A&A 551, A79 (2013)
DOI: 10.1051/0004-6361/201220509
© ESO 2013

Astronomy & Astrophysics

Signals embedded in the radial velocity noise

Periodic variations in the τ Ceti velocities*

M. Tuomi[1,2], H. R. A. Jones[1], J. S. Jenkins[3], C. G. Tinney[4,5], R. P. Butler[6], S. S. Vogt[7], J. R. Barnes[1], R. A. Wittenmyer[4,5], S. O'Toole[7], J. Horner[4,5], J. Bailey[4], B. D. Carter[3], D. J. Wright[4,5], G. S. Salter[4,5], and D. Pinfield[1]

[1] University of Hertfordshire, Centre for Astrophysics Research, Science and Technology Research Institute, College Lane, AL10 9AB, Hatfield, UK
e-mail: mikko.tuomi@utu.fi; m.tuomi@herts.ac.uk
[2] University of Turku, Tuorla Observatory, Department of Physics and Astronomy, Väisäläntie 20, 21500 Piikkiö, Finland
[3] Departamento de Astronomía, Universidad de Chile, Camino del Observatorio 1515, Las Condes, Santiago, Chile
[4] School of Physics, University of New South Wales, 2052 Sydney, Australia
[5] Australian Centre for Astrobiology, University of New South Wales, 2052 Sydney, Australia
[6] Department of Terrestrial Magnetism, Carnegie Institute of Washington, Washington, DC 20015, USA
[7] UCO/Lick Observatory, Department of Astronomy and Astrophysics, University of California at Santa Cruz, Santa Cruz, CA 95064, USA

Received 5 October 2012 / Accepted 13 December 2012

ABSTRACT

Context. The abilities of radial velocity exoplanet surveys to detect the lowest-mass extra-solar planets are currently limited by a combination of instrument precision, lack of data, and "jitter". Jitter is a general term for any unknown features in the noise, and reflects a lack of detailed knowledge of stellar physics (asteroseismology, starspots, magnetic cycles, granulation, and other stellar surface phenomena), as well as the possible underestimation of instrument noise.

Aims. We study an extensive set of radial velocities for the star HD 10700 (τ Ceti) to determine the properties of the jitter arising from stellar surface inhomogeneities, activity, and telescope-instrument systems, and perform a comprehensive search for planetary signals in the radial velocities.

Methods. We performed Bayesian comparisons of statistical models describing the radial velocity data to quantify the number of significant signals and the magnitude and properties of the excess noise in the data. We reached our goal by adding artificial signals to the "flat" radial velocity data of HD 10700 and by seeing which one of our statistical noise models receives the greatest posterior probabilities while still being able to extract the artificial signals correctly from the data. We utilised various noise components to assess properties of the noise in the data and analyse the HARPS, AAPS, and HIRES data for HD 10700 to quantify these properties and search for previously unknown low-amplitude Keplerian signals.

Results. According to our analyses, moving average components with an exponential decay with a timescale from a few hours to few days, and Gaussian white noise explains the jitter the best for all three data sets. Fitting the corresponding noise parameters results in significant improvements of the statistical models and enables the detection of very weak signals with amplitudes below 1 m s^{-1} level in our numerical experiments. We detect significant periodicities that have no activity-induced counterparts in the combined radial velocities. Three of these signals can be seen in the HARPS data alone, and a further two can be inferred by utilising the AAPS and Keck data. These periodicities could be interpreted as corresponding to planets on dynamically stable close-circular orbits with periods of 13.9, 35.4, 94, 168, and 640 days and minimum masses of 2.0, 3.1, 3.6, 4.3, and 6.6 $M_\oplus$, respectively.

Key words. methods: statistical – methods: numerical – techniques: radial velocities – stars: individual: HD 10700

1. Introduction

The improving instrumental precision and the rapidly increasing number of measurements of radial velocity (RV) target stars of different surveys are enabling the discoveries of smaller planets on longer orbits in increasing numbers (e.g. Lovis et al. 2011; Mayor et al. 2009, 2011). Currently, however, the stellar noise, sometimes referred to as stellar jitter (e.g. Wright 2005), presents the greatest obstacle in reaching the ultimate goal of being able to detect Earth-mass planets in stellar habitable zones (Barnes et al. 2012). While the magnitude of this jitter is still very uncertain for all the target stars, its other properties, such as shape of the noise distribution, its variability as a function of time, and dependence on the stellar properties have not received much attention.

When analysing RV data, the common strategy is to bin the measurements made within, say, intervals of an hour, and use the resulting binned data as a basis of further statistical analyses. The motivation for this operation is the possibility of reducing the amount of short-term noise in the data (e.g. O'Toole et al. 2009). While this approach certainly mitigates against asteroseismological signals, it results in a loss of information about the properties of stellar noise and may also prevent the detection of the faintest signals in the data because it corresponds to artificially decreasing the number of measurements. In general, RV signals are published in the literature as received from binned data but with no information about the binning techniques used,

* Radial velocities are only available in electronic form at the CDS via anonymous ftp to cdsarc.u-strasbg.fr (130.79.128.5) or via http://cdsarc.u-strasbg.fr/viz-bin/qcat?J/A+A/551/A79

i.e. binning time-scale and uncertainty estimation. To circumvent this approach and its possible shortcomings, we study the properties of the jitter by comparing different noise models: different autoregressive (AR) and moving average (MA) models, their combinations (ARMA), and different additive white noise models.

Because it is not possible to generate artificial RV data with realistic noise properties (for the obvious reason that these properties are not known), we study a quiet star with a large amount of data and no published signals. We analyse the RVs of the nearby Solar-type star HD 10700 (τ Ceti). It is known to be a very inactive and quiescent star and its long high-precision RV data set from the High Accuracy Radial Velocity Planet Searcher (HARPS) spectrograph (Mayor et al. 2003) has not been reported to contain planetary signatures despite more than 4000 spectral observations (Pepe et al. 2011). Also, there are two other large RV data sets of HD 10700 available from the High Resolution Echelle Spectrograph (HIRES; Vogt et al. 1994) on the Keck telescope and from the U.C.L Echelle Spectrograph at the Anglo Australian Telescope (AAT). We refer to the RV data from the AAT as Anglo-Australian Planet Search (AAPS) data. To verify the trustworthiness of our noise models in extracting weak signals from the data, we first test their performance by adding artificial signals to the HARPS data set for HD 10700. The models that enable the recovery of the artificial signals are then compared using the Bayesian model selection techniques to find the most accurate descriptions of these HARPS RVs. Finally, we search for periodic signatures of planetary companions in the HARPS velocities.

We also find the best noise models for the AAPS and HIRES RVs and use them in combination with the HARPS velocities in further analyses because signals that are not detected in any of these data sets might be available from the combined set because of its greater size and better phase-coverage relative to the individual data sets.

We start by describing the properties of the target star HD 10700 and the RV data in Sect. 2 and the statistical tools (methods and models) in Sect. 3. Model selection methods and our criteria for detecting signals in the RV data are presented in Sects. 4 and 5, respectively. After this, we proceed by determining which noise model performs the best in extracting artificial signals introduced into the data (Sect. 6). The best noise models are then used to analyse the HARPS data (Sect. 7). To make sure that the signals we detect do not correspond to any activity-related phenomena, we analyse the timeseries of HARPS activity indices in Sect. 7.5. We also perform analyses of the AAPS and HIRES data sets (Sect. 8) and the combined data (Sect. 9). Finally, we assess briefly the dynamical stability of the system of planet candidates we find in the combined data (Sect. 10) and present the conclusions and discussion in Sect. 11.

2. HD 10700 and radial velocity measurements

2.1. Stellar properties

HD 10700 is a very nearby Solar-type (G8.5 V; Gray et al. 2006) star with a large Hipparcos parallax of 273.96 ± 0.17 mas (van Leeuwen 2007) implying a distance of only 3.650 ± 0.002 pc. Because only 18 stellar systems (single, double, or triple stars) are located closer to the Sun, HD 10700 is in the immediate neighbourhood of our own system. While it can readily be called a Sun-like star, HD 10700 is somewhat lighter with a mass of $0.783 \pm 0.012\ M_\odot$ (Teixeira et al. 2009) and less luminous and also has a sub-Solar metallicity of -0.55 ± 0.05 (Pavlenko et al. 2012, and references therein), which could potentially make it an ideal target for searches for low-mass planets due to the relatively higher frequency of low-mass planets around low-metallicity stars (Jenkins et al. 2013).

Table 1. Estimated stellar properties of HD 10700.

Parameter	Value	Reference
Spectral type	G8.5 V	(i)
$\log R'_{\rm HK}$	−4.995	(h)
π [mas]	273.96 ± 0.17	(a)
$L_{\rm star}$ [$L_\odot$]	0.488 ± 0.010	(b)
$R_{\rm star}$ [$R_\odot$]	0.793 ± 0.004	(b)
$M_{\rm star}$ [$M_\odot$]	0.783 ± 0.012	(b)
$T_{\rm eff}$ [K]	5344 ± 50	(f)
[Fe/H]	−0.55 ± 0.05	(j)
Age [Gyr]	5.8	(d)
$v \sin i$ [km s^{-1}]	0.90	(g)
$P_{\rm rot}$ [days]	34	(e)

References. Data from: (a) van Leeuwen (2007); (b) Teixeira et al. (2009); (c) Soubiran et al. (1998); (d) Mamajek & Hillebrand (2008); (e) Baliunas et al. (1996); (f) Santos et al. (2004); (g) Santos et al. (2002); (h) Pepe et al. (2011); (i) Gray et al. (2006); (j) Pavlenko et al. (2012).

Despite the fact that HD 10700 is a target star of several RV searches for planets around nearby stars (e.g. the HARPS, Keck-HIRES, and AAPS, described in the next section), no planetary or other sub-stellar companions have been reported orbiting it (Wittenmyer et al. 2006; Pepe et al. 2011). However, the star has a bright debris disc, i.e. a circumstellar disc estimated to be an order of magnitude greater than the mass of Edgeworth-Kuiper belt in the Solar System, extending out to ~55 AU (Greaves et al. 2004) and has been observed to have warm dust orbiting it (Di Folco et al. 2007). These findings promote the possibility of HD 10700 having some of this circumstellar matter in the form of planets orbiting the star as well. Combining these arguments, and noting that HD 10700 is a very inactive star (a "flat activity" star; Judge et al. 2004) with $\log R_{\rm HK} = -4.955$ (Pepe et al. 2011), it seems likely that either HD 10700 is pole-on (as also suggested by Gray & Baliunas 1994, though there is currently no evidence in favour of this hypothesis), which, given co-planarity, prevents the detections of planets orbiting HD 10700 using the RV method (and indeed using the transit photometry), or its companions are very small and do not induce observable periodic Doppler variations to the stellar spectra. The obvious third option is that there simply are no planets orbiting HD 10700. While a plausible hypothesis, every revision of the frequency of planets around nearby stars seems to indicate that their frequency increases rapidly as their mass decreases and any estimates of their frequencies seem to be revised towards higher values as the amount of data accumulates (e.g. O'Toole et al. 2009; Howard et al. 2010; Mayor et al. 2011; Wittenmyer et al. 2011a).

2.2. Radial velocity data

The HARPS spectra of HD 10700 were extracted from the ESO archive in a wavelength-calibrated form. This calibration was made using the HARPS data reduction software (HARPS-DRS). After the spectral calibration, the RVs (Fig. 1) were calculated using the cross-correlation function (CCF) technique presented in Pepe et al. (2002).

As the HARPS data ($N = 4864$, 205 epochs) contained some velocities that had significantly greater, i.e. more than 100 times greater, uncertainties than the HARPS velocities had in general,

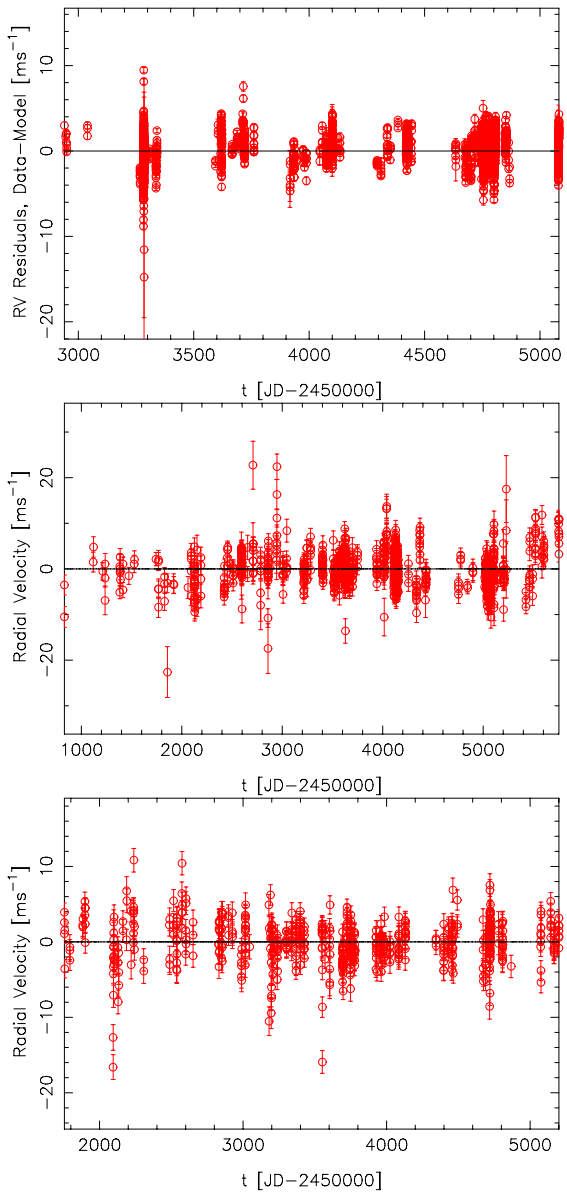

Fig. 1. HARPS (*top*), AAPS (*middle*), and HIRES (*bottom*) RVs with their respective mean estimates subtracted.

we simply neglected them and dropped them out of the data set prior to any further analyses. We also removed some spurious epochs from the set because they had velocities that differed by a few km s^{-1} from the data mode. As a result, there were 4398 RV points (202 epochs) with a baseline of 2142 days. By visual inspection this data set indeed seemed "flat", as also described by Pepe et al. (2011), and is unlikely to contain periodic RV signals with amplitudes in excess of few m s^{-1}. The standard deviation of these data was found to be 1.7 m s^{-1}, which implies that there indeed could not be signals in the data in excess of roughly 2.0 m s^{-1}.

Calculating the Lomb-Scargle periodogram (Lomb 1976; Scargle 1982) of the HARPS data revealed that this data set contains a "jungle of peaks" (Fig. 2, top panel) in excess of the analytical 0.1% false alarm probabilities (FAPs). The reason is that the noise in this "raw" velocity data is probably not Gaussian nor white and thus violates the assumptions underlying the periodogram analyses. From this periodogram, it is thus difficult to interpret the significance of the corresponding peaks.

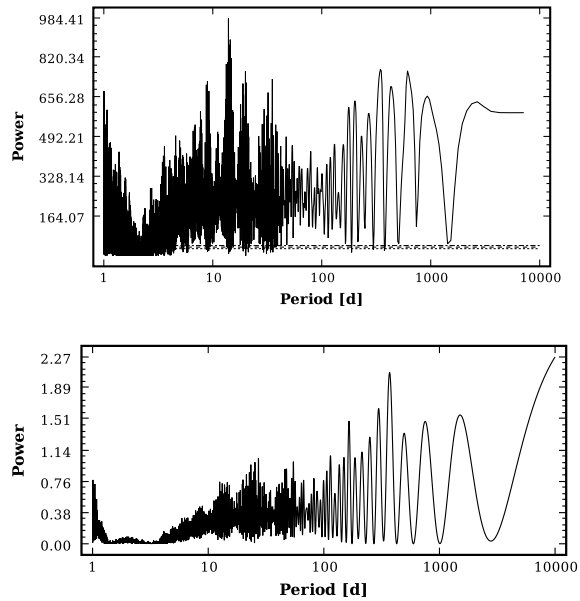

Fig. 2. Lomb-Scargle periodogram (*top*) with 0.1%, 1%, and 10% FAPs and the window function (*bottom*) of the HARPS data.

