Implementing the BAHAMAS Model using Stan

Vatsal Patel
Project Code: ASTR-HEAVENS-1
Supervisor: Professor Alan Heavens
Assessor: Professor Andrew H Jaffe
Word count: 4500

6th May, 2018

Abstract

The BAyesian HierArchical Model for the Analysis of Supernova cosmology (BAHAMAS model) developed by March et al (2016) was implemented to infer the cosmological and Phillips colour and stretch parameters. The Hamiltonian Monte Carlo sampling method was used through a probabilistic programming language called Stan [1]. The results were compared with those obtained by March et al (2016), which employed Gibbs type samplers. Furthermore Stan's performance was evaluated in terms of both the ease of implementation of the model as well as the computational resources required. 311 out of the 740 spectroscopically confirmed supernovae type-Ia (SNIa) from the Joint Light-curve Analysis (JLA) dataset were analysed. The matter density, Ω_m, the curvature parameter, Ω_κ, and the cosmological constant parameter, Ω_Λ, were found to be 0.29 ± 0.13, 0.29 ± 0.31 and 0.43 ± 0.21 respectively. The α and β values which govern the strength of the Phillips stretch and colour corrections were found to be 0.14 ± 0.01 and 2.91 ± 0.13 respectively. These results are consistent with that of March et al (2016) within 1σ. The residual scatter of the intrinsic magnitudes of the supernovae in the sample, σ_{res}, was found to be 0.08 ± 0.005 and inconsistent with March et al (2016) within 1σ. The reason for the discrepancy was not determined, but was argued to be most likely due to the truncated sample size. The implementation of the BAHAMAS model into Stan was found to be relatively straightforward, except during debugging. Running Stan with the full 740 supernovae was found to be too time consuming on a modern average laptop (2.9 GHz Intel Core i5 processor and 8GB RAM) but with further optimisation of code recommended.

Contents

Summary of Report

This projects attempts to implement the BAHAMAS model with Stan to infer the cosmological and Phillips stretch and colour parameters. The introduction section introduces the context of this investigation detailing the historical use of Supernova Type-1a in cosmology and the development of the model itself. The next section explains the theoretical aspect of the model and its hierarchical Bayesian structure. The section also covers the specific SNIa dataset used. Thereafter, in section 3 the statistical techniques employed by Stan including Hamiltonian Monte Carlo sampling and Leapfrog integration are reviewed. Section 3.2 discusses the verification technique used to ensure the model had been correctly implemented. The results are presented in Section 4 and are compared to previous implementations of the BAHAMAS model and wider literature in section 5. Finally a brief summary of the findings are listed in the conclusion section.

1 Introduction

In the late 90's two independent projects, the Supernova Cosmology Project and the High Z Supernova Search Team utilised Supernova type 1a (SNIa) as standard candles, objects of approximate uniform intrinsic luminosity, to present the initial evidence of the accelerated expansion of the universe [2, 3]. Since then the SNIa sample has risen to over a thousand spectroscopically confirmed SNIa due to world-wide observational efforts and they remain one of the key tools in probing the characteristics of this accelerated expansion with time [4].

Presently, the accelerated expansion of the universe is widely attributed to the existence of dark energy, which is compatible with Einstein's cosmological constant [5]. The precise measurement of how the expansion changes with time are key to characterising this dark energy or even establishing if general relativity needs to be modified on cosmological scales. Therefore, better understanding of how the properties of SNIa correlate with their environment is necessary to increase the proficiency of their use as standard candles.

While there is still debate on the progenitor scenarios of SNIa, they are believed to occur when the material accreting onto a white dwarf from a companion, drives the mass of the dwarf above the Chandrasekhar limit, around 1.4 solar masses [6]. This is the maximum mass that can be supported by electron degeneracy pressure and is the point at which the star collapses, triggering the explosion that powers its light curve (their observed apparent magnitudes as a function of time). They are observationally identified by the absence of Hydrogen lines and the presence of strong Silicon II lines in their spectrum.

The critical mass threshold underpins the underlying assumption in the use of SNIa as standard candles; that their intrinsic magnitudes are sufficiently homogeneous. Even so, variability in several factors including composition, rotation rate and accretion rate can lead to significant differences between the different SNIa [7]. Most importantly, the variation in the intrinsic magnitude of the SNIa means that observational data, in its raw form, cannot be used for estimating cosmological parameters. However, empirical

observations have found that intrinsically brighter SN1a have greater decay times for their light curves [8, 9]. Furthermore, it has also been found that brighter SN1a are redder in colour [10]. Collectively known as the Phillips corrections, these empirical corrections can reduce the residual scatter in the intrinsic magnitude to typically $\sim 0.1 - 0.15$ magnitude enabling them to be used as standardised candles.

As the SNIa sample size continues to grow, so does the relative importance of the systematic errors compared to the statistical errors. In fact, the collective SN1a sample is currently of a size where measurements of cosmological parameters using SNIa are actually limited by the systematics. Consequently the discrepancies in their modelling are becoming increasingly scrutinised. One of the main methods used to deduce the empirical corrections from the light-curves is the SALT- II methodology. This splits the process in two steps. First, the Phillips corrections are derived from the light-curves and the cosmological parameters are then constrained in a separate inference step. Despite the increasing sample size and measurements of SNIa, the crucial inference step of deriving cosmological constraints from the SALT-II light-curve fits remained largely unchanged. It suffers from several drawbacks, such as not allowing for rigorous model checking, and not providing a rigorous framework for the evaluation of systematic uncertainties [7]. Consequently in 2011 March et al developed the BAHAMAS model, a statistically principled and rigorous Bayesian hierarchical model, which fitted the cosmological parameters to the SNIa data using the SALT-II light-curve fits [7]. This was then further developed in 2016 by March et al [1].

