The Astrophysical Journal, 824:16 (19pp), 2016 June 10 doi:10.3847/0004-637X/824/1/16

BAYESIAN ANALYSIS OF COSMIC RAY PROPAGATION: EVIDENCE AGAINST HOMOGENEOUS DIFFUSION

G. Jóhannesson[1], R. Ruiz de Austri[2], A. C. Vincent[3], I. V. Moskalenko[4,5], E. Orlando[4,5], T. A. Porter[4,5], A. W. Strong[6], R. Trotta[7,8], F. Feroz[9], P. Graff[10,11], and M. P. Hobson[9]
[1] Science Institute, University of Iceland, Dunhaga 3, IS-107 Reykjavik, Iceland
[2] Instituto de Física Corpuscular, IFIC-UV/CSIC, Valencia, Spain
[3] Institute for Particle Physics Phenomenology (IPPP), Department of Physics, Durham University, Durham DH1 3LE, UK
[4] Hansen Experimental Physics Laboratory, Stanford University, Stanford, CA 94305, USA
[5] Kavli Institute for Particle Astrophysics and Cosmology, Stanford University, Stanford, CA 94305, USA
[6] Max-Planck-Institut für extraterrestrische Physik, Postfach 1312, D-85741 Garching, Germany
[7] Astrophysics Group, Imperial Centre for Inference and Cosmology, Imperial College London, Blackett Laboratory, Prince Consort Road, London SW7 2AZ, UK
[8] Data Science Institute, William Penney Laboratory, Imperial College London, SW7 2AZ London, UK
[9] Astrophysics Group, Cavendish Laboratory, University of Cambridge, J.J. Thomson Avenue, Cambridge CB3 0HE, UK
[10] The Johns Hopkins University Applied Physics Laboratory, 11100 Johns Hopkins Road, Laurel, MD 20723, USA
[11] Department of Physics, University of Maryland, College Park, MD 20742, USA
Received 2016 February 8; accepted 2016 March 29; published 2016 June 3

ABSTRACT

We present the results of the most complete scan of the parameter space for cosmic ray (CR) injection and propagation. We perform a Bayesian search of the main GALPROP parameters, using the MultiNest nested sampling algorithm, augmented by the BAMBI neural network machine-learning package. This is the first study to separate out low-mass isotopes (p, $\bar{p}$, and He) from the usual light elements (Be, B, C, N, and O). We find that the propagation parameters that best-fit p, $\bar{p}$, and He data are significantly different from those that fit light elements, including the B/C and ^{10}Be/^{9}Be secondary-to-primary ratios normally used to calibrate propagation parameters. This suggests that each set of species is probing a very different interstellar medium, and that the standard approach of calibrating propagation parameters using B/C can lead to incorrect results. We present posterior distributions and best-fit parameters for propagation of both sets of nuclei, as well as for the injection abundances of elements from H to Si. The input GALDEF files with these new parameters will be included in an upcoming public GALPROP update.

Key words: astroparticle physics – cosmic rays – diffusion – Galaxy: general – ISM: general – methods: statistical

1. INTRODUCTION

Cosmic ray (CR) physics has entered a data-driven era. Until recently, CR observations were not accurate enough to warrant sophisticated studies of the propagation model parameter space, although some attempts have been made using mostly analytical propagation codes (e.g., Maurin et al. 2001, 2002, 2010; Donato et al. 2002; Putze et al. 2010). The launch of Payload for Antimatter Matter Exploration and Light-nuclei Astrophysics (PAMELA) in 2006 (Picozza et al. 2007), followed by the *Fermi* Large Area Telescope (*Fermi*-LAT) in 2008 (Atwood et al. 2009), and finally the Alpha Magnetic Spectrometer-02 (AMS-02) in 2011 have changed the landscape dramatically. The technologies employed by these space missions have enabled measurements with unmatched precision and data sets orders of magnitude larger than earlier experiments, which allow for searches of subtle signatures of new phenomena in CR and γ-ray data. For example, the claimed precision of AMS-02 data reaches 1%-3%. This requires propagation models of comparable accuracy in order to take full advantage of such high-quality data. Other missions have just launched (e.g., the CALorimetric Electron Telescope—CALET) or are awaiting launch (the Cosmic-Ray Energetics and Mass investigation—ISS-CREAM).

Our understanding of CR propagation in the Milky Way comes from a combination of substantial observational data and a strong theoretical basis. These include exhaustive maps of the distribution of gas in the Galaxy, interstellar dust, radiation field, and magnetic field, in addition to the latest data and codes describing particle and nuclear cross sections. Incorporation of such information is not possible using analytic methods and a fully numerical modeling for the treatment of CR propagation in the Galaxy is required. This was realized about 20 years ago, when some of us started to develop the most advanced fully numerical CR propagation code, called GALPROP[12], which is also a *de facto* standard in astrophysics of CRs (Moskalenko & Strong 1998; Strong & Moskalenko 1998). GALPROP makes use of information from astronomy, particle, and nuclear physics to predict CRs, γ-rays, synchrotron, and other observables in a self-consistent manner (a review can be found in Strong et al. 2007). The code's output includes CR spectra and intensities in every spatial grid point (in two dimensions and three dimensions; 2D and 3D, respectively) in the Galaxy, as well as the associated diffuse emissions from CR interactions with the interstellar gas, radiation, and magnetic fields.

The first successful attempt to embed such a detailed and fully numerical propagation code within a Bayesian statistical approach was made in 2011 (Trotta et al. 2011, hereafter Paper I). This became possible because of an extensive optimization and parallelization of the GALPROP code (necessary for the fast evaluation of the likelihood function) combined with highly efficient sampling algorithms employed in the SuperBayeS (de Austri et al. 2006) package. The advantages of such analysis are many. First, the Bayesian sampling method used enables a statistical analysis of the entire parameter space, rather than being limited to scanning a

[12] http://galprop.stanford.edu

reduced number of dimensions at a time. Crucially, this allows all relevant CR parameters to be fit simultaneously. Second, the parameters that are not of immediate relevance can be marginalized (integrated) over, without sacrificing computing time. Correlations in the global parameter space are thus fully accounted for in the resulting probability distributions. Third, we recover statistically well defined errors for each parameter in addition to the best fit; this constitutes one of the most important achievements of our earlier work. Finally, a large number of "nuisance" parameters can be incorporated, leading to an overall more robust fit. These parameters include the modulation potentials and experimental error rescaling parameters, and allow us to mitigate the effect of potential systematic errors that could arise from the data or the theoretical model.

This paper is novel in three ways. First, it builds on the framework established in Paper I and improves it in several directions. We demonstrate a first application of machine-learning techniques to speed up the computationally expensive inference from fully numerical codes in an automatically supervised manner. We introduce neural network (NN) training in the form of the BAMBI algorithm, which reduces computational effort by 20%. The ensuing trained NN can then in principle be used to conduct a (usually more costly) profile likelihood analysis with almost no computational effort. Second, we now constrain both the CR propagation model parameters and the source abundances, using an iterative scheme to convergence. Third, for the first time we split the data sets used into low-mass isotopes (p, $\bar{p}$, and He) and light nuclei (Be–Si). Significantly different inelastic cross sections of protons (and antiprotons) and heavier nuclei (~40 mb for protons versus ~250 mb for carbon) result in different CR species propagating from different distances in the Galaxy. Treating them separately allows us to directly probe different diffusion length scales in the Galaxy for the first time.

This paper is organized as follows: In Section 2 we give an overview of CR propagation and the GALPROP code, Bayesian inference, and the BAMBI/SkyNet codes. In Section 3, we discuss the propagation model used, its parameters and prior ranges (including source abundances), the iterative procedure we adopt to scan both the propagation parameters and the abundances, and the data sets adopted (including the likelihood function used). In Section 4, we present our results in terms of Bayesian posterior probability distributions and give the posterior mean and best-fit parameters, along with associated errors. In Section 5 we discuss our findings. Section 6 gives our conclusions. In the Appendix, we validate our NNs/SkyNet approach against a full (non-accelerated) scan.

2. THEORY AND ALGORITHMS

2.1. CR Propagation (GALPROP Code)

A brief review of CR production and propagation relevant to the present paper is given in Paper I and more information can be found in a review by Strong et al. (2007).

The theoretical understanding of the CR propagation in the interstellar medium (ISM) became a framework that the GALPROP code for CR propagation is built around. GALPROP numerically solves the system of partial differential equations describing the particle transport with a given source distribution and boundary conditions for all species of CRs.

In spite of its relative simplicity, the diffusion equation is remarkably successful at modeling transport processes in the ISM. Processes which enter propagation include diffusive reacceleration and, for nuclei, nuclear spallation, secondary particle production, radioactive decay, electron K-capture, and stripping, in addition to energy loss from ionization and Coulomb interactions.

The GALPROP source (injection) abundances are taken first as the solar system abundances, which are iterated (Moskalenko et al. 2008) to achieve an agreement with the propagated abundances as provided by *ACE* at ~200 MeV/nucleon (Wiedenbeck et al. 2001) assuming a propagation model. The source abundances derived for two propagation models, diffusive reacceleration and plain diffusion, were used in many GALPROP runs.

Galactic properties on large scales, including the diffusion coefficient, halo size, Alfvén velocity, and/or convection velocity, as well as the mechanisms and sites of CR acceleration, can be probed by measuring stable and radioactive secondary CR nuclei. The ratio of the halo size to the diffusion coefficient can be constrained by measuring the abundance of stable secondaries such as $_5$B). Radioactive isotopes ($^{10}_{4}$Be, $^{26}_{13}$Al, $^{36}_{17}$Cl, $^{54}_{25}$Mn) then allow the resulting degeneracy to be lifted (e.g., Ptuskin & Soutoul 1998; Strong & Moskalenko 1998; Webber & Soutoul 1998; Moskalenko et al. 2001). However, the interpretation of the peaks observed in the secondary-to-primary ratios (e.g., $_5$B/$_6$C, [$_{21}$Sc+$_{22}$Ti +$_{23}$V]/$_{26}$Fe) around energies of a few GeV/nucleon remains model-dependent.

Closely connected with the CR propagation, but not related to the present paper, is the production of the Galactic diffuse γ-rays and synchrotron emission (Orlando & Strong 2013). Proper modeling of the diffuse γ-ray emission, including the disentanglement of the different components, requires well developed models for the interstellar radiation field (ISRF) and gas densities, together with the CR propagation (see, e.g., Strong et al. 2007; Ackermann et al. 2012). Global CR-related properties of the Milky Way galaxy are calculated in Strong et al. (2010).

The Parker (1965) equation models the modulation of CRs as they propagate in the heliosphere. The modulated fluxes significantly differ from the interstellar spectra below energies of ~20-50 GeV/nucleon, but correspond to the ones actually measured by balloon-borne and spacecraft instruments.

Spatial diffusion, convection with the solar wind, drifts, and adiabatic cooling are the main mechanisms that determine transport of CRs to the inner heliosphere. These effects have been incorporated into realistic (time-dependent, three-dimensional) models (e.g., Florinski et al. 2003; Potgieter & Langner 2004; Langner et al. 2006). The "force field" approximation that is ordinarily used (Gleeson & Axford 1968) instead characterizes the modulation effect as it varies over the solar cycle using a single parameter—the "modulation potential." Despite having no predictive power, the force field approximation is a useful low-energy parameterization of the modulated spectrum for a given interstellar spectrum. A new stochastic 2D Monte Carlo (HelMod) code (Bobik et al. 2012) is being developed that would allow an accurate calculation of the heliospheric modulation for an arbitrary epoch and is fully compatible with GALPROP.

The GALPROP project now has nearly 20 years of development behind it. The key idea behind GALPROP is

that all CR-related data, including direct measurements, γ-rays, sychrotron radiation, etc., are subject to the same galactic physics and must therefore be modeled simultaneously. The original FORTRAN90 code has been public since 1998, and a rewritten C++ version was produced in 2001. The latest major public release is v54 (Vladimirov et al. 2011). It is also possible to configure and run GALPROP through a browser with WebRun at the dedicated website.[13] The website also contains links to all galprop publications and has detailed information on CR propagation and the GALPROP code.

We refer to Paper I and the dedicated website for a detailed description of the code. In this work we use a development version of the GALPROP code which is described in Moskalenko et al. (2015), and references therein. The development version has the possibility to vary the injection spectrum independently for each isotope. It also includes the dependency tree pre-built from the nuclear reaction network and made for each species to ensure that its dependencies are propagated before the source term is generated. This way, special cases of β^--decay (e.g., $^{10}\mathrm{Be}\rightarrow{}^{10}\mathrm{B}$) are treated properly in one pass of the reaction network instead of the two passes required before. This reduces the runtime of the GALPROP code by up to a half.

2.2. Statistical Framework

Here we summarize briefly the underlying statistical framework, referring the reader to Paper I for full details (see, e.g., Trotta 2008 for an overview of Bayesian methods). Bayesian inference works by evaluating the posterior probability distribution function (PDF) for the parameters of interest, which is the normalized product of the prior PDF (summarizing our state of knowledge before we see the data) and the likelihood function (which contains the information supplied by the data). Denoting by Θ the vector of parameters and by $\boldsymbol{D}$ the data, Bayes Theorem reads

$$P(\Theta|\boldsymbol{D}) = \frac{P(\boldsymbol{D}|\Theta)P(\Theta)}{P(\boldsymbol{D})}, \quad (1)$$

where $P(\Theta|\boldsymbol{D})$ is the posterior, $P(\boldsymbol{D}|\Theta) = \mathcal{L}(\Theta)$ is the likelihood function (when considered as a function of Θ for the observed data $\boldsymbol{D}$), and $P(\Theta)$ is the prior. The quantity in the denominator of Equation (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.

Together with the model for the data (entering the likelihood, possibly specified hierarchically; see, e.g., Shariff et al. 2015), the priors for the parameters which enter Bayes' theorem, Equation (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, possibly informed by the posterior from a previous measurement.

The problem is then fully specified once we give the likelihood function (see Section 3.3 below). The posterior distribution $P(\Theta|\boldsymbol{D})$ is determined numerically by drawing samples from it using an appropriate sampling scheme (see Section 2.3).