The 978 AAPS RVs have a similar character to the HARPS data (Fig. 1, middle panel). This data set, too, can be described as being flat and no periodic signals have been reported by the AAPS group despite an extensive baseline of the time-series of 4923 days (Wittenmyer et al. 2011b). This data does not have such clear annual gaps as the HARPS data (Fig. 1) and deviates from the mean by approximately 5.0 m s^{-1} on average.

The smallest set of RVs was that measured using the HIRES (Fig. 1, bottom panel). The 567 HIRES RV measurements have a baseline of 3446 days and nothing has been reported about this data set in the literature. The velocities deviate roughly 2.9 m s^{-1} from the mean and do not have significant gaps apart from relatively narrow annual gaps corresponding to the visibility of HD 10700 from the Keck telescope's northern location in Hawaii.

3. Statistical analysis and modelling

We modelled the RVs with a statistical model that contains three additive terms. These terms are (1) the superposition of k Keplerian signals and a reference velocity, (2) Gaussian noise with two components, namely the instrument noise and all the excess noise in the measurements, and (3) an autoregressive (AR) and/or moving average (MA) term that accounts for the possible correlations between the subsequent velocities and/or their errors. The MA term accounts for the dependence of a measurement on the deviation of the last measurement from the mean – this term corresponds effectively to binning measurements made within a certain timescale in order to remove variations within that timescale. While the first part of this model is described in detail in Tuomi (2011) and Tuomi et al. (2011), we describe the rest of our statistical models and modelling approaches in the following subsections.

3.1. Posterior samplings

We analysed each model in our collection of candidate models using the adaptive Metropolis algorithm (Haario et al. 2001). This algorithm works well for typical models of RVs and can

be used to receive statistically representative samples from the posterior densities of the parameters of these models and with relatively little computational cost (Tuomi 2011, 2012; Tuomi et al. 2011). The proposal density of the adaptive Metropolis algorithm is a multivariate Gaussian density, which poses problems and possibly slows down the convergence rate of the chains when the parameter posterior density is non-Gaussian with high non-linear correlations between the parameters. We overcome this problem by simply increasing the chain lengths sufficiently in each sampling to ensure that we obtain statistically representative samples. Typically, samples of roughly 10^5–10^6, depending on the dimension of the parameter vector, are sufficient when the posterior density can be approximated as a multivariate Gaussian density. However, when there are several Keplerian signals in the model and the posterior density necessarily has non-linear correlations, we increase the chain lengths by factors of up to 100 to ensure that the samples we obtain are statistically representative.

The fact that the adaptive Metropolis algorithm is not exactly Markovian means that the sample drawn from the posterior density might not be close enough to the actual posterior density to draw conclusions on the parameters, and, on the model probabilities. We approximate these probabilities using the truncated posterior mixture (TPM) estimate of Tuomi & Jones (2012) that requires a statistically representative sample from the posterior. For this reason, when the chains we calculate have converged close to the posterior density and are found to move randomly around in the vicinity of the maximum a posteriori (MAP) estimate, we fixed the proposal density of the chain to its current value. This essentially makes this algorithm equivalent to the Metropolis-Hastings Markov chain Monte Carlo (MCMC) algorithm (Metropolis et al. 1953; Hastings 1970). This is the same procedure that has been used in e.g. Tuomi (2012).

We estimate all the parameters simply by using the MAP estimates and the 99% Bayesian credibility sets (as defined in e.g. Tuomi & Kotiranta 2009), i.e. 99% credibility intervals in a single dimension. For a definition of the MAP estimates, and the corresponding caveats of relying on point estimates in the first place, we refer the interested reader to any introductory text of Bayesian statistics (e.g. Berger 1980).

3.2. First order AR model

To be able to use autoregression in the statistical modelling and analysis, we arrange the RVs such that for epochs t_i and t_j, $t_i < t_j$ if $i < j$, i.e. put the measurements in chronological order. Now, the measurement (r_i) at epoch t_i depends on that at t_{i-1} according to

$$\begin{aligned} r_i &= f_k(t_i) + \epsilon_i + \epsilon_J, \text{ for } i = 1 \\ r_i &= f_k(t_i) + \phi_{i,i-1} r_{i-1} + \epsilon_i + \epsilon_J, \text{ for } i > 1, \end{aligned} \tag{1}$$

where function f_k is the superposition of k Keplerian signals with an additive constant parameter γ_l, corresponding to the reference velocity of the lth telescope-instrument combination, ϵ_i is a Gaussian random variable with a zero mean and known variance corresponding to the estimated instrument noise σ_i^2 at t_i, the Gaussian random variable ϵ_J with zero mean and unknown variance (σ_J^2) represents all the excess (white) noise in the data.

The function $\phi_{i,j}$ describes the magnitude of the correlation between two subsequent measurements made at t_i and t_j. We assume that the correlation described by this function depends on the time difference between the two consequent observations and write it as

$$\phi_{i,j} = \phi_i \exp\left[\alpha(t_j - t_i)\right], \tag{2}$$

where the positive numbers ϕ_j, for all j, and α are free parameters of the model. It can be seen that the function $\phi_{i,j}$ decreases exponentially as the time difference between the two consequent epochs increases. Also, when $|\phi_j| < 1$ for all j and $i > j$, its value is always less than unity, which makes the model stationary.

3.3. First order MA model

To apply the MA model, we arrange the RVs again in chronological order. The RV measurement made at epoch t_i is then modelled as

$$\begin{aligned} r_i &= f_k(t_i) + \epsilon_i + \epsilon_J, \text{ for } i = 1 \\ r_i &= f_k(t_i) + \omega_{i,i-1}(\epsilon_{i-1} + \epsilon_J) + \epsilon_i + \epsilon_J, \text{ for } i > 1, \end{aligned} \tag{3}$$

The MA part is defined using the function $\omega_{i,j}$ multiplied by the deviation of the previous observation from f_k. This function has the same form as $\phi_{i,j}$, with parameters ω_j and β instead of ϕ_j and α in Eq. (2), because we expect the dependence of the ith measurement on the deviation of the previous one to decrease as a function of the time difference of the corresponding measurements.

3.4. General ARMA model

It may be the case in practice that taking into account the correlations between measurements at epochs t_i and t_{i-1} is not sufficient but a better model can be constructed by taking into account the correlations between the measurement at t_i and those at t_{i-j}, $j = 1, ..., q$, where $q < i$. We write the corresponding qth order AR model as

$$\begin{aligned} r_i &= f_k(t_i) + \epsilon_i + \epsilon_J, \text{ for } i = 1 \\ r_i &= f_k(t_i) + \epsilon_i + \epsilon_J + \sum_{j=1}^{q} \phi_{j,i-j} r_{i-j}, \text{ for } i > 1. \end{aligned} \tag{4}$$

Similarly, the pth order MA model is defined as

$$\begin{aligned} r_i &= f_k(t_i) + \epsilon_i + \epsilon_J, \text{ for } i = 1 \\ r_i &= f_k(t_i) + \epsilon_i + \epsilon_J + \sum_{j=1}^{p} \omega_{j,i-j}(\epsilon_{i-j} + \epsilon_J), \text{ for } i > 1. \end{aligned} \tag{5}$$

The general ARMA model is then written by including both AR and MA terms in the statistical model together with the function f and the additive random variables representing the white noise components in the data. We denote this general model as AR(q)MA(p).

3.5. White noise models

To determine the shape of the additive white noise in our statistical models, i.e. the distribution of the sum $\epsilon_i + \epsilon_J$, we compare some different distributions by relaxing the assumption that these two random variables (and consequently the sum) have Gaussian probability distributions. The Gaussian distribution is written simply as

$$\mathcal{N}(\mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}, \tag{6}$$

and is well known for its property that independent random variables $X_1 \sim \mathcal{N}(\mu_1, \sigma_1^2)$ and $X_2 \sim \mathcal{N}(\mu_2, \sigma_2^2)$ satisfy $X_1 + X_2 \sim \mathcal{N}(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$.

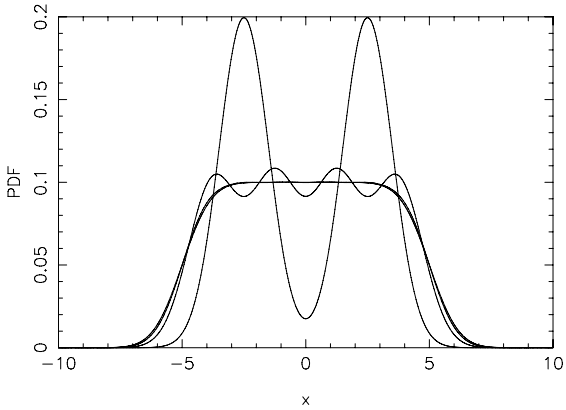

Fig. 3. Estimated probability distribution $\mathcal{S}(x|a,\sigma^2)$ in Eq. (8), with $n = 1, ..., 6$ (the corresponding functions have $2n$ maxima). The parameters are selected as $a = 5$ and $\sigma^2 = 1$.

The first alternative distribution we use is the Cauchy distribution $C(x|\mu,\gamma) = C(\mu,\gamma)$ defined as

$$C(\mu,\gamma) = \frac{1}{\pi}\left[\frac{\gamma}{(x-\mu)^2+\gamma^2}\right], \tag{7}$$

which has longer tails than the Gaussian distribution and satisfies the convenient property that for independent $X_1 \sim C(\mu_1,\gamma_1)$ and $X_2 \sim C(\mu_2,\gamma_2)$ it holds that $X_1 + X_2 \sim C(\mu_1+\mu_2,\gamma_1+\gamma_2)$. This distribution is selected to see if the noise is dominated by outliers that cannot be explained by the relatively short and "light" tails of the Gaussian distribution.

Our second alternative model assumes that $\epsilon_J \sim \mathcal{U}(-a,a) * \mathcal{N}(0,\sigma_J^2)$, where $\mathcal{U}(-a,a)$ is a uniform distribution of interval $[-a,a]$, and $\epsilon_i \sim \mathcal{N}(0,\sigma_i^2)$ are independent. In this case, their sum is distributed according to the convolution of the densities which we approximate as

$$\mathcal{S}(x|a,\sigma^2) = \lim_{n\to\infty}\frac{1}{2n}\sum_{k=0}^{2n-1}\mathcal{N}\left(x\middle|\frac{a}{2n}[1-2n+2k],\sigma^2\right), \tag{8}$$

where $\sigma^2 = \sigma_i^2 + \sigma_J^2$. In practice, when the values of a and σ are of the same magnitude, approximating the above infinite sum with a choice of $n = 6$ provides an accurate estimate for the distribution. This distribution has a "flat" maximum but tails according to the Gaussian function, as seen in Fig. 3 for $a = 5$ and $\sigma = 1$. Even when parameter a is five times greater than σ, the approximation converges very rapidly and is an accurate description of the convolution for $n = 3$ – for obvious reasons decreasing a or increasing σ or n improves this accuracy. As seen in Fig. 3, the curves for $n = 3, ..., 6$ are practically indistinguishable from one another. We choose this model to investigate if the peak of the white noise component is not as sharp as in the Gaussian (or indeed in the Cauchy) model because a range of values at the vicinity of zero have roughly equal probabilities.

To summarise the above, our white noise models consist of (1) a Gaussian distribution because – according to our knowledge – it is the only distribution that has been used to model the measurement noise when analysing RV data, (2) a Cauchy distribution with longer tails, and (3) a flatter distribution that accounts for jitter with small deviations from the mean equally likely regardless of their exact magnitude. While these distributions are by no means a comprehensive collection of white noise models, we expect the comparisons of these models to provide information on the overall shape of the white noise component caused by the telescope-instrument combination and the stellar surface.

3.6. Prior choice

Because we take advantage of Bayesian inference, the choice of priors needs to be addressed briefly. Essentially, we use the same prior probability densities and prior probabilities that were used in Tuomi (2012), there with the much more restricted data set of HD 10180. In particular, we chose $\pi(e) \propto \mathcal{N}(0, 0.3^2)$, with the corresponding scaling, which penalises eccentricities more as they get closer to unity, yet, allows them if the higher eccentricities are needed to better describe the data. Because it is a scale invariant parameter, we adopt the logarithm of the period as a parameter and use a uniform prior with cutoffs $T_{\rm min}$ and $T_{\rm max}$ corresponding to one day and 10 $T_{\rm obs}$, respectively, where $T_{\rm obs}$ is the baseline of the data analysed. We choose $T_{\rm max} > T_{\rm obs}$ because signals with periods in excess of the data baseline can be detected in RV data sets (Tuomi et al. 2009).

The noise models have additional parameters as well, and their priors were chosen as follows. The parameters ϕ_j (and ω_j) were set to have uniform priors in the interval $[-1, 1]$, for all j, to allow both positive and negative correlations to occur. Although, we expected these parameters to receive mostly positive values that were at least consistent with zero given their uncertainties. The parameter α (β), which determines the time-scale of the AR (MA) effect in the noise, was chosen to have a uniform prior in the logarithmic scale, i.e. the Jeffreys prior, with cutoffs at $\alpha_{\rm min} = 1$ min^{-1} and $\alpha_{\rm max} = 1$ year^{-1}. Though we expected this effect to take place on the time-scale of hours, we chose a wider interval limited by the minimum time-difference between two subsequent measurements (~1 min) and a maximum of one year to account for possible noise correlations on longer time-scales as well.

We choose the prior probabilities, $P(\mathcal{M}_i)$, of our noise models ($\mathcal{M}_i$) to be equal. However, when comparing models with different numbers of Keplerian signals we do not assume equal prior probabilities but set them such that they favour the model with one less signal by a factor of two. These priors have been applied in e.g. Tuomi (2012).

Given the unprecedented amount of data, we expect the priors to be overwhelmed in the case of HD 10700 velocities and that they do not have a detectable effect on the results. For this reason, we do not test the dependence of our results on the choice of prior densities.

4. Model selection

We take advantage of the Bayesian model selection framework, in which each model is equipped with a number that describes its relative posterior probability given the measurements. The fact that this probability is relative means that it only tells how probable it is that the measurements, as random variables, have been drawn from the random process described by the model instead of being drawn from those described by the other models in the model set. Bayesian model selection has been used in determining the most probable number of planetary signals in the RV data (e.g. Gregory 2005, 2007a,b, 2011; Ford & Gregory 2007; Tuomi & Kotiranta 2009; Tuomi 2011, 2012; Tuomi et al. 2011) but it can be readily applied to any other model comparison problem because of its generality.

We compare the models by calculating the posterior probabilities as

$$P(\mathcal{M}_i|m) = \frac{P(m|\mathcal{M}_i)P(\mathcal{M}_i)}{\sum_{j=1}^{k} P(m|\mathcal{M}_j)P(\mathcal{M}_j)}, \quad (9)$$

where $P(m|\mathcal{M}_i)$ are the marginal integrals whose values need to be calculated for each model given the measurements m. We estimate these integrals using the TPM estimate but verify the results using the simple Akaike information criterion (Akaike 1973) for small sample size (AICc, see e.g. Burnham & Anderson 2002) because the data sets we analyse are large enough so that the number of data points well exceeds the number of model parameters.

To ensure that the TPM estimates we obtain are trustworthy, we perform several samplings for each data set and statistical model with different initial states. We perform three such samplings to check that the obtained TPM estimates are consistent. If we find that the sample sizes are insufficient, we typically increase them by a factor of 10 and again obtain three independent samples. We found that even when there were several Keplerian signals in the model, sample sizes of the order of 10^7 were sufficient and independent samplings yielded TPM estimates within roughly 0.05 from one another in the log-scale. We note that the parameters λ and h in the algorithm of the TPM estimate (Tuomi & Jones 2012) were chosen to be 10^{-5} and 10^5, respectively, because the chain members θ_n and θ_{n+h} became independent for roughly $h \approx 10^3$–10^4 and parameter λ of 10^{-5} was found to converge rather rapidly but still provide TPM estimates within roughly 0.05 from one another in the log-scale.

