

Bayesian Approach to Galactic Cosmic Ray Propagation

**G. Jóhannesson,^a R. Ruiz De Austri,^b A. C. Vincent,^c I. V. Moskalenko*,^d
E. Orlando,^d T. A. Porter,^d A. W. Strong,^e R. Trotta^f**

^aScience Institute, University of Iceland, Dunhaga 5, IS-107 Reykjavík, Iceland

^bInstituto de Física Corpuscular, E-46980 Paterna, Spain

^cInstitute For Particle Physics Phenomenology, Durham University, Durham DH1 3LE, UK

^dHansen Experimental Physics Laboratory and Kavli Institute for Particle Astrophysics and Cosmology, Stanford University, Stanford, CA 94305, U.S.A.

^dMax-Planck-Institut für extraterrestrische Physik, Postfach 1312, D-85741 Garching, Germany

^fImperial College London, Astrophysics, Blackett Laboratory, London SW7 2AZ, UK

E-mail: gudlaugu@glast2.stanford.edu, r Ruiz@ific.uv.es,
aaron.vincent@durham.ac.uk, imos@stanford.edu,
eorlando@stanford.edu, tporter@stanford.edu, aws@mpe.mpg.de,
r.trotta@imperial.ac.uk

The fully Bayesian approach to the problem of deriving constraints for cosmic ray (CR) model parameters has several advantages. These are: (i) an efficient global scan of the whole parameter space allowing us to explore and take into account parameter correlations and degeneracies, (ii) a best-fit point and statistically well-defined errors on the parameters, (iii) the ability to include and marginalize over "nuisance" parameters (such as modulation potential and error rescaling parameters) making the analysis more robust. For this study, we use the latest version of the CR propagation code GALPROP together with the BAMBI code, the most efficient Bayesian analysis code available to date that combines MultiNest with Neural networks. The results of the analysis will be reported during the conference.

*The 34th International Cosmic Ray Conference,
30 July- 6 August, 2015
The Hague, The Netherlands*

*Speaker.

1. Introduction

Considerable achievements in astrophysics of CRs in recent years have become possible due to superior instrumentation launched into space and to the top of the atmosphere. Such advances stimulate hopes for significant advances in many related areas, such as indirect searches for dark matter, the origin and propagation of CR, particle acceleration in putative CR sources (SNRs) and the interstellar medium (ISM), studies of our local Galactic environment, CR propagation in the heliosphere, and the origin of extragalactic diffuse emissions.

Getting most of the accurate astrophysical measurements requires *reliable and detailed calculations*. Our current knowledge of CR propagation in the Galaxy is based on a large body of observational data together with substantial theoretical background: the latest developments in CR acceleration and transport mechanisms, detailed maps of the three-dimensional Galactic gas distribution, detailed studies of the interstellar dust, radiation field, and magnetic field, as well as up-to-date particle and nuclear cross section data and codes. However, the number of parameters in realistic models incorporating all of this information is large, and using the available data to perform statistical inference on the models' free parameters is a highly non-trivial task. In the past this has only been possible with analytical or semi-analytical models where the computation is fast [1]-[5]. But, such models necessarily require many simplified assumptions to allow the problem to be analytically tractable and to reduce the computational load, making the estimation of the confidence level of their results difficult. More realistic treatments using the analytic approach lead to a growing complexity of the formulae, thus removing any computational advantage over the purely numerical approach [6].

The Bayesian methods are more and more often used in multi-dimensional statistical analyses in physics and astrophysics, where traditional fitting methods have failed or require unrealistic resources. The fully Bayesian approach to the problem of deriving constraints for CR propagation model parameters has several advantages. Firstly, the higher efficiency of Bayesian methods allows us to carry out a global statistical analysis of the whole parameter space, rather than be limited to scanning a reduced number of dimensions at the time. This is important in order to be able to fit simultaneously all relevant CR parameters and to explore degeneracies. Secondly, we can marginalize (i.e., integrate over) the parameters one is not interested in at almost no additional computational costs, thus obtaining probability distributions for the parameters of interest that fully account for correlations in the global parameter space. Thirdly, our method returns not only a global best fit point, but also statistically well-defined errors on the parameters, which is one of the most important achievements of this work. Finally, we are able to include in our analysis a large number of “nuisance” parameters (such as modulation potentials and experimental error rescaling parameters, see below for details) that mitigate the impact of potential systematic errors in the data and/or in the theoretical model, thus making our fits much more robust. Bayesian inference however requires to choose priors for the parameters involved. This is done very carefully in the present work, and our results do not depend strongly on the choice of priors, which again is an hallmark of a robust statistical analysis.