[13] http://galprop.stanford.edu/webrun

2.3. The BAMBI Algorithm

In order to efficiently explore the propagation model parameter space for a higher spatial and energy resolution than adopted in Paper I (hence with a higher computational cost per likelihood evaluation; see Section 3.1), in this work we upgrade our sampling techniques. We use the BAMBI algorithm (Graff et al. 2012), which implements the nested sampling algorithm MultiNest (Feroz & Hobson 2008; Feroz et al. 2009; Feroz et al. 2013), as described by Skilling (2004, 2006), and the NN training algorithm SkyNet (Graff et al. 2014) to learn the likelihood function online during the sampling and thus accelerate the sampling procedure. We briefly describe each algorithm below.

2.3.1. MultiNest

MultiNest is a highly efficient implementation of 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 (for details, see Feroz & Hobson 2008). In nested sampling, a set of "live" points is initially sampled from the prior distribution. The point with lowest (log-)likelihood, $\mathcal{L}_{\mathrm{min}}$, is then replaced by a new point that is sampled from the prior under the constraint that $\mathcal{L}_{\mathrm{new}} > \mathcal{L}_{\mathrm{min}}$. To facilitate this sampling from the constrained prior, MultiNest encloses the set of live points within a set of (possibly overlapping) ellipsoids from which new samples can be taken analytically. The ellipsoidal decomposition is chosen in order to minimize the sum of the volumes and is well suited to sampling from posterior distributions that exibit curving degeneracies and/or multi-modality. If subsets of the ellipsoid set do not overlap in parameter space, these can be identified as separate modes and evolved independently. The sampling converges when the Bayesian evidence is computed to within a user-specified tolerance.

MultiNest takes advantage of parallel computing architectures by allowing each CPU to compute a proposal replacement point simultaneously. As the run progresses, the actual sampling efficiency (fraction of accepted samples over number of proposal replacements) drops as the ellipsoidal approximation is less accurate and the likelihood constraint on the prior is harder to meet. By computing N proposal samples concurrently, we can obtain speed increases of up to a factor of $\simeq N$. This linear speed up, however, flattens once $N \approx 1/\epsilon$, where ϵ is the efficiency of the algorithm (i.e., the number of accepted samples over the number of likelihood evaluations). Past this point, a further increase in the number of CPUs does not result in any appreciable speed advantage. The actual values used in our scans are given in Section 2.3.2.

In addition to providing the log-likelihood and prior, the user only needs to tune a few parameters for any specific implementation. These are the number of live points (higher for higher dimensional parameter spaces, and/or multi-modal posteriors), the target efficiency (controlling the degree of shrinkage of the ellipsoids), and the tolerance (controlling the precision to be achieved on the evidence). The number of live points needs to be sufficient that all posterior modes are sampled (ideally with at least one live point in the initial set) and we use 2000 for our analyses (which does not suffer from multi-modality). The target efficiency affects how conservatively the ellipsoidal decomposition is made and a value of 0.5 was found to be sufficient; smaller values will produce more

accurate evidence values (irrelevant to the present study) but require more samples. Lastly, we chose a tolerance of 0.5 in the evidence calculation, as recommended in Feroz et al. (2009).

2.3.2. SkyNet and BAMBI

SkyNet (Graff et al. 2014) is an algorithm for training of artificial NNs—computational models that are used to approximate one or several target functions and that depend on a number of free input parameters. In our application, the input parameters are the free parameters in the model, Θ, and the target function is the log-likelihood, $\mathcal{L}$. SkyNet implements a feed-forward NN, where scalar values are passed from one layer to the next over weighted connections with nonlinear activation functions. BAMBI is a framework that joins up MultiNest with SkyNet: accepted samples from the MultiNest run are fed as training samples into SkyNet, which uses them to train the NN online (i.e., as the posterior sampling progresses).

Training is performed using a fast, approximate second-order algorithm to find the NN weights that best approximate a value of $\mathcal{L}$ for a given input Θ. 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 Graff et al. (2014).

The user must specify the size of the network, both in the number of hidden layers and the number of nodes in each. We use a network with a single hidden layer of 200 nodes. This was verified to give a sufficiently accurate approximation, as shown in the Appendix. The sigmoid activation function, $f(x) = 1/(1 + \exp(-x))$, is used for the hidden layer and a linear activation function, $f(x) = x$, is used for the output layer.

Once SkyNet's training has reached sufficient accuracy on likelihood values provided by MultiNest, within BAMBI 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 likelihood calls by MultiNest. Since the trained network is effectively an analytic interpolating function, calls to the neural-network approximated likelihood are almost instantaneous, thus greatly reducing the computational cost. If the predictions are insufficiently accurate, then more samples will be generated using the full likelihood function and training will resume once enough new samples have been collected. This setup is explained further and examples are provided in Graff et al. (2012). Setting the accuracy threshold too low will require more samples from the original likelihood and longer network training, while setting it too high can produce unreliable likelihood approximations that affect the accuracy of the posterior sampling. We use a tolerance of 0.8, which led to convergence in an acceptable amount of time, although it also led to some spurious maxima in the likelihood. These were removed by post-processing the posterior samples with a full evaluation of GALPROP (which can be done in an exact parallel way at a post-processing stage, and thus can benefit from massive parallel processing). Our two main BAMBI scans (see Section 3) were doubly parallelized. For the light elements (Be–Si) we used 96 CPUs, split over 12 MPI nodes, with each GALPROP evaluation using 8 openMP thread. For the second scan, over p, $\bar{p}$, and He only, we were able to use 144 CPUs, with 18 MPI processes using 8 openMP cores each. In this configuration, full convergence of these scans required approximately 2 million GALPROP calls each, totaling 35 CPU years in the light element case, and 5.5 CPU years for protons and helium. Over 99% of the computing power was used for GALPROP likelihood evaluations, with the remaining $\sim$1% spent on BAMBI training. In both cases, the NNs performed approximately 20% of the likelihood evaluations, saving around 10 CPU years, or 4.5 months of real computing time. More details are presented in the Appendix.

3. METHOD

3.1. Propagation Model and Parameters

The aim of this study is to simultaneously constrain the propagation parameters, as in Paper I, as well as the CR source abundances, since the latter are model-dependent.

Our benchmark model for this study is the diffusive reacceleration (hereafter DR) model, which is by far the most commonly used propagation model used with GALPROP (e.g., Moskalenko et al. 2002; Strong et al. 2004; Abdo et al. 2009; Ackermann et al. 2012, 2015; Vladimirov et al. 2012; Ajello et al. 2016, and references therein). The distribution of Galactic CR sources is based on pulsars (Lorimer 2004). For this study, we use $f_{\rm CR}(R) = (R/R_0)^{\alpha} e^{-\beta(R-R_0)}$, i.e., normalized to 1 at $R = R_0 = 8.5$ kpc, where $\alpha = 0.475$ and $\beta = 2.166$. The profile is constant for $R > 10$ kpc and cuts off at $R = 15$ kpc. The flattening in the outer Galaxy is suggested from Fermi studies (Abdo et al. 2010; Ackermann et al. 2011).

In this model, the spatial diffusion coefficient is given by

$$D_{xx} = \beta D_0 \left(\frac{\rho}{\rho_0}\right)^{\delta}, \tag{2}$$

where D_0 is a free normalization at the fixed rigidity $\rho_0 = 4 \times 10^3$ MV. For Kolmogorov diffusion, the power-law index is $\delta = 1/3$; however, we allow δ to freely vary. Reacceleration is modeled as a momentum-space diffusion where the coefficient D_{pp} is related to the spatial coefficient D_{xx} (Berezinskii et al. 1990; Seo & Ptuskin 1994) with

$$D_{pp} D_{xx} = \frac{4 p^2 v_{\rm Alf}^2}{3\delta(4 - \delta^2)(4 - \delta) w}, \tag{3}$$

where w characterizes the level of turbulence (we take $w = 1$ since only the quantity $v_{\rm Alf}^2/w$ is relevant); the Alfvén velocity $v_{\rm Alf}$ is allowed to vary freely.

The CR injection spectrum is modeled as a broken power law, with indices below $(-\nu_0)$ and above $(-\nu_1)$ the break as free parameters. This is known to be necessary in DR models in order to compensate for the large bump at low rigidities, a consequence of the large Alfvèn velocities needed to fit the B/C ratios below 1 GeV. While the location of this break is typically fixed for a given study (around $\rho_{\rm br} = 10$ GV), we allowed it to vary in our scan. Other models are able to reproduce the B/C ratio without a low-energy break in the injection spectra, but at the cost of an ad hoc break in the diffusion coefficient; these will be examined in detail in an upcoming study. Because we are using high-energy (>TeV) data, we must also include a second break in the injection spectrum at 220 GV, and thus a third freely varying index $-\nu_2$ (see a discussion of the possible origin of this break in

Table 1
Summary of Input Parameters and Prior Ranges

Quantity	Symbol	Prior range	Prior type
Propagation model parameters Θ_P			
Proton normalization (10^{-9} cm^2 sr^{-1}s^{-1} MeV^{-1})	N_p	[2, 8]	Uniform
Diffusion coefficient[a] (10^{28} cm^2 s^{-1})	D_0	[1, 12]	Uniform
Rigidity power-law index	δ	[0.1, 1.0]	Uniform
Alfvén speed (km s^{-1})	$v_{\rm Alf}$	[0, 50]	Uniform
Diffusion zone height (kpc)	z_h	[0.5, 20.0]	Uniform
Rigidity of first injection break (10^4 MV)	$\rho_{\rm br}$	[1, 30]	Uniform
Nucleus injection index below $\rho_{\rm br}$	ν_0	[1.00, 2.50]	Uniform
Nucleus injection index above $\rho_{\rm br}$	ν_1	[ν_0, 3.00]	Uniform
Nucleus injection index above 220 GV	ν_2	[1.5, ν_1]	Uniform
Difference between p and heavier inj. indices	δ_ν	[0.0, 1.0]	Uniform
Injection abundance parameters Θ_X[a]			
Proton normalization (10^{-9} cm^2 sr^{-1}s^{-1} MeV^{-1})	N_p	[2, 8]	Uniform
Helium	$X_{\rm He}$	[0.1, 2] $\times 10^5$	Uniform
Carbon	$X_{\rm C}$	[0.1, 6] $\times 10^3$	Uniform
Nitrogen	$X_{\rm N}$	[0.1, 5] $\times 10^2$	Uniform
Oxygen	$X_{\rm O}$	[0.1, 10] $\times 10^3$	Uniform
Neon	$X_{\rm Ne}$	[0.0, 1] $\times 10^3$	Uniform
Sodium	$X_{\rm Na}$	[0.0, 5] $\times 10^2$	Uniform
Magnesium	$X_{\rm Mg}$	[0.0, 1.5] $\times 10^3$	Uniform
Aluminum	$X_{\rm Al}$	[0.0, 5] $\times 10^2$	Uniform
Silicon	$X_{\rm Si}$	[0.0, 1.5] $\times 10^5$	Uniform
Experimental nuisance parameters			
Modulation parameters ϕ (MV)			Log-normal prior[b]
HEAO-3	$m_{\rm HEAO-3}$		[0, 1250]
ACE	$m_{\rm ACE}$		[0, 1125]
CREAM	$m_{\rm CREAM}$		Fixed (no modulation)
ISOMAX	$m_{\rm ISOMAX}$		[0, 1075]
PAMELA	$m_{\rm PAMELA}$		[0, 1000]
Variance rescaling parameters ($j = 1, ..., 5$)	$\log \tau_j$	[−1.5, 0.0]	Log-uniform on $\log \tau_j$

Notes.
[a] The hydrogen abundance is fixed to $X_{\rm H} \equiv 1.06 \times 10^6$.
[b] We use a log-normal distribution, where $\sigma = 50\%$ of the central value. Quoted limits correspond to 3σ.

Vladimirov et al. 2012). We also allow for a different injection spectrum for protons and heavier elements by setting the power-law indices of the proton injection spectrum to $\nu_i' = \nu_i + \delta_\nu$ for $i \in \{0, 1, 2\}$.

The other free model parameters are the halo height z_h and the normalization N_p of the propagated CR proton spectrum at 100 GeV. This yields a total of 10 free propagation parameters, summarized in Table 1; we label these

$$\Theta_P = \{N_p, D_0, \delta, v_{\rm Alf}, z_h, \nu_0, \nu_1, \nu_2, \rho_{\rm br}, \delta_\nu\}. \quad (4)$$

For each parameter in Θ_P we use a uniform prior distribution whose range is informed by the results of Paper I. Although we let more parameters vary in this analysis[14], we do not expect the posterior distributions to stray very far from the determination of Paper I. The prior ranges (informed by the results we obtained in Paper I) are given in Table 1, and are discussed in greater detail in Section 3.4.

Thanks to the speed up from BAMBI and the GALPROP code improvements mentioned in Section 2.1, we are able to use a finer grid in this work than in Paper I, giving better accuracy. We found that a spatial resolution of $\Delta r = 1$ kpc and $\Delta z = 0.2$ kpc and an energy grid with $E_{i+1}/E_i = 1.2$ was a reasonable compromise between accuracy and speed. The full set of numerical parameters that we adjusted is shown in Table 2.

Table 2
GALPROP Resolution Parameters Used in This Study

Variable name	Parameter	Value
`dr`	radial spacing (kpc)	1.0
`dz`	height spacing (kpc)	0.2
`Ekin_factor`	(log) kinetic energy spacing	1.2
`timestep_factor`	rescaling factor when reducing timesteps	0.5
`start_timestep`	size of initial timestep (s)	10^8
`end_timestep`	size of final timestep (s)	10^2
`timestep_repeat`	repeats per timestep	20
`max_Z`	number of elements	14

Due to their smaller inelastic cross sections, secondary antiprotons probe different length scales than the light elements; the diffusion parameters that characterize their propagation can therefore be different, and indeed we found in our test scans that the same parameter set would not allow a good simultaneous fit to the high- and low-mass data. We therefore split the propagation scan into two: one, propagating only protons, antiprotons, and helium; and one "light element"

[14] Specifically, $\rho_{\rm br}$, ν_0, δ_ν—see Table 1, as well as the 10 abundance parameters.