This investigation's aim is to implement the simplest version of the BAHAMAS model, the Baseline model, using the same dataset. However, instead of using the Gibbs type samplers which March et al employed for the inference step, a probabilistic programming language called Stan is used. Stan uses the Hamiltonian Monte Carlo sampling method (discussed in detail in section 3.1) and offers a supposedly easier and quicker approach to Bayesian model implementation. After obtaining the Phillips corrections and cosmological parameters with Stan the results are then compared to March et al (2016) and wider literature. Stan is then evaluated for practicality, ease of use and the results it produces for this model.

2 The BAHAMAS Model

The BAHAMAS modelling strategy relies on the homogeneity of the SN1a absolute (intrinsic) magnitudes which allows the estimation of their distance moduli from their apparent magnitudes. This can then be used in conjunction with their measured redshifts to obtain estimates of the cosmological parameters which dictate the relationships between them.

2.1 Cosmology

The distance modulus, $\mu(z_i; \mathcal{L})$, is related to the apparent magnitude, $m_i^\star$, and the intrinsic magnitude, M_i, by the following

$$m_i^\star = \mu(z_i; \mathcal{L}) + M_i \quad \text{for } i = 1, ..., n\,, \tag{1}$$

where i refers to the i$^{\text{th}}$ SN1a, z_i is the redshift and $\mathcal{L}$ represents the cosmological parameters.

The distance modulus is also given by

$$\mu(z;\mathcal{L}) = 25 + 5\log\frac{d_L(z;\mathcal{L})}{\text{Mpc}}, \tag{2}$$

where $d_L(z;\mathcal{L})$ is the luminosity distance to redshift. Under the ΛCDM model the luminosity distance is given by

$$\mu(z;\mathcal{L}) = \frac{c(1+z)}{H_0\sqrt{|\Omega_k|}}\text{sinn}_{\Omega_k}\{\sqrt{|\Omega_k|}\int_0^z dz'[(1+z')^3\Omega_m + \Omega_{\text{DE}}(z') + (1+z')^2\Omega_\kappa]^{-0.5}\}, \tag{3}$$

where

$$\text{sinn}_{\Omega_k}(x) = \begin{cases} x, & \text{if } \Omega_\kappa = 0 \\ \sin(x), & \text{if } \Omega_\kappa < 0 \\ \sinh(x), & \text{if } \Omega_\kappa > 0 \end{cases}, \tag{4}$$

Ω_κ is the curvature parameter, Ω_m is the total matter density parameter which includes both baryonic and dark matter, c is the speed of light in km and H_0 is the Hubble parameter today, taken to be 67.3 km/s/Mpc from Planck [11]. Ω_{DE} can be expressed as

$$\Omega_{\text{DE}}(z) = \Omega_\Lambda \exp\left[3\int_0^z \frac{1+w(x)}{1+x}\right], \tag{5}$$

where Ω_Λ is the dark energy density parameter and $w(z)$ is the general dark energy equation of state as a function of redshift. For this investigation a curved universe is assumed with a cosmological constant such that $w(z) = -1$. This leads to w being a time independent constant such that the following relation holds

$$\Omega_\Lambda = 1 - \Omega_\kappa - \Omega_m \tag{6}$$

2.2 Bayesian Hierarchical Structure

The SALT2 fit of the multi-colour light curve observation of SN1a i produces the following measured quantities; the heliocentric redshift, $\hat{z}_i$, the B-band apparent magnitude, $\hat{m}^\star_{Bi}$, the stretch correction parameter, $\hat{x}_{1i}$, the colour correction parameter, $\hat{c}_i$, and the (3×3) variance-covariance matrix, $\hat{C}_i$, describing the measurement errors of $\hat{m}^\star_{Bi}$, $\hat{x}_{1i}$ and $\hat{c}_i$. It has been shown that the observational error of the redshifts do not lead to any appreciable difference in results and thus $\hat{z}_i = z_i$ is set throughout [7].

At the highest level of the hierarchical model the SALT2 fits are modelled as independent Gaussian variables[1] centred at their true values,

$$\begin{pmatrix} \hat{m}^\star_{Bi} \\ \hat{x}_{1i} \\ \hat{c}_i \end{pmatrix} \overset{indep}{\sim} \mathcal{N}\left[\begin{pmatrix} m^\star_{Bi} \\ x_{1i} \\ c_i \end{pmatrix}, \hat{C}_i\right], \quad \text{for } i = 1, ..., n\ , \tag{7}$$

[1] $\mathcal{N}(\mu, \Sigma)$ is used to denote a (multivariate) Gaussian distribution of mean, μ, and variance- covariance matrix ,Σ. For the 1-dimensional case, Σ reduces to the variance, σ^2.

The true values, $m^{\star}_{Bi}$, x_{1i} and c_i, are treated as latent variables.

