

A Bayesian approach to the modelling of α Cen A

M. Bazot,^{1*} S. Bourguignon² and J. Christensen-Dalsgaard³

¹*Centro de Astrofísica da Universidade do Porto, Rua das Estrelas, 4150-762 Porto, Portugal*

²*Institut de Recherche en Communication et Cybernétique de Nantes, 1 rue de la Noë, BP 92101, 44321 Nantes Cedex 3, France*

³*Stellar Astrophysics Centre, Department of Physics and Astronomy, Aarhus University, Ny Munkegade 120, DK-8000 Aarhus C, Denmark*

Accepted 2012 July 28. Received 2012 July 27; in original form 2012 July 2

ABSTRACT

Determining the physical characteristics of a star is an inverse problem consisting of estimating the parameters of models for the stellar structure and evolution, and knowing certain observable quantities. We use a Bayesian approach to solve this problem for α Cen A, which allows us to incorporate prior information on the parameters to be estimated, in order to better constrain the problem. Our strategy is based on the use of a Markov chain Monte Carlo (MCMC) algorithm to estimate the posterior probability densities of the stellar parameters: mass, age, initial chemical composition, etc. We use the stellar evolutionary code *ASTEC* to model the star. To constrain this model both seismic and non-seismic observations were considered. Several different strategies were tested to fit these values, using either two free parameters or five free parameters in *ASTEC*. We are thus able to show evidence that MCMC methods become efficient with respect to more classical grid-based strategies when the number of parameters increases. The results of our MCMC algorithm allow us to derive estimates for the stellar parameters and robust uncertainties thanks to the statistical analysis of the posterior probability densities. We are also able to compute odds for the presence of a convective core in α Cen A. When using core-sensitive seismic observational constraints, these can rise above ~ 40 per cent. The comparison of results to previous studies also indicates that these seismic constraints are of critical importance for our knowledge of the structure of this star.

Key words: methods: numerical – methods: statistical – stars: evolution – stars: fundamental parameters – stars: individual: α Cen A – stars: oscillations.

1 INTRODUCTION

In stellar physics, α Cen A holds a particular place. Located at 1.3 pc from the Sun, it is the closest star to the Solar system. Its apparent magnitude is $V = -0.01$, making it a privileged target to obtain precise and accurate data from spectroscopy, photometry, astrometry and interferometry. Having a G2V spectral type, it is physically close to the Sun and is thus of importance for physicists interested in the processes and mechanisms governing the structure and evolution of solar analogues. Finally, it is part of a triple system and forms a close, visual, binary with α Cen B. This allows us to directly measure the mass of the star, which is an information of critical importance in stellar physics.

For all these reasons, α Cen A has been the focus of many observational campaigns and theoretical studies. The former have allowed us to obtain very precise data for the star. We possess measurements of its atmospheric parameters (effective temperature, surface gravity, chemical abundances; Chmielewski et al. 1992; Neuforge-Verheecke & Magain 1997), luminosity (Söderhjelm 1999), radius (Kervella et al. 2003) and oscillation frequencies (Bouchy &

Carrier 2002; Bedding et al. 2004). All these have been combined to constrain theoretical models for the stellar structure and evolution, leading to a significant improvement in our knowledge of the physical properties of α Cen A, its mass, age and the initial chemical composition.

Despite all these advances, it is obvious that there is no perfectly unified picture yet either of the internal structure of α Cen A or of its evolutionary state. Even though its mass is well known, there exist discrepancies in the ages found in the literature (Turcotte & Christensen-Dalsgaard 1998; Guenther & Demarque 2000; Morel et al. 2000; Eggenberger et al. 2004; Miglio & Montalbán 2005). It indeed depends very much on the considered observational constraints and the particular physics included in the models that are used to reproduce these quantities.

Furthermore, as it turned out, α Cen A is in a very peculiar state with regard to stellar structure and evolution. Indeed, stars with mass $\gtrsim 1.1 M_{\odot}$ can be convectively unstable in their cores. This is very important because it affects the energy transport in the star and thus its entire evolution. However, it is a difficult task to identify such small convective cores. First, they might not influence much the stellar observables (with the notable exception of its oscillation frequencies) and could remain undetected. Secondly, the physics of convection governing these structures are poorly known, and it is

*E-mail: bazot@astro.up.pt

likely that the models used to describe α Cen A are incorrect with respect to the treatment of its inner convection. This is reflected by the fact that various studies have led to different claims regarding the existence of a convective core in α Cen A (Thévenin et al. 2002; Thoul et al. 2003; Eggenberger et al. 2004; Miglio & Montalbán 2005; de Meulenaer et al. 2010).

A further problem appears when one needs to quantify precisely the uncertainties on the physical stellar parameters estimated relatively to the uncertainties on the observations. This is not a problem specific to α Cen A: it has always been challenging in stellar physics to obtain robust and credible uncertainties on the parameters estimated from modelling. This is mostly due to the substantial computational cost of the stellar evolutionary codes, which is often incompatible with the intensive use of the models required by statistical methodologies. It does not come as a surprise then that most of the efforts in stellar physics have long been directed towards improving the codes themselves, either numerically or by including more accurate physics, rather than developing tools for parameter estimation.

This situation has progressively changed during the past decades. Many studies have been published providing extensive grids of models, spanning large ranges of values of physical parameters such as the stellar age, mass or the initial chemical composition (Schaller et al. 1992; Girardi et al. 2000; Pietrinferni et al. 2004; Quirion, Christensen-Dalsgaard & Arentoft 2010). These can help to proceed to statistical studies without having to compute too many models at once. Simultaneously, there has been a growing interest for methodologies for parameter estimation, which have been applied to stellar physics. Among them, optimization methods have been introduced first; in some cases great care has been taken to combine this procedure with uncertainty analysis such as singular value decomposition (SVD) (Brown et al. 1994). The methods adopted for optimization are many, from straightforward grid analysis to solving the non-linear least-squares problem using a Levenberg–Marquardt algorithm (Miglio & Montalbán 2005) or stochastic sampling via genetic algorithms (Metcalf 2003). Other studies have focused on the application of Bayesian statistics with some success. Here again, the methodologies differ, some using grid-based strategies (Jørgensen & Lindegren 2005; da Silva et al. 2006; Takeda et al. 2007) and others Markov chain Monte Carlo (MCMC) algorithms (Bazot, Bourguignon & Christensen-Dalsgaard 2008).

In this article, we present a Bayesian modelling of α Cen A based on such an MCMC methodology. In Section 2, defining our inverse problem, we briefly recall the essentials of the Bayesian approach. We also give a brief account of the MCMC theory and present the algorithm we used to solve this inverse problem. In Sections 3 and 4 we present the physical context to which these tools are to be applied. We review briefly the characteristics of our stellar evolution code and give an overview of the current observational knowledge on α Cen A. In order to assess the potential gains one can expect from the use of MCMC algorithms with respect to grid-based strategies we give two solutions to our inverse problem. In the first one we let two stellar parameters vary and make assumptions to fix three others. In the second one, we estimate all five. We report the results from both numerical experiments in Section 5. In particular, we concentrate on two points. One is physical: what is the general picture of the structural and evolutionary state of α Cen A? The other is numerical: how efficiently were these results obtained? From the methodological point of view, our results show the interest of MCMC approaches for such estimation problems, when no explicit mapping exists between observables and the parameters to be estimated. In particular, compared to grid-based computations, the

MCMC approach reveals better efficiency by concentrating most of the computational effort on the regions of interest. They also naturally provide crucial information about uncertainties associated with the estimated values. We briefly compare our results to previous studies in Section 5.3.

2 THE BAYESIAN APPROACH

2.1 A difficult inverse problem

Stellar models include several parameters that have to be estimated by matching observational data, which represent a typical inference problem. Stellar observable quantities usually considered are the luminosity L , the effective temperature T_{eff} , the surface metallicity Z , the radius R and the frequencies of stellar pulsation modes in a given range of radial orders n and angular degrees l , say $(v_{n,l})$. α Cen A is one of the very important cases in stellar physics for which all of these measurements are available to high precision. From such observations, one usually wants to estimate stellar parameters such as the mass of the star, M , its age, t_* , the mixing-length parameter, α , and the initial mass-fractions of helium and heavy elements, respectively Y_0 and Z_0 .¹ This is by no means an exhaustive set of parameters, in particular if non-standard physics are to be included.

Unfortunately, there is no closed-form equation linking stellar parameters together with observables, which could be *inverted*. For a given set of interest stellar parameters, corresponding values of L , T_{eff} , Z , R and $(v_{n,l})$ can be computed by stellar evolution codes (which may also require additional parameters). Let $\theta = (\theta_1, \dots, \theta_K)$ regroup the unknown stellar parameters and $X = (X_1, \dots, X_N)$ collect the observed quantities. We shall call the parameter space, \mathbb{P} , and the observable space, \mathbb{O} , the spaces containing respectively θ and X . We have $\dim(\mathbb{P}) = K$.

We shall now consider a mapping from \mathbb{P} to \mathbb{O} and write

$$\tilde{X} = \mathcal{S}(\theta),$$

where elements in \tilde{X} are the values predicted by the stellar code for these quantities, denoted here by \mathcal{S} . We consider that the observations are related to \tilde{X} by

$$X = \tilde{X} + \epsilon,$$

where ϵ is an unknown error term accounting for instrumental errors on the measurements and possible model errors between \tilde{X} and $\mathcal{S}(\theta)$. The problem addressed here concerns the estimation of θ from

$$X = \mathcal{S}(\theta) + \epsilon, \quad (1)$$

which defines an *inverse problem* (Idier 2008). A classical approach to solve this problem consists of minimizing an appropriate distance, in \mathbb{O} , between X and \tilde{X} , such as the usual χ^2 error $\|X - \mathcal{S}(\theta)\|^2 = \sum_{n=1}^N (X_n - \tilde{X}_n)^2$. This is a hard task in our case. First, $\mathcal{S}(\theta)$ is not an explicit function, so no analytical formulation can be exploited by efficient optimization methods [e.g. no expression of the derivatives of $\mathcal{S}(\theta)$ with respect to each variable in θ can be used for optimization]. Secondly, stellar models computed by \mathcal{S} make use of highly non-linear functions, so that classical optimization methods may be stacked in local minima or yield degenerate solutions. The complexity of the mapping from the space of parameters to the space of observables has already been pointed out

¹ Or X_0 , the initial hydrogen fraction, remembering that the relation $X_0 + Y_0 + Z_0 = 1$ holds.

(Brown et al. 1994; Jørgensen & Lindegren 2005) and represents one of the major challenges in stellar parameter estimation. It has to be noted that obtaining $\mathcal{S}(\theta)$ for a given set of parameters θ is generally computationally expensive. Therefore, computing the model on a grid of discrete values of the parameters may be extremely time-consuming if one wants to explore the whole parameter space with high precision, especially as its dimension increases.

We consider the use of MCMC methods as an alternative to classical optimization methods, where model computations are concentrated around the interest regions of the parameter space, and are in that sense ‘optimized’. Basically, the distance between observations and their prediction by $\mathcal{S}(\theta)$ is linked to a probability distribution, and MCMC algorithms aim at approximating it by generating random samples distributed according to such distribution. Then, they allow one not only to obtain an estimate of the interest parameters, but also to associate characterizations such as uncertainties or higher-dimensional statistics. MCMC are also naturally oriented towards a Bayesian setting, in which prior information on the unknown parameters is also taken into account by means of probability distributions. Furthermore, being stochastic in nature, they might help to avoid confinement in local minima.

2.2 Probabilistic framework

2.2.1 Likelihood function

Equation (1) has a statistical interpretation, where the perturbation term ϵ can be seen as the realization of a random process. We suppose hereafter that the error ϵ_n on each observation X_n is normally distributed, with variance σ_n^2 and that the errors affecting the different observations are independent of each other. In our case, we assume that σ_n corresponds to the uncertainty on X_n , which is determined by the observation.

Then, the probability density of all perturbation terms taken together is

$$p_\epsilon(\epsilon) = \prod_{n=1}^N \frac{1}{(2\pi)^{1/2}\sigma_n} \exp\left(-\frac{\epsilon_n^2}{2\sigma_n^2}\right),$$

which defines the *likelihood function* of data X given a model parametrized by θ using (1)

$$\begin{aligned} \mathcal{L}(\theta; X) &= p_\epsilon(X - \mathcal{S}(\theta)) \\ &= \prod_{n=1}^N \frac{1}{(2\pi)^{1/2}\sigma_n} \\ &\quad \times \exp\left(-\frac{1}{2} \sum_{n=1}^N \left[\frac{\mathcal{S}_n(\theta) - X_n}{\sigma_n}\right]^2\right). \end{aligned} \quad (2)$$

Then, maximizing the likelihood (2) is equivalent to minimizing the usual least-squares term

$$\min_{\theta} \chi^2(\theta) = \sum_{n=1}^N \left[\frac{\mathcal{S}_n(\theta) - X_n}{\sigma_n}\right]^2. \quad (3)$$

2.2.2 Incorporating prior information

Suppose one has some – even weak – prior information about the unknown parameters θ , such as an acceptable range of possible values or previous measurements with associated uncertainties. One may want to take it into account in order to constrain the solution of (3). Bayesian statistics provide a natural way to do this (Idier

2008) by considering such information through a prior probability distribution on θ , say $\pi(\theta)$.

Suppose, for example, that the age t_* is bounded between t_{\min} and t_{\max} , without more information. Then one can use

$$\pi(t_*) = \frac{1}{t_{\max} - t_{\min}} \mathbb{1}_{[t_{\min}, t_{\max})}(t_*), \quad (4)$$

where $\mathbb{1}_{\mathcal{D}}(u)$ is the indicator function of the domain \mathcal{D} , which equals 1 if $u \in \mathcal{D}$ and 0 otherwise. That is, $\pi(t_*)$ is a uniform distribution on \mathcal{D} . Or suppose that the mass M has been previously measured at M_{est} , with uncertainty $\sigma_{M_{\text{est}}}$. Then, one may use a Gaussian prior distribution $\pi(M)$, centred at M_{est} and with standard deviation $\sigma_{M_{\text{est}}}$ (or $2\sigma_{M_{\text{est}}}$, $3\sigma_{M_{\text{est}}}$ for a weaker prior distribution).

Once such priors are defined, θ can be estimated from the posterior probability distribution (PPD), which reads, by Bayes’ theorem,

$$\pi(\theta|X) = \frac{\mathcal{L}(\theta; X)\pi(\theta)}{m(X)}. \quad (5)$$

$\pi(\theta|X)$ represents the probability distribution of all unknown parameters constrained by observed values of X and by prior distributions. In the latter equation, $m(X)$ is a normalization constant ensuring that the integral of $\pi(\theta|X)$ over the entire space of parameters is 1.

2.2.3 Exploiting the posterior distribution

Suppose we are able to access the whole PPD $\pi(\theta|X)$. It contains all statistical information available, given a set of related observed quantities and (possibly) under the given prior assumptions. If no prior distribution is used, then $\pi(\theta|X)$ identifies with the likelihood function $\mathcal{L}(\theta; X)$. We are then theoretically able to exploit such a distribution in order to derive estimates of interest parameters θ . In this paper, we focus on two most common ones in Bayesian analysis. The most intuitive one is the maximum a posteriori (MAP) estimate, which is the set of parameters that maximize the PPD (i.e. ‘the most probable value’ of θ given observations X)

$$\hat{\theta}_{\text{MAP}} = \arg \max_{\theta} \pi(\theta|X). \quad (6)$$

The Posterior Mean (PM) estimate corresponds to the expected value of θ under distribution (5)

$$\hat{\theta}_{\text{PM}} = \mathbb{E}[\theta|X] = \int \theta \pi(\theta|X) d\theta. \quad (7)$$

It minimizes the mean square error $\mathbb{E}[(\hat{\theta}(X) - \theta)^2]$.

