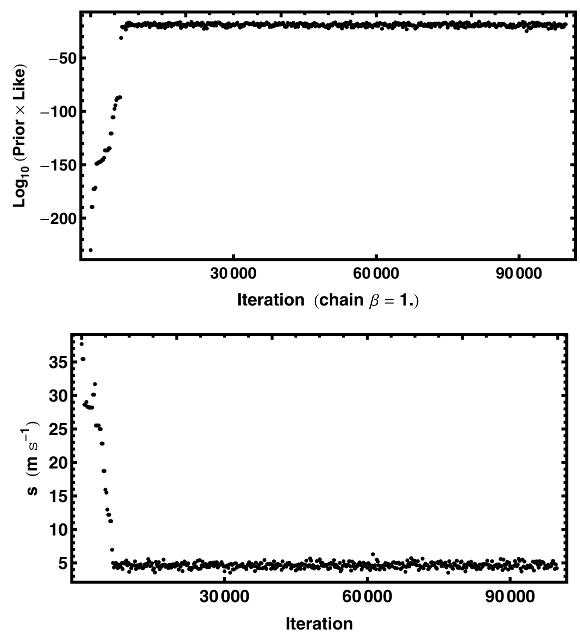

Figure 2. The upper panel is a plot of the Log$_{10}$[Prior $\times$ Likelihood] versus MCMC iteration. The lower panel is a similar plot for the extra-noise term s. Initially s is inflated and then rapidly decays to a much lower level as the best-fitting parameter values are approached.

The goal of the second stage is to achieve a set of proposal σ's that equalizes the MCMC acceptance rates when new parameter values are proposed separately and achieves the desired acceptance rate when they are proposed jointly. Details of the second-stage CS were given in Gregory (2007b).

The first stage is run only once at the beginning, but the second stage can be executed repeatedly, whenever a significantly improved parameter solution emerges. Frequently, the burn-in period occurs within the span of the first-stage CS, i.e. the significant peaks in the joint parameter probability distribution are found, and the second stage improves the choice of proposal σ's based on the highest probability parameter set. Occasionally, a new higher (by a user-specified threshold) target probability parameter set emerges after the first two stages of the CS are completed. The CS has the ability to detect this and automatically reactivate the second stage. In this sense, the CS is adaptive. If this happens, the iteration corresponding to the end of the CS is reset. The useful MCMC simulation data is obtained after the CS are switched off.

The adaptive capability of the CS can be appreciated from an examination of Fig. 1. The upper left portion of the figure depicts the MCMC iterations from the eight parallel chains, each corresponding to a different tempering level β as indicated on the extreme left. One of the outputs obtained from each chain at every iteration (shown at the far right) is the log prior + log likelihood. This information is continuously fed to the CS which constantly updates the most probable parameter combination regardless of which chain the parameter set occurred in. This is passed to the 'Peak parameter set' block of the CS. Its job is to decide if a significantly more probable parameter set has emerged since the last execution of the second-stage CS. If so, the second-stage CS is rerun using the new more probable parameter set which is the basic adaptive feature of the CS.

The CS also includes genetic algorithm block which is shown in the bottom right of Fig. 1. The current parameter set can be treated as a set of genes. In the present version, one gene consists of the parameter set that specifies one orbit. On this basis, a three-planet model has three genes. At any iteration, there exists within the CS the most probable parameter set to date $\boldsymbol{X}_{\max}$ and the most probable parameter set from the eight chains for the most recent iteration $\boldsymbol{X}_{\mathrm{cur}}$. At regular intervals (user specified) each gene from $\boldsymbol{X}_{\mathrm{cur}}$ is swapped for the corresponding gene in $\boldsymbol{X}_{\max}$. If either substitution leads to a higher probability it is retained and $\boldsymbol{X}_{\max}$ updated. The effectiveness of this operation can be tested by comparing the number of times the gene crossover operation gives rise to a new value of $\boldsymbol{X}_{\max}$ compared to the number of new $\boldsymbol{X}_{\max}$ arising from the normal PT MCMC iterations. The gene crossover operations prove to be very effective and give rise to new $\boldsymbol{X}_{\max}$ values $\approx$3 times more often than MCMC operations. Of course, most of these swaps lead to very minor changes in probability but occasionally big jumps are created.

Gene swaps from $\boldsymbol{X}_{\mathrm{cur2}}$, the parameters of the second most probable current chain, to $\boldsymbol{X}_{\max}$ are also utilized. This gives rise to new values of $\boldsymbol{X}_{\max}$ at a rate approximately half that of swaps from $\boldsymbol{X}_{\mathrm{cur}}$ to $\boldsymbol{X}_{\max}$. Crossover operations at a random point in the entire parameter set did not prove as effective except in the single planet case where there is only one gene. Further experimentation with this concept is ongoing.

3 DATA AND ANALYSIS

Our initial analysis was based on data obtained at the Lick Observatory and spans a period of 21.6 yr. The data are listed in Tables 1 and 2. In addition to the observation time, RV and velocity error (ΔRV), the detector dewar number used is also included. We originally analysed the data ignoring possible residual velocity offsets associated with dewar changes (Case A). To investigate how robust the results were, we subsequently repeated the analysis incorporating the dewar velocity offsets as additional unknown parameters (Case B). In Case A, the data from all six dewars are used. For Case B, we excluded dewar 1 because with only a single measurement the analysis is unable to separate the offset from the model velocity contribution which reduces the time base by 235 d. Results for the two cases follow in subsequent sections labelled accordingly. In Section 6, we extend the analysis to include the Wittenmyer et al. (2009) data from the 9.2-m Hobby-Eberly Telescope (HET) and 2.7-m Harlam J. Smith (HJS) telescopes of the McDonald Observatory. In the rest of this section, we describe the model fitting equations and the selection of priors for the model parameters. We also characterize a noise-induced eccentricity bias that leads to a second choice for an eccentricity prior.

We have investigated the 47 UMa data using models ranging from a single planet to five planets. For a one-planet model the predicted RV is given by

$$v(t_i) = V + K\{\cos[\theta(t_i + \chi P) + \omega] + e\cos\omega\}, \tag{2}$$

and involves the six unknown parameters:

V = a constant velocity,
K = velocity semi-amplitude,
P = the orbital period,
e = the orbital eccentricity,
ω = the longitude of periastron,
χ = the fraction of an orbit, prior to the start of data taking, that periastron occurred at. Thus, χP = the number of days prior to $t_i = 0$ that the star was at periastron, for an orbital period of P days.
$\theta(t_i + \chi P)$ = the true anomaly, the angle of the star in its orbit relative to periastron at time t_i.

We utilize this form of the equation because we obtain the dependence of θ on t_i by solving the conservation of angular momentum equation

$$\frac{\mathrm{d}\theta}{\mathrm{d}t} - \frac{2\pi[1 + e\cos\theta(t_i + \chi\ P)]^2}{P(1-e^2)^{3/2}} = 0. \tag{3}$$

Our algorithm is implemented in MATHEMATICA and it proves faster for MATHEMATICA to solve this differential equation than solve the equations relating the true anomaly to the mean anomaly via the eccentric anomaly. MATHEMATICA generates an accurate interpolating function between t and θ so the differential equation does not need to be solved separately for each t_i. Evaluating the interpolating function for each t_i is very fast compared to solving the differential equation, so the algorithm should be able to handle much larger samples of RV data than those currently available without a significant increase in computational time. For example, an increase in the data by a factor of 6.5 resulted in only an 18 per cent increase in execution time.

As described in more detail in Gregory (2007a), we employed a reparameterization of χ and ω to improve the MCMC convergence speed motivated by the work of Ford (2006). The two new parameters are $\psi = 2\pi\chi + \omega$ and $\phi = 2\pi\chi - \omega$. Parameter ψ is well determined for all eccentricities. Although ϕ is not well determined for low eccentricities, it is at least orthogonal to the ψ parameter. We use a uniform prior for ψ in the interval 0 to 4π and uniform prior for ϕ in the interval -2π to $+2\pi$. This ensures that a prior that is wraparound continuous in (χ, ω) maps into a wraparound continuous distribution in (ψ, ϕ). The big (ψ, ϕ) square holds two copies of the probability patch in (χ, ω) which does not matter. What matters is that the prior is now wraparound continuous in (ψ, ϕ).

Table 1. Radial velocities (RV) for 47 UMa. The ΔRV column gives the RV uncertainty and the next column gives the detector dewar number.