To incorporate the empirical Phillips corrections, the absolute magnitude, M_i, in equation 1 is modelled as

$$M_i = -\alpha x_{1i} + \beta c_i + M^{\epsilon}_i \,, \tag{8}$$

where α and β are parameters which dictate the strength of the Phillips corrections, x_{1i} and c_i respectively, and M^{ϵ}_i is the empirically corrected absolute magnitude after the Phillips corrections have been applied. This modifies equation 1 so that it becomes

$$m^{\star}_i = \mu(z_i; \mathcal{L}) - \alpha x_{1i} + \beta c_i + M^{\epsilon}_i \quad \text{for } i = 1, ..., n \,. \tag{9}$$

At the lowest level M^{ϵ}_i, x_{1i}, c_i are modelled as Gaussian distributions with unknown hyper-parameters governing the mean and variance of each

$$M^{\epsilon}_i | M^{\epsilon}_0, \sigma_{res} \sim \mathcal{N}(M^{\epsilon}_0, \sigma_{res}) \tag{10}$$

$$x_{1i} | x^{\star}_1, R_{x1} \sim \mathcal{N}(x^{\star}_1, R^2_{x1}) \tag{11}$$

$$c_{\star} | c_{\star}, R_c \sim \mathcal{N}(c_{\star}, R^2_c) \tag{12}$$

The above details of the Bayesian BAHAMAS model can be depicted visually through the graph shown in figure 1.

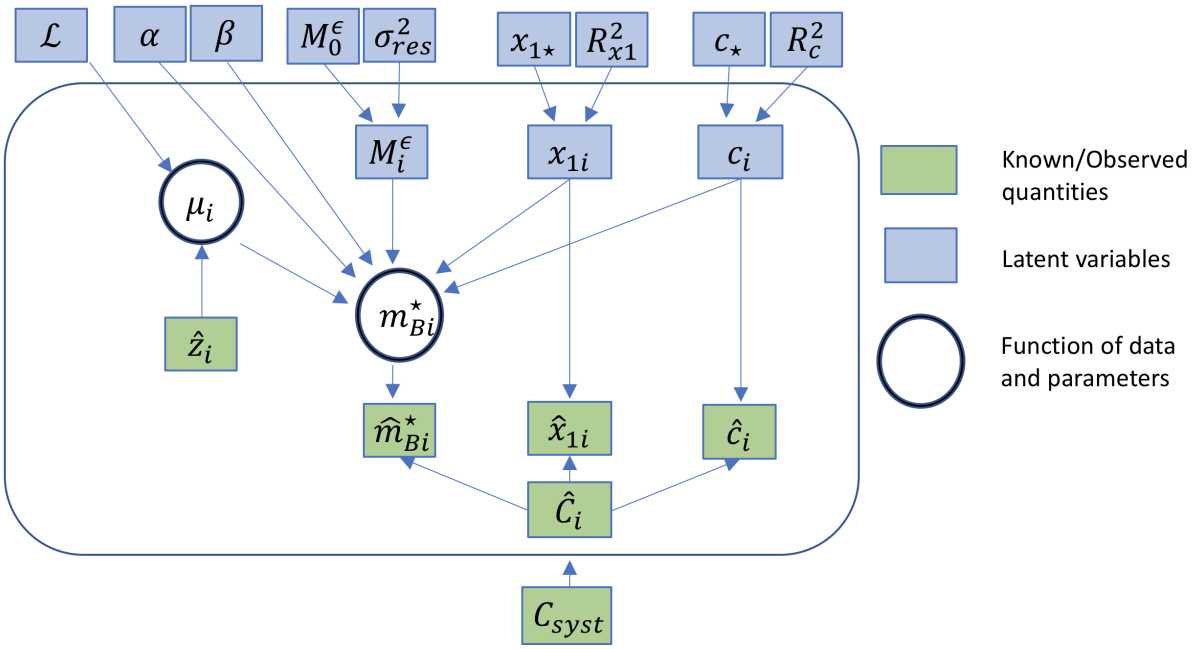

Figure 1: Graphical representation of the BAHAMAS model. The meaning of the symbols for the latent variables are given in section 2.2. C_{syst} represents the full systematic variance- covariance matrix as described in section 2.3

2.3 Covariance Matrix

For this model it is assumed that the SALT2 measurements for each SNIa i are conditionally independent. Therefore the full $(3n \times 3n)$ variance-covariance matrix, $C_{stat} \equiv \text{diag}(\hat{C}_1, \ . \ . \ . \ , \hat{C}_n)$, is block diagonal. A systematic variance-covariance matrix, C_{syst}, with correlations among the SNIa which includes contributions from calibration, model uncertainty, bias correction, host, dust, peculiar velocities and contamination has already been calculated by Betoule et al [12]. This is added onto C_{stat} to give the full $(3n \times 3n)$ matrix which includes both the statistic and systematic errors of the measured quantities.