The latter one is also often used in Bayesian estimation since, knowing $\pi(\theta|X)$, one can associate variances, that is, uncertainties, to the estimated values.

In practice, of course, we do not have access to the whole PPD and computation of the former estimates is a hard task. Optimization of $\pi(\theta|X)$ is at least as difficult as optimization of likelihood (2), for the reasons explained in Section 2.1: $\pi(\theta|X)$ does not have an exploitable closed-form expression. Computing $\hat{\theta}_{\text{PM}}$ requires an integration over all possible values of θ and is also impossible analytically.

2.3 Markov chain Monte Carlo algorithm

MCMC methods (for a general review see Robert & Casella 2005) are powerful techniques that allow us to approximate complex probability densities. Roughly speaking, MCMC aims at generating samples randomly distributed according to the interest distribution (5), and then approximating such distributions by histograms. MCMC

methods are powerful in the sense that they are able to generate random samples distributed (at least, asymptotically) according to a target distribution only by evaluating this distribution at a certain number of parameter values.

In our case, suppose we have T samples $\theta^{(1)}, \dots, \theta^{(t)}, \dots, \theta^{(T)}$ distributed according to $\pi(\theta|X)$ – we use the notation $\theta^{(t)} \sim \pi(\theta|X)$. Then, the PM estimate (7) of all parameters in θ can be approximated by

$$\hat{\theta}_{\text{PM}} \simeq \frac{1}{T} \sum_{t=1}^T \theta^{(t)}, \quad (8)$$

and the corresponding unbiased variance estimator by

$$\hat{\sigma}^2(\hat{\theta}_{k,\text{PM}}) \simeq \frac{1}{T-1} \sum_{t=1}^T (\theta_k^{(t)} - \hat{\theta}_{k,\text{PM}})^2, \quad (9)$$

with θ_k the k th component of θ . We can also approximate the MAP estimate by taking, among all $\theta^{(t)}$, the set of parameters for which $\pi(\theta^{(t)}|X)$ is the highest (note that the corresponding values are naturally obtained as a by-product of the MCMC algorithm).

MCMC methods work by generating a *Markov chain*, that is, a sequence $\theta^{(1)}, \dots, \theta^{(T)}$ such that the rule for generating $\theta^{(t)}$ depends only on the previous value $\theta^{(t-1)}$. The *Monte Carlo* reference obviously concerns the random nature of the process. Giving a full account of the theory of MCMC is obviously far beyond the scope of this paper. In the following, we will only comment on some essential practical aspects necessary for the comprehension of the subsequent parameter estimation. This approach was first suggested in the framework of stellar modelling by Bazot et al. (2008).

2.3.1 The Metropolis–Hastings algorithm

In this paper, we construct a Metropolis–Hastings (MH) algorithm (Metropolis 1953; Hastings 1970), which is the most generic class of MCMC methods. The generality of this method allows us to estimate probability densities of all forms, with the only assumption that they can be evaluated at every point in the parameter space. The PPD expressed in (5) falls within this category.

The Markov chain is constructed through a two-step iterative procedure. Suppose that $\theta^{(1)}, \dots, \theta^{(t)}$ have been obtained. First, a trial value of the set of parameters, θ^* , is generated according to an *instrumental distribution* $q(\theta^*|\theta^{(t)})$ that may depend on the previous value of the chain $\theta^{(t)}$. Then a selection process operates, where $\theta^{(t+1)}$ is set to θ^* with some probability ρ , and to $\theta^{(t)}$ otherwise. The convergence of the distribution of $\{\theta^{(t)}\}$ towards the target distribution is then ensured by the analytical expression of ρ (see the algorithm description below). We give more details about monitoring of the convergence of MCMC in Appendix A.

Algorithm [A.1] MH algorithm

For $t = 1$, choose an initial value $\theta^{(1)}$. Then, at iteration t :

1. Generate: $\theta^* \sim q(\theta|\theta^{(t)})$
2. Select:

$$\theta^{(t+1)} = \begin{cases} \theta^* & \text{with probability } \rho(\theta^{(t)}, \theta^*), \\ \theta^{(t)} & \text{with probability } 1 - \rho(\theta^{(t)}, \theta^*), \end{cases}$$

where

$$\rho(\theta^{(t)}, \theta^*) = \min \left\{ \frac{\pi(\theta^*|X) q(\theta^{(t)}|\theta^*)}{\pi(\theta^{(t)}|X) q(\theta^*|\theta^{(t)})}, 1 \right\}.$$

The choice of θ^* at step 1 is detailed in Section 2.3.2. In practice, accepting $\theta^{(t+1)} = \theta^*$ with probability ρ is performed by generating u according to a uniform distribution in $[0, 1]$, and $\theta^{(t+1)} = \theta^*$ is accepted if $u < \rho$.

Note that if q is a symmetrical distribution [i.e. $q(\theta^{(t)}|\theta^*) = q(\theta^*|\theta^{(t)})$], which will be the case here, then the ratio in probability ρ is nothing but the PPD ratio evaluated at θ^* and $\theta^{(t)}$. Hence, parameters increasing the PPD are always accepted while parameters decreasing the PPD are randomly accepted, the probability of acceptance decreasing with the PPD ratio.

2.3.2 Instrumental distribution

A key point conditioning the efficiency of the MH algorithm is the choice of the instrumental distribution. First, it should be computationally easy to draw random samples distributed according to q . Secondly, q should concentrate the proposed values θ^* around the interest regions of $\pi(\theta|X)$. Finally, q must also explore the whole support of possible values in order not to miss plausible parameter sets. It should be noted that the first condition represents one of the major computational interests of using MCMC algorithms: it becomes possible to simulate complex distributions using only simpler ones.

Trying to account for these issues and following former similar choices (e.g., Andrieu & Doucet 1999), we choose $q(\theta^*|\theta^{(t)})$ as a mixture of two Gaussian distributions centred at $\theta^{(t)}$ and the uniform distribution on the definition domain of all parameters. That is, for Step 1 in the MH algorithm, we generate, for all parameters θ_k^* , with $k = 1, \dots, K$,

$$\begin{cases} \theta_k^* \sim \mathcal{N}_{\mathcal{D}_{\theta_k}}(\theta_k^{(t)}, \zeta_{k,1}^2) & \text{with probability } p_1, \\ \theta_k^* \sim \mathcal{N}_{\mathcal{D}_{\theta_k}}(\theta_k^{(t)}, \zeta_{k,2}^2) & \text{with probability } p_2, \\ \theta_k^* \sim \mathcal{U}_{\mathcal{D}_{\theta_k}} & \text{with probability } p_3, \end{cases} \quad (10)$$

where $\mathcal{N}_{\mathcal{D}_{\theta_k}}(\theta_k^{(t)}, \zeta_{k,1}^2)$ is the Gaussian distribution with mean $\theta_k^{(t)}$ and variance $\zeta_{k,1}^2$, truncated to the definition domain \mathcal{D}_{θ_k} of θ_k , and $\mathcal{U}_{\mathcal{D}_{\theta_k}}$ is the uniform distribution on \mathcal{D}_{θ_k} – see Sections 5.1 and 5.2 in which we specify the $(\zeta_{k,j})$, p_i and the acceptable definition domain for θ .

The use of Gaussian distributions centred on the current value $\theta^{(t)}$ defines random walks in the space of parameters. Having two of them helps to perform local scans of the parameter space at different scales. We added the uniform distribution to avoid possible trapping in local extrema: if the chain is being stuck in such an area, there is a chance for the algorithm to propose, from time to time, values of θ well outside this region.

3 STELLAR MODELS

3.1 Physics

In equation (1) appears the quantity $\mathcal{S}(\theta)$, describing the predicted values for the observable quantities. These values were computed using the AARHUS STELLAR EVOLUTION CODE (ASTEC). A thorough description is given by Christensen-Dalsgaard (1982, 2008b); we thus only briefly recall the main outline of the physics used in our models. The oscillation frequencies were computed with the ADIPLS package (Christensen-Dalsgaard 2008a). We restrict ourselves to the so-called standard physics. This implies a spherical, non-rotating

star. Microscopic diffusion, magnetic field and mixing processes other than convection (overshooting, rotational mixing, . . .) are neglected. This latter is described using the mixing-length theory (Böhm-Vitense 1958) that requires the setting of a mixing-length parameter, α , defined as $\alpha \equiv \ell_m/H_p$, where ℓ_m is the mixing-length and H_p is the pressure scale-height.

The physics used in the stellar code are relatively basic. It is not the goal of this study to adopt the state-of-the-art physics for the stellar evolution code. The emphasis was rather put on the need to use a fast, robust and already tested version of the code. This is the reason why we used the EFF tables for the equation of state (Eggleton, Faulkner & Flannery 1973) instead of the latest OPAL tables (Rogers, Swenson & Iglesias 1996; Rogers & Nayfonov 2002), noting however that they agree ‘reasonably well’. It also appears from Miglio & Montalbán (2005) that, in the framework of the modelling of α Cen A, using EFF does not give results departing significantly from those using OPAL. Other input physics included the NACRE nuclear reaction rates (Angulo et al. 1999) and OPAL opacity tables (Iglesias & Rogers 1996). The relative abundances of heavy elements are those prescribed by Grevesse, Noels & Sauval (1993, GNS abundances).

This latter point should be briefly commented since these abundances are, at the moment, subject to debate. Using three-dimensional simulations, Asplund et al. (2004) have derived new solar abundances (AGS abundances). The most remarkable features are the low values found for C, N, O. Abundances of other heavy elements such as Na, Mg or Fe were also revised to lower values. The overall impact is a reduction of ~ 30 per cent of the estimated solar metallicity. However, including these values in solar models has led to sound-speed profiles in disagreement with helioseismology (e.g. Montalbán et al. 2004; Turck-Chièze et al. 2004; Antia & Basu 2005). Although several attempts have been made, some of them including non-standards physics (e.g. Guzik, Watson & Cox 2005; Castro, Vauclair & Richard 2007), no satisfactory agreement between theoretical models and seismology could be reached. In contrast, models using the GNS abundances, computed using one-dimensional models, can reproduce with a remarkable precision the solar sound-speed profiles. It should be noted that recent studies (e.g. Caffau et al. 2008; Asplund et al. 2009) have led to upward revisions of these abundance. Discriminating between AGS and GNS abundances is beyond the scope of this study. We thus keep on using the GNS heavy-element abundances in our models of α Cen A. This rather conservative course of action is motivated by the fact that we focus on testing a new technique for statistical inference of stellar parameters. That should be done with input physics that have already led to coherent results. At the very least, the inferred parameters for α Cen A could be compared to other stellar models that also used GNS abundances.

Given these physical prescriptions, it is important to note that the ‘natural’ parameter space is such that $\dim(\mathbb{P}) = 5$, with the typical associated vector $\theta = (M, t_*, \alpha, Z_0, X_0)$.

4 OBSERVATIONAL CONSTRAINTS

In the following we briefly review the values we selected for the different observables for α Cen A, their uncertainties and their potential bias. We present both the non-seismic and the seismic observations. Table 1 sums up the former ones. For the oscillation frequencies, the reader is referred to Bazot et al. (2007).

Table 1. Observational constraints on α Cen A. The last column specifies the reference for the selected quantity.

Observable	Value	Reference
T_{eff}	5810 ± 50 K	Eggenberger et al. (2004)
L/L_{\odot}	1.522 ± 0.030	Eggenberger et al. (2004)
R/R_{\odot}	1.224 ± 0.003	Kervella et al. (2003)
Z/X	0.039 ± 0.006	Thoul et al. (2003)
$\langle \Delta \nu \rangle$	105.9 ± 0.1 μHz	Bazot et al. (2007)
$\langle \delta \nu \rangle$	6.9 ± 0.4 μHz	Bazot et al. (2007)
M/M_{\odot}	1.105 ± 0.007	Pourbaix et al. (2002)

4.1 Spectrophotometric observables

Several spectroscopic determinations of the effective temperature of α Cen A have been carried out, by fitting either the wings of the $H\alpha$ line or the Fe I lines (Neuforge-Verheecke & Magain 1997). However, some discrepancies appear between the value used in the modelling by Morel et al. (2000), 5790 ± 30 K obtained by fitting the wing of the $H\alpha$ line and the value given by Neuforge-Verheecke & Magain (1997), 5830 ± 30 K. To solve this problem, Eggenberger et al. (2004) have adopted an average value of these two measurements and 1σ error bars encompassing their uncertainties; the same value was used by Miglio & Montalbán (2005). Since it can help to compare our results with these works, we also use $T_{\text{eff}} = 5810 \pm 50$ K, even though reduced error bars of ± 30 K could have helped in reducing the uncertainties on the estimated parameters.

Although the proximity of α Cen A should allow a precise measurement of its parallax, discrepancies also subsist between different studies, affecting the subsequent values of the luminosity. We select the parallax determined by Söderhjelm (1999). The corresponding luminosity is $L/L_{\odot} = 1.522 \pm 0.030$ (Eggenberger et al. 2004). Note that this value of the parallax has also been used to derive the radius and the mass we describe below.

4.2 Radius

The radius of α Cen A has been measured by Kervella et al. (2003) with VINCI (the VLTI Commissioning Instrument), which operates in the K band (2.0–2.4 μm). They measured an angular diameter, after correcting for limb-darkening effects, $d = 8.511 \pm 0.020$ mas. Combined with the parallax derived by Söderhjelm (1999), it led to $R = 1.224 \pm 0.003 R_{\odot}$. For the reference solar radius value, we used the value measured from helioseismology by Brown & Christensen-Dalsgaard (1998), $R_{\odot} = 695.508$ Mm. It is noteworthy that the authors argued that this value, measured in the near-IR, can be compared to models computed with the stellar evolution code CESAM (Morel 1997) since the radius is there defined as the point where $T = T_{\text{eff}}$, which is close to the temperature of the layer at which the continuum at 2.2 μm is formed. Since ASTEC uses the same atmospheric boundary conditions (Christensen-Dalsgaard 1982), we consider that this assumption still holds in our case.

This radius value has been used in Eggenberger et al. (2004), along with seismic data from Bouchy & Carrier (2002), to constrain their model. They concluded that they cannot reproduce at the same time the seismic measurements and the radius within their 1σ error bars. However, they noted that this could be done when considering a 2σ interval. This confirmed the 1.3σ agreement Kervella et al. (2003) found between the observed value of the radius and a previous theoretical estimate using CORALIE frequencies (Thévenin

et al. 2002). Miglio & Montalbán (2005) did not use the observed radius as a constraint for their models but always obtained values higher than the measured one. Montalbán & Miglio (2006), using this constraint in combination with seismic data from Bedding et al. (2004), were able to reproduce all observations within 1σ .

4.3 Seismic observations

4.3.1 Individual frequencies

The frequencies used in this paper are those from Bazot et al. (2007). As far as their accuracy is concerned, they can be considered as reliable. The high signal-to-noise ratio that can be achieved in the radial-velocity spectrum of α Cen A allows us to find easily the most prominent modes, whose amplitude can reach up to 14 times the average noise level. A comparison between the studies by Bouchy & Carrier (2002), Bedding et al. (2004) and Bazot et al. (2007) offers a convincing empirical evidence that the modes are indeed real. The identifications presented in these studies are consistent (which of course does not prove them correct).