JD 244 0000	RV ($m\,s^{-1}$)	ΔRV ($m\,s^{-1}$)	Dewar	JD 244 0000	RV ($m\,s^{-1}$)	ΔRV ($m\,s^{-1}$)	Dewar	JD 244 0000	RV ($m\,s^{-1}$)	ΔRV ($m\,s^{-1}$)	Dewar
6959.7372	−40.70	14.00	1	11607.9163	−17.77	4.51	18	12722.8295	−20.88	3.13	24
7194.9122	−33.96	7.49	6	11626.7707	−34.76	6.65	18	12737.7703	−10.01	2.47	24
7223.7982	−18.31	6.14	6	11627.7539	−29.07	5.87	18	12793.7298	1.53	2.41	24
7964.8927	20.40	8.19	6	11628.7275	−34.86	5.71	18	12794.7134	−5.06	2.20	24
8017.7302	−8.18	10.57	6	11629.8320	−32.06	4.48	18	12834.6981	21.08	2.83	24
8374.7707	−20.25	9.37	6	11700.6937	−2.83	4.80	18	12991.0537	57.90	3.94	24
8647.8971	62.95	11.41	8	11861.0498	36.20	5.53	18	12992.0732	55.57	4.69	24
8648.9100	51.93	11.02	8	11874.0684	39.39	5.34	18	13009.0525	53.57	2.70	24
8670.8777	74.56	11.45	8	11881.0443	32.79	4.41	18	13009.9546	51.65	2.88	24
8745.6907	71.89	8.76	8	11895.0663	33.89	4.28	18	13018.9971	55.32	4.48	24
8992.0612	23.42	11.21	8	11906.0148	34.69	3.91	18	13020.9531	39.96	5.42	24
9067.7708	4.86	7.00	8	11907.0112	37.74	4.24	18	13022.0027	46.17	5.15	24
9096.7339	−6.19	6.79	8	11909.0420	39.07	3.76	18	13044.9198	58.89	3.33	24
9122.6909	−27.90	7.91	8	11910.9537	36.96	4.13	18	13068.8447	54.81	5.38	24
9172.6855	−18.68	10.55	8	11914.0674	34.35	5.17	18	13069.8323	48.36	3.34	24
9349.9122	−32.93	9.52	8	11915.0473	41.14	3.72	18	13072.8875	45.63	2.93	24
9374.9638	−29.14	8.67	8	11916.0335	40.99	3.47	18	13078.8069	52.75	3.30	24
9411.8387	−16.88	12.81	8	11939.9703	42.47	4.72	18	13079.8275	52.69	3.18	24
9481.7197	−33.01	13.40	8	11946.9598	42.21	4.19	18	13080.7919	52.88	3.27	24
9767.9184	64.68	5.34	39	11969.9024	48.36	4.29	18	13081.8171	48.72	2.99	24
9768.9072	62.32	4.79	39	11971.8934	52.56	4.80	18	13100.8148	53.34	3.83	24
9802.7911	63.99	3.61	39	11998.7785	49.07	3.81	18	13107.7773	35.01	4.43	24
10058.0797	32.21	3.18	39	11999.8203	48.13	3.98	18	13119.7426	50.57	4.24	24
10068.9773	36.13	4.01	39	12000.8587	50.97	4.16	18	13120.6914	42.03	2.83	24
10072.0117	38.76	4.10	39	12028.7386	60.65	4.39	18	13131.6826	51.29	4.13	24
10088.9932	23.38	3.54	39	12033.7461	49.37	4.93	18	13132.7334	38.98	5.04	24
10089.9473	26.18	3.19	39	12040.7593	47.52	3.54	18	13147.6943	44.95	4.94	24
10091.9004	18.37	4.23	39	12041.7192	49.30	3.37	18	13155.7006	38.98	2.92	24
10120.9179	17.53	3.91	39	12042.6957	45.95	3.88	18	13156.7062	40.23	2.77	24
10124.9042	23.41	3.69	39	12071.7291	53.86	9.39	18	13157.6869	43.30	2.98	24
10125.8234	18.49	3.61	39	12073.7217	44.45	4.61	18	13339.0682	9.72	3.31	24
10127.8979	13.98	3.77	39	12101.6865	59.74	6.62	18	13363.0139	2.22	4.52	24
10144.8770	13.75	4.67	39	12103.6875	41.81	5.71	18	13363.9655	−9.51	4.21	24
10150.7964	12.07	3.89	39	12104.6855	47.90	5.78	18	13383.9778	−11.43	9.04	24
10172.8289	4.69	4.13	39	12105.6836	41.69	5.71	18	13385.0057	−23.57	4.17	24
10173.7627	9.36	5.29	39	12216.0355	27.62	4.56	18	13385.9946	−25.20	3.94	24
10181.7425	−2.47	3.18	39	12222.0432	28.69	4.35	18	13388.0012	−19.02	10.53	24
10187.7390	7.94	4.22	39	12278.0718	−6.78	4.79	18	13389.9276	−32.39	4.48	24
10199.7291	5.49	3.62	39	12279.0680	−2.81	4.54	18	13390.9468	−18.25	4.75	24
10203.7330	1.63	4.23	39	12283.0395	2.35	7.53	18	13391.9987	−30.29	4.56	24
10214.7308	−2.09	3.54	39	12286.0614	−4.09	3.53	18	13392.9238	−31.99	4.95	24
10422.0176	−32.32	4.05	39	12288.0176	−3.07	4.86	18	13402.9585	−13.79	4.74	24
10438.0010	−23.92	4.30	39	12306.9303	−20.54	6.38	18	13403.9527	−23.51	4.68	24
10442.0273	−26.34	3.84	39	12314.9275	−15.06	3.57	18	13404.9472	−24.04	5.42	24
10502.8535	−15.99	3.86	39	12315.9273	−12.71	2.72	18	13436.7878	−24.91	5.44	24
10504.8594	−19.78	4.24	39	12316.9996	−0.12	6.13	18	13437.8865	−40.32	5.46	24
10536.8441	−3.96	4.58	39	12348.8617	−22.34	4.03	18	13438.8413	−31.99	3.91	24
10537.8426	−6.81	3.81	39	12375.7996	−26.80	3.35	24	13439.8543	−36.26	4.13	24
10563.6734	−0.73	3.76	39	12376.7234	−28.65	3.71	24	13440.7724	−32.85	5.39	24
10579.6952	11.11	3.55	39	12380.7568	−28.60	3.90	18	13441.8656	−34.10	5.38	24
10610.7188	12.05	3.34	39	12388.7530	−34.84	3.12	24	13460.8047	−39.69	4.47	24
10793.9570	58.79	3.97	39	12389.7036	−43.22	3.76	24	13475.7043	−39.74	4.67	24
10795.0391	62.55	4.07	39	12577.0504	−37.38	3.58	24	13476.7068	−39.51	4.68	24
10978.6848	55.48	4.51	18	12599.0475	−33.96	2.53	24	13477.7253	−38.11	4.41	24
11131.0654	37.48	6.35	18	12609.0665	−38.85	3.56	24	13478.7598	−43.63	4.14	24
11175.0273	21.32	7.24	18	12631.9926	−26.01	3.61	24	13479.7748	−47.47	4.21	24
11242.8418	1.34	4.82	18	12657.0184	−41.11	2.29	24	13511.7132	−53.62	4.11	24
11303.7119	−25.20	4.20	18	12687.8597	−26.21	3.48	24	13512.6881	−40.62	4.35	24
11508.0703	−36.52	8.34	18	12688.9015	−34.23	5.04	24	13744.0283	−38.93	4.31	24
11536.0640	−43.83	4.75	18	12705.8382	−25.21	2.74	24	13744.9815	−40.34	4.30	24

Table 2. Radial velocities (RV) for 47 UMa. The ΔRV column gives the RV uncertainty and the next column gives the detector dewar number.

JD 244 0000	RV ($m\,s^{-1}$)	ΔRV ($m\,s^{-1}$)	Dewar	JD 244 0000	RV ($m\,s^{-1}$)	ΔRV ($m\,s^{-1}$)	Dewar	JD 244 0000	RV ($m\,s^{-1}$)	ΔRV ($m\,s^{-1}$)	Dewar
13753.0361	−53.52	2.95	24	14135.8630	24.39	2.32	24	14598.7489	−40.52	3.10	24
13755.8982	−41.94	5.04	24	14165.8471	30.47	3.34	24	14622.7505	−41.29	2.42	24
13773.8466	−51.29	4.91	24	14196.8162	35.39	3.14	24	14623.7115	−34.55	2.52	24
13866.7278	−21.49	4.38	24	14219.7662	24.68	3.08	24	14784.0515	−31.13	5.10	24
13867.7226	−27.25	4.67	24	14220.7881	33.23	3.26	24	14785.0826	−34.05	4.85	24
13868.7523	−25.55	4.56	24	14253.6937	27.72	2.73	24	14845.0201	−30.72	1.74	24
13869.7295	−13.48	4.15	24	14254.7002	24.49	2.65	24	14847.9355	−26.93	3.42	24
14074.0693	34.48	3.23	24	14427.0782	−6.25	4.40	24	14848.9727	−30.86	2.74	24
14099.0854	40.26	3.15	24	14450.0617	−10.65	3.41	24	14849.9710	−27.88	3.09	24
14100.0667	32.10	3.21	24	14462.0257	−16.81	2.42	24	14850.9698	−31.85	3.06	24
14102.0466	36.94	3.38	24	14547.9127	−27.51	3.24	24	14863.9813	−27.92	4.26	24
14104.0288	36.91	4.44	24	14574.8034	−52.51	1.79	24	14864.9193	−29.72	5.10	24
14133.9656	32.61	4.66	24	14578.8416	−41.13	2.11	24	14865.9624	−19.55	5.54	24
14134.9264	25.80	2.71	24								

Table 3. Prior parameter probability distributions.

Parameter	Prior	Lower bound	Upper bound
Orbital frequency	$p(\ln f_1, \ln f_2, \ldots \ln f_n \vert M_n, I) = \frac{n!}{[\ln(f_H/f_L)]^n}$ (n = number of planets)	1/1.5 (d)	1/1000 (yr)
Velocity K_i	Modified Jeffreys[a]	0 ($K_0 = 1$)	$K_{\rm max}\left(\frac{P_{\rm min}}{P_i}\right)^{1/3}\frac{1}{\sqrt{1-e_i^2}}$
($m\,s^{-1}$)			
	$\frac{(K+K_0)^{-1}}{\ln\left[1+\frac{K_{\rm max}}{K_0}\left(\frac{P_{\rm min}}{P_i}\right)^{1/3}\frac{1}{\sqrt{1-e_i^2}}\right]}$		$K_{\rm max} = 2129$
V ($m\,s^{-1}$)	Uniform	$-K_{\rm max}$	$K_{\rm max}$
e_i Eccentricity	a) Uniform	0	1
	b) Ecc. noise bias correction filter	0	0.99
ω_i Longitude of periastron	Uniform	0	2π
s Extra noise ($m\,s^{-1}$)	$\frac{(s+s_0)^{-1}}{\ln\left(1+\frac{s_{\rm max}}{s_0}\right)}$	0 ($s_0 = 1$)	$K_{\rm max}$

[a]Since the prior lower limits for K and s include zero, we used a modified Jeffreys prior of the form

$$p(X|M, I) = \frac{1}{X + X_0}\,\frac{1}{\ln\left(1 + \frac{X_{\rm max}}{X_0}\right)} \quad (4)$$

For $X \ll X_0$, $p(X|M, I)$ behaves like a uniform prior and for $X \gg X_0$ it behaves like a Jeffreys prior. The $\ln(1 + \frac{X_{\rm max}}{X_0})$ term in the denominator ensures that the prior is normalized in the interval 0 to $X_{\rm max}$.

In a Bayesian analysis, we need to specify a suitable prior for each parameter. These are tabulated in Table 3. For the current problem, the prior given in equation (1) is the product of the individual parameter priors. Detailed arguments for the choice of each prior were given in Gregory (2007a).

Gregory (2007a) discussed two different strategies to search the orbital frequency parameter space for a multiplanet model: (i) an upper bound on $f_1 \le f_2 \le \ldots \le f_n$ is utilized to maintain the identity of the frequencies, and (ii) all f_i are allowed to roam over the entire frequency range and the parameters re-labelled afterwards. Case (ii) was found to be significantly more successful at converging on the highest posterior probability peak in fewer iterations during repeated blind frequency searches. In addition, case (ii) more easily permits the identification of two planets in 1:1 resonant orbits. We adopted approach (ii) in the current analysis.

All of the models considered in this paper incorporate an extra-noise parameter, s, that can allow for any additional noise beyond the known measurement uncertainties.[2] We 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 σ_i is the standard deviation of the known noise for ith data point. 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. Other factors like star spots and chromospheric activity can also contribute to this extra velocity noise term which is often referred to as stellar jitter. Several researchers have attempted to estimate stellar jitter for individual

[2]In the absence of detailed knowledge of the sampling distribution for the extra noise, we pick a Gaussian because for any given finite noise variance it is the distribution with the largest uncertainty as measured by the entropy, i.e. the maximum entropy distribution (Jaynes 1957; Gregory 2005a, section 8.7.4).