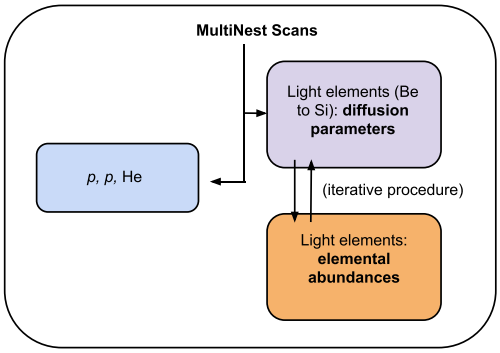

Figure 1. Sets of neural-network-assisted nested sampling scans that we perform in this work. We separate p, $\bar{p}$, and He (left) from the light elements (right) into two separate runs. For the light elements, we also vary the elemental abundances in separate, faster runs, which are performed iteratively with the propagation parameter scans. Since the GALPROP output is linear in the injection abundances, this allows extremely rapid convergence of the abundances. We will keep the same color code throughout the text: blue for p, $\bar{p}$, and He results, magenta for light elements, and orange for the abundances.

scan, propagating elements from beryllium up to silicon. This has the further advantage that the p, $\bar{p}$, and He scans do not require computation of the full nuclear network for each likelihood evaluation, allowing them to run quickly and in parallel with the light element scans. Thus, at the chosen resolution, our light element scan took approximately 9.8 CPU minutes per evaluation (or 1.22 minutes when parallelized over 8 CPUs), while the p, $\bar{p}$, and He case was sped up to 1.25 CPU minutes (or 9.4 seconds in real time).

The nuclear chain that we use for the light element scans begins at ^{30}Si and proceeds down to protons. The source abundances of nuclei $6 \leqslant Z \leqslant 14$ have an important influence on the B/C and ^{10}Be/^{9}Be ratios used in this study. We therefore let the abundances of the 10 most important elements vary freely, with prior ranges determined by the measured CR abundances from *ACE* data at a few 100 MeV/nucleon (George et al. 2009). The isotopes that are allowed to vary are ^{1}H, ^{4}He, ^{12}C, ^{14}N, ^{16}O, ^{20}Ne, ^{22}Ne, Na, ^{24}Mg, ^{25}Mg,^{26}Mg, ^{27}Al, ^{28}Si, ^{29}Si, and ^{30}Si; their prior ranges are presented in Table 1. The abundances X_i are scaled to the proton injection abundance $X_{\rm H}$, whose absolute normalization, N_p, is fixed by its *final* flux at Earth, at the reference energy $E_{\rm ref} = 10^2$ GeV. We label the 10-dimensional abundance parameters set Θ_X, defined with respect to $X_{\rm H} \equiv 1.06 \times 10^6$.

For each part of the experiment that provides data below a few GeV/nucleon, we must introduce an additional nuisance parameter ϕ_j ($j = 1,...,5$) to account for solar modulation. Furthermore, we introduce a set of parameters τ_j ($j = 1,...,5$) designed to mitigate the possibility that the fit is dominated by unknown systematic errors in the data, as explained in detail in Section 3.3, and following the procedure introduced in Paper I. We denote the joint set of nuisance parameters by ϑ.

Adding the abundance parameters constitutes a significant enlargement of the parameter space to be sampled: our full parameter space has 30 dimensions, and it would be computationally very costly to sample it simultaneously, even with MultiNest and BAMBI. Instead, we take advantage of the fact that for a given set of propagation parameters the final CR composition depends *linearly* on the injection abundance of each isotope. Thus, the likelihood as a function of Θ_X for fixed Θ_P is obtained quickly by linear rescaling of the CR spectra with Θ_X. This requires us to run GALPROP only once per nuclear species ($O(10)$ runs) and therefore the posterior for Θ_X conditional on Θ_P can be explored very quickly. Then we fix the abundances to their posterior mean, and sample from the posterior of Θ_P conditional on Θ_X. In all cases, we leave free the applicable nuisance parameters (solar modulation potentials m_j and error rescaling parameters τ_j). This procedure is then iterated with new abundances determined using propagation parameters fixed to the posterior mean of the scan over propagation parameters (Figure 1). This effectively amounts to implementing a Gibbs sampling scheme as follows:

$$\Theta'_X \sim P(\Theta_X|\boldsymbol{D}, \Theta_P) \quad (5)$$

$$\Theta'_P \sim P(\Theta_P|\boldsymbol{D}, \Theta'_X), \quad (6)$$

where a prime denotes the updated value of the parameter set.

We start our procedure with a scan over the abundance parameters Θ_X, fixing the propagation parameters to the posterior means of a low-resolution test scan over the propagation parameters Θ_P using the same isotopic abundances as in Paper I. This was followed by a propagation parameter scan at full numerical precision using the results of the first abundance scan, after which we performed a final abundance scan, which yielded no significant variation with respect to the first scan—and thus no need for a third iteration. The structure of our three scans is illustrated in Figure 1.

3.2. CR Data and Modulation

The data selection is based on similar principles as in Paper I. We use the most accurate CR data sets available, preferably taken near the solar minimum to reduce the effect of solar modulation. Table 3 lists the data we use in the analysis (obtained from a database by Maurin et al. 2013).

To reduce the number of nuisance parameters, we limit our data to instruments which cover many different CR species. As in the first paper, we use data from *ACE*-CRIS (George et al. 2009) for the lowest energies. Those data agree well with data from other instruments while providing better statistics and elemental coverage. At intermediate energies, the HEAO3-C2 data (Engelmann et al. 1990) provides good statistics while also agreeing with observations of other instruments. The recent elemental data observed by PAMELA (Adriani et al. 2014) has better statistics, but was not available at the start of this analysis, nor was the recent determination of the p flux by AMS-02 (Aguilar et al. 2015). Our results should not be affected by the former because the data from the *HEAO* and PAMELA instruments are compatible. We return to the recently released proton data from AMS-02 in Section 4.

For higher energies we decided to use only CREAM data (Ahn et al. 2008, 2009; Yoon et al. 2011), since its small energy binning was compatible with our method of evaluating the likelihood from a single energy point per bin, in contrast with the wide binning of, e.g., TRACER (Ave et al. 2008; Obermeier et al. 2011, 2012). For additional constraints on the propagation injection spectrum we also use H and He data from PAMELA at intermediate energies (Adriani et al. 2011) and CREAM at higher energies (Yoon et al. 2011). The PAMELA data was used because of its superior statistics and the high-energy CREAM data

Table 3
Data Used in This Analysis

Element	Experiment	Energy Range
	Data used in the p, $\bar{p}$, and He scan	
H	PAMELA ('06–'08)[a]	0.44–1000 GeV/n
	CREAM-I ('04–'05)[b]	3–200 TeV/n
H	PAMELA ('06–'08)[c]	0.28–128 GeV/n
He	PAMELA ('06–'08)[a]	0.13–504 GeV/n
	CREAM-I ('04–'05)[b]	0.8–50 TeV/n
	Data used in the light element scan and the abundance scan	
B/C	*ACE*-CRIS ('97–'98)[d]	72–170 MeV/n
	HEAO3-C2 ('79–'80)[e]	0.62–35 GeV/n
	CREAM-I ('04–'05)[f]	1.4–1450 GeV/n
^{10}Be/^{9}Be	*ACE*-CRIS ('97-'99)[g]	81–132 MeV/n
	ISOMAX ('98)[h]	0.51–1.51 GeV/n
B	*HEAO*-3 ('79–'80)[e]	0.62–35 GeV/n
C	*HEAO*-3 ('79–'80)[e]	0.62–35 GeV/n
	CREAM-II ('05–'06)[i]	86–7415 GeV/n
N	*HEAO*-3 ('79–'80)[e]	0.62–35 GeV/n
	CREAM-II ('05–'06)[i]	95–826 GeV/n
O	*HEAO*-3 ('79–'80)[e]	0.62–35 GeV/n
	CREAM-II ('05–'06)[i]	64–7287 GeV/n
	Data used in the abundance scan	
Ne	*ACE*-CRIS ('97–'98)[d]	85–240 MeV/n
	HEAO3-C2 ('79–'80)[e]	0.62–35 GeV/n
	CREAM-II ('05–'06)[i]	47–4150 GeV/n
Na	*ACE*-CRIS ('97–'98)[d]	100–285 MeV/n
	HEAO3-C2 ('79–'80)[e]	0.8–35 GeV/n
Mg	*ACE*-CRIS ('97–'98)[d]	100–285 MeV/n
	HEAO3-C2 ('79–'80)[e]	0.8–35 GeV/n
	CREAM-II ('05–'06)[i]	27–4215 GeV/n
Al	*ACE*-CRIS ('97–'98)[d]	100–285 MeV/n
	HEAO3-C2 ('79–'80)[e]	0.8–35 GeV/n
Si	*ACE*-CRIS ('97–'98)[d]	120–285 MeV/n
	HEAO3-C2 ('79–'80)[e]	0.8–35 GeV/n
	CREAM-II ('05–'06)[i]	27-2418 GeV/n

Notes.
[a] Adriani et al. (2011).
[b] Yoon et al. (2011).
[c] Adriani et al. (2010).
[d] George et al. (2009).
[e] Engelmann et al. (1990).
[f] Ahn et al. (2008).
[g] Yanasak et al. (2001).
[h] Hams et al. (2004).
[i] Ahn et al. (2009).

can be used without an additional modulation nuisance parameter.

For ^{10}Be/^{9}Be, we include *ACE* (Yanasak et al. 2001), which yields the most accurate measurements at low energies. These data are in agreement with Voyagers 1 and 2 (Lukasiak et al. 1999), as well as Ulysses (Connell 1998) data. ISOMAX (Hams et al. 2004) has given two data points at higher energies (per nucleon), with very large error bars, which we nonetheless include.

As in Paper I, we fit to the CR data in the whole energy range from few tens MeV/n to a few TeV/n and account for heliospheric modulation below ~20 GeV/n. We employ here the same method as in Paper I and use a simple force field approximation (Gleeson & Axford 1968), which is characterized with the value of the modulation potential. To avoid the uncertainty associated with the specific choice of the modulation potential, we allow some flexibility to the fits and include it as a free nuisance parameter (one free parameter per experiment). Gaussian priors with mean and standard deviation motivated by ballpark estimates of the modulation potential are used to avoid unphysical or implausible values. Because CREAM data start above 20 GeV/n, we do not include a modulation parameter for that experiment as it is irrelevant.

3.3. The Likelihood Function

We denote by $\Theta = \{\Theta_P, \Theta_X\}$ the joint set of CR propagation parameters and abundances, and by $\vartheta = \{\phi, \tau\}$ the joint set of nuisance parameters. For a given value of $\{\Theta, \vartheta\}$, we use GALPROP to compute the CR spectrum as a function of energy, $\Phi_Y(E, \Theta, \vartheta)$, for species Y. To mitigate against

Table 4
Summary of Constraints on All Propagation Parameters

	p, $\bar{p}$, and He scan			Light element (Be, ..., Si) scan		
Quantity	Best-fit value	Posterior mean and standard deviation	Posterior 95% range	Best-fit value	Posterior mean and standard deviation	Posterior 95% range
Diffusion model parameters Θ_P						
D_0 (10^{28} cm^2 s^{-1})	6.330	6.102 ± 1.662	[2.138, 8.205]	6.188	9.030 ± 1.610	[5.743, 11.256]
δ	0.466	0.461 ± 0.065	[0.343,0.586]	0.375	0.380 ± 0.018	[0.349, 0.412]
v_{Alf} (km/s)	8.922	8.970 ± 1.244	[7.036, 11.254]	32.573	30.017 ± 2.461	[25.484, 34.465]
z_h (kpc)	9.507	10.358 ± 4.861	[2.461, 19.034]	4.900	10.351 ± 4.202	[4.544, 19.078]
$\rho_{\rm br}$ (GV)	2.486	2.345 ± 0.344	[1.870,2.739]	15.782	16.687 ± 1.498	[14.051, 19.849]
ν_0	1.854	1.765 ± 0.229	[1.230, 2.133]	2.012	2.025 ± 0.073	[1.885, 2.155]
ν_1	2.352	2.358 ± 0.063	[2.230, 2.468]	2.549	2.548 ± 0.050	[2.452, 2.642]
ν_2	2.182	2.186 ± 0.068	[2.062, 2.308]	2.195	2.197 ± 0.088	[2.042, 2.374]
$10^9 N_p$ (cm^{-2} sr^{-1} s^{-1} MeV^{-1})	4.798	4.791 ± 0.066	[4.672, 4.913]	4.511	4.482 ± 0.220	[4.055, 4.884]
δ_ν	0.045	0.047 ± 0.009	[0.030, 0.064]	–	–	—
$X_{\rm He} \times 10^{-4}$	10.261	10.294 ± 0.505	[9.416, 11.240]	–	–	—
Experimental nuisance parameters						
Modulation parameters m_j						
PAMELA08 mod (MV)	637.625	645.740 ± 26.694	[601.226, 696.164]	–	–	—
HEAO80 mod (MV)	–	–	–	622.201	611.039 ± 93.229	[438.307, 789.523]
ACECRIS99 mod (MV)	–	–	–	445.975	421.682 ± 48.797	[330.972, 509.777]
ISOMAX98 mod (MV)	–	–	–	380.722	492.036 ± 206.243	[184.184, 958.214]
Variance rescaling parameters τ						
PAMELA08 $\log\tau$	−0.237	−0.277 ± 0.053	[−0.370, −0.181]	–	–	—
HEAO80 $\log\tau$	–	–	–	−0.516	−0.571 ± 0.089	[−0.740, −0.407]
ACECRIS99 $\log\tau$	–	–	–	0.000	−0.263 ± 0.209	[−0.780, −0.015]
CREAM05 $\log\tau$	−0.973	−1.014 ± 0.260	[−1.440, −0.480]	−0.704	−0.764 ± 0.140	[−1.053, −0.516]
ISOMAX98 $\log\tau$	–	–	–	−0.115	−0.604 ± 0.378	[−1.380, −0.045]

undetected systematics, we follow the procedure described in, e.g., Barnes et al. (2003). For each data set we introduce in the likelihood a parameter τ_j ($j = 1,...,5$), whose function is to rescale the variance of the data points in order to account for possible systematic uncertainties (see Paper I for a more detailed description). The role of the set of parameters $\tau = \{\tau_1,...,\tau_5\}$, which we call "error bar rescaling parameters," is to allow for the possibility that the error bars reported by each of the experiments underestimate the true noise. Furthermore, τ also takes care of all aspects of the model that are not captured by the reported experimental error: this includes also theoretical errors (i.e., the model not being completely correct), errors in the cross section normalizations, etc.