Of much greater difficulty is the task of estimating confidence intervals on the estimated frequencies. The assumption made in Bazot et al. (2007) may be viewed as the crudest one: considering that the modes are unresolved, the frequency resolution is chosen as the uncertainty on the frequency. The generic uncertainty given, 1.3 μHz , may be viewed as conservative. At the time our MCMC simulations were first performed, this was as good an estimate as one could find. We note, however, that methods are currently under development to allow better and more trustworthy estimation of the confidence intervals on the oscillation frequencies of solar-like stars (Brewer et al. 2007; Stahn & Gizon 2008; Benomar et al. 2010; Handberg & Campante 2011). A recent discussion on these methods has shown that they are efficient even for ground-based gaped and irregularly sampled data (Bazot et al. 2012). While keeping in mind that room for potential improvement exists, we nevertheless used the original uncertainties in the present study, since their values may at the very least represent a realistic upper bound. We should also note that a new seismic study of α Cen A using the previous data of Bouchy & Carrier (2002) and Bedding et al. (2004) has been published by de Meulenaer et al. (2010) during our study. They have not been included here, but we note that their results partially agree with those of Bazot et al. (2007).

Finally, we should note that fitting the individual frequencies is difficult because they are affected by surface effects that we cannot model. There have been attempts to bypass this problem (Kjeldsen, Bedding & Christensen-Dalsgaard 2008) and thus allow us to use full sets of frequencies as an observational constraint. These methods have gained in popularity and are often used in the framework of Kepler for the modelling of stars. However, these strategies are empirical, and we chose not to use them in the present study. We instead rely on the frequency combinations that we present below. A possible development to be tested along these lines would consist of including in our estimation the Bayesian formalism given by Gruberbauer, Guenther & Kallinger (2012) that deals with this problem. This is left for future studies.

4.3.2 Large separations

The first asymptotic relation was given in the limit of low-degree, high-order p modes by Vandakurov (1967). Tassoul (1980) further

expanded the eigenfrequency equation to higher orders

$$\nu_{n,l} = \left(n + \frac{l}{2} + \epsilon\right)\Delta\nu_0 + (l(l+1)A - B)\frac{\Delta\nu_0^2}{\nu} + \mathcal{O}\left(\frac{1}{\nu^2}\right) \quad (11)$$

with

$$\Delta\nu_0 = \left(2 \int_0^{R_t} \frac{dr}{c}\right)^{-1}, \quad (12)$$

$$A = \frac{1}{4\pi^2\Delta\nu_0} \left[\frac{c(R_t)}{R_t} - \int_0^{R_t} \frac{1}{r} \frac{dc}{dr} \frac{dr}{r} \right], \quad (13)$$

and with R_t the upper turning point of the mode and B a surface phase term.

If we assume, to the lowest order of approximation, that the phase-related term quantity ϵ_s does not depend on the frequency, we immediately see that modes of consecutive order and same degree are equally spaced. The difference $\Delta\nu_{n,l} = \nu_{n,l} - \nu_{n-1,l}$ is called the large separation and may be considered roughly as a constant. This has been a useful tool in helio- and asteroseismology to identify pulsation modes (e.g. Grec, Fossat & Pomerantz 1983; Bouchy & Carrier 2002).

The large separation is, according to equation (12), the inverse of twice the acoustic radius (i.e. the travelling time of an acoustic wave from the centre to the surface). We thus obtain a new global quantity that can be used to constrain stellar models. As noted by Deubner & Gough (1984), this integral depends mostly on the external regions of the star. This can be easily seen if one considers that the sound speed in the stellar interior is a roughly monotonic, smoothly decreasing function of the radius, and that the ratio of its values at the centre and at the surface for a solar-like star is ~ 70 . Therefore, any acoustic wave will spend a much longer time closer to the surface than to the centre and will be predominantly sensitive to this region.

It has been noted that the large separation depends on the frequency. Allowing for the surface-phase term to vary with frequency, in an equation similar to (11), Vorontsov (1991) pointed out that it will control the overall behaviour of $\Delta\nu_0$ when seen as a function of ν . Once more, some problems arise when it comes to stellar models: as noted by Vorontsov & Zharkov (1989), they usually provide poor values for this surface term and its frequency derivative, which is directly related to $\Delta\nu_0$.

From the point of view of stellar parameter estimation, this is nevertheless an interesting quantity. Besides the new (and relatively easy to extract) global constraint it provides, it is possible to compare the stellar parameters inferred using the large separations and other seismic indicators, obtained using different frequency combinations sensitive to deeper layers. This may in turn give us some insight into the modelling errors occurring below the surface. The quantity we used is the average large separation that we define as $\langle\Delta\nu\rangle = \Delta\nu_0$ and compute as explained in Bazot et al. (2007) and note that the estimate they give should read $105.9 \pm 0.1 \mu\text{Hz}$ and not $105.9 \pm 0.3 \mu\text{Hz}$; we therefore adopt the former value.

Numerically, we computed the average large separation by performing a linear regression, for each degree $l = 0, 1, 2$, of the relation $\nu = \nu(n)$, only selecting in the code output the modes with orders corresponding to the values observed for α Cen A. This gives us three slopes, which are the average large separation for each degree. We then averaged over l to obtain the average large separation.

4.3.3 Small separations

The second quantity used as a seismic indicator comes also straightforwardly from equation (11). It immediately appears that modes separated by one order and two degrees, $\nu_{n,l}$ and $\nu_{n,-l+2}$, are degenerate at the first order. The difference between these frequencies is a second order quantity in $1/\nu$. It is called the small separation.

Contrary to the large separations, the small ones depend mostly on the deep interior of the star. This can be seen from equation (13). The integral term behaves as c/r which is large at the centre and negligible near the surface. An in-depth analysis by Gough (1986), Gabriel (1989) and Van Hoolst & Smeyers (1991) described much more rigorously this dependence.

Small separations are obviously a very important tool for stellar parameter estimation using seismology. It is noteworthy that Christensen-Dalsgaard (1993) has shown that, in principle, the knowledge of the average values of the large and small separations is sufficient to characterize fully the evolutionary state of the star. We performed test estimation using the individual and the average large separations. The main difference is that the small uncertainty on the latter provides a better constraint to our stellar model; therefore, we adopted the value provided by Bazot et al. (2007), $\langle \delta\nu \rangle = 6.9 \pm 0.4 \mu\text{Hz}$. The average small separation in our models was simply computed by averaging our theoretical individual separations over the mode order (selecting only the observed ones).

Finally, we also tested the use of the ratio describe by Roxburgh & Vorontsov (1994), which should be completely free of surface effects.² However, the large uncertainty on these quantities, in the range 22.4 to 30.2 per cent, made them poor constraints in the case of this α Cen A data set.

4.4 Using a mass measurement as a prior

α Cen A being part of a close binary, its mass can be measured using astrometry. Although its period is somewhat long, observations have been carried out throughout the last century, culminating with the precise measurements of Murdoch & Hearnshaw (1993), Endl et al. (2001) and Pourbaix et al. (2002), using high-precision spectrometers. This led to the derivation of a precise mass by Pourbaix et al. (2002), $M_{\text{obs}} = 1.105 \pm 0.007 M_{\odot}$ (the error being given to the 1σ level). This mass was also used by Eggenberger et al. (2004), Miglio & Montalbán (2005) and Montalbán & Miglio (2006).

We used this value to obtain prior information on the mass. The assumption is made that the uncertainties are Gaussian and reflect truly the 1σ deviation from the mean value, therefore

$$\pi(M) \propto \exp\left(-\frac{(M - M_{\text{obs}})^2}{\sigma_M^2}\right). \quad (14)$$

It should be noted that Pourbaix et al. (2002) issued a warning concerning the accuracy of this value. Even though the mass ratio between α Cen A and α Cen B can be accurately constrained, the same is not true for their individual masses, because of the combined effects of the gravitational redshift and convective blueshift (this latter being especially difficult to correct for).

² Whereas there is still a contribution of the surface layers to the small separations. In practice, our test runs with this so-called Roxburgh's ratio on α Cen A gave us the same estimates as those using the small separations, albeit with much larger error bars.

5 RESULTS ON α CEN A

5.1 MCMC simulations with $\dim(\mathbb{P}) = 2$

5.1.1 Implementation details

Our first set of MCMC simulations aimed at sampling the PPD in a parameter space of small dimension. This has some advantages, in particular, for the comparison with procedures involving direct integration of the normalization factor in equation (5) from fixed grids of models (Section 5.1.4). It is indeed possible, in that case, to compute very dense grids of stellar models. We thus limit ourselves to a parameter space such as $\dim(\mathbb{P}) = 2$, with the stellar mass and age the only two free quantities.

In order to fix the other parameters, we had to make a certain number of assumptions. The initial hydrogen fraction was considered roughly solar $X_0 = X_{0,\odot} \sim 0.7$. Combining with the [Fe/H] ratio for α Cen A, and because we neglected diffusion, this implies $Z_0 = 0.027$. This of course is an oversimplification since we intentionally neglect the effect of the uncertainty on [Fe/H]. However, such a configuration may happen if one is, for instance, using grids of stellar models only sparsely sampling the initial metallicity parameter. The mixing-length parameter is somewhat more problematic since there is no fundamental physical justification for introducing this parameter. It can be related to the gradient of entropy in the considered convective zone, but there is no physical prescription that can be unequivocally applied to this quantity. It is possible to use a canonical value, calibrated from the solar case. This is the approach we used here, setting $\alpha = \alpha_{\odot} \sim 1.8$. Some numerical simulations suggest that the value of the mixing-length parameter in the solar neighbourhood in the $(M - t_*)$ plane is close to constant (Trampedach 2007).

The setup was similar for all our MCMC simulations. For each combination, X , of the observables we considered, we launched $m = 3$ independent Markov chains of length $T = 50\,000$. The use of parallel chains is often a good method to control how dependent the results are on the initial guesses. The length of the chains was chosen, after some test runs, so that it would be always longer than the effective number of iterations required to reasonably conclude that convergence has been reached. For each run, convergence was tested using graphical indicators such as the sequential mean and variance. A typical burn-in sequence of 5000 iterations was discarded at the beginning of each Markov chain realization in order to ensure that the sample was indeed generated according to the target PPD (see Appendix A).

Based on test runs, we set the Gaussian components of the instrumental law $q(\theta^*|\theta^{(i)})$, given by equation (10), to $\zeta_1 = (0.005 M_{\odot}, 0.01 \text{ Gyr})$ and $\zeta_2 = (0.01 M_{\odot}, 1 \text{ Gyr})$. At this point, we should note that the joint PPDs for (M, t_*) show strong correlations between the two parameters, as can be seen in Fig. 1 in which we display them. Such a behaviour was expected from a theoretical point of view, but it certainly affects the efficiency of the MCMC algorithm, the instrumental laws used to sample the PPDs being symmetrical. This effect is difficult to correct manually, since the correlation coefficient appears to change from one combination X to the other and thus cannot be known a priori.

Empirically, we notice that the composite structure of the instrumental law helps us to converge if compared to simple ones, either Gaussian or uniform. This can indeed be seen from the acceptance rates for the different components of $q(\theta^*|\theta^{(i)})$ from one run to the other. Whereas it is always low for the uniform distribution,

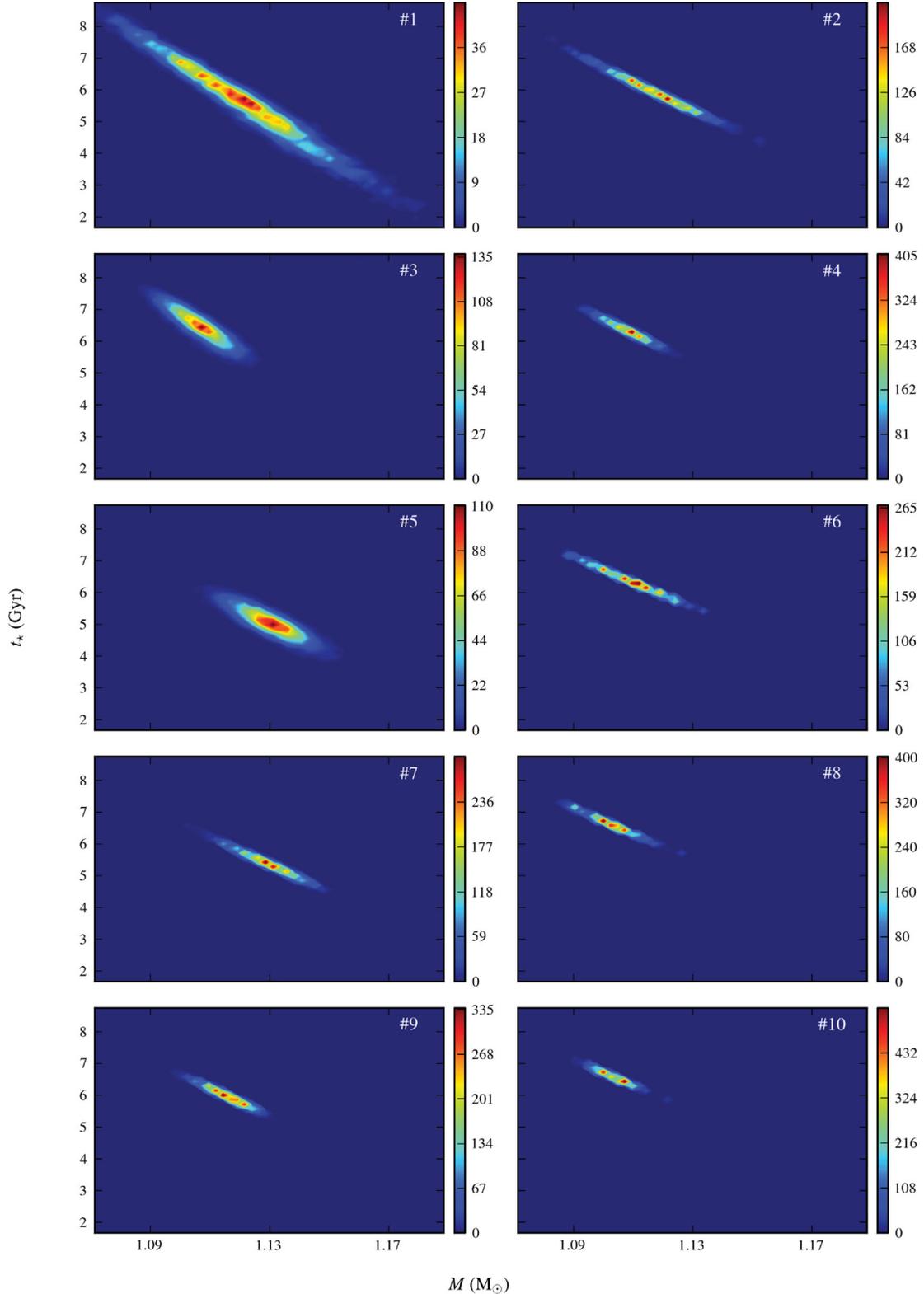


Figure 1. Joint posterior probability densities for the couple (M, t_*) obtained from our MCMC simulations. Each one is labelled with the number given in the first column of Table 2.

it varies significantly between runs for the large and the medium Gaussian. Therefore, the possibility to switch from one law to the other compensates the fact that it was not adjusted from simulation to simulation. The probabilities for law selection in equation (10)

were set to $p_1 = p_2 = 2p_3 = 0.4$. Of course, this is only one possible solution for the choice of the instrumental law and it might not be the optimal one. We have merely selected it for its convenience of implementation and handling.