The GALPROP¹ code is the most advanced of its kind (see Moskalenko et al., these Proceedings). GALPROP uses astronomical information and other data as input to self-consistently

¹<http://galprop.stanford.edu>

predict CRs, γ -rays, synchrotron and other observables. The code provides a full numerical calculation of the CR spectra and intensities, together with the diffuse emissions associated with the CRs interacting with the interstellar gas, radiation, and magnetic fields. A considerable optimization and parallelization of the fully numerical GALPROP code made possible a complete inference for propagation parameters in a Bayesian framework [7].

The purpose of this work is to build on the framework established in [7] and improve it in several ways. We introduce an automated neural network training in the form of the BAMBI algorithm, which speeds up the convergence of our scans by a factor of ~ 2 . The ensuing trained neural network can then be used to conduct a profile likelihood analysis, which (being typical 10 times more computationally expensive than a Bayesian posterior) would otherwise be computationally impossible. The speed-up at this stage is of a factor $\sim 10^5$ or more. From the physics point of view, we now constrain both the CR propagation model parameters and the source abundances, using an iterative scheme to convergence.

2. Bayesian Inference

Bayesian inference is based on the posterior probability distribution function (pdf) for the parameters, which updates our state of knowledge from the prior by taking into account the information contained in the likelihood. Denoting by Θ the vector of parameters one is interested in constraining, and by \mathbf{D} the available observations, Bayes Theorem reads

$$P(\Theta|\mathbf{D}) = \frac{P(\mathbf{D}|\Theta)P(\Theta)}{P(\mathbf{D})}, \quad (2.1)$$

where $P(\Theta|\mathbf{D})$ is the posterior distribution on the parameters (after the observations have been taken into account), $P(\mathbf{D}|\Theta) = \mathcal{L}(\Theta)$ is the likelihood function (when considered as a function of Θ for the observed data \mathbf{D}) and $P(\Theta)$ is the prior distribution, which encompasses our state of knowledge about the value of the parameters before we have seen the data. Finally, the quantity in the denominator of eq. (2.1) is the Bayesian evidence (or model likelihood), a normalizing constant that does not depend on Θ and can be neglected when interested in parameter inference. The evidence is obtained by computing the average of the likelihood under the prior (so that the r.h.s. of eq. [2.1] is properly normalized),

$$P(\mathbf{D}) = \int P(\mathbf{D}|\Theta)P(\Theta)d\Theta. \quad (2.2)$$

The evidence is the prime quantity for Bayesian model comparison, which aims at establishing which of the available models is the “best” one, i.e., the one that fits the data best while being the most economical in terms of parameters, thus giving a quantitative implementation of Occam’s razor [8].

Together with the model, the priors for the parameters which enter Bayes’ theorem, eq. (2.1), must be specified. Priors should summarize our state of knowledge and/or our theoretical prejudice about the parameters before we consider the new data, and for the parameter inference step the prior for a new observation might be taken to be the posterior from a previous measurement (for model comparison issues the prior is better understood in a different way, see [9]).

The problem is then fully specified once we give the likelihood function for the observations. The posterior distribution $P(\Theta|\mathbf{D})$ is determined numerically by drawing samples from it and Markov Chain Monte Carlo (MCMC) techniques can be used for this purpose.

3. The BAMBI algorithm

In this work we use the BAMBI algorithm [10] for Bayesian sampling. BAMBI implements MULTINEST [11]-[13] to perform nested sampling, as described in [14, 15], and the neural network training algorithm SKYNET [16] to learn the likelihood function and thus accelerate the sampling procedure.

