

AMPERE: A tool to fit heterogeneous observations consistently

P. Scicluna^{1,2,3*}, S. Zeegers^{4,5,3}, J. P. Marshall³, F. Kemper^{6,7,8}, S. Srinivasan⁹,
T. Dharmawardena¹⁰, L. Fanciullo^{11,12,3}, O. Morata⁶, A. Trejo-Cruz^{9,3}

¹Centre for Astrophysics Research, Department of Physics, Astronomy and Mathematics, College Lane Campus, University of Hertfordshire, Hatfield AL10 9AB, UK

²Space Science Institute, 4750 Walnut Street, Suite 205, Boulder, CO 80301, USA

³Institute of Astronomy and Astrophysics, Academia Sinica, 11F of AS/NTU Astronomy-Mathematics Building, No.1, Sec. 4, Roosevelt Rd, Taipei 106319, Taiwan

⁴European Space Agency / ESTEC Keplerlaan 1, 2201 AZ, Noordwijk, The Netherlands

⁵SRON Space Research Organisation Netherlands, Niels Bohrweg 4, 2333 CA Leiden, The Netherlands

⁶Institut de Ciències de l'Espai (ICE, CSIC), Can Magrans, s/n, E-08193 Cerdanyola del Vallès, Barcelona, Spain

⁷ICREA, Pg. Lluís Companys 23, E-08010 Barcelona, Spain

⁸Institut d'Estudis Espacials de Catalunya (IEEC), E-08034 Barcelona, Spain

⁹Instituto de Radioastronomía y Astrofísica, UNAM. Apdo. Postal 72-3 (Xangari), Morelia, Michoacán 58089, Michoacán, Mexico

¹⁰Center for Cosmology and Particle Physics, Department of Physics, New York University, 726 Broadway, New York, NY 10003, USA

¹¹National Chung Hsing University, 145 Xingda Rd., South Dist., Taichung City 402, Taiwan, R.O.C.

¹²Tamkang University, 151 Yingzhuan Rd., Tamsui Dist., New Taipei City 251301, Taiwan, R.O.C.

Accepted XXX. Received YYY; in original form ZZZ

ABSTRACT

As astronomy advances and data becomes more complex, models and inference also become more expensive and complex. In this paper we present AMPERE, which aims to solve this problem using modern inference techniques such as flexible likelihood functions and likelihood-free inference. AMPERE can be used to do Bayesian inference even with very expensive models (hours of CPU time per model) that do not include all the features of the observations (e.g. missing lines, incomplete descriptions of PSFs, etc). We demonstrate the power of AMPERE using a number of simple models, including inferring the posterior mineralogy of circumstellar dust using a Monte Carlo Radiative Transfer model. AMPERE reproduces the input parameters well in all cases, and shows that some past studies have tended to underestimate the uncertainties that should be attached to the parameters. AMPERE can be applied to a wide range of problems, and is particularly well-suited to using expensive models to interpret data.

Key words: methods: statistical – dust, extinction

1 INTRODUCTION

Modern astronomy depends heavily on interpreting data using models. This is particularly relevant as data grows and becomes more complex, as both the volume of data to be modelled and the complexity of the models in use also grow (e.g. SKA, Vera Rubin Observatory). Analyses routinely involve combining diverse observations such as images, spectra, visibilities or datacubes, but even the relatively simple problem of, for example, simultaneously interpreting the global spectral energy distribution (SED) and the mid-infrared spectrum of a source is complicated by the need to interpret the continuum, lines and solid-state features in a single model. As a result, over the last two decades the astronomical community has increasingly coalesced around Bayesian approaches as the standard for analysis (see e.g. Eadie et al. 2023, and references therein for a useful summary, some terminology and recommendations). Bayesian inference is an approach to inference that applies Bayes' Theorem to learn from data by encoding the information in probability distribu-

tions. This has the advantages of properly accounting for uncertainty in data and parameters, illuminating degeneracies between model parameters, giving a principled pathway to include subjective information, and making assumptions explicit (amongst others). However, the need to deal with probability distributions makes all but the simplest problems very computationally expensive. As the availability of computing power grows, so too does the range of datasets and problems that Bayesian inference can be applied to, while more advanced inference algorithms become feasible.

There are also powerful reasons to take the most advanced inference approaches available. As new facilities become more and more expensive, the *cost*¹ of observing time grows, making it ever more important to maximise the return on investment. This places an obligation on scientists to perform the most complete and careful analysis possible of datasets, both to maximise the output of a single observation, and to ensure that future time is not wasted on a blind alley because of incorrect inference. However, since powerful statis-

* E-mail: p.scicluna@herts.ac.uk

¹ where cost could be interpreted in terms of money, time, resources or energy

tical methods have high computational (and hence energy) demands, this must be balanced against the need to use the least resources (particularly fossil fuels) possible to protect the climate (Martin et al. 2022; Knödseder et al. 2022).

This use of Bayesian approaches relies on the availability of a *generative* model – that is, one that is able to explain the observations – to solve the forward problem of inference (see e.g. Hogg et al. 2010). This forward-modelling approach is particularly important for understanding degeneracies between parameters. In astronomy, however, a model containing all the physics that results in the observable may be impractical or impossible, either because there is no complete solution or because it requires too much computing time to reach it. If a model is incomplete in such a way that it fails to produce key parts of the observed data (model misspecification), it may bias the inference. For example, if emission lines are missing from a model of a spectrum, the continuum may then be overestimated to compensate, with knock-on effects to other parameters of interest.

In recent years, there have been several developments which help to alleviate model misspecification. In particular, the development of *flexible likelihood functions* (see e.g. Czekala et al. 2015, for an application to astronomy), which enable self-consistent down-weighting of features that the model is incapable of explaining. This simultaneously enables the best inference possible with incomplete models, while also identifying how to improve future models. This in turn reduces the required computing power without biasing inference and hence wasting effort.

Nevertheless, many problems can only be solved with models that are still numerically intensive, even in an incomplete form. For example, understanding the mineralogy of newly-formed dust in a variety of environments relies on solving the radiative-transfer problem to predict the emission of the dust, particularly in the mid-infrared regime where many dust species have features. The shape and strength of these features are influenced by e.g. the size, shape, stoichiometry, or mineralogy of the dust grains, as well as radiative-transfer effects, making this a highly-degenerate problem (e.g. Srinivasan et al. 2017); as a result, many studies make large numbers of simplifying assumptions to reduce the size of the parameter space, and hence fail to quantify the impact of these degeneracies on their results. Nevertheless, even an incomplete solution to the radiative-transfer problem (e.g. considering stars and dust only, not the coupling with gas, and treating the geometry in a reduced symmetry such as 1D or 2D instead of arbitrary dust distributions) may take up to several minutes even with GPU acceleration (e.g. Juvela 2019). Similarly, computing the cross-sections of a distribution of dust grains from their refractive indices alone takes seconds even in simplified cases such as Mie theory (compact spherical grains Mie 1908). In these cases, further optimisation of the inference is required beyond that which can be achieved with flexible likelihood functions. Some works have opted to accelerate this by using neural networks to train fast surrogate models on a small number of simulations of the exact model (e.g. Ribas et al. 2020) and then deploying this approximate model with Markov Chain Monte Carlo (MCMC) for “exact” inference. However, in fields such as cosmology there is a long history of instead continuing to use the exact but expensive model and adopting approximate inference methods like Approximate Bayesian Computing (e.g. Beaumont 2019; Grazian & Fan 2019, and citations therein). In recent years there has been particular growth in the use of neural estimation methods, which exploit a neural network to directly infer the likelihood, posterior or likelihood-to-evidence ratio from simulations of the observations (see e.g. Cranmer et al. 2020 for a recent review and Hahn & Melchior 2022 for a recent example in astronomy).

In this paper, we present a new software package, AMPERE, to handle the issues identified in the preceding paragraphs. AMPERE is specifically designed to handle automatic inference on astronomical data with misspecified and/or very expensive models. In Sect. 2 we discuss the structure and capabilities of AMPERE, which are then demonstrated on synthetic observables and literature datasets in Sect. 3. Finally, we summarise the manuscript in Sect. 4.

2 AMPERE

AMPERE, available at <https://github.com/ICSM/ampere.git>, is a framework for automatic inference on incomplete astrophysical models. It prioritises comparing physical models to astronomical observations e.g. spectra, but can in principle be adapted to a wide range of problems. By providing a means for users to define their model in terms of a data-generating process and a prior on its parameters, AMPERE aims to make Bayesian inference easy even for complex datasets. In this section, we provide a description of the key approaches underlying AMPERE, along with a description of the code, its capabilities and features.

2.1 Flexible likelihood functions for model misspecification

AMPERE employs a flexible likelihood approach (e.g. Czekala et al. 2015) to handle imperfect forward models by treating structured residuals as arising from correlated noise in the data.

Model misspecification produces structured residuals that are statistically indistinguishable from correlated observational noise (e.g. Fig. 5 in Czekala et al. 2015). Consider the case of a spectrum with fringing (see Fig. 1). The fringes themselves can be considered a physical, quasi-periodic component of the model that has to be introduced, or a source of quasi-periodic correlated noise. However, from the perspective of the residuals produced by a model that does not consider fringing, either of these possibilities are equivalent.

Since these two effects are indistinguishable, we can therefore absorb model deficiencies into the data covariance matrix. Given observations d (which could be a spectrum, SED, or other observable, with length n)² with nominal uncertainties $\tilde{\sigma}$, we assume that the data are drawn from a multivariate Normal distribution such that

$$d \sim \mathcal{N}(m(\vec{\theta}), \mathbf{C}(\vec{\phi})), \quad (1)$$

where $m(\vec{\theta})$ is the model prediction. Note that this prescription naturally separates the physical parameters $\vec{\theta}$, and the parameters of the noise model $\vec{\phi}$. This approach arises from the assumption that the uncorrelated noise is drawn from a multivariate Normal such that $d \sim \mathcal{N}(m(\vec{\theta}), \sigma^2)$ and that either correlated noise or model misspecification must be incorporated. However, it does not distinguish between these two potential causes of structure.