Table 2. Inferred stellar parameters in the $(M - t_*)$ plane of the parameter space constrained using various combination, X , of the observables. First column: the number of the run; second column: observational constraints included in X ; third column: prior used on the mass. The next columns give the estimates for these quantities in the form of the MP estimate and the associated credible 1σ credible interval. Columns 4 and 5: estimates obtained with our MCMC algorithm. Columns 6 and 7: estimates obtained by computing the PPD on (216×216) grid sampling regularly the $(M - t_*)$ plane. Columns 8 to 13: same with (108×108) , (32×32) and (16×16) subgrids in the $(M - t_*)$ plane. Note that the dimension of the grid is indicative; the effective number of points used is lower since some models could not be computed (see text).

Run no.	X	$\pi(M)$	MCMC		Grid 1 (216×216)		Grid 2 (108×108)	
			\hat{M} (M_\odot)	\hat{t}_* (Gyr)	\hat{M} (M_\odot)	\hat{t}_* (Gyr)	\hat{M} (M_\odot)	\hat{t}_* (Gyr)
1	(L, T_{eff})	Uniform	1.121 ± 0.019	5.63 ± 1.14	1.120 ± 0.020	5.67 ± 1.20	1.120 ± 0.020	5.68 ± 1.20
2	(L, T_{eff}, R)	Uniform	1.116 ± 0.011	5.98 ± 0.52	1.116 ± 0.012	5.99 ± 0.53	1.116 ± 0.012	5.99 ± 0.53
3	(L, T_{eff})	Gaussian	1.107 ± 0.007	6.47 ± 0.44	1.107 ± 0.007	6.48 ± 0.44	1.107 ± 0.007	6.48 ± 0.44
4	(L, T_{eff}, R)	Gaussian	1.108 ± 0.006	6.34 ± 0.28	1.108 ± 0.006	6.34 ± 0.28	1.108 ± 0.006	6.34 ± 0.28
5	$(L, T_{\text{eff}}, \langle \delta v \rangle)$	Uniform	1.130 ± 0.008	5.03 ± 0.39	1.131 ± 0.008	5.00 ± 0.39	1.130 ± 0.008	5.03 ± 0.39
6	$(L, T_{\text{eff}}, \langle \Delta v \rangle)$	Uniform	1.109 ± 0.009	6.35 ± 0.38	1.111 ± 0.009	6.28 ± 0.37	1.111 ± 0.009	6.28 ± 0.37
7	$(L, T_{\text{eff}}, R, \langle \delta v \rangle)$	Uniform	1.130 ± 0.008	5.36 ± 0.34	1.130 ± 0.008	5.35 ± 0.35	1.130 ± 0.008	5.36 ± 0.35
8	$(L, T_{\text{eff}}, R, \langle \Delta v \rangle)$	Uniform	1.103 ± 0.007	6.62 ± 0.29	1.103 ± 0.007	6.60 ± 0.27	1.103 ± 0.007	6.60 ± 0.27
9	$(L, T_{\text{eff}}, R, \langle \delta v \rangle)$	Gaussian	1.116 ± 0.006	5.98 ± 0.25	1.115 ± 0.006	5.99 ± 0.25	1.115 ± 0.006	6.00 ± 0.25
10	$(L, T_{\text{eff}}, R, \langle \Delta v \rangle)$	Gaussian	1.103 ± 0.005	6.58 ± 0.20	1.104 ± 0.005	6.56 ± 0.20	1.104 ± 0.005	6.56 ± 0.20

Run no.	X	$\pi(M)$	Grid 3 (32×32)		Grid 4 (16×16)	
			\hat{M} (M_\odot)	\hat{t}_* (Gyr)	\hat{M} (M_\odot)	\hat{t}_* (Gyr)
1	(L, T_{eff})	Uniform	1.120 ± 0.020	5.68 ± 1.19	1.120 ± 0.020	5.70 ± 1.21
2	(L, T_{eff}, R)	Uniform	1.116 ± 0.012	5.99 ± 0.53	1.115 ± 0.012	6.04 ± 0.55
3	(L, T_{eff})	Gaussian	1.107 ± 0.007	6.48 ± 0.44	1.106 ± 0.007	6.45 ± 0.44
4	(L, T_{eff}, R)	Gaussian	1.108 ± 0.006	6.35 ± 0.29	1.102 ± 0.005	6.64 ± 0.23
5	$(L, T_{\text{eff}}, \langle \delta v \rangle)$	Uniform	1.131 ± 0.008	5.03 ± 0.39	1.131 ± 0.008	5.02 ± 0.39
6	$(L, T_{\text{eff}}, \langle \Delta v \rangle)$	Uniform	1.113 ± 0.009	6.19 ± 0.37	$1.101 \pm 3 \times 10^{-5}$	$6.70 \pm 1 \times 10^{-3}$
7	$(L, T_{\text{eff}}, R, \langle \delta v \rangle)$	Uniform	1.129 ± 0.008	5.38 ± 0.33	1.124 ± 0.002	5.65 ± 0.12
8	$(L, T_{\text{eff}}, R, \langle \Delta v \rangle)$	Uniform	1.102 ± 0.004	6.66 ± 0.18	$1.101 \pm 2 \times 10^{-5}$	$6.70 \pm 9 \times 10^{-4}$
9	$(L, T_{\text{eff}}, R, \langle \delta v \rangle)$	Gaussian	1.115 ± 0.005	5.99 ± 0.24	1.120 ± 0.008	5.81 ± 0.34
10	$(L, T_{\text{eff}}, R, \langle \Delta v \rangle)$	Gaussian	1.101 ± 0.002	6.69 ± 0.07	$1.101 \pm 5 \times 10^{-7}$	$6.70 \pm 2 \times 10^{-5}$

5.1.2 Estimates of the parameters

The results for our MCMC simulations are reported in Table 2. We considered several possible combinations for X and give them in the second column. The third column shows the prior we used on the mass. We test the effect of uniform, equation (4), and Gaussian, equation (14), priors on the mass. Based on the discussion from Section 4.4, we shall insist on the fact that using the latter is equivalent to make a strong assumption on the value of the mass compared to the non-informative case involving a uniform distribution. We always considered a uniform prior on the age defined on the 2–9 Gyr domain.³

From the MCMC simulation we directly obtain the joint PPDs for the parameters. This is done by simply representing the $M^{(t)}$ and $t_*^{(t)}$, $t = 1, \dots, T$ in the form of a two-dimensional histogram (see Fig. 1). It is also possible to obtain the marginal densities for each parameter. These are usually computed by integrating the PPD over the other parameters (sometimes called nuisance parameters). For instance, the marginal PPD for the mass is given by

$$\pi(M|X) = \int \pi(M, t_*|X) dt_* \quad (15)$$

This integral can be approximated in the same way we did for the joint PPD. A quick inspection shows that they are all close to Gaussian distributions (which can be checked straightforwardly by computing the skewness and the kurtosis of the distribution).

We can use the PM estimate and compute the Posterior Standard Deviation (PSD; square root of the Posterior Variance, noted σ) as the associated credible intervals. In this case, the PSD has the standard interpretation that a realization of the considered variable will deviate from the mean of the distribution by $\pm\sigma$ with probability ~ 68 per cent. These values are given in Table 2, in Column 4 for the mass and in Column 5 for the age. In the context of stellar physics, these statistical summaries are perhaps the most useful products of a Bayesian analysis. The derivation of such error bars allows us to quantify precisely how the various observables and their associated uncertainties impact the stellar parameters. In our 10 runs, we explored the impact of various observables on the final estimates. We always included the effective temperature and the luminosity. We then combined them with other constraints.

The first important result that stands out is the effect of including a Gaussian prior in equation 5. In this case, the estimates of the mass are systematically lower than when a uniform prior is used. Furthermore, they are always extremely close to the value M_{obs} , which reflects the impact of this prior. This, of course, leaves open a classical question that arises in Bayesian studies: is the used prior to be trusted? It is critical since including (14) definitely has a strong influence on the final estimates. Given the general mass–age anticorrelation, the estimated ages are also higher when the prior is Gaussian.

This picture is different when the prior on the mass is uniform and the effects of the other observed quantities appear much more clearly. To understand this, we shall first briefly examine the effect of the various combination we used for X . The three others observational constraints we studied in detail are the radius, R , the average

³ With the exception of Run 1 in Table 2, which required a 1 Gyr lower limit for the uniform prior.

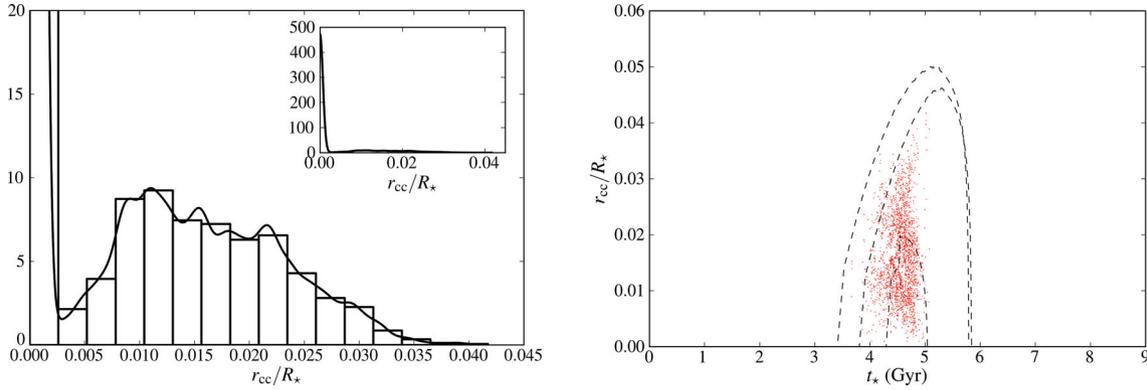


Figure 2. Left-hand panel: posterior probability density for the radius of the convective core of α Cen A estimated using $X = (T_{\text{eff}}, L, \langle \delta\nu \rangle)$ and a uniform prior on the mass (Run 5). The PPD is represented in the form of a histogram and as a kernel-estimated continuous density. The same plot is represented with a different ordinate scale in the inset. Right-hand panel: radius of the convective core as a function of the stellar age. The dots are the values obtained for α Cen A from the same MCMC simulation. The dashed lines represent the theoretical values for stars of mass, from bottom to top, $M = 1.14 M_{\odot}$, $1.15 M_{\odot}$ and $1.16 M_{\odot}$.

large separation, $\langle \Delta\nu \rangle$, the average small separation, $\langle \delta\nu \rangle$. They all lead to more precise estimates with respect to Run 1. This effect is more important for the seismic observable than for the radius. However, it is less than when a Gaussian prior on the mass is used.

It is clear from Table 2 that two groups emerge among these quantities: the small separation on one hand and the radius and large separation on the other hand. The former, when included, leads to sensibly different estimates of the stellar parameters than the latter two. Using $\langle \delta\nu \rangle$ we find the highest mass for all our runs. Conversely, the estimated age is the lowest. This seismic indicator is supposed to be mostly sensitive to the central regions of the star, which concentrate most of the total mass and are the locus of the nuclear burning of hydrogen, and hence it is the best marker for stellar age. Based on these quantities, we obtain a general picture for α Cen A of a star with a mass $\sim 1.13 M_{\odot}$ and an age roughly in the range ~ 5 – 5.3 Gyr; of course this is only true for the physics described in Section 3 and is susceptible of change for different prescriptions.

Both the radius and the average large separation imply lower estimates of the mass and higher estimates for the age (with respect to the estimates obtained including the average small separation in X). For some cases, the estimates vary significantly, for instance between Runs 5 and 6 and Runs 7 and 8, and do not agree within their 1σ error bars (although all runs agree if we consider 3σ error bars). This discrepancy is likely to be caused, as we shall see in Section 5.2, by the assumptions made on the physics of α Cen A, including in this case the setting of approximate values for X_0 and α . Another source of disagreement, as discussed in Section 4, is that $\langle \Delta\nu \rangle$ and R are both affected by the surface layers of the star and reproducing them would imply the use of a stellar evolution code that models accurately these regions. Finally, we also note that the effect of $\langle \Delta\nu \rangle$ is the same as R on the PM estimates, although lesser in magnitude than including a Gaussian prior on the mass, whose observational determination might also suffer from a bias related to the poorly known stellar surface (Section 4.4).

5.1.3 Structure of α Cen A

These variations on the estimated mass and age with respect to X affect the general picture we can draw of the internal structure of α Cen A. A good example concerns the search for a convective core that has been a longstanding problem for this star (Turcotte &

Christensen-Dalsgaard 1998; Guenther & Demarque 2000; Morel et al. 2000; Thévenin et al. 2002; Thoul et al. 2003; Eggenberger et al. 2004; Miglio & Montalbán 2005; Montalbán & Miglio 2006). It is possible to obtain as an output of an MCMC simulation the PPD for any function of the parameters.⁴ We can thus estimate the probability of existence of a convective core $P(r_{\text{cc}} > 0)$ [or similarly $P(m_{\text{cc}} > 0)$], with r_{cc} the radius of the convective core (and m_{cc} its mass). The distribution for the radius of the convective core obtained from Run 5 is shown as an example in Fig. 2. We plotted the distribution in the form of a histogram and as a continuous density obtained from kernel estimation. However, we note that it is difficult to find an adequate representation since it is obviously discontinuous at $r_{\text{cc}} = 0$.

In the runs not including seismic constraints, we find a probability of having a convective core of 10 per cent and 2.5 per cent for Runs 1 and 2, respectively. In Run 3, for which the Gaussian prior on the mass was used, this probability becomes null. We shall note right now that, regardless of the observation vector chosen, we never identify a model with a convective core when using this prior. For a uniform prior on the mass, we can again draw a line between estimates, depending whether $\langle \delta\nu \rangle$ or $\langle \Delta\nu \rangle$ has been included in X . In the former case, Runs 5 and 7, we find a probability for the existence of a convective core of 16.5 per cent and 11.2 per cent, respectively. In the latter case, it becomes null. As we shall see in Section 5.2.3 this is partly due to the sensitivity of $\langle \delta\nu \rangle$ to the innermost regions of the star, but it is also a bias due to our assumptions on Z_0 and X_0 .

These results also allow us to derive an upper limit on the radius and the mass of a possible convective core. If we base ourselves on Run 5, we find $r_{\text{cc}} \lesssim 0.040 R_*$ and $m_{\text{cc}} \lesssim 0.015 M_*$ (these values are compatible with the results from Run 7). Note also that the distribution of r_{cc} is not Gaussian anymore. We can see here an obvious advantage of the Bayesian approach: if one wants to estimate uncertainties on such physical parameters, it is not possible to rely on point-estimation method such as likelihood maximum and use, for instance, Hessian inversion to derive them. One has to sample the entire distribution for this quantity to be in position to provide with robust statistics.

⁴ Once we have estimated the joint PPD for θ , we also have an estimate of the PPD of any quantity h such that $h = h(\theta)$.