Assuming Gaussian noise on the observations, we take the following likelihood function for each observation of species Y at energy E_i

$$P(\hat{\Phi}_Y^{ij}|\Theta, \vartheta) = \frac{\sqrt{\tau_j}}{\sqrt{2\pi}\,\sigma_{ij}} \exp\left(-\frac{1}{2}\frac{(\Phi_Y(E_i, \Theta, \phi) - \hat{\Phi}_Y^{ij})^2}{\sigma_{ij}^2/\tau_j}\right), \quad (7)$$

where $\Phi_Y(E_i, \Theta, \vartheta)$ is the prediction from the CR propagation model for species Y at energy E_i, $\hat{\Phi}_Y^{ij}$ is the measured spectrum, and σ_{ij} is the reported standard deviation. The index i runs through the data points within each data set j. We assume bins are independent, such that the full likelihood function is given by the product of terms of the form given above:

$$P(\boldsymbol{D}|\Theta, \vartheta) = \prod_{j=1}^{5} \prod_{i=1}^{N_j} P(\hat{\Phi}_Y^{ij}|\Theta, \phi). \quad (8)$$

3.4. Choice of Priors

The full posterior distribution for the CR propagation model parameters Θ, the variance rescaling parameters τ, and the modulation parameters ϕ is written

$$P(\Theta, \vartheta|\boldsymbol{D}) \propto P(\boldsymbol{D}|\Theta, \phi, \tau)P(\Theta)P(\tau)P(\phi). \quad (9)$$

The likelihood $P(\boldsymbol{D}|\Theta, \tau, \phi)$ is given by Equations (7) and (8).

The priors $P(\Theta)$, $P(\phi)$, and $P(\tau)$ in Equation (9) determined in the following way. Priors on the model parameters $P(\Theta)$ are taken as uniform on Θ, with ranges given in Table 1. As shown below, the posterior is close to Gaussian and well constrained for Θ; the results should thus be fairly independent of the choice of priors.

We take a Gaussian prior on each of the modulation parameters. This is informed by the values provided by each experiment (see Table 1), in order to avoid physically

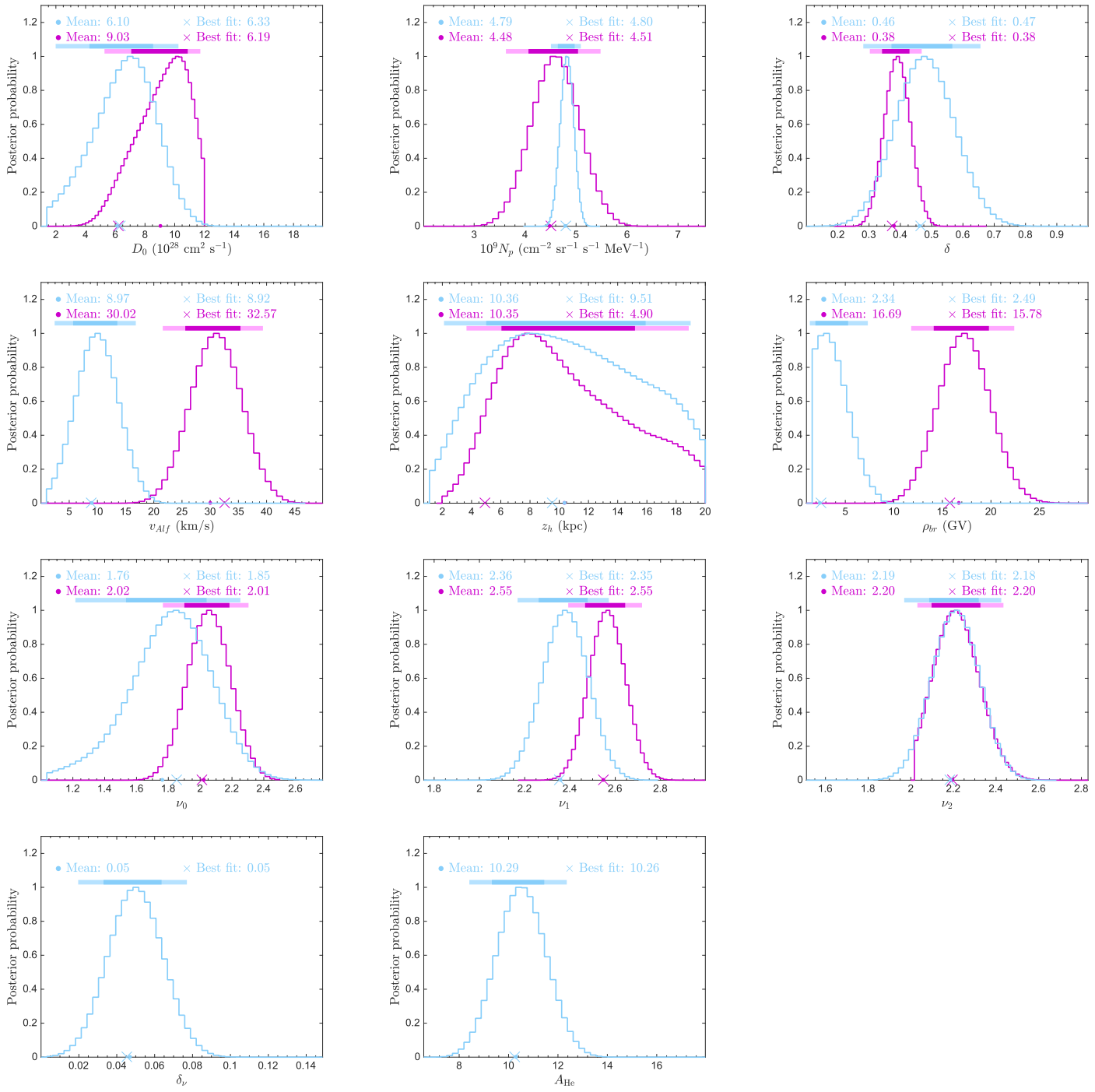

Figure 2. 1D marginalized posterior distributions, showing 1- and 2-sigma credible intervals, for the propagation parameters that were varied in the propagation scan. Light blue: the constraints from p, $\bar{p}$, and He scan, using PAMELA and CREAM data only; Purple: light element scan, fitting Be, B, C, N, and O data. (Given in Table 3.) While most of the propagation parameters overlap between runs, there is a clear ($>2\sigma$) separation seen in the Alfvén speed and in the low-energy injection break rigidity $\rho_{\rm br}$. Differences in the $D_0 - z_h$ plane can be clearly seen in Figure 3. The injection index for p, $\bar{p}$, and He is also consistently lower below the 220 GV break, suggesting a harder source injection spectrum.

unreasonable values. A description of the CR data sets is given in Section 3.2.

The τ_j are scaling parameters in the likelihood; the applicable prior is therefore given by the Jeffreys' prior, which is uniform on $\log \tau_j$ (see Barnes et al. 2003 or Jaynes & Bretthorst 2003). We thus adopt the proper prior

$$P(\log \tau_j) = \begin{cases} 2/3 \text{ for } & \text{for } -3/2 \leqslant \log \tau_j \leqslant 0 \\ 0 & \text{otherwise} \end{cases} \quad (10)$$

corresponding to a prior on τ_j of the form

$$P(\tau_j) \propto \tau_j^{-1}. \quad (11)$$

Including the nuisance parameters ϕ and τ (which are then marginalized over) in our analysis yields a more robust fit (as τ can absorb the effects of potential systematic effects in the data and ϕ incorporates solar modulation), while simultaneously giving more conservative constraints on the CR parameter

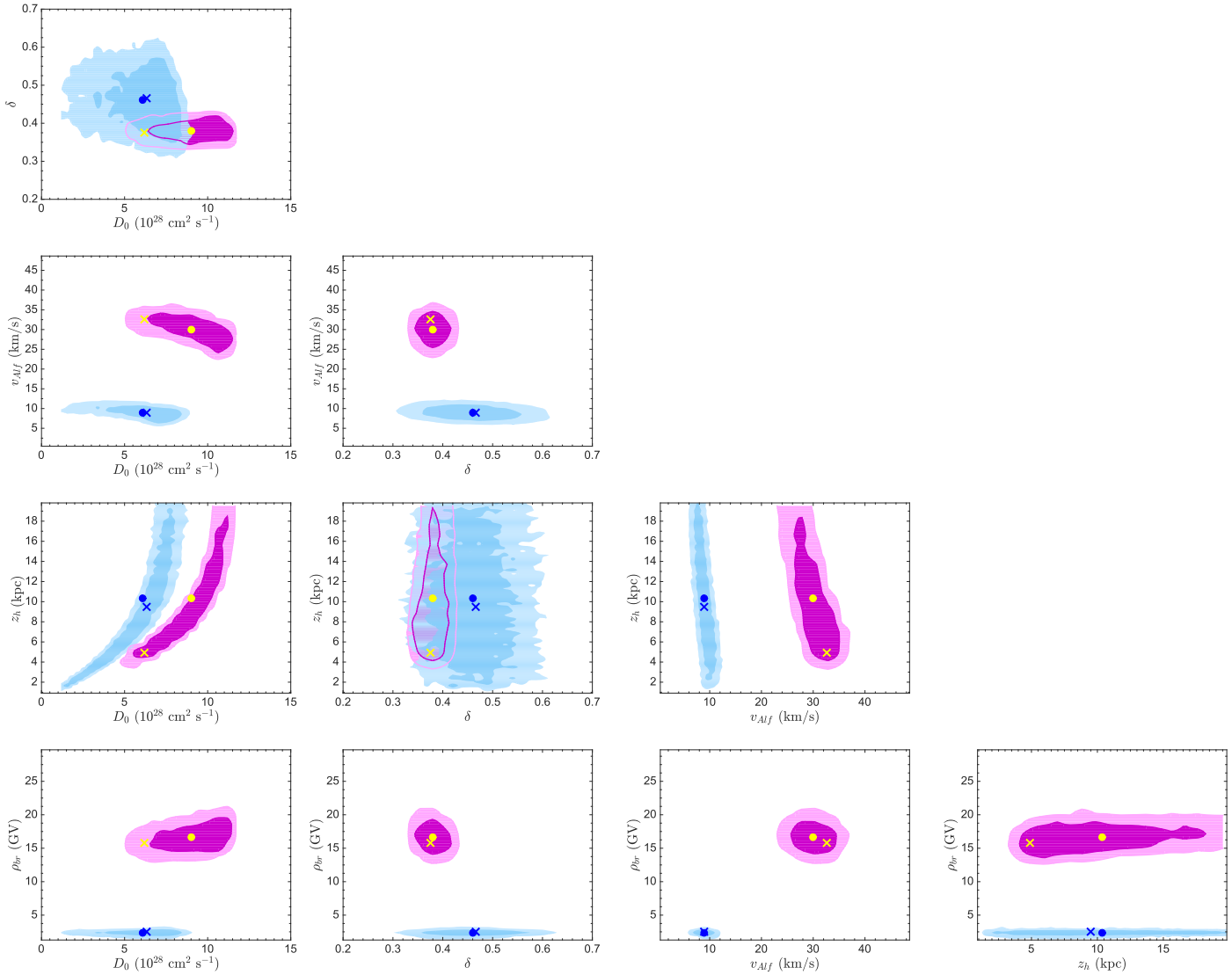

Figure 3. 2D posterior distributions, showing 1- and 2-sigma credible intervals for the p, $\bar{p}$, and He scan (blue), and for the light element (Be–Si, magenta). The posterior mean in each case is shown as a dot and the best fit as a cross.

space, since we fully account for degeneracies with all values of the nuisance parameters that are compatible with the data.

4. RESULTS

We report the best fit and posterior mean locations, along with confidence intervals, in Table 4. In Figure 2 we show the one-dimensional (1D) posterior distributions obtained for the propagation parameters in our full Multinest/BAMBI scan. We present 2D 68% and 95% (highest posterior density) credible regions for the most relevant propagation parameters in Figure 3. One can see that parameters are generally well constrained. However, it is apparent that the measurements of the ^{10}Be/^{9}Be ratio used here are not sufficient to break the well known degeneracy between D_{0xx} and z_h. Indeed, values of the halo height z_h can range between about 4 and 20 kpc, while the diffusion parameter normalization can be in the range [5, 11] $\times 10^{28}$ cm^2 s^{-1} in the light element scan. Comparing to the p, $\bar{p}$, and He scan, we can see that the inclusion of the radioactive-to-stable secondary ratio only marginally improves the constraint on the halo height, mostly from below.

While the 1D posterior distribution for D_{xx} and z_h from the two propagation runs contain a significant overlap, the 2D distributions show a clear separation between the two scans. There is therefore a significant tension between using $\bar{p}$ and B for the determination of propagation parameters. This is also evident in the $v_{\rm Alf}$ posterior distributions that are clearly separated for the two scans. These results thus strongly suggest that the propagation parameters are not constant over the entire Galaxy and using only the B/C ratio to determine the propagation parameters can significantly bias the results.

The reason behind this separation can partially be gleaned from the $z_h - D_0$ posterior distributions shown in Figure 3. For a fixed diffusion parameter, p, $\bar{p}$, and He probe a halo height that is approximately twice as large as the light elements. Since the inferred propagation parameters represent a volume-averaged quantity, these results indicate that p, $\bar{p}$, and He are probing a significantly larger volume than the light elements, and that the ISM properties vary quickly enough on these large distances to yield a significantly different Alfvèn speed, diffusion coefficient, and its index.