3.1 MULTINEST

MULTINEST is a generic algorithm that implements the nested sampling technique. This technique is aimed at computing the Bayesian evidence, but is able to produce samples from the posterior in the process of doing so. MULTINEST is able to take advantage of parallel computing architectures by allowing each CPU to compute a new proposal point. As the run progresses, the actual sampling efficiency (fraction of accepted samples from total samples proposed) will drop as the ellipsoidal approximation is less exact and the likelihood constraint on the prior is harder to meet. By computing N samples concurrently, we can obtain speed increases of up to a factor of N with the largest increase coming when the efficiency drops below $1/N$.

3.2 SKYNET

SKYNET is an algorithm for training artificial neural networks. Training is performed using a fast, approximate second-order algorithm to find the neural network weights that best approximate a value of $\mathcal{L} = \log(P(\mathbf{D}|\Theta, \phi))$ for a given input $\{\Theta, \phi\}$. This method efficiently finds an optimal set of weights and is designed to minimize overfitting to the training data. ℓ_2 norm regularization aids the algorithm in finding the global optimum. A test data set, distinct from the training data, is used to stop training when the algorithm begins to overfit to the training data. The algorithm is described in more detail in [16].

3.3 Combining into BAMBI

Once SKYNET's training has converged on likelihood samples provided by MULTINEST, within the BAMBI framework the network is tested for the accuracy of its predictions. If the root-mean-square error is below a user-defined threshold, the network will be used for calculating future likelihoods by MULTINEST. If not, then more samples will be generated using the original likelihood function and training will resume once enough new samples have been collected. This setup is explained in more detail and examples are provided in [10]. Setting the accuracy threshold too low will require more samples from the original likelihood and more network training, while setting it too high can produce unreliable likelihood calculations that affect the evidence and posterior samples.

4. Propagation Model and Parameters

Our benchmark model for this study is the diffusion-reacceleration (DR) model, which is by far the most commonly used propagation model used with GALPROP [17]-[29]. The aim of this study is to simultaneously constrain the propagation parameters, as in [7], as well as the primary CR abundances, since these have a non-negligible effect on the propagation parameters that best fit the observed CR fluxes. For each parameter in Θ_D we used a uniform prior distribution, with endpoints that were informed by the results of [7].

The nuclear chain that we use begins at ^{30}Si and proceeds down to protons. The source abundances of nuclei $6 \leq Z \leq 14$ have an important influence on the B/C and $^{10}\text{Be}/^9\text{Be}$ ratios used in this study. We therefore let the abundances of the ten most important elements vary freely, with prior ranges determined by the measured CR abundances from ACE data at a few 100 MeV/nucleon [30]. The elements that are allowed to vary are H, He, C, N, O, Ne, Na, Mg, Al and Si. The abundances A_i are defined with respect to the proton injection abundance A_H , whose absolute normalization is fixed by its *final* flux at Earth, N_p , at the reference energy $E_{\text{ref}} = 10^2$ GeV. We label the abundance parameters Θ_A , and define $A_H \equiv 1.06 \times 10^6$.

Since this constitutes a significant enlargement of the parameter space, we split the analysis into two sets of scans: the first over abundances, keeping the propagation parameters fixed, and the second over propagation parameters, fixing the abundances to the posterior mean of the first scan. We describe this procedure in more detail in Section 4.1.

4.1 Iterative Procedure

Given the large parameter space (18 physical parameters, in addition to the 7 nuisance parameters), it is highly advantageous to split the search into two distinct runs, which we alternated in order to obtain an iteratively-converged result. This is facilitated by the fact that, for a fixed set of propagation parameters, the final CR composition of the ISM depends linearly on the injection abundance of each isotope. It thus becomes necessary to run GALPROP only once per nuclear species ($O(10)$) rather than once per likelihood evaluation ($O(10^5)$). We thus ran a first, fast scan over the abundance parameters Θ_A , using the results of a first low-resolution scan to fix the propagation parameters. We then ran our full MULTINEST/BAMBI scan over the propagation parameters Θ_D , this time fixing the abundances to the posterior means obtained in the first scan. A second abundance scan was then performed with the updated propagation parameters. In order to avoid a second, computationally expensive scan over Θ_D , we performed the second iteration by importance sampling of a thinned version of the first chain that we obtained. This allowed the likelihood evaluations to be massively parallelized, which is not possible with MultiNest.