In practice, we first compare different AR and MA models to find out the most reliable description of the nature and amount of autoregression and noise correlation in the data. After that, we compare the different models for the remaining white noise in the data. We perform the comparisons by using the HARPS RVs of HD 10700. We generate a total of fourteen data sets from the 4398 HARPS RVs by adding Keplerian signals to the time-series to see which ones of the AR and MA models yield the correct RV amplitudes for these artificial signals. Out of the models that yield results consistent with the parameters of the artificially added signals, we select a model that has the greatest posterior probability according to Eq. (9). We do not simply choose blindly the best model according to this Eq. but require primarily that the model yields results consistent with the Keplerian signals introduced into the data. The reason for this choice is that it is possible that a signal, or at least part of it, gets interpreted as being a consequence of significant autoregression in the data, or is generated by pure noise by some other means. Therefore, especially with respect to the AR models, we analyse the reliability of the different noise models carefully.

After having found the best descriptions out of the AR and MA models, especially the ones that do not result in biases in the artificial signals, we perform a comparison of our white noise models.

5. Signal detection criteria

Before analysing the RV data with and without artificially added signals, we discuss briefly the requirements for a positive detection of a periodic signal in the data. Because these criteria are essential in the detection of weak signals in, not only RVs, but in any measurements aiming at detecting periodic signals, we describe our criteria in this section.

The regular approach to detections of Keplerian signals in RVs is based on the Lomb-Scargle periodogram of residuals that are assumed to have a Gaussian distribution (Lomb 1976; Scargle 1982). While we studied the periodogram powers in our analyses, we do not use them as an indication of whether there is a periodic signal in the data or not. The reason for this choice is that calculating the periodogram of model residuals assumes the remaining noise is Gaussian and that there are no additional signals – if there are, the assumptions are violated and the test cannot be considered trustworthy (see e.g. Tuomi 2012).

The analytical detection threshold of Tuomi et al. (2009) can be used to receive a rough estimate for the detectability of various signals in a given data set. According to this criterion, a periodic signal with period P can be detected if the square of its amplitude exceeds the threshold given by

$$K_T^2 = \frac{9.22\sigma^2}{N}\left[f_c^2(\psi) + f_s^2(\psi)\right], \text{ where} \quad (10)$$

$$f_c = 2[1 - \cos \pi\psi]^{-1} \text{ if } \psi < 1 \text{ and unity otherwise,}$$

$$f_s = [\sin \pi\psi]^{-1} \text{ if } \psi < 0.5 \text{ and unity otherwise,}$$

where N is the number of measurements, σ is the average noise level of the data, and $\psi = TP^{-1}$, where T is the baseline of the data. This criterion is only a very rough estimate because it does not take into account the various effects that data sampling might have on the detectability of signals. However, it does take into account the ratio of the data baseline and the period of the signal, which means that the criterion is applicable even in cases where the period of the signal exceeds the data baseline.

Using the criterion in Eq. (10), we find that signals with periods less than roughly 1000 days can be detected in the HARPS RVs of HD 10700 if their amplitudes exceed 0.1 m s^{-1}. In practice, the amplitude of a signal has to be significantly above this threshold, i.e. in excess of the 99% Bayesian credibility interval. While this threshold appears to be very low, it results from the fact that the data set contains a large number of high-precision RVs.

Because the threshold in Eq. (10) is only a rough approximation, we use more robust criteria to determine whether signals are present in a data set. We say that $k + 1$ signals are detected if (1) the posterior probability of a model with $k + 1$ signals is at least 150 times greater than that of a model with k signals (Kass & Raftery 1995; Feroz et al. 2011; Tuomi 2012), if (2) the RV amplitudes of all signals are statistically significantly greater than zero, and if (3) the periods of all signals are well constrained from above and below (Tuomi 2012). We adopt these criteria but require also that the signal amplitude exceeds the threshold presented in Eq. (10).

6. Artificial signals

We added twelve different Keplerian signals to the HARPS RV data of HD 10700 to see which models could extract their parameters from these artificial data sets the most accurately. These signals were set to have periods of 20, 50, 100, and 200 days and amplitudes of 1, 2, and 5 m s^{-1}, which resulted in a collection of twelve data sets. We also generated two additional sets by adding a 200 days periodicity with amplitudes of 0.5 and 0.3 m s^{-1} to see whether such extremely low-amplitude signals could be retrieved from the data. We chose the 200 days period because the power spectrum of the HARPS velocities had the fewest peaks between roughly 150 and 300 days (Fig. 2). We state that a signal is detected reliably if (1) the detection criteria in the previous

Table 2. Artificial (first column) and retrieved (subsequent columns) signals (their MAP parameter estimates) in terms of parameters K, P, and e using a model without any of the AR or MA components (0) and with selected AR or MA models for each data set.

Signal	0	AR(1)	MA(1)	AR(3)	MA(3)	MA(5)	MA(7)	MA(10)
$K = 5$ [m s^{-1}]	5.50	2.87	5.49	2.03	5.32	5.36	5.39	5.38
$P = 20$ [days]	19.99	19.99	19.99	19.99	**20.00**	**20.00**	**20.00**	**20.00**
$e = 0$	0.05	0.06	0.05	**0.00**	**0.01**	**0.03**	**0.00**	**0.02**
$K = 2$ [m s^{-1}]	2.60	1.25	2.57	0.75	2.39	2.41	2.37	2.40
$P = 20$ [days]	19.98	19.98	19.98	19.97	**19.99**	**20.00**	**20.00**	**20.00**
$e = 0$	0.08	0.08	**0.07**	0.10	**0.08**	**0.02**	**0.05**	**0.03**
$K = 1$ [m s^{-1}]	1.68	0.80	1.65	0.44	1.41	1.39	1.39	1.40
$P = 20$ [days]	19.96	19.97	19.97	19.96	**19.98**	**20.00**	**20.00**	**20.01**
$e = 0$	0.11	**0.06**	**0.10**	**0.01**	**0.07**	**0.07**	**0.03**	**0.03**
$K = 5$ [m s^{-1}]	5.32	2.63	5.33	1.78	5.30	5.31	5.27	5.31
$P = 50$ [days]	50.08	50.08	50.07	50.06	50.05	50.05	50.04	50.05
$e = 0$	0.07	0.07	0.07	**0.00**	0.05	**0.04**	**0.00**	**0.00**
$K = 2$ [m s^{-1}]	2.42	1.13	2.45	0.65	2.40	2.37	2.33	2.33
$P = 50$ [days]	50.12	50.13	50.14	50.17	50.11	50.09	50.09	50.11
$e = 0$	0.14	0.13	0.13	**0.03**	0.11	**0.08**	**0.07**	**0.07**
$K = 1$ [m s^{-1}]	1.58	0.72	1.56	0.39	1.41	1.37	1.36	1.36
$P = 50$ [days]	50.21	50.21	50.20	50.21	50.16	50.16	50.17	50.19
$e = 0$	0.25	0.22	0.20	**0.21**	**0.15**	**0.11**	**0.10**	**0.05**
$K = 5$ [m s^{-1}]	5.46	2.64	5.45	1.76	5.39	5.40	5.36	5.35
$P = 100$ [days]	**99.98**	99.85	**99.92**	100.07	**99.96**	**99.99**	**99.98**	**99.99**
$e = 0$	0.09	**0.00**	0.07	0.04	0.05	**0.04**	**0.05**	**0.03**
$K = 2$ [m s^{-1}]	2.67	1.30	2.58	0.68	2.43	2.36	2.39	2.36
$P = 100$ [days]	100.47	100.43	100.38	**100.02**	**100.00**	**99.99**	**100.02**	**100.00**
$e = 0$	0.30	0.30	0.30	**0.05**	0.16	**0.11**	**0.10**	**0.10**
$K = 1$ [m s^{-1}]	1.72	0.80	1.63	0.44	1.48	1.41	1.42	1.40
$P = 100$ [days]	100.48	100.45	100.41	**100.44**	**100.06**	**100.11**	**100.07**	**100.13**
$e = 0$	0.30	0.29	0.29	**0.27**	**0.27**	**0.24**	**0.17**	**0.19**
$K = 5$ [m s^{-1}]	**4.84**	2.34	**4.86**	1.46	**4.91**	**5.00**	**5.03**	**5.03**
$P = 200$ [days]	200.74	200.69	200.62	200.66	**200.34**	**200.15**	**200.14**	**200.09**
$e = 0$	0.08	0.06	0.07	**0.02**	0.07	0.07	0.07	0.07
$K = 2$ [m s^{-1}]	**1.95**	0.94	**1.96**	0.59	**2.03**	**2.00**	**2.05**	**2.09**
$P = 200$ [days]	201.97	201.73	201.71	**201.33**	**199.92**	**201.04**	**200.13**	**199.84**
$e = 0$	0.14	**0.12**	0.14	**0.09**	**0.16**	0.13	**0.12**	0.12
$K = 1$ [m s^{-1}]	**1.13**	0.54	**1.06**	0.33	**1.05**	**1.10**	**1.11**	**1.12**
$P = 200$ [days]	203.55	203.18	203.55	**202.14**	**201.92**	**201.68**	**200.57**	198.06
$e = 0$	0.20	**0.14**	0.15	**0.12**	**0.14**	**0.15**	**0.15**	**0.14**
$K = 0.5$ [m s^{-1}]	–	–	–	–	–	**0.66**	**0.64**	**0.65**
$P = 200$ [days]	–	–	–	–	–	**202.08**	**201.97**	**201.36**
$e = 0$	–	–	–	–	–	**0.09**	**0.13**	**0.11**
$K = 0.3$ [m s^{-1}]	–	–	–	–	–	**0.50**	**0.48**	**0.48**
$P = 200$ [days]	–	–	–	–	–	**201.84**	**202.62**	**203.16**
$e = 0$	–	–	–	–	–	**0.17**	**0.11**	**0.09**

Notes. Signals detected reliably, i.e. whose 99% Bayesian credibility intervals contain the parameter values of the artificially added signals, are emphasised using boldface.

section are satisfied and if (2) the 99% Bayesian credibility sets (BCSs; as defined in e.g. Tuomi & Kotiranta 2009), i.e. intervals in one dimension, of the period and amplitude parameters contain the correct values of the added artificial signals.

According to the results presented in Table 2, the pure white noise model and both first order AR and MA models could not be used to determine the parameter values of the artificial signals reliably. This was found to be the case even with the signals with amplitudes as high as 5 m s^{-1} that should be detectable from high-precision data easily, especially, given the large number of data points. We also found this to be the case for higher order AR models that yielded severe biases in the signals because the signals were interpreted, in part, as noise-related correlations in the data. Therefore, despite the addition of AR components to the noise model improving the goodness of the model, we do not consider the AR models trustworthy for our purposes. According to the results in Table 2, the AR models underestimate the signal amplitudes significantly.

The MA models were found to be reasonably reliable in quantifying the properties of the signals in the data. While they

overestimated slightly the amplitudes of the 20, 50, and 100 day signals, they were accurate for the 200 day signal though again somewhat overestimated the signal amplitudes when recovering the 0.3 or 0.5 m s^{-1} injected signals. Also, apart from the 50 day signal, the MA models of order seven and ten yielded the best results for periods and eccentricities of the injected signals.

The reason the 50 day signal could not be extracted correctly and the fact that the MA models, even the most accurate (i.e. that have the greatest posterior probabilities) seventh and tenth order ones, yielded biased estimates for the amplitudes of the 20 and 100 day day signals warrants an explanation. We are essentially studying the properties of the noise in the HARPS data of HD 10700. However, there is a possibility that the noise models we use lack some important features that impinges on the ability to relialbly recover injected signals. Another possibility is that there are already signals present in the data and we actually detect the superpositions of these real signals and the artificial ones. If this is the case, the above considerations depend on how much the existing signals affect the artificial ones. We expect that there are no significant signals at or around 200 days in the data because the 200 day signals were extracted the most accurately with the best MA models. However, in Sect. 7 we show that there are genuine low-amplitude signals in the data near the periods that provided relatively poor recovery of signals and so the lack of complete recovery of injected signals is not surprising.

The artificial signals at 200 days with the lowest amplitudes received slightly biased amplitudes. The amplitudes of the recovered artificial signals were systematically around 0.15−0.20 m s^{-1} greater than their real values. While these values were within the 99% BCSs of the obtained estimates, this over-estimation is a rather awkward feature and implies that the models are not as good descriptions of the data as they should be. Yet, this might again be a caused by the fact that there are signals – or their aliases – at or around 200 days in the HARPS data that cause biases to our estimates.

In Table 2, the estimates are only shown for models MA(5), MA(7), and MA(10), because the other models did not identify any significant periodicities at or around 200 days.

When sampling the posterior density, we observed that the joint posterior density of the noise parameters and reference velocity was close-Gaussian in all the samplings we performed. Therefore, we are confident that the adaptive Metropolis algorithm enables fast convergence and enables us to draw statistically representative samples. We illustrate this by plotting the equiprobability contours of parameters β, ω_1, γ, and σ_J in Fig. 4 as obtained using the HARPS data and a model without Keplerian signals.

7. HARPS radial velocities of HD 10700

7.1. Noise model

We analysed the HARPS RVs using several different noise models. We chose the same AR or MA models as in Table 2 and calculated their posterior probabilities to compare their performances in explaining the data. We started with pure noise models without any Keplerian signals. As was the case with the datasets with injected artificial signals, the AR models (AR(1) and AR(3)), not to mention the pure white noise model, did not perform as well as the MA models (Table 3). According to our probabilities in Table 3, the MA(10) model was the best description of the data because it had greater posterior probability than the MA(7) model. It can be seen that adding components to the MA model improves its performance until there are roughly 5−10 components. Therefore, we did not try more components and continued analysing the data with the MA(10) noise model that had the greatest posterior probability.

Table 3. Relative posterior probabilities and log-Bayesian evidences for different noise models using the HARPS RVs.

Model	$P(\mathcal{M}\|d)$	$\ln P(d\|\mathcal{M})$
G	$\sim 10^{-673}$	−8431.0
G+AR(1)	$\sim 10^{-245}$	−7446.0
G+MA(1)	$\sim 10^{-245}$	−7447.3
G+AR(3)	3.2×10^{-53}	−7004.5
G+MA(3)	1.1×10^{-48}	−6994.0
G+MA(5)	2.8×10^{-14}	−6914.8
G+MA(7)	1.9×10^{-9}	−6903.7
G+MA(10)	0.61	−6884.1
C+MA(10)	$\sim 10^{-167}$	−7267.1
G+U+MA(10)	0.39	−6884.5

Notes. No Keplerian signals were included in the model. The noise models are Gaussian (G), Cauchy (C), Uniform (U), AR(q), and MA(p).

Table 4. Relative posterior probabilities and log-Bayesian evidences of models $\mathcal{M}_k$ with $k = 0, ..., 3$ Keplerian signals, given the HARPS RVs.

k	$P(\mathcal{M}_k\|d)$ TPM	$\ln P(d\|\mathcal{M}_k)$ TPM	$P(\mathcal{M}_k\|d)$ AICc	$\ln P(d\|\mathcal{M}_k)$ AICc
0	4.7×10^{-45}	−6884.1	1.0×10^{-39}	−6886.7
1	8.4×10^{-29}	−6846.0	1.6×10^{-24}	−6851.1
2	8.1×10^{-18}	−6820.0	7.4×10^{-14}	−6825.8
3	~ 1	−6780.0	~ 1	−6794.9

We also tested the relative performance of the two different white-noise models, namely the Cauchy and the convolution of Gaussian and uniform density (C and G+U in Table 3, respectively), in explaining the HARPS velocities. According to our results, the Cauchy noise model does not perform well with respect to the Gaussian white noise model (Table 3). The uniform component, containing one extra parameter that penalises the probability of this model in accordance with the principle of parsimony, does not increase the performance of the noise model enough to receive a greater posterior probability than the Gaussian model. This indicates that the white noise component of the HARPS velocities has a shape that resembles the Gaussian density and can be modelled well using the density in Eq. (6) if the variance is written as the sum of the variances of the instrument noise (σ_i^2) and the excess noise (σ_J^2) for every measurement r_i. In practice this also means that the additional parameter a in Eq. (8) is statistically indistinguishable from zero in this case.