2.4 Priors

Parameters	**Prior Distribution**
Cosmological Parameters	
Matter density parameter	$\Omega_m \sim \text{Uniform}(0, 2)$
Curvature parameter	$\Omega_\kappa \sim \text{Uniform}(-3, 1)$
Hubble Parameter	$H_0/\text{km/s/Mpc} = 67.3$
Covariates	
Coefficient of stretch covariate	$\alpha \sim \text{Uniform}(0, 1)$
Coefficient of color covariate	$\beta \sim \text{Uniform}(0, 4)$
Population-level Distributions	
Mean of absolute magnitude	$M_0^\epsilon \sim \mathcal{N}(-19.3, 2^2)$
Residual scatter after corrections	$\sigma_{res}^2 \sim \text{InvGamma}(0.003, 0.003)$
Mean of Phillips stretch	$x_{1\star} \sim \mathcal{N}(0, 10^2)$
Standard deviation of Phillips stretch	$R_{x_1} \sim \text{LogUniform}(-5, 2)$
Mean of Phillips colour	$c_\star \ \sim \mathcal{N}(0, 1^2)$
Standard deviation of Phillips colour	$R_c \sim \text{LogUniform}(-5, 2)$

Table 1: The different priors used for the baseline BAHAMAS model in this investigation. These are equivalent to the those used by March et al (2016) except Ω_κ is used here instead of Ω_Λ. Equation 6 was used in conjunction with the Ω_Λ and Ω_m priors given in March et al (2016) to produce a minimum value of -3 and maximum value of 1 for Ω_κ

The priors used for this model are shown in table 1 and are equivalent to those used by March et al (2016) except for Ω_κ. This is because Ω_κ was found to be easier to use

than Ω_Λ when implementing the model in Stan. However the two are degenerate, as can be seen from equation 6. To calculate the constraints on the prior of Ω_κ, equation 6 was used in conjunction with the priors on Ω_Λ and Ω_m given in March et al (2016) to produce a minimum value of -3 and maximum value of 1 for Ω_κ.

2.5 SNIa Data

The data used in this investigation is the SNIa sample from the Joint Light-Curve Analysis (JLA) dataset which is comprised of 740 spectroscopically confirmed SNIas obtained by the SDSS-II and SNLS collaborations [12]. It includes several low-redshift samples ($z < 0.1$), all three seasons from SDSS-II ($0.05 < z < 0.4$), and three years from SNLS ($0.2 < z < 1$). The JLA data set is available on-line [2] together with the Csyst and Cstat variance-covariance matrices.

3 Implementing the model in Stan

The Bayesian Hierarchical model described in section 2 and depicted in figure 1 was implemented using the probabilistic programming language Stan. With Stan the user specifies a fully Bayesian model and inference follows automatically [13]. In this investigation the PyStan interface was utilised to allow use of the friendlier Python environment [14].

3.1 How Stan Works

To perform inference Stan uses the Hamiltonian Monte Carlo (HMC) sampling method. This takes the unknown parameters, the latent variables at the top of figure 1 in this investigation, and assigns them as the vector position of a fictional particle. It then introduces a random momentum variable for each parameter. The potential of the particle is then assigned to be the negative log of the likelihood function and the kinetic energy to be the classical mechanics energy with mass = 1. A random momentum is then generated at each iteration and Hamiltonian dynamics is simulated for a short time, ϵ, and steps, L, using the Leapfrog method (the Leapfrog method is a numerical integration technique which uses discrete steps) [13]. This directs the Markov Chain to the lowest potential, where the log likelihood is the highest, allowing samples to be drawn with greater efficiency. Then a Metropolis acceptance step is applied, and a decision is made whether to update to the new state variables or keep the existing state. The Metropolis decision making process includes an element of randomness that ensures the Markov Chains do not become stuck at local minima and thereby sample from a wide region of the parameter space.

In practice, a major hurdle encountered when employing HMC methods is that its performance is highly sensitive to the two parameters in the Leapfrog method: the step time ϵ and number of steps L, both of which are usually user-specified. In particular, if

[2]The JLA data set can be downloaded at http://supernovae.in2p3.fr/sdss_snls_jla/ReadMe.html. The specific folder is the jla_likelihood_v6.tgz file. This includes the measured heliocentric redshift, apparent magnitude, colour an stretch values. The Csyst and Cstat variance-covariance matrices are located in the covmat_v6.tgz file

L is too small then the algorithm exhibits undesirable random walk behaviour, while if L is too large the algorithm wastes computation. Therefore, these parameters have to be fine-tuned with computationally costly user-tuned runs to ensure the desired region in parameter space is sampled. Stan avoids this complication and automatically tunes these parameters, without user intervention, with the use of the No-U-Turn Sampler [15]. This works by calculating the optimum values of ϵ and L so that the Leapfrog steps do not trace back on themselves.

The Stan programme code is split into functional blocks which makes the implementation of the BAHAMAS model in figure 1 relatively straightforward. The only complexity is in performing the integral for equation 3. To do this the Simpson's rule for integration was coded into the Functions block of Stan and tested using a step size of 0.0001. This worked well in terms of the accuracy however the computational time was excessive even for a small 30 SNIa sample size. Therefore an alternative method was sought. Stan provides an ODE solver which uses the Runge-Kutta 45 (RK45) method. This is a fourth order ODE integrator and because it has been configured by the Stan Team it is also optimised. However, to use the RK45 function provided by Stan, the redshift have to be in ascending order which caused an issue as the JLA data was not in order. To solve this problem a index list was created which would allow Stan to map an ordered set of the redshifts to the original order in the JLA dataset. Both these ordered lists were in-putted in the data Functional Block. Using the RK45 gave a considerable speed boost, approximately 10×, allowing computation of a much larger sample size.