5. Conclusion

This work is still in progress. The results will be reported during the conference.

Acknowledgments

GALPROP development is supported through NASA Grants No. NNX10AE78G, NNX13A092G, and NNX13AC47G.

References

- [1] F. Donato, D. Maurin, and R. Taillet, *beta -radioactive cosmic rays in a diffusion model: Test for a local bubble?*, *A&A* **381** (2002) 539–559, [[astro-ph/0108079](#)].
- [2] D. Maurin, F. Donato, R. Taillet, and P. Salati, *Cosmic Rays below $Z=30$ in a Diffusion Model: New Constraints on Propagation Parameters*, *ApJ* **555** (2001) 585–596, [[astro-ph/0101231](#)].
- [3] D. Maurin, R. Taillet, and F. Donato, *New results on source and diffusion spectral features of Galactic cosmic rays: I B/C ratio*, *A&A* **394** (2002) 1039–1056, [[astro-ph/0206286](#)].
- [4] D. Maurin, A. Putze, and L. Derome, *Systematic uncertainties on the cosmic-ray transport parameters. Is it possible to reconcile B/C data with $\delta = 1/3$ or $\delta = 1/2$?*, *A&A* **516** (2010) A67, [[arXiv:1001.0553](#)].
- [5] A. Putze, L. Derome, and D. Maurin, *A Markov Chain Monte Carlo technique to sample transport and source parameters of Galactic cosmic rays. II. Results for the diffusion model combining B/C and radioactive nuclei*, *A&A* **516** (2010) A66, [[arXiv:1001.0551](#)].
- [6] A. W. Strong, I. V. Moskalenko, and V. S. Ptuskin, *Cosmic-Ray Propagation and Interactions in the Galaxy*, *Ann. Rev. Nucl. Part. Sci.* **57** (2007) 285–327, [[astro-ph/0701517](#)].
- [7] R. Trotta, G. Jóhannesson, I. V. Moskalenko, T. A. Porter, R. Ruiz de Austri, and A. W. Strong, *Constraints on Cosmic-ray Propagation Models from A Global Bayesian Analysis*, *ApJ* **729** (2011) 106, [[arXiv:1011.0037](#)].
- [8] R. Trotta, *Applications of Bayesian model selection to cosmological parameters*, *MNRAS* **378** (2007) 72–82, [[astro-ph/0504022](#)].
- [9] R. Trotta, F. Feroz, M. Hobson, L. Roszkowski, and R. Ruiz de Austri, *The impact of priors and observables on parameter inferences in the constrained MSSM*, *Journal of High Energy Physics* **12** (2008) 24, [[arXiv:0809.3792](#)].
- [10] P. Graff, F. Feroz, M. P. Hobson, and A. Lasenby, *BAMBI: blind accelerated multimodal Bayesian inference*, *MNRAS* **421** (2012) 169–180, [[arXiv:1110.2997](#)].
- [11] F. Feroz and M. P. Hobson, *Multimodal nested sampling: an efficient and robust alternative to Markov Chain Monte Carlo methods for astronomical data analyses*, *MNRAS* **384** (2008) 449–463, [[arXiv:0704.3704](#)].
- [12] F. Feroz, M. P. Hobson, and M. Bridges, *MULTINEST: an efficient and robust Bayesian inference tool for cosmology and particle physics*, *MNRAS* **398** (2009) 1601–1614, [[arXiv:0809.3437](#)].
- [13] F. Feroz, M. P. Hobson, E. Cameron, and A. N. Pettitt, *Importance Nested Sampling and the MultiNest Algorithm*, *ArXiv e-prints* (2013) [[arXiv:1306.2144](#)].
- [14] J. Skilling, *Nested Sampling*, in *AIP Conf. Ser.* (R. Fischer, R. Preuss, and U. V. Toussaint, eds.), vol. 735, pp. 395–405, 2004.
- [15] J. Skilling, *Nested sampling for general Bayesian computation*, *Bayesian Analysis* **1** (2006) 833–861.
- [16] P. Graff, F. Feroz, M. P. Hobson, and A. Lasenby, *SKYNET: an efficient and robust neural network training tool for machine learning in astronomy*, *MNRAS* **441** (2014) 1741–1759, [[arXiv:1309.0790](#)].
- [17] I. V. Moskalenko and A. W. Strong, *Production and Propagation of Cosmic-Ray Positrons and Electrons*, *ApJ* **493** (1998) 694–707, [[astro-ph/9710124](#)].