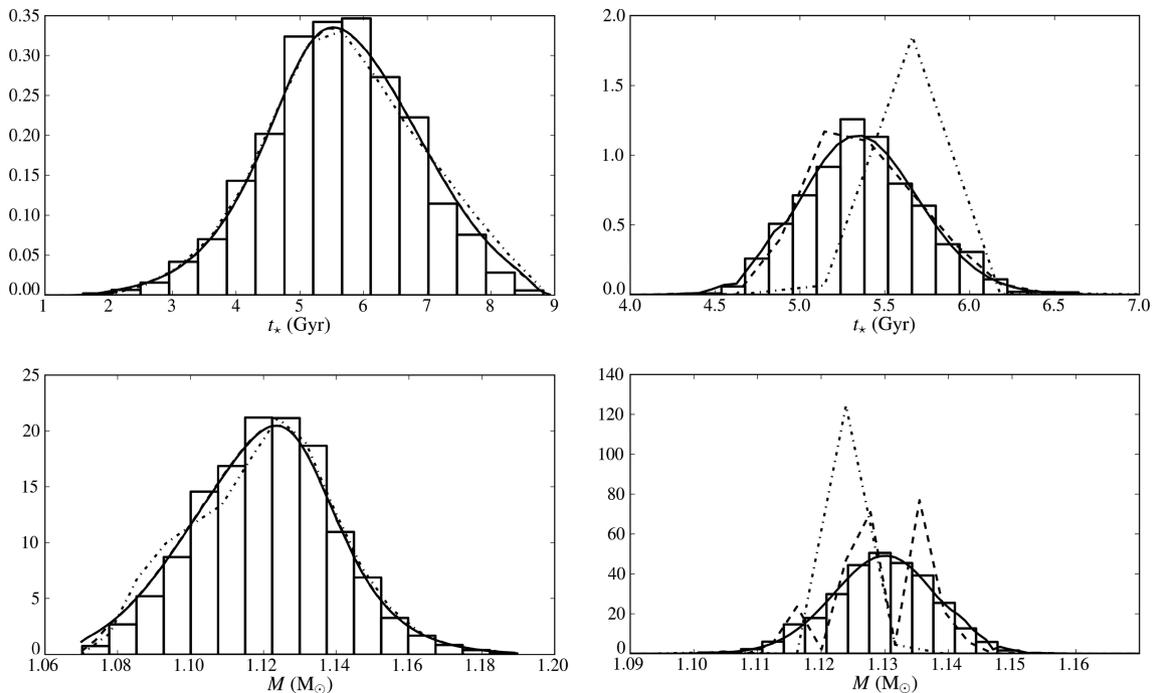


Figure 3. Left-hand column: marginal PPDs for the mass (bottom graph) and the age (top graph) of α Cen A. The observational constraints are $X = (T_{\text{eff}}, L)$ and the prior on the mass is uniform. The histogram represents the results from the MCMC simulation. The lines represent the same marginal PPD computed from grids 2 (full line), 3 (dashed line) and 4 (dash-dotted line). Right-hand column: same as the left-hand column with $X = (T_{\text{eff}}, L, \langle \delta v \rangle)$.

From Fig. 2, we can also obtain a general picture of the possible evolutionary state of the convective core of α Cen A. It displays the size of the convective core as a function of time of the stellar age, the dots are the results from the MCMC simulation and the dashed line gives $r_{\text{cc}} = r_{\text{cc}}(t_*)$ for $M = 1.14 M_{\odot}$, $1.15 M_{\odot}$ and $1.16 M_{\odot}$. It appears that the convective cores are in a growing phase. This mirrors the fact that we tend to find younger age estimate when using $\langle \delta v \rangle$. The growing phase is relatively longer than the receding one, explaining the significant probability obtained for the existence of a convective core in Run 7 is explained by the higher age and only a handful of models are being selected which are in the shorter time-scale receding phase of the convective core.

5.1.4 Comparison with grid-based integration

Sampling only M and t_* allowed us to compute a very dense grid in the region of interest of the space of parameters. Using such a grid, it is straightforward to integrate directly the denominator in the right-hand side of equation (5), which simply is

$$m(X) = \int_{\mathbb{P}} \pi(\theta) \mathcal{L}(X|\theta) d\theta, \quad (16)$$

using the classical quadrature formula (the trapezoidal rule in this case). Similarly, we can integrate out the nuisance parameters to obtain the marginal PPDs using the formula of the same type as (15).

To proceed to this comparison, we computed a grid of approximate size $N_{\text{g}} = 216 \times 216$ in the $(M - t_*)$ plane. It was chosen so that the total number of computed models is close to 45 000, which is the size of our MCMC sample when the typical burn-in sequence of 5000 models is removed. In fact, the total number of computed models is lower than N_{g} because not all points in the

$(M - t_*)$ plane correspond to stable solutions for our stellar evolutionary code. Stars massive enough might indeed reach the giant branch well before the required t_* . Models producing numerical convergence errors during these phases of the stellar evolution were discarded. The effective number of computed models in our grid is $N_{\text{g}}^{\text{eff}} = 45\,533$.

Fig. 3 shows a comparison between the marginal PPDs obtained with our MCMC simulations and from direct integration from this grid. We selected two example cases with $X = (T_{\text{eff}}, L)$ and $X = (T_{\text{eff}}, L, \langle \delta v \rangle)$. The MCMC and grid-integrated distributions are almost identical. This can be checked graphically or by comparing the first moments of the marginal densities given in Table 2. They agree within a few per cent at most. This was expected and is a good test of the self-consistency of our Bayesian approach that two radically different numerical strategies converge. At such a level of agreement, the choice of the method should be made on the basis of the computational performances. Producing dense grids of stellar models in two-dimensional parameter spaces is relatively inexpensive; therefore this should be the approach of choice compared to the MCMC approach.

The fact that we are using a dense grid is critical. When considering five-dimensional parameter spaces we will not have this possibility. Therefore, we use the computational flexibility we have in a lower-dimensional parameter space to illustrate the effect of sparser sampling on the calculation of the PPD. For this, we have selected subgrids in \mathbb{P} by removing points from our original one. They have dimensions 216×216 , 108×108 , 32×32 and 16×16 . As mentioned before, some models could not be computed with *ASTEC* and the effective numbers of points are, respectively, 11 397, 936 and 245. We repeated the integration of the PPDs to obtain the corresponding marginal densities and their first two moments. We note that the original grid covers the $(M - t_*)$ plane with regular cells of dimension $5.5 \times 10^{-4} M_{\odot} \times 37 \text{ Myr}$. The sparsest subgrid has

cells of size $7.7 \times 10^{-3} M_{\odot} \times 518 \text{ Myr}$. Even these large-sampling units offer very acceptable coverage of the space of parameters, in particular with regard to commonly used grids which use steps in mass⁵ $\gtrsim 0.01 M_{\odot}$ (see e.g. Quirion et al. 2010).

The main result is that downsizing the sampling can significantly affect the outcome of the estimation process. The values of the PMs and standard deviations for all selected subgrids are reported in Table 2. For the first two subgrids, the results agree fairly well with the MCMC simulations and the dense-grid integration. However, the results for the sparsest grid show significant discrepancies with these reference estimates. The estimated PMs can depart from the reference ones by a few per cent (around 3–5 per cent for the worst cases). We thus have a numerically induced loss of accuracy. This down-sampling also affects the estimates of the uncertainties (which can be seen as the precision on the parameters and should be distinguished from the accuracy). In some cases, they can be overestimated. For instance, in Run 1 the uncertainty on the age is 30 per cent larger when using grid 4. Conversely some uncertainties can also be underestimated. This is significant since we see departures ranging from the MCMC and dense-grid estimates ranging from a factor of ~ 3 (Run 5) to several orders of magnitudes (Runs 6, 8 or 10). These very low uncertainties sometimes obtained with grid 4 simply mean that only few points in our sampling scheme correspond to non-zero values of the PPD. In these cases, the numerical integral we perform to compute the second moment returns unrealistically low values. This can be seen as a numerical error introduced by the sampling itself.

In general, and unsurprisingly, the sharper the peaks of the densities, the more important this phenomenon is. This reflects that the sampling in the parameter space is not dense enough to capture properly the variations of the PPD. This in turn impacts the numerical approximation to the integrals as can be seen in Fig. 3. It appears clearly that the estimated densities depart strongly from the correctly sampled cases (MCMC or dense grids) for $\mathbf{X} = (T_{\text{eff}}, L, \langle \delta \nu \rangle)$, less so for $\mathbf{X} = (T_{\text{eff}}, L)$.

Note that we chose the smallest subgrid because it has $N_{\text{g}}^{\text{eff}} \sim 16^2 = 256$ points. This is representative of the density of points we have in our five-dimension grids. It is also noteworthy that this behaviour is only indicative of the problem we might encounter when dealing with higher dimension. It is not possible to extrapolate fully the results of the two-dimensional to the five-dimensional case.

5.2 MCMC simulations with $\dim(\mathbb{P}) = 5$

5.2.1 Implementation details

Our next step is to increase the number of parameters left free to vary. As noted in Section 5.1.1, considering only the stellar mass and age as free parameters implies strong assumptions on other parameters. In particular, we needed to extrapolate the mixing-length and the initial hydrogen mass fraction from their solar values, which is always a delicate procedure that should ideally be restricted to stars lying in a close vicinity of the Sun in the space of parameters (such as solar twins; see e.g. Bazot et al. 2011). Most of the grids of

stellar parameters offer more than two values to sample from (for a recent example, see Quirion et al. 2010). However, increasing the dimension of \mathbb{P} comes at a computational cost, namely that, all other things held constant, the number of values to consider in a volume element grows exponentially.

Our goal here is twofold. We first want to estimate M , t_* , Z_0 , α and X_0 together for α Cen A and determine the impact of relaxing the last three with regard to the results from Section 5.1. We then try to evaluate whether we can gain in efficiency when using the MCMC algorithm with respect to the grid approach in the five-dimensional case.

The physical settings of *ASTEC* remain the same. The initial metallicity, the mixing-length parameter and the initial hydrogen mass fraction are now left to vary. The relevant domains are $[0.015, 0.036]$ for Z_0 , $[1.0, 3.0]$ for α and $[0.6, 0.8]$ for X_0 . These limits were set after several trial runs. The Gaussian components of the instrumental law $q(\boldsymbol{\theta}^* | \boldsymbol{\theta}^{(i)})$ were set to $\boldsymbol{\zeta}_1 = (5 \times 10^{-3} M_{\odot}, 0.1 \text{ Gyr}, 5 \times 10^{-5}, 1 \times 10^{-2}, 1 \times 10^{-3})$ and $\boldsymbol{\zeta}_2 = (2.5 \times 10^{-2} M_{\odot}, 0.5 \text{ Gyr}, 5 \times 10^{-4}, 5 \times 10^{-2}, 1 \times 10^{-2})$.

In order to limit the computational costs, we restricted the combinations we considered in the observable space for \mathbf{X} . We used the full set of non-seismic constraints and tested them together with $\langle \delta \nu \rangle$ and $\langle \Delta \nu \rangle$. For each case, we considered a Gaussian prior on the mass. For $\mathbf{X} = (T_{\text{eff}}, L, R, \langle \delta \nu \rangle)$ we also studied the effect of a uniform prior on the mass. The prior on the age, the metallicity, the mixing-length parameter and the initial hydrogen fraction are always uniform. Our Markov chains are approximately three to four time longer than in the two-dimensional case, ranging within 150 000–200 000 iterations. The burn-in sequence is also longer. After testing several values for this latter, we chose to discard a generic 10 000-iteration sequence. We also used eight parallel chains for each run, increasing the computational cost but providing with a possibility to check the convergence of our simulations with some robustness.

5.2.2 Estimates of the parameters

First and foremost, it should be noted that the efficiency of all the strategies decreases when $\dim(\mathbb{P})$ increases. This has practical consequences on the subsequent statistical analysis. It becomes difficult to correctly approximate the PPD for five stellar parameters in a reasonable amount of time. This is due to two entangled factors: (i) stellar models take relatively long to compute and this can become a limiting factor when increasing the number of points to sample in \mathbb{P} , (ii) with increasing dimension, the shape of probability densities might become complex; this is expected to be even more critical for stellar models, which are highly non-linear in their parameters, hence being more difficult to sample properly from the simpler distributions used for the instrumental law.

A simple manifestation of these effects can first be seen in Fig. 5. We plotted the marginal densities of the five parameters for Run 13. It is obvious by simply comparing these densities that one has to explore a much larger volume in \mathbb{P} . This translates into a much higher uncertainty on the mass obtained from Run 12 (compared to Run 7) or on the age from Run 13 (compared to Run 9). Secondly, the departures from Gaussianity become more important than what was observed in the two-dimensional case. Thirdly, one has to consider that these values have been obtained for chains of length $\sim 150\,000$. We controlled our results using multiple parallel chains and noticed that, in some cases, convergence does not seem to be reached. This might be due to one particular chain being stuck in

⁵ The problem of the stellar age sampling is a bit particular. Most of the time, these grids are data base of *stellar evolutionary tracks*. This means that the stellar age is imposed by the stellar evolution code itself and not the user. We still needed to maintain a strict control of this parameter in order to compare properly with MCMC algorithm.

Table 3. Inferred stellar parameters constrained using seismic and non-seismic observables. First column: the number of the run; second column: observational constraints included in X ; third column: prior on the mass; Columns 4–8: PM estimates and associated credible intervals (the distributions are assumed Gaussian) for M , t_* , Z_0 , α and X_0 respectively.

Run no.	X	$\pi(M)$	\hat{M} (M_\odot)	\hat{t}_* (Gyr)	Z_0	α	X_0
12	$[L, T_{\text{eff}}, R, \langle \delta\nu \rangle]$	Uniform	1.12 ± 0.03	4.8 ± 0.5	0.027 ± 0.003	1.7 ± 0.1	0.69 ± 0.02
13	$[L, T_{\text{eff}}, R, \langle \delta\nu \rangle]$	Gaussian	1.105 ± 0.007	4.8 ± 0.5	0.026 ± 0.003	1.62 ± 0.08	0.68 ± 0.02
14	$[L, T_{\text{eff}}, R, \langle \delta\nu \rangle, \langle \Delta\nu \rangle]$	Gaussian	1.102 ± 0.005	4.8 ± 0.5	0.026 ± 0.003	1.60 ± 0.10	0.69 ± 0.02

a local maximum or not exploring efficiently enough the parameter space. Furthermore, the variance between the multiple chains is higher than in the two-dimensional case (as could be seen using the between-chain variance B_T defined in Appendix A). Practically, this is reflected by the rounding of our estimates in Table 3. In more intuitive terms, this can be interpreted as a loss of numerical precision between the two- and five-dimensional cases, Run 12 being a clear example of this effect.

Table 3 also offers good insight into the impact of observational data. The most striking fact is the huge effect the prior on the mass has. With a uniform prior, the precision on the mass decreases by almost an order of magnitude. We should also note the numerical advantage offered by the use of the Gaussian prior: by simplifying the target PPD it greatly favours the convergence properties of the MCMC algorithm.

Since the mass measurements can be interpreted in terms of prior density, their impact is relatively straightforward to evaluate. The relation between the other observables and the stellar parameters is much more difficult to unveil. Theoretically, it is expected that the average small separation should be a good constraint for the stellar age. To test this hypothesis without performing as many runs as we did in Section 5.1 (and also because the use of $\langle \Delta\nu \rangle$ alone was impractical, see below), we performed another MCMC run (Run 14) with the observable vector $X = (T_{\text{eff}}, L, R, \langle \delta\nu \rangle, \langle \Delta\nu \rangle)$. The comparison of results with Run 13 is particularly interesting. It appears clearly that the impact of the small separation is strong, and in fact much more constraining than the large separation. Comparing Run 12 and Run 13, it is clear that the mass parameter is mostly constrained by the prior, and then it appears that the small separation is the observable that strongly controls the estimates of the other parameters.

