New Markov-Chain Monte Carlo analyses for the evaluation of the antiproton background

P De la Torre Luque¹,², F Gargano², F Loparco¹,², M N Mazziotta² and D Serini¹,²
¹ Dipartimento di Fisica “M. Merlin” dell’Università e del Politecnico di Bari, via Amendola 173, I-70126 Bari, Italy
² Istituto Nazionale di Fisica Nucleare, Sezione di Bari, via Orabona 4, I-70126 Bari, Italy
E-mail: pedro.delatorreluque@ba.infn.it

Abstract. Current measurements of the cosmic ray spectra have reached unprecedented accuracy thanks to the new generation of experiments, and in particular the AMS-02 mission. At the same time, significant progress has been made in the propagation models of galactic cosmic rays. Nevertheless, the current knowledge on spallation cross sections is very poor, impeding a more precise estimation of the diffusion coefficient. In this work we show a new Markov-Chain Monte Carlo algorithm able to derive the propagation parameters from the flux ratios of light secondary cosmic rays (Li, Be, B) to C and O and a new procedure able to combine the flux of these secondary cosmic rays in order to get rid of the uncertainties associated to their production cross sections. Then, we show that the antiproton spectrum inferred from this diffusion model match experimental data much better than with earlier analyses, discarding the excess of data previously explained as a possible signature of antiproton production from dark matter.

1. Introduction
Propagation of Galactic cosmic rays (GCRs) is governed by the magnetic collision-less interactions they suffer with interstellar plasma waves, making them wander inside the Galaxy following a diffusive motion [1]. GCRs are accelerated inside astrophysical sources, like supernovae or pulsars, with their energy spectra following power laws that span several order of magnitude in energy, as explained by the diffusive shock acceleration mechanism (DSA) [2]. The primary CR spectra, usually called "injection spectra" are given by: \( Q_{\text{prim}}(E) = KE^{-\gamma} \) [3]. The parameter \( \gamma \) is known as the spectral index. Nevertheless, their diffusive propagation modifies this power-law, making their spectrum at Earth to have the form \( J_{\text{prim}}(E) \propto E^{-(\gamma+\delta)} \), where the diffusion parameter \( \delta \) is related to the CR confinement time in the Galaxy.

During the propagation time, GCRs interact with the interstellar gas (which consists mainly of hydrogen with about a 11% of helium and traces of heavier elements [4]) producing lighter nuclei (secondary CRs) that otherwise would be found in tiny proportions (mainly boron, lithium, beryllium and the so-called sub-Fe nuclei) since they are hardly produced in stellar fusion or other thermonuclear processes. The spectrum of any secondary CR, \( k \), formed via these interactions (spallation reactions) with gas particles in the interstellar medium (ISM) has the form: \( J_k(E) \sim \rho_g \sum_i J_i(E) \sigma_{ik}(E) \). Here, the density of the interstellar medium is \( \rho_g \) and \( \sigma_{ik}(E) \) is the inclusive cross section describing the production of the nucleus \( k \) in the interactions of the nucleus \( i \) with the ISM.
Therefore, a good knowledge on the cross sections of production of secondary CRs is crucial given that their fluxes can provide valuable information about their propagation time and, hence, to evaluate the diffusion coefficient associated to their movement in the Milky Way (since $D \propto t_{\text{prop}}^{-1}$). Nonetheless, the current knowledge on spallation cross sections is poor and based on a few data points that hardly reach energies above a few GeV.

In this way, a precise evaluation of the propagation of GCRs is very important, for instance, to study the expected antiproton flux at Earth, which could be used to constrain the possible existence of dark matter. Recent CR antiproton studies have claimed the possibility of an excess of data over the predicted flux, which can be the signature of annihilation or decay of a dark matter particle with a mass around 80 GeV into antiprotons. We derive in this work the antiproton spectra predicted by the propagation parameters inferred from Markov-Chain Monte Carlo (MCMC) analyses of the CRs secondary-over-primary flux ratios and a new strategy able to get rid of the uncertainties associated to the cross sections of production of secondary CRs.