- [18] A. W. Strong and I. V. Moskalenko, *Propagation of Cosmic-Ray Nucleons in the Galaxy*, *ApJ* **509** (1998) 212–228, [[astro-ph/9807150](#)].
- [19] I. V. Moskalenko and A. W. Strong, *Anisotropic Inverse Compton Scattering in the Galaxy*, *ApJ* **528** (2000) 357–367, [[astro-ph/9811284](#)].
- [20] A. W. Strong, I. V. Moskalenko, and O. Reimer, *Diffuse Continuum Gamma Rays from the Galaxy*, *ApJ* **537** (2000) 763–784, [[astro-ph/9811296](#)].
- [21] I. V. Moskalenko, A. W. Strong, J. F. Ormes, and M. S. Potgieter, *Secondary Antiprotons and Propagation of Cosmic Rays in the Galaxy and Heliosphere*, *ApJ* **565** (2002) 280–296, [[astro-ph/0106567](#)].
- [22] I. V. Moskalenko, A. W. Strong, S. G. Mashnik, and J. F. Ormes, *Challenging Cosmic-Ray Propagation with Antiprotons: Evidence for a “Fresh” Nuclei Component?*, *ApJ* **586** (2003) 1050–1066, [[astro-ph/0210480](#)].
- [23] A. W. Strong, I. V. Moskalenko, and O. Reimer, *Diffuse Galactic Continuum Gamma Rays: A Model Compatible with EGRET Data and Cosmic-Ray Measurements*, *ApJ* **613** (2004) 962–976, [[astro-ph/0406254](#)].
- [24] V. S. Ptuskin, I. V. Moskalenko, F. C. Jones, A. W. Strong, and V. N. Zirakashvili, *Dissipation of Magnetohydrodynamic Waves on Energetic Particles: Impact on Interstellar Turbulence and Cosmic-Ray Transport*, *ApJ* **642** (2006) 902–916, [[astro-ph/0510335](#)].
- [25] A. E. Vladimirov and et al., *GALPROP WebRun: An internet-based service for calculating galactic cosmic ray propagation and associated photon emissions*, *Computer Phys. Comm.* **182** (2011) 1156–1161, [[arXiv:1008.3642](#)].
- [26] I. Moskalenko, *Isotopic Production Cross Sections for CR Applications (ISOPROCS Project)*, *ICRC* **6** (2011) 283.
- [27] M. Ackermann et al., *Fermi-LAT Observations of the Diffuse γ -Ray Emission: Implications for Cosmic Rays and the Interstellar Medium*, *ApJ* **750** (2012) 3, [[arXiv:1202.4039](#)].
- [28] M. Kachelriess, I. V. Moskalenko, and S. S. Ostapchenko, *Nuclear Enhancement of the Photon Yield in Cosmic Ray Interactions*, *ApJ* **789** (2014) 136, [[arXiv:1406.0035](#)].
- [29] M. Kachelriess, I. V. Moskalenko, and S. S. Ostapchenko, *New Calculation of Antiproton Production by Cosmic Ray Protons and Nuclei*, *ApJ* **803** (2015) 54, [[arXiv:1502.0415](#)].
- [30] I. V. Moskalenko, A. W. Strong, and T. A. Porter, *Isotopic composition of cosmic-ray sources*, *ICRC* **2** (2008) 129–132, [[arXiv:0708.1974](#)].