The use of just $\langle \Delta\nu \rangle$ introduces a new problem. Studying the marginal PPD for t_* , it appears that most of the models lie close to the upper limit of the domain we defined. This means that our prior is very likely to be incorrect. However, when increasing the range for t_* , the MCMC algorithm does not converge anymore. This is seen by the fact that our eight chains are stuck in very different regions of the space of parameters. Note that the lack of convergence when sampling a larger volume of \mathbb{P} could be identified thanks to our use of several parallel Markov chains. Indeed, if considered individually, chains getting stuck in a local maximum could appear as the solution of our inverse problem. This shows the importance of a careful monitoring of the convergence of MCMC algorithm. This issue might be resolved with more sophisticated version of the MCMC algorithm (Robert & Casella 2005; Gregory 2005; Benomar et al. 2010; Handberg & Campante 2011).

5.2.3 Structure of α Cen A

The picture of the stellar interior we obtain from the MCMC simulations in the five-dimensional parameter space is sensibly different from those deduced from the two-dimensional exercise. This

is particularly true when looking at the potential presence of the convective core. The mass prior and the atmosphere-dependent observables do not play such a critical role as they did in Section 5.1.3 in discarding the possibility of convection at the centre of α Cen A.

As before, we find a non-negligible probability of having a convective core when the only seismic observable included in X is the average small separation. We note that this probability might vary slightly from one chain to the other, but we still obtain an approximate probability of existence for the convective core $P(m_{\text{cc}}, r_{\text{cc}} \neq 0 | X) \sim 0.40$. This is much higher than the value obtained in the two-dimensional case. Adding the average large separation does not change this number; the probability being in that case ~ 42 per cent. We can again put an upper limit on the mass and radius of a possible convective core. For Run 12 we obtain $r_{\text{cc}} \lesssim 0.059 R_*$ and $m_{\text{cc}} \lesssim 0.035 M_*$, and for Run 13 $r_{\text{cc}} \lesssim 0.055 R_*$ and $m_{\text{cc}} \lesssim 0.030 M_*$. The difference being small, it shows that it is really the use of a precise value of $\langle \delta\nu \rangle$ that controls the estimated characteristics of the convective core. In Fig. 4 we show the posterior probability densities obtained for r_{cc} from Runs 12 and 13. We see some small departures from Gaussianity. If we consider the distribution for non-null radius and mass and compute the MAP and its associated credible set,⁶ we can conclude that an existing convective core in α Cen A would have $r_{\text{cc}} = 0.044^{+0.007}_{-0.011} R_*$ and $m_{\text{cc}} = 0.014^{+0.005}_{-0.002} M_*$ from Run 12. Again, this is very close to the values we obtain from Run 13, $r_{\text{cc}} = 0.044^{+0.007}_{-0.08} R_*$ and $m_{\text{cc}} = 0.016^{+0.002}_{-0.003} M_*$ from.

In Fig. 5, we represent, for each individual parameters the range within which are observed convective cores in Run 13. We see that they span the entire probability densities for the mass, the age and the mixing-length parameter. We can on the other hand determine lower and upper limits for, respectively, Z_0 and X_0 for the onset of convection. The resulting condition is $Z_0 \gtrsim 0.23$ and $X_0 \lesssim 0.69$. It is interesting to note that this upper limit for X_0 does not correspond to the value we adopted in the two-dimensional case. Whereas in Section 5.1.3 we obtain some models with $X_0 = 0.7$ and a convective core, it is never the case in the five-dimensional case. This means that when these parameters are anticorrelated and that when the initial hydrogen fraction is as high as we assumed when we fixed it, then Z_0 is much lower than the value of 0.027 we assumed above. Therefore, the models with convective cores that we observed in both cases do not recover the same regions of the parameter space \mathbb{P} . This shows how difficult it is to make a priori assumptions about the stellar parameters and how misleading this can be in some cases.

⁶ In this case, we could not assume that the densities were Gaussian, therefore we used the MAP as the estimate of the age and defined the $(1 - \eta)$ per cent credible set as $C_k = \{\theta_k \in \mathbb{P}_k : \pi(\theta_k | X) \geq q(\eta)\}$, with q the smallest constant such as $P(C_k | X) > 1 - \eta$, and \mathbb{P}_k the subspace of \mathbb{P} corresponding to parameter θ_k .

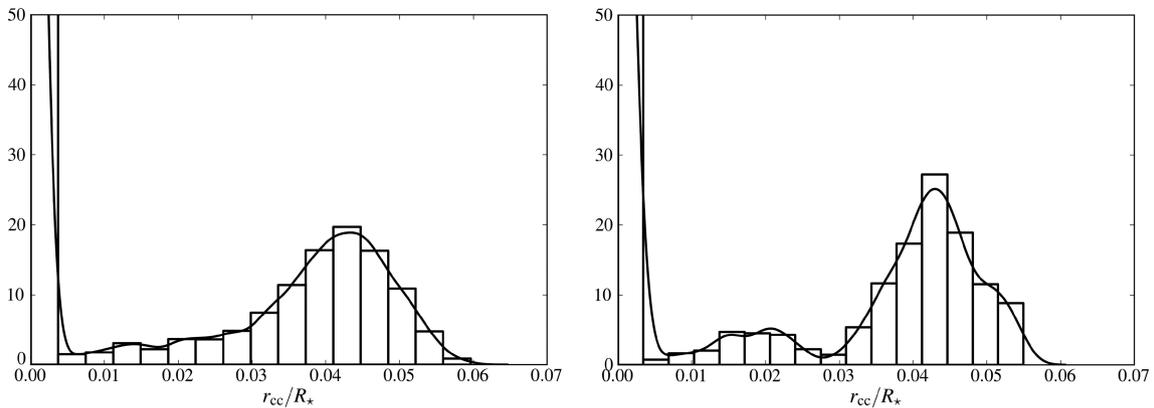


Figure 4. Left-hand panel: posterior probability density for the radius of the convective core of α Cen A estimated using $X = (T_{\text{eff}}, L, R, \langle \delta v \rangle)$ and a Gaussian prior on the mass (Run 12). The PPD is represented in the form of a histogram and as a kernel-estimated continuous density. Right-hand panel: same as the left-hand panel for $X = (T_{\text{eff}}, L, R, \langle \delta v \rangle, \langle \Delta v \rangle)$ (Run 13).

The general picture obtained from the five-dimensional modelling is that the critical parameters for the existence of a convective core are those controlling the chemical composition. Interestingly enough, the precise value of the mass is not relevant for the formation of these structures, in the sense that the distribution of the m_{cc} and r_{cc} for α Cen A is independent of M . This is slightly surprising since the appearance of a convective core is linked to the transition from the pp cycle to the CNO cycle as the main source of energy production at the stellar centre. On the zero age main sequence, when the temperature is high enough, the energy production rate of the latter (ε_{CNO}) overcomes the one from the former (ε_{pp}) by several orders of magnitude. For a given chemical composition, ε_{CNO} increases very rapidly with the temperature and even the modest rise in central temperature with mass is enough to trigger the burning of hydrogen through the CNO cycle. The substantial increase in the nuclear energy rate that results leads in turn to an increase in ∇_{rad} , hence to convectively unstable layers. However, in our results, we observe no correlation between the mass of α Cen A and the existence (or size or mass) of a convective core. It is noteworthy that this result also holds for Run 12. This means that even though the absolute value of the mass is important (low-mass stars will almost always be in the solar or M-dwarf regimes; higher-mass stars will always have a convective core), at the precision level obtained on this parameter,⁷ its potential variations will not affect the development of the inner convective instabilities.

We thus need to examine the influence of the other stellar parameters on the convective core formation process. As for the mass, we observe them for all considered values of t_* and α . However, the situation is different for the two parameters. The mixing-length parameter is essentially decoupled from the mass and radius of the convective core. For the age, we need to look more closely to the distribution $\pi(t_*, m_{\text{cc}} | X)$, which is represented in Fig. 6(a). We can see that, for $m_{\text{cc}} > 0$, it is skewed towards older ages. One thus cannot only rely on the extremum age values at which convective cores are observed and given in Fig. 5, since the upper limit is only representative of a tail of the two-dimensional distribution $\pi(t_*, m_{\text{cc}} | X)$. We will discuss below the implication of this preliminary remark.

We now turn our attention towards X_0 and Z_0 . It must first be noted that these parameters are strongly anti-correlated (see Fig. 6b).

Therefore, in the following discussion, the effect of an increase in Z_0 can be directly translated into a decrease in X_0 . We should also recall that on the main sequence, since only hydrogen is burned and helium produced, and because we have neglected microscopic diffusion, the initial metallicity corresponds to the final metallicity. The main physical variables governing the energy rate (see e.g. Kippenhahn & Weigert 1994) are the temperature, T , the density ρ , the hydrogen mass fraction, X , and, in the case of the CNO cycle, the CNO-element mass fraction X_{CNO} (when we refer to their central values, we label them with the subscript ‘c’). It is our goal here to understand how they are affected by the initial chemical composition of α Cen A. It can be seen in Fig. 6(c) that ultimately Z_0 is correlated to ε_c (which is $\varepsilon_{\text{pp,c}} + \varepsilon_{\text{CNO,c}}$) and this is the relation we want to elucidate.

Figs 6(c) and (d) clearly show that two families of models exist for α Cen A. In the first one, the metallicity is between approximately 0.020 and 0.026. In this interval, the nuclear energy production rate does not increase much whereas the central temperature is positively correlated with the metallicity. In the second regime, i.e. for $0.034 \gtrsim Z_0 \gtrsim 0.026$, ε_c increases with Z_0 . In contrast, T_c does not seem to increase anymore and an absolute maximum in the marginal PPD is reached at temperatures $\sim 1.86 \times 10^7$ K.

We can interpret these two regimes as reflecting the energy generating cycle at play in the central layers of α Cen A. The first one corresponds to the pp cycle. The increase in temperature with Z_0 corresponds to an opacity effect in the higher layer of α Cen A. On the one hand, enhancing the metallicity increases the opacity in these regions.⁸ On the other hand, we observe that the surface luminosity is correlated with neither Z_0 , ε_c nor T_c . Therefore, to produce the same surface luminosity the temperature near the core needs to increase so that ε_c does too. This rise in T_c has relatively modest impact on the energy production rate, which is expected for the pp cycle.

Models of α Cen A with metallicities higher than $\gtrsim 0.026$ have clearly a different behaviour. We consider that they are stars for which the CNO cycle has already become the main source of nuclear energy in the central regions. The fact that the temperature does not increase with metallicity anymore is precisely due to the sensitivity

⁷ Even in the case of stars other than α Cen A for which the presence of a convective core is dubious: precisions on the mass of the order of what we report in Run 12 are now routinely claimed from asteroseismic studies.

⁸ Most likely through an increase in H^- opacity at the surface, which is sensitive to the metallicity. Other sources, like bound-bound transitions or ionization of heavy atoms in deeper layers, might also have an effect, albeit smaller.

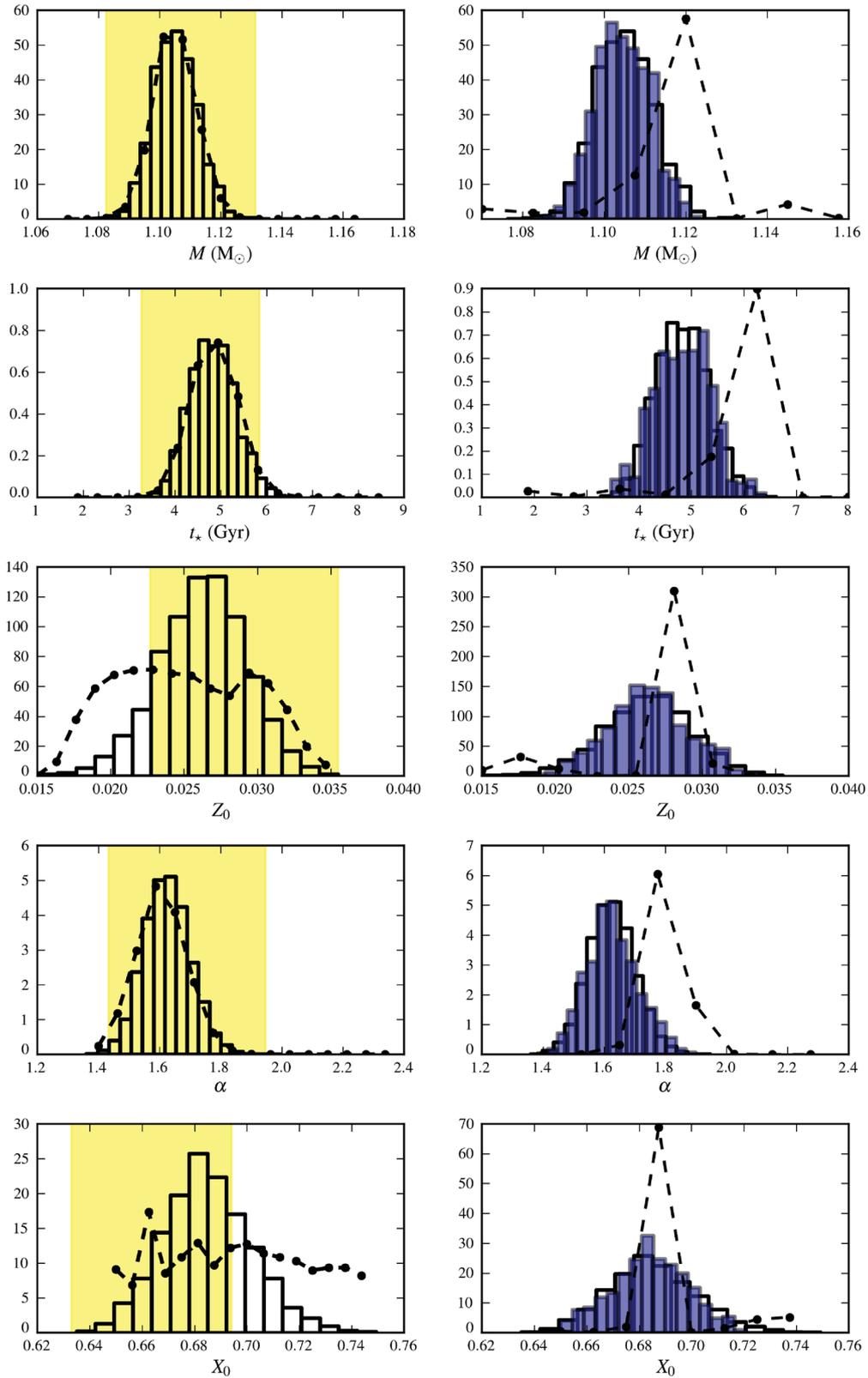


Figure 5. Left-hand column: the histograms represent the marginal PPDs for M , t_* , Z_0 , α and X_0 obtained from Run 13 [$X = (T_{\text{eff}}, L, R, \langle \delta v \rangle)$, Gaussian prior on the mass]. The dots represent the same distribution but obtained from the integration of a grid of size $\sim 16^5$. The yellow-shaded area marks the range for each parameter within which models with convective cores have been obtained. Right-hand column: the white histograms are the same as in the left-hand column. The blue histograms are the marginal PPDs but with a reduced Markov chain. The dots represent the same marginal PPDs but obtained from the integration of a subgrid of $\sim 8^5$ models.