3.2 Verification

The primary method used to verify the Stan implementation was to compare the results obtained to March et al (2016). The complexity of the baseline BAHAMAS model is such that most, if not all, incorrect implementations would be easily identified through large discrepancies in the results. However to avoid large computational times during this verification stage, the Stan implementation was tested on a subset of 30 SNIa samples. These were selected to ensure coverage of the full range of redshifts in the JLA dataset. The verification run returned sensible values for all the parameters of interest when compared to results from March et al (2016). Subsequently, a programme run using the full 740 Supernovae JLA dataset was attempted. However the computational time required was excessive (over 5 weeks) on a machine with a 2.9 GHz Intel Core i5 processor and 8GB RAM. Hence due to the limited time availability a truncated sample size of 311 Supernovae , spanning the full range of the JLA dataset, was used instead.

4 Results

A total of 4 Markov chains were used with each running for 2000 iterations. The first 500 of these were assigned as 'burn-in' iterations. The burn-in iterations enable the chains to find the general area of interest and determine the optimum values of ϵ and L as mentioned in 3.1. These samples are not included in the final results by default in Stan. The number of Markov chains was chosen to ensure that the results could be checked

sufficiently for convergence and mixing, as described below. The number of iterations were the maximum possible within the time constraints.

Parameter	**Mean**	**Standard Deviation**	N_{eff}	$\hat{R}$
Ω_m	0.29	0.13	1013	1.0
Ω_κ	0.29	0.31	685	1.0
Ω_λ	0.43	0.21	-	-
M_0^ϵ	-19.13	0.03	113	1.01
σ_{res}	0.08	0.005	847	1.0
α	0.14	0.01	2951	1.0
β	2.91	0.13	1897	1.0

Table 2: The marginalized posterior constraints on the cosmological and Phillips correction parameters, assuming $H_0 = 67.3$ km/s/Mpc. A truncated version of the JLA dataset, comprising of 311 spectroscopically confirmed SNIa, was used. The SNIa were selected to span the full range of the JLA dataset. The Stan programme was run for 4 Markov chains with each chain running 2000 iterations. 500 of these were burn-in iterations and were not included in the final results. N_{eff} represents the effective independent sample size and $\hat{R}$ the potential reduction scale statistic. Ω_λ was not directly sampled by Stan and thus there are no N_{eff} and $\hat{R}$ statistics associated with it. However estimates of these can be taken to be the statistics of Ω_κ as a result of the degeneracy given in equation 6.

The results for some of the parameters of interest are shown in figure 2 and table 2. Table 3 compares the results of this investigation to those obtained by March et al (2016).

However, before the results from Stan can be accepted it is important to verify if the Markov chains were producing the expected behaviour, especially with regards to convergence. With MCMC methods, such as the one employed by Stan, each draw depends on the one before which can result in highly autocorrelated posterior samples. Autocorrelation simply gives a measure of independence between different samples, with low autocorrelation representing greater independence. The N_{eff} statistic measures autocorrelation and estimates the number of truly independent samples (i.e., the number of effective samples). From the official Stan manual this value should be much greater than $\frac{1}{1000}$ of the total sample size , N [13]. Thus $\frac{1}{1000} \times 6000 = 6$ is the minimum threshold for this experiment. As can be seen from table 2, which displays the results obtained, the effective sample size is well above this minimum threshold for all the parameters of interest. Another statistic, $\hat{R}$, more commonly known as the potential reduction scale statistic, quantifies the convergence of the Markov Chains. The convergence is assessed by comparing the estimated between-chains and within-chain variances for each param-