Thus we proceed to model the noise in the HARPS velocities using the tenth order MA model and Gaussian white noise.

7.2. Signals in the HARPS data

After removing the MAP estimated MA(10) components of the noise from the data and calculating the Lomb-Scargle periodogram of the residuals, most of the peaks appearing in the "raw" velocity data (Fig. 2) seem to disappear from the power spectrum (Fig. 5, top panel). However, there is one strong peak that exceeds the 1% FAP at 35.3 days and four others that exceed the 10% FAPs at 13.9, 20.1, 363, and 595 days. Since we did not perform binning of the data, the measurements likely contain more information than the binned data from which significant

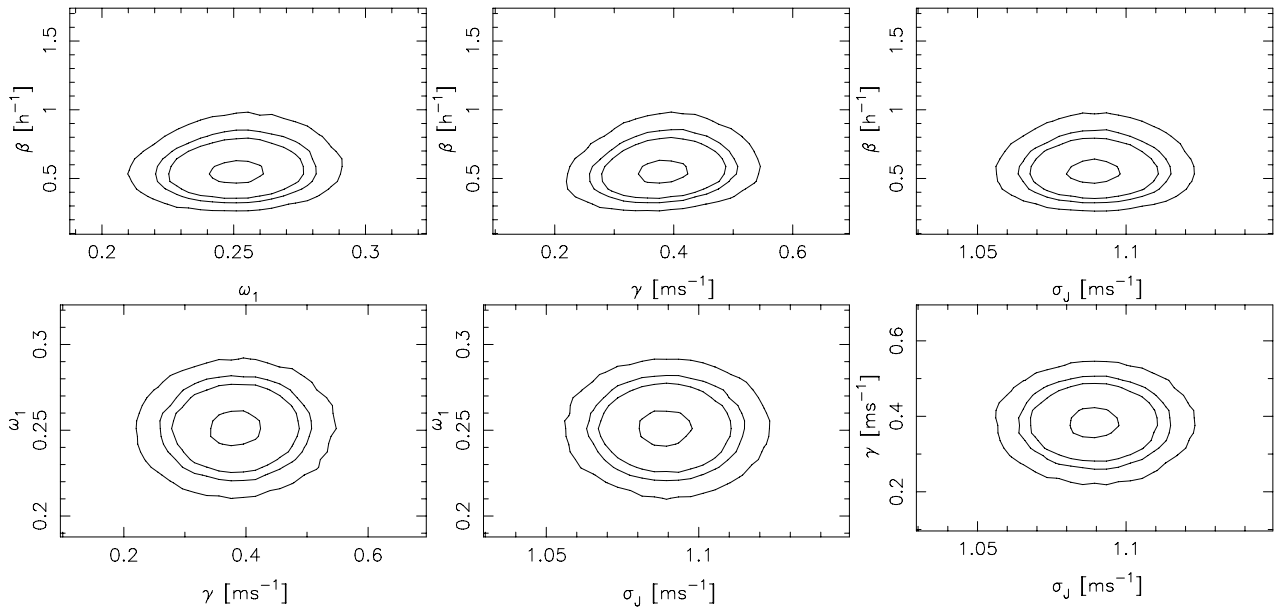

Fig. 4. Equiprobability contours containing 50%, 10%, 5%, and 1% of the probability density of all the combinations of parameters β, ω_1, γ, and σ_J from a single Markov chain using a model without Keplerian signals and a MA(10) noise description.

periodicities have not been found (Pepe et al. 2011). Therefore, we added one Keplerian signal to our model and calculated its posterior probability to see if the strongest peaks in the periodogram were significant signals according to our criteria.

The one-Keplerian model increased the model probability by a factor of 1.8×10^{16} (Table 4), which makes the corresponding periodicity of 35 days significantly present in the data. However, the residuals of the one-Keplerian model showed an additional peak exceeding the 1% FAP at 14 days (Fig. 5, second panel), and we continued analysing the data with a two-Keplerian model. This model further increased the model probability by a factor of 9.6×10^{10}. Finally, the residuals of this two-Keplerian model contained one more signal at roughly 94 days that exceeded the 10% FAP level (Fig. 5, third panel). Again, the corresponding periodicity in a three-Keplerian model increased the model probability significantly by a factor of 1.2×10^{17}. After removing this signal there were no strong powers in the power spectrum (Fig. 5, bottom panel) and the samplings of the parameter space of a four-Keplerian model failed to identify any additional significant periodicities. Therefore, there appear to be three Keplerian signals in the HARPS RVs of τ Ceti. The posterior probabilities in Table 4 indicate that taking these signals into account improves the model very significantly, which implies that there is strong evidence for three periodic signatures in the τ Ceti data. The probabilities based on the simple AICc imply the same qualitative results (Table 4).

We show the MAP phase-folded orbits of our three-Keplerian solution in Fig. 6. While these plots are not visually very impressive, they do indicate that the large amount of data together with the improved modelling of the noise enables the detection of these signals. The RV amplitudes of all three signals were found to have MAP estimates below 1.0 m s^{-1} level but they were still well constrained and differ from zero very significantly (Table 5).

With respect to the artificial data sets we analysed in the previous section, we suspect that the artificial signals at 50 days did not get detected reliably because there already were periodicities at 35 and 94 days in the data. The strongest periodicity in the data, namely at 35 days with an amplitude of roughly 1 m s^{-1}, is likely preventing the reliable detection of the artificial 50 days signals. While we still could extract them significantly out of the data, their amplitudes were biased, which is likely caused by the fact that the superposition of the artificial and real signals (and/or their aliases) was actually the one that we detected. To test this hypothesis, we analysed one of the data sets with injected signal (K = 1.0 m s^{-1}, P = 50 days) while simultaneously modelling the existing signals in the data. While we still obtained biased estimates for the parameters P and K, taking into account all three periodicities at periods of approximately 14, 35, and 94 days (Table 5) enabled us to retrieve the injected 50 day signal correctly given the uncertainties of the parameters. This indicates that the existing periodicities might indeed cause biases to the process of recovering the artificial signals. We expect this is the reason the 20 and 100 days signals were poorly recovered from our artificial test data. However, there do not appear to be very strong periodicities around 200 days (Fig. 5), which made it possible to extract the corresponding artificial signals from the data reasonably reliably given a sufficiently sophisticated noise model. Yet, even the artificial signals at 200 days received amplitudes that were systematically greater than the values used to generate them. This suggests that despite the fact that we could not detect additional signals in the HARPS data, there might still be some low-amplitude signals hidden in the measurement noise at longer periods.

We note that, with the samplings of the parameter space of models with more than one Keplerian signal, the non-linear correlations slow the convergence rate of the Markov chains. This effect arises because the multivariate Gaussian proposal density does not resemble the posterior very well. For this reason, we adopted a brute force approach and simply increased the chain lengths suficiently to obtain samples that are statistically representative. In practice, we increased the chain lengths to up to few 10^7 to ensure that they had indeed converged to the posterior density.

The signals in the HARPS velocities are of low amplitude and it is therefore relevant to ask whether they can be retrieved from the HARPS data with different noise models. We tested the dependence of the extracted signals on the noise models by

Table 5. MAP estimates and the corresponding 99% BCSs of the parameters of the three-Keplerian model for the HARPS RVs.

Parameter	HD 10700 b	HD 10700 c	HD 10700 d
P [d]	13.927 [13.901, 13.957]	35.314 [35.142, 35.434]	94.32 [93.62, 95.12]
e	0.17 [0, 0.41]	0.11 [0, 0.28]	0.21 [0, 0.42]
K [m s^{-1}]	0.60 [0.30, 1.02]	0.89 [0.52, 1.32]	0.89 [0.42, 1.22]
ω [rad]	2.2 [0, 2π]	2.8 [0, 2π]	3.9 [0, 2π]
M_0 [rad]	3.4 [0, 2π]	3.9 [0, 2π]	5.1 [0, 2π]
$m_p \sin i$ [$M_\oplus$]	2.0 [1.0, 3.3]	4.0 [2.2, 5.7]	5.5 [2.6, 7.5]
a [AU]	0.106 [0.099, 0.110]	0.197 [0.184, 0.204]	0.379 [0.356, 0.393]
σ_J [m s^{-1}]	1.05 [1.01, 1.12]		

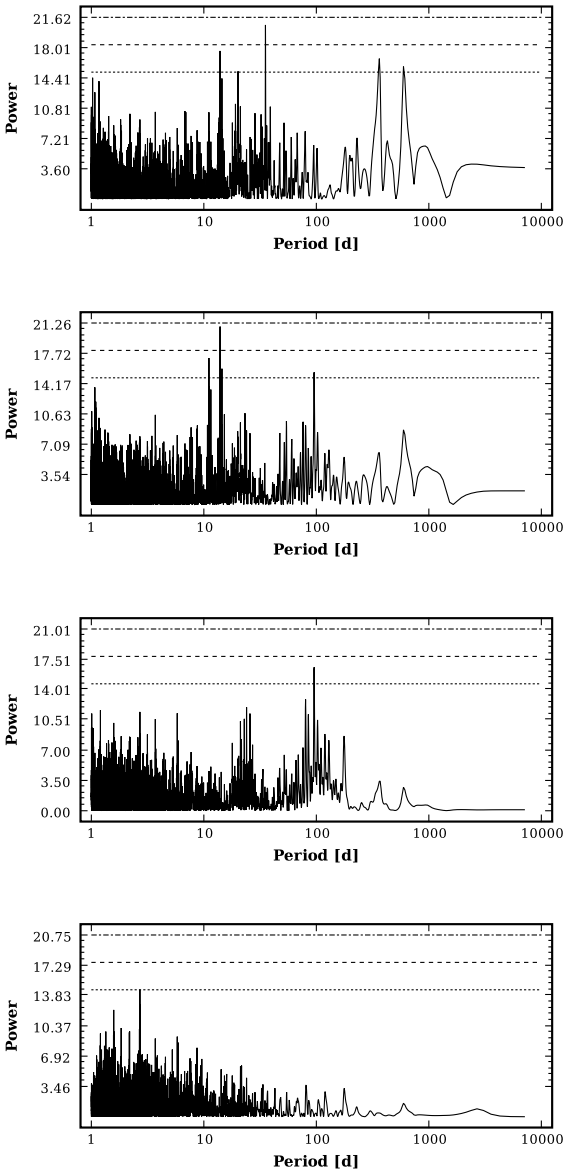

Fig. 5. Lomb-Scargle periodograms of the HARPS data residuals after removing moving average components from the noise (*top panel*) and removing the 35, 14, and 95 day signals (subsequent panels). The dotted, dashed, and dot-dashed lines indicate the analytical 0.1%, 1%, and 10% FAPs, respectively.

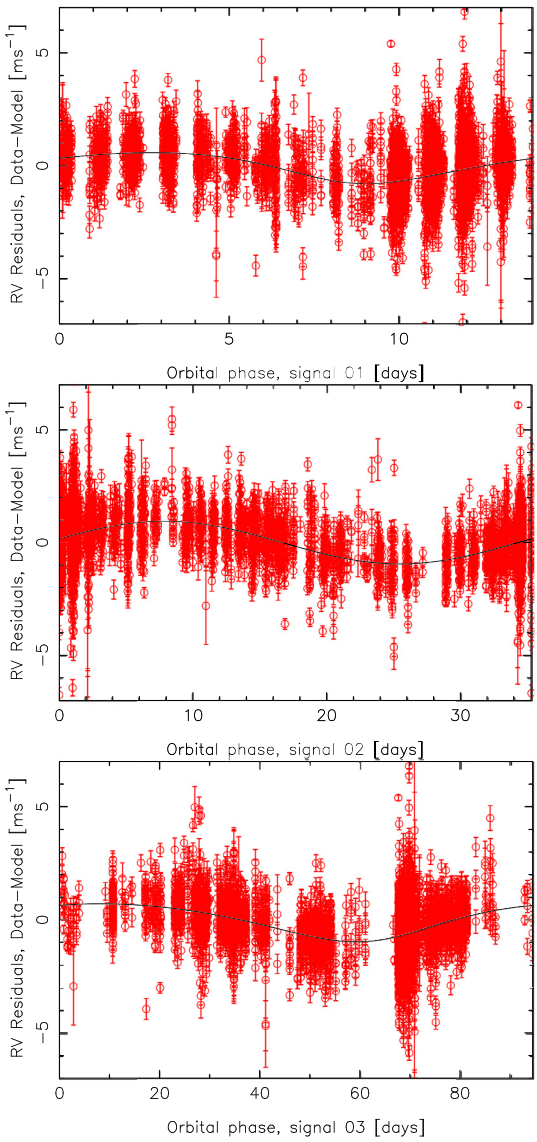

Fig. 6. Phase-folded signals of the three-Keplerian solution with the other two signals and the moving average component (MA(10)) of the noise removed.

seeing whether the same probability maxima existed in the period space. The exact amount of MA components was found to impact the resuts little and we could retrieve the three signals using models with less (5) or more (12) MA components. With the MA(3) model, however, the period space was found to contain much more maxima likely arising from noise correlations that were not accounted for in the model. We could also obtain two of the signals (at periods of 14 and 35 days) by using an AR(5)MA(5) model and observed a maximum periodogram power in the residual periodogram at 95 days but could not constrain the corresponding signal by samplings. We could also detect the three signals by using the flat "Gaussian" likelihood

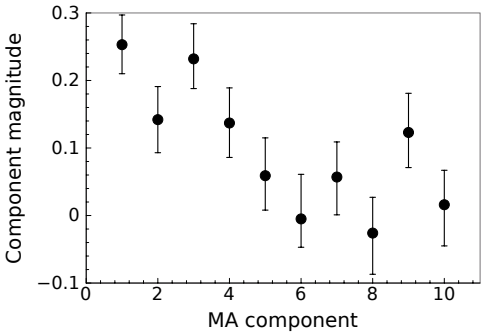

Fig. 7. Estimated MA components ϕ_i for the HARPS RVs.

Table 6. MAP estimates and the corresponding 99% BCSs of the MA(10) noise parameters.

Parameter	HD 10700 b
ω_1	0.212 [0.168, 0.261]
ω_2	0.120 [0.075, 0.170]
ω_3	0.230 [0.182, 0.282]
ω_4	0.150 [0.097, 0.208]
ω_5	0.071 [0.017, 0.126]
ω_6	0.016 [−0.043, 0.074]
ω_7	0.066 [0.007, 0.126]
ω_8	−0.027 [−0.092, 0.031]
ω_9	0.146 [0.080, 0.213]
ω_{10}	0.010 [−0.054, 0.073]

model in Eq. (8). These results indicate that the signals in the HARPS data are rather independent of the exact noise model.

7.3. Noise properties

The best noise model of the HARPS data was found to be the tenth order MA model with Gaussian white noise component. This model contains 13 free parameters, namely, the magnitude of the Gaussian white noise (σ_J), reference velocity about the data mean γ, a parameter describing the MA time-scale β, and ten MA components ω_j, $j = 1, ..., 10$. While the MA components ω_j received values between roughly 0.3 and 0.0 indicating that the dependence of the noise of the ith measurement on the 10 previous ones was only moderate at most, the time-scale parameter, with units of h^{-1}, received an interesting value close to unity with a MAP estimate of 1.19 h^{-1} and a 99% BCS of [0.98, 1.40] h^{-1}. This means that the noise correlations exist on the time-scale of an hour. Also, the MA components roughly decreased as a function of their order in a natural way (Fig. 7), which indicates that the correlations between the measurement errors were the greatest the closer these measurements were in time, both quantitatively and qualitatively. We show the MAP estimates of all the MA parameters in Table 6.