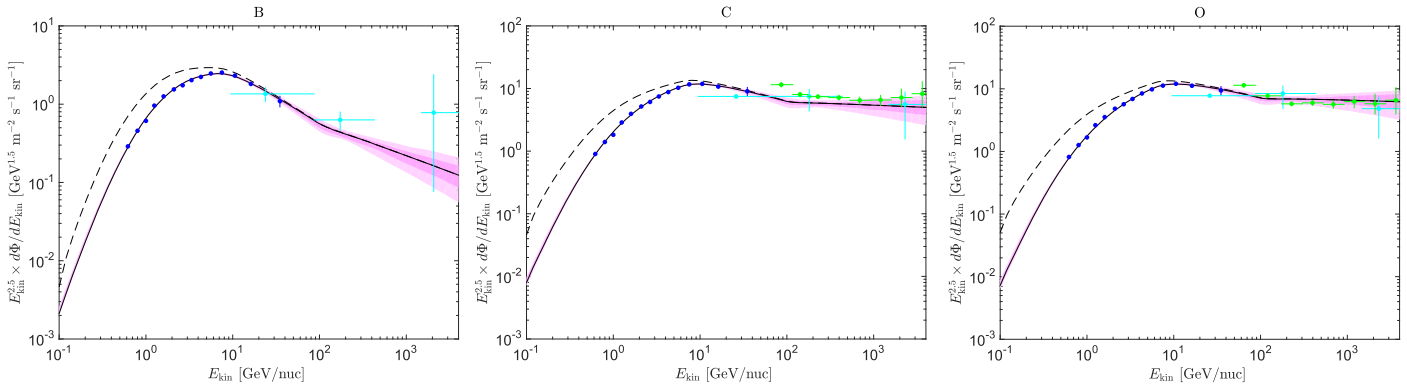

Figure 4. Spectral fluxes with 68% and 95% posterior regions from the posteriors of our light element (Be–Si) scan, shown in magenta in Figure 1, and using the HEAO modulation posteriors. Data shown are HEAO (blue), CREAM (green), and TRACER (cyan). The best fit is shown as a black line, and the dashed lines correspond to the LIS (unmodulated) spectra.

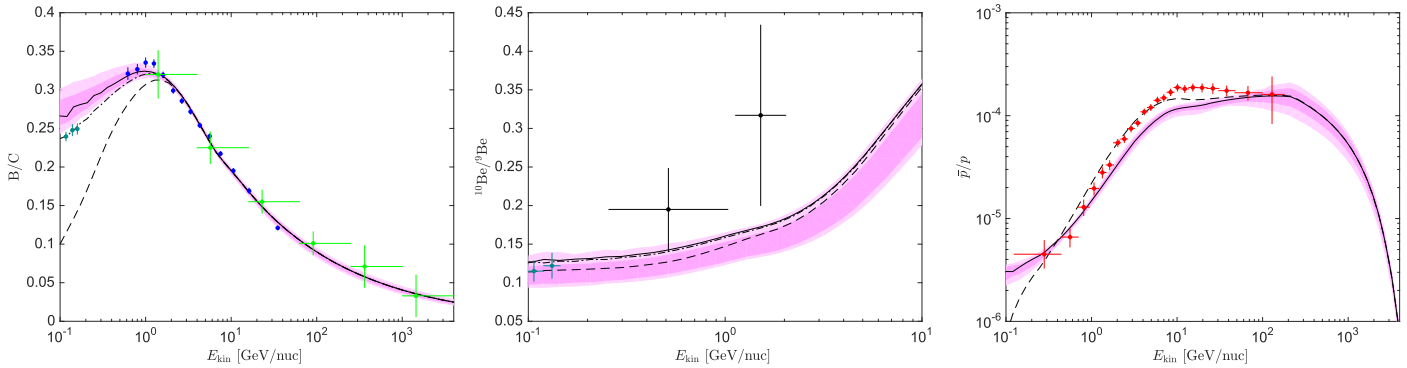

Figure 5. Secondary-to-primary ratio 68% and 95% posterior bands from our light element (Be–Si) scan, shown in magenta in Figure 1. The $\bar{p}/p$ ratio is shown to indicate that using the same propagation parameters for hydrogen yields a very bad fit to the data. Data shown are HEAO (blue), CREAM (green), ACE (light blue), ISOMAX (black), and PAMELA (red). The best fit is shown as a black line, and the dashed lines correspond to the LIS (unmodulated) ratios. In the left-hand panel, we use the HEAO modulation posterior, and the solid line uses the HEAO best-fit modulation potential. The dash-dotted line is the modulated spectrum using the best fit to the ACE-CRIS modulation potential; for clarity we do not show the posterior intervals for this case. Correspondingly, the central plot uses the ACE modulation (BF in black), and we show the best fit using the ISOMAX best-fit modulation potential with a dash–dotted line.

The posterior distributions for the injection spectrum for the two scans are very similar at high energies, but the spectrum of He is systematically harder than H below the break at 220 GV. The low-energy break of the proton and He spectrum is also lower than that of the heavier elements. Given that the power-law indices of the proton injection spectrum are δ_ν larger than that of the He injection spectrum, the results indicate that the injection spectrum of heavier elements are closer to that of the protons rather than that of He.

Figures 4–6 show 68% and 95% posterior intervals for our models overlaid on some of the data used in the analysis. The data-model agreement is very good in all cases. The need for the high-energy break is evident in the spectrum of protons and He in Figure 6 but we can also see from the spectra of heavier elements in Figure 4 that the high-energy break improves the agreement between model and data. The prediction for the $\bar{p}/p$ ratio of the light element scan in Figure 5 further illustrates the tension between the two data sets because there is a clear and significant mismatch between the data and model prediction. Indeed, a preliminary scan which included all data sets was not able to find an acceptable fit, yielding very large error rescaling parameters with $\tau_{\rm PAMELA}$ reaching the prior box boundary at $-\log \tau_{\rm PAMELA} = 1.5$. This is an indication that the model cannot simultaneously fit the light element and p, $\bar{p}$, and He data.

Note that newer $\bar{p}$ production cross sections (Kachelriess et al. 2015) yield better description of the $\bar{p}$ production in proton-proton, proton-nucleus, and nucleus-nucleus interactions, but were not available at the start of this analysis. They provide a higher $\bar{p}$ yield above $E_{\bar{p}} > 100$ GeV. Meanwhile, the parameterizations used in the present paper (Tan & Ng 1983a, 1983b; Moskalenko et al. 2002) were tuned to the $\bar{p}$ data at moderate energies providing a reasonable description in that energy range. The new cross sections are now incorporated into the GALPROP code to be used in our future calculations.

In Table 5 we provide the best fit, posterior means, and confidence intervals for the abundance parameters. These are compared with solar data in Figure 7. We also show the previously recommended values from GALPROP (Moskalenko et al. 2008). Abundances are in generally good agreement with the solar values, with partial volatiles (C, N, O, Ne) being depleted with respect to the solar abundances. This is a well known result, as CRs are likely preferentially accelerated from refractory-rich dust grains (Ellison et al. 1997; Meyer et al. 1997; Rauch et al. 2009; Ahn et al. 2010). The only major change versus previous GALPROP determinations is a higher sodium abundance, which is now brought in line with solar system measurements.

In Figure 8 we show the posterior distributions for the modulation potentials of the three experiments that we used

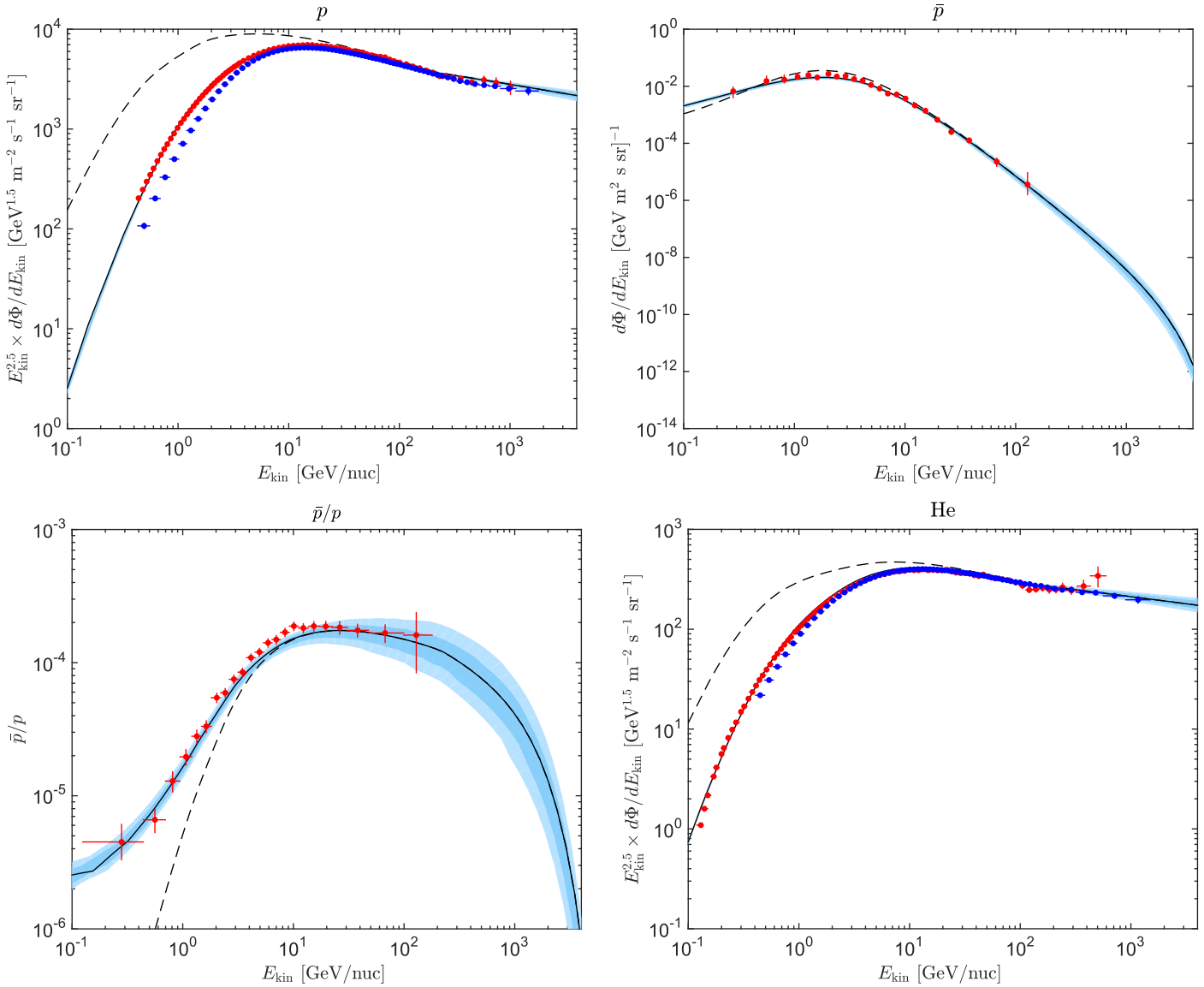

Figure 6. Spectra and $\bar{p}/p$ ratio 68% and 95% posterior bands of our $\bar{p}$, p, He scan, shown in blue in Figure 2. The best fit is plotted in black, and the dashed lines correspond to the LIS (umodulated) spectra. PAMELA data are shown in red. We also show recent AMS-02 (Aguilar et al. 2015, blue) for the available proton and helium flux data, which were not available at the time of our analysis (and hence are not included in the likelihood).

whose energy range was low enough to be affected by solar modulation. The posterior mean values are in good agreement with those estimated using ground-based neutron monitors (Usoskin et al. 2011).

The τ error bar rescaling parameters for each experiment are shown in Figure 9. These are mainly skewed toward $\log \tau = 0$, indicating no rescaling is necessary and thus good agreement between data sets. Some tension can be seen in the CREAM data (green points in our figures), possibly owing to the wide binning. Finally, the ISOMAX rescaling parameter was effectively consistent with the entire prior range, due to the paucity of available data (two data points).

5. DISCUSSION

A considerable underprediction of the $\bar{p}$ flux calculated in reacceleration models that are tuned to the B/C ratio was first noticed by Moskalenko et al. (2002). It has been shown that accurate antiproton measurements during the solar minimum of 1995–1997 by the BESS instrument (Orito et al. 2000) are inconsistent with existing propagation models at the ∼40% level at about 2 GeV, while the stated measurement uncertainties in this energy range were ∼20%. Using local CR measurements, simple energy dependence of the diffusion coefficient, and uniform CR source spectra throughout the Galaxy, conventional models failed to simultaneously reproduce both the secondary/primary nuclei ratio and $\bar{p}$ flux. The reacceleration model designed to match secondary/primary nuclei ratios (e.g., B/C) produces too few antiprotons because matching the B/C ratio at all energies requires the diffusion coefficient to be too large. The models without reacceleration can reproduce the $\bar{p}$ flux; however, the low-energy decrease in the B/C nuclei ratio requires an *ad hoc* break in the diffusion coefficient. The diffusion-convection model was constructed specifically to reproduce the $\bar{p}$ data, but required fine tuning. These results were later confirmed by Sina et al. (2005).