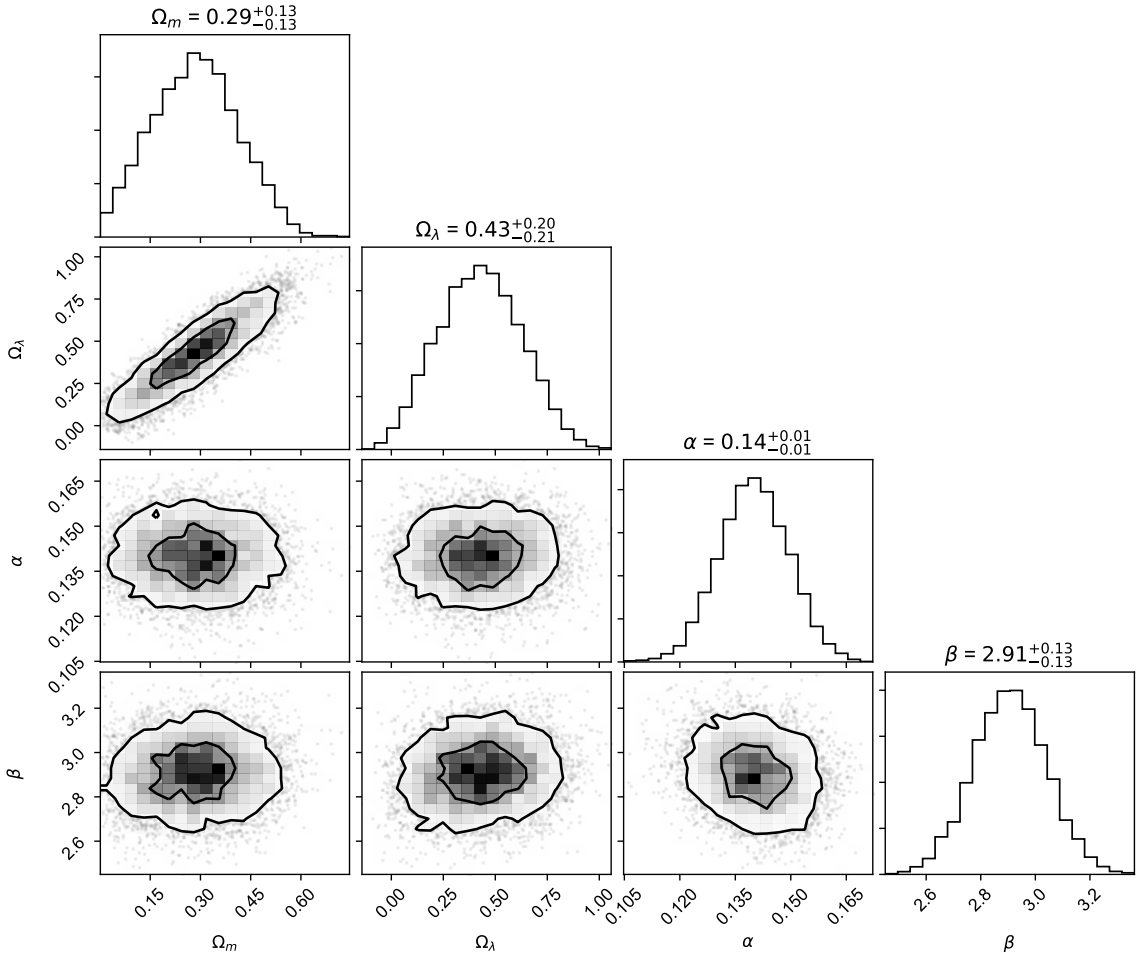

Figure 2: 1D and 2D marginal posterior distributions for the cosmological parameters, and the colour and stretch correction parameters. The two contours represent the 1σ and 2σ values. In most cases, the marginal posterior distributions are symmetric and approximately Gaussian, in which case the reported errors on the histograms are the posterior standard deviation. The exception is the intervals for Ω_Λ which is reported with asymmetric positive and negative error due to the non-Gaussian shape of the posterior distribution.

eter with a $\hat{R}$ value of 1 indicating ideal convergence. Although, in practice a number less than 1.1 is regarded as a good indicator of convergence [13]. From table 2 it is seen that the $\hat{R}$ values are all well below the 1.1 threshold, indicating convergence. Another test is to check how well the chains are mixing, or moving around in parameter space. If the chains take many iterations to move around then it is indicative of poor convergence. This check is performed through visual inspection of the traceplot for each parameter: plots which depict the sampled value of the parameter at each iteration.

Figure 3 displays the traceplots for each parameter of interest. All the plots fluctuate well around a stable value and do not become stuck in small regions indicating that the

Parameter	Experiment	March et al
Ω_m	0.29 ± 0.13	0.340 ± 0.101
Ω_κ	0.29 ± 0.31	0.119 ± 0.249
Ω_λ	$0.43 \pm^{0.20}_{0.21}$	0.542 ± 0.157
M_0^ϵ	-19.13 ± 0.03	-19.140 ± 0.022
σ_{res}	0.08 ± 0.005	0.104 ± 0.005
α	0.14 ± 0.01	0.137 ± 0.006
β	2.91 ± 0.13	3.058 ± 0.085

Table 3: Comparison of the results obtained in this experiment with those from March et al (2016) for the baseline BAHAMAS model with $w = -1$. All the parameters are consistent within one sigma level except for σ_{res}.