The covariance matrix decomposes as

$$C_{ij}(\vec{\phi}) = \mathcal{K}(\lambda_i, \lambda_j; \vec{\phi}) + \delta_{ij} \sigma_i^2, \quad (2)$$

where δ_{ij} is the Kronecker delta, λ_i is the wavelength of the i -th data point, σ_i are the known observational uncertainties, and $\mathcal{K}$ is a kernel function encoding correlations. Multiple effects may be represented by adding kernels:

$$\mathcal{K}(\lambda_i, \lambda_j; \vec{\phi}) = \sum_{\ell=1}^L \mathcal{K}_{\ell}(\lambda_i, \lambda_j; \vec{\phi}_{\ell}). \quad (3)$$

² This section assumes the data is a vector, but the problem can be trivially extended to higher-order cases e.g. images, multiple vectors, etc.

Comparison of Model Specification and Flexible Likelihood Approach

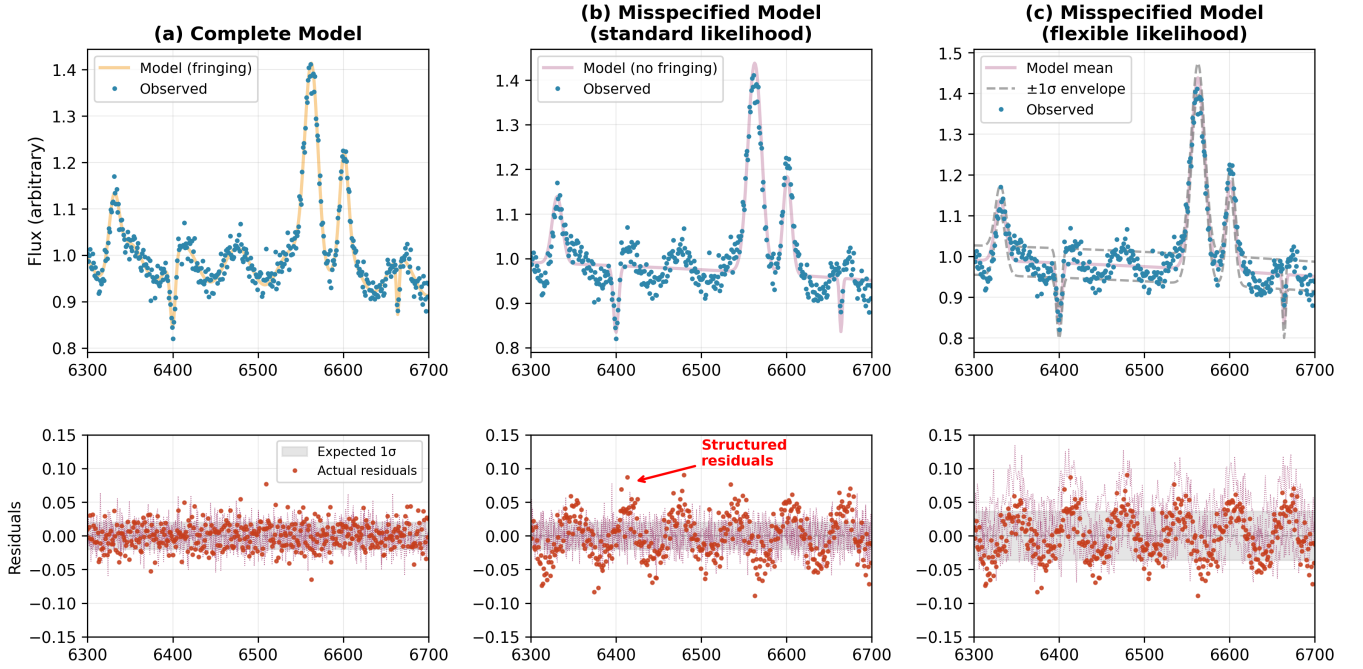


Figure 1. Demonstration of concepts of model misspecification and flexible likelihoods. If we start from a spectrum that contains emission and absorption lines but is contaminated by fringing, we can attempt to interpret it by modelling the lines and the fringing simultaneously (*a, top*). When using this model, the residuals (red points, *a, bottom*) follow the expected i.i.d. Gaussian distribution (three samples from which are shown as pink dotted lines). However, if we take a misspecified model which does not include the fringing (*b, top*), the resulting residuals (*b, bottom*) show quasi-periodic structure, which is inconsistent with the samples from the i.i.d. Gaussians we would normally expect (pink dotted lines), resulting in a dramatic misunderstanding of the uncertainty and potentially the parameters as well. However, when we change the likelihood function (*c, top*), the uncertainty distribution becomes more representative. Indeed, the simplified uncertainty band is overly pessimistic, since it neglects the dependence between points that has been introduced (*c, bottom*). The samples from the correlated distribution in this case show not just similar amplitude to the observed residuals, but also oscillations of similar quasi-period and amplitude. This highlights the equivalence between model misspecification and correlated noise.

This construction ensures $\mathbf{C}$ remains positive (semi-)definite while limiting the dimensionality of $\vec{\phi}$ (e.g., [Rasmussen & Williams 2006](#)).

The resulting log-likelihood is

$$\ln \mathcal{L}(\mathbf{d} | \vec{\theta}, \vec{\phi}) = -\frac{1}{2} \left[n \ln(2\pi) + \ln |\mathbf{C}(\vec{\phi})| + r(\vec{\theta})^T \mathbf{C}(\vec{\phi})^{-1} r(\vec{\theta}) \right], \quad (4)$$

where $r = d - m(\vec{\theta})$ are the residuals. The astute observer will notice that this is equivalent to the log marginal likelihood of a Gaussian process whose mean function is the physical model i.e., $m(\vec{\theta})$, kernel $\mathcal{K}$ and hyperparameters $\vec{\phi}$.

Unconstrained maximization over $\vec{\phi}$ would allow us to trivially inflate $\mathcal{K}$ to allow for any residual structure. It therefore behaves us to impose priors $p(\vec{\phi})$ on the hyperparameters and marginalize:

$$p(\vec{\theta} | \mathbf{d}) \propto p(\vec{\theta}) \int p(\mathbf{d} | \vec{\theta}, \vec{\phi}) p(\vec{\phi}) d\vec{\phi}. \quad (5)$$

By marginalising over these hyperparameters, we ensure that the distribution of correlated noise is accounted for without allowing the noise to grow arbitrarily large. This in turn automatically implies that we should perform Bayesian inference over $\vec{\theta}$ since we are already paying the computational cost involved.

2.2 AMPERE architecture & features

AMPERE is designed to be modular and extensible, and hence employs object-oriented approaches throughout the code. It defines three types of objects: 1) model; 2) data; and 3) inference objects (Figure 2). Through them, AMPERE defines a uniform interface to perform inference on arbitrary (user-defined) models. This allows users to solve their own problems while making it easy to extend and maintain the codebase. The following sections expound further on the implementation and usage of these objects.

2.2.1 Models

Model classes define a physical model (M above) for which the user wishes to perform inference. The model packages the metadata along with a method which forward-models (i.e. simulates) realisations of the model free from observational effects (e.g. noise, PSFs, etc.). Models must also provide methods to compute the prior probability density of a set of model parameters, and a method to transform from samples drawn from the Uniform distribution to those drawn from the prior. The choice of inference method may place constraints on how the prior is defined. Models may be deterministic or stochastic, but care must be taken with stochastic models - not all inference methods are appropriate in these cases. While a number of basic models are

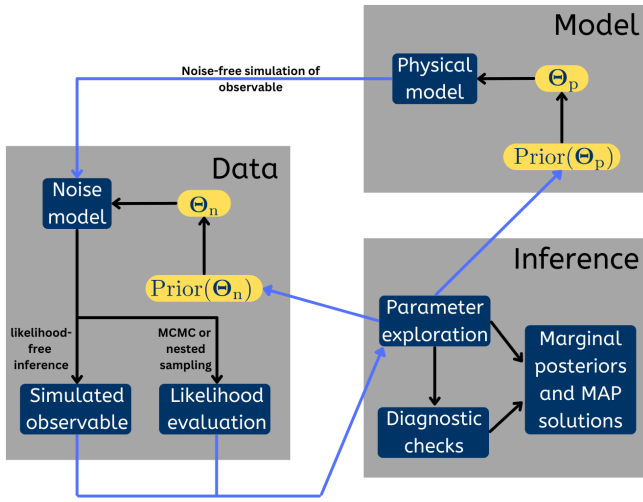


Figure 2. AMPERE architecture. See text for details.

included with AMPERE³, it is generally assumed that users will define their own models suitable to their problem.

2.2.2 Data

Data classes encapsulate the processes required to go from the physical model to the observables, i.e. they contain the noise model, including any additional correlated noise as described above, as well as any transformations that the instrument applies to the data. At present, AMPERE implements classes to handle photometric and spectroscopic observations. Processes implemented in data classes might include convolution with a point-spread function or line-spread function, convolution with a filter profile, resampling in wavelength or angular coordinates, etc. It should also include effects such as zero-point offsets between different observations (e.g. spectral orders). All of these effects should either be precisely known, or treated as nuisance parameters to be integrated out. Finally, data objects evaluate the log-likelihood of the data given an evaluation of the model (for likelihood-based inference), or generate sample (i.e. noisy) observables given a model realisation (for simulation-based, or likelihood-free, inference).

Photometry are assumed to have uncertainties supplied by the user that include any e.g. calibration uncertainty, and that the data points are uncorrelated. However, the user is free to supply a fixed covariance matrix if the instruments are sufficiently well characterised. The only transformation applied is therefore to convolve the model spectrum with filter profiles to obtain appropriate model flux estimates. Synthetic photometry is implemented using *pyphot* (Fouesneau 2022).

Spectra, on the other hand, are treated with a slightly more complex noise model. The uncertainties on the data are assumed to be uncorrelated, and include only the relative uncertainty between points, i.e. no calibration uncertainty. The user is therefore able to supply a prior on the calibration (assumed to be drawn from a log-normal with $\mu = 0$, $\sigma = 0.01$ otherwise), and the observed spectra are multiplied by this factor before comparing to the model. As a result, each order of a spectrum should be treated as a separate observation to optimise their combination and correctly propagate the stitching uncertainty

into the model results. Spectrum objects also implement the flexible likelihood function described above. By default, this features a single Gaussian kernel (also known as a radial basis function or RBF kernel) per spectrum object to account for additional correlated noise. Future improvements will enable a broader range of kernels, including user-specified ones.