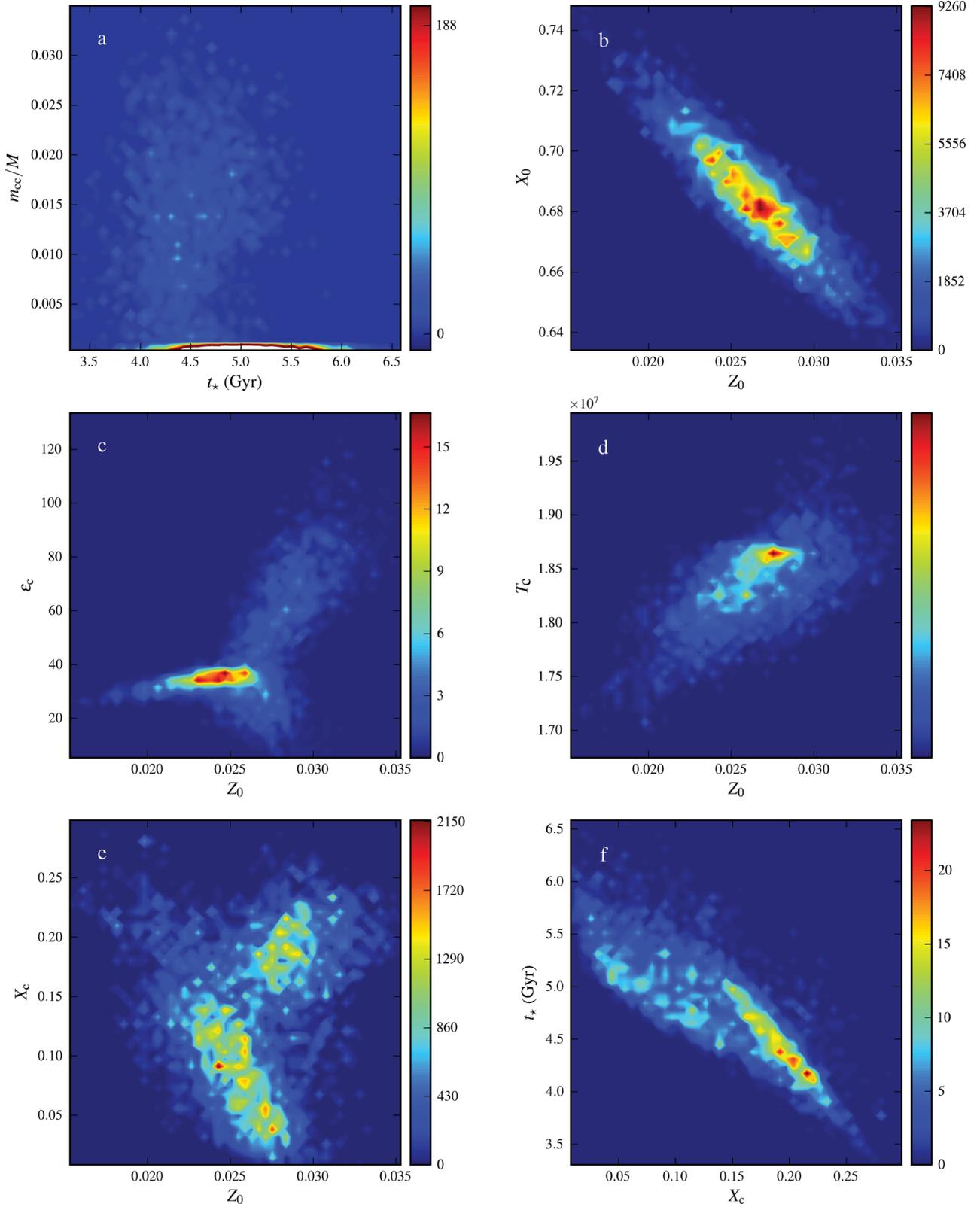


Figure 6. Two-dimensional marginal posterior probability densities for (a) the age of α Cen A, and relative mass of its convective core (note that the PPD has been truncated in order to obtain sufficient contrast where $m_{cc} > 0$), (b) the initial metallicity and hydrogen mass fraction, (c) the initial metallicity and central energy production rate, (d) the initial metallicity and the central temperature, (e) the initial metallicity and central hydrogen mass fraction and (f) the central hydrogen mass fraction and age.

Table 4. Same as Table 3 but with estimates obtained from the grid described in Section 5.2.4.

Run no.	X	$\pi(M)$	$\hat{M} (M_{\odot})$	\hat{t}_{\star} (Gyr)	Z_0	α	X_0
12	$[L, T_{\text{eff}}, R, \langle \delta\nu \rangle]$	Uniform	1.12 ± 0.03	4.9 ± 0.5	0.026 ± 0.005	1.6 ± 0.1	0.70 ± 0.02
13	$[L, T_{\text{eff}}, R, \langle \delta\nu \rangle]$	Gaussian	1.105 ± 0.007	4.9 ± 0.5	0.025 ± 0.005	1.61 ± 0.08	0.70 ± 0.03
14	$[L, T_{\text{eff}}, R, \langle \delta\nu \rangle, \langle \Delta\nu \rangle]$	Gaussian	1.102 ± 0.005	5.0 ± 0.5	0.024 ± 0.004	1.58 ± 0.08	0.70 ± 0.02

of ε_{CNO} to the temperature. Models with higher temperatures would have had much too large luminosities for α Cen A. Conversely, the fact that ε_{C} becomes strongly correlated with the initial metallicity can be explained by an increase in $X_{\text{CNO,c}}$, on which the energy production rate depends linearly. Indeed, this quantity is considered to be proportional to Z_0 and hence Z . Since X_{CNO} can approximately double in the considered range for the metallicity, we are merely seeing the effect of this increase in the right part of the distribution in Fig. 6(c). It is extremely interesting to note that convective cores are present in α Cen A models almost immediately as the CNO cycle becomes dominant. It is the great concentration of the energy generation rate near the centre that triggers convection. Therefore, the CNO peaks of distributions in Figs 6(c) and (d) can be related straightforwardly to the peak for $m_{\text{cc}} > 0$ (Fig. 6a).

Finally, Figs 6(e) and (f) show that these two regimes also correspond to an age effect and we observe a clear bimodality in the distribution for (Z_0, X_c) . Models with initial metallicity $\gtrsim 0.026$ have a higher X_c and are thus younger than those with lower metallicity. This is confirmed by Fig. 6(f), which, besides displaying the well-known age–central hydrogen fraction correlation, also shows that in the CNO mode X_c decreases less rapidly with age than in the pp regime. This is related to the presence of a convective core in these models, which allows the less hydrogen-depleted regions of the star to act like a reservoir for the centre. We can also note that in the CNO regime, X_c seems to be positively correlated with Z_0 . Therefore, there is a slight trend for models with a higher metallicity to be younger. This is clearly the opposite in the pp regime.

5.2.4 Comparison with grid-based integration

The last step of our study was to compare the results from the MCMC simulations, in a similar fashion to what has been done in Section 5.1.4, to the those obtained from integration on a grid of models. According to our preliminary remarks, we had to compute a sparser grid than what we did for the two-dimensional case. We chose to sample the space of parameters using 16 points in each direction. This amounts to 914 749 points in \mathbb{P} , after discarding those which do not lead to a convergent solution using `ASTEC`.

In Fig. 5, we represented the marginal probability densities as obtained from direct integration based on our grid. The results are fairly similar to the MCMC simulation for the mass, the age and the mixing-length parameter. They differ significantly for Z_0 and X_0 . Based on our discussion from Section 5.1.4, we interpret this as a numerical effect stemming from the sparse sampling.⁹ Table 4 also gives the first moment of the distribution obtained by the grid-integration procedure. Note that we chose to display them in the standard form of an estimate and the limits of an associated symmetrical credible interval. We did with the pedagogical purpose

in mind to show that providing such statistical summaries without knowledge of the underlying density might be misleading. Even though the PSDs are close, one can see from Fig. 5 that they do not recover the same reality of our state of knowledge. Therefore, one has to consider some of these values with care: the Gaussian approximation can very well be wrong, especially when dealing with the highly non-linear stellar evolutionary models. In this case of the grid-based determinations of Z_0 and X_0 the PM is not even the best estimator one can choose. One may rather want to use the MAP or the posterior median. The case for the metallicity is even more extreme. We see that the improper sampling leads to a multimodal distribution, and one might be tempted to distinguish two possible scenarios.

These departures between the MCMC and grid-sampling approach are, once more, of much lesser magnitude when the target PPD is broader. Therefore, we conclude that, since no predefined rule for a regular sampling of \mathbb{P} exists, the MCMC approach is able, when the set of observable allows us to constrain enough the PPD of the stellar parameters, to obtain more information on these densities, at an inferior computational cost.

However, one might still be content with the results from the grid-based estimation and consider that they are close enough to what seems to be the correct PPD. Thus, the question we are now facing is to determine, even roughly, how much more efficient the MCMC could be in this five-dimension parameter space and if this performance can be approached by a grid using the same number of model. To that effect, we selected a subgrid by rejecting one point out of two for each parameter. This leads us to a total of $N \sim 8^5$ points sampling the volume of interest in \mathbb{P} . In Fig. 5 (right-hand column), we represented the marginal PPDs obtained from the integration of this subgrid and those from the MCMC simulation using a Markov chain of length equal to the number of points in the subgrid. The result is spectacular, in the sense that the grid approach fails severely to approximate the densities. On the other hand, the reduced MCMC sample is extremely close to the large one. This gives a convincing illustration that, for $\text{dim}(\mathbb{P}) = 5$, the stochastic sample approach offered by the MCMC methodology is much more efficient than the grid strategy. Our conjecture is that this remains valid for $\text{dim}(\mathbb{P}) > 5$.

It indeed is a well-known property of stochastic sampling methodologies that they tend to perform better than more systematic sampling schemes for high-dimensional parameter spaces. However, it was our goal to verify that this was already satisfied for stellar parameter spaces of dimension of at least 5. We also wanted to verify if this could translate into significant computational gains. This example is a very good indication that it is the case and that stellar parameter estimation can benefit greatly from the use of MCMC strategies.

5.3 Comparison with other studies

In this section we briefly compare our results to previous studies using seismic data for α Cen A (Eggenberger et al. 2004; Miglio

⁹ But of course, for real data, one cannot rule out with absolute certitude the possibility that the MCMC algorithm got stuck in a local maximum and could not sample properly the PPD. However, given the fact that our eight independent chains do agree, it seems highly unlikely here to be the case.

& Montalbán 2005; de Meulenaer et al. 2010). It is important to first stress upon the methodological differences between these previous studies and ours. They both rely on optimization approaches, and none of them in a Bayesian framework. This poses obvious problems when comparing the results. First, it is not clear how fixing the mass compares with our use of a Gaussian prior. More importantly, without access to the PPDs of the parameters, it is not possible to assess whether or not the maximum values reported in these studies correspond to local or global extrema. We already saw in Section 5.2 that some PPDs were difficult to sample [in the five-dimensional case, with $X = (L, T_{\text{eff}}, R, \langle \Delta \nu \rangle)$ and a Gaussian prior on the mass], therefore such a configuration might not be uncommon.

Different physics were also used in the stellar codes. The major point is that both studies used microscopic diffusion, which we have neglected. This should have an observable effect, but its nature remains difficult to anticipate. Gravitational settling will deplete the external convective zone of its heaviest elements. The initial metallicity might thus be increased, or the initial hydrogen mass fraction decreased, in order to reproduce the observed Z/X . The impact of this additional mechanism shall be studied with care in future studies using the MCMC approach.

Finally, α Cen B has been modelled simultaneously in these studies. This introduces additional constraints on the solution to the inverse problem. We did not retain this strategy for logistical reasons. Modelling two stars at the same time would have implied significant modifications in our code. We thus preferred to focus on the star with the most reliable data (in particular the seismic data).

We should also consider the effect of the value they used for the average small separation, $\langle \delta \nu \rangle = 8.7 \pm 0.8 \mu\text{Hz}$ used in these studies. To evaluate rapidly the impact this has on the estimate of the age, we can use the grid described in Section 5.2.4. We obtain a relatively Gaussian marginal PPD and a lower age $t_* = 3.84 \pm 0.61 \mu\text{Hz}$ for $X = (L, T_{\text{eff}}, R, \langle \delta \nu \rangle)$ and the Gaussian prior on the mass. Note that although we use the value provided by de Meulenaer et al. (2010), $\langle \delta \nu \rangle = 5.8 \pm 0.1 \mu\text{Hz}$ we obtained $t_* = 6.04 \pm 0.35 \mu\text{Hz}$ for the same X and prior. This agrees with the theoretical behaviour of the small separation, decreasing with age. We should, however, recall that although these grid-based strategies are convenient because the estimates are easily computed once the grid has been produced, we are still affected by the drawbacks we mentioned in Section 5.2.4. The marginal PPDs for the Z_0 , α and X_0 are obviously problematic and a proper MCMC simulation would be necessary (or a denser grid on larger scales) to provide acceptable estimates of these parameters.

Keeping in mind these remarks, we see that the age of α Cen A we estimate in the five-dimensional parameter space (Runs 12 and 13) is significantly lower than those quoted in Miglio & Montalbán (2005). Their A1e model seems to be the more relevant for a comparison with our own results. They obtain an age of 6.7 ± 0.5 Gyr. This is the main departure between the two studies. The other parameters are all within their respective error bars (which are, as a general rule commensurate). However, given the complex interplay between the stellar parameters and other variables such as X_c , a change in one of them, even within the quoted error bars, might very well lead to a very significant departure of another parameter. Only a homogeneous comparison using both models, with the same physical prescriptions, can give us more information on the differences between our studies. The same kind of picture can be drawn from a comparison between our work and the results of Eggenberger et al. (2004).

The problem of the uncertainties is also difficult to tackle. As a general rule, those from Eggenberger et al. (2004) seem to be under-

estimated with respect to ours. However, their method for estimating credible intervals is much less robust and statistically sound than what is allowed from the study of the marginal PPDs of the stellar parameters. Miglio & Montalbán (2005) used a Hessian-based approach and obtained uncertainties commensurate with ours. Our study also provides a confirmation of theirs since they need to make the implicit assumption that they are dealing with normally distributed parameters before inverting their Hessian matrix. Note also that if this is not the case, this methodology is known to underestimate the uncertainties.

One could also contend that the differences come from the physics in our models, in particular the use of the EFF equation of state. This is unlikely since we see in Miglio & Montalbán (2005) that a change of equation of state only modestly affects the final estimates (see for instance their models A1e and A1r). Therefore, we suggest that most of the difference should be attributed to the adopted value for $\langle \delta \nu \rangle$.

Finally, we note that gyrochronology allows us to obtain some measurements of the age of the stars. One should note that these could have been used as a prior on the age. We chose not to do so, in particular because we preferred to constrain the age using seismic quantities rather than by the use of gyrochronology. Nevertheless, we note that our estimates using the average small separations give an age in agreement with the results from Barnes (2007). On the other hand, the estimates from Delorme et al. (2011) are much higher and agree more closely with our estimates using $\langle \Delta \nu \rangle$. It is an interesting fact that the same technique provides such different estimates.

6 CONCLUSION

In this paper, we applied a Bayesian methodology to solve the inverse problem that consists of estimating the stellar parameters of α Cen A, some observable quantities being known. To that purpose, we used an MCMC algorithm in its classical MH form. We repeated the experiment twice, first fitting the data by varying only the mass and the age, and using assumptions to fix the mixing-length parameter and the initial hydrogen fraction and metallicity, and, secondly, letting all five parameters free.