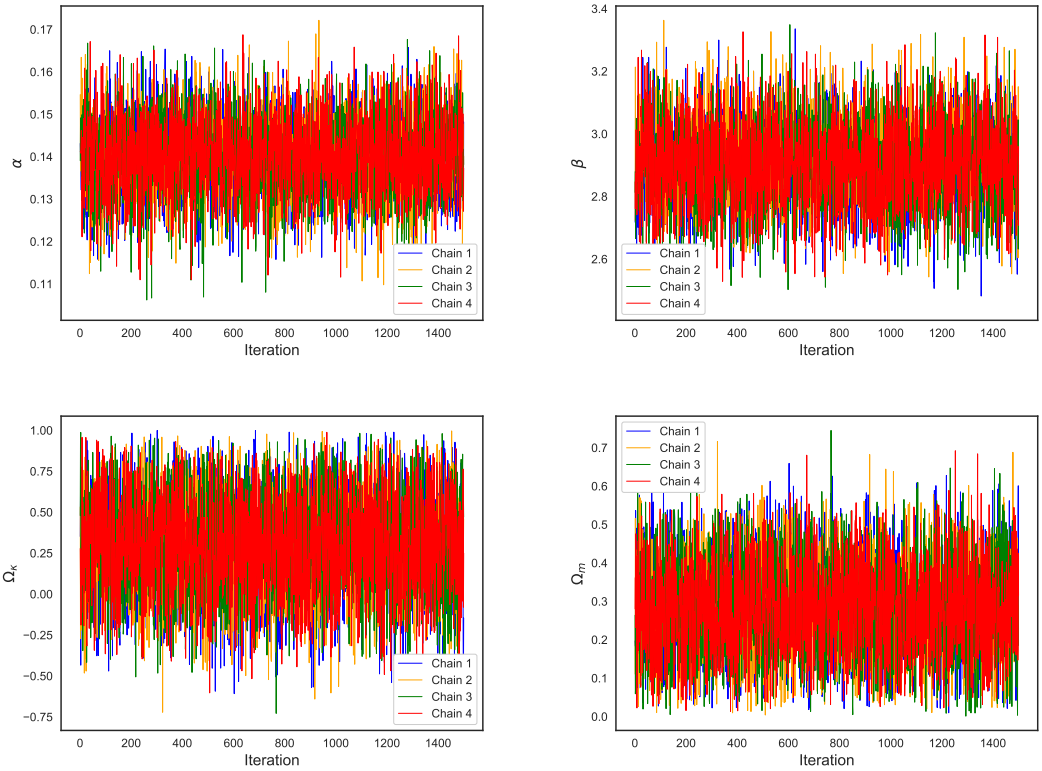

Figure 3: Traceplots for 4 of the parameters of interest. All 4 chains for each of the parameters fluctuate around a stable value, indicating good mixing. The other parameter traceplots, not shown here, indicated similar mixing.

chains are sampling the full range of the parameter space.

Overall these three checks give fairly high confidence of convergence and that the programme is running as it should.

5 Discussion

As can be seen in table 3 the results obtained using Stan are consistent with those obtained by March et al (2016), expect for σ_{res} which is explained below. Hence, it can be concluded that the Stan implementation is correct and can be used as an alternative method to model the baseline BAHAMAS model.

The residual intrinsic dispersion, σ_{res}, is found to be 0.074 ± 0.008. This is the average residual scatter in the (post-correction) intrinsic magnitudes, M_0^ϵ. This value has an approximately equal meaning to the σ_{coh} in Betoule et al (2014) which ranges from 0.08 to 0.12 [12]. The results from this investigation are consistent with this range within 1σ. However, it is not consistent with what March et al (2016) obtained within 1σ. This is most likely due to the truncated data set used, as this could have inadvertently resulted in choosing supernovae with a slightly decreased residual scatter.

The curvature parameter, Ω_κ is found to be 0.29 ± 0.31 which excludes a flat universe ($\Omega_\kappa = 0$) at $\sim 1\sigma$ level.

In terms of Stan's performance, there were no major difficulties in implementing the model. In fact the ease to which a complex Bayesian model, such as the baseline BAHAMAS model, can be implemented is encouraging. However, debugging in Stan can become tedious and complex because of the way the language is complied and executed in C++. Furthermore, with Stan being a relatively new language, the documentation and resources available are limited. This drawback was especially evident when trying to implement the complex RK45 ODE solver in this investigation. Even so, with an active development team and new capabilities being added regularly, Stan is recommended to use with other cosmological models and Bayesian models in general.

To improve this investigation for future work the full 740 sample size should be used and the number of total iterations increased. The Stan development team recommend around $2000 - 3000$ N_{eff} sample size for each parameter of interest. Currently, Stan does not offer a method to automatically stop the sampling when a minimum N_{eff} is reached but a crude estimate can be obtained using table 2 [16]. Noticing that 6000 iterations resulted in the lowest N_{eff} of 113, the total number of iterations would have to be increased almost 20 fold to achieve the minimum recommended. Therefore, it would be necessary to either optimise the code or use high performance computers to ensure sensible run-times.

For further improvement, bias with the data itself needs to be checked and if present, accounted for. In astronomical observations the problem with data gathered through a magnitude limited survey such as this is that the survey is incomplete at fainter magnitudes. Certain faint objects are simply not observed due to the sensitivity limit of the telescope camera and various environmental factors, leaving a truncated dataset. This missing data introduces an effect knows as the Malmquist bias which can result in biases in parameter inference if not properly accounted for [17]. Therefore, the flux limits of

the surveys of the SN1a's need to be checked to see if Malmquist bias could be important. March et al (2018) have shown that the bias with truncated data sets in Bayesian inference models can be avoided by using a suitable inclusion model [18].

6 Conclusion

The BAHAMAS model developed by March et al. was implemented to infer the cosmological and Phillips colour and stretch parameters using the Hamiltonian Monte Carlo sampling method, through a probabilistic programming language called Stan. 311 out of the 740 spectroscopically confirmed supernovae type Ia (SNIa) from the Joint Light-curve Analysis" dataset were analysed. 4 Markov chains were used with 2000 iterations each including 500 burn-in iterations. The results obtained were consistent with those of March et al (2016) which employed Gibbs type samplers. The only discrepancy was the residual scatter of the intrinsic magnitudes of the supernovae, σ_{res}, which was found to be 0.08 ± 0.005 and not consistent with March et al. The reason for the discrepancy was not determined, but was argued to be most likely due to the truncated sample size. The matter density, Ω_m, the curvature parameter, Ω_κ, and the cosmological constant parameter, Ω_Λ, were found to be 0.29 ± 0.13, 0.29 ± 0.31 and 0.43 ± 0.21 respectively. The α and β values which govern the strength of the Phillips stretch and colour corrections were found to be 0.14 ± 0.01 and 2.91 ± 0.13 respectively.

Furthermore Stan's performance was evaluated in terms of both the ease of implementation of the model as well as the computational resources required. The implementation was relatively straightforward, however the C++ compilation and execution culminated in debugging being tedious and complex. Furthermore, running the programme with 740 supernovae was found to be too time consuming on a modern average laptop. However with an active Stan development team and new capabilities being added regularly, Stan is recommended to use with other cosmological models and Bayesian models in general.