Another interesting parameter, the magnitude of Gaussian white noise, received a low MAP value of 1.06 $\mathrm{m\,s^{-1}}$ (with 99% BCS of $[1.02, 1.11]^{-1}$) This value is considerably lower than the original 1.7 $\mathrm{m\,s^{-1}}$ we obtained when modelling the noise as pure Gaussian white noise. This indicates that removing the MA components and the three signals from the data reduces the deviation of the data from the mean considerably. It also explains why the three signals can be detected in the data. The reason is that the amplitudes of these signals are only slightly below the noise magnitude enabling the detections, whereas they are considerably below the noise level when the noise correlations are not accounted for. Based on these results and the estimated detection thresholds from the analytical criterion in Eq. (10), we estimate that with the current precision of HARPS velocities it is already possible to detect signals with RV amplitudes of 0.3−0.5 $\mathrm{m\,s^{-1}}$ if the properties of the noise in the velocities are taken into account.

7.4. Analysis of partial HARPS data

To check the robustness of our solution to the HARPS RVs, we tested whether it was possible to receive consistent results with only part of the HARPS velocities.

As our partial data set, we chose the last 2763 HARPS velocities. This choice, while rather arbitrary, was made because it corresponds to excluding the first two observation periods distinguished clearly in Fig. 1 (top panel) because of the annual gaps. We decided to exclude the first two observational periods because the overall stability of HARPS has likely been improved after the first two years of operation. We therefore expect, not only that can we extract the same signals from this partial HARPS data, but that we might be able to constrain them better and see the noise levels in the data decrease because of a better stability of the partial data between epochs 3593 and 5082 [JD-2 450 000].

With this smaller dataset and the same procedures as in the previous section, we identified the same three signals that we found in the full HARPS data set. We could not find a clear probability maximum for a four-Keplerian model. This confirms that there are three statistically significant signals in the data. When we increased the number of Keplerian signals in the statistical model from zero to three, the posterior probabilities increased by factors of 3.9×10^{16}, 3.5×10^{10}, and 3.4×10^{9}, which indicates that these signals are again very significant. However, comparing these values to the factors received for the full data set indicates that, while they are very similar when moving from $k = 0$ to $k = 2$, the significance of the third signal decreases considerably for the partial HARPS data.

These results are consistent with the ones received for the full HARPS data set but we found a significantly different noise level in the partial HARPS data set. The parameter σ_J that describes the magnitude, i.e. standard deviation, of the Gaussian excess white noise in the data, received a MAP estimate of 0.82 $\mathrm{m\,s^{-1}}$ (with a 99% BCS of [0.78, 0.87] $\mathrm{m\,s^{-1}}$). This can be compared to the MAP estimate of 1.05 $\mathrm{m\,s^{-1}}$ for the full HARPS data set (Table 5). Thus, the first two observing periods indeed contain considerably more noise than the subsequent ones. This could be due to improvements in the instrumental stability of the HARPS spectrograph over the years, but we cannot know this for sure as it might also be caused by a more quiescent period of the target star with e.g. a lower amount of starspots. Either way, it looks like the first two observing periods are contaminated by a significantly greater amount of noise than the subsequent periods.

7.5. HARPS activity indicators

Our Bayesian analyses identify three clear periodic signals – possibly caused by periodic Doppler shifts – in the RV data acquired using HARPS, but the possibility of these signals being caused by activity-related phenomena on the stellar surface must be accounted for. Therefore, we extracted the HARPS S chromospheric activity indices using methods honed using other high

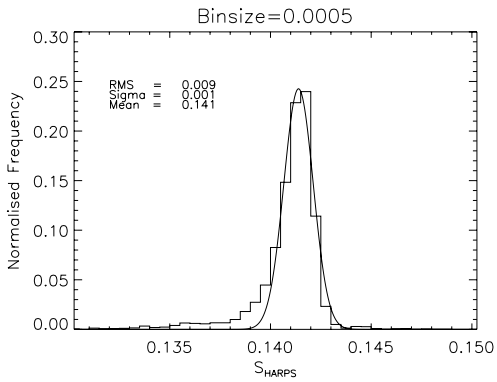

Fig. 8. Distribution of HARPS S-index. The solid curve indicates the fitted Gaussian curve.

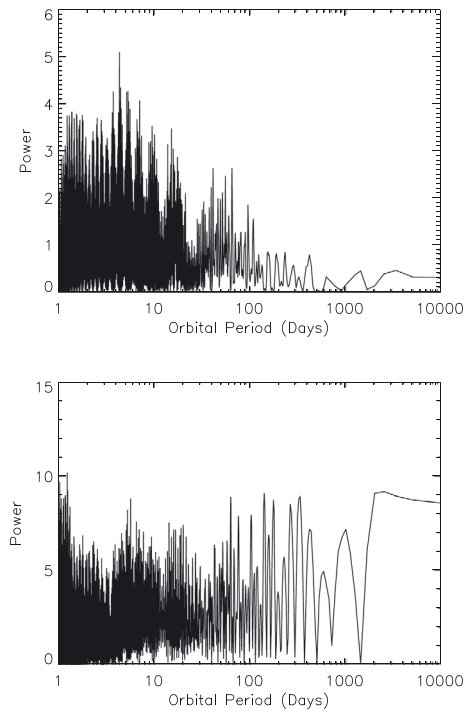

Fig. 9. Lomb-Scargle periodograms of the S-indices (*top panel*) and the BIS values (*bottom panel*).

resolution spectrographs (Jenkins et al. 2006, 2011; Tuomi et al. 2012) to search for periodicities in the activities, and/or correlations between the activity of HD 10700 and the RV signals we find.

The distribution of HARPS S-indices for HD 10700 was found to have an approximately Gaussian profile, as shown in Fig. 8. We can see that HD 10700 has a very tight spread of chromospheric activities which helps us to realise that this star is magnetically very stable. The standard deviation of the S-index distribution (Fig. 8) is only 0.001 dex, which is very stable in comparison to other typical G-dwarf stars. For instance, the Sun can exhibit activity index changes at the level of 0.005 dex over long timescales (Livingston et al. 2007). This suggests that HD 10700 is exhibiting some period of sustained magnetic stability, or its orientation in space is such that it appears from our vantage point that the spot patterns on the stellar surface are of very low number and do not change considerably over the baseline of the HARPS measurements.

Another feature worth noting is that the distribution of the S-index is well described by a Gaussian function except in the lower wing region. There appears to be a small excess of low activity values for HD 10700 that may require an additional modification of the best fit distribution. It may be that a double Gaussian is needed, similar to the distribution seen for HD 114613 (Tuomi et al. 2012) but at a much more reduced level. This may indicate that double Gaussian distributions of S-indices are common for F, G, and K dwarf and subgiant stars over these timescales of few years at these levels of precision.

The top panel in Fig. 9 shows a Lomb-Scargle periodogram analysis of the S-indices and there only appears to be one region in the period space that exhibits strong signals compared to the rest of the data. The strongest power in this periodogram has a period of 4.34 days and the eight strongest powers are found in a range between 3.7 and 5.5 days. Interestingly, we detected a weak signal, i.e. a clear probability maximum but not a statistically significant periodicity, in the partial HARPS data with a period of 3.70 days. Taking the features in the activity data into account, we expect that this short period signal is likely caused by activity-related features in the RV data and not by a genuine Doppler shift of planetary origin.

The bottom panel in Fig. 9 shows the periodogram for the bisector span (BIS) values of the HARPS data set. These BIS values were drawn directly from the HARPS-DRS analysis and details of their usefulness can be found in Santos et al. (2010). There are no significant periodicities in the BIS values data either. These results, i.e. the lack of significant periodicities in both, the S indices and BIS values, indicate that activity or line asymmetries are not the cause of the periodicities we detect in the HARPS RVs. Nevertheless, we should note that our 35.362 day signal is close to the 34 day period (Table 1) reported by Baliunas et al. (1996) and is something we discuss further in Sect. 11.

Since the signals we have discovered in the RVs are of very low amplitude we can investigate whether activity indicators, i.e. the S-index and the BIS value, correlate with the RV variations in the HD 10700 data. We calculated the phase-folded signals (as shown in Fig. 6) and searched for such correlations between each of the signals and the activity indicators. We could not find any strong linear correlations between these indices and any of the signals. None of these correlations were found to be significant and the corresponding Pearson correlation coefficients do not exceed ±0.08. This supports the argument that spot modulation, or some other periodic stellar phenomena that affects the stellar line profiles, is not the root cause of the periodicities that we find in the data for HD 10700, reinforcing the possibility that the three signals might be induced by planets orbiting the star.

8. AAPS and HIRES radial velocities

The AAPS and HIRES velocities had significantly more noise – roughly three times more – than the HARPS precision velocities, which suggests that detecting the same signals might not be possible from them. We started by finding the best noise model for the AAPS and HIRES RVs. We compared the performance of different MA models because the AR models could not be considered trustworthy descriptions of RVs based on the tests with artificially injected signals. According to our results, the fifth order MA model was the best description of the AAPS data and increasing the number of MA components did not result in a model with a greater posterior probability. For HIRES data, the first order MA model was already a reasonably accurate description and increasing the order above three resulted in only a minor

Table 7. Relative posterior probabilities of noise models for AAPS and HIRES RVs.

Model	$P(\mathcal{M}\|d)$ (AAPS)	$P(\mathcal{M}\|d)$ (HIRES)
G	7.8×10^{-104}	4.5×10^{-28}
G+MA(1)	2.7×10^{-28}	0.04
G+MA(3)	9.8×10^{-5}	0.24
G+MA(5)	~1	0.72

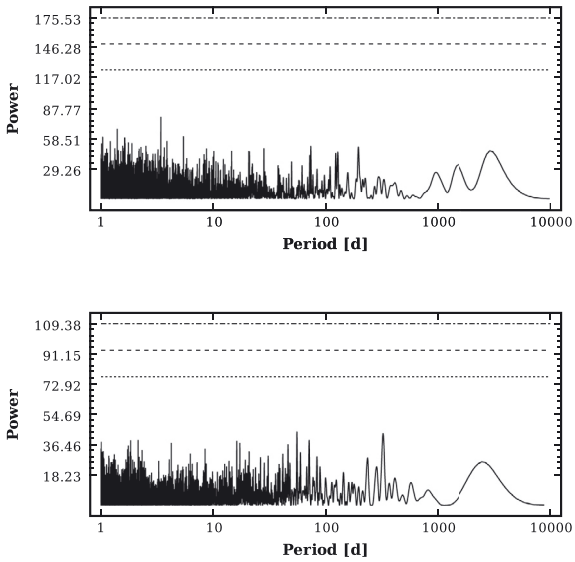

Fig. 10. Lomb-Scargle periodograms of the AAPS (*top*) and HIRES (*bottom*) data sets after removing the noise correlations from them. Dotted, dashed, and dash-dotted lines indicate the analytical 10%, 1%, and 0.1% FAPs, respectively.

improvement that we do not consider significant. The posterior probabilities of the statistical models are shown in Table 7 up to a fifth order MA model. Based on these posterior probabilities, we use the MA(5) and MA(3) models for the AAPS and HIRES data in the following analyses.

After removal of the identified noise components, we could neither identify any significant periodic signals in the AAPS and HIRES RVs nor significant powers in their periodograms (Fig. 10). Despite several samplings of the parameter space of a one-Keplerian model, our Markov chains did not converge to any periodicities for either data set. This indicates, that the signals detected from the HARPS velocities are below the detection thresholds of these two data sets and cannot be obtained from them.

The nature of the noise in the AAPS and HIRES data sets turned out to be different from that observed in the HARPS RV data. The noise in AAPS and HIRES data sets did not require as many MA components as HARPS data did and the time-scale of the correlations turned out to be different as well. We find that the α parameter in Eq. (2) received values of 0.0026 [0.0005, 0.0080] h^{-1} and 0.0244 [0.0006, 0.0740] h^{-1} for the AAPS and HIRES velocities, respectively. These values correspond to noise correlations on the time-scale of days to weeks, not the few hour time-scale of the HARPS data.

This different time-scale of noise correlations in the AAPS and HIRES data sets could be caused by e.g. instability of the instruments from night-to-night that exceeds any noise correlations on shorter time-scales. Indeed asterioseismology analyses suggest that the short-term performance of the spectrographs are relatively similar (Bedding et al. 2007). The HARPS spectrograph is shown to have good short-term stability (the reader is referred to the long series of publications titled "The HARPS search for southern extra-solar planets" of which two recent ones are: Dumusque et al. 2011; Ségransan et al. 2011), which could enable the detection of noise correlations over time-scales of a few hours – possibly arising from stellar surface phenomena – from HARPS data.

Whether caused by the telescope-instrument combinations or the surface of the stellar target, the noise correlations need to be taken into account when analysing AAPS and HIRES RVs. While their inclusion in the statistical model improves the performance of the model considerably (Table 7), it also enables us to remove these correlations from the data, which could reveal signals otherwise hidden in the noise. For a pure Gaussian noise model, the magnitude of the excess jitter was found to be 3.18 [2.90, 3.48] $\mathrm{m\,s^{-1}}$ and 2.45 [2.17, 2.74] $\mathrm{m\,s^{-1}}$ for the AAPS and HIRES data sets, respectively. These values decreased to 2.16 [1.93, 2.44] $\mathrm{m\,s^{-1}}$ and 2.11 [1.86, 2.42] $\mathrm{m\,s^{-1}}$ for the best MA models, which indicates that a considerable amount of noise that results from correlations was interpreted as Gaussian white noise in the pure Gaussian white noise model. Therefore, as was the case for the HARPS RV data, accurate noise modelling can improve the potential of the AAPS and HIRES programmes to find lower amplitude signals.

We note that the white noise component appeared to be close to Gaussian in the AAPS and HIRES data sets. The Cauchy model provided a much worse description of these data sets because they lack considerable outliers. Also, the noise model in Eq. (8) was worse than the pure Gaussian model because it has one extra parameter that appears to be unnecessary because the noise distribution does not have a flat maximum.

9. Combined radial velocities

We combined the HARPS, AAPS, and HIRES RV data in an attempt to see if the signals detected in the HARPS data alone could be extracted from this combined set as well, and especially, whether their significances would change with respect to those found in the HARPS data alone. For obvious reasons, the combined data set had a better phase coverage than any of the individual data sets and a baseline corresponding to that of the AAPS data of approximately 13.5 years. We modelled the noise properties of this combined set by using the best noise descriptions found for each data set. While the periodogram of the combined data did not show any signatures of periodic signals after removing the correlations in the noise, we performed samplings of models with k Keplerian signals, with $k = 1, 2, ...$, anyway.

Identifying the 35 day periodicity was again straightforward and the corresponding one-Keplerian model received a posterior probability that was 8.1×10^{13} times greater than that of the model without Keplerian signals (see Table 8). Similarly, adding more Keplerian signals in the statistical model we could identify the 13.9 and 94 day periodicities rather easily. Taking these periodicities into account increased the model probabilities further by factors of 2.6×10^{11} and 2.5×10^{15}, respectively, which means that the three Keplerian signals detected from the HARPS data alone were present in the combined data set as well with high significances. However, we note that with the three-Keplerian model, these three signals did not receive exactly the same parameter estimates as they did for the HARPS data alone. Their RV amplitudes received values of approximately 0.10 $\mathrm{m\,s^{-1}}$ lower than for the HARPS data and their eccentricities received slightly lower values better consistent with circular orbits. Yet, given their 99% BCSs, all parameter values

Table 8. Relative posterior probabilities and log-Bayesian evidences of models $\mathcal{M}_k$ with k Keplerian signals given the combined HARPS, AAPS, and HIRES RV data.

k	$P(\mathcal{M}_k\|d)$ TPM	$\ln P(d\|\mathcal{M}_k)$ TPM	$P(\mathcal{M}_k\|d)$ AICc	$\ln P(d\|\mathcal{M}_k)$ AICc	P_s [days]
0	5.0×10^{-53}	−10 689.4	6.5×10^{-49}	−10 6699.0	–
1	4.0×10^{-39}	−10 6656.7	2.3×10^{-36}	−10 6669.4	35.2
2	1.1×10^{-27}	−10 6629.7	2.3×10^{-26}	−10 6645.6	13.9
3	2.6×10^{-12}	−10 6593.6	1.4×10^{-11}	−10 6610.9	95.4
4	3.3×10^{-7}	−10 6581.1	8.3×10^{-6}	−10 6596.9	630
5	0.937	−10 565.6	~ 1	−10 584.5	168
5	0.063	−10 568.3	3.1×10^{-4}	−10 592.6	315

Notes. P_s denotes the MAP period estimate of the signal added to the solution.

were consistent with the estimates in Table 5. We note that these observed differences in the MAP estimates from HARPS data and the combined data could be caused by insufficient noise modelling.