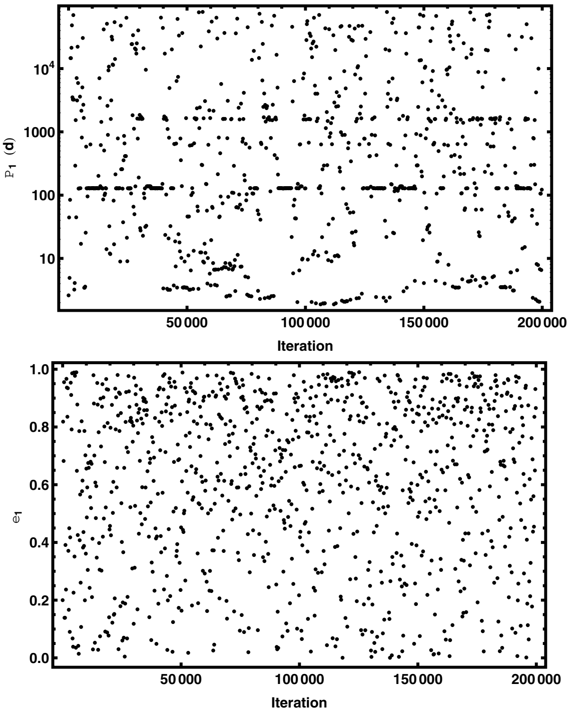

Figure 3. The upper panel shows MCMC period parameter versus iteration for a one-planet model fit to residuals (with randomized velocity values) from a three-planet model fit. The lower panel is the same for the eccentricity parameter.

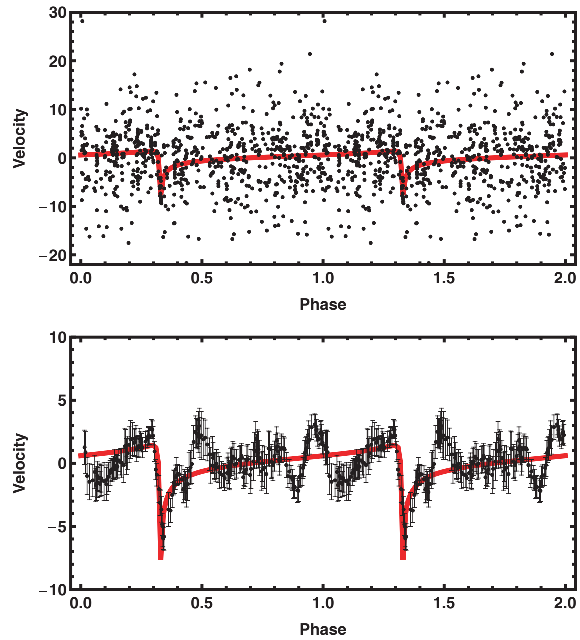

Figure 4. A typical high-eccentricity orbit (in this case $e = 0.93$) found from an MCMC fit of a one-planet model to residuals with randomized velocities. The upper panel shows the raw data points plotted versus two cycles of period phase and the lower panel shows binned averages.

stars based on statistical correlations with observables (e.g. Saar & Donahue 1997; Saar, Butler & Marcy 1998; Wright 2005). 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 cannot be explained by the model and known measurement errors as noise, leading to conservative estimates of orbital parameters. See sections 9.2.3 and 9.2.4 of Gregory (2005a) 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 upper limit on s was set equal to $K_{\max}$. We employed a modified Jeffrey's prior for s with a knee, $s_0 = 1$ m s^{-1}.

We used two different choices of priors for eccentricity, a uniform prior and eccentricity noise bias correction filter that is described in the next section.

3.1 Eccentricity bias

When searching for low-amplitude orbits, any true signal has to compete against spurious orbital signals arising from noise. It was observed that the majority of the probability peaks detected in low signal-to-noise ratio residuals exhibited high eccentricities. The upper panel in Fig. 3 shows MCMC period parameter versus iteration for a one-planet model fit to residuals (with randomized velocity values) from a three-planet model fit. The lower panel is the same for the eccentricity parameter. The HMCMC finds many probability peaks spread over the full period range. There is no significance to the concentration of periods around 100 and 1500 d as the location of period concentrations changes markedly in other realizations of the velocity randomization. The concentration of eccentricity towards higher values is a regular feature. The corresponding plot of eccentricity shows a preponderance of high-eccentricity values. Fig. 4 shows a phase plot for one of these high-eccentricity orbits which provides further insight into why high-eccentricities orbits are favoured. It is clear that for most of the orbit ($e = 0.93$) the predicted shape is relatively flat providing an agreeable fit to points that fluctuate in an uncorrelated noise-like fashion about some mean. Only for a small portion of the orbit does the noise have to conspire to give rise to the rapidly changing orbital velocity peak. To mimic a circular velocity orbit, the noise points would have to appear correlated over a larger fraction of the orbit. For this reason, it is more likely that noise will give rise to spurious highly eccentric orbits than low-eccentricity orbits.

To explore this effect more quantitatively, we analysed a large number of real data sets where the observing times were kept fixed but the velocity residual data was randomly reorganized. In each trial, we fit a one-planet orbit model which explored eccentricities in the range 0 to 0.99 using the one-planet Bayesian Kepler periodogram. In the first instance the data used was the five-planet fit residuals for 55 Cancri. The data for 55 Cancri were a mixture of Lick and Keck Observatory data. When the residual velocities were randomized, the error associated with a particular velocity was shifted with its velocity because the quoted errors were very different for the two observatories. The red curve in the left-hand panel of Fig. 5 is the average of five different 55 Cancri randomized residual trials. The green curve is the average of four trials of randomized residuals from a two-planet 47 UMa model fit, and the blue curve the average of eight trials of randomized residuals from a three-planet 47 UMa model fit. All three curves are very similar and indicate a strong noise-induced eccentricity bias towards high eccentricities.

To increase the chance of detecting and defining the parameters of low- and moderate-eccentricity orbits, we have constructed an eccentricity noise bias correction filter from the reciprocal of the average of the three eccentricity bias curves just mentioned. The lower panel of Fig. 5 shows the best-fitting polynomial (dashed curve) to the reciprocal of the mean of the three eccentricity bias

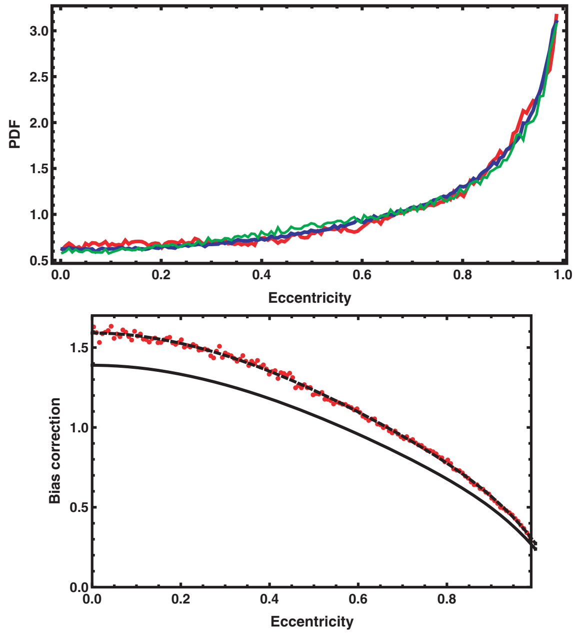

Figure 5. The upper panel shows the marginal probability densities for the eccentricity parameter obtained from MCMC one-planet fits to randomized residuals from 47 UMa two-planet model fits (green), three-planet (blue) and 55 Cancri five-planet (red) fits. The green curve is the average of four trials, the blue curve is the average of eight trials and the red curve is the average of five trials. The lower panel shows the best-fitting polynomial (dashed curve) to the reciprocal of the mean of the three eccentricity bias curves (red points). After normalization, this yields the eccentricity noise bias correction filter (solid black curve).

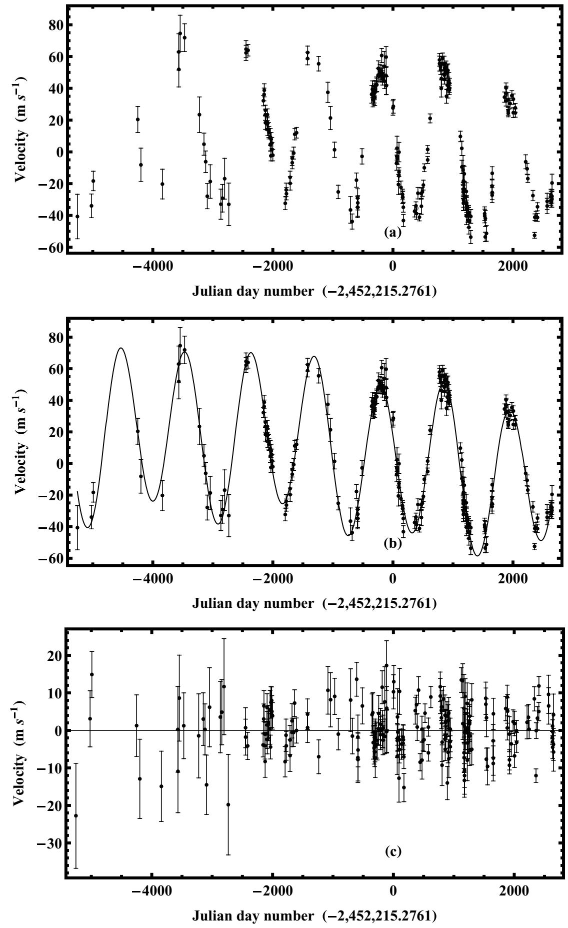

Figure 6. Panel (a) shows the Lick Observatory observations of 47 UMa. Panel (b) shows the final Case A best three-planet model fit compared to the data and panel (c) shows the residuals.

curves (red points). After normalizing the best-fitting polynomial so that the integral is equal to unity over the search range ($e = 0$ to 0.99), we obtain the eccentricity noise bias correction filter (solid black curve). This becomes our second option for a choice of prior for eccentricity. The probability density function for this filter (solid black curve) is given by

$$\mathrm{pdf}(e) = 1.3889 - 1.5212e^2 + 0.53944e^3 - 1.6605(e - 0.24821)^8. \quad (5)$$

On the basis of our understanding of the mechanism underlying the noise-induced eccentricity bias, we expect the eccentricity prior filter to be generally applicable to searches for low-amplitude orbital signals in other precision RV data sets.

An obvious further question that remains to be explored is to what extent the observed distribution of published orbital eccentricities is influenced by such a bias.

4 RESULTS (CASE A)

For Case A, the dewar velocity offsets with respect to our reference dewar 24 are assumed to be zero.

4.1 Parameter estimation

In this section, we present the results of an exploration of the 47 UMa data with the multiplanet HMCMC Kepler periodogram starting with a one-planet model and extending to a five-planet model. The data for 47UMa are shown in Fig. 6, panel (a). Panel (b) shows our final best three-planet model fit compared to the data, and panel (c) shows the residuals.

The one-planet model turned up the 1080 d period which is clearly visible by eye in the raw data. We do not show any results for that model except to compute the marginal likelihood for model-selection purposes which is presented in Section 4.2.

Fig. 7 shows a plot of Log_{10}[Prior × Likelihood] (upper) and period (lower) versus HMCMC iteration (every 200th point) for a two-planet model. The starting periods of 4.7 and 1080 d are shown on the left-hand side of the lower plot at a negative iteration number. The burn-in period of approximately 70 000 iterations is clearly discernable.

Fig. 8 shows a plot of eccentricity versus period for a sample of the HMCMC parameter samples for the two-planet model. Since the duration of the data set is only 7906 d, it is not surprising that uncertainties on the parameters of the second orbit are very large. On the basis of a two-planet model, the parameters of the second planet are $P_2 = 7952^{+388}_{-348}$ d and $e_2 = 0.43^{+0.05}_{-0.08}$. It is clear that e_2 has a low eccentricty tail which reaches zero for a value of $P_2 \approx 9500$ d. This agrees with the value of $P_2 = 9660$ d found by Wittenmyer et al. (2009) in their best-fitting two-planet model where they fixed $e_2 = 0.005$, the values proposed by Fischer et al. (2002).