Future development will implement data classes to handle a wider range of observations, such as (spectro-)imaging, polarimetry, or interferometric observables. Furthermore, we expect to implement higher-level data containers (e.g. datasets, samples) to facilitate hierarchical inference.

2.2.3 Inference

AMPERE uses inference objects to provide an interface to a wide range of packages for Bayesian inference. Rather than reinventing the wheel and performing inference ourselves, we wrap other packages and ensure that they receive the information they expect. At present, AMPERE supports inference with MCMC using *emcee* (Foreman-Mackey et al. 2013) and *zeus* (Karamanis & Beutler 2020; Karamanis et al. 2021), as well as nested sampling with *dynesty* (Speagle 2020). These methods all rely on directly evaluating the likelihood, and so may give strange behaviour for stochastic models⁴. We have therefore also implemented likelihood-free inference using *sbi* (Simulation-based inference; Tejero-Cantero et al. 2020); *sbi* provides access to a number of different likelihood-free approaches which are often orders of magnitude more efficient for models which are stochastic or have very high computational cost. AMPERE provides an interface to the “(Sequential) Neural Posterior Estimation” ((S)NPE.; Papamakarios & Murray 2016) approach as implemented in *sbi*, which uses a masked autoregressive flow neural network to directly learn the relationship between the parameters of the model (including noise) and observations. In particular, NPE provides access to *amortized inference*, in which an approximate posterior can be learned from a set of simulated observations and then applied to many observations without re-training, providing almost instantaneous inference for any number of sources (Zammit-Mangion et al. 2025). Through these 4 packages, AMPERE provides tools suitable for a wide range of astronomical inference problems. The characteristics of the different approaches are summarised in Table 1.

Inference objects also provide post-processing capabilities. AMPERE automatically computes some statistics commonly used to identify non-convergence in MCMC samplers, such as the effective sample size (ESS), autocorrelation, Rubin-Gelman split- $\hat{R}$ (Gelman & Rubin 1992) and Geweke- z drift score (Geweke 1992). AMPERE will issue warnings when these statistics clearly indicate non-convergence. AMPERE also produces a number of diagnostic plots, including posterior-predictive plots, traces, and corner plots. Some examples of these can be seen in the following sections.

3 PRACTICAL APPLICATIONS OF AMPERE

What kind of thing can AMPERE do? In this section, we demonstrate AMPERE with a number of test models, demonstrating that it recovers

³ including all models used to demonstrate and validate AMPERE in this work

⁴ The stochasticity of the model means that any single likelihood evaluation is only one of a distribution for that set of parameters. It is therefore necessary to estimate a marginal log-likelihood for stochastic models by computing multiple realisations of the model. While this will be implemented in AMPERE in future, it is rather inefficient compared to approaches which are designed for stochastic inference.

Table 1. Comparison of currently-implemented inference approaches in AMPERE

approach	affine-invariant MCMC	slice-sampling MCMC	nested sampling	likelihood-free inference
package	emcee	zeus	dynesty	sbi
max # parameters	few tens	few hundred	few hundred	depends*
correlated parameters	✓	✓	✓	✓
multimodal posteriors	X	✓	✓	✓
requires gradients	X	X	X	X
stochastic models	X	X	X	✓
amortized inference	X	X	X	✓

*The LFI approaches used in AMPERE are insensitive to nuisance parameters - the algorithm simply “sees” them as a source of noise, which therefore do not impact the scaling of the inference.

the input parameters for a small selection of common astronomical inference problems. Code to reproduce all examples is available at <https://github.com/ICSM/ampere/tree/master/examples>.

3.1 A modified blackbody

The modified blackbody (or Planck) function is one of the most widely applied in astronomy. It is particularly useful for inferring the temperature and emissivity index of cold, featureless dust. However, it suffers from a number of issues that require care: the model is not linear in its parameters, making many simplifying assumptions useless, and there is a degeneracy between two of the parameters, the temperature T and the spectral index β . This degeneracy makes it imperative to understand the posterior distributions of the parameters under this model.

As a simple test, we therefore define an arbitrary *true* modified blackbody with parameters $(T, \beta, \log_{10} \frac{M}{M_{\odot}}, d) = (30 \text{ K}, -2, 1, 0.11 \text{ kpc})$ where M is the dust mass and d is the distance to the object, assuming a fixed mass-absorption coefficient at the wavelength $250 \mu\text{m}$ $\kappa_{250} = 10 \text{ cm}^2 \text{ g}^{-1}$. From the true model, we generate synthetic photometry in the AKARI FIS and *Herschel* PACS & SPIRE bands (i.e. ten bands in total), and add Gaussian noise to them assuming an uncertainty of 10% of the true flux value and no correlated noise. We adopt independent priors over each parameter of $P(T) \sim U(10, 50)$, $P(\beta) \sim U(-3, 0)$, $P(\log \frac{M}{M_{\odot}}) \sim U(-3, 3)$ and $P(d) \sim U(0.05, 0.15)$ and retrieve the known input parameters using each inference method implemented in ampere.

As seen in fig. 3, all four approaches recover the input parameters. All models were run for approximately the same amount of wall time on the same machine, and made similar estimates of the median of the posterior 1-D marginal distributions of the parameters. However, *sbi* produced a larger estimate of the credible interval, as expected for an approximate inference method. This typically occurs when there are not enough sample models to accurately describe the distribution of the data. This could be remedied by producing more sample models or by instead using multiple rounds of sampling. However, given the simple nature of this model and data, *sbi* is unnecessary for this case, and is only used here for demonstration.

3.2 Astrophysical parameters of stars

A second key problem in astronomy is to infer the properties of stars. In the age of massive photometric and spectroscopic surveys, doing this *quickly and efficiently* is doubly important. The properties of stars are used to understand the structure, composition and history of galaxies (particularly our own).

To demonstrate this problem, we create a synthetic dataset and

Table 2. Results of fitting synthetic photometric and spectroscopic observations with AMPERE

Parameter	Truth	Phot	Spec
$\log(L)$	0.67	0.68 ± 0.10	0.61 ± 0.02
T_{eff}	6750	6940^{+600}_{-540}	6527^{+49}_{-50}
$\log(g)$	4.37	4.5 ± 0.3	4.1 ± 0.1
A_V	1.0	1.0 ± 0.3	0.81 ± 0.07
R_V	3.2	$3.4^{+0.4}_{-0.3}$	$3.44^{+0.29}_{-0.28}$

attempt to recover the input parameters with AMPERE. The star is generated from PHOENIX model spectra for a given $(L, T_{\text{eff}}, \log g)$ combination, and foreground extinction (R_V, A_V) . The parameter recovery is tested with both a typical set of photometric filters alone (Gaia: G, BP & RP; 2MASS: JHK_s; WISE: W1 & W2), and in combination with a synthetic spectrum similar to a typical *Gaia* RVS spectrum ($842 \text{ nm} \leq \lambda \leq 872 \text{ nm}$; $R \sim 11000$; $\text{SNR} \sim 100$). The model then uses STARFISH (Czekala et al. 2015) to interpolate in the space of model spectra and EXTINCTION (Barbary 2016) to compute foreground extinction. Inference on this model is performed with *sbi* and 10,000 model evaluations and the default masked autoregressive flow posterior for NPE.

The results are shown in Fig. A1 and Table 2. We see that AMPERE is highly effective when working with either photometry or both photometry and spectroscopy. It efficiently recovers the input parameters, albeit with a small bias in the case of spectroscopy. The final parameter uncertainties are comparable to published values from *Gaia* RVS spectra (e.g. Recio-Blanco et al. 2023). Given that it took approximately 20 minutes to generate the models and train this amortised posterior, this shows that the adoption of amortised approaches can be used to dramatically accelerate inference for large photometric and spectroscopic surveys without loss of precision.

3.3 A star with circumstellar dust

Now, we consider a combination of the above two problems - a star orbited by dust whose emission follows a modified blackbody. This is analogous to the detection and characterisation of debris disc host stars, or other cases of stars with cold, featureless circumstellar dust. In this particular case, we attempt to fit observations of HD 105, a young Sun-like star surrounded by a circumstellar debris disc (Marshall et al. 2018). We model the stellar photosphere following the method laid out in Sect. 3.2, and include the dust emission as a modified blackbody. Since the dust component radiates at much longer wavelengths than the star, we add longer wavelength synthetic photometry to the model in the *Herschel*/PACS and SPIRE filters, although we omit the ALMA and ATCA photometry included in Marshall et al. (2018) from the modelling presented here.

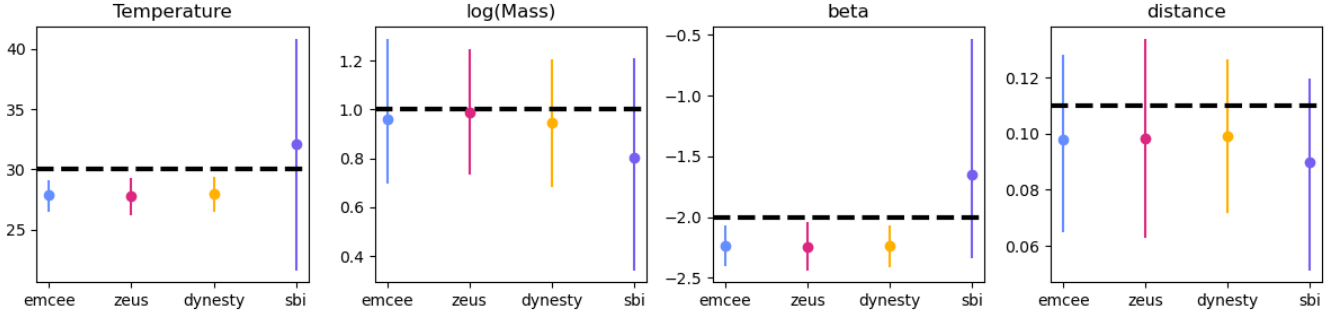


Figure 3. Composite figure showing the values recovered by the different inference approaches for the modified blackbody example. The black dashed line in each panel denotes the 'true' values of the input blackbody function, whilst the coloured dots denote the output for the four inference methods implemented in AMPERE.