2. Propagation setup

Simulations of CR propagation are performed numerically using the 2-Dimensional model of the Galaxy offered by the newest version of the DRAGON code [5], assuming cylindrical symmetry. In this model, the Galaxy is described as a thin disk with radius $R \sim 20kpc$ and the Sun at a distance of $8.3kpc$ from its center. The disk is surrounded by the halo, with the same radius as the disk and a height, $H$, of a few kpc (two-zone model). A diffusion-reacceleration model is considered, without convection, since recent analyses support it (see [6]).

The general formula describing this propagation model is given by the following equation:

$$-D\nabla^2 N_i + \frac{\partial}{\partial p} \left[ p^2 D_{pp} \frac{\partial}{\partial p} \left( \frac{N_i}{p^2} \right) \right] = Q_{\text{source}}^i + \frac{\partial}{\partial p} \frac{\hat{p}}{\tau_i} N_i - \frac{N_i}{\tau_i} + \sum_j \sum_{j\rightarrow i} \Gamma_{j\rightarrow i} (N_j) - \frac{N_i}{\tau_i} + \sum_j \frac{N_j}{\tau_{j\rightarrow i}} \tag{1}$$

In the previous equation, $D\nabla N_i = \vec{J}_i$ indicates the CR diffusive flux of the $i$-th species and $N_i$ is its density per unit momentum. The diffusion coefficient is considered independent of position. The second term in the left-hand side accounts for the diffusion in momentum space, related to the spatial coefficient via the Alfvén velocity, $V_A$, as $D_{pp} = \frac{4}{3} \gamma^2 (1 - \gamma^2)^{-1} \nu^2 \lambda^2 / D$ [2].

Then, the first term in the right-hand side of equation (1), $Q_{\text{source}}$, represents the distribution and energy spectra of particle emitted by sources (called injection spectra). In this work we inject $^4He$, $^{12}C$, $^{14}N$, $^{16}O$, $^{20}Ne$, $^{24}Mg$ and $^{28}Si$ as primary nuclei, such that they fit the most recent AMS-02 [8] experimental data. Finally, the second term of equation (1) describes the momentum (energy) losses and the remaining four terms describe production and destruction of nuclei due to decays and nuclear reactions. Total inelastic cross sections are associated with nuclei fragmentation from interactions with ISM gas and the inclusive cross sections are associated to production of (secondary) nuclei production [9]. Although inelastic cross sections are known with relatively good accuracy, production (inclusive) cross sections are subject to large uncertainties. We implement the GALPROP cross sections [10] parametrizations [4] as our reference cross sections for nuclei production.

Simple approximations of the propagation equation allow us to find approximate relations between the flux of CRs and the propagation parameters make us find that the best way to constrain propagation parameters is by means of the secondary-over-primary CR ratios [11]. As the flux of a primary CR is $J_{\text{pri}}(E) \propto \tau_{\text{diff}}(E)Q_{\text{source}}(E)$ and for a pure secondary CR is $J_{\text{sec}}(E) \propto \tau_{\text{diff}}(E)J_{\text{pri}}(E)$, we get a simple approximate relation between the diffusion coefficient and these ratios: $\frac{J_{\text{sec}}(E)}{J_{\text{pri}}(E)} \propto \sigma(E)\tau_{\text{diff}}(E) \propto \sigma(E)/D(E)$. Although, the accuracy showed in the

1 https://github.com/cosmicrays/DRAGON2-Beta_version
2 Publicly available at https://dmaurin.gitlab.io/USINE/input_xs_data.html#nuclei-xs-nuclei
last experimental results of the AMS-02 collaboration is of the order of 1-5\%, the uncertainties on the cross sections reach levels of 20-50\% in some channels (see [12]) which makes clear the necessity of combining information from secondary CRs to get rid of systematic uncertainties [13] related to cross sections parametrizations used.