An attempt to find an acceptable solution for the reacceleration models was made by Moskalenko et al. (2003). They

Table 5
Summary of Constraints on Abundance Parameters

Quantity	Best-fit value	Posterior mean and standard deviation	Posterior 95% range
$10^9 N_p$ (cm^{-2} sr^{-1} s^{-1} MeV^{-1})	4.512	4.544 ± 0.097	[4.369,4.715]
$X_{\rm He} \times 10^{-4}$	9.044	8.975 ± 0.264	[8.499, 9.508]
$X_{\rm C}$	2578.407	2553.666 ± 66.318	[2442.083, 2666.097]
$X_{\rm N}$	210.667	221.389 ± 12.245	[199.314, 246.589]
$X_{\rm O}$	3372.090	3335.543 ± 82.290	[3184.869, 3492.503]
$X_{^{20}\rm Ne}$	304.155	306.029 ± 26.345	[259.181, 357.127]
$X_{^{22}\rm Ne}$	97.767	94.118 ± 22.321	[50.997, 137.982]
$X_{\rm Na}$	33.578	35.931 ± 2.812	[31.065, 41.583]
$X_{^{24}\rm Mg}$	583.254	548.250 ± 40.044	[472.095, 623.988]
$X_{^{25}\rm Mg}$	80.104	87.010 ± 28.553	[35.980, 143.917]
$X_{^{26}\rm Mg}$	85.998	100.340 ± 23.765	[55.965, 147.898]
$X_{\rm Al}$	79.410	78.102 ± 3.211	[72.186, 83.727]
$X_{^{28}\rm Si}$	643.797	629.755 ± 21.512	[589.202, 669.806]
$X_{^{29}\rm Si}$	44.661	47.725 ± 10.524	[27.996, 67.989]
$X_{^{30}\rm Si}$	32.996	38.987 ± 8.010	[23.997, 54.992]
Experimental nuisance parameters			
HEAO80 mod (MV)	593.085	591.606 ± 11.074	[573.154, 610.848]
ACECRIS99 mod (MV)	329.543	340.231 ± 14.137	[315.514, 371.142]
PAMELA08 mod(MV)	664.817	671.463 ± 21.223	[630.303, 708.612]
Variance rescaling parameters τ			
HEAO80 $\log\tau$	−0.615	−0.594 ± 0.062	[−0.721, −0.478]
ACECRIS99 $\log\tau$	−1.162	−1.269 ± 0.120	[−1.465, −1.037]
CREAM05 $\log\tau$	−1.039	−1.008 ± 0.087	[−1.184, −0.853]
PAMELA08 $\log\tau$	−1.500	−1.499 ± 0.001	[−1.499, −1.494]
TRACER06 $\log\tau$	−1.712	−1.563 ± 0.185	[−1.921, −1.228]

showed that the spectra of *primary nuclei* as measured in the heliosphere may contain a fresh, local, "unprocessed" component at low energies. The latter leads to an effective decrease in both the B/C ratio at low energies and the diffusion coefficient, thus increasing the production of antiprotons. The paper associated the fresh component with the Local Bubble and independent evidence for supernova activity in the solar vicinity in the last few million years was taken as a support to this idea.

Ptuskin et al. (2006) found that the diffusive reacceleration model with Iroshnikov-Kraichnan spectrum of interstellar turbulence $\delta = 0.5$ (Iroshnikov 1964; Kraichnan 1965) and wave damping helps to alleviate the problem, though does not solve it completely. The main idea of that paper is that the dissipation of waves due to the resonant interaction with CR particles may terminate the slow Kraichnan-type cascade below wavelengths 10^{13} cm thus leading to the increase in the diffusion coefficient at low rigidities. No significant effect of CR damping was found in the case of the fast Kolmogorov cascade.

These early papers (Moskalenko et al. 2002, 2003; Ptuskin et al. 2006) compared the predicted $\bar{p}$ flux to the data collected during the two balloon flights of the BESS instrument (Orito et al. 2000). The total number of collected antiprotons was between 51 and 64 per energy bin in four bins ranging from 1.52 to 3.00 GeV. Some of these antiprotons could be secondaries produced in the atmosphere above the instrument. The discrepancy with the predictions of the reacceleration model could also imply possible unaccounted systematic errors of the data analysis. However, direct measurements in space by the PAMELA experiment (Adriani et al. 2010) made during the next solar minimum confirmed the earlier BESS measurements with doubled statistics in the same energy range. Simultaneously, the PAMELA measurements of the B/C ratio (Adriani et al. 2014) yield a value of $\delta = 0.397 \pm 0.007$ for the index of the diffusion coefficient that is close to the classical value of $\delta = 1/3$, hinting at the

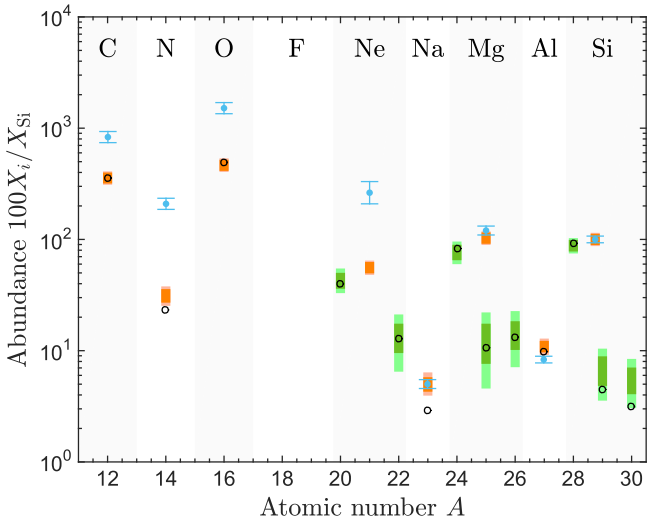

Figure 7. 95% (light bars) and 68% (dark bars) posterior intervals from our final abundance study. Total elemental abundances are in orange, while individual isotopes are in green. We show the latest determination of the solar photospheric (blue dots) elemental abundances and errors from Asplund et al. (2009), with updated heavier ($A \geqslant 23$) elemental abundances from Scott et al. (2015). We also show previously used values from GALPROP (Moskalenko et al. 2008) with open black circles.

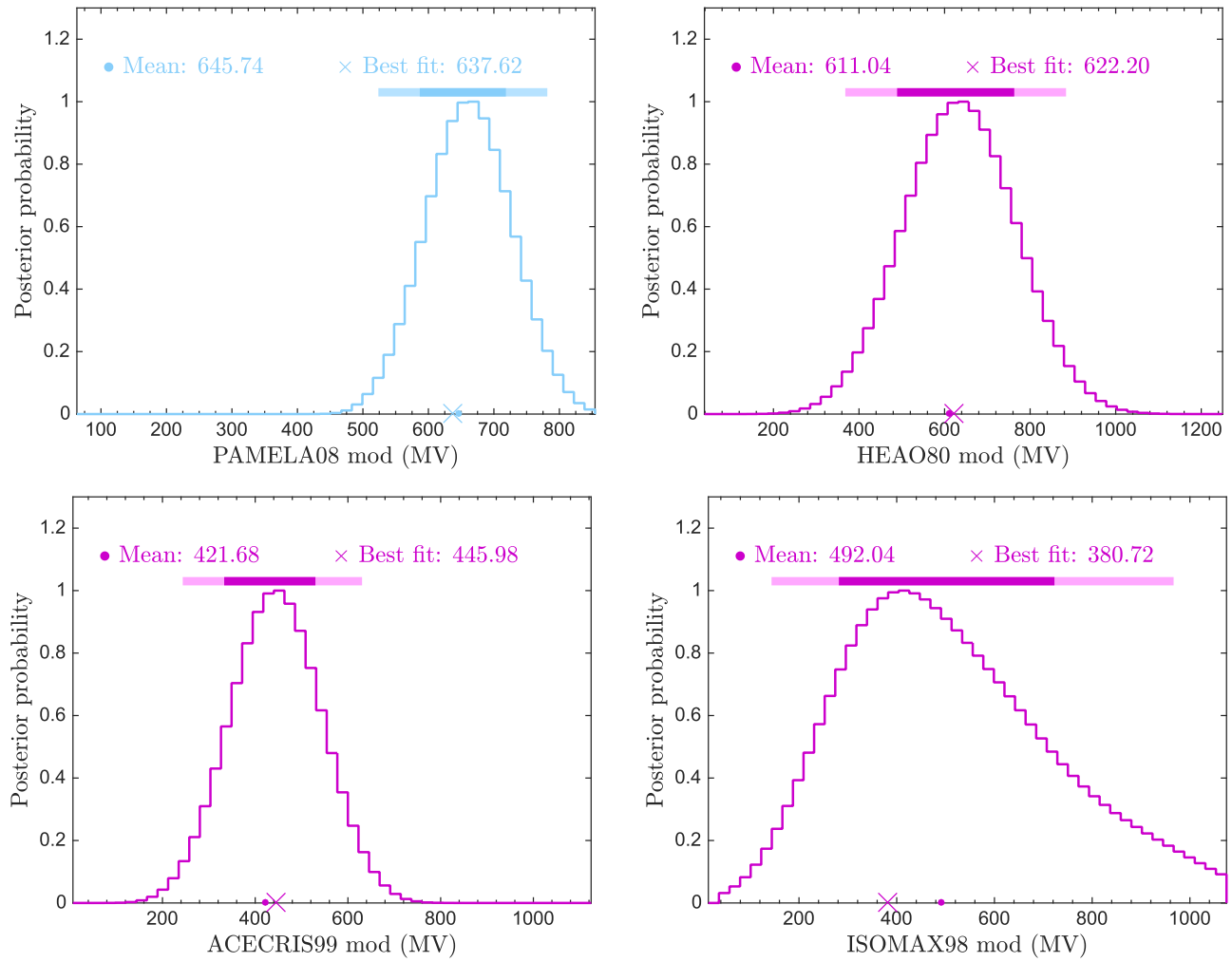

Figure 8. Posterior distributions of the modulation parameters for each experiment used in the fit, with 1- and 2-sigma credible intervals.

Kolmogorov spectrum of interstellar turbulence. Furthermore, the preliminary AMS-02 results for the B/C ratio reported at the AMS days meeting[15] agree with the PAMELA data and indicate a somewhat flatter index.

Agreement between BESS and PAMELA on the $\bar{p}$ measurement and a hint that the index of the diffusion coefficient is close to the Kolmogorov value support the idea that the discrepancy with the predicted $\bar{p}$ flux is inherent and not due to experimental uncertainty. Our first scan of the parameter space (see e.g., Paper I) quantitatively confirms this finding. Our new results (Section 4) show significant tension between a set of propagation parameters derived from a standard secondary-to-primary ratio B/C, and those derived from p, $\bar{p}$, and He data, as can be explicitly seen in Figure 5. This tension may, in fact, reflect the properties of significantly different Galactic volumes probed by different species.

To illustrate this point, let us calculate the effective propagation distance for different CR species. For the interaction timescale we have

$$\tau \sim [\sigma_r n c]^{-1}, \tag{12}$$

where σ_r is the total reaction cross section, $n \sim 1$ cm^{-3} is the average gas number density in the Galactic disk, and c is the speed of light. The effective propagation distance can be estimated as

$$\langle x \rangle \sim \sqrt{6D\tau} \sim \left(\frac{6D_0}{\sigma_r n c}\right)^{1/2} \left(\frac{\rho}{\rho_0}\right)^{\delta/2}. \tag{13}$$

In the case of nuclear species, the total reaction cross section is approximately

$$\sigma_r(A) \approx 250 \text{ mb } (A/12)^{2/3}, \tag{14}$$

which is made to roughly reproduce the cross sections published by Wellisch & Axen (1996), and corrected by H.P. Wellisch & D. Axen (1999, private communication). and we took $\sigma_r(^{12}\text{C}) \approx 250$ mb. In the case of p and $\bar{p}$, $\sigma_r^p \approx \sigma_r^{\bar{p}} \approx 40$ mb. The exact numbers are not very important as we are seeking for a rough estimate of the diffusion volume for different species at the rigidity of a few GV.

Table 4 gives the results of the propagation parameters scan. For the p, $\bar{p}$, and He scan, we have $D_0^p \approx 6 \times 10^{28}$ cm^2 s^{-1} at $\rho_0 = 4$ GV, and $\delta \approx 0.46$. For the light elements (Be–Si), we have $D_0^A \approx 9 \times 10^{28}$ cm^2 s^{-1} at ρ_0, and $\delta \approx 0.38$. The superscripts p and A are added to distinguish between the values derived from $\bar{p}$ and B propagation parameters scans.

[15] https://indico.cern.ch/event/381134/timetable/#20150415

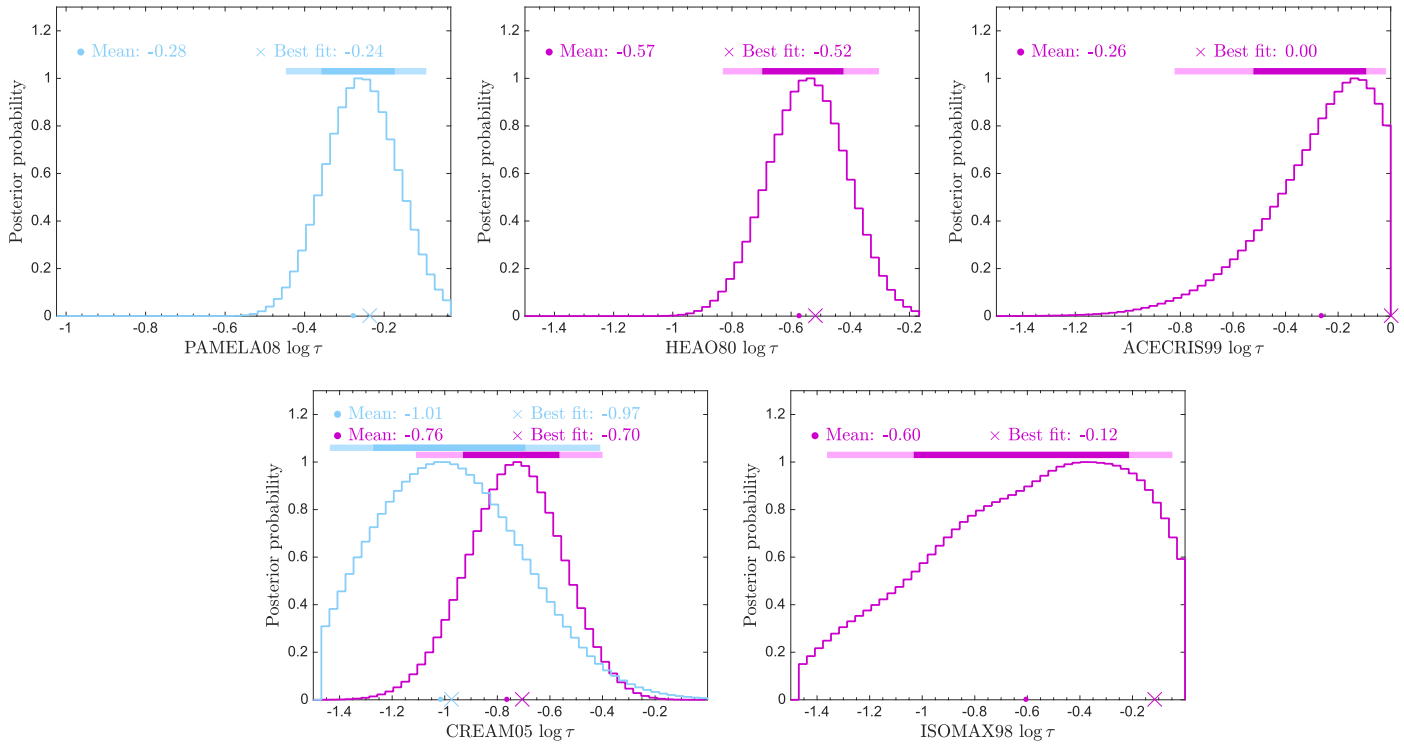

Figure 9. Posterior distributions of the τ rescaling parameters, with 1- and 2-sigma credible intervals.