We continued by sampling the parameter space of a four-Keplerian model, and despite the fact that none of the residual periodograms of the individual data sets or the full set showed any powers in addition to the three aforementioned signals, we discovered a significant probability maximum in the parameter space corresponding to another signal with a period of 630 days and a RV amplitude of 0.52 [0.24, 0.85] m s^{-1}. This signal satisfied all our detection criteria by making the four-Keplerian model 1.3×10^5 times more probable than the three-Keplerian one. Also, all the MCMC samplings we performed with different initial states converged to this same solution (Fig. 11) and we could safely conclude that the 630 days periodicity corresponds to a reasonably high probability maximum and is one of the periods in the four-Keplerian model. We note that this 630 day signal shows in the periodogram of the HARPS data as well before removing any periodicities from the data as a peak exceeding the 10% FAP (Fig. 5, top panel).

When sampling the parameter space of a five-Keplerian model, we chose the four-Keplerian one as a starting point and set the initial parameter values corresponding to the four signals close to their MAP estimates. However, the period of the fifth one was chosen randomly either between the 95 and 640 days or above 630, such that its velocity amplitude was set close to zero. This division of the period-space into two parts was made because we could then treat the corresponding models with the fifth periodicity in either subspace as separate models. We chose this division because our initial samplings indicated that both subspaces contained probability maxima and sampling a parameter space with well-separated modes is computaionally very demanding. The division enabled us to perform the corresponding samplings much more rapidly than performing the samplings of the full period space.

Samplings of the parameter space identified three additional probability maxima in the period space at roughly 170, 320, and 1300 days. These periods correspond to orbital distances where low-mass planets would likely remain on stable orbits because they retain a sufficient orbital spacing. Also, we note that the 320 day signal shows as a peak exceeding the 10% FAP in Fig. 5 (top panel) but because of the fact that this peak is so close to one year periodicity, which also shows in the window function as a result of HARPS data sampling (Fig. 2, bottom panel), we cannot conclude that it actually corresponds to a genuine signal.

According to our samplings of the parameter space of the five-Keplerian model, the probability maximum around

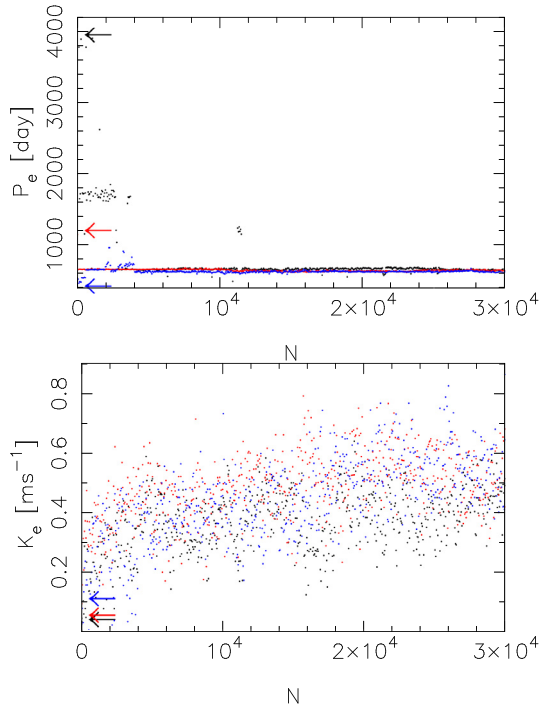

Fig. 11. Convergence of the parameters of the fourth Keplerian as a function of Markov chain length to a period of 630 days (*top*) and a RV amplitude of 0.56 m s^{-1} (*bottom*) for three different Markov chains (denoted using different colours). Arrows indicate the randomly selected initial values (K is chosen to be initially close to zero). The chains have been thinned by a factor of ten.

1300 days did not turn out to be significant. The amplitude of the signal corresponding to this probability maximum was not found distinguishable from zero, and we could conclude that the data, when interpreted using the five-Keplerian model, did not support the existence of a signal at or around 1300 days. Also, though there are other lower probability maxima beyond 1300 day period (especially around 2000 days) up to the maximum period in the parameter space of ten times the data baseline, none of them was found significant either.

Instead, the period space between 95 and 630 days provided some interesting solutions for the fifth signal. Especially, we found a global probability maximum for the fifth period at 168 days which increased the model probability of the five-Keplerian model a factor of 3.0×10^6 greater than that of the four-Keplerian one (Table 8). In addition to this solution, we identified another possible solution for the fifth signal at 315 days. This local solution was found to have a significantly

Table 9. Five-planet solution of HD 10700 RVs. MAP estimates of the parameters and their 99% BCSs.

Parameter	HD 10700 b	HD 10700 c	HD 10700 d
P [days]	13.965 [13.941, 13.982]	35.362 [35.256, 35.450]	94.11 [93.48, 94.81]
e	0.16 [0, 0.38]	0.03 [0, 0.31]	0.08 [0, 0.34]
K [m s^{-1}]	0.64 [0.40, 0.88]	0.75 [0.48, 0.99]	0.59 [0.34, 0.87]
ω [rad]	1.5 [0, 2π]	3.0 [0, 2π]	4.0 [0, 2π]
M_0 [rad]	2.6 [0, 2π]	3.2 [0, 2π]	5.8 [0, 2π]
t_0 [days]*	4.17 [−]	20.62 [−]	2.31 [−]
$m_p \sin i$ [$M_\oplus$]	2.0 [1.2, 2.8]	3.1 [2.0, 4.5]	3.6 [1.9, 5.3]
a [AU]	0.105 [0.099, 0.110]	0.195 [0.184, 0.204]	0.374 [0.354, 0.391]
	HD 10700 e	HD 10700 f	
P [days]	168.12 [166.01, 170.44]	642 [615, 679]	
e	0.05 [0, 0.27]	0.03 [0, 0.29]	
K [m s^{-1}]	0.58 [0.30, 0.86]	0.58 [0.26, 0.85]	
ω [rad]	5.5 [0, 2π]	3.9 [0, 2π]	
M_0 [rad]	0.5 [0, 2π]	1.6 [0, 2π]	
t_0 [days]*	37.42 [−]	168.49 [−]	
$m_p \sin i$ [$M_\oplus$]	4.3 [2.2, 6.3]	6.6 [3.1, 10.1]	
a [AU]	0.552 [0.522, 0.575]	1.35 [1.26, 1.43]	
$\sigma_{J,1}$ [m s^{-1}] (HIRES)	2.14 [1.92, 2.41]		
$\sigma_{J,2}$ [m s^{-1}] (AAPS)	2.13 [1.88, 2.41]		
$\sigma_{J,3}$ [m s^{-1}] (HARPS)	1.06 [1.02, 1.10]		

Notes. Parameter t_0 denotes the time of periastron. ($\star$) days after JD = 2 450 000 that was used as the zero-point in our analyses.

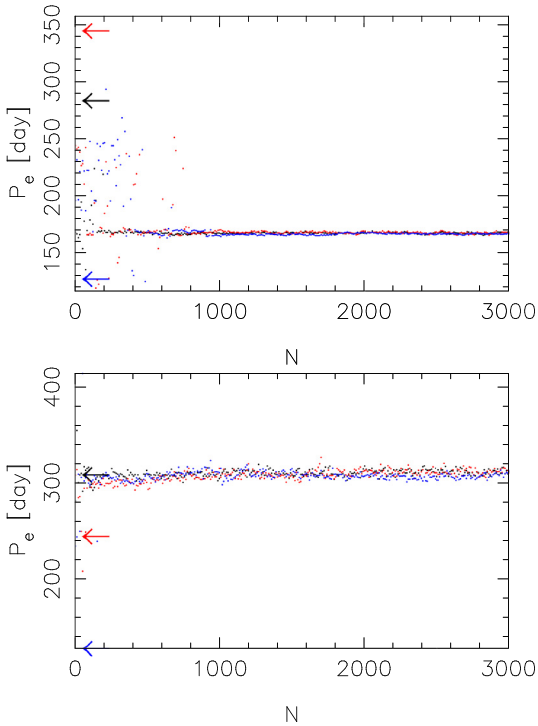

Fig. 12. Convergence of the period of the fifth Keplerian as a function of Markov chain length to periods of 168 (*top*) and 315 days (*bottom*). The details of the panels are as in Fig. 11 but the chains have been thinned by a factor of ten.

greater probability than the four-Keplerian model did, but it was not as probable as the global solution corresponding to the 168 day periodicity as a fifth signal (Table 8). Yet, both these solutions satisfied all the detection criteria by having RV amplitudes statistically significantly greater than zero and by having well-constrained periods and the Markov chains converged to either one of them very rapidly regardless of the initial choice of the fifth period (Fig. 12). We note that 168 and 315 day periodicities are actually each other's one-year aliases. We could easily verify this by sampling a six-Keplerian model with these two periodicities as initial states of the periods of two signals. These samplings did not converge to two signals but altered between either of the periodicities in the sense that the RV amplitudes of the 168 and 315 day signals had a negative correlation with either reaching a maximum when the other approached zero.

The five-Keplerian solution with the 168 day periodicity is clearly favoured by the data according to the model probabilities in Table 8. However, the solution with the 315 day periodicity still has a considerable probability of 0.06 (though it is not significant based on the AICc), which means that we cannot completely rule out this alternative solution. Because of this alternative solution and a local maximum in the period space of the fifth Keplerian signal at roughly 1300 days, we carried on sampling the parameter space of a six-Keplerian model further.

The global solution of the six-Keplerian model was found to correspond to periodicities at 14, 35, 94, 168, 640, and 1300 days. However, despite the fact that the global solution contained the 1300 day signal, this signal was found to have an amplitude that was not statistically significantly different from zero. Also, We could not identify a solution where the sixth signal had a period between the 168 and 640 day ones. In the six-Keplerian solution, the periodicity of 1300 days was well constrained from above and below, but the RV amplitude of this signal was only barely constrained with a MAP estimate of 0.36 m s^{-1} and a 99% BCS of [0.00, 0.66] m s^{-1}, which demonstrates that there is a fair chance that this signal in fact has a negligible amplitude, i.e. there is no evidence in favour of its existence. Because we could not distinguish this signal in the six-Keplerian model from zero, we did not calculate the probability for this model because we could not be sure whether the corresponding Markov chains had converged.

To make sure that we did not miss any additional signals, we performed samplings of the parameter space of a seven-Keplerian model. These samplings did not help identifying any other significant probability maxima in the parameter space. Therefore, we conclude that the best description of the combined velocities of HD 10700 contains five Keplerian signals. We

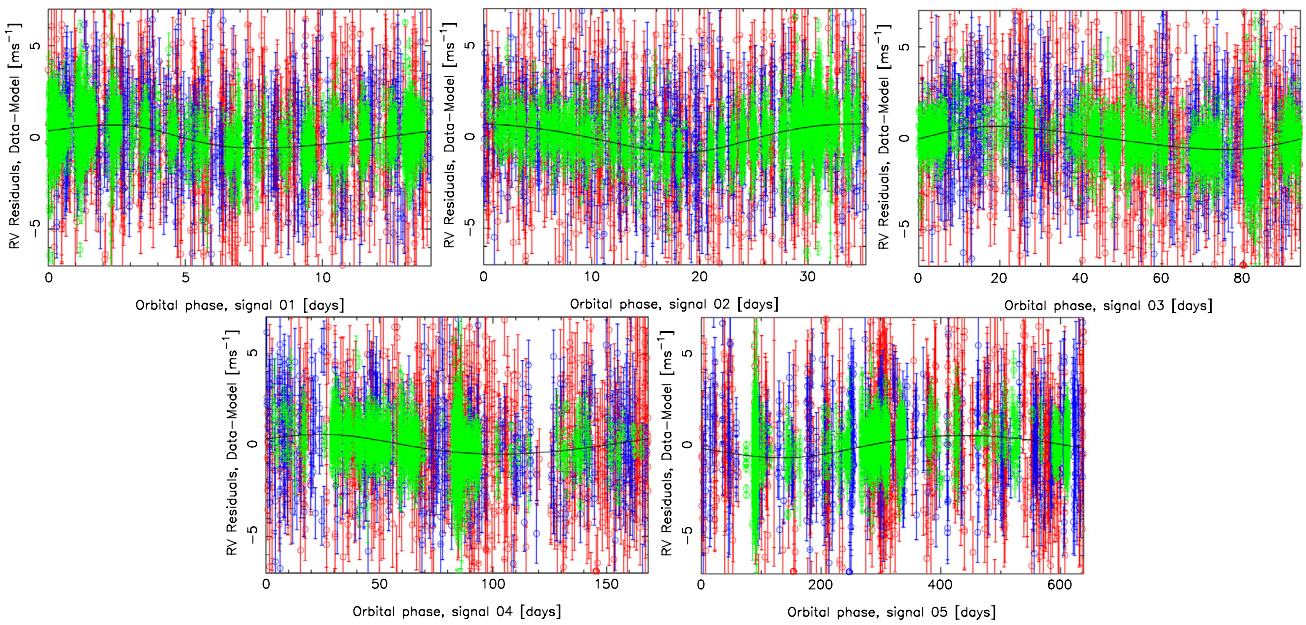

Fig. 13. Phase-folded Keplerian signals in the combined HARPS (green), AAPS (red), and HIRES (blue) RVs of HD 10700. Vertical axis is scaled to the interval of [−7, 7] m s^{-1} for clarity and the RVs deviating more than that from zero are therefore not shown.

present this orbital solution in Table 9 and show the phase-folded signals in Fig. 13. We also show the distributions of orbital periods, RV amplitudes, and eccentricities in Fig. 14 to demonstrate that the periods and amplitudes are well constrained and, especially, fully comply with our detection criteria. To visualise the significance of the signals, we plotted the phase-folded orbits by dividing the phase of each signal into 200 bins and by calculating the means and standard deviations for each bin. These plots are shown in the Appendix A for the combined data set and for the HARPS data alone (Figs. A.1 and A.2). When comparing these two figures, it can also be seen how the phase coverages of the signals are improved when using the combined data set instead of the HARPS data alone.

The five signals we observe in the combined data of HD 10700 are all consistent with circular orbits and have MAP estimated amplitudes below 1 m s^{-1}. Despite these low amplitudes, their detection is robust in the sense that (1) their inclusion in the statistical model increases the model probabilities significantly, (2) they have well-constrained periods and amplitudes that differ statistically from zero. We have shown in Sect. 7.5 that these signals do not have counterparts in the S and BIS activity indicators. It also seems unlikely that they arise from data sampling since the phase-folded RV curves have a good phase-coverage (Fig. 13) and the signals are constrained reasonably accurately (Fig. 14 and Table 9).

9.1. Signals in periodograms

To examine whether the three most significant signals are indeed present in the data and whether the two longer periodicities of 168 and 640 days are distinguishable from the data, we subtracted the MA components of the best model in our analyses from the combined data and analysed the residuals using the Lomb-Scargle periodograms. The sequence of residual periodograms is shown in Fig. 15.

In Fig. 15, the top left panel shows the residual periodogram after removing the MA components from the data. This periodogram shows several features exceeding the 0.1% FAP level. The highest peak corresponds to a periodicity at 35 days and has an analytical FAP of 2.2×10^{-69}. Similarly, after removing the most significant periodicities and again calculating the residual periodograms the one- and two-signal residuals show power at 13.9 and 95 days (left middle and bottom panels) exceeding the 1.5×10^{-27} and 1.4×10^{-19} FAPs, respectively. These three signals correspond to the ones detected from the HARPS data alone.