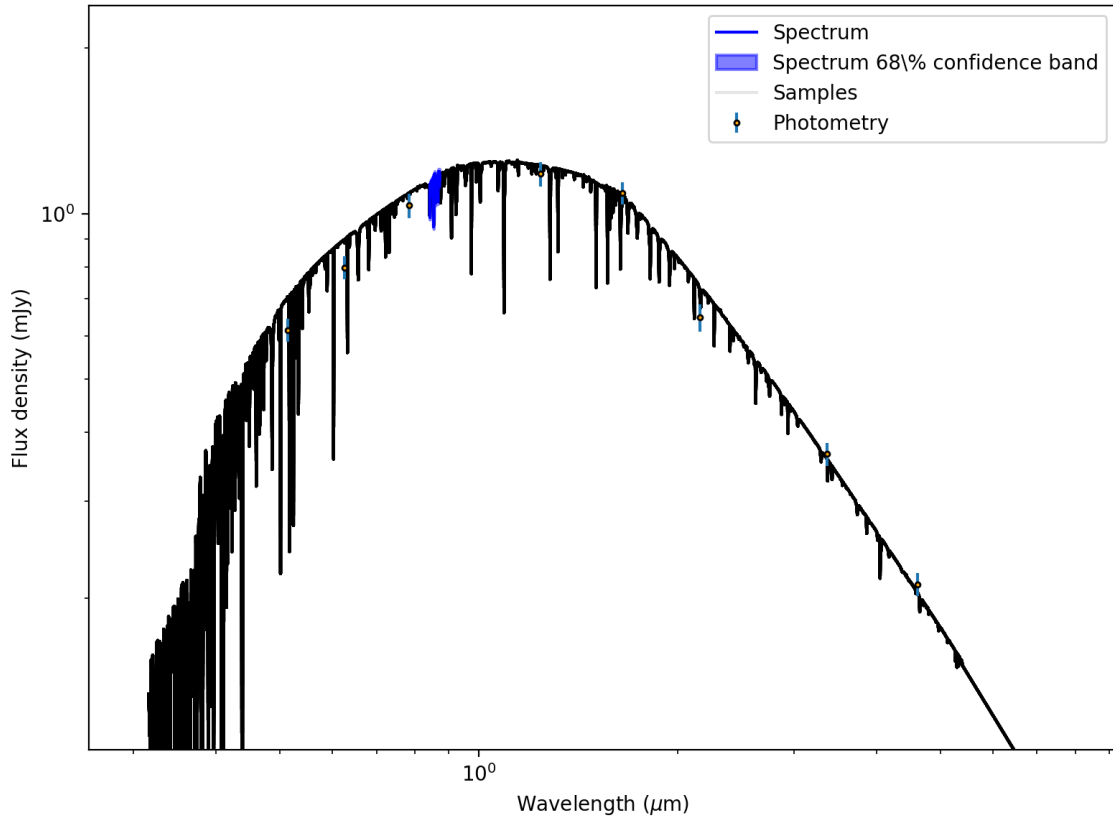


Figure 4. Posterior samples from models fitted to synthetic photometry and a synthetic Gaia spectrum for inferring astrophysical parameters. The parameters are so tightly constrained by this combination that the model samples form a single band.

The parameters of the model are therefore $L_\star$ (in $L_\odot$), $T_\star$ (in K), $\log g$, and metallicity $[\text{Fe}/\text{H}]$ (in dex, relative to Solar) for the star, and the dust temperature, T_{dust} (in K), emitting area A_{dust} (in au^2), break wavelength λ_0 (in μm) where the slope of the dust emissivity changes and spectral index β for the dust component. We adopt uniform priors for all parameters of the model, restricting the ranges to physically meaningful values. For the stellar parameters, we first fixed the distance for the system to that inferred from the Gaia DR3 parallax measurement. The constraint on the distance strongly curtails the set of $L_\star$, $T_\star$ values that can then fit the photometry, reducing the impact of the inherent degeneracy of those parameters. The stel-

lar luminosity was constrained within a range 1.0 to 1.4 $L_\odot$, and the stellar temperature within a range 5,500 to 6,500 K, bracketing the expected values of 1.2 $L_\odot$ and 6034 K. The log stellar surface gravity was given a range of 4.0 to 5.0, and finally the stellar metallicity was constrained between 0.0 and 0.5 $Z_\odot$, again with expected values of 4.478 and 0.02, respectively. For the dust parameters, we restrict the range of T_{dust} between 30 to 100 K, the area scaling factor A_{dust} has an allowed range between 10^{-3} and 10^3 au^2 , the break wavelength λ_0 is restricted to between 50 and 500 μm , and the power law index β value is assumed to lie between 0 and 4. The ranges of these parameters are motivated by either the observational constraints, or

Table 3. Results of star with circumstellar dust.

Parameter	Marshall et al. (2018)	This work
$L_\star$ ($L_\odot$)	1.216 ± 0.005	$1.298^{+0.001}_{-0.001}$
$T_\star$ (K)	6034 ± 8	$5784^{+0.2}_{-0.2}$
$\log g$	4.478 ± 0.006	$4.32^{+0.03}_{-0.04}$
[Fe/H] (dex)	0.02 ± 0.04	$0.07^{+0.02}_{-0.01}$
A_{dust} (au^2)	0.15	$0.16^{+0.21}_{-0.10}$
T_{dust} (K)	50.0	51.3 ± 0.2
λ_0 (μm)	100.0	$99^{+0.1}_{-0.1}$
β	1.0	1.37 ± 0.06

the known properties of debris discs (e.g. Backman & Paresce 1993; Wyatt 2008). Due to its large dynamic range, the scaling factor A_{dust} is explored in log-space to help with convergence.

Fitting was carried out using emcee to explore parameter space using 40 walkers over 10,000 steps to generate 400,000 realisations of the model. A burn-in of 160,000 realisations was discarded before construction of the posterior probability distribution and estimation of parameter values and uncertainties from the remaining 240,000 realisations. Realisations of the star plus disc model were generated by AMPERE using the QuickSED script held under the directory: `./ampere/models/QuickSED.py`. The modelling results are held in the AMPERE repository under the directory `./examples/star_disc/` with a timestamp: 2023-08-18_11-07-25. Total run time for AMPERE to produce this example was ≈ 27 hours using 32 Intel Xeon 2.1 GHz processors, including the downloading and training of the STARFISH grid of model stellar atmospheres.

The results of the fitting process reproduce the observations well. The maximum *a posteriori* parameters for the system are broadly consistent with the values presented in Marshall et al. (2018), where they can be directly compared with their visualisation in Figure 5. The posterior probability distributions for the model run included with the repository is shown in Figure A2, and the numerical results of the fitting are presented in Table 3. The clear exception to the general consistency between these approaches is the β value in the AMPERE fit being significantly steeper than the fit determined in the original paper, but that may be attributed to the absence of millimetre wavelength data points.

3.4 Mineralogy of LMC carbon-star envelopes

Following from above, we expect an effective use-case for AMPERE to be determining mineralogy of dusty environments, and we illustrate this by examination of the modelling of an asymptotic giant branch star. The outflow of the LMC carbon star OGLE LMC LPV 28579 was modelled in detail by Srinivasan et al. (2010), however, they performed the inference by-eye to determine the optimal abundance ratio of amorphous carbon to silicon carbide (SiC) dust. Here we repeat their analysis using AMPERE to perform inference instead. Since this is only a demonstration, we fix all the model parameters except the relative abundances of the dust, the stellar luminosity, and inner and outer radii of the dust shell. There are also two nuisance parameters, the stellar mass and r_0 for the power-law density distribution which are sampled but do not contribute to the output spectrum. For the fixed parameters (e.g. the parameters of the stellar photosphere), we adopt the best-fitting parameters found by Srinivasan et al. (2010). We stress that this choice is simply made to make the model presentation easier, and there is no reason why all parameters could not be inferred simultaneously. We use the Monte Carlo Radiative Transfer code HYPERION (Robitaille 2011) for the dust emission calculations. This was a somewhat arbitrary choice, as any appropriate radiative

Table 4. OGLE LMC LPV 28579 results

Parameter	Srinivasan et al. (2010)	This work
$L_\star$ [$L_\odot$]	4810	4909^{+240}_{-244}
$\log \frac{R_{\text{in}}}{R_\star}$	0.64	0.34 ± 1.15
$\log \frac{R_{\text{out}}}{R_{\text{in}}}$	3	2.82 ± 0.35
$\log \frac{M_d}{M_\odot}$	$-5.85^\dagger$	-6.02 ± 0.21
$\log \frac{M_d}{M_\odot \text{yr}^{-1}}$	$-8.60^{+0.06}_{-0.02}$	$-8.60^\dagger$
C fraction	$0.88^{+0.02}_{-0.04}$	0.949 ± 0.056
SiC fraction	$0.12^{+0.04}_{-0.02}$	$0.045^{+0.043}_{-0.042}$

† Converted from the point estimate using $\dot{M}_d = \frac{M_d}{R_{\text{out}} R_{\text{in}}} v_{\text{exp}}$, where

we assume $v_{\text{exp}} = 10 \text{ km s}^{-1}$.

transfer code that can be called in a scripted fashion would be appropriate using the modular nature of AMPERE. Since this model relies on calculating dust radiative-transfer, we use NPE to keep the inference time reasonable; as this is necessarily a stochastic model NPE is better able to handle the inference than AMPERE's other options.