In this study, we compare two parametrizations of the spatial diffusion coefficient able to explain CR experimental data at high energy: in the Galaxy and is usually parametrized as

\[ D = D_0 \beta^n \left( \frac{R}{R_0} \right)^\delta \]  

**Source hypothesis**  

\[ D = D_0 \beta^n \frac{(R/R_0)^\delta}{1 + (R/R_0)^{\Delta \delta/s}} \]  

**Diffusion hypothesis**

Equation (2) explains the high energy trend of CRs as a change in the spectrum of the primary sources at those energies (source hypothesis), while equation (3) justify this trend as a change in the power-law describing the diffusion coefficient (diffusion hypothesis) [14]. In these equations, \( \beta = \nu/c, R \) is the rigidity of the particle, \( D_0 \) is the constant diffusion coefficient at the reference rigidity \( R_0 \) (it is set to 4 GV) and \( \eta \) is left free since negative values of this parameter seem to be theoretically expected [15]. The other parameters are the break position in rigidity (\( R_b \)), the change in spectral index after the break position and a smoothing parameter used to allow a soft transition around the hardening position. In our analyses, we fix these parameters to the ones found in [16]: \( \Delta \delta = 0.14, R_b = 312 \) and \( s = 0.040 \).

3. MCMC analyses

In order to infer the main CR propagation parameters (\( D_0, V_A, \eta \) and \( \delta \)), a Markov-Chain Monte Carlo (MCMC) procedure, relying on Bayesian inference, is developed to get the probability distribution (PDF) for a set of propagation parameters to describe AMS-02 secondary-over-primary flux ratios and their confidence intervals. Posterior probability, \( \mathcal{L'} \), is calculated by

\[ \mathcal{L'}(\theta | \hat{D}) \propto \Pi(\theta) \cdot \mathcal{L}(\hat{D} | \theta) \]

where priors, \( \Pi(\theta) \), are set to follow a uniform distribution, the likelihood, \( \mathcal{L} \), is set to be a Gaussian function and \( \theta = \{ \theta_1, \theta_2, ..., \theta_m \} \) and \( \hat{D} \) stand for the set of parameters we are evaluating and the experimental data set, respectively. Instead of the classical Metropolis-Hastings algorithm, a modified version of the *Goodman & Weare* algorithm [17] is used by means of the *emcee* Python’s module.

In a first step, the MCMC algorithm is applied to find the propagation parameters that best describe AMS-02 ratios of Li, Be and B to C and O, independently. The expression used to parametrize the injection of primary nuclei is a doubly broken power law when we study the source hypothesis (equation (2)), and a simple broken power law, when dealing with the diffusion hypothesis (equation (3)). This serves, on one side, to compare which of the diffusion parametrizations describe better the data and, on the other side, to understand the discrepancies arising from the uncertainties in their production cross sections. The latter can be better visualized by another important observable, the flux ratios among these secondary CRs deriving from the uncertainties in their production cross sections. The latter can be better visualized by another important observable, the flux ratios among these secondary CRs deriving from the uncertainties in their production cross sections. The latter can be better visualized by another important observable, the flux ratios among these secondary CRs deriving from the uncertainties in their production cross sections. The latter can be better visualized by another important observable, the flux ratios among these secondary CRs deriving from the uncertainties in their production cross sections. The latter can be better visualized by another important observable, the flux ratios among these secondary CRs deriving from the uncertainties in their production cross sections. The latter can be better visualized by another important observable, the flux ratios among these secondary CRs deriving from the uncertainties in their production cross sections.

\[ \frac{J_k}{J_j}(E) \propto \sum_{\alpha \to k} J_\alpha(E) \sigma_{\alpha \to k}(E) \quad \text{high energies} \]  

\[ \propto \sum_{\alpha \to j} \sum_{\alpha \to k} J_\alpha(E) \sigma_{\alpha \to k}(E) C_\alpha E^{-\gamma_\alpha} \sigma_{\alpha \to j}(E) \]  

\[ \sum_{\alpha \to j} C_\alpha E^{-\gamma_\alpha} \sigma_{\alpha \to j}(E) \]  

Figure 1 displays the simulated Li, Be and B ratios to carbon (a) and to oxygen (b) from both parametrizations of the diffusion coefficient compared to AMS-02 data. As we see, experimental
data favor the diffusion hypothesis in every ratio and, therefore, from now on we will be using this parametrization of the diffusion coefficient. Then, we must highlight that our simulations are able to perfectly reproduce their shape.