Fig. 9 shows plots of the three period parameters versus HMCMC iteration for a three-planet model[3] with Log_{10}[Prior × Likelihood]

[3]The HMCMC runs shown here used the eccentricity prior based on the eccentricity noise bias correction filter discussed in Section 3.1. The results obtained using a uniform eccentricity prior are qualitatively the same.

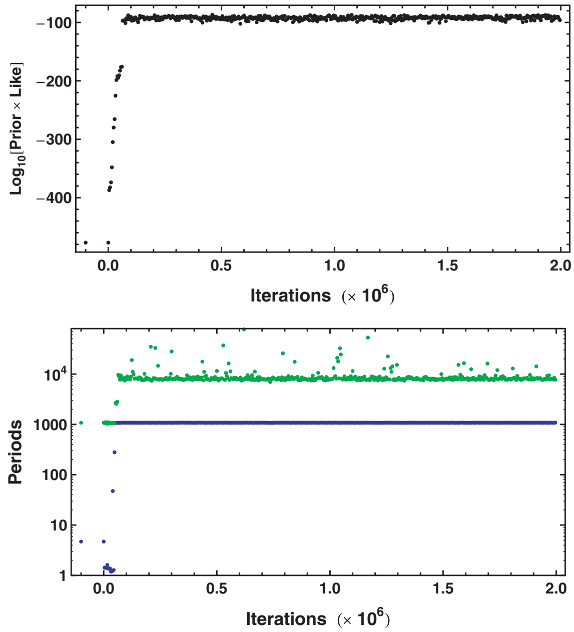

Figure 7. Plot of Log_{10}[Prior × Likelihood] (upper) and period (lower) versus MCMC iteration for a two-planet model.

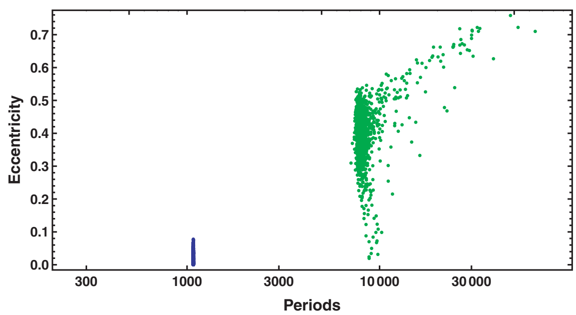

Figure 8. A plot of eccentricity versus period for the two-planet fit (Case A).

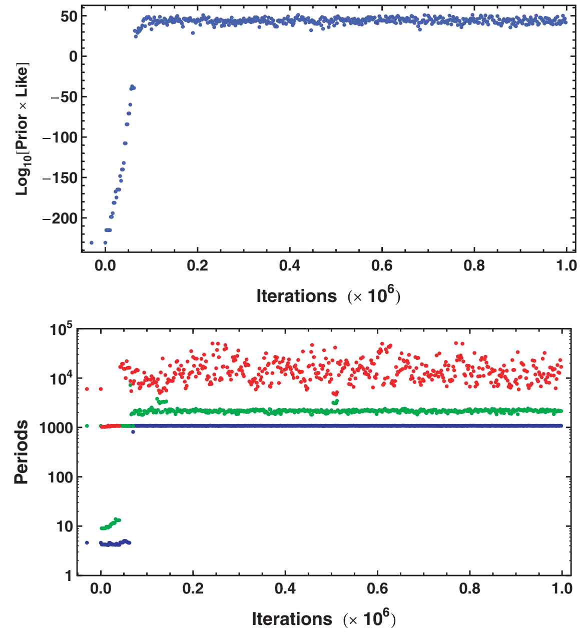

Figure 9. Plot of Log_{10}[Prior × Likelihood] (upper) and period (lower) versus HMCMC iteration for a three-planet fit.

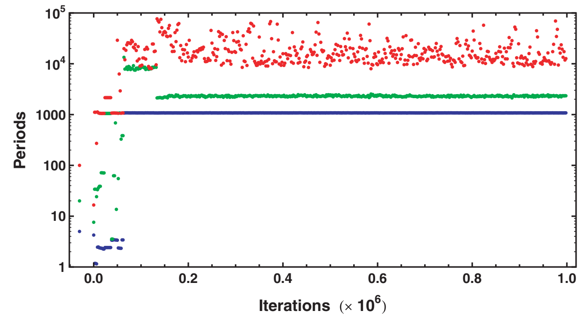

Figure 10. Plot of period versus HMCMC iteration for a three-planet fit. In this case the start periods were 5, 20, 100 d.

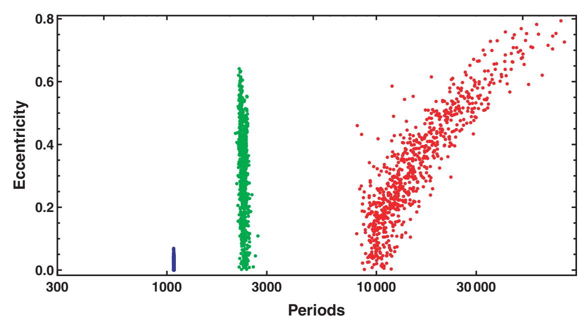

Figure 11. A plot of eccentricity versus period for the three-planet HMCMC (Case A).

plotted above. A new period of 2300 d has emerged and the longest period has shifted from 7952 to ∼10 000 d, and this feature is considerably broader. The starting periods of 89, 1080 and 7200 d are shown on the left at a negative iteration number. Previous experience with the HMCMC periodogram (Gregory 2009) indicates that it is capable of finding a global peak in a blind search of parameter space for a three-planet model. Fig. 10 shows the results of a blind search starting from three very different periods of 5, 20 and 100 d. The algorithm readily finds the same set of final periods in both cases.

Fig. 11 shows a plot of eccentricity versus period for a sample of the HMCMC parameter samples for the three-planet model. There is a large uncertainty in the eccentricity of the two largest periods which extends down to very low eccentricities.

Fig. 12 shows the marginal probability distributions for the periods, eccentricities and K values for the three orbits found. The 10th plot is s, the σ of the added white noise term. A summary of the three-planet model parameters and their uncertainties are given in Table 4. The parameter value listed is the median of the marginal probability distribution for the parameter in question, and the error bars identify the boundaries of the 68.3 per cent credible

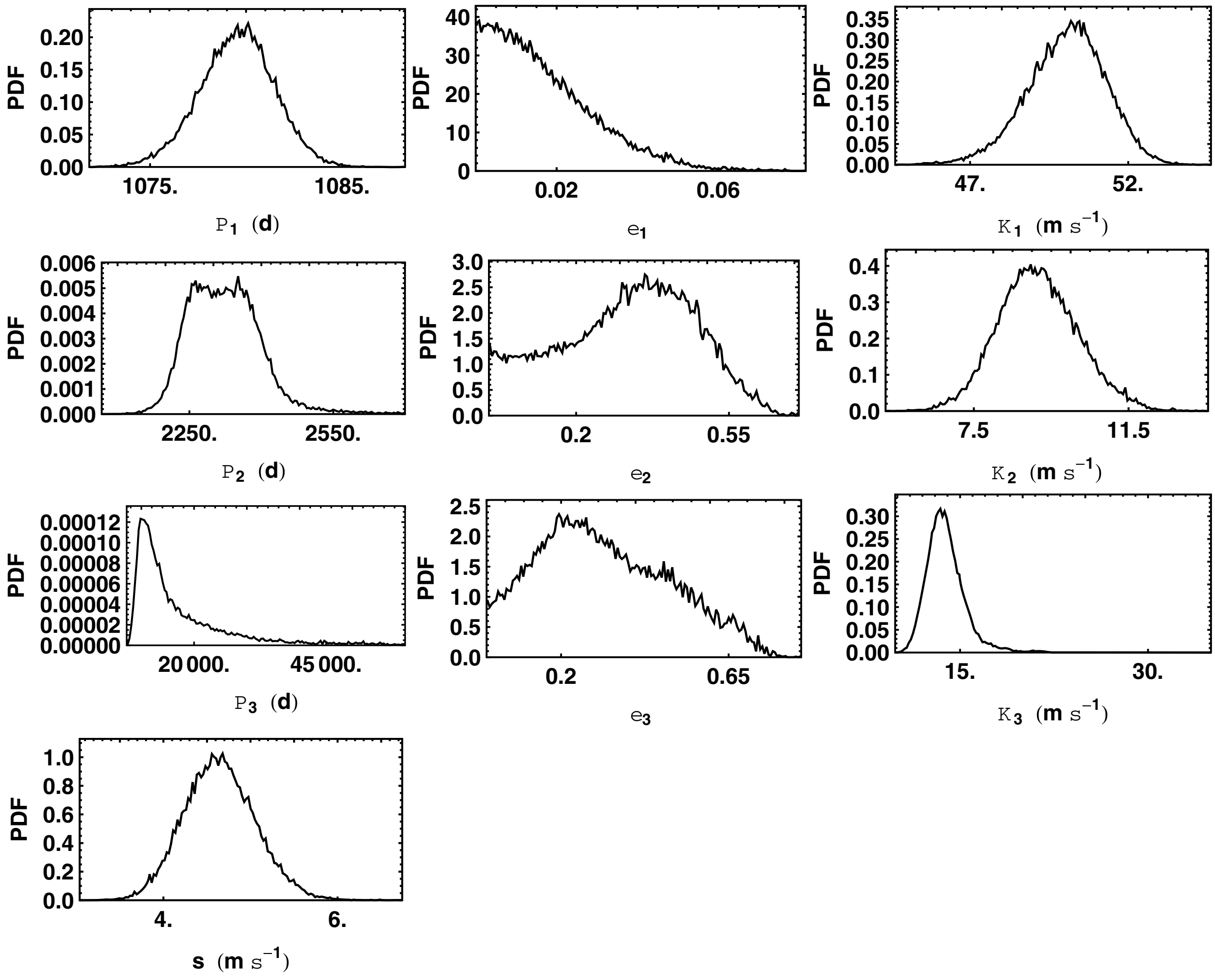

Figure 12. A plot of parameter marginal distributions for a three-planet HMCMC (Case A).

Table 4. Three-planet model parameter estimates (Case A).