We use the same photometry and spectra presented in Srinivasan et al. (2010), and perform NPE using a masked autoregressive flow density estimator. We use 10 000 model evaluations to train the posterior. The priors on all parameters are uniform, except for the relative abundances which are drawn from a Dirichlet distribution with concentration parameter $\alpha = 1$. The results of our fit are shown in Table 4, alongside the results from Srinivasan et al. (2010); the 1D and 2D marginal posterior distributions are shown in Fig. A3. We see that our results are in general consistent with those of Srinivasan et al. (2010), however our uncertainties suggest that the approach of Srinivasan et al. (2010) tended to overestimate the precision of their results by a factor of a few. One discrepancy is seen in the relative abundance of SiC and amorphous carbon. This is probably caused by the use of Rouleau & Martin (1991) optical constants here, which produce dust grains a few times less emissive than the Zubko et al. (1996) optical constants used in Srinivasan et al. (2010); hence, a larger fraction of amorphous carbon is required to maintain the same feature-to-continuum ratio. We also see a strong correlation between R_{out} and M_d ; this is to be expected, given that increasing the envelope size increases the mass of the shell, but it bears emphasising that *the uncertainty on the outer radius is the main driver of uncertainty in the dust mass*.

Generating these results with sbi takes only a matter of days (most of which is spent generating the radiative-transfer models), while experiments with MCMC took many weeks merely for burn-in. If this were used for amortised inference, the investment in generating models and training could be used to generate posterior samples for huge samples of AGB stars in a fraction of the time required for MCMC to sample for even a single source.

3.5 The detection of carbonates in NGC 6302

In 2002, Kemper et al. discovered carbonates outside the Solar System in planetary nebula NGC 6302. Carbonates are normally associated with aqueous alteration (weathering) of silicates (rocks), which requires the presence of liquid water on a parent body with an atmosphere. Due to the partial pressure from the atmosphere, gaseous CO_2 is dissolved in the water as CO_3^{2-} , which can then react with the rocky material to form carbonates such as calcite (CaCO_3) and dolomite ($\text{CaMg}(\text{CO}_3)_2$) (the species detected in NGC 6302). The

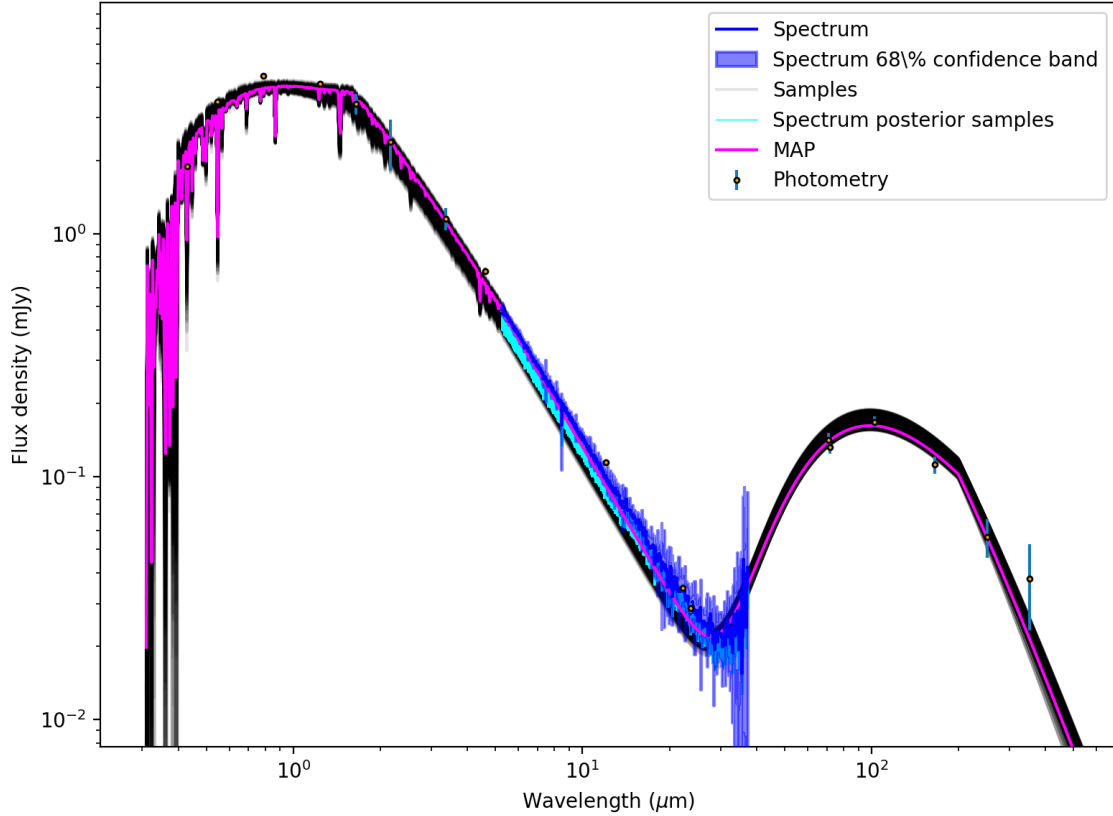


Figure 5. Plot of wavelength vs flux density for the star with circumstellar dust example. Photometric data points are denoted by the circular data points with $1-\sigma$ uncertainties, the *Spitzer*/IRS spectrum is denoted by the blue line and its $1-\sigma$ uncertainty as the shaded region. Realisations of the stellar photosphere and dust modified blackbody model are shown by the black lines, with the maximum amplitude probability model denoted by the magenta line.

amount of carbonates detected in NGC 6302 is significant, and cannot be explained by the destruction of carbonate-containing planets to micron-sized grains. This leads [Kemper et al. \(2002a\)](#) to conclude that an alternate formation mechanism for carbonates must exist that is not based on the weathering of rock in the presence of liquid water.

In an accompanying paper, [Kemper et al. \(2002b\)](#) present the data and describe the analysis that was done to come to this conclusion. The best fit is found through trial and error by varying parameters and visual comparison with the target (observed) spectrum, the so-called *chi-by-eye* method. Here, we repeat this analysis, using AMPERE to sample from the posterior distribution of the model parameters. [Kemper et al. \(2002b\)](#) assume that the circumstellar environment is composed of a warm (*w*) and a cold (*c*) component, both in the form of spherically symmetric shells, and both contributing to the spectrum obtained with the Infrared Space Observatory (ISO):

$$\mathcal{F}_{v,w,i} = \int_{T_{out,w}}^{T_{in,w}} \frac{1}{D^2} \frac{4\pi a^2 r_{0,w}^3 n_{0,i}}{3-p} \left(\frac{T(r)}{T_{out,w}} \right)^{-\frac{3-p}{q}} Q_{v,w,i} B_v(T) dT \quad (6)$$

$$\mathcal{F}_{v,c,i} = \int_{T_{out,c}}^{T_{in,c}} \frac{1}{D^2} \frac{4\pi a^2 r_{0,c}^3 n_{0,i}}{3-p} \left(\frac{T(r)}{T_{out,c}} \right)^{-\frac{3-p}{q}} Q_{v,c,i} B_v(T) dT \quad (7)$$

$$\mathcal{F}_v = \sum_i^{n_w} \mathcal{F}_{v,w,i} + \sum_i^{n_c} \mathcal{F}_{v,c,i} \quad (8)$$

with $Q_{v,w,i}$ and $Q_{v,c,i}$ are the opacities of each species in the warm and cold components, respectively, where i corresponds to the index of the species in the warm and cold component. The warm and cold components contain different subsets of the species, hence we separately sum over them in equation 8. Further, D is the distance to the object, a is the grain size, and r_0 is the inner radius of the temperature component. p is the index of the power law of the density distribution with radius in each shell, and q is the index of the power law of the temperature distribution with radius in each shell. The value $n_{0,i}$ is the density for each species at the inner radius of the shell, where the cold and warm shells are considered separately, and have different $n_{0,i}$ for the same species. Every dust component has its own value for $n_{0,i}$, and the ratios between these values yield the relative abundances.

We have translated the original IDL code to calculate the specific flux into python. This code is included in `./examples/NGC6302.py`. We use the same opacity files as [Kemper et al. \(2002a,b\)](#), which are stored in `./examples/NGC6302/`.

We resolved the parameters in this model using AMPERE `./examples/NGC6302.py` with the logarithm of the abundances (n_0), rather than the values themselves, as the values are very small but cannot be negative, and numerical errors are likely to occur otherwise. We used flat priors, with limits from -6 to 0 for the logarithm of the

abundances, 10 and 80 K for the cold temperature component, and 80 and 180 K for the warm temperature component. The values for the indices p and q were fixed at $1/2$. The distance was set to 910 pc and the grain size to $0.1 \mu\text{m}$, following [Kemper et al. \(2002b\)](#). For the initial guess, the fit values from 2002 were used. The cold model component consists of amorphous olivine, forsterite, clino-enstatite, water ice, diopside, calcite and dolomite; while the warm component consists of amorphous olivine, metallic iron, forsterite, and clino-enstatite. [Kemper et al. \(2002b\)](#) did not detect water ice, diopside, calcite and dolomite and placed upper limits on the abundance of these species in the warm component.

We ran AMPERE with 50 walkers, and used 50000 samples. The burn-in value was set to 40000. Runs with smaller numbers of samples did not give satisfactory results because the abundances introduce strong correlations between parameters; this results in very long autocorrelation times for the chains. In order to achieve reasonable convergence, the parameter `calUnc` was set to 10^{-10} , and `scale-lengthPrior` to 0.1 when defining the observed spectrum using the class `Spectrum`. The fit was performed over the 12 to $120 \mu\text{m}$ wavelength range, while the observed spectrum was resampled using linear spacing of $0.1 \mu\text{m}$ up to $35 \mu\text{m}$; and a spacing of 0.0085 in $1/\lambda$ units at longer wavelengths. This sampling prevents fringing in the longer wavelength segments of the laboratory data from dominating the fit to the observed spectrum.