![Graphs](image-url)

**Figure 1.** Flux ratios of the secondary CRs B, Be and Li to the primary CRs C (a) and O (b) for the diffusion and source hypotheses confronted to AMS-02 data as derived from the propagation parameters found in the MCMC analysis.

Nevertheless, the uncertainty related to their production cross sections (mostly in the normalization of its parametrizations) make us unable to reproduce, for same propagation parameters, the spectra of these secondary CRs at the same time, as shown from their flux ratios in figure 2.

These ratios are extremely sensitive to the cross sections and their main dependence at low energy is the value of the halo size, since the flux of the radioactive isotope $^{10}$Be (which decays in $^{10}$B) depends on this parameter, as shown in [18]. Nevertheless, at energies above 30 GeV the halo size does not matter any more and the only way of fitting the sec/sec ratios is by adjusting the cross sections parametrizations. Therefore, we adjust the Li, B and Be production cross simultaneously to get a fit of the their respective secondary-to-secondary flux ratios. In order to do this, the same MCMC algorithm is applied to these ratios above 25 GeV, with a scaling factor for Li, B and Be cross sections as the parameters to be determined. In addition, we include a penalty factor that penalizes high scaling factors, ensuring that the values found are the smallest possible. In essence, this means that, although all possible scaling factors that fit simultaneously all the Sec/Sec AMS-02 data and lay inside the experimental cross sections uncertainties are allowed candidates, the preferred candidate is that which involves the minimum change from the original parametrization.

The novel idea of combining Li, B and Be at the same time from their ratios allow us to get rid of systematic uncertainties in their production cross sections, enabling a more robust determination of the diffusion coefficient. In figure 3 we show these scaled secondary-over-secondary fluxes, where the grid lines of 2% residual are highlighted in a different colour. The resulting scaling factors able to achieve this simultaneous fit are 0.94 (~ 6% scaling) for the B production cross sections, and 0.88 (-12% scaling) and 1.18 (+18% scaling) for the Be and Li production cross sections, respectively. These values are reasonable with the expectations: the B production cross sections are more accurately known since its production depends mainly on O and C, which are the channels better described by experimental data, thus implying a smaller scaling factor needed. Be and Li production depend on a larger number of channels (also N, Ne, Mg and Si are important), and specially Li production presents the fewest number of experimental data (see [12]). Then, our MCMC analysis is applied to determine the propagation
parameters with these cross sections and both diffusion models are compared in the derivation of the antiproton, \( \bar{p} \), flux in section 4.

![Figure 2.](image)

**Figure 2.** Flux ratios among B, Be and Li derived from the model found by the MCMC algorithm with the original GALPROP cross sections. Residuals with respect to data are also displayed.

4. The antiproton flux

In this section, we show our derivation for the antiproton flux using the diffusion models found from the original cross sections and the "corrected" ones. In particular, the models derived from the B/O analyses are used, since it is considered to be less affected by cross sections uncertainties. We use the \( \bar{p}/p \) spectrum to visualize it since it is hardly sensitive to uncertainties in the fit of protons and also modulation uncertainties are mitigated. Antiprotons are particularly interesting since recent studies indicate the possible presence of a dark matter signal (since a WIMP particle could decay or annihilate into antiprotons) seen as an excess of data over predictions at around 10 GeV. However, how significant is this excess depends a lot on the interpretation of the systematic uncertainties. The cross sections of \( \bar{p} \) production are those derived in [19].

![Figure 3.](image)

**Figure 3.** Same as in figure 2, but at energies bigger than 10 GeV (since other parameters, as the halo size value, are needed to be studied to make a complete fit) and for the "corrected" cross sections.