Parameter	Planet 1	Planet 2	Planet 3
P (d)	$1079.6^{+2.0}_{-1.8}$ (1079.2)	2319^{+63}_{-76} (2278)	13346^{+4030}_{-4940} (21342) mode = 9991
K (m s^{-1})	$50.1^{+1.3}_{-1.2}$ (50.3)	$9.1^{+1.0}_{-1.0}$ (9.6)	$13.7^{+1.3}_{-1.4}$ (13.2)
e	$0.014^{+.008}_{-.014}$ (0.012)	$0.33^{+.2}_{-.17}$ (0.48)	$0.29^{+.21}_{-.21}$ (0.44)
ω (°)	350^{+84}_{-69} (345)	222^{+21}_{-21} (222)	162^{+40}_{-50} (111)
a (au)	$2.10^{+.02}_{-.02}$ (2.10)	$3.50^{+.07}_{-.08}$ (3.46)	$11.3^{+2.2}_{-2.8}$ (15.4)
$M \sin i$ (M_J)	$2.63^{+.09}_{-.07}$ (2.64)	$0.575^{+.052}_{-.056}$ (0.566)	$1.58^{+.17}_{-.18}$ (1.69)
Periastron passage (JD 2440 000)	11967^{+252}_{-202} (11 943)	11914.6^{+166}_{-131} (11 930)	12655^{+5144}_{-4543} (12 047)

region.[4] The value immediately below in parenthesis is the maximum a posteriori (MAP) value, the value at the maximum of the joint posterior probability distribution. It is not uncommon for the MAP value to fall close to the borders of the credible region. In one case, the period of the third planet, the MAP value falls outside the 68.3 per cent credible region which is one reason why we prefer to quote median values as well. The marginal for P_3 is so asymmetric we also give the mode which is 9991 d. The semimajor axis and $M \sin i$ values are derived from the model parameters assuming a stellar mass of $1.063^{-0.022}_{+0.029}\,\mathrm{M}_\odot$ (Takeda et al. 2007). The quoted errors on the semimajor axis and $M \sin i$ include the uncertainty in the stellar mass.

The Gelman & Rubin (1992) statistic is typically used to test for convergence of the parameter distributions. In PT MCMC, new widely separated parameter values are passed up the line to the $\beta = 1$ simulation and are occasionally accepted. Roughly every 100 iterations the $\beta = 1$ simulation accepts a swap proposal from

[4]In practice, the probability density for any parameter is represented by a finite list of values p_i representing the probability in discrete intervals X. A simple way to compute the 68.3 per cent credible region, in the case of a marginal with a single peak, is to sort the p_i values in descending order and then sum the values until they approximate 68 per cent, keeping track of the upper and lower boundaries of this region as the summation proceeds.

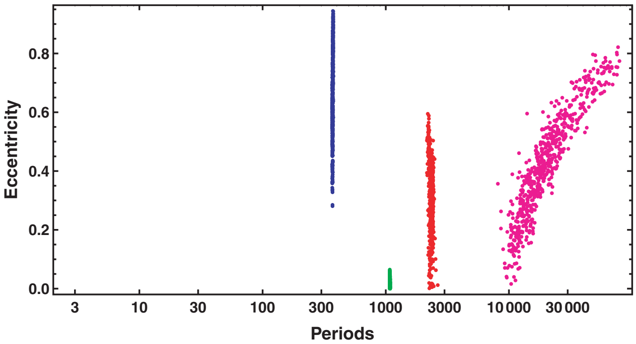

Figure 13. A plot of eccentricity versus period for the four-planet HMCMC (Case A).

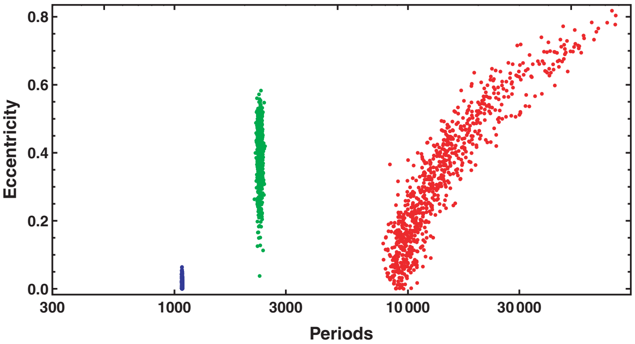

Figure 14. A plot of eccentricity versus period for a three-planet HMCMC fit of the three-planet simulation.

its neighbouring simulation. The final $\beta = 1$ simulation is thus an average of a very large number of independent $\beta = 1$ simulations. What we have done is divide the $\beta = 1$ iterations into 12 equal time intervals and inter-compared the 12 different essentially independent average distributions for each parameter using a Gelman-Rubin test. For all of the three-planet model parameters the Gelman-Rubin statistic was ≤ 1.07.

Fig. 13 shows a plot of eccentricity versus period for a four-planet model. A well-defined fourth period of $370.8^{+2.4}_{-2.0}$ d and eccentricity of $0.57^{+0.22}_{-0.15}$ was detected in repeated HMCMC trials. The amplitude was $K = 5.0^{+1.0}_{-1.1}$ m s^{-1}. The significance of this period is discussed further in Sections 4.2 and 6.

Finally, a five-planet model was also attempted. In addition to the four periods found by the four-planet model, a variety of probability peaks at other periods were observed but none were deemed significant.

4.1.1 Simulation test

As a test of our overall methodology, we simulated data for a three-planet model based on the MAP values from the fit to the real data for the Case A analysis. The data were sampled at the real observation times and had added independent Gaussian noise with a $\sigma = \sqrt{(e_i)^2 + s^2}$, where e_i is the quoted measurement error for the ith point and s, the extra-noise parameter, was 4.4 m s^{-1}. Fig. 14 shows a plot of eccentricity versus period for a sample of the HMCMC parameter samples for the three-planet model fit to the simulated data set. Again, the starting period values for the HMCMC were 5, 20 and 100 d, a long way from the expected values. Comparison with Fig. 11 indicates that the results for the actual data and three-planet simulation are qualitatively very similar.

To test whether the fourth period in the Lick data (period = $370.8^{2.4}_{-2.0}$ d) is a window function artefact of the sampling times, we analysed two three-planet simulations with a four-planet model. In both cases the HMCMC found the three periods expected from the simulation. No well-defined fourth period was found and the peak amplitude was $K = 3$ m s^{-1} compared with a $K = 5$ m s^{-1} for the real data set. This suggests that the fourth period is not simply a window function artefact. However, later HMCMC fits of a combination of Lick and Mcdonald Observatory data did not confirm this period.

4.2 Model selection

One of the great strengths of Bayesian analysis is the built-in Occam's razor. 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 (see for example Gregory 2005a, p. 45). The analysis yields the relative probability of each of the models explored.

To compare the posterior probabilities of the ith planet model to the one-planet models, we need to evaluate the odds ratio, $O_{i1} = p(M_i \mid D, I)/p(M_1 \mid D, I)$, the ratio of the posterior probability of model M_i to model M_1. Application of Bayes's theorem leads to

$$O_{i2} = \frac{p(M_i|I)}{p(M_1|I)} \frac{p(D|M_i, I)}{p(D|M_1, I)} \equiv \frac{p(M_i|I)}{p(M_1|I)} B_{i2} \tag{6}$$

where the first factor is the prior odds ratio and the second factor is called the *Bayes factor*, B_{i2}. The Bayes factor is the ratio of the marginal (global) likelihoods of the models. The marginal likelihood for model M_i is given by

$$p(D|M_i, I) = \int \mathrm{d}\boldsymbol{X}\, p(\boldsymbol{X}|M_i, I) \times p(D|\boldsymbol{X}, M_i, I). \tag{7}$$

Thus, Bayesian model selection relies on the ratio of marginal likelihoods, not maximum likelihoods. The marginal likelihood is the weighted average of the conditional likelihood, weighted by the prior probability distribution of the model parameters and s. This procedure is referred to as marginalization.

The marginal likelihood can be expressed as the product of the maximum-likelihood and Occam penalty (see Gregory & Loredo 1992 and Gregory 2005a, p. 48). The Bayes factor will favour the more complicated model only if the maximum-likelihood ratio is large enough to overcome this penalty. In the simple case of a single parameter with a uniform prior of width ΔX, and a centrally peaked likelihood function with characteristic width δX, the Occam factor is $\approx \delta X/\Delta X$. If the data is useful then generally $\delta X \ll \Delta X$. For a model with m parameters, each parameter will contribute a term to the overall Occam penalty. The Occam penalty depends not only on the number of parameters but also on the prior range of each parameter (prior to the current data set, D), as symbolized in this simplified discussion by ΔX. If two models have some parameters in common then the prior ranges for these parameters will cancel in the calculation of the Bayes factor. To make good use of Bayesian model selection, we need to fully specify priors that are independent of the current data D. The sensitivity of the marginal likelihood to the prior range depends on the shape of the prior and is much greater for a uniform prior than a Jeffreys prior (e.g. see Gregory 2005a, p. 61). In most instances we are not particularly interested in the Occam factor itself, but only in the relative probabilities of the competing models as expressed by the Bayes factors. Because the Occam factor arises automatically in the marginalization procedure, its effect will be present in any model-selection calculation (Note that no Occam factors arise in parameter estimation problems. Parameter estimation can be viewed as model selection where the competing models have the same complexity so the Occam penalties are identical and cancel out.)

The MCMC algorithm produces samples which are in proportion to the posterior probability distribution which is fine for parameter estimation but one needs the proportionality constant for estimating the model marginal likelihood. Clyde (2007) recently reviewed the state of techniques for model selection from a statistics perspective and Ford & Gregory (2007) have evaluated the performance of a variety of marginal-likelihood estimators in the extrasolar planet context.

Gregory (2007a), in the analysis of velocity data for HD 11964, compared the results from three marginal-likelihood estimators: (a) PT, (b) ratio estimator and (c) restricted MC (RMC). MC integration can be very inefficient in exploring the whole prior parameter range because it randomly samples the whole volume. The fraction of the prior volume of parameter space containing significant probability rapidly declines as the number of dimensions increases. For example, if the fractional volume with significant probability is 0.1 in one dimension then in 17 dimensions the fraction might be of the order of 10^{-17}. In RMC integration, this should be much less of a problem because the volume of parameter space sampled is restricted to a region delineated by the outer borders of the marginal distributions of the parameters. For HD 11964, the three methods were compared for one-, two- and three-planet models. For the one-planet model, all three methods agreed within 15 per cent. For the two-planet model, the three methods agreed within 28 per cent with the RMC giving the lowest estimate. For the three-planet model, the estimates were very different. The RMC estimate was 16 times smaller than the PT estimate and the ratio estimator was 18 times larger than the PT estimate. The PT method is very compute intensive. For a three-planet model, 40 tempering levels and 10^7 iterations were required. The problem becomes more difficult for larger numbers of planets. Thus for three or more planet models, accurately computing the marginal likelihood is a very big challenge.

In this work, we consider only RMC marginal-likelihood estimates. This method is expected to underestimate the marginal likelihood in higher dimensions, and this underestimate is expected to become worse the larger the number of model parameters, i.e. increasing number of planets. When we conclude, as we do, that the RMC computed odds in favour of the three-planet model compared to the two-planet model is $\sim 10^{17}$, we mean that the true odds is $\geq 10^{17}$.

In earlier work, we defined the outer boundary of parameter space for RMC integration based on the 99 per cent credible region. One problem is that if there is a significant contribution to the integral within say the 30 per cent credible region, the volume in this region can be such a small fraction of the total that no random sample lands in that region. In this work, we use a nested version of RMC integration. Multiple boundaries were constructed based on credible regions ranging from 30 per cent to $\geq$ 99 per cent, as needed. We are then able to compute the contribution to the total RMC integral from each nested interval and sum these contributions. For example, for the interval between the 30 and 60 per cent credible regions, we generate random parameter samples within the 60 per cent region and reject any sample that falls within the 30 per cent region. Using the remaining samples, we can compute the contribution to the RMC integral from that interval.