In each case we were able to obtain the posterior probability densities for all considered parameters and their derived values, in particular the mass and radius of the convective core. We observed that different families of solutions emerged depending on whether the observations we used are sensitive to the innermost regions of the star (the average small separation) or to its outer layers (all the others). This remarkable fact allows us in particular to reopen the case for the presence of a convective core in α Cen A. Expressed in probabilistic terms, we obtained non-negligible odds for the existence of such a structure, sometimes reaching over 40 per cent.

From a numerical standpoint, we also produced evidence that the MCMC approach offers a good and efficient alternative to methods generally used in stellar physics, either optimization or grid-based Bayesian approaches. More precisely, when the dimension of the parameter space increases, it becomes less computationally expensive than the straightforward grid integration approach (assuming one wishes to obtain the same statistical output). It also preserves the advantage of Bayesian methodologies with respect to optimization methods in the sense that we have access to the full posterior probability density. Considering that it is already efficient and that there exist other variants of the MH algorithm that can improve the

efficiency of our own, we deem this method as very interesting for the future of stellar parameter estimation.

Finally, it seems clear that further progress on the understanding of the inner structure of α Cen A will require more precise and accurate data. It is in particular our hope that asteroseismology can help to unravel the convective core puzzle. We have already seen that precise values of $\langle \delta\nu \rangle$ greatly affect our answer to this problem. For this, one will need very long and precise radial velocity measurements, if possible continuous. The MCMC methodology will then be of great use to provide robust estimates of the stellar parameters and their uncertainties, knowing these new data.

ACKNOWLEDGMENTS

Michaël Bazot would like to thank S. Hannestad for providing him access to the Grendel cluster at the Danish Center for Scientific Computing, Aarhus University, of which important use has been made during this work. He would also like to thank H. Carfantan for his kind support and J. P. Marques for very interesting and useful discussions. This work was co-supported by grants SFRH/BPD/47994/2008 and PTDC/CTE-AST/098754/2008 from FCT/MCTES and FEDER, Portugal. Funding for the Stellar Astrophysics Centre is provided by The Danish National Research Foundation. The research is supported by the ASTERISK project (ASTERoseismic Investigations with SONG and Kepler) funded by the European Research Council (Grant agreement no.: 267864). Finally, the authors would like to thank the anonymous referee for the suggestions that improved the content of the article.

REFERENCES

- Andrieu C., Doucet A., 1999, *IEEE Trans. Signal Process.*, 47, 2667
- Angulo C. et al., 1999, *Nucl. Phys. A*, 656, 3
- Antia H. M., Basu S., 2005, *ApJ*, 620, L129
- Asplund M., Grevesse N., Sauval A. J., Allende Prieto C., Kiselman D., 2004, *A&A*, 417, 751
- Asplund M., Grevesse N., Sauval A. J., Scott P., 2009, *ARA&A*, 47, 481
- Barnes S. A., 2007, *ApJ*, 669, 1167
- Bazot M., Bouchy F., Kjeldsen H., Charpinet S., Laymand M., Vauclair S., 2007, *A&A*, 470, 295
- Bazot M., Bourguignon S., Christensen-Dalsgaard J., 2008, *Mem. Soc. Astron. Italiana*, 79, 660
- Bazot M. et al., 2011, *A&A*, 526, L4
- Bazot M. et al., 2012, *A&A*, 544, A106
- Bedding T. R., Kjeldsen H., Butler R. P., McCarthy C., Marcy G. W., O'Toole S. J., Tinney C. G., Wright J. T., 2004, *ApJ*, 614, 380
- Benomar O., Baudin F., Marques J. P., Goupil M. J., Lebreton Y., Deheuvels S., 2010, *Astron. Nachr.*, 331, 956
- Böhm-Vitense E., 1958, *Z. für Astrophys.*, 46, 108
- Bouchy F., Carrier F., 2002, *A&A*, 390, 205
- Brewer B. J., Bedding T. R., Kjeldsen H., Stello D., 2007, *ApJ*, 654, 551
- Brown T. M., Christensen-Dalsgaard J., 1998, *ApJ*, 500, L195
- Brown T. M., Christensen-Dalsgaard J., Weibel-Mihalas B., Gilliland R. L., 1994, *ApJ*, 427, 1013
- Caffau E., Ludwig H.-G., Steffen M., Ayres T. R., Bonifacio P., Cayrel R., Freytag B., Plez B., 2008, *A&A*, 488, 1031
- Castro M., Vauclair S., Richard O., 2007, *A&A*, 463, 755
- Chmielewski Y., Friel E., Cayrel de Strobel G., Bentolila C., 1992, *A&A*, 263, 219
- Christensen-Dalsgaard J., 1982, *MNRAS*, 199, 735
- Christensen-Dalsgaard J., 1993, in Brown T. M., ed., *ASP Conf. Ser. Vol. 42, GONG 1992. Seismic Investigation of the Sun and Stars*. Astron. Soc. Pac., San Francisco, p. 347
- Christensen-Dalsgaard J., 2008a, *Ap&SS*, 316, 113
- Christensen-Dalsgaard J., 2008b, *Ap&SS*, 316, 13
- da Silva L. et al., 2006, *A&A*, 458, 609
- de Meulenaer P., Carrier F., Miglio A., Bedding T. R., Campante T. L., Eggenberger P., Kjeldsen H., Montalbán J., 2010, *A&A*, 523, A54
- Delorme P., Collier Cameron A., Hebb L., Rostron J., Lister T. A., Norton A. J., Pollacco D., West R. G., 2011, *MNRAS*, 413, 2218
- Deubner F.-L., Gough D., 1984, *ARA&A*, 22, 593
- Eggenberger P., Charbonnel C., Talon S., Meynet G., Maeder A., Carrier F., Bourban G., 2004, *A&A*, 417, 235
- Eggleton P. P., Faulkner J., Flannery B. P., 1973, *A&A*, 23, 325
- Endl M., Kürster M., Els S., Hatzes A. P., Cochran W. D., 2001, *A&A*, 374, 675
- Gabriel M., 1989, *A&A*, 226, 278
- Gelman A., Rubin D. B., 1996, *Stat. Sci.*, 7, 457
- Girardi L., Bressan A., Bertelli G., Chiosi C., 2000, *A&AS*, 141, 371
- Gough D. O., 1986, in Osaki Y., ed., *Hydrodynamic and Magnetodynamic Problems in the Sun and Stars*. Tokyo Univ., Tokyo, p. 117
- Grec G., Fossat E., Pomerantz M. A., 1983, *Sol. Phys.*, 82, 55
- Gregory P. C., 2005, *Bayesian Logical Data Analysis for the Physical Sciences: A Comparative Approach with 'Mathematica' Support*. Cambridge Univ. Press, Cambridge
- Grevesse N., Noels A., Sauval A. J., 1993, *A&A*, 271, 587
- Gruberbauer M., Guenther D. B., Kallinger T., 2012, *ApJ*, 749, 109
- Guenther D. B., Demarque P., 2000, *ApJ*, 531, 503
- Guzik J. A., Watson L. S., Cox A. N., 2005, *ApJ*, 627, 1049
- Handberg R., Campante T. L., 2011, *A&A*, 527, A56
- Hastings W. K., 1970, *Biometrika*, 57, 97
- Idier J., ed., 2008, *Bayesian Approach to Inverse Problems*. ISTE Ltd, London/John Wiley & Sons, Inc., Hoboken, NJ
- Iglesias C. A., Rogers F. J., 1996, *ApJ*, 464, 943
- Jørgensen B. R., Lindgren L., 2005, *A&A*, 436, 127
- Kervella P., Thévenin F., Ségransan D., Berthomieu G., Lopez B., Morel P., Provost J., 2003, *A&A*, 404, 1087
- Kippenhahn R., Weigert A., 1994, *Stellar Structure and Evolution*. Springer-Verlag, Berlin
- Kjeldsen H., Bedding T. R., Christensen-Dalsgaard J., 2008, *ApJ*, 683, L175
- Metcalfe T. S., 2003, *Ap&SS*, 284, 141
- Metropolis N., 1953, *J. Chem. Phys.*, 21, 1087
- Miglio A., Montalbán J., 2005, *A&A*, 441, 615
- Montalbán J., Miglio A., 2006, *Memorie della Societa Astronomica Italiana*, 77, 472
- Montalbán J., Miglio A., Noels A., Grevesse N., di Mauro M. P., 2004, in Danesy D., ed., *ESA Special Publication, Vol. 559, SOHO 14 Helio and Asteroseismology: Towards a Golden Future*, ESA, Noordwijk, p. 574
- Morel P., 1997, *A&AS*, 124, 597
- Morel P., Provost J., Lebreton Y., Thévenin F., Berthomieu G., 2000, *A&A*, 363, 675
- Murdoch K., Hearnshaw J. B., 1993, *Observatory*, 113, 79
- Neuforge-Verheecke C., Magain P., 1997, *A&A*, 328, 261
- Pietrinferni A., Cassisi S., Salaris M., Castellì F., 2004, *ApJ*, 612, 168
- Pourbaix D. et al., 2002, *A&A*, 386, 280
- Quirion P.-O., Christensen-Dalsgaard J., Arentoft T., 2010, *ApJ*, 725, 2176
- Robert C. P., Casella G., 2005, *Monte Carlo Statistical Methods* (Springer Texts in Statistics). Springer-Verlag New York, Inc., Secaucus, NJ
- Rogers F. J., Nayfonov A., 2002, *ApJ*, 576, 1064
- Rogers F. J., Swenson F. J., Iglesias C. A., 1996, *ApJ*, 456, 902
- Roxburgh I. W., Vorontsov S. V., 1994, *MNRAS*, 267, 297
- Schaller G., Schaerer D., Meynet G., Maeder A., 1992, *A&AS*, 96, 269
- Söderhjelm S., 1999, *A&A*, 341, 121
- Stahn T., Gizon L., 2008, *Sol. Phys.*, 251, 31
- Takeda G., Ford E. B., Sills A., Rasio F. A., Fischer D. A., Valenti J. A., 2007, *ApJS*, 168, 297
- Tassoul M., 1980, *ApJS*, 43, 469
- Thévenin F., Provost J., Morel P., Berthomieu G., Bouchy F., Carrier F., 2002, *A&A*, 392, L9
- Thoul A., Scuflaire R., Noels A., Vatoz B., Briquet M., Dupret M.-A., Montalbán J., 2003, *A&A*, 402, 293

- Trampedach R., 2007, in Stancliffe R. J., Houdek G., Martin R. G., Tout C. A., eds, AIP Conf. Ser. Vol. 948, *Unsolved Problems in Stellar Physics: A Conference in Honor of Douglas Gough*. Am. Inst. Phys., Melville, p. 141
- Turck-Chièze S., Couvidat S., Piau L., Ferguson J., Lambert P., Ballot J., García R. A., Nghiem P., 2004, *Phys. Rev. Lett.*, 93, 211102
- Turcotte S., Christensen-Dalsgaard J., 1998, in Korzennik S., ed., *ESA Special Publication, Vol. 418, Structure and Dynamics of the Interior of the Sun and Sun-like Stars*. ESA, Noordwijk, p. 407
- Vandakurov Y. V., 1967, *AZh*, 44, 786
- Van Hoolst T., Smeyers P., 1991, *A&A*, 248, 647
- Vorontsov S. V., 1991, *SvA*, 35, 400
- Vorontsov S. V., Zharkov V. N., 1989, *Astrophys. Space Phys. Rev.*, 7, 1

APPENDIX A: MONITORING OF CONVERGENCE

A1 Algorithm monitoring

Assessing the convergence of an MCMC algorithm to its target distribution is a vast and complex problem. There is no clear-cut answer to this issue and most of the current methods are empirical ones. We present in this section the few simple indicators we used to monitor convergence in the present study.

A1.1 Burn-in sequence

As can be seen in Algorithm [A.1], the initialization of the process involves the input of an initial value for the parameter. Hence, during a first phase, traditionally called the burn-in, the random generation will not be done according to the target PPD. It is then necessary to discard the first iterations of the Markov chain in order to yield the *stationary* regime. It is difficult to define with precision the length of the burn-in sequence, which obviously depends on the target distribution and on the initial guess.

For an MCMC sample to approximate the targeted PPD, it is necessary that the $\theta^{(t)}$ are generated randomly¹⁰ and, hence, are independent of each other.

To verify this, one can compute the autocorrelation function (ACF), which can be written, for a Markov chain of length N , as

$$\rho^{\theta_i}(k) = \frac{\sum_k^{N-(k+1)} (\theta_i^{(t)} - \bar{\theta})(\theta_i^{(t+k)} - \bar{\theta})}{\sqrt{\sum_k^{N-(k+1)} (\theta_i^{(t)} - \bar{\theta})^2} \sqrt{\sum_k^{N-(k+1)} (\theta_i^{(t+k)} - \bar{\theta})^2}}, \quad (\text{A1})$$

the denominator being a normalization factor. A value close to zero is a good indication that the values of the sequence have indeed been generated randomly. The behaviour of the ACF is grossly that of

¹⁰ Not sufficient though: one has still to determine if they are indeed generated according to the desired PPD.

a decreasing exponential $\exp(-k/\ell_\rho)$, ℓ_ρ being the characteristic length for decorrelation (Gregory 2005). For each run, we tested burn-in sequences of size $\sim \ell_\rho$ and higher. Usually, discarding a number of iteration twice as large as ℓ_ρ guarantees that the sample is indeed stationarily generated.

A1.2 Convergence testing

As mentioned above, MCMC algorithms rely on the convergence properties of the Markov chains. Therefore, the number of iterations has to be large enough in order to guarantee that the Markov chain asymptotic regime is a good approximation to the target PPD. Assuming that an adequate burn-in sequence has been removed, the question arises of the stopping rule.

A classical way to address the problem of convergence is to use empirical graphical indicators. It is particularly interesting to follow the evolution of the empirical cumulative average estimator

$$S_T = \frac{1}{T} \sum_{t=1}^T h(\theta^{(t)}), \quad (\text{A2})$$

where h is an arbitrary function of θ and T the length of the Markov chain. In practice, we will be mostly interested in the function $h(\theta) = \theta_i$ for $i = 1, \dots, n_p$, so that S_T approximates the PM estimate (7). At convergence, variations in S_T (as a function of T) should be negligible.

Depending on the dimension of the parameter space we considered, we ran $M = 3$ or $M = 8$ chains with different initializations. This was done in order to explore fully the space of parameter and eventual dependence on the initial guess. In this configuration, instead of S_T , we consider the average over the M chains:

$$\langle h(\theta) \rangle = \frac{1}{M} \sum_{m=1}^M \frac{1}{T} \sum_{t=1}^T h(\theta_m^{(t)}), \quad (\text{A3})$$

where m indexes the number of the chain, and the associated average variance (*within-chain variance*)

$$W_T = \frac{1}{M} \sum_{m=1}^M \frac{1}{T-1} \sum_{t=1}^T \left(\theta_m^{(t)} - \frac{1}{T} \sum_{t=1}^T \theta_m^{(t)} \right)^2. \quad (\text{A4})$$

One can also define the *between-chain variance*

$$B_T = \frac{1}{M-1} \sum_{m=1}^M \left(\frac{1}{T} \sum_{t=1}^T \theta_m^{(t)} - \frac{1}{M} \sum_{m=1}^M \frac{1}{T} \sum_{t=1}^T \theta_m^{(t)} \right)^2; \quad (\text{A5})$$

this other indicator can be used to check convergence as explained in Gelman & Rubin (1996).

This paper has been typeset from a $\text{\TeX}/\text{\LaTeX}$ file prepared by the author.