The range of solutions shows a good resemblance to the observed spectrum (Fig. 6; the corresponding output files are included in the file repository, and are labelled with a timestamp: 2022-12-19_10-33-51). The temperature ranges for the two temperature components are resolved as 57 ± 2 to 36 ± 2 K and 123 ± 3 to 105 ± 3 K. In comparison, [Kemper et al. \(2002a,b\)](#) used 60 to 30 K and 118 to 100 K, and gave these numbers without error bars. Given that the priors allowed for a much broader range, the temperatures that were converged on with AMPERE are remarkably close to, although not quite consistent with, the temperature ranges used by [Kemper et al. \(2002b\)](#). The dust masses have been derived from n_0 using the script `./examples/NGC6302-calculate-dust-mass.py`. When comparing the resulting dust composition (Tab. 5) with the composition derived by [Kemper et al. \(2002b\)](#), we see that the overall dust masses in the cold component are quite similar between both studies ($0.048^{+0.020}_{-0.010} M_{\odot}$ for the current work, versus $0.050 M_{\odot}$ for [Kemper et al. \(2002b\)](#)), the vast majority of which consists of amorphous olivine. We find we need a factor of four less cold forsterite in the present work, and other species in the cold component are even less abundant, compared to what [Kemper et al. \(2002b\)](#) find. In particular, the mass contained in the species giving rise to the far-infrared part spectral signature – calcite, dolomite, diopside and water ice – is a factor of about ten lower in the present work (20 for water ice), but the features are still well reproduced in the solution. [Kemper et al. \(2002b\)](#) cite an uncertainty of 50% on the relative abundances, and a factor of 3 on the absolute dust masses, meaning that except for the overall dust mass and the relative abundances of amorphous and crystalline olivine, the results are not consistent with the present work. We thus conclude that while AMPERE reproduces both the overall shape and the detailed characteristics of the spectrum very well, the determination of trace abundances deviates from the chi-by-eye approach. We consider the dust masses derived in the present work to be more robust, and they are also more comfortably reconciled with the available Calcium budget, compared to the work by [Kemper et al. \(2002a,b\)](#).

4 SUMMARY

In this paper, we have demonstrated AMPERE, a new package to interpret heterogeneous datasets. We extend existing approaches to handle model misspecification based on flexible likelihood functions. AMPERE implements general noise models which exploit kernel functions to produce a wide variety of correlated noise, making it straightforward to apply to a range of problems of interest to the astronomical community. By combining these approaches with cutting-edge inference techniques, we demonstrate that even the most expensive astrophysical models, such as dust radiative transfer, can be used for Bayesian inference on large datasets.

We have demonstrated this with a series of examples representing common astronomical problems. These problems include fitting a modified blackbody to photometry, determining stellar parameters from photometric and spectroscopic data, and calculating infrared excesses for circumstellar dust models.

In each case, AMPERE recovers values and uncertainties for the physical parameters of the model that are consistent with the ‘truth’ provided by previous fitting using standard methods from the literature. The comparison of AMPERE with these previous analyses demonstrates its utility in obtaining robust estimates of parameters from model fitting. In particular, AMPERE illuminated the underlying or inherent biases in the previously adopted model fitting procedures that used limited data sets. Further, AMPERE revealed that uncertainties were generally underestimated by previous fitting procedures and in some cases substantially so (e.g. Sects 3.3 & 3.4). This last point is essential to the interpretation of large data sets particularly with the deployment of statistical techniques to infer the robustness of signal detection(s) amongst large samples.

AMPERE is suitable for analyses on large samples, for example interpreting the large volume of spectra produced by missions such as JWST, SPHEREx, SDSS-V, and many more. We are particularly interested in e.g., inferring the properties of evolved stars and their dust, to understand how they contribute to the chemical evolution of galaxies. However, AMPERE is much more widely applicable, and could for example be applied to problems as diverse as extracting the atmospheric parameters of exoplanets from transit spectroscopy, modelling high-redshift galaxy spectra to infer stellar populations or AGN fractions, or modelling emission-line spectra of ionised nebulae to infer abundances, temperatures and densities.

Future development will extend the existing framework with data classes appropriate for other widely used astrophysical data sets, such as (spectro-)imaging and interferometry. We also plan to integrate the popular JAX framework to accelerate AMPERE, exploiting just-in-time compilation, GPU acceleration, and automatic differentiation. The modular nature of the Model classes and interaction with a wide variety of Inference tools through a framework of common interfaces enables AMPERE to be flexible and extensible to the needs of users and future software developments. We look forward to working with the wider astronomical community to realise the potential of AMPERE, and hope that others will join us as users, collaborators and fellow developers.

ACKNOWLEDGEMENTS

All the authors previously worked together at Academia Sinica Institute of Astronomy and Astrophysics (ASIAA). We would like to thank ASIAA for bringing us together and supporting the early stages of AMPERE’s development.

JPM, FK and LP acknowledge support by the Ministry of Sci-

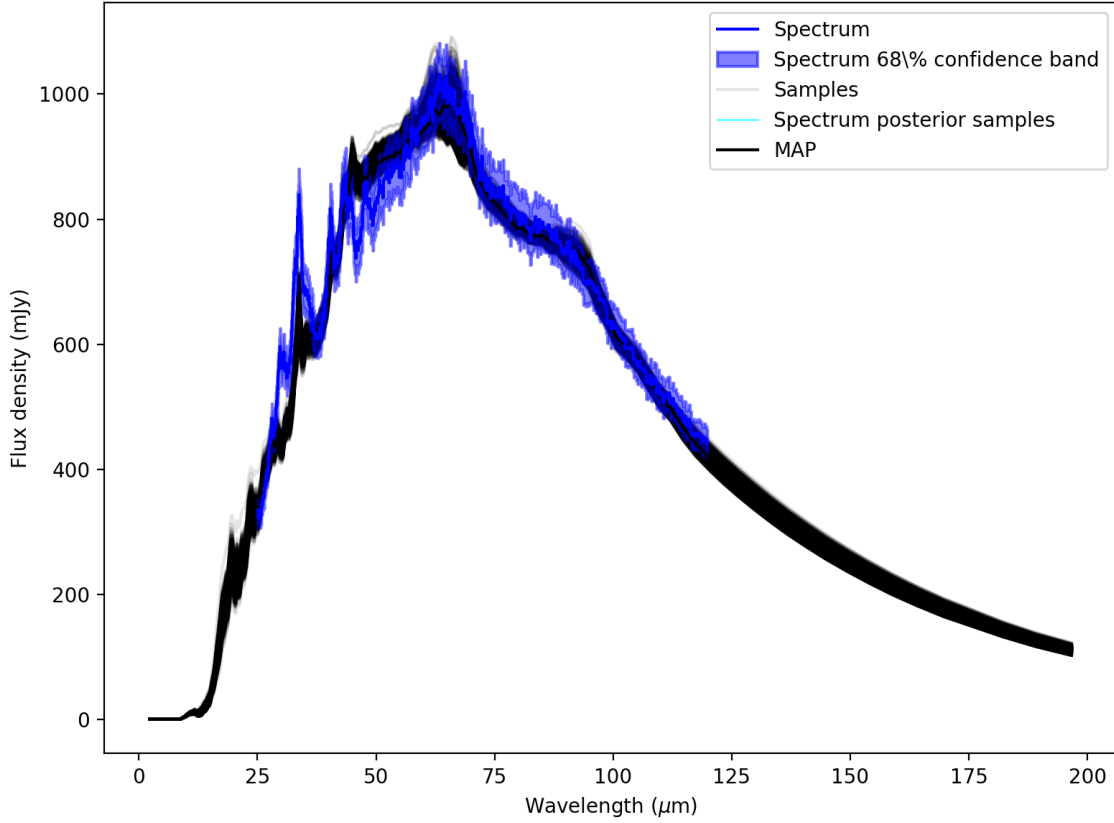


Figure 6. Output plot from AMPERE for the NGC 6302 model. The blue line shows the observed spectrum plotted over the spectral range used to evaluate the fit. The blue shaded area shows the 68% confidence band. The collection of black lines shows the various model outcomes.

Table 5. Dust masses in NGC 6302

species	present work			Kemper et al. (2002b)		
	$M_{d,cold}/M_{\odot}$	mass fraction (cold)	$M_{d,warm}/M_{\odot}$	$M_{d,cold}/M_{\odot}$	mass fraction (cold)	$M_{d,warm}/M_{\odot}$
amorphous olivine	$4.8^{+2.0}_{-1.0} \cdot 10^{-2}$	98.7%	$6.8^{+5.2}_{-3.1} \cdot 10^{-6}$	$4.7 \cdot 10^{-2}$	94 %	$6.1 \cdot 10^{-6}$
metallic iron			$1.5^{+1.1}_{-0.5} \cdot 10^{-5}$			$1.2 \cdot 10^{-4}$
forsterite	$5.7^{+4.0}_{-1.6} \cdot 10^{-4}$	1.2%	$8.7^{+7.3}_{-1.2} \cdot 10^{-8}$	$> 2.0 \cdot 10^{-3}$	> 4.0 %	$3.7 \cdot 10^{-7}$
clino-enstatite	$3.9^{+5.0}_{-2.3} \cdot 10^{-6}$	0.008%	$8.8^{+6.4}_{-3.8} \cdot 10^{-8}$	$> 5.5 \cdot 10^{-4}$	> 1.1 %	$3.1 \cdot 10^{-7}$
water ice	$1.5^{+1.7}_{-0.9} \cdot 10^{-5}$	0.03%		$3.6 \cdot 10^{-4}$	0.72 %	$< 1.5 \cdot 10^{-8}$
diopside	$3.0^{+1.9}_{-1.5} \cdot 10^{-5}$	0.06%		$2.8 \cdot 10^{-4}$	0.56 %	$< 1.2 \cdot 10^{-7}$
calcite	$1.5^{+0.8}_{-0.5} \cdot 10^{-5}$	0.03%		$1.3 \cdot 10^{-4}$	0.26 %	$< 1.0 \cdot 10^{-7}$
dolomite	$5.5^{+6.1}_{-3.2} \cdot 10^{-6}$	0.01%		$7.9 \cdot 10^{-5}$	0.16 %	$< 3.0 \cdot 10^{-8}$

ence and Technology of Taiwan under grant MOST 107-2119-M-001-031-MY3, and Academia Sinica under grant AS-IA-106-M03. JPM acknowledges research support by the National Science and Technology Council of Taiwan under grants NSTC 109-2112-M-001-036-MY3 and NSTC 112-2112-M-001-032-MY3. SS acknowledges support from UNAM-PAPIIT Programs IA 104822 and IA 104824. This work was also partly supported by the Spanish program Unidad de Excelencia María de Maeztu CEX2020-001058-M, financed by MCIN/AEI/10.13039/501100011033. FI-B-00617. S.Z. acknowledges support from the European Space Agency (ESA) as an ESA Research Fellow. L.F. acknowledges support from the Ministry of Science and Technology / National Science and Technology

Council of Taiwan under grants no. 111-2811-M-005-007, 109-2112-M-005-003-MY3, and 114-2811-M-032-005.