The left-hand panel of Fig. 15 shows the contributions from the individual intervals for five repeats of the nested RMC evaluation for the two-planet model. The right-hand panel shows the summation of the individual contributions versus the volume of the credible region. The credible region listed as 9995 per cent is defined as follows. Let X_{U99} and X_{L99} correspond to the upper and lower boundaries of the 99 per cent credible region, respectively, for any of the parameters. Similarly, X_{U95} and X_{L95} are the upper and lower boundaries of the 95 per cent credible region for the parameter. Then $X_{U9995} = X_{U99} + (X_{U99} - X_{U95})$ and $X_{L9995} = X_{L99} + (X_{L99} - X_{L95})$. Similarly, $X_{U9984} = X_{U99} + (X_{U99} - X_{U84})$.

Table 5 gives the marginal-likelihood estimates, Bayes factors and false alarm probabilities (FAPs) for zero-, one-, two-, three- and four-planet models which are designated $M_0, \ldots, M_4$. The last two columns list the MAP value of extra-noise parameter, s, and the RMS residual. For each model the RMC calculation was repeated five times and the quoted errors give the spread in the results, not the standard deviation. The Bayes factors that appear in the third column are all calculated relative to Model 2. Examination of a plot like the one shown in Fig. 15, but for the four-planet model, indicates that RMC is probably seriously underestimating the marginal likelihood. A better method of computing this quantity is sorely needed.

We can readily convert the Bayes factors to a Bayesian FAP. For example, in the context of claiming the detection of a three-planet model the FAP is the probability that there are actually two or less

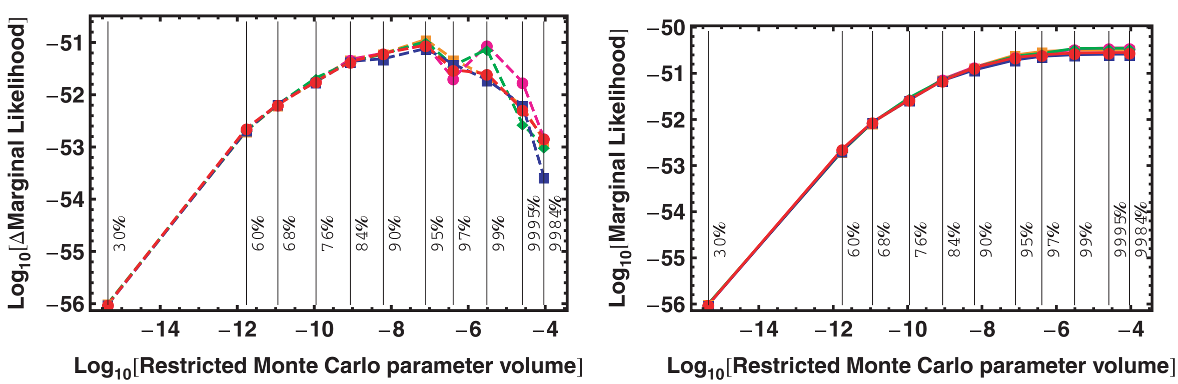

Figure 15. Left-hand panel shows the contribution of the individual nested intervals to the RMC marginal likelihood for the two-planet model. The right-hand panel shows the integral of these contributions versus the parameter volume of the credible region. Note that only the relative values of the units on the vertical axes of these two plots are meaningful.

Table 5. Marginal-likelihood estimates, Bayes factors and FAPs for (Case A) zero-, one-, two-, three- and four-planet models which are designated $M_0, \ldots, M_4$. The last two columns list the MAP value of extra-noise parameter, s, and the rms residual.

Model	Periods (d)	Marginal likelihood	Bayes factor nominal	FAP	s (m s^{-1})	rms residual (m s^{-1})
M_0		2.63×10^{-481}	10^{-127}		34.8	35.3
M_1	(1080)	$(7.51 \pm 0.07) \times 10^{-394}$	10^{-39}	10^{-88}	11.2	12.5
M_2	(1080, 8000)	$(4.1 \pm 0.5) \times 10^{-355}$	1.0	10^{-39}	6.1	8.1
M_3	(1080, 2300, ∼10000)	$(4^{\times 2}_{\times 1/5}) \times 10^{-338}$	10^{17}	10^{-17}	4.4	6.5
M_4	(371, 1080, 2300, ∼10000)	$(4^{\times 7}_{\times 1/2}) \times 10^{-338}$	10^{17}	0.5	3.7	6.1

planets

$$\mathrm{FAP} = \sum_{i=0}^{2} (\text{prob. of } i \text{ planets}). \tag{8}$$

If we assume a priori (absence of the data) that the prob of one-planet model = prob. of two-planet model, etc., then probability of each model is related to the Bayes factors by

$$p(M_i \mid D, I) = \frac{B_{i2}}{\sum_{j=0}^{N_{\rm mod}} B_{j2}}, \tag{9}$$

where $N_{\rm mod}$ is the total number of models considered, and of course $B_{22} = 1$. Given the Bayes factors in Table 5 and substituting into equation (8) gives

$$\mathrm{FAP} = \frac{(B_{02} + B_{12} + B_{22})}{\sum_{j=0}^{3} B_{j2}} \approx 10^{-17}. \tag{10}$$

For the three-planet model, we obtain a very low FAP $\approx 10^{-17}$. The Bayesian FAPs for one-, two-, three- and four-planet models are given in the fourth column of Table 5.

In the context of claiming the detection of a four-planet model, the FAP is ≈0.5. This is very high and does not justify a claim for the detection of a fourth planet. The fourth period is also suspiciously close to 1 yr to be of concern.

5 RESULTS (CASE B)

For Case B, we incorporated four additional parameters to allow for the unknown residual velocity offsets of dewars 6, 8, 39 and 18 relative to dewar 24. These are labelled V_6, V_8, V_{39} and V_{18}, where the subscript denotes the detector dewar. In a Bayesian analysis, these are treated as additional nuisance parameters which we can marginalize. Additionally, since they are of interest to the observers we also provide a summary of each residual offset parameter. In the RV data processing pipeline, every effort was made to ensure that the dewar velocity offsets were allowed for, so the residuals are expected to be small. For the Case B analysis, we have assumed a Gaussian prior for each V_i centred on zero with a $\sigma = 3$ km s^{-1}.

5.1 Parameter estimation

In this section, we redo the analysis of the 47 UMa data with the multiplanet HMCMC Kepler periodogram starting with a one-planet model and extending to a four-planet model. The data is same as shown in Fig. 6, panel (a), with the exception of the first point corresponding to dewar 1.

Fig. 16 shows a plot of eccentricity versus period for a sample of the HMCMC parameter samples for the two-planet model for Case B. The two-planet model again favours a second period in the range 8100–15 000 d (68 per cent credible region) with a long

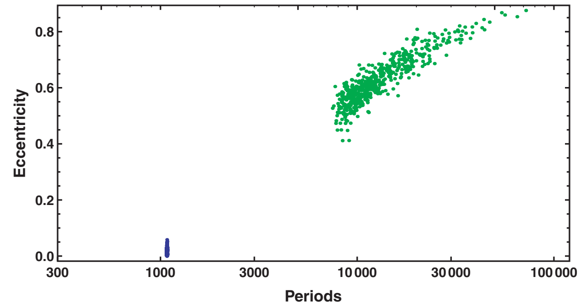

Figure 16. A plot of eccentricity versus period for the two-planet fit (Case B).

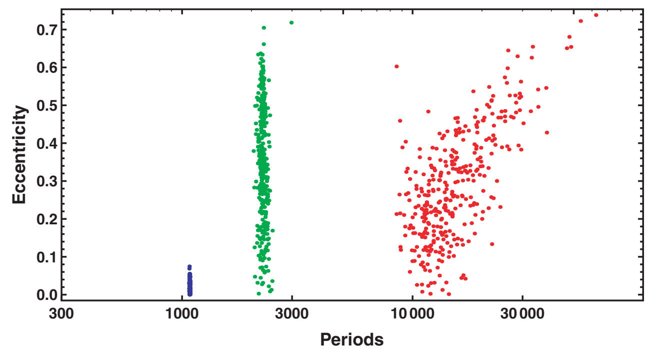

Figure 17. A plot of eccentricity versus period for the three-planet HMCMC (Case B).

higher eccentricity tail extending to much longer periods. In Case B, the time base is 235 d shorter than Case A so the lower eccentricity/lower period end is less well defined, but otherwise there is general agreement. This model was run twice using different starting periods, but the two-planet HMCMC run did not favour a period around 2240 d even when the two starting periods used were 1078 and 2240 d, respectively. This is not that surprising given the relative sizes of the K values for planets 2 and 3 in Table 4.

Fig. 17 shows a plot of eccentricity versus period for a sample of the HMCMC parameter samples for the three-planet model for Case B. Again, we see the emergence of a period of ∼2250 d, and the third longer period appears better defined (compared to the two-planet model) and extends to much lower eccentricities. Qualitatively, there is general agreement with the Case A results shown in Fig. 11. The dewar residual offset velocities were $V_6 = 0.07^{+2.7}_{-2.6}$, $V_8 = 1.7^{+3.0}_{-2.3}$, $V_{39} = -3.2^{+2.5}_{-2.4}$ and $V_{18} = -1.1^{+2.0}_{-2.0}$ m s^{-1}.

The four-planet HMCMC analysis again showed a clear fourth period of $372^{1.9}_{-1.3}$ with an eccentricity of 0.73 ± 0.14. We did not compute the marginal likelihood for the four-planet model, but based on the Case A results the FAP for a four-planet model is expected to be very high.

5.2 Model selection (Case B)

We repeated the FAP for the three-planet model as described in Section 4.2 for the Case B analysis which incorporates the dewar residual offset parameters

$$\mathrm{FAP} = \frac{(B_{02} + B_{12} + B_{22})}{\sum_{j=0}^{3} B_{j2}}. \quad (11)$$

The computed Bayes factors are $B_{02} = 1.6 \times 10^{-141}$, $B_{12} = 2.0 \times 10^{-28}$, $B_{22} = 1.0$ and $B_{32} = 2.0 \times 10^{5}$. This gives a FAP of 5.0×10^{-6}. Even though this is much greater than the value found in Case A, it still argues strongly for favouring a three-planet model.

6 DISCUSSION

On the basis of the model-selection results, we can conclude that there is strong evidence for three planets although the longest period orbital parameters are still not well defined. The results for the Lick only analysis do not rule out low-eccentricity orbits for all three planets. The major difference produced by including the dewar residual offset parameters was to reduce the FAP for a three-planet model from $\sim 10^{-17}$ to $\sim 10^{-5}$. A significant part of this reduction might be a consequence of the reduced span of the data set by 235 d for the Case B analysis.