The spectral indices are somewhat different, but we can use a single index of $\delta \approx 0.4$ in our estimates.

Substitution of these values into Equation (13) gives:

$$\langle x \rangle_A \sim 2.7 \text{ kpc} \left(\frac{A}{12}\right)^{-1/3} \left(\frac{\rho}{\rho_0}\right)^{\delta/2}, \quad (15)$$

$$\langle x \rangle_p \sim 5.6 \text{ kpc} \left(\frac{\rho}{\rho_0}\right)^{\delta/2}. \quad (16)$$

Even though the value of the diffusion coefficient derived from p, $\bar{p}$, and He is a factor of 1.5 smaller than that for the light nuclei, the former probes an area ($\propto \langle x \rangle^2$) of the Galaxy that is four times larger. This ratio does not depend on δ. The volume probed by the lighter species includes a considerable area in the inner Galaxy, where the supernova remnant (SNR) rate and the OB star distribution reach their maxima (at a distance of about 5 kpc from the Galactic center). It is thus natural to expect that more turbulent ISM has a smaller diffusion coefficient.

This is only an estimate, but it gives some idea of the typical distances. Even though CRs can in principle come from larger distances, their number density would be negligible compared to locally produced CRs of the same species. This estimate is consistent with the typical lifetime of CRs in the Galaxy assuming a uniform diffusion coefficient in the disk and halo. The best-fit halo size derived from the ^{10}Be/^{9}Be ratio is $z_h \approx 5$ kpc in the case of the light elements, and $z_h \approx 10$ kpc from the p, $\bar{p}$, and He scan (Table 4), i.e., larger than the effective distances given by Equations (15) and (16). Their posterior means are even larger, $z_h \approx 10.35$ kpc with 1σ error bars of 4.2 kpc and 4.9 kpc correspondingly.

Our results are therefore the first to definitively show that by separating the two data sets, one can fit them with two different reacceleration parameter sets. The significantly lower Alfvèn speed $v_{\rm Alf} \propto B/\sqrt{\rho_{\rm ISM}}$, 8.9 ± 1.2 km s^{-1} (p, $\bar{p}$, He) versus 30.0 ± 2.5 km s^{-1} (Be–Si), may hint at a smaller $B/\sqrt{\rho_{\rm ISM}}$, possibly owing to a denser ISM plasma as one approaches the inner Galaxy.

Variations of the propagation parameters throughout the Galaxy is not the only possible reason of the discussed differences. Source (SNe) stochasticity (Strong & Moskalenko 2001) may contribute to the local fluctuations in fluxes of individual CR species. Freshly accelerated CR particles from relatively recent supernova (SN) explosions may or may not lead to the increased local production of secondary species. As was already mentioned, the presence of local sources of low-energy primary nuclei could lead to effects that mimic the propagation parameters variations (Moskalenko et al. 2003). In particular, the value of the effective diffusion coefficient D_0^A could be reduced, i.e., made consistent with D_0^p, by invoking an additional component of the locally produced *primary* CR nuclei. Equations (15) and (16) indicate that such sources should be located within 1–2 kpc. Besides the Local Bubble, other obvious candidates are the local (Orion) arm and the Perseus arm, where the SN rate is higher than in the interarm region (Dragicevich et al. 1999).

If instead, the value of D_0^p is required to be made consistent with D_0^A, then CR sources should produce additional antiprotons. Production of secondary nuclei in the SNR shocks was proposed by Berezhko et al. (2003). Antiprotons are also secondary and thus can be produced in the same process (Blasi & Serpico 2009; see also (Kachelrieß& Ostapchenko 2013; Cholis & Hooper 2014; Mertsch & Sarkar 2014). However, this argument is circular unless we assume that there is a distinct type of CR sources that is nuclei ($Z > 2$) deficient and that this type of source has enough material nearby to produce additional antiprotons in significant amounts. The first

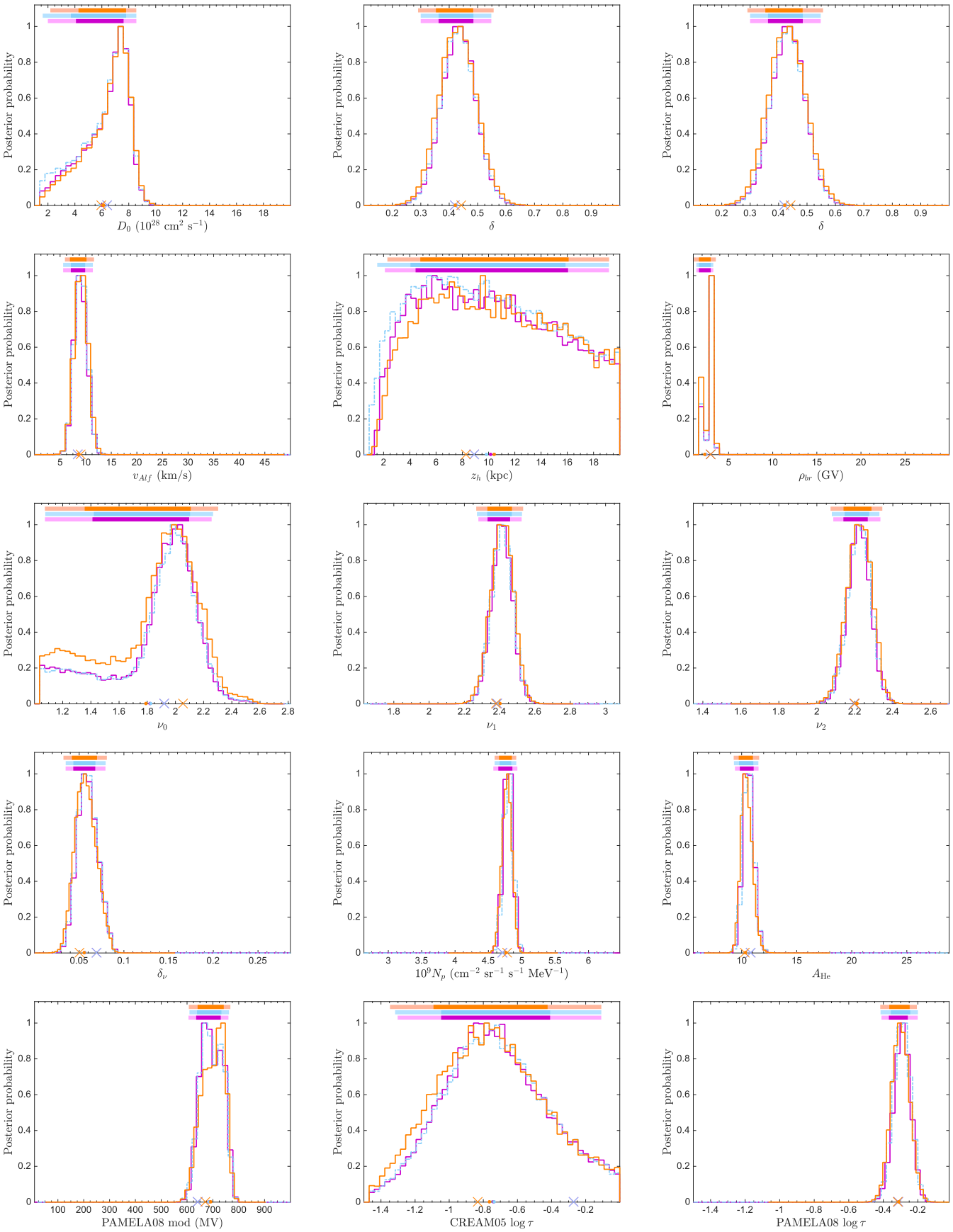

Figure 10. 1D posterior distributions (with 68% and 95% credible intervals) for the different CR propagation parameters in a low-resolution, $\{p, \bar{p}$, and He$\}$ propagation scenario using MultiNest as a sampler (no neural network speed up, magenta) and from BAMBI runs with two different values for the neural network input parameter σ. Light blue: $\sigma = 0.5$; Orange: $\sigma = 0.8$. All BAMBI chains have been post-processed in the same way as in our main paper runs.

hypothesis of the local sources producing mostly *primary* CR nuclei therefore appears more reasonable.

Even though the structure of our Galaxy has been the subject of research since the invention of the telescope in the beginning of the seventeenth century, only now we are starting to learn about its effects on CR fluxes.

The Galaxy is clearly not axially symmetric, yet the quality of the CR and diffuse emission data available until recently did not warrant propagation calculations beyond simple cylindrically symmetric geometry (Strong et al. 2007). The full 3D setup was available in GALPROP since the very beginning (Strong & Moskalenko 1998, 2001), but it was mostly used to test the cylindrically symmetric 2D solution. In fact, the entirely uniform, so-called the Leaky-Box model, was completely dominating CR modeling in the twentieth century.

Not surprisingly, the discussions on the influence of the Galactic structures on the intensity of CRs began about a decade ago.

The effects of the solar system's passage through the spiral arms on the global climate (ice ages) were discussed by Shaviv (2003). These ideas were further developed in Shaviv et al. (2009) in connection with the so-called positron excess reported by PAMELA (Adriani et al. 2009) and earlier by the HEAT experiment (Barwick et al. 1997). Clear evidence of the increased CR density in the spiral arms is provided by the *Fermi*-LAT residual maps (Ackermann et al. 2012), produced by subtracting the GALPROP diffuse γ-ray predictions from the *Fermi*-LAT skymaps. The most significant excesses coincide with the tangential directions to the spiral arms, which presumably contain freshly accelerated CRs. There is currently no fully consistent model that would incorporate the details of the Galactic structure. This is mostly connected with the difficulty of recovering the 3D structure of our Galaxy, such as the distributions of gas, magnetic field, SNRs, and star-forming regions using astrophysical observations. Dependence on the temperature gradient in the ISM is discussed in Erlykin et al. (2016), and possible effects of the details of the Galactic structure on CR propagation are actively discussed in the literature using a simplified description (e.g., Becker Tjus et al. 2015; Jóhanneson et al. 2015; Kissmann et al. 2015; Orlando et al. 2015; Porter et al. 2015; Benyamin et al. 2016). A nearby source (see, e.g., Erlykin & Wolfendale 2015) would also lead to similar problems.

The most complete ever scan of the parameter space for CR injection and propagation is another landmark of the present paper. Calculations of the CR source abundances were done in the past (e.g., Engelmann et al. 1990; Duvernois & Thayer 1996; Wiedenbeck et al. 2001, 2008). However, such calculations were usually made for elemental abundances[16] using the Leaky-Box model or its equivalent. By current standards, the models and data sets (e.g., semi-empirical cross sections) used in such calculations in the past were not detailed enough, but reflected the current state of knowledge at that time. The first successful attempt to find the source abundances and propagation parameters in a self-consistent way using a proper propagation code GALPROP was made by Moskalenko et al. (2008). The source (injection) abundances were taken first as the solar system abundances, which were then iterated to achieve an agreement with the propagated abundances as provided by *ACE* at ~200 MeV/nucleon (Wiedenbeck et al. 2001) assuming a propagation model, such as diffusive reacceleration or plain diffusion. The propagation parameters were then re-adjusted to reflect the final source abundances. Even though the resulting abundances are fairly close to the previous calculation (Figure 7), the current paper accomplishes a significantly more challenging task by performing a full NN-assisted scan over the 20 propagation and abundance parameters. Ten more nuisance parameters were included into the scan to account for possible systematic errors of different experiments. The result is the full set of best-fit values, posterior means, and standard deviations. The latter allows the detailed propagation calculations with meaningful constrains for related areas and for possible signatures of new physics.

Thus far, we have only considered the reacceleration model. Other models will be analyzed in forthcoming papers.

[16] Wiedenbeck et al. (2001, 2008) and other *ACE* team publications calculate *isotopic* source abundances.

6. CONCLUSIONS

We have performed the largest ever—in terms of number of free parameters, data, resolution, and computing time—study of CR propagation using a fully numerical state-of-the art computer code. By combining GALPROP with the BAMBI package, we were able to perform a full NN-assisted scan over the 20 propagation and abundance parameters, as well as 10 nuisance parameters. Two GALDEF input files based on the best fits found here will be included in an upcoming update of the publicly available GALPROP code[11].

Our results have highlighted two important conclusions. (1) available measurements of the radioactive species ^{10}Be are not sufficient to significantly remove the degeneracy between the halo height z_h and the diffusion parameter normalization D_0; and (2) The propagation parameters derived from the CR p, $\bar{p}$ and He data are *not* compatible with those found from fitting light elements Be–Si.

We take these results as a probable indication that the ISM properties differ significantly enough over kiloparsec scales to affect propagation of CRs, though we have mentioned other interpretations. This fact has important consequences for CR propagation studies: it is customary to use propagation parameters calibrated to local B/C data to predict fluxes of other CR species including electrons and positrons, both locally and as far away from the Earth as the Galactic center, or otherwise to assume an *ad hoc* functional form for the spatial dependence of the diffusion coefficient. Such approaches are particularly misleading in the search for physics beyond the Standard Model, such as signals of dark matter annihilation. An excess in antiprotons, positrons or γ-rays could indeed be an indication of a mischaracterized ISM, rather than a need for new physics.