References

[1] Hikmatali Shariff, Xiyun Jiao, Roberto Trotta, and David A Van Dyk. Bahamas: New Analyis Of Type Ia Supernovae Reveals Inconsistencies With Standard Cosmology. *The Astrophysical Journal*, 827(1), 2016. doi: 10.3847/0004-637X/827/1/1. URL http://iopscience.iop.org/article/10.3847/0004-637X/827/1/1/pdf.

[2] Adam G. Riess, Alexei V. Filippenko, and Peter Challis et al. Observational Evidence from Supernovae for an Accelerating Universe and a Cosmological Constant. *The Astronomical Journal*, 116(3):1009–1038, 9 1998. ISSN 00046256. doi: 10.1086/300499. URL http://stacks.iop.org/1538-3881/116/i=3/a=1009.

[3] S. Perlmutter, G. Aldering, and G. Goldhaber et al. Measurements of Ω and Λ from 42 HighRedshift Supernovae. *The Astrophysical Journal*, 517(2):565–586, 6 1999. ISSN 0004-637X. doi: 10.1086/307221. URL http://stacks.iop.org/0004-637X/517/i=2/a=565.

[4] NASA. HubbleSite - Dark Energy - Type Ia Supernovae, . URL http://hubblesite.org/hubble_discoveries/dark_energy/de-type_ia_supernovae.php.

[5] NASA. HubbleSite - Dark Energy - Did Einstein Predict Dark Energy?, . URL http://hubblesite.org/hubble_discoveries/dark_energy/de-did_einstein_predict.php.

[6] Silvia Toonen, Gijs Nelemans, and Simon Portegies Zwart. Supernova Type Ia progenitors from merging double white dwarfs: Using a new population synthesis model. *Astrophysics*, 8 2012. doi: 10.1051/0004-6361/201218966. URL http://arxiv.org/abs/1208.6446.

[7] M. C. March, R. Trotta, and P. Berkes et al. Improved constraints on cosmological parameters from SNIa data. *Monthly Notices of the Royal Astronomical Society*, 418 (4):23082329, 2 2011. doi: 10.1111/j.1365-2966.2011.19584.x. URL http://arxiv.org/abs/1102.3237http://dx.doi.org/10.1111/j.1365-2966.2011.19584.x.

[8] M. M. Phillips, P. Lira, and N. B. Suntzeff et al. The Reddening-Free Decline Rate Versus Luminosity Relationship for Type Ia Supernovae. *THE ASTRONOMICAL JOURNAL*, 118:1766–1776, 7 1999. doi: 10.1086/301032. URL http://arxiv.org/abs/astro-ph/9907052.

[9] M. M. Phillips and M. M. The absolute magnitudes of Type IA supernovae. *The Astrophysical Journal*, 413:L105, 8 1993. ISSN 0004-637X. doi: 10.1086/186970. URL http://adsabs.harvard.edu/doi/10.1086/186970.

[10] Adam Riess, William Press, and Robert Kirshner. A Precise Distance Indicator: Type Ia Supernova Multicolor Light Curve Shapes. *The Astrophysical Journal*, 473(88), 4 1996. doi: 10.1086/178129. URL http://arxiv.org/abs/astro-ph/9604143.

[11] P. A. R. Ade, N. Aghanim, and C. Armitage-Caplan et al. Planck 2013 results. XVI. Cosmological parameters. *Astronomy & Astrophysics*, 571:A16, 11 2014. doi: 10.1051/0004-6361/201321591. URL http://www.aanda.org/10.1051/0004-6361/201321591.

[12] M. Betoule, R. Kessler, and J. Guy et al. Improved cosmological constraints from a joint analysis of the SDSS-II and SNLS supernova samples. *Astronomy & Astrophysics*, 568:A22, 8 2014. doi: 10.1051/0004-6361/201423413. URL http://www.aanda.org/10.1051/0004-6361/201423413.

[13] Stan Development Team. Stan - Documentation Version 2.17.0, 2017. URL http://mc-stan.org/users/documentation/.

[14] Stan Development Team. PyStan: the Python interface to Stan, 2014. URL https://pystan.readthedocs.io/en/latest/.

[15] Matthew D. Homan and Andrew Gelman. The No-U-turn sampler: adaptively setting path lengths in Hamiltonian Monte Carlo. *The Journal of Machine Learning Research*, 15(1):1593–1623, 2001. URL https://dl.acm.org/citation.cfm?id=2638586.

[16] Stan Development Team. Setting a minimum effective sample size, 2017. URL http://discourse.mc-stan.org/t/setting-a-minimum-effective-sample-size/1824.

[17] R. Kessler and D. Scolnic. Correcting Type Ia Supernova Distances for Selection Biases and Contamination in Photometrically Identified Samples. *The Astrophysical Journal*, 836(1):56, 2 2017. doi: 10.3847/1538-4357/836/1/56. URL http://stacks.iop.org/0004-637X/836/i=1/a=56?key=crossref.d0bddb5607f63ec4347f7e1b0143aef5.

[18] Marisa Cristina March and Marisa Cristina. A Bayesian approach to truncated data sets: An application to Malmquist bias in Supernova Cosmology. *American Astronomical Society*, 231, 2018. URL http://adsabs.harvard.edu/abs/2018AAS...23115315M.

Declaration of Work

The investigation in this report was produced jointly and equally by Vatsal Patel and Zoltan Kiss.