Our results appear to be entirely consistent with the latest analysis of Wittenmyer et al. (2009). Their best-fitting two-planet model now calls for $P_2 = 9660$ d. They note that to fit a second planet, they fixed the parameters e_2 and ω_2 at the values proposed by Fischer et al. (2002): $e_2 = 0.005$ and $\omega_2 = 127$. In our Case A two-planet fit (Fig. 8), in which all parameters were free, the eccentricity versus period plot exhibits a low-eccentricity tail which occurs at a period between 9000 and 10 000 d, directly comparable to their 9660 d period. The $\sim$2300 d period in the Lick data only shows up in our three-planet and higher models. This is probably because the longer period signal with a $K = 13.8$ m s^{-1} dominates over the 2300 d period signal with a $K = 8.0$ m s^{-1} (see Table 6). Wittenmyer et al. (2009) did not report any results on fitting a three-planet model.

To test this further, we combined the Lick data with the Wittenmyer et al. (2009) data from the 9.2-m HET and 2.7-m HJS telescopes of the McDonald Observatory. We subtracted initial offset velocities of 23.3 and 25.4 m s^{-1} based on a comparison of plots of the HET and HJS data sets to the Lick data. We then included a free parameter for each telescope to allow for an unknown residual velocity offset compared with the Lick dewar 24 in the same way as we had done for the other Lick dewars in Case B.

Fig. 18 shows a plot of eccentricity versus period for our three-planet HMCMC fit to the combined data set. The three starting periods used for the HMCMC run were 10, 1078, and 6000 d. The residual velocity offset parameters for the HET and HJS telescopes were $1.5^{+1.0}_{-1.1}$ and -0.2 ± 1 m s^{-1}, respectively. It is clear from the figure that the same three periods appear as before, but with the extra data the results now favour low-eccentricity orbits for all three periods. This is a particularly pleasing result as low-eccentricity orbits are more likely to exhibit long-term stability than high-eccentricity orbits. The preference for low-eccentricity orbits is more apparent in the marginal distributions shown in Fig. 19.

Our final orbital parameters are summarized in Table 6 together with the residual offset velocities and the extra-noise term s. Again, the parameter value listed is the median of the marginal probability distribution for the parameter in question and the error bars identify the boundaries of the 68.3 per cent credible region. The value immediately below in parenthesis is the MAP value, the value at the maximum of the joint posterior probability distribution.

The final period phase plots are shown in Fig. 20. The top-left panel shows the data and model fit versus 1078 d orbital phase after removing the effects of the two other orbital periods. The red and green curves are the mean HMCMC model fit +1 standard deviation and mean model fit −1 standard deviation, respectively. The dashed curve is the mean HMCMC fit. The other two panels correspond to phase plot for the other two periods. In each panel, the quoted period is the mode of the marginal distribution. The P_2 and P_3 phase coverage for the combined HET and HJS data (not shown) is not sufficient to warrant a fully independent search for these two periods.

Table 6. Final three-planet model parameter estimates from the HMCMC fit of the combined Lick, HET and HJS telescope data set.

Parameter	Planet 1	Planet 2	Planet 3
P (d)	1078^{+2}_{-2} (1078)	2391^{+100}_{-87} (2430)	14002^{+4018}_{-5095} (47 831) mode = 11251
K (m s^{-1})	$48.4^{+0.8}_{-0.9}$ (48.2)	$8.0^{+1.0}_{-1.0}$ (8.3)	$13.8^{+2.2}_{-2.9}$ (13.5)
e	$0.032^{+.014}_{-.014}$ (0.038)	$0.098^{+0.047}_{-0.096}$ (0.020)	$0.16^{+.09}_{-.16}$ (0.67)
ω (°)	334^{+23}_{-23} (324)	295^{+114}_{-160} (356)	110^{+132}_{-160} (110)
a (au)	$2.100^{+.02}_{-.02}$ (2.10)	$3.6^{+.1}_{-.1}$ (3.6)	$11.6^{+2.1}_{-2.9}$ (26.3)
$M \sin i$ (M_J)	$2.53^{+.07}_{-.06}$ (2.53)	$0.540^{+.066}_{-.073}$ (0.567)	$1.64^{+.29}_{-0.48}$ (1.86)
Periastron passage (JD 244 0000)	11917^{+63}_{-76} (11888)	12441^{+628}_{-825} (12778)	11736^{+6783}_{-5051} (11736)
V_6 (m s^{-1})	$1.1^{+2.8}_{-2.9}$ (4.0)	V_8 (m s^{-1})	$-0.6^{+2.6}_{-2.6}$ (1.0)
V_{39} (m s^{-1})	$-5.0^{+2.8}_{-2.7}$ (-0.5)	V_{18} (m s^{-1})	$-5.1^{+1.7}_{-1.6}$ (-4.6)
V_{HET} (m s^{-1})	$1.5^{+1.0}_{-1.1}$ (1.3)	V_{HJS} (m s^{-1})	$-0.2^{+1.0}_{-1.0}$ (0.1)
s (m s^{-1})	$5.7^{+0.3}_{-0.3}$ (5.3)		

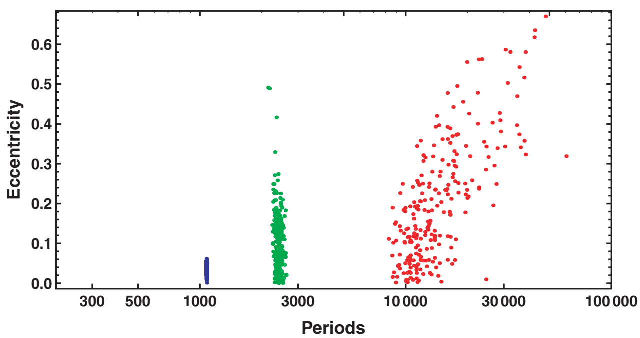

Figure 18. A plot of eccentricity versus period for a three-planet HMCMC fit of the combined Lick, HET and HJS telescope data set.

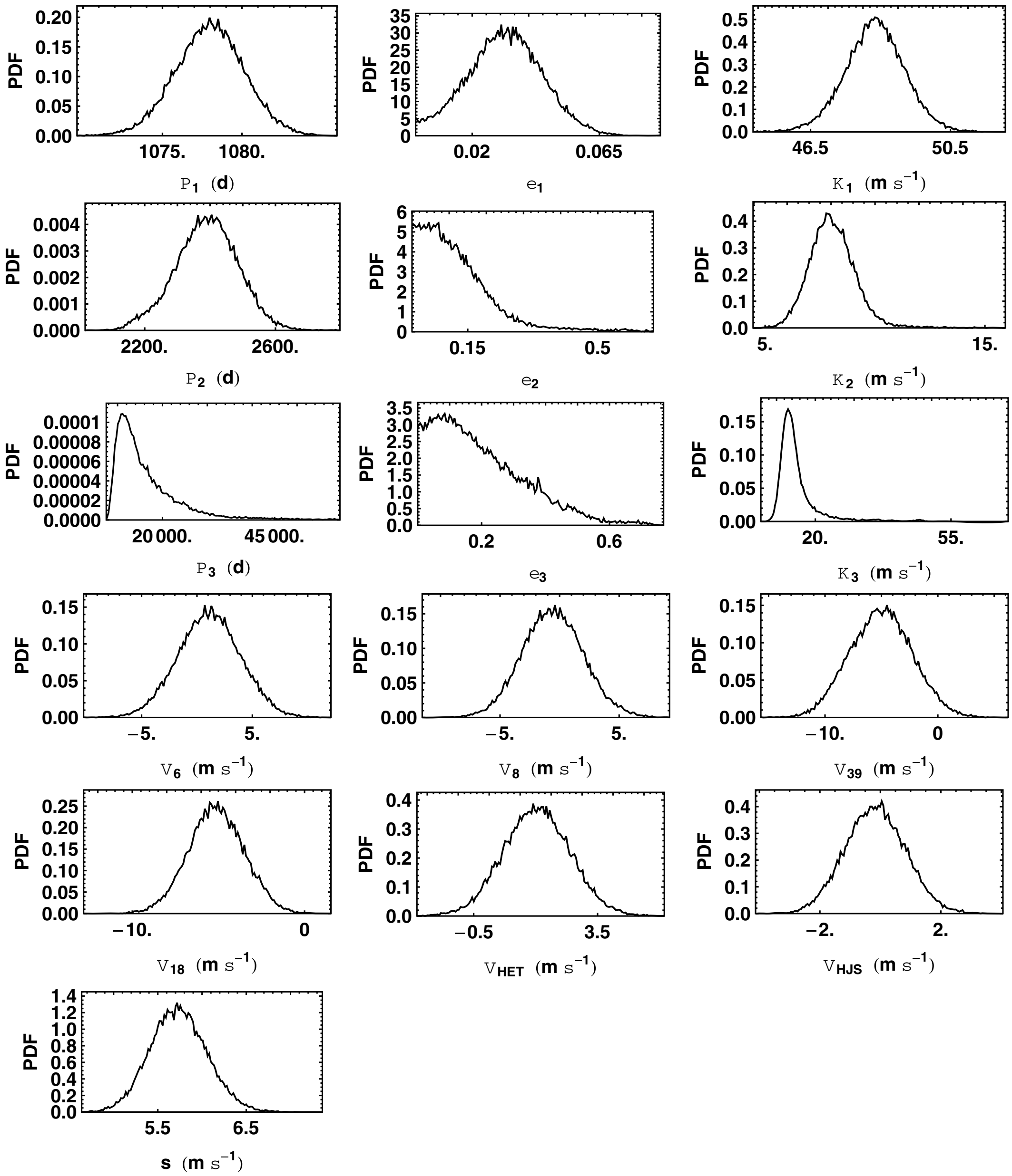

Figure 19. A plot of parameter marginal distributions for a three-planet HMCMC of the combined Lick, HET and HJS telescope data set. The residual offset velocity parameters are relative to the Lick dewar 24. They are designated V_j, where $j = 6, 8, 39, 18$ correspond to the other Lick dewars and subscripts HET and HJS refer to the HET and HJS telescopes (Wittenmyer et al. 2009).

HMCMC fits of a four-planet model to the combined Lick, HET and HJS data set failed to detect a well-defined fourth period casting doubt on the validity of the $370.8^{+2.4}_{-2.0}$ d period detected in the Lick only data. Even though this period was well defined in the Lick only data, the FAP of $\approx$0.5 is much too high to warrant any claim of significance. The period is also suspiciously close to 1 yr and might be an artefact of the data reduction.

6.1 Eccentricity bias

In Section 3.1, we showed that HMCMC periodogram peaks exhibit a well-defined statistical bias towards high eccentricity in the absence of a real periodic signal. To mimic a circular velocity orbit, the noise points need to be correlated over a larger fraction of the orbit than they do to mimic a highly eccentric orbit. For this reason, it is more likely that noise will give rise to spurious highly eccentric orbits than low-eccentricity orbits. Is there a similar or stronger bias when there is a real periodic signal? Based on the above explanation of the bias, we would expect noise to conspire to increase the eccentricity of detected periodogram peaks associated with the real periodic signals. Our expectation is that the importance of this bias will be dependent on the strength of the signal and possibly on the number of observed periods.[5] For very strong signals like the 1078 d period, we would expect the bias to be very small. For very weak signals, the bias might well be approximated by the no real periodic signal eccentricity bias which we quantified earlier. As we have seen, in the case of the 47 UMa ∼2300 d period, the Lick

[5] This will be the subject of a future investigation.