This work has made use of the Python programming language (van Rossum 1995), and the packages ASTROPY (Astropy Collaboration et al. 2013, 2018), CORNER (Foreman-Mackey 2016), DYNesty (Speagle 2020), EMCEE (Foreman-Mackey et al. 2013), EXTINCTION (Barbary 2016), HYPERION (Robitaille 2011), MATPLOTLIB (Hunter 2007), MIEPYTHON, NUMPY (Harris et al. 2020), PYPHOT (Fouesneau 2022), SBI (Tejero-Cantero et al. 2020), SciPy (Virtanen et al. 2020), SPECTRES (Carnall 2017), STARSFISH (Czekala et al. 2015), and ZEUS (Karamanis et al. 2021).

DATA AVAILABILITY

All data in this work is either in the public domain, or can be generated with scripts available at <https://github.com/ICSM/ampere/tree/master/examples>.

REFERENCES

- Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, *A&A*, 558, A33, doi: [10.1051/0004-6361/201322068](https://doi.org/10.1051/0004-6361/201322068)
- Astropy Collaboration, Price-Whelan, A. M., Sipőcz, B. M., et al. 2018, *AJ*, 156, 123, doi: [10.3847/1538-3881/aabc4f](https://doi.org/10.3847/1538-3881/aabc4f)
- Backman, D. E., & Paresce, F. 1993, in *Protostars and Planets III*, ed. E. H. Levy & J. I. Lunine, 1253
- Barbary, K. 2016, extinction v0.3.0, Zenodo, doi: [10.5281/zenodo.804967](https://doi.org/10.5281/zenodo.804967)
- Beaumont, M. A. 2019, *Annual Review of Statistics and Its Application*, 6, 379, doi: [10.1146/annurev-statistics-030718-105212](https://doi.org/10.1146/annurev-statistics-030718-105212)
- Carnall, A. C. 2017, arXiv e-prints, arXiv:1705.05165. <https://arxiv.org/abs/1705.05165>
- Cranmer, K., Brehmer, J., & Louppe, G. 2020, *Proceedings of the National Academy of Science*, 117, 30055, doi: [10.1073/pnas.1912789117](https://doi.org/10.1073/pnas.1912789117)
- Czekala, I., Andrews, S. M., Mandel, K. S., Hogg, D. W., & Green, G. M. 2015, *ApJ*, 812, 128, doi: [10.1088/0004-637X/812/2/128](https://doi.org/10.1088/0004-637X/812/2/128)
- Eadie, G. M., Speagle, J. S., Cisewski-Kehe, J., et al. 2023, arXiv e-prints, arXiv:2302.04703, doi: [10.48550/arXiv.2302.04703](https://doi.org/10.48550/arXiv.2302.04703)
- Foreman-Mackey, D. 2016, *The Journal of Open Source Software*, 24, doi: [10.21105/joss.00024](https://doi.org/10.21105/joss.00024)
- Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, *PASP*, 125, 306, doi: [10.1086/670067](https://doi.org/10.1086/670067)
- Fouesneau, M. 2022, pyphot. <https://github.com/mfouesneau/pyphot>
- Gelman, A., & Rubin, D. B. 1992, *Statistical Science*, 7, 457, doi: [10.1214/ss/1177011136](https://doi.org/10.1214/ss/1177011136)
- Geweke, J. 1992, in *Bayesian Statistics 4: Proceedings of the Fourth Valencia International Meeting, Dedicated to the memory of Morris H. DeGroot, 1931–1989* (Oxford University Press), doi: [10.1093/oso/9780198522669.003.0010](https://doi.org/10.1093/oso/9780198522669.003.0010)
- Grazian, C., & Fan, Y. 2019, arXiv e-prints, arXiv:1909.02736, doi: [10.48550/arXiv.1909.02736](https://doi.org/10.48550/arXiv.1909.02736)
- Hahn, C., & Melchior, P. 2022, *ApJ*, 938, 11, doi: [10.3847/1538-4357/ac7b84](https://doi.org/10.3847/1538-4357/ac7b84)
- Harris, C. R., Millman, K. J., van der Walt, S. J., et al. 2020, *Nature*, 585, 357, doi: [10.1038/s41586-020-2649-2](https://doi.org/10.1038/s41586-020-2649-2)
- Hogg, D. W., Bovy, J., & Lang, D. 2010, arXiv e-prints, arXiv:1008.4686, doi: [10.48550/arXiv.1008.4686](https://doi.org/10.48550/arXiv.1008.4686)
- Hunter, J. D. 2007, *Computing in Science and Engineering*, 9, 90, doi: [10.1109/MCSE.2007.55](https://doi.org/10.1109/MCSE.2007.55)
- Juvela, M. 2019, *A&A*, 622, A79, doi: [10.1051/0004-6361/201834354](https://doi.org/10.1051/0004-6361/201834354)
- Karamanis, M., & Beutler, F. 2020, arXiv preprint arXiv: 2002.06212
- Karamanis, M., Beutler, F., & Peacock, J. A. 2021, arXiv preprint arXiv:2105.03468
- Kemper, F., Jäger, C., Waters, L. B. F. M., et al. 2002a, *Nature*, 415, 295, doi: [10.1038/415295a](https://doi.org/10.1038/415295a)
- Kemper, F., Molster, F. J., Jäger, C., & Waters, L. B. F. M. 2002b, *A&A*, 394, 679, doi: [10.1051/0004-6361:20021119](https://doi.org/10.1051/0004-6361:20021119)
- Knödseder, J., Brau-Nogué, S., Coriat, M., et al. 2022, *Nature Astronomy*, 6, 503, doi: [10.1038/s41550-022-01612-3](https://doi.org/10.1038/s41550-022-01612-3)
- Marshall, J. P., Milli, J., Choquet, É., et al. 2018, *ApJ*, 869, 10, doi: [10.3847/1538-4357/aaec6a](https://doi.org/10.3847/1538-4357/aaec6a)
- Martin, P., Brau-Nogué, S., Coriat, M., et al. 2022, *Nature Astronomy*, 6, 1219, doi: [10.1038/s41550-022-01771-3](https://doi.org/10.1038/s41550-022-01771-3)
- Mie, G. 1908, *Annalen der Physik*, 330, 377, doi: <https://doi.org/10.1002/andp.19083300302>
- Papamakarios, G., & Murray, I. 2016, in *Advances in Neural Information Processing Systems*, ed. D. Lee, M. Sugiyama, U. Luxburg, I. Guyon, & R. Garnett, Vol. 29 (Curran Associates, Inc.). https://proceedings.neurips.cc/paper_files/paper/2016/file/6aca97005c68f1206823815f66102863-Paper.pdf
- Rasmussen, C. E., & Williams, C. K. I. 2006, *Gaussian Processes for Machine Learning* (MIT Press)
- Recio-Blanco, A., de Laverny, P., Palicio, P. A., et al. 2023, *A&A*, 674, A29, doi: [10.1051/0004-6361/202243750](https://doi.org/10.1051/0004-6361/202243750)
- Ribas, Á., Espaillat, C. C., Macías, E., & Sarro, L. M. 2020, *A&A*, 642, A171, doi: [10.1051/0004-6361/202038352](https://doi.org/10.1051/0004-6361/202038352)
- Robitaille, T. P. 2011, *A&A*, 536, A79, doi: [10.1051/0004-6361/201117150](https://doi.org/10.1051/0004-6361/201117150)
- Rouleau, F., & Martin, P. G. 1991, *ApJ*, 377, 526, doi: [10.1086/170382](https://doi.org/10.1086/170382)
- Speagle, J. S. 2020, *MNRAS*, 493, 3132, doi: [10.1093/mnras/staa278](https://doi.org/10.1093/mnras/staa278)
- Srinivasan, S., Sargent, B. A., Matsuura, M., et al. 2010, *A&A*, 524, A49, doi: [10.1051/0004-6361/201014991](https://doi.org/10.1051/0004-6361/201014991)
- Srinivasan, S., Kemper, F., Zhou, Y., et al. 2017, *Planet. Space Sci.*, 149, 56, doi: [10.1016/j.pss.2017.08.012](https://doi.org/10.1016/j.pss.2017.08.012)
- Tejero-Cantero, A., Boelts, J., Deistler, M., et al. 2020, *Journal of Open Source Software*, 5, 2505, doi: [10.21105/joss.02505](https://doi.org/10.21105/joss.02505)
- van Rossum, G. 1995, *Python tutorial*, Tech. Rep. CS-R9526, Centrum voor Wiskunde en Informatica (CWI), Amsterdam
- Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2020, *Nature Methods*, 17, 261, doi: [10.1038/s41592-019-0686-2](https://doi.org/10.1038/s41592-019-0686-2)
- Wyatt, M. C. 2008, *ARA&A*, 46, 339, doi: [10.1146/annurev.astro.45.051806.110525](https://doi.org/10.1146/annurev.astro.45.051806.110525)
- Zammit-Mangion, A., Sainsbury-Dale, M., & Huser, R. 2025, *Annual Review of Statistics and Its Application*, 12, 311, doi: <https://doi.org/10.1146/annurev-statistics-112723-034123>
- Zubko, V. G., Mennella, V., Colangeli, L., & Bussoletti, E. 1996, *MNRAS*, 282, 1321, doi: [10.1093/mnras/282.4.1321](https://doi.org/10.1093/mnras/282.4.1321)