In Fig. 15(right top panel) we calculate the residual periodogram of the three-signal model.This periodogram has two significant powers exceeding the 1% FAP level at 15 and 168 days. Because we did not find signals at 15 days in our analyses, we adopted the 168 day periodicity as a fourth signal and carried on by calculating the residual periodogram of the four-signal model (Fig. 15, right top panel). Removing the 168 day signal also removed the peak at 15 days, which indicates that the latter was an alias of the former. The periodogram shows the 640 day periodicity as a last significant power (with a FAP of 2.5×10^{-3}, Fig. 15 right middle panel). Based on these results, the periodogram analyses support the results of the Bayesian analyses that there are five periodic signals in the combined RV data of HD 10700 on the basis that the noise in the data is modelled by using the MA models and Gaussian white noise.

10. Dynamical stability of the Keplerian solution

The discovery of these periodic signals has only been possible as a result of noise modelling and indeed we can not be confident of this procedure because we cannot find the signals independently in the different datasets. We nonetheless consider it instructive to assess whether these periodic signals correspond to a planetary system that is dynamically stable in long-term. Thus, following Tuomi (2012), we investigated the stability of the planetary system corresponding to the five signals by plotting the approximate Lagrange stability thresholds for each subsequent pair of planets in the system (Barnes & Greenberg 2006). According to our results shown in Fig. 16, the system appears to be long-term stable if the orbital eccentricities are not much greater than their MAP estimates (see Table 9). Figure 16 also indicates that there

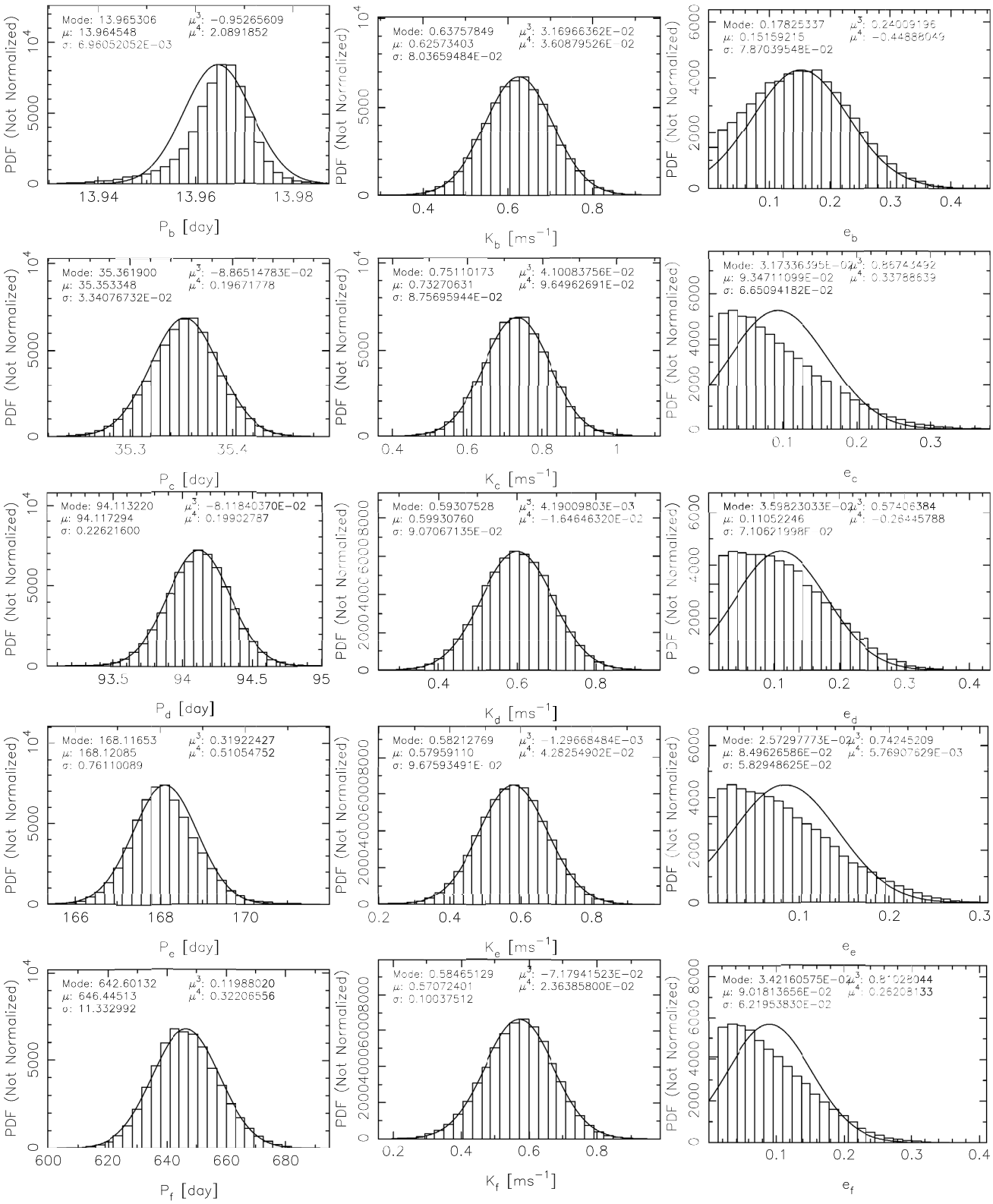

Fig. 14. Distributions estimating the posterior densities of orbital periods (P_x), eccentricities (e_x), and RV amplitudes (K_x) of the five Keplerian signals. The solid curve is a Gaussian density with the same mean (μ) and variance (σ^2) as the parameter distribution has. Additional statistics, mode, skewness (μ^3) and kurtosis (μ^4) of the distributions are also shown.

are orbital distances where additional low-mass planets could remain on stable orbits.

We tested the predictions of the stability thresholds by assuming that the putative planets in the HD 10700 system have co-planar orbits. With this assumption, we increased the masses corresponding to the inclination of the orbital plane approaching zero. According to our results, if the eccentricities of the five planets are at or below their MAP estimates, the stability thresholds are still in favour of a dynamically stable system when the inclination of the orbital plane is below 3°. This corresponds

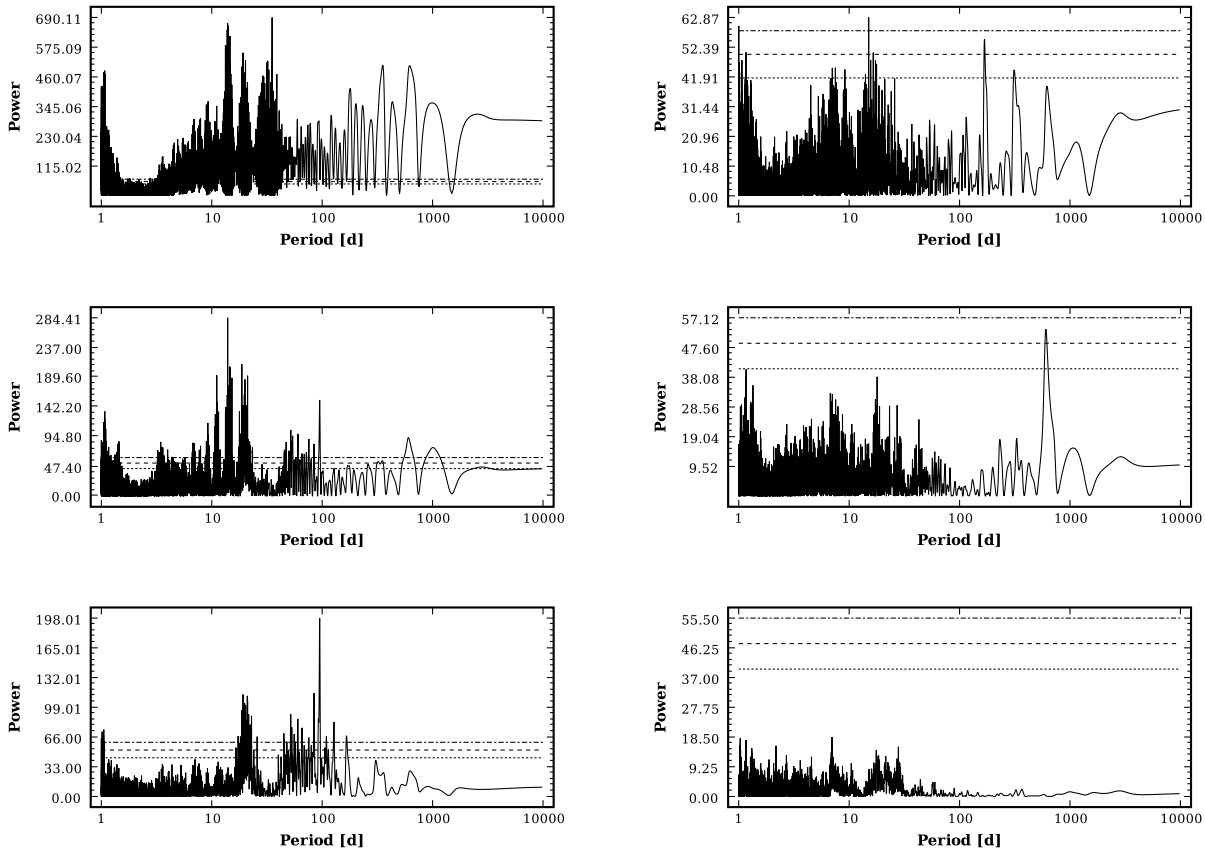

Fig. 15. Lomb-Scargle periodograms of the combined data residuals after removing MA components from the noise (*top left panel*) and removing the 35 (*middle left panel*), 14 (*bottom left panel*), 95 (*top right panel*), 168 (*middle right panel*), and 640 day signals (*right bottom panel*). The dotted, dashed, and dot-dashed lines indicate the analytical 10%, 1%, and 0.1% FAPs, respectively.

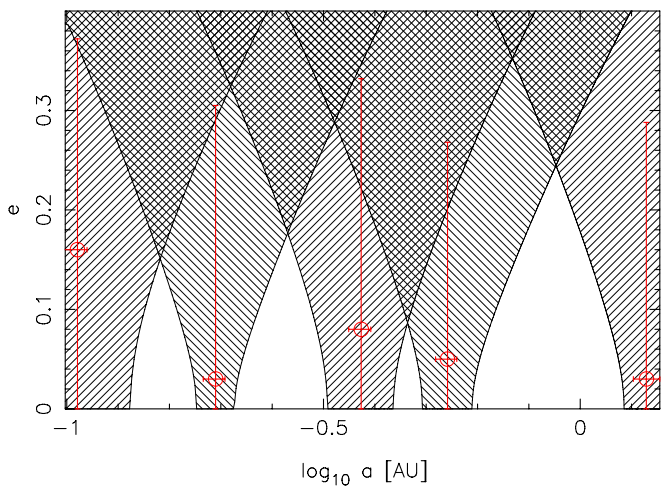

Fig. 16. Approximate Lagrange stability thresholds calculated for the HD 10700 signals. The hatched areas surrounding the orbital MAP estimates (red circles) indicate the parameter space where the neighbouring planets would make the system unstable.

to actual masses of roughly 40 times the minimum masses reported in Table 9. Therefore, the orbital spacing of the putative five-planet system around HD 10700 enables almost any orbital inclinations, and thus the system can still be dynamically stable on long timescales for masses considerably greater than the minimum masses. We note that full-scale numerical integrations of the orbits of the putative planets are needed to verify the above results and are beyond the scope of this paper.

11. Discussion

Signals with amplitudes lower than 1 m s^{-1} have been reported in quiescent HARPS targets such as HD 20794 and HD 85512 (Pepe et al. 2011). The lowest reported RV amplitude in the literature appears to be the 0.56 m s^{-1} one of HD 20794 c (Pepe et al. 2011). In practice, such accuracy corresponds to being able to find habitable zone super-Earths and hot rocky planets even less massive than the Earth orbiting nearby Solar-type stars.

In the current work, we have considered several noise models for high-precision velocities using a large data set for HD 10700 from the HARPS, AAPS and Keck exoplanet surveys. Our tests indicate that noise models containing a moving average component and assuming Gaussian white noise is the most effective modelling strategy for such data. Our simulations of the HARPS data for HD 10700 indicate that even the recovery of signals with amplitudes of 0.3 m s^{-1} with a period of 200 days is possible. Our tests also indicate that noise modelling can greatly improve the sensitivity of RV searches to low-amplitude signals. For instance, while the practise of nightly binning might remove correlations from the noise and make the excess noise appear "whiter", it also artificially decreases the amount of data and may lead to the undesirable side-effect that low-amplitude signals cannot be detected in the data after all.

We find that after removal of the moving average noise components in the HARPS data for HD 10700, three signals can be discerned. Although, these signals can not be confirmed independently in the AAPS and Keck datasets, they are found confidently in the combined HARPS, AAPS, and HIRES data set as well. These three signals can also be found when using slightly

modified (sub-optimal w.r.t. the preferred noise model of the HARPS data with MA(10) term and Gaussian white noise component) noise models, such as the MA(5) and MA(7). Also, two additional signals become statistically significant in the combined dataset. These signals discerned with Bayesian techniques and periodogram analysis are found at periods of 13.9, 35.4, 94, 168, and 630 days and do not have activity-induced counterparts based on the S and BIS indices from HARPS. The strongest of these signals at 35.4 days is close to the 34 day rotational period of the star which is a cause for concern. However, in addition to the lack of activity-induced signals near the rotation period, we also note that the period distribution in the posterior densities gives $\sigma \sim 0.0334$ (see Fig. 14) for the 35.362 day signal. If this signal was indeed induced by stellar rotation and activity, we would expect it to receive a much larger uncertainty under the reasonable assumption that a magnetic cycle and differential rotation were present. Any spot induced RV signal would be expected to change periodicity depending on its emergent latitude (which varies between 0 and 40 degrees on the Sun during an 11 year magnetic cycle), due to the latitude dependent differential rotation. We note that the HARPS data set has a baseline of 5.9 years (13.5 years for the combined data set), a significant proportion of the *Solar* activity cycle. The degree of rotational shear can be written as $\Omega(\theta) = \Omega_{\rm eq} - \Delta\Omega \sin^2(\theta)$, where θ is the stellar surface latitude, Ω the angular velocity, and Ω_{eq} the equatorial angular velocity. Following Barnes et al. (2005), who found that $\Delta\Omega$ scales with stellar mass but is only weakly dependent on rotation rate, we obtain $\Delta\Omega = 0.06$ rad day^{-1} or equivalently, $\Delta P = 11.04$ days. This represents the maximum rotation period variation (i.e. between equator and polar regions), but for the solar-like case with spots limited to 0 and 40 degrees we expect to measure a rotation period variation of $11.04 \sin^2(40) = 4.56$ days. This would lead to a considerably larger σ in the posterior density distribution than we find, even if we take into consideration that the HD 10700 data only span the equivalent of approximately half a solar cycle where the maximum variation on latitude might be closer to 3.3 days. This is still an order of magnitude greater than the uncertainty in the obtained period. In other words, our posterior period distribution width for the 35.4 day signal is *not* consistent with an activity induced signal that a solar-like star might be expected to produce.

Similar arguments can be applied to the constrained nature of the other signals in the posterior densities (see Fig. 14) and thus we do not consider that rotation and spots are a ready explanation for the signals. Thus it is plausible that the signals are caused by a system of five planets with minimum masses of 2.0, 3.1, 3.6, 4.3, and 6.6 $M_\oplus$, respectively. Indeed, such a system would be dynamically stable. However, before going farther with a planetary explanation, it should be emphasised that (1) the periodic signals that we are discussing appear following the removal of the "moving average" noise apparent in the data and with the assumption of Gaussian errors and that (2) with the available data we cannot discover any of the signals independently in more than one different dataset. It is worth noting that all the signals have similar MAP amplitudes between 0.58 and 0.75 m s^{-1}. Given the nature of these signals appearing after removal of noise it is unsurprising that they all have low amplitudes. However, for a planetary interpretation it might require some coincidence that the planets in this system have such an arrangement that the masses and periods conspire to correspond to similar Doppler amplitudes. Although, our simulations indicate that the recovery of such low-amplitude signals is not particularly dependent on the details of the moving average noise removal, the actual recovered amplitude of such small signals in our simulations might be overestimated. Thus, care must be taken in the interpretation of the nature of the signals.