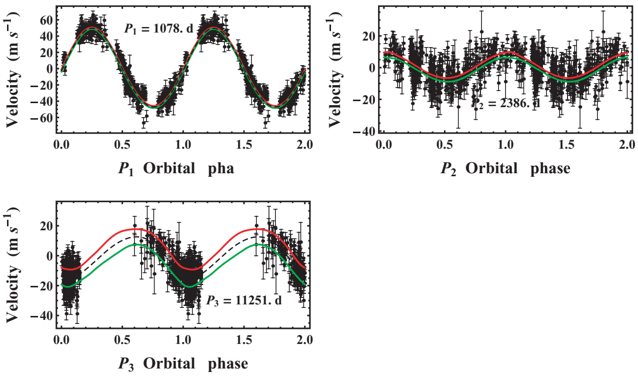

Figure 20. The top-left panel shows the data and model fit versus 1078 d orbital phase after removing the effects of the two other orbital periods. The red and green curves are the mean HMCMC model fit +1 standard deviation and mean model fit −1 standard deviation, respectively. The dashed curve is the mean HMCMC fit. The other two panels correspond to phase plot for the other two periods.

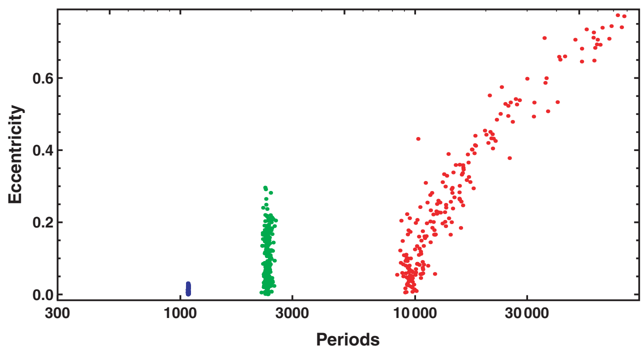

Figure 21. A plot of eccentricity versus period for the three-planet HMCMC fit of the three-planet simulation.

data alone favour an eccentricity of ≈0.3, even when we include the eccentricity bias filter. When we added more data, the eccentricity was noticeably reduced. What if we simulated a Lick only data set for a three-planet model based on the MAP three-planet model parameters for the combined Lick, HET and HJS analysis. Would the HMCMC analysis of the simulated data favour higher eccentricities, possibly indicating that there is some additional eccentricity bias operating. To test for this, we carried out this simulation but modified the MAP parameter values so all three eccentricities were identically zero and $P_3 = 10\,000$ d. Also, no residual offsets were included for this test so the analysis corresponds to Case A.

Fig. 21 shows a plot of eccentricity versus period for the simulation. The starting period values for the HMCMC were 5, 20 and 100 d, a long way from the expected values. Again, all three simulated periods were detected and the preferred eccentricities are all close to zero but with significant tails extending to higher eccentricity. Based on this test, there does not appear to be any clear additional eccentricity bias operating. The fact that the real Lick data alone favour (in Case A and Case B) somewhat larger eccentricities for P_2 and P_3 suggests there may be something else present in the real data, possiblly some low-level systematic effect or other real signals. In this regard, the eccentricity of the longer period was considerably higher in the two-planet models than when allowance was made for an additional period in the three-planet models.

7 CONCLUSIONS

In this paper, we have demonstrated that a Bayesian adaptive HMCMC analysis of a challenging data set has helped clarify the number of planets present in 47UMa. HMCMC integrates the advantages of PT, simulated annealing and the genetic algorithm. Each of these techniques was designed to facilitate the detection of a global minimum in χ^2. Combining all three in an adaptive HMCMC greatly increases the probability of realizing this goal. The adaptive Bayesian HMCMC is very general and can be applied to many different non-linear modelling problems. It has been implemented in GRIDMATHEMATICA on an 8 core PC. The increase in a speed for the parallel implementation is a factor of 6.6. When applied to the Kepler problem, it corresponds to a multiplanet Kepler periodogram which is ideally suited for detecting signals that are consistent with Kepler's laws. However, it is more than a periodogram because it also provides full marginal posterior distributions for all the orbital parameters that can be extracted from RV data. The execution time for a one-planet blind fit (seven parameters) is 10^6 iterations per hour. The program scales linearly with the number of parameters being fit.

The 47UMa data has been analysed using one-, two-, three-, four- and five-planet models. On the basis of the model-selection results, we can conclude there is strong evidence for three planets based on an FAP of 5.0×10^{-6}, however, the longest period orbital parameters are still not well defined. The measured periods (based on the combined data set) are 1078 ± 2, 2391^{+100}_{-87} and 14002^{+4018}_{-5095}d, and the corresponding eccentricities are 0.032 ± 0.014, $0.098^{+.047}_{-.096}$ and $0.16^{+.09}_{-.16}$. The results favour low-eccentricity orbits for all three. Note that the longer time base of the full Lick data set favours a value for P_3 at the lower end of the 68 per cent credible region of ∼10 000 d. Assuming the three signals (each one consistent with a Keplerian orbit) are caused by planets, the corresponding limits on planetary mass ($M \sin i$) and semimajor axis are ($2.53^{+.07}_{-.06} M_J$, 2.10 ± 0.02 au), ($0.54 \pm 0.07 M_J$, 3.6 ± 0.1 au), and ($1.6^{+0.3}_{-0.5} M_J$, $11.6^{+2.1}_{-2.9}$ au), respectively. Based on our three-planet model results, the remaining unaccounted for noise (stellar jitter) is 5.7 m s^{-1}.

A four-planet model fit to the Lick data yielded a well-defined fourth period of $370.8^{+2.4}_{-2.0}$ d and eccentricity of $0.57^{+0.22}_{-0.15}$, but the combined data set did not yield a well-defined fourth period. Even though this period was well defined in the Lick only data, the FAP of ≈0.5 is much too high to warrant any claim of significance. The period is also suspiciously close to 1 yr and might be an artefact of the data reduction.

The velocities of model fit residuals were randomized in multiple trials and processed using a one-planet version of the HMCMC Kepler periodogram. In this situation, periodogram probability peaks are purely the result of the effective noise. The orbits corresponding to these noise-induced periodogram peaks exhibit a well-defined statistical bias towards high eccentricity. We have characterized this eccentricity bias and designed a correction filter that can be used as an alternate prior for eccentricity to enhance the detection of planetary orbits of low or moderate eccentricity. On the basis of our understanding of the mechanism underlying the eccentricity bias, we expect the eccentricity prior filter to be generally applicable to searches for low-amplitude orbital signals in other precision RV data sets.

ACKNOWLEDGMENTS

One of us (Gregory) would like to thank Wolfram Research for providing a complementary license to run GRIDMATHEMATICA that was used in this research. DAF acknowledges research support from NASA grant NNX08AF42G. The authors thank Matt Giguere for calculating dewar offsets and the many student observers who have contributed to the 47 UMa data set over the years. They particularly thank Geoff Marcy and R. Paul Butler who began the Lick planet survey in 1989.

REFERENCES

Bretthorst G. L., 1988, Bayesian Spectrum Analysis and Parameter Estimation. Springer-Verlag, New York
Butler R. P., Marcy G. W., 1996, ApJ, 464, L153
Clyde M. A., Berger J. O., Bullard F., Ford E. B., Jefferys W. H., Luo R., Paulo R., Loredo T., 2007, in Babu G. J., Feigelson E. D., eds, ASP Conf. Ser. Vol. 371, Statistical Challenges in Modern Astronomy IV. Astron. Soc. Pac., San Francisco, p. 224
Cumming A., 2004, MNRAS, 354, 1165
Cumming A., Dragomir D., 2010, MNRAS, 401, 1029
Fischer D. A., Marcy G. W., Butler R. P., Laughlin G. L., Vogt S. S., 2002, ApJ, 564, 1028
Ford E. B., 2005, AJ, 129, 1706
Ford E. B., 2006, ApJ, 620, 481
Ford E. B., Gregory P. C., 2007, in Babu G. J., Feigelson E. D., eds, ASP Conf. Ser. Vol. 371, Statistical Challenges in Modern Astronomy IV. Astron. Soc. Pac., San Francisco, p. 189
Gelman A., Rubin D. B., 1992, Statistical Sci., 7, 457
Gregory P. C., 2005a, Bayesian Logical Data Analysis for the Physical Sciences: A Comparative Approach with Mathematica Support. Cambridge Univ. Press, Cambridge
Gregory P. C., 2005b, ApJ, 631, 1198
Gregory P. C., 2005c, in Abbas A. E., Morris R. D., Castle J. P., eds, AIP Conf. Proc. Vol. 803, Bayesian Inference and Maximum Entropy Methods in Science and Engineering. Am. Inst. Phys., New York, p. 139
Gregory P. C., 2007a, MNRAS, 374, 1321
Gregory P. C., 2007b, in Knuth K. H., Caticha A., Center J. L., Giffin A., Rodríguez C. C., eds, AIP Conf. Proc. Vol. 954, Bayesian Inference and Maximum Entropy Methods in Science and Engineering. Am. Inst. Phys., New York, p. 307
Gregory P. C., 2007c, MNRAS, 381, 1607
Gregory P. C., 2009, JSM Proc., Denver. American Statistical Association (arXiv:0902.2014v1)
Gregory P. C., Loredo T. J., 1992, ApJ, 398, 146
Jaynes E. T., 1957, Stanford University Microwave Laboratory Report 421. Reprinted in 1988 in Erickson G. J., Smith C. R., eds, Maximum Entropy and Bayesian Methods in Science and Engineering. Kluwer, Dordrecht, p. 1
Jaynes E. T., 1987, in Smith C. R., Erickson G. L., eds, Maximum Entropy and Bayesian Spectral Analysis and Estimation Problems. Reidel, Dordrecht, p. 1
Loredo T., 2004, in Erickson G. J., Zhai Y., eds, AIP Conf. Proc. Vol. 707, Bayesian Inference and Maximum Entropy Methods in Science and Engineering. Am. Inst. Phys., New York, p. 330
Loredo T. L., Chernoff D., 2003, in Feigelson E. D., Babu G. J., eds, Statistical Challenges in Modern Astronomy III. Springer, New York, p. 57
Naef D., Mayor M., Beuzit J. L., Perrier C., Queloz D., Sivan J. P., Udry S., 2004, A&A, 414, 351
Roberts G. O., Gelman A., Gilks W. R., 1997, Ann. Applied Probability, 7, 110
Saar S. H., Donahue R. A., 1997, ApJ, 485, 319
Saar S. H., Butler R. P., Marcy G. W., 1998, ApJ, 498, L153
Takeda G., Ford E. B., Sills A., Rasio F. A., Fischer D. A., Valenti J. A., 2007, ApJS, 168, 297
Udry S. et al., 2007, A&A, 469, L43
Wittenmyer R., Endl M., Cochran W., 2007, ApJ, 654, 625
Wittenmyer R. A., Endl M., Cochran W. D., Levison H. F., Henry G. W., 2009, ApJS, 182, 97
Wright J. T., 2005, PASP, 117, 657

This paper has been typeset from a TEX/LATEX file prepared by the author.