APPENDIX A: CORNER PLOTS

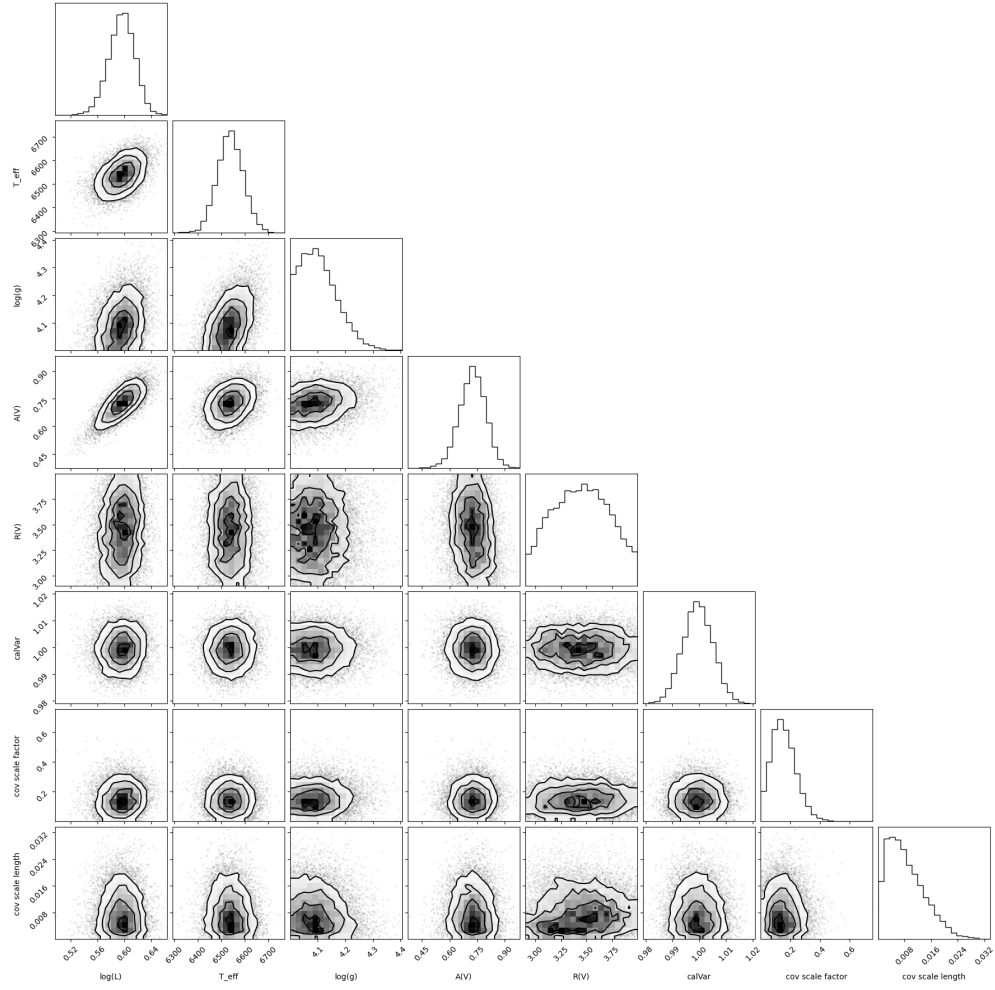


Figure A1. Corner plot for photometry+spectrum inference of astrophysical parameters. As expected, there are mild correlations between $\log(L)$, T_{eff} and A_V .

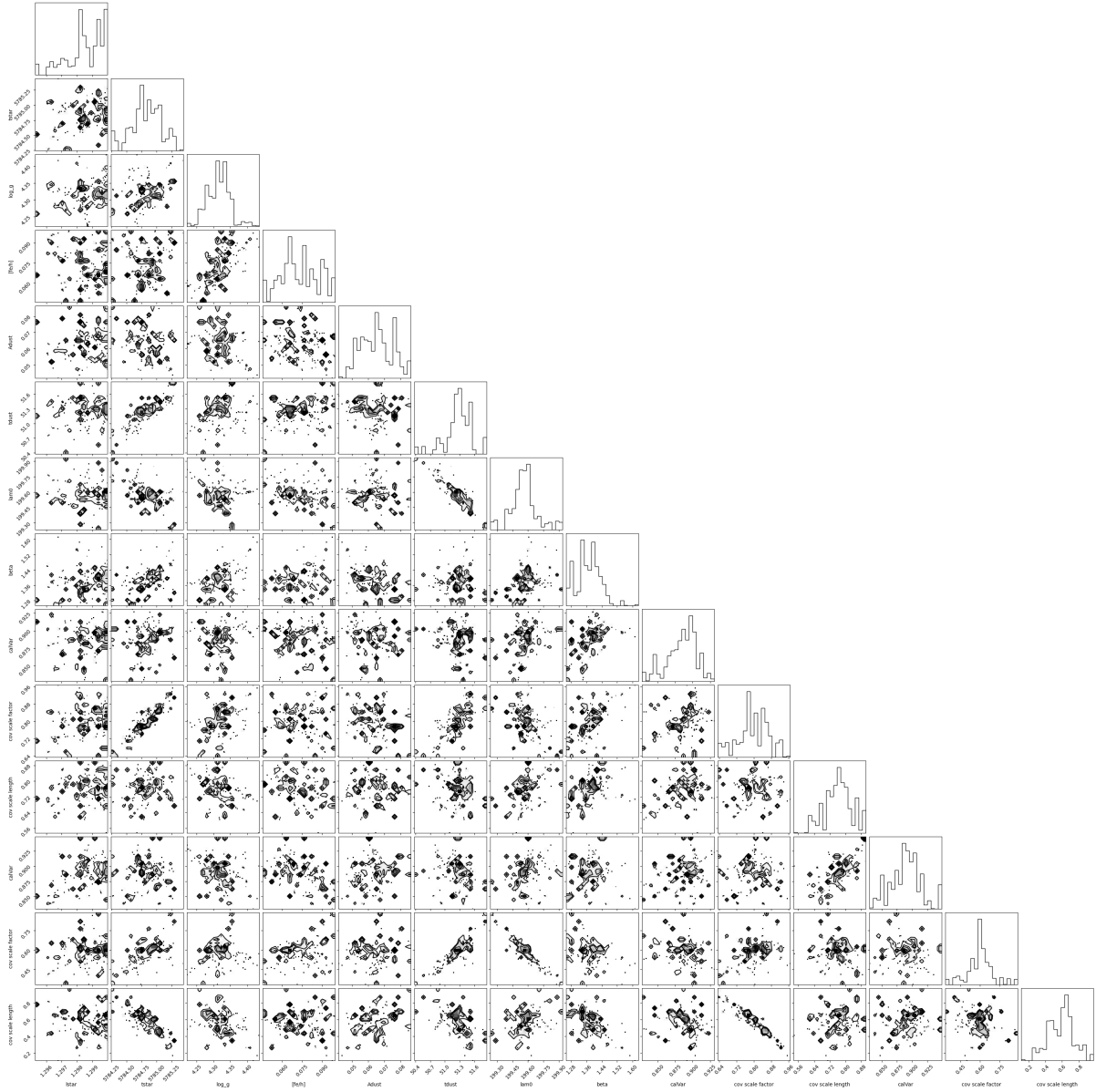


Figure A2. Corner plot for the star plus disc model showing posterior probability distribution of the model fitting as described in the text. The posterior probability distributions are comprised of 6,000 realisations of the model.

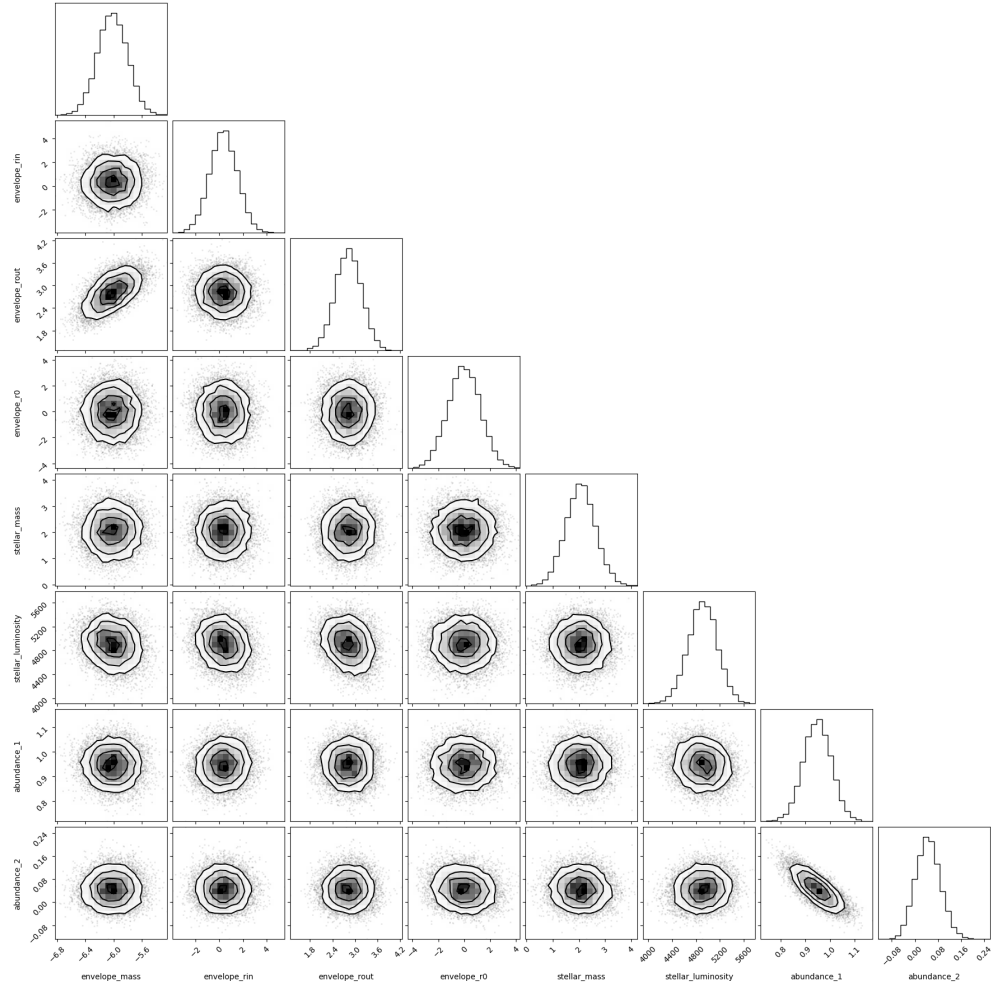


Figure A3. Corner plot for the carbon-star model, showing only the parameters of the physical model and excluding the noise-model parameters.