We note that there are many possible ways to model the noise in radial velocity data and that in this work we have only explored a few of them. Indeed, the signals we detect may also result from the combination of insufficient noise modelling and our lack of understanding of stellar physics: asteroseismology, starspots, magnetic cycles, granulation and other phenomena of stellar surfaces.

In the future, we hope to expand the quantity of high precision radial velocity data of HD 10700 and to consider a wider range of noise models and incorporate an improved understanding of stellar phenomena. While our noise model choices serve to account for asteroseismological noise on timescales at or near the exposure times they do not necessarily correspond physically to e.g. granulation noise in the data. Comparisons of a more extensive collection of statistical models taking into account non-white features in the RV noise are needed to verify or falsify the existence of the signals we report in this work. Also, future data will be of essence in determining the nature of the signals we detect. If the five periodicities can be detected independently in different datasets, their genuine nature as signals of stellar origin will be verified. While even this will not imply that they are definitely Doppler signatures of low-mass planets, it will help ruling out spurious periodicities that insufficient modelling and instrumental instability might cause.

With a distance of only 3.7 pc, HD 10700 is the third closest star reported to be a host to a putative planetary system after ϵ Eridani (Hatzes et al. 2000) with a distance of 3.2 pc and α Centauri B (Dumusque et al. 2012) with a distance of 1.3 pc, though both of these remain to be confirmed and Zechmeister et al. (2013) have casted considerable doubt on the existence of a planet around ϵ Eridani. This makes HD 10700 an ideal target for future direct-imaging missions. The signals we find, which suggests the presence of low-mass planets, are consistent with both current theoretical models for low-mass planet formation and extant observational evidence for the presence of low-mass planets in the immediate Solar neighbourhood. Assuming the signals are indeed of planetary origin, the orbital periods of the two innermost candidates appear to suggests a 5:2 commensurability that would likely enable the long-term stability of the system (see e.g. a similar stabilising resonance of NN Serpentis Horner 2012). Given this assumption, we also note that a candidate corresponding to the signal with a period of 168 days would have an orbit inside the liquid water habitable zone as defined by (Selsis et al. 2007). However, these issues remain merely speculative until the planetary origin of the signals can be verified by an independent detection.

Acknowledgements. M.T. is supported by RoPACS (Rocky Planets Around Cool Stars), a Marie Curie Initial Training Network funded by the European Commission's Seventh Framework Programme. J.S.J. acknowledges funding by Fondecyt through grant 3110004 and partial support from Centro de Astrofísica FONDAP 15010003, the GEMINI-CONICYT FUND and from the Comité Mixto ESO-GOBIERNO DE CHILE. S. Vogt gratefully acknowledges support from NSF grant AST-0908870. J.H. and C.G.T. gratefully acknowledge the financial support of the Australian government through ARC Grant DP0774000. The authors acknowledge the significant efforts of the HARPS-ESO team in improving the instrument and its data reduction pipelines that made this work possible. We also acknowledge the efforts of all the individuals that have been involved in observing the target star with HARPS, Keck, and the AAT. The work herein is based in part on observations obtained at the W. M. Keck Observatory, which is operated jointly by the University of California and the California Institute of Technology. Finally, the authors would like to thank the anonymous referee for comments and suggestions that enabled considerable improvements of the manuscript.

Appendix A: Phase-folded orbits

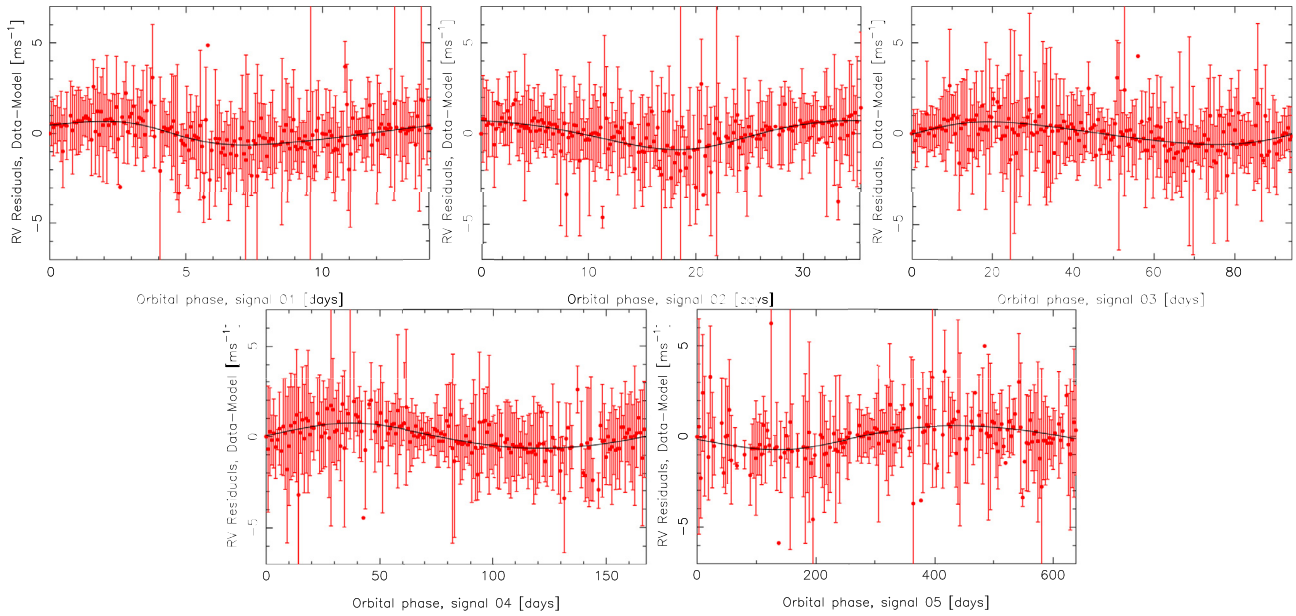

Fig. A.1. Phase-folded Keplerian signals for the combined HARPS, HIRES, and AAPS data. Orbital phase of each signal is divided into 200 bins and the means and the corresponding standard deviations of the data in each bin are plotted together with the Keplerian signal.

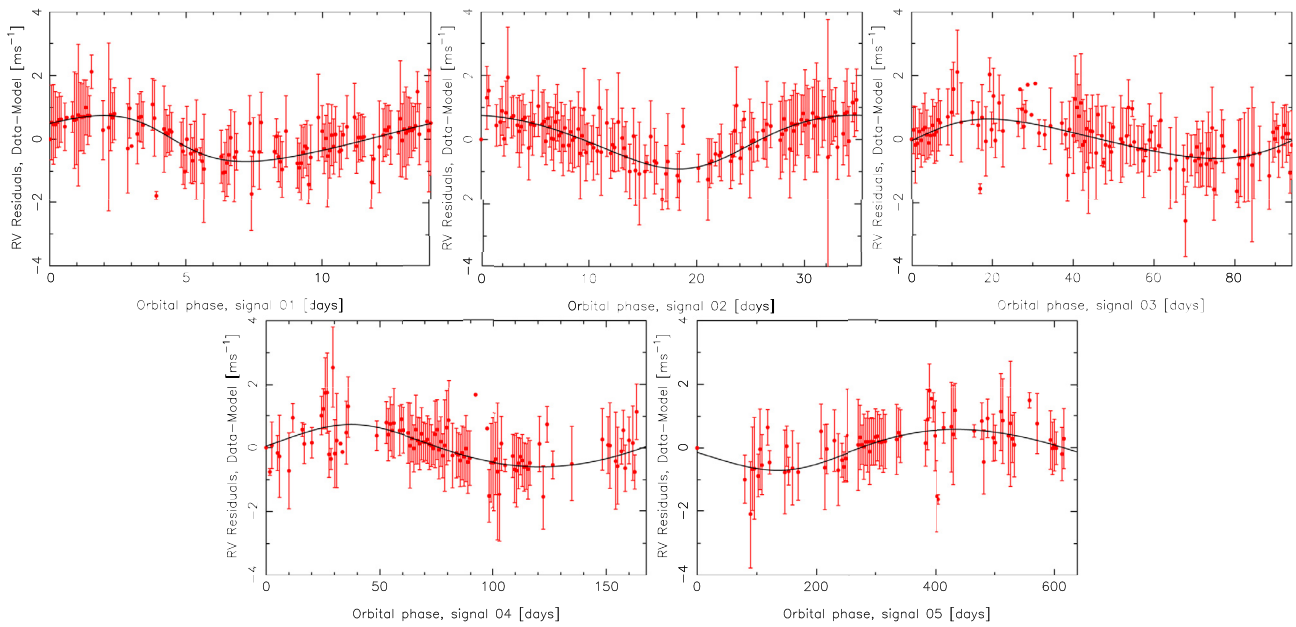

Fig. A.2. As in Fig. A.1 but for the high-precision HARPS data alone.

References

Akaike, H. 1973, in Second International Symposium on Information Theory. Akadémiai Kiadó, eds. B. N. Petrov, F. & Csaki, 267, 1
Baliunas, S., Sokoloff, D., & Soon, W. 1996, ApJ, 457, L99
Barnes, R., & Greenberg, R. 2006, ApJ, 647, L163
Barnes, J. R., Collier Cameron, A., Donati, J.-F., et al. 2005, MNRAS, 357, L1
Barnes, J. R., Jenkins, J. S., Jones, H. R. A., et al. 2012, MNRAS, 424, 59 1
Bedding, T. R., Kjeldsen, H., Arentoft, T., et al. 2007, ApJ, 663, 1315
Berger, J. O. 1980, Statistical Decision Theory and Bayesian Analysis (Springer)
Burnham, K. P., & Anderson, D. R. 2002, Model selection and multimodel inference: A practical information-theoretic approach (Springer-Verlag)
Di Folco, E., Absil, O., Augereau, J.-C., et al. 2007, A&A, 475, 243
Dumusque, X., Lovis, C., Ségransan, D., et al. 2011, A&A, 535, A55
Dumusque, X., Pepe, F., Lovis, C., et al. 2012, Nature, 491, 207
Feroz, F., Balan, S. T., & Hobson, M. P. 2011, MNRAS, 415, 3462
Ford, E. B., & Gregory, P. C. 2007, Statistical Challenges in Modern Astronomy IV, eds. G. J. Babu, & E. D. Feigelson, ASP Conf. Ser., 371, 189
Gautier III, T. N., Charbonneau, D., Rowe, J. F., et al. 2012, ApJ, 749, 15
Gray, D. F., & Baliunas, S. L. 1994, ApJ, 427, 1042
Gray, R. O., Corbally, C. J., Garrison, R. F., et al. 2006, AJ, 132, 161
Greaves, J. S., Wyatt, M. C., Holland, W. S., & Dent W. R. F. 2004, MNRAS, 351, L54
Gregory, P. C. 2005, ApJ, 631, 1198
Gregory, P. C. 2007a, MNRAS, 381, 1607
Gregory, P. C. 2007b, MNRAS, 374, 1321
Gregory, P. C. 2011, MNRAS, 415, 2523
Haario, H., Saksman, E., & Tamminen, J. 2001, Bernoulli, 7, 223
Hastings, W. 1970, Biometrika, 57, 97
Hatzes, A., Cochran, W. D., McArthur, B., et al. 2000, ApJ, 544, L145
Horner, J., Wittenmyer, R. A., Hinse, T. C., & Tinney, C. G. 2012, MNRAS, 425, 749

Howard, A. W., Marcy, G. W., Johnson, J. A., et al 2010, Science, 330, 653
Jenkins, J. S., Jones, H. R. A., Tinney, C. G., et al. 2006, MNRAS, 372, 163
Jenkins, J. S., Murgas, F., Rojo, P., et al. 2011, A&A, 531, A8
Jenkins, J. S., Jones, H. R. A., Tuomi, M., et al. 2013, ApJ, accepted [arXiv:1207.1012]
Judge, P. G., Saar, S. H., Carlsson, M., & Ayres, T. R. 2004, ApJ, 609, 392
Kass, R. E., & Raftery, A. E. 1995, J. Am. Stat. Ass., 430, 773
Lissauer, J. J., Fabrycky, D. C., Ford, E. B., et al. 2011, Nature, 470, 53
Livingston, W., Wallace, L., White, O. R., & Giampapa, M. S. 2007, ApJ, 657, 1137
Lomb, N. R. 1976, Ap&SS, 39, 447
Lovis, C., Ségransan, D., Mayor, M., et al. 2011, A&A, 528, A112
Mamajek, E. E., & Hillebrand, L. A. 2008, ApJ, 687, 1264
Mayor, M., Pepe, F., Queloz, D., et al. 2003, ESOMe, 114, 20
Mayor, M., Bonfils, X., Forveille, T., et al. 2009, A&A, 507, 487
Mayor, M., Marmier, M., Lovis, C., et al. 2011, A&A, submitted [arXiv:1109.2497]
Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., et al. 1953, J. Chem. Phys., 21, 1087
O'Toole, S. J., Tinney, C. G., Jones, H. R. A., et al. 2009a, MNRAS, 392, 641
O'Toole, S. J., Jones, H. R. A., Tinney, C. G., et al. 2009b, ApJ, 701, 1732
Pavlenko, Ya. V., Jenkins, J. S., Jones, H. R. A., et al. 2012, MNRAS, 422, 542
Pepe, F., Mayor, M., Galland, F., et al. 2002, A&A, 388, 632
Pepe, F., Lovis, C., Ségransan, D., et al. 2011, A&A, 534, A58
Santos, N. C., Mayor, M., Naef, D., et al. 2002, A&A, 392, 215
Santos, N. C., Israelian, G., García-López, R. J., et al. 2004, A&A, 427, 1085
Santos, N. C., Gomes da Silva, J., Lovis, C., & Melo, C. 2010, A&A, 511, A54
Scargle, J. D. 1982, ApJ, 263, 835
Ségransan, D., Mayor, M., Udry, S., et al. 2011, A&A, 535, A54
Selsis, F., Kasting, J. F., Levrard, B., et al. 2007, A&A, 476, 1373
Soubiran, C., Katz, D, & Cayrel, R., 1998, A&AS, 133, 221
Teixeira, T. C., Kjeldsen, H., Bedding, T. R., et al. 2009, A&A, 494, 237
Tuomi, M. 2011, A&A, 528, L5
Tuomi, M. 2012, A&A, 543, A52
Tuomi, M., & Jones, H. R. A. 2012, A&A, 544, A116
Tuomi, M., & Kotiranta, S. 2009, A&A, 496, L13
Tuomi, M., Kotiranta, S., & Kaasalainen, M. 2009, A&A, 494, 769
Tuomi, M., Pinfield, D., & Jones, H. R. A. 2011, A&A, 532, A116
Tuomi, M., Jones, H. R. A., Jenkins, J. S., et al. 2012, MNRAS, submitted
van Leeuwen, F. 2007, A&A, 474, 653
Vogt, S. S., Allen, S. L., Bigelow, B. C., et al. 1994, Proc. SPIE, 2198, 362
Wittenmyer, R. A., Endl, M., Cochran, W. D., et al. 2006, AJ, 132, 177
Wittenmyer, R. A., Tinney, C. G., Butler, R. P., et al. 2011a, ApJ, 738, 81
Wittenmyer, R. A., Tinney, C. G., O'Toole, S. J., et al. 2011b, ApJ, 727, 102
Wright, J. T. 2005, PASP, 117, 657
Zechmeister, M., Kürster, M., Endl, M., et al. 2013, A&A, in press DOI: 10.1051/0004-6361/201116551