This work has been supported by the Royal Society under the International Exchange Scheme, grant number IE120221 as well as by STFC grant ST/N000838/1.I.V.M., E.O., T.A.P. acknowledge support from NASA Grant No. NNX13AC47G, E.O. additionally acknowledges support from NASA Grant Nos. NNX16AF27G and NNX15AU79G, and T.A.P. additionally acknowledges support from NASA Grant No. NNX10AE78G. R. R.d.A. is supported by the Ramón y Cajal program of the Spanish MICINN and also acknowledges the support of the Spanish MICINN's Consolider-Ingenio 2010 Programme under the grant MULTIDARK CSD2209-00064, the *Invisibles* European ITN project (FP7-PEOPLE-2011-ITN,

PITN-GA-2011-289442-INVISIBLES, the "SOM Sabor y origen de la Materia" (2014-57816), the "Fenomenologia y Cosmologia de la Fisica mas alla del Modelo Estandar e lmplicaciones Experimentales en la era del LHC" (FPA2013-44773) MEC projects, and Severo Ochoa del MINECO: SEV-2014-0398. A.C.V. was supported by FQRNT (Québec) and *Invisibles*. The use of Imperial College High Performance Computing cluster is gratefully acknowledged.

APPENDIX
VALIDATION OF BAMBI/SkyNet

Before launching our high-resolution physics scans, we performed a series of validation scans, with the goal of optimizing the BAMBI framework with MultiNest as well as determining the reliability of the trained neural nets (NNs). In order to determine the optimal input settings for the network, (i.e., those that maximize speed up while predicting the likelihood function reliably), several runs were carried out with different values for the two main settings that determine the efficiency and accuracy of the NN training: $n_{\rm hid}$, the number of hidden nodes; and σ, which sets the desired accuracy for the predicted likelihood value before the network takes over as an interpolator. These tests were carried out with fixed elemental abundances and with a low GALPROP resolution (`dr` = 1.0, `dz` = 0.1, `Ekin` = 2.0, `starttimestep` = 1.0e9, `endtimestep` = 1.0e2, `timestepfactor` = 0.25, `timesteprepeat` = 20) in order to rapidly obtain trained networks.

We found that a training parameter value $\sigma = 0.5$ accurately reproduced the results obtained using MultiNest as a sampler (and no BAMBI acceleration). However, in this case only 3% of the likelihood evaluations were performed by the neural nets, hence with a very minimal speed up in the computational time. In contrast, $\sigma = 0.8$ led to a good convergence with 21% of the likelihood evaluations performed by the nets. Since some of the resulting samples gave spurious high-likelihood regions, we further post-processed them to remove any residual inaccuracy. The posterior distributions from these test runs are shown in Figure 10, where they are compared with the posterior resulting from a full MultiNest run.

The analysis for $\sigma = 0.5$ was carried out for both `nhid` = 200 and `nhid` = 300. Both runs led to good parameter inference results, and the number of likelihood evaluations computed using the network was very similar. Based on these results, we decided to fix the input network settings to `nhid` = 200 and $\sigma = 0.8$, leading to reliable parameter inference with a speed up of ∼20%.

REFERENCES

Abdo, A. A., Ackermann, M., Ajello, M., et al. 2009, PhRvL, 103, 251101
Abdo, A. A., Ackermann, M., Ajello, M., et al. 2010, ApJ, 710, 133
Ackermann, M., Ajello, M., Atwood, W. B., et al. 2012, ApJ, 750, 3
Ackermann, M., Ajello, M., Albert, A., et al. 2015, ApJ, 799, 86
Ackermann, M., Ajello, M., Baldini, L., et al. 2011, ApJ, 726, 81
Adriani, O., Barbarino, G. C., Bazilevskaya, G. A., et al. 2009, Natur, 458, 607
Adriani, O., Barbarino, G. C., Bazilevskaya, G. A., et al. 2010, PhRvL, 105, 121101
Adriani, O., Barbarino, G. C., Bazilevskaya, G. A., et al. 2011, Sci, 332, 69
Adriani, O., Barbarino, G. C., Bazilevskaya, G. A., et al. 2014, ApJ, 791, 93
Aguilar, M., Aisa, D., Alpat, B., et al. 2015a, PhRvL, 115, 211101
Aguilar, M., Aisa, D., Alpat, B., et al. 2015b, PhRvL, 114, 171103
Ahn, H. S., Allison, P. S., Bagliesi, M. G., et al. 2008, APh, 30, 133
Ahn, H. S., Allison, P. S., Bagliesi, M. G., et al. 2009, ApJ, 707, 593
Ahn, H. S., Allison, P. S., Bagliesi, M. G., et al. 2010, ApJ, 715, 1400
Ajello, M., Albert, A., Atwood, W. B., et al. 2016, ApJ, 819, 44
Asplund, M., Grevesse, N., Sauval, A. J., & Scott, P. 2009, ARA&A, 47, 481
Atwood, W. B., Abdo, A. A., Ackermann, M., et al. 2009, ApJ, 697, 1071
Ave, M., Boyle, P. J., Gahbauer, F., et al. 2008, ApJ, 678, 262
Barnes, T. G., III, Jefferys, W. H., Berger, J. O., et al. 2003, ApJ, 592, 539
Barwick, S. W., Beatty, J. J., Bhattacharyya, A., et al. 1997, ApJL, 482, L191
Becker Tjus, J., Eichmann, B., Kroll, M., & Nierstenhöfer, N. 2015, e-print (arXiv:1510.07801)
Benyamin, D., Nakar, E., Piran, T., & Shaviv, N. j. 2016, e-print (arXiv:1601.03072)
Berezhko, E. G., Ksenofontov, L. T., Ptuskin, V. S., Zirakashvili, V. N., & Völk, H. J. 2003, A&A, 410, 189
Berezinskii, V. S., Bulanov, S. V., Dogiel, V. A., & Ptuskin, V. S. 1990, Astrophysics of Cosmic Rays, ed. V. L. Ginzburg (Amsterdam: North-Holland)
Blasi, P., & Serpico, P. D. 2009, PhRvL, 103, 081103
Bobik, P., Boella, G., Boschini, M. J., et al. 2012, ApJ, 745, 132
Cholis, I., & Hooper, D. 2014, PhRvD, 89, 043013
Connell, J. J. 1998, ApJL, 501, L59
de Austri, R. R., Trotta, R., & Roszkowski, L. 2006, JHEP, 05, 002
Donato, F., Maurin, D., & Taillet, R. 2002, A&A, 381, 539
Dragicevich, P. M., Blair, D. G., & Burman, R. R. 1999, MNRAS, 302, 693
Duvernois, M. A., & Thayer, M. R. 1996, ApJ, 465, 982
Ellison, D. C., Drury, L. O., & Meyer, J.-P. 1997, ApJ, 487, 197
Engelmann, J. J., Ferrando, P., Soutoul, A., Goret, P., & Juliusson, E. 1990, A&A, 233, 96
Erlykin, A. D., & Wolfendale, A. W. 2015, JPhG, 42, 115202
Erlykin, A. D., Wolfendale, A. W., & Dogiel, V. A. 2016, AdSpR, 57, 519
Feroz, F., & Hobson, M. P. 2008, MNRAS, 384, 449
Feroz, F., Hobson, M. P., & Bridges, M. 2009, MNRAS, 398, 1601
Feroz, F., Hobson, M. P., Cameron, E., & Pettitt, A. N. 2013, e-print (arXiv:1306.2144)
Florinski, V., Zank, G. P., & Pogorelov, N. V. 2003, JGRA, 108, 1228
George, J. S., Lave, K. A., Wiedenbeck, M. E., et al. 2009, ApJ, 698, 1666
Gleeson, L. J., & Axford, W. I. 1968, ApJ, 154, 1011
Graff, P., Feroz, F., Hobson, M. P., & Lasenby, A. 2012, MNRAS, 421, 169
Graff, P., Feroz, F., Hobson, M. P., & Lasenby, A. 2014, MNRAS, 441, 1741
Hams, T., Barbier, L. M., Bremerich, M., et al. 2004, ApJ, 611, 892
Iroshnikov, P. S. 1964, Soviet Ast., 7, 566
Jaynes, E. T., & Bretthorst, G. L. (ed.) 2003, Probability Theory (Cambridge: Cambridge Univ. Press)
Jóhanneson, G., Moskalenko, I. V., Orlando, E., Porter, T. A., & Strong, A. W. 2015, Proc. ICRC, 517
Kachelriess, M., Moskalenko, I. V., & Ostapchenko, S. S. 2015, ApJ, 803, 54
Kachelrieß, M., & Ostapchenko, S. 2013, PhRvD, 87, 047301
Kissmann, R., Werner, M., Reimer, O., & Strong, A. W. 2015, APh, 70, 39
Kraichnan, R. H. 1965, PhFl, 8, 1385
Langner, U. W., Potgieter, M. S., Fichtner, H., & Borrmann, T. 2006, ApJ, 640, 1119
Lorimer, D. R. 2004, ed. F. Camilo, & B. M. Gaensler IAU Symp. 218, Young Neutron Stars and Their Environments, ed. F. Camilo, & B. M. Gaensler (San Francisco, CA: ASP), 105
Lukasiak, A., McDonald, F. B., & Webber, W. R. 1999, Proc. ICRC, 3, 41
Maurin, D., Donato, F., Taillet, R., & Salati, P. 2001, ApJ, 555, 585
Maurin, D., Melot, F., & Taillet, R. 2013, e-print (arXiv:1302.5525)
Maurin, D., Putze, A., & Derome, L. 2010, A&A, 516, A67
Maurin, D., Taillet, R., & Donato, F. 2002, A&A, 394, 1039
Mertsch, P., & Sarkar, S. 2014, PhRvD, 90, 061301
Meyer, J.-P., Drury, L. O., & Ellison, D. C. 1997, ApJ, 487, 182
Moskalenko, I. V., Jóhanneson, G., Orlando, E., et al. 2015, Proc. ICRC, 492
Moskalenko, I. V., Mashnik, S. G., & Strong, A. W. 2001, Proc. ICRC, 5, 1836
Moskalenko, I. V., & Strong, A. W. 1998, ApJ, 493, 694
Moskalenko, I. V., Strong, A. W., Mashnik, S. G., & Ormes, J. F. 2003, ApJ, 586, 1050
Moskalenko, I. V., Strong, A. W., Ormes, J. F., & Potgieter, M. S. 2002, ApJ, 565, 280
Moskalenko, I. V., Strong, A. W., & Porter, T. A. 2008, Proc. ICRC, 2, 129
Obermeier, A., Ave, M., Boyle, P., et al. 2011, ApJ, 742, 14
Obermeier, A., Boyle, P., Hörandel, J., & Müller, D. 2012, ApJ, 752, 69
Orito, S., Maeno, T., Matsunaga, H., et al. 2000, PhRvL, 84, 1078
Orlando, E., & Strong, A. 2013, MNRAS, 436, 2127
Orlando, E., Strong, A. W., Moskalenko, I. V., et al. 2015, Proc. ICRC, 547
Parker, E. N. 1965, P&SS, 13, 9
Picozza, P., Galper, A. M., Castellini, G., et al. 2007, APh, 27, 296
Porter, T. A., Jóhanneson, G., & Moskalenko, I. V. 2015, Proc. ICRC, 908

Potgieter, M., & Langner, U. 2004, AnGeo, 22, 3729
Ptuskin, V. S., Moskalenko, I. V., Jones, F. C., Strong, A. W., & Zirakashvili, V. N. 2006, ApJ, 642, 902
Ptuskin, V. S., & Soutoul, A. 1998, A&A, 337, 859
Putze, A., Derome, L., & Maurin, D. 2010, A&A, 516, A66
Rauch, B. F., Link, J. T., Lodders, K., et al. 2009, ApJ, 697, 2083
Scott, P., Grevesse, N., Asplund, M., et al. 2015, A&A, 573, A25
Seo, E. S., & Ptuskin, V. S. 1994, ApJ, 431, 705
Shariff, H., Jiao, X., Trotta, R., & van Dyk, D. A. 2015, arXiv:1510.05954
Shaviv, N. J. 2003, NewA, 8, 39
Shaviv, N. J., Nakar, E., & Piran, T. 2009, PhRvL, 103, 111302
Sina, R., Ptuskin, V. S., & Seo, E. S. 2005, AdSpR, 35, 147
Skilling, J. 2004, in AIP Conf. Ser. 735, Bayesian Inference and Maximum Entropy Methods in Science and Engineering: 24th International Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering, ed. R. Fischer, R. Preuss, & U. V. Toussaint (New York: AIP), 395
Skilling, J. 2006, BayAn, 1, 833
Strong, A. W., & Moskalenko, I. V. 1998, ApJ, 509, 212
Strong, A. W., & Moskalenko, I. V. 2001, Proc. ICRC, 5, 1964
Strong, A. W., Moskalenko, I. V., & Ptuskin, V. S. 2007, ARNPS, 57, 285
Strong, A. W., Moskalenko, I. V., & Reimer, O. 2004, ApJ, 613, 962
Strong, A. W., Porter, T. A., Digel, S. W., et al. 2010, ApJL, 722, L58
Tan, L. C., & Ng, L. K. 1983a, JPhG, 9, 227
Tan, L. C., & Ng, L. K. 1983b, JPhG, 9, 1289
Trotta, R. 2008, ConPh, 49, 71
Trotta, R., Jóhannesson, G., Moskalenko, I. V., et al. 2011, ApJ, 729, 106
Usoskin, I. G., Bazilevskaya, G. A., & Kovaltsov, G. A. 2011, JGRA, 116, 2104
Vladimirov, A. E., Digel, S. W., Jóhannesson, G., et al. 2011, CoPhC, 182, 1156
Vladimirov, A. E., Jóhannesson, G., Moskalenko, I. V., & Porter, T. A. 2012, ApJ, 752, 68
Webber, W. R., & Soutoul, A. 1998, ApJ, 506, 335
Wellisch, H. P., & Axen, D. 1996, PhRvC, 54, 1329
Wiedenbeck, M. E., Binns, W. R., Cummings, A. C., et al. 2008, Proc. ICRC, 2, 149
Wiedenbeck, M. E., Yanasak, N. E., Cummings, A. C., et al. 2001, SSRv, 99, 15
Yanasak, N. E., Wiedenbeck, M. E., Mewaldt, R. A., et al. 2001, ApJ, 563, 768
Yoon, Y. S., Ahn, H. S., Allison, P. S., et al. 2011, ApJ, 728, 122