![Figure 4.](image)

**Figure 4.** Antiproton to proton flux ratio derived from the diffusion model found for the original cross sections (a) and for the "corrected" cross sections (b) compared to AMS-02 data. Uncertainties related to the determination of the propagation parameters are also shown.
In figure 4 we show the derived $\bar{p}/p$ spectrum with propagation parameters found for the B/O MCMC analysis with the original (a) and corrected (b) cross sections including its related uncertainties. As we see in figure 4 (a), the same excess of data over predictions arises, peaking at 10 GeV. Nevertheless, this bump completely disappears when using the model derived using the corrected cross sections. This means that, on one hand, the uncertainties on the estimation of the antiproton flux are underestimated and, on the other hand, the excess of data over predictions is not significant enough to account be taken seriously yet. In short, this means that we must take into account these uncertainties in order to put correct constraints on possible dark matter or primordial black holes signals.

5. Conclusions

With this work, we have implemented a new MCMC analysis able to infer the propagation parameters from different diffusion and cross sections parametrizations, with a customized version of the DRAGON code. We have shown the novel idea of using the flux ratios among the secondary CRs Li, B and Be in combination to their cross sections parametrization to adjust their production cross sections and enable a better determination of the propagation parameters. We have shown that applying this procedure highly improves the predicted $\bar{p}/p$ spectrum w.r.t. AMS-02 data even though we did not include it to evaluate the propagation parameters, leading to a prediction which completely lays inside the experimental error bars. At minimum, it is demonstrated that the excess of data around 10 GeV depends on the spallation cross sections used to predict the diffusion coefficient, so that it must be taken into account in order to put constraints on possible dark matter or primordial black holes signals.

References

[1] Berezinskii V, Dogiel V, Bulanov S and Ginzburg V L 1990 Astrophysics of cosmic rays vol 393 (Amsterdam: North-Holland)
[2] Bell A 1978 Mon. Notices Royal Astron. Soc. 182 443–55
[3] Blasi P 2013 Astron. Astrophys. Rev. 21 70 (Preprint 1311.7346)
[4] Asplund M, Grevesse N and Sauval J 2005 ASP Conf. Ser. 336 25 (Preprint astro-ph/0410214)
[5] Evoli C, Gaggero D, Grasso D and Maccione L 2008 J. Cosmol. Astropart. Phys. 2008 018
[6] Niu J S and Li T 2018 Phys. Rev. D 97 023015 (Preprint 1705.11089)
[7] Seo E S and Ptuskin V S 1994 Astrophys. J. 431 705–14
[8] Battiston R 2020 Riv. Nuovo Cim. 43 319–84
[9] Mazzotta M, Cerutti F, Ferrari A, Gaggero D, Loparco F and Sala P 2016 Astropart. Phys. 81 21–38
[10] Moskalenko I V, Strong A W and Mashnik S G 2005 AIP Conf. Proc. 769 1612–7
[11] Reinert A and Winkler M W 2018 J. Cosmol. Astropart. Phys. 2018 055
[12] Genolini Y, Maurin D, Moskalenko I V and Unger M 2018 Phys. Rev. C 98 034611 (Preprint 1803.04686)
[13] Maurin D, Putze A and Derome L 2010 Astron. Astrophys. 516 A67
[14] Vladimirov A, Jóhannesson G, Moskalenko I and Porter T 2012 Astrophys. J. 752 68 (Preprint 1108.1023)
[15] Ptuskin V S, Moskalenko I V, Jones F C, Strong A W and Zirakashvili V N 2006 Astrophys. J. 642 902
[16] Génolini Y et al. 2017 Phys. Rev. Lett. 119 241101
[17] Goodman J and Weare J 2010 Comm. App. Math. Comp. Sci. 5 65–80
[18] Evoli C, Morlino G, Blasi P and Aloisio R 2020 Phys. Rev. D 101 023013 (Preprint 1910.04113)
[19] Tan L C and Ng L K 1982 Phys. Rev. D 26(